The Digital City

by

Michael Grant

Presented for the Degree of
Doctor of Philosophy

Department of Architecture and Building Science
University of Strathclyde, Glasgow
May 2002

Abstract:

This thesis outlines the experiences of the author in conducting various research and development projects addressing different means of representing the urban environment. These projects all fall within a fifteen year period that has been characterised by the most rapid growth and diversification of any technology in history.

The document steps through four eras in the progress of these projects and, while addressing only a single viewpoint, attempts to follow the developmental thread that has linked all these activities over the years.

As with all retrospective treatments of a single technology there is always a time period when a single snapshot represents the expensive state-of-the-art which, some time later, becomes derided as worthless and outdated before finally entering a phase where it may be regarded with nostalgia and perhaps new found worth. This cycle is true to all aspects of computing technology, hardware, software and applications.

The rapid pace of progress with the computing industry has distorted this time frame allowing ground breaking applications of only a few years of age to be treated with derision by some of those who have only experienced the latest cutting edge of the technology. Unfortunately this temporal distortion has forced much of our computing history towards an early grave without providing a sufficient period within which fond memories might grow. This is lamentable not just for emotional reasons but mainly because many of today's techniques and technologies are based on yesterday's precedents.

In order to appreciate the reasoning behind any one developmental phase of the project it is necessary to place it in its context of the available computing infrastructure both in terms of hardware and software. To this end each chapter seeks to identify the key enabling technological foundations on which the work is constructed

Contents:

Figures:

Tables:

Introduction

The following chapters outline the experiences of the author in conducting a number of research and development projects addressing different systems of representing the urban environment. These projects all fall within a fifteen year period that has been characterised by the most rapid growth and diversification of any technology in history. This document steps through five eras in the progress of these projects and by maintaining only a single viewpoint attempts to follow the developmental thread that has linked all these activities over the years.

While a conventional thesis treatment dictates that the research should be contextualised within a formal literature survey, in this instance the continuity provided by reporting on a single application area allows further insights to be gained by considering only the circumstances of the technological environment within which each generation has transpired. By necessity this is an inward looking process and while similar activities have been conducted by numerous research groups the restriction of this narrative to the activities of one individual enables the trail to be followed as a consequence of the availability of evolving technology. The choice of a single application area provides a common denominator against which to gauge the influence of new developments and their effect on progress within this field. A selection of the author's published papers from the period is given in the appendix.

As with all retrospective treatments of a single technology there is always a time period when a single snapshot represents the expensive state-of-the-art which, some time later, becomes derided as worthless and outdated before finally entering a phase where it may be regarded with nostalgia and perhaps new found worth. This cycle is true to all aspects of computing technology, hardware, software and applications. The rapid pace of progress with the computing industry has distorted this time frame allowing ground breaking applications of only a few years of age to be treated with derision by some of those who have only experienced the latest cutting edge of the technology. Unfortunately this temporal distortion has forced much of our computing history towards an early grave without providing a sufficient period within which fond memories might grow. This is lamentable not just for emotional reasons but mainly because many of today's techniques and technologies are based on yesterday's precedents. In order to appreciate the reasoning behind any one phase of the project it is necessary to place it in its context of the available computing infrastructure both in terms of hardware and software. To this end each chapter seeks to identify the key enabling technological foundations on which the work is constructed.

These projects have been based on the premise that it is possible to construct a virtual representation of an urban environment which would then be useful in a range of application areas. That the configuration and construction of this model may take many forms is evidenced by the range of examples in the subsequent chapters. The difference between these representations is not just due to changes in technology or technique but is critically influenced by the goals which they are intended to address. The concept of the Digital City is loosely applied and can be taken to relate to the representation of any large scale conjunction of geometry and their associated information attributes. The city is the epitome of this concept and has served as the focus for a variety of investigation over the period.

The foundations of computer graphics in the service of architectural design were laid during the period between 1970 and 1980. This was a decade when computing facilities were far from ubiquitous with limited power, availability and access. Research was directed towards the determination of how to address design issues in order to explore the wider envelope of design criteria for the prediction of economic and functional performance. While the competence of the computer in performing repetitive calculations had never been in doubt the ability to communicate graphically had still to be fully realised and yet this functionality lies at the core of the design process. The pursuit of this goal opened a new and demanding field of research which required not just innovative thinking but also the founding of an entirely novel approach to design. Developments within this period sought to overcome the limitations of the computer environment in the quest for some form of experiential output that was capable of communicating the experience of the design and not just enumerating its characteristics.

A decade of progress initially brought slow changes in computing capabilities but the 1980s heralded the introduction of an escalating rise in power and performance. This change could not only be measured in performance metrics but also in availability. Off-site access was replaced with local resources and a growing market in personal computers was set to revolutionise the technology. At the same time the acceptance of computer graphics coincided with a growing interest in the visual quality of the built environment, this lead to an increasing need for graphical applications targeting specific design problems. The ability to address ever larger groups of buildings brought the realisation that the quality of information content could increase with the quantity. This in turn lead to the desire to create models on a city wide scale, an endeavour that brought new problems in data capture, manipulation and storage. Research and development targeted on these issues resulted in the compilation of a model that was at that time unique in its size and scale.

Towards the end of the 1980s the success of computer graphics in providing new design insights opened up new markets in all areas of application, stimulating hardware development and driving down prices. The advent of the engineering workstation with fast raster based colour graphics brought an opportunity to open a new era in modelling technology. Previously the constraints of affordable computer graphics had reinforced the formal static view of the city at the expense of addressing the dynamics of the city. The arrival of hardware acceleration provided the ability to interact with the model and colour graphics with solid shading brought realism and greater detail. Now that the fundamental difficulties in interacting with large quantities of geometrical information had been eased attention was directed at the concept of an urban information system. Linking information attributes with the geometrical form of the buildings brought the promise of a system that could represent the city with the buildings as the reference and layers of information accessible through a simple interface.

The advantage in an urban information system lies in the facility to combine the form and description of the city within an interface that promotes the ability to interact with the data. However the esoteric nature of a relatively highly priced engineering workstation and its complicated operating system mitigates against the widespread adoption of such a concept. During the early 1990s a new technology came to prominence allowing ubiquitous, low cost machines at the bottom end of the computational hardware scale to integrate all aspects of diverse media types. The generic category of multimedia encompasses a wide range of capabilities offering a platform for an innovative spread of applications. In respect to the Digital City this was a major advance in that for the first time developers had access to a platform that was essentially "good at everything" and that would encourage not just progress but also diversity

The distinction between the multimedia era and that heralded by the increased competence of the Internet is blurred by the enthusiasm with which each technology has been embraced. What the Internet did both offer and deliver would be the focus of the collective wisdom of a wide cross section of like minded individuals willing to share knowledge and software tools. This enlightened attitude resulted in a new medium for the deployment of the Digital City. The provision of an abstraction layer between the client and a wide range of information providers ensures a hardware independent set of standards that would guarantee the widest audience for any developments. This in turn would spur the industry to cater for the performance requirements of an increasingly discerning and demanding market forcing the pace of hardware development particularly in the graphics arena.

By the turn of the millennium progress is no longer marked by developments within any one sphere of technology rather it is the conjunction of mutually supportive techniques that advance the concept. The pursuit of an ever greater degree of connection between the user and the system has been fundamental to the use of interactive computing since its first inception. While the early pioneers could only imagine what future developments might offer (which they did with remarkable accuracy) many of the key concepts were already proven and just waiting to be rediscovered.

Chapter 1 In the Beginning

This chapter seeks to explore the factors influencing the adoption of computer graphics in the service of architectural design during the period between 1970 and 1980. This era saw the introduction of many of the innovations in computing technology that laid the foundations on which today's applications are built. In order to establish this groundwork the early adopters of the technology had not only to devise totally new and innovative methods of addressing architectural design problems but also found themselves faced with the problems of creating practical implementations within a machine environment that, by today's standards, would be considered impossibly primitive.

1.1 Agents of Change

Today, at the turn of the millennium, computers and graphics are inextricably linked, software without graphical output or at least a graphical user interface (GUI) occupies an ever diminishing niche market. Consequently there is little demand for a mainstream hardware platform incapable of supporting graphical software. However this was not always the case and a complex and intertwined relationship between the desires of the users, the capabilities of the hardware and the availability of software have formed the stimuli to bring the technology to today's status.

No one aspect of the advance of technology or market pressure can be credited as the sole driving force behind this period of change. Each incremental advance has been a combination of factors in one area that have in turn inspired further evolution in another. Traditionally the role of computers had been regarded as addressing a narrow domain targeted at purely numerical problem solving. Within these limits the progression to faster processors, greater availability of system memory and more capable operating systems was clearly self justifying. However, as computing became an ever more ubiquitous tool directed at a widening range of applications then additional stimuli appeared from new directions.

In the field of architectural computing applications a number of evolutionary trends could be identified, perhaps the most important among them being the requirement for some form of graphical capability. This need stemmed from an obligation to service three fundamental sectors within the design process. The challenge was to meet these obligations while still reconciling the needs of the designer with the capabilities of the computer.

Within the design process these needs could be identified under the following points:

- In the early design stages there is a need for a graphical display of geometrical data which could be used as an interactive aid to the optimisation of a design hypothesis.
- Within the entire design cycle there is a desire for some form of experiential output from the computer, capable of simulating the experience of the design rather than just enumerating it's characteristics.
- At the end of the design process there is a need for the provision of graphics addressing the purpose of documenting the design: particularly directed towards the production of working drawings.
- While not a facet of the design activity in its own right, the above functionality created a need for the ability to interact graphically with the computer and through this interface gain both a ready interaction with graphical quantities and also a more intuitive control over program functions.

Each of the above are as strong a stimulus today as they were thirty years ago. However, progress within these areas can be directly linked to the availability of suitable computer hardware and associated peripherals. Therefore, before any of these developments can be investigated it would be profitable to establish the nature of the computational environment within which this initial evolution took place. This will encompass the availability of computing resources, the standard modes of usage and the costs and capabilities of the available technology.

1.2 The 1970s Computing Environment

Computers of this period were generally large, self contained machines located remotely from most users and continually fed batches of programs and data by operators who endeavoured to keep the machines productive. Initially this "batch" mode would be the only available option for a user's interaction. Paper tape or card decks could be prepared off line and would be sent to some centralised computer resource for processing. Any output would then be prepared at the computer site and eventually returned to the users. As demand grew more users became remotely connected to the computing resource through the telephone network, as long as they had the benefits of diect dialling, and would then be able to make greater usage of the facilities.

During this time period the degree of outright ownership of large scale computing resources was small. Only the largest multi-national companies could afford not only to purchase and operate computer facilities but also to fund the research and development of the necessary software. For the majority, access to computing resources was mainly obtained through the services of a bureau where the user would purchase units of time on a large time sharing computer, this presupposes that the required programs did both exist and if so were capable of being supported by that hardware configuration. Academic institutes, such as ABACUS, relied on one of a number of national computer centres that serviced the needs of all users within the catchment area for that region. Research projects had to restrict computer usage to fall within a pre-allocated budget of "Accounting Units" (AUs) which determined the amount of computer resources that could be used. The worth of an AU was not just a factor of processor time but could also be consumed at a variable rate depending on memory usage. An additional economic implication would be dictated by the mode of usage, variable accounting rates applying to "long" or "short" batch queues as well as "demand" mode for interactive access.

The main computing resource used by ABACUS during this period was the UNIVAC 1108 mainframe at the National Engineering Laboratory, East Kilbride. This was the largest computer configuration in Scotland at the time and provided extensive hardware and software facilities for both Government staff at NEL and remote users from industry and universities in both Scotland and England.

The following are representative prices for UNIVAC 1108 components as of 1968. These costings were described as "Basic" and were quoted to customers who performed all their own maintenance and support.

Model	Description	Price
3011-95	1108 CPU	$566,460
7005-52	131k word Core Memory	$823,500
5009-00	FASTRAND Controller	$41,680
6010-00	FASTRAND II Storage Unit	$134,400
5012-00	FH-432 / FH-1782 Drum Controller	$67,360
6016-00	FH-432 Drum (capacity 262,144 words	$34,640
6015-00	FH-1782 Drum (capacity 2,097,152 words)	$95,680
4009-99	Console (TTY-35)	$29,365

Table 1.1 UNIVAC specification and price (1968).

In comparison with today's hardware budgets these prices appear extreme even before factoring in inflationary costs. However it is interesting to note the disparity in terms of relative component costs that is highlighted in the following quote from the UNIVAC Memories Web Site (Walker 2000).

"The depreciation of the U.S. dollar since 1968 makes a benchmark of the value of the dollar in those days useful. In 1968, a sporty domestic automobile, the Chevrolet Malibu Sport Coupe with a 307 cid V-8 engine cost US$2663, 696 times less than a megabyte of UNIVAC core memory."

The Central Processing Unit had 192k words of core storage memory, 128 high speed arithmetic and indexing registers and an extensive instruction set. Sixteen independent input-output channels formed the interconnect between the CPU and a range of storage devices, standard peripherals and communication interfaces.

FASTRAND Drums were used as the primary mass storage media with six high speed devices giving a total capacity of 1.4 million words and two low speed units with a capacity of 46 million words. The high speed sets were utilised as a backing store for computational processes while the low performance units stored permanent user files. Six magnetic tape units were also available for off-line storage.

Storage capacity	22,020,096 36-bit words 132,120,576 6-bit characters approximately 90 megabytes
Drum speed	880 revolutions per minute
Read / write heads	64, on a movable boom between the drums.
Average access time	92 milliseconds
Transfer rate	26,286 36-bit words a second approx. 100 Kb per second
Recording density	1000 bits per inch
Track density	105 tracks per inch
Price (1968 dollars)	Controller : $41,684 Storage unit : $134,400 Maximum 8 storage units per controller

Table 1.2 Fastrand II drum storage specifications.

The system employed two card reader-line printer units and a separate high speed line printer as well as a high speed paper tape input-output sub-system. Communications interfaces connect the UNIVAC to thirty-two on site teletypewriters, eight GPO auto-answering lines capable of handling off-site terminals, two local graphics display terminals and other front end computers used for remote job entry and display drivers. Two Calcomp drum plotters were available, one permanently on-line and the other for off-line jobs.

The UNIVAC 1108 was controlled by the EXECUTIVE VIII (Exec 8) operating system that provided a multi-programming, multi-access, time sharing environment within which a number of compute jobs could share processing time. Jobs could be submitted either in batch mode or in the higher priority demand mode that would be the normal mode for interactive or graphical programs initiated and controlled from an on-line terminal.

ABACUS made use of the facility through terminal equipment connected over dedicated GPO lines to the remote site. These terminals included two standard teletypes, ASR-33s, with automatic paper tape readers and punches and two Tektronix 4010 direct view storage tube graphics displays. Local hard copy output from the graphics terminals was available either on demand through keyboard commands or executed direct by the computer program.

Despite the huge capital investment devoted to the UNIVAC at the National Engineering Centre the state of the art still imposed many fundamental limitations on the user. These limitations were evidenced not just in regard to the processing capabilities but also in the restrictions inherent to the available devices that could be used to take input from the user and produce output from the system.

In the event, the adoption of computer graphics as a generic means of built form appraisal would lag accepted numerical methods by almost ten years. This was not due to the limitations of computer graphics themselves as many contemporary packages were capable of producing wire line views in both plan and perspective. However this form of output was viewed mainly as a means of validating data input and not as adding any value to existing appraisal metrics. While the desire to progress beyond this point was evident there would still be significant obstacles to overcome.

As stated previously this state of affairs was mainly due to the difficulty experienced in interacting with the computer. For most users the accepted method of interaction with the computer was in a batch mode. This obviously meant that the advantage of using graphics as an immediate medium for feedback and evaluation were lost. It also meant that without the benefits of interactive graphics all input was entered numerically, allowing no opportunity to visually check for errors before the simulation was run. Given that the data entry procedure could itself take several days this was universally acknowledged as being both tedious and frustrating. Although purely numerical simulation programs laboured under the same data entry restrictions the input values required were generally more meaningful and the operator could make use of a paper based, form filling procedure that made errors easier to spot. In a pure graphics context the large numbers of Cartesian co-ordinates required were next to impossible to check without some means of visual feedback.

The nature of any interaction between the designer and the computer is determined, to a greater extent, by the available hardware. "Drawing" by typing co-ordinates on a teletype was a process guaranteed to be alien to most design professionals. At the time, this limitation was held as just another constraint to be added to the design problem but the future of computer based design tools would be dependent on the availability of a hardware interface that would be amenable to the working practices of architectural professionals.

All interface hardware can be classified under three main headings depending on their capabilities:

- Input only.
- Output only.
- Input and Output combined.

1.3 Interface Hardware - Input Only

Within the parameters of this discussion the one peripheral in the "input only" class and capable of interactive operation was the tablet digitiser. This device was capable of accepting, or recording for subsequent input, a series of two dimensional co-ordinates from a manually operated stylus or cross hair. The pointing device could be mounted on a tracking transport, in which case the points were measured electro-magnetically, or the "puck" could be tether free and the co-ordinates were measured by capacitive or inductive means. Digitising tables were available in a range of physical sizes from A0 to A4. The sophistication of such systems varied considerably but a typical A1 size device would cost between £3,000 [£27,930] and £7000 [£65,170] while a complex system incorporating a mini computer and associated application programs could retail for as much as £50000. [£465,500] (Gott and Tilbrook, 1973) The need for an enhanced capability such as that provided by the above could come from two of the main drawbacks of digitising tables.

Firstly, there was a usually need for the co-ordinate data to be assigned and attributed with alpha numeric quantities thus requiring some measure of control and keyboard assistance to the input process. Secondly, a fundamental disadvantage of using a digitiser is the difficulty in keeping track of which parts of the input had already been processed. As a result it was common to slave the digitiser to a local mini-computer driving a plotter or graphical VDU so as to echo the current stage of the input.

While these limitations were a disadvantage, a digitiser still provided the most natural interface to application programs and so helped resolve the problems inherent in entering large amounts of graphical descriptions to the computer. A further advantage, to be discussed later, was the availability of an extensive area of "real estate" on the digitiser that could be utilised to provide a limited graphical interface and so improve the man machine communications barrier.

1.4 Interface Hardware - Output Only

At the end of a program run there was an inevitable need to receive the output in some human readable form. This was especially true in the case of batch mode operation where the printed output was the only available medium through which the computer could communicate with the user. There were other devices that have been omitted from this discussion such as laser plotters and microfilm printers. However, the huge cost of the former and esoteric nature of the later made these modes the exception rather than the rule and relegated them to niche markets. The two most generic devices were the line printers and incremental plotters.

1.4.1 Line Printers.

Although the term "line printer" has now become depreciated and is commonly used to refer to any typewriter style printing device, it was originally so named because each entire line of printed output was composited before being printed. The advantage of this class of device lay in its printing speed. It was possible to print up to 120 characters per line at speeds of 300 to 1000 lines per minute although these rates were only achieved via direct connection to the host. This was purely an output device having no capability to send data to the host.

Despite being nominally restricted to a purely alpha-numeric function both teletypes and line printers have had a long history of use as graphical output devices. Crude diagrams and graphs could be formed using the standard set of characters. For example a row of hyphens, '-', form a horizontal line, the letter 'I' representing a vertical stroke, as in Figure 1.1

```
I----------I -------------I                       I--------------I
I          I              I-- I------------------ I              I
I          I              I   I                   I              I
I          I              I   I                   I--------------I
I          I -------------I   I------------------ I              I
I----------I                                         I-----------I
I          I -----------------I------------------I   I           I
I          I                  I                  I   I           I
I          I                  I                  I   I           I
I----------I -----------------I                  I   I           I
                              I                  I   I           I
                              I------------------I -- I----------I
```

Figure 1.1 ASCII diagram of a floor plan formed from characters.

Alternately, diagrams could be formed from blocks of characters such as the room layout depicted in Figure 1.2. Although restricted to simple rectilinear forms these techniques enjoyed a measure of popularity as it was cheap, fast and employed the lowest common denominator of output device. The disadvantages lay in the limited palette of graphical forms available and the complexities involved in programming the correct sequence of output characters.

```
AAAAAAAAAAAA AAAAAAAAAAAA                       HHHHHHHHHHHH
AAAAAAAAAAAA AAAAAAAAAAAA CC FFFFFFFFFFFFFFFFHHHHHHHHHHHH
AAAAAAAAAAAA AAAAAAAAAAAA CC FFFFFFFFFFFFFFFFHHHHHHHHHHHH
AAAAAAAAAAAA AAAAAAAAAAAA CC FFFFFFFFFFFFFFFFHHHHHHHHHHHH
AAAAAAAAAAAA CCCCCCCCCCCC CC CCCCCCCCCCCCCCCCCC CCCCCCCCCCCC
CCCCCCCCCCCC CCCCCCCCCCCC CC CCCCCCCCCCCCCCCCCC CCCCCCCCCCCC
CCCCCCCCCCCC EEEEEEEEEEEEE  GGGGGGGGGGGGGGGGCC DDDDDDDDDDD
DDDDDDDDDDD EEEEEEEEEEEEE  GGGGGGGGGGGGGGGGCC DDDDDDDDDDD
DDDDDDDDDDD EEEEEEEEEEEEE  GGGGGGGGGGGGGGGGCC DDDDDDDDDDD
DDDDDDDDDDD EEEEEEEEEEEEE  GGGGGGGGGGGGGGGGCC DDDDDDDDDDD
                            GGGGGGGGGGGGGGGGCC DDDDDDDDDDD
                            GGGGGGGGGGGGGGGGCC DDDDDDDDDDD
```

Figure 1.2 Graphical room layout as depicted on a line printer.

1.4.2 Incremental Plotter

As its name suggests an incremental plotter draws discrete vectors by using a moving pen on either a stationary paper sheet or by bearing on a continuous roll mounted on a transport system. The limitations lie in the speed at which the device can be driven, being restricted by the available acceleration, and the inherent constraint of drawing all shapes as a series of interconnected straight line segments. Generally these devices were very slow, an averagely complex plot taking up to an hour to complete, and where many alpha-numeric characters were required, much longer. Since charges were based on connection time this form of output was disproportionately expensive. Economy of operation could be achieved by writing the plot file to tape and either producing the drawing when computer time was less expensive or by transferring the output file on tape to a less powerful machine which could then drive the plotter. As previously, contemporary sources give the cost as lying between £1000 [£7,372] *"for the smallest and simplest"* and £100,000 [£737,200] *"for the large and highly sophisticated"*

1.5 Interface Hardware - Input/Output Devices

The classic input/output device was undoubtedly the teletypewriter but as the demand for interactive access grew and generated an attendant need for faster modes of interaction with graphical capabilities then the demand for Visual Display Units (VDU) grew in proportion.

1.5.1 The Teletypewriter.

In operation a teletype is much like an electronic typewriter but with the ability to host bi-directional communication between the user and the computer. The user may type instructions at the keyboard, the characters being sent serially to the computer, in response the computer can both echo the input and send any output back down the line to the teletype. All input/output (IO) is printed on continuous paper sheet. When connected remotely, communication speeds were initially limited to 110 baud, only rising to 2400 baud after some years. (110 baud equates to about 10 characters per second or 100 words per minute.) A response time of a second or two between transmit and receive was generally accepted as "good" when connected to a time sharing service. Some more advanced teletype devices were capable of supporting their own limited local stores, these then being able to produce and read punch tape and latterly magnetic tape.

1.5.2 Video Terminals.

While all video display units (VDUs) operated on the same basic principles, there were basically two broad classifications within the group, raster scan displays and vector displays. In either case the basic mode of operation was similar in that a phosphor film on the screen is activated by an electronic cathode ray and the activated points emit light.

1.5.2.1 Raster Scan Terminals.

Raster scan VDUs can be regarded as belonging to one of two categories, graphical or alpha-numeric. Although this concept may seem foreign the principle is still incumbent within most modern window managers where shell windows can be invoked as either textports or in a graphical mode but not both. The cheapest and most popular VDUs were those capable of only displaying characters. The reason for this differentiation in capabilities is due, in part, to the speed of communications between computer and terminal and hence the need to buffer the computers output on the terminal.

A raster display functions in much the same way as a domestic television. An electron beam of varying intensity scans the whole screen along a fixed raster or pattern, text being painted on the display because points in the raster are activated to a greater or lesser extent. In order that the display is kept illuminated the raster must be scanned at a vertical refresh rate of not less than 30 Hz. Unless the computer and the communications infrastructure could sustain the required transfer rate then the entire terminal output must be stored on the terminal itself. A standard computer quality monitor could display around 30 lines of 80 characters, each of which would require a local buffer. Given this local storage capacity some higher specification VDUs offered the ability to enter and edit a screen full of text before submitting it for processing. An intelligent terminal, with minimal buffer and some editing capabilities would cost up to £3,000. [£27,930]

The pinnacle of development of the raster scan device was the true bitmapped display. In this mode each individually addressed picture element, a pixel, could be stored in local graphics memory and mapped onto the screen. However during this period, such displays were the preserve of large companies and well founded research laboratories. To store and display the intensity of each single pixel on a low resolution display of 640 by 480 at a depth of 8 bits per colour component would consume 640 x 480 x 24 bits of memory, a total of just under a megabyte. Given that a 8k upgrade to the core memory of a Data General Nova mini computer cost £3000 [£9510] in 1979 and that the main memory of a general purpose mainframe was in the order of 128k words a high resolution 24 bit frame buffer was a rare item. A further restriction was the available bandwidth between the host and the terminal, this restricted the application of bit mapped displays to that of a static frame store local to the host. Even by the end of the decade a low resolution colour frame buffer could cost around £40,000. [£107,200].

Figure 1.3 Data General Nova 820 at ABACUS in the mid 1970s.

All vector displays operate by commanding the electron beam to move directly to point A and then draw a line directly to point B. A "dark move" is used to reposition the electron beam to a new address where it is then turned on to draw in a “bright move” to the subsequent position. Because there is no rasterisation the display exhibits no aliasing artefacts thus providing a very clean image. Once again the predominant technologies fall into two classes, refresh screens and direct view storage tubes (DVST).

The standard or refreshing vector display combined high persistence phosphors with a continually scanning electron beam to keep the screen illuminated as the computer generated the graphics and traversed the display list . As the number of elements increased there was a limit to the complexity of the display before the screen started to flicker. Given a minimum refresh rate of 16Hz the practical limit to the number of vectors simultaneously displayed was in the order of 300. The principle drawbacks were the large amount of computing power required to process the display list and the limiting band width of the interconnect. This initially restricted the use of such displays to a direct connection local to the host computer. Later, more intelligent versions of refresh terminals would emerge that employed their own dedicated processing capability, effectively becoming a slave of the host. Typical cost in 1972 was in excess of £10,000. [£93,100].

1.5.2.2 (Direct View) Storage Tube Displays

Within a DVST there are a pair of electron guns, one of which draws the elements on screen while the other illuminated the entire display maintaining the illumination of the phosphors that have been activated. The drawing is stored within a charge maintained on a grid of very fine wires behind the phosphor screen. In this way there was no limit to the number of vectors that could be simultaneously displayed. The down side was that the screen could not be selectively refreshed, the correct methodology being to divide the program's output into pages of into pages of information, the entire screen being erased between pages. This meant that any representation of motion or animation was impossible, similarly the display characteristics were limited to the use of vectors and outlining. The crudity of attempting anything other than the most basic 2-D hatching or polygon fills precluded their use for any visually realistic shading.

Taking the Tektronix 4010 and 4014 as typical of this class of technology, and period of operation, the following characteristics can be noted. DVST graphical terminals were fully teletype compatible providing the normal alphanumerical input and output facilities of a teletype on a display screen capable of holding 35 lines of up to 72 characters. The terminal could also be toggled to and from a graphics mode, either programmaticaly or via a keypress on the terminal keyboard, in which characters sent to the terminal are used to address a position on the screen. Each instruction was comprised of a character followed by four bytes of binary information. For instance a point could be defined by the character "p" where the next four bytes represent the x and y co-ordinate each value being an unsigned integer.

In graphics mode the terminal had 1024 by 1024 addressable points, of which only 1024 by 780 were visible on screen, and a program could draw vectors, either visibly or invisible, to or from any given point. The 4014 terminal could be upgraded with the addition of an Enhanced Graphics Module (EGM) which then provided an addressable area of 4096 by 3120. While the illuminated phosphor on the screen was considerably bigger than this resolution suggests, the inherent precision made this technology ideally suitable for CAAD applications. Even today this screen resolution is only equalled by plotters and similar hard copy devices. A Tektronix 4010 with associated hard copy unit cost £4100 [£30,223] in 1973

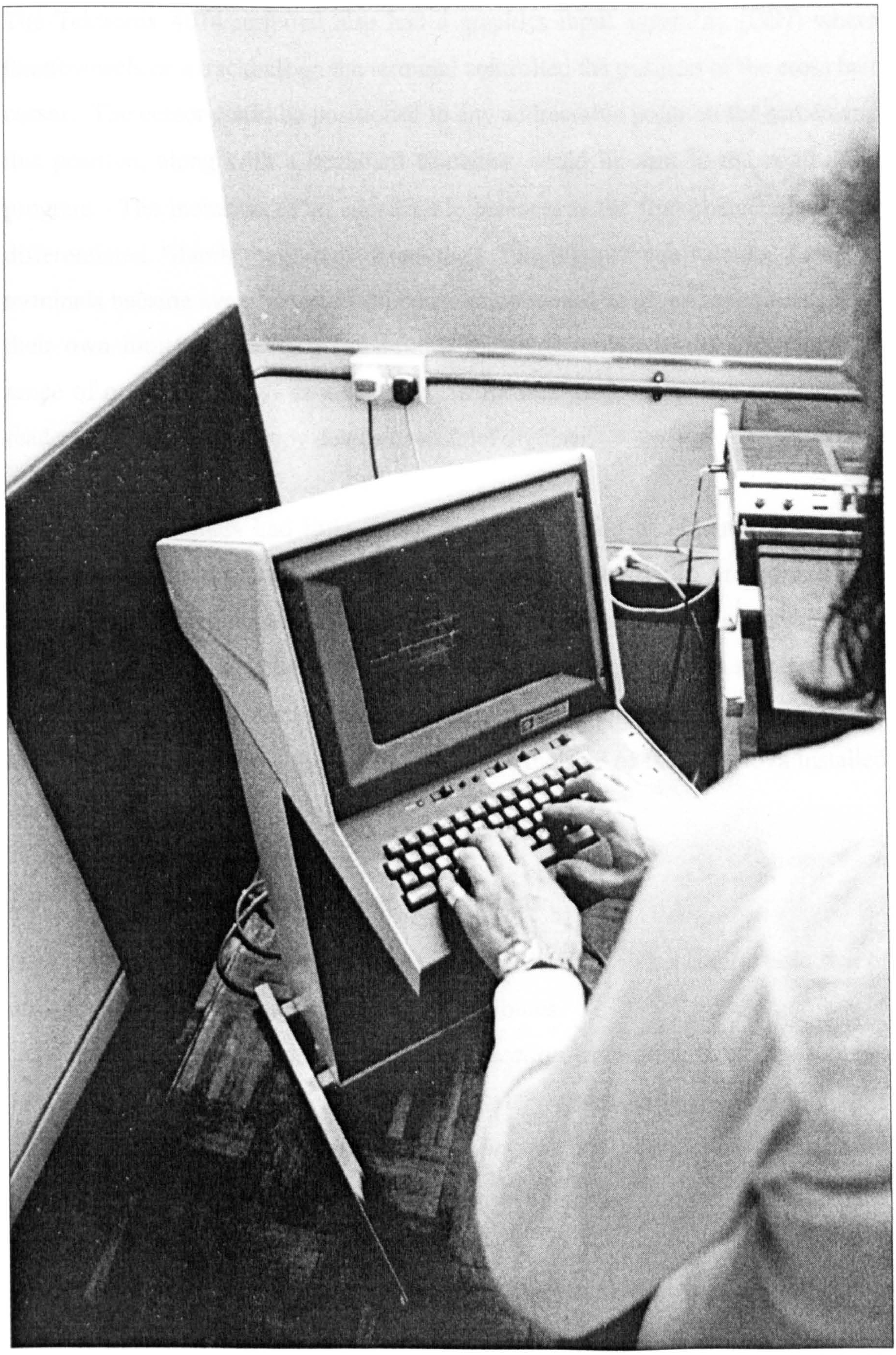


Figure 1.4 Tektronix 4010 DVST terminal at ABACUS in the early 1970s.

The Tektronix 4014 terminal also had a graphics input capability (GIN) where thumbwheels or a trackball on the terminal controlled the position of the cross hair cursor. The cursor could be positioned to any addressable point on the screen and this position, along with a keyboard character, could be sent to the controlling program. The inclusion of an addressable cursor was the first characteristic that differentiated "dumb" terminals from their "intelligent" equivalent. Later as terminals became even "smarter" this distinction would be given to terminals with their own limited processing capabilities. Smart terminals could also support a range of peripherals such as a slave screen for dual page operation, a paper tape reader or punch, a hard copy device or a tablet digitiser.

Even smart terminals had little useful functionality and to obtain any leverage from the use of interactive graphics it was often necessary to provide a mini-computer as an interface. The prime function of the mini computer would to: service communications between the remote mainframe and the on site terminals, to buffer display files for the local displays, and to allow some local calculation and modification of the display files. This was the role of the DG Nova installed at ABACUS, see Figure 1.3.

1.6 Architectural Computing Applications

If a computer aided design system is to help the designer in his central task, that of design, then it must include the following attributes.

> *"It must focus on form, nurturing its generation, permitting its review through visual simulation as well as through other modes, aiding its modification. Communications with it must be graphic, and its analytic operations must relate to form changes." (Kraus et al, 1968)*

Given the above it can been seen that the implementation of this class of design tool could only be enabled by the availability of a computing infrastructure that supported access to remote resources through a low cost graphical terminal. Notwithstanding these purely pragmatic considerations it must also be evident that there is a requirement for the developers of such applications to exercise the vision required to map the embryonic capabilities of an emerging technology to the needs of design problem.

In the period leading up to the 1970s the Department of Architecture and Building Science at the University of Strathclyde had developed a theme of research addressing specific problems in the design of large public buildings, notably schools and hospitals. Subsequently ABACUS developed a number of computer programs which drew on this research, SECS (Marcus 1967) which was capable of generating schedules of accommodation for educational establishments and PACE, (Maver 1972) which was the prototype for the next generation of comprehensive programs for the appraisal of built form.

Building on the experience gained from these initial studies, and supported by grant aid from the Science Research Council, a suite of programs were developed which formed much of the research focus within ABACUS for the next ten years. The philosophical basis for these evaluation programs is shown in Figure 1.5.

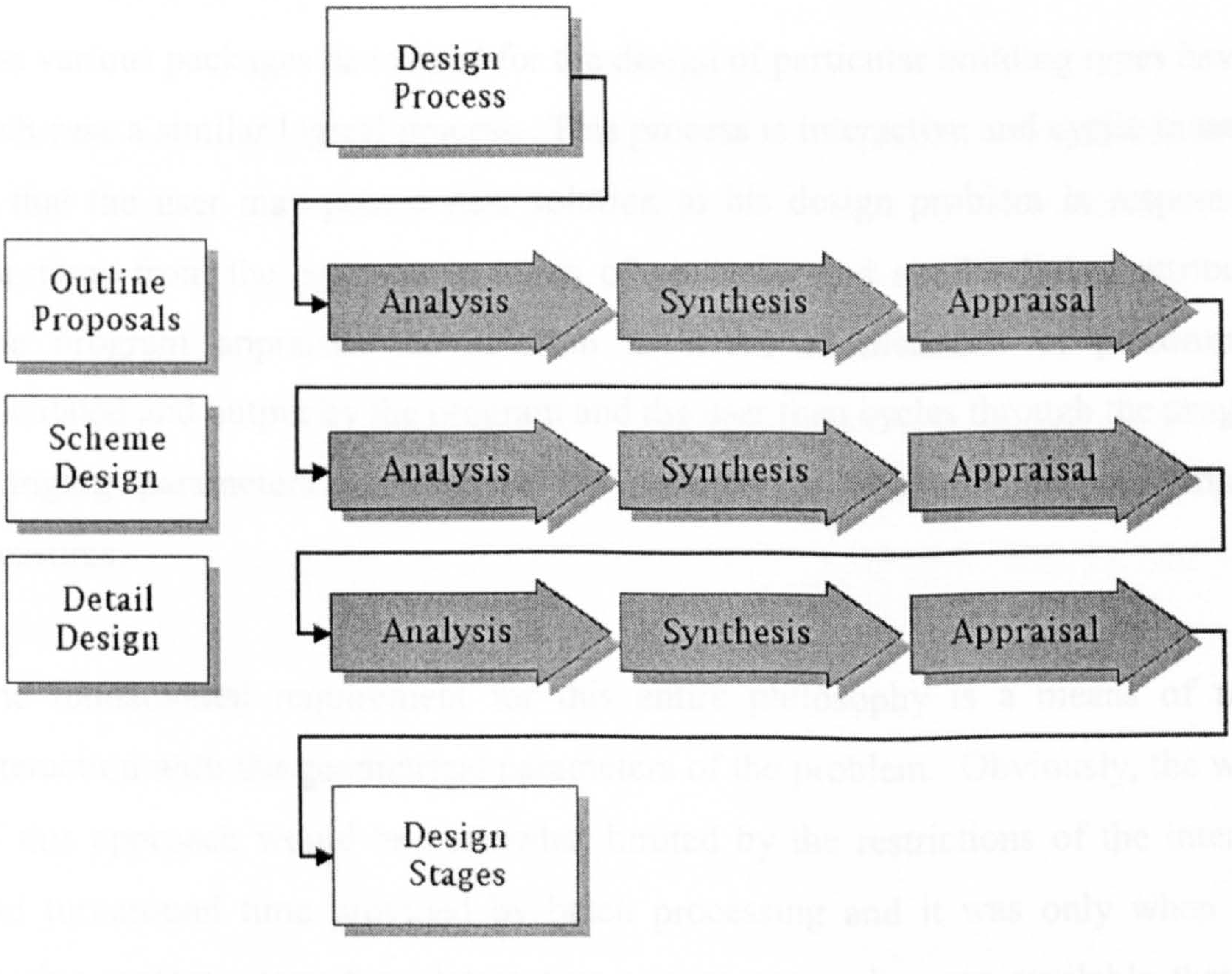


Figure 1.5 A model of the design activity

This diagram models the design process as a function of the RIBA's practice and management strategy. (RIBA 1965) Marcus and Maver maintained that the design process could be regarded as a sequence of three progressive steps, analysis, synthesis and appraisal, and that this sequence may be concluded by a decision either to reiterate the design problem with varied parameters or to proceed to the next stage of the proposal where the same sequence is again repeated at increasing levels of detail.

In the context of schools, the main tasks are as follows:

- analysis involves the translation of curricular, educational policy and teaching methods into a schedule of accommodation.
- synthesis involves the generation of formal layouts based on the accommodation schedule and certain explicit functional and planning criteria.
- appraisal involves the representation, measurement and evaluation of the cost and performance attributes of any particular built form layout.

The various packages developed for the design of particular building types have in each case a similar logical process. This process is interactive and cyclic in nature in that the user may pose a new solution to his design problem in response to questions from the program in terms of geometry and similar linked attributes. The program appraises the solution in terms of measures of performance calculated and output by the program and the user then cycles through the program changing parameters to improve his solution in terms of the performance measures.

The fundamental requirement for this entire philosophy is a means of ready interaction with the geometrical parameters of the problem. Obviously, the worth of this approach would be somewhat limited by the restrictions of the interface and turnaround time provided by batch processing and it was only when time sharing systems permitting interactive remote access became available that this style of programming became practical.

However, there was still much to be accomplished before this potential could be realised. As all aspects of this technology were new, a potential developer had not only to create the high level application code but was also required to evolve methods of linking to the latent functionality of novel hardware systems. In addition to these factors the entire methodology of utilising interactive computer programs for architectural design was yet to be tried and tested.

1.6.1 Graphical Interface Development

Within the "Architects' Specification for a Computer System" mentioned in the previous section, Kraus et al state that the system must:

> *".. allow the designer to select what question to probe next, to operate at any scale, to check any criteria at any time during the design process. The designer must be able to look at part, whole or enlarge the context of the problem proceeding with operations in the order he judges best."*
> *(Kraus et al 1968)*

A central restriction was the limited capabilities of graphics functionality. In today's computing environment the developer has a choice of numerous high level libraries providing functionality that addresses a wide spread of fundamental tasks. These high level libraries form an abstraction layer between the low level capabilities of the hardware and the more commonly required functionality required by the application. In the period under discussion the intrinsic functionality of the hardware was minimal and many of these libraries had yet to be developed

1.6.2 Software Functionality

The software requirements are therefor to develop aids to this process which provide the program user with the means to control his path through the program and to display or change in particular geometrical data. This had been achieved by developing a set of common graphical and menuing routines which could be called and used as separate modules by any applications program. (Sussock 1973) Thus any program which has a modular structure comprising in general:

- a controlling monitor program.
- graphical manipulation routines.
- base graphics and menu routines.
- algorithmic routines for calculating performance metrics.
- other routines e.g. for data structuring.

The limited nature of the base hardware functionality addressing graphical and terminal control routines was provided by the hardware manufacturer, in the case under discussion Tektronics, for use on any compatible host and performed the following functions.

- screen control.
- mode setting.
- setting co-ordinate scales, display window and viewport definition.
- vector plotting.
- cursor input control.
- alphanumeric character output.

These low level tasks were augmented by ABACUS to provide a second library layer combining low level routines to allow the display of a generic range of output primitives such as rectangles, ellipses and circles and to allow the specification and display of menus of commands from a program that permitted the picking of an item from the menu using the cross hair cursor.

Using routines from this basic set, a program could display a combination of pictures and text such as bubble diagrams, axonometric views or hidden line perspective views and could allow the user to exercise control over the program with minimal typing at the keyboard. The basic routines described above are relatively unsophisticated and do not greatly enhance the interactive interface required by the program user although they do provide some visual rather than alphanumeric feedback. The design of these applications requires the user to interact with the visual information by changing the building geometry in a natural and efficient manner. To this end a further library layer was constructed GRAMP (Graphics Manipulation Package) to enable these sets of tasks within the limitations of the storage tube display terminal.

1.6.3 A Graphical Manipulation Package

GRAMP uses the menu routine to display a menu of allowable commands to the package and displays a vector diagram of the current geometry, Figure 1.6. Using the terminals cross hair cursor the user can operate on the geometry to create edit and manipulate both the size and shape of the geometric primitives and their relation to one another.

In normal operation of the package the graphics display occupies most of the screen area but it is also possible to manipulate the viewport so that a re-scaled image shares the screen with supplementary text or additional graphics. The utilities provided by this library are enumerated in Table 1.3.

menu	action
HEAD	To produce a picture heading giving the display scale of the picture
GRID	To draw a tick grid along the x, y axes of the picture at user specified spacing
NUMBER	To number the elements, either in input order, or with user specified numbers
CREATE	To generate a new element at a given position and size
ERASE	To erase an existing element
MOVE	To change the position of an element
SHAPE	To change the shape of an element, preserving its area
MOVSHP	To move and re-shape an element
DRAW	To draw the picture at full scale
REDRAW	To draw the picture at the current scale
ZOOMIN	To re-draw at a scale increased by a given factor
ZOOMOT	To re-draw at a reduced scale
SCALE	To redraw at a user specified scale
PUSH	To push one element against another
BOUND	To push elements on a line by moving the notional boundaries of the picture
CANCEL	To cancel the previous change of geometry

Table 1.3 Graphical functionality provided by the GRAMP interface.

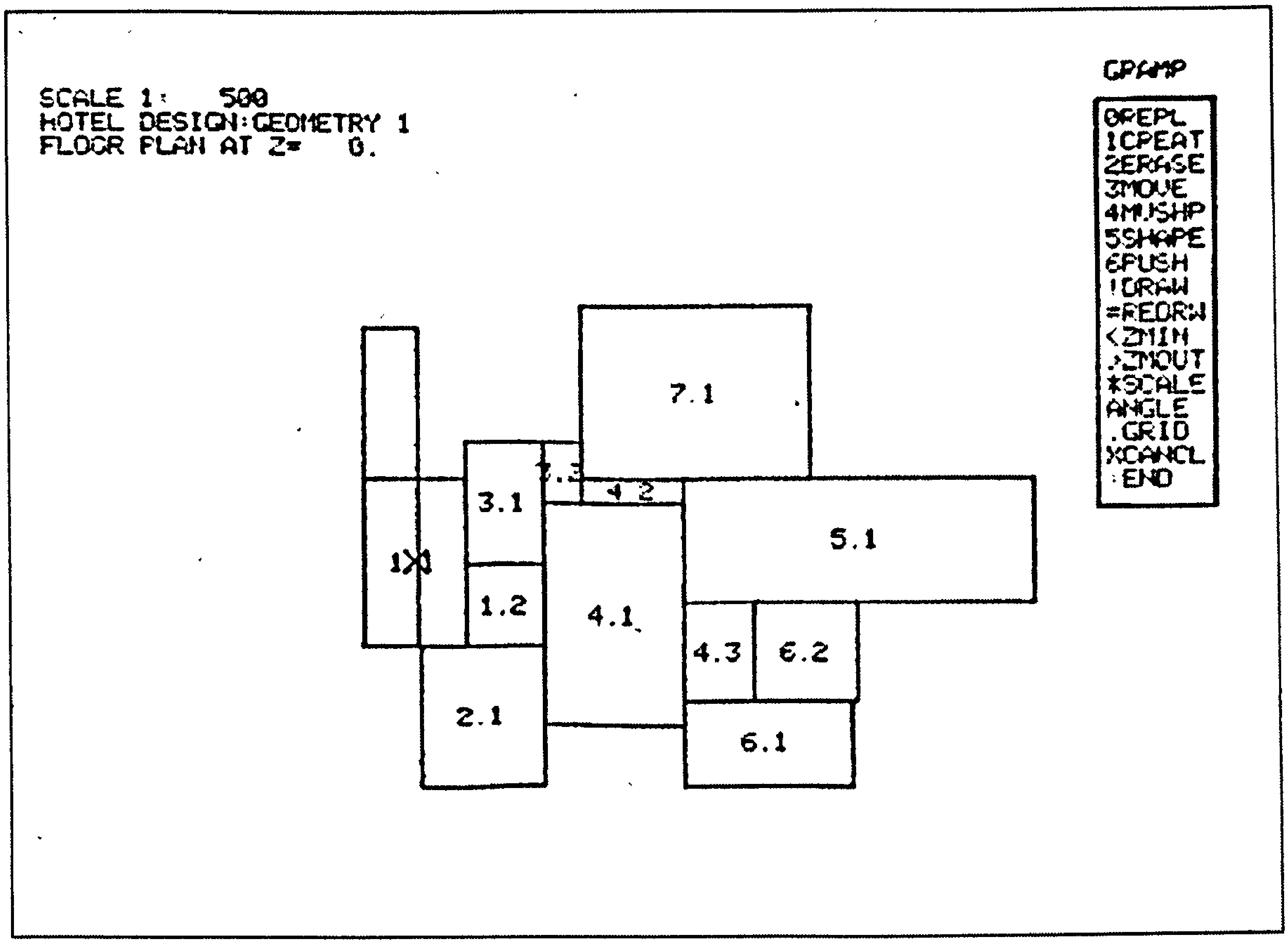
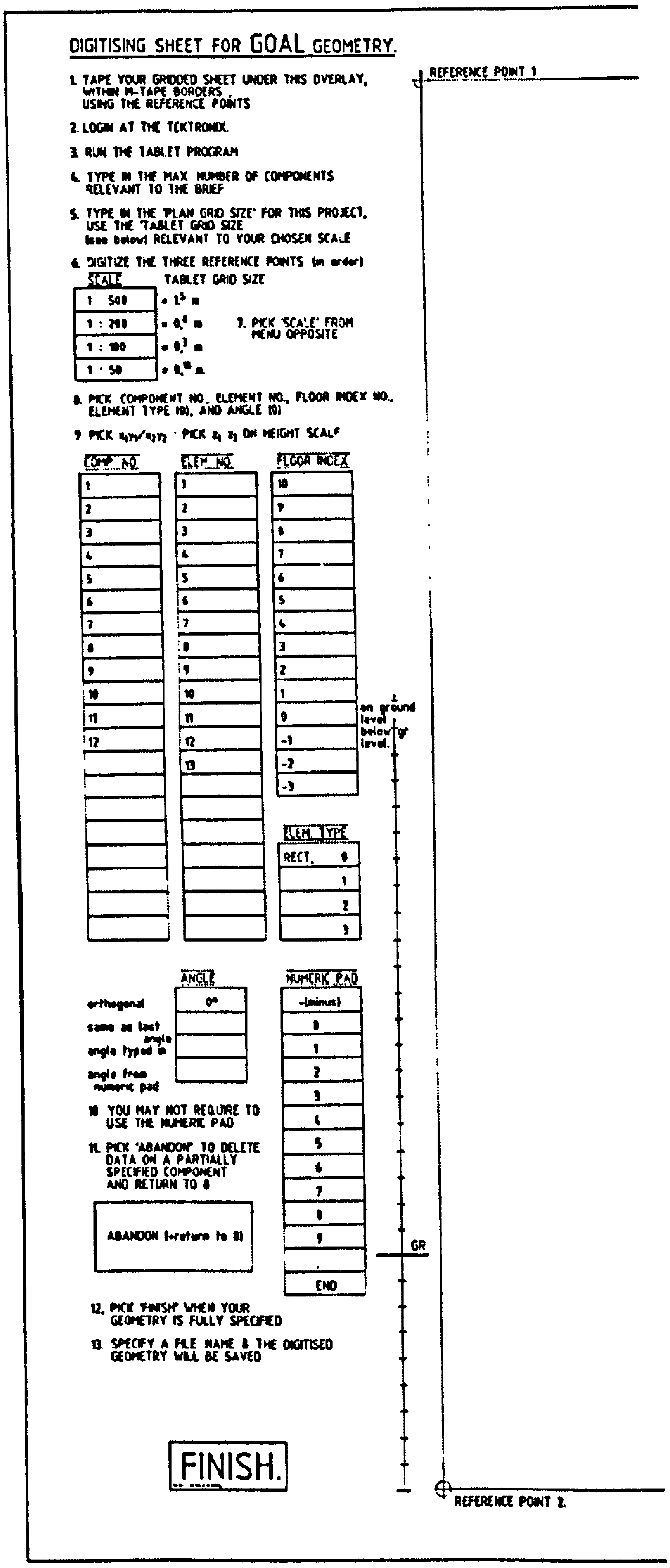


Figure 1.6 Geometry editing on a storage tube terminal

While the introduction of this functionality signalled a landmark in the development of interactive graphics the application was still limited by the technology employed. These limitations were primarily due to the speed of the round trip between a user's interaction and the processor completing the task and delivering a screen update. This was due in equal part to the restrictions imposed by time sharing, available processing power and the bandwidth of the interconnect. These constraints were ameliorated by designing the software so as to maximise the use of graphics while ensuring that no redundant draw commands required a refresh of the display. For example, if the user changed a geometric primitive depicting a room, perhaps by deleting it, then the update would be signalled by drawing a cross through the original entity. If it were to be moved, then only the new geometry would be required to be drawn, the old instance being deleted as in the above procedure.

Similarly the complexity of the geometry was restricted not just by these factors but also the availability of processing power and core memory. This reduced the description of a project to a simple plan view and any attempt to reproduce the third dimension was limited to the introduction of a floor and ceiling height.

As the programs evolved there was a need to provide more comprehensive interaction with the application and for a time it was popular to utilise a digitising tablet as the input mechanism. This allowed a greater physical area to be used for menu controls which in turned allowed more options to be displayed without having to page between screens with the attendant delay for a terminal refresh. Equally important was the ability to provide a more natural language textual description of menu functions. Figure 1.7 displays the tablet overlay for GOAL illustrating how the entire data capture procedure can be reduced to a sequential menu with associated description of functions. (Sussock 1981)



DIGITISING SHEET FOR GOAL GEOMETRY.

1. TAPE YOUR GRIDDED SHEET UNDER THIS OVERLAY, WITHIN M-TAPE BORDERS, USING THE REFERENCE POINTS
2. LOGIN AT THE TEKTRONIX.
3. RUN THE TABLET PROGRAM
4. TYPE IN THE MAX NUMBER OF COMPONENTS RELEVANT TO THE BRIEF
5. TYPE IN THE 'PLAN GRID SIZE' FOR THIS PROJECT. USE THE 'TABLET GRID SIZE (see below) RELEVANT TO YOUR CHOSEN SCALE
6. DIGITIZE THE THREE REFERENCE POINTS (in order)

SCALE	TABLET GRID SIZE
1 500	= 1.5 m
1 : 200	= 0.6 m
1 : 100	= 0.3 m
1 · 50	= 0.15 m

7. PICK 'SCALE' FROM MENU OPPOSITE
8. PICK COMPONENT NO., ELEMENT NO., FLOOR INDEX NO., ELEMENT TYPE (0), AND ANGLE (0)
9. PICK x_1y_1/x_2y_2 · PICK z_1 z_2 ON HEIGHT SCALE
10. YOU MAY NOT REQUIRE TO USE THE NUMERIC PAD
11. PICK 'ABANDON' TO DELETE DATA ON A PARTIALLY SPECIFIED COMPONENT AND RETURN TO 8
12. PICK 'FINISH' WHEN YOUR GEOMETRY IS FULLY SPECIFIED
13. SPECIFY A FILE NAME & THE DIGITISED GEOMETRY WILL BE SAVED

Figure 1.7 Tablet overlay, providing graphical data entry for GOAL

1.7 Experiential Output

During the period spanning the 1970s the main emphasis in the development of applications targeted at Computer Aided Architectural Design (CAAD) had been centred on appraisal, addressing the representation, measurement and numerical evaluation of built form layout and performance. The majority of appraisal packages output numerical measures associated with the economic, functional, spatial and environmental performance metrics of the design. Not withstanding the proven worth of this class of program, there was a growing acceptance of the fact that numerical descriptions of performance give the designer only a limited insight into the performance of their design. The significance of such criteria to non-designers, such as end users and the client, was even more doubtful.

Modest but growing attention began to be paid to devising a means for simulating the experience of the design rather than just enumerating it's characteristics. In an architectural context the most immediate and meaningful route to experiencing a design is through some form of graphical medium. Within this medium the degree of depicted realism is in direct proportion to the benefits accrued. Although the numerical basis and algorithmic design of computer programs capable of generating perspective views of three dimensional form had been known for some time previous, the cost and the lack of any elegance in interfacing to the machine had inhibited their use in any effective experiential appraisal.

Many of the restrictions facing developers have previously been mentioned but with the desire to desire to visualise 3-D geometry came a new challenge. Previously, computer generated perspectives had been produced as one of a number of forms of output from more general purpose appraisal packages, The ability to extrude 2-D floor plans by assigning a floor and ceiling height was relatively straight forward, to move beyond this and implement a method of defining generic geometry for an independent perspective drawing program was both novel and challenging.

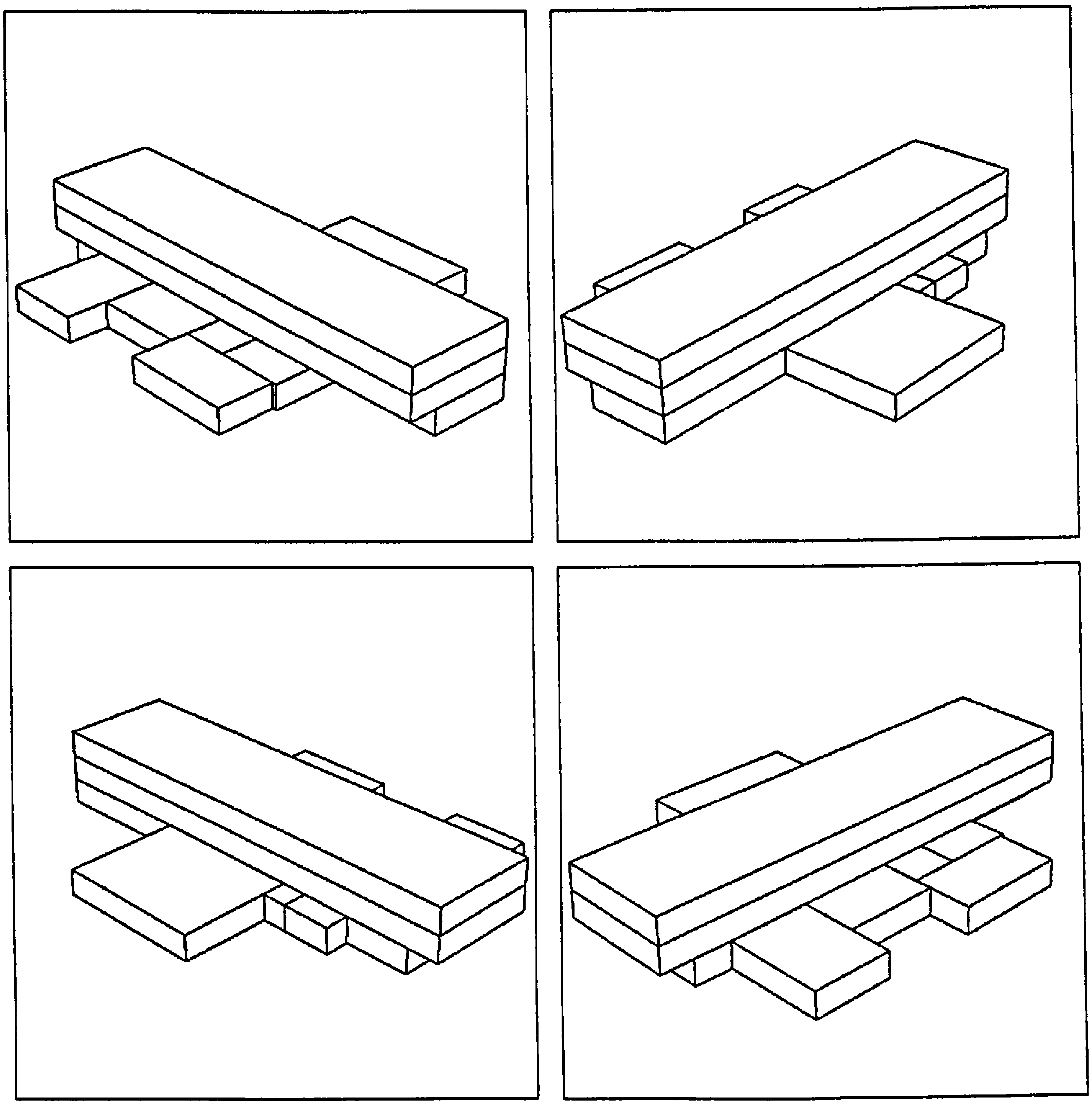


Figure 1.8 Automatic generation of perspective views from spatial data.

1.7.1 Defining Geometric Primitives

The technological constraints of the period lead to the adoption of a number of methods for defining geometry. In some instances it was avoided altogether, geometric quantities being defined purely as areas, volumes and adjacencies but where the 3-D form was critical a number of short hand definitions were devised. One such method was to formulate geometry through the use of FORTRAN callable subroutines that enable the user to define complex objects as combinations of basic shapes such as "cube", "sphere", “box” or "pyramid". Each subroutine represents a shape with parameters passed as arguments defining position, scaling and rotation. At run time this list of functions must be compiled and linked before the program could be executed. A cube might then be specified with the following arguments:

```
SUBROUTINE BOX(OX,OY,OZ,DX,DY,DZ,THETA)
```

```
where ox = the displacement of the object origin in the X axis
where oy = the displacement of the object origin in the Y axis
where oz = the displacement of the object origin in the Z axis
where dx = the dimension of the object along the X axis
where dy = the dimension of the object along the Y axis
where dz = the dimension of the object along the Z axis
where theta = the angular rotation around the Z axis
```

A typical sequence of commands might form the following code segment, generating the output shown in Figure 1.9.

```
CALL BOX(0.0,0.0,0.0, 1.0, 1.0, 1.0, 0.0)
CALL BOX(1.0,0.0,0.0, 1.0, 1.0, 1.0, 0.0)
CALL BOX(2.0,0.0,0.0, 1.0, 1.0, 1.0, 0.0)
CALL BOX(0.0,1.0,0.0, 1.0, 1.0, 1.0, 0.0)
CALL BOX(0.0,2.0,0.0, 1.0, 1.0, 1.0, 0.0)
CALL BOX(0.0,0.0,1.0, 1.0, 1.0, 1.0, 0.0)
CALL BOX(1.0,0.0,1.0, 1.0, 1.0, 1.0, 0.0)
CALL BOX(0.0,1.0,1.0, 1.0, 1.0, 1.0, 0.0)
CALL BOX(0.0,0.0,2.0, 1.0, 1.0, 1.0, 0.0)
```

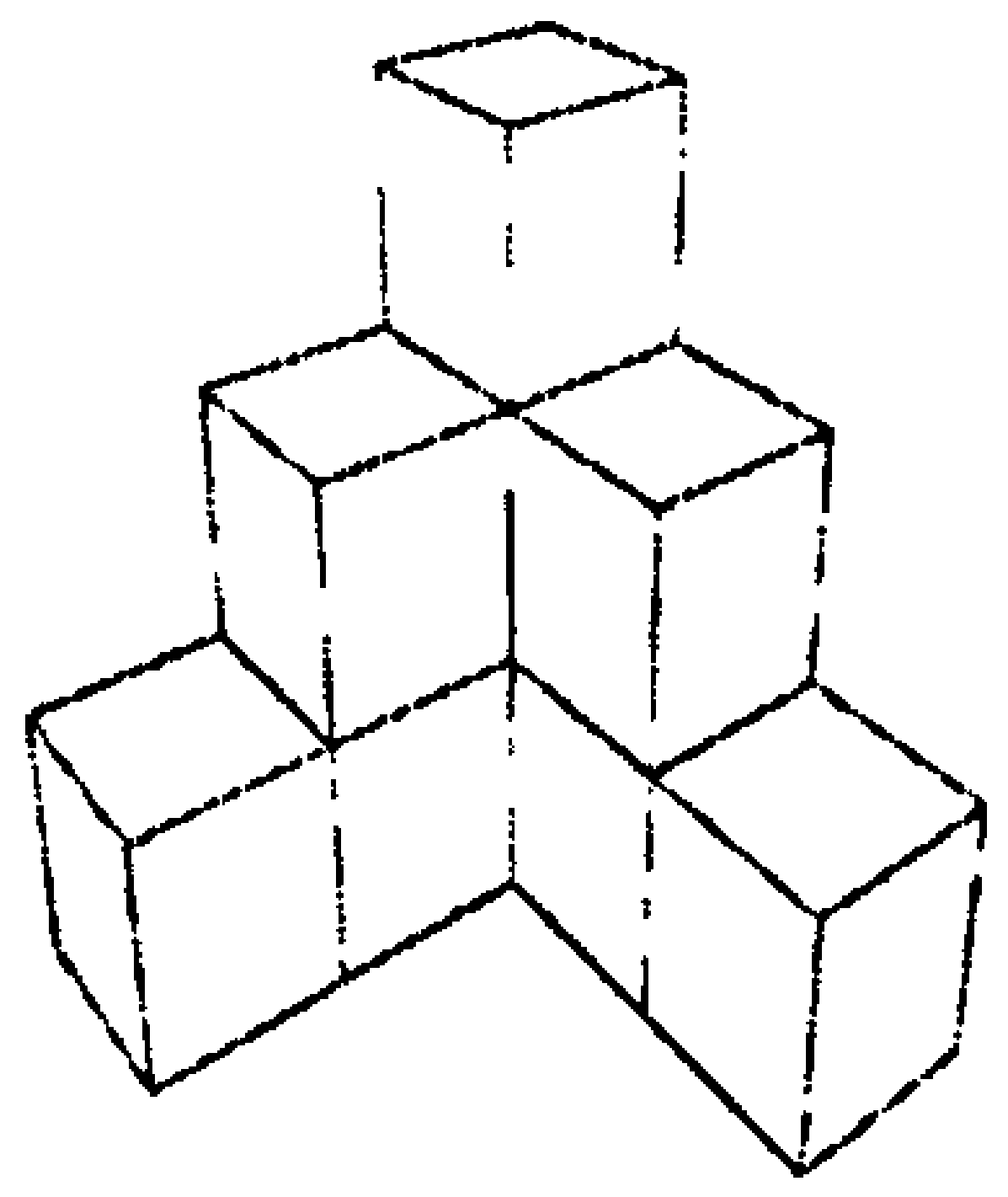


Figure 1.9 Computer generated perspective with hidden line removal

A more popular methodology was to parameterise a limited set of primitives that could then be defined with relative ease. For example spatial blocks may be described to the computer in a number of ways. In the case of right parallelepipeds (i.e. a regular shape, the opposite ends of which form three pairs of equal rectangles) some of the choices are:

- the Cartesian co-ordinates of two opposing vertices.
- the Cartesian co-ordinates of the origin vertex together with the length, width and height of the block.
- the polar co-ordinates of two opposing vertices.
- the polar co-ordinates of the origin vertex together with the length, width and height of the block.

Unless some descriptive format similar to the above were to be compiled within the executable code of a program, the software would typically look for an input file containing a description of the data. Patently, describing complex architectural form using only rectangular prisms would be both inefficient and unlikely to yield a convincing end product. As applications became more competent so did the range of descriptive formats employed.

With reference to ABACUS applications this data file consisted of an input stream within which the primitive entities (bodies) were flagged by a three character mnemonic followed by a description of their format. In general these entities were by default solid polyhedral shapes in that they were bounded by plane surfaces. Body topography could be either convex or concave and within the format distinction was made between regular and general bodies.

The data format for the rectangular (REC) block corresponds to a block specification at a given `x,y,z` position with an angular rotation defined positively from the x-axis. A regular (REG) body was a vertical prism comprised of two horizontal surfaces, a base and a roof, which are parallel and equal polygons and the remaining faces being vertical and joining corresponding edges of the base and roof. A general (GEN) body could be any shape, but must be completely bounded by planes. Curved shapes have to be approximated by polygons. General bodies may contain holes, but the holes must be properly bounded by planes as well.

```
REC                  REG                  GEN
0.0, 0.0, 0.0        4, 0, 1,             8, 6,
1.0, 1.0, 1.0        0.0, 0.0, 0.0,       0.0, 0.0, 0.0,
0.0                  1.0, 0.0, 0.0,       1.0, 0.0, 0.0,
                     1.0, 1.0, 0.0,       1.0, 1.0, 0.0,
                     0.0, 1.0, 0.0,       0.0, 1.0, 0.0,
                                          0.0, 0.0, 1.0,
                                          1.0, 0.0, 1.0,
                                          1.0, 1.0, 1.0,
                                          0.0, 1.0, 1.0,
                                          4, 1, 2, 6, 5
                                          4, 2, 3, 7, 8
                                          4, 3, 4, 3, 4
                                          4, 4, 8, 5, 1
                                          4, 1, 2, 3, 4
                                          4, 8, 7, 6, 5
```

Table 1.4 A single unit cube aligned with the co-ordinate axes

The above formats had the advantages of being terse, with attendant benefits in both storage requirements and transmission times, they were also human readable and easily created and edited. Inevitably there was a trade-off between the ease of creating primitives in this fashion and the complexity of the end product. In order to simplify the input procedure it became possible not just to feed the computer input data but also to enlist it's aid in defining that data.

1.7.2 Conversational Data Entry Programs

With the advent of time sharing computers accessed through a local teletype terminal interaction had improved considerably. However, local graphics terminals were still relatively rare and pen plotters were still the most favoured output device. Although this eased some of the problems mentioned above, live interaction and hence connection with the computer was critically expensive. The output of graphics to a local device meant that the terminal had to stay on line for some considerable time and connection charges via the GPO, CPU time on the remote machine and the accounting strategy for demand mode operations made this procedure expensive if not totally uneconomical. In spite of this, the ability to conduct a dialogue with the computer allowed developers to code conversational programs that made the correct definition of complex geometry a much more simple task.

The description of the topography and topology of shapes is, with minor variations universal, and in general only the syntax varies. Regardless of the means of input, most application programs will expect a complex shape to be formed by one or more bodies, where a body is a single construct defined by a set of Cartesian co-ordinates that define points and some definition of a list of lines formed by their terminal points that in turn form planes defined by the ordered lines around their boundary.

The accepted mode of modelling geometry was, and to a certain extent still is, to define a set of axes giving a local origin and rotation and then to take dimensions off plans, sections and elevations provided by a set of working drawings. Geometrical form is then represented by the construction of a number of bodies which can then be shaped, sized and oriented in the local co-ordinate system. For an inexperienced operator this method of defining "raw" geometry could be a daunting task. The use of a conversational data entry program could help alleviate these difficulties in two major ways. Firstly, by providing error checking on procedural definitions and input syntax and secondly adding the ability to instance previously defined bodies by name so that a number of bodies could be placed by specifying only the name, location, scale and angle of rotation.

1.7.3 Hidden Line Elimination

Naturally the purpose of defining three dimensional geometry is to prepare input for some set of routines that will allow the designer to visualise the proposed structure. One such program developed at this time was an implementation of one of the classic hidden line elimination algorithms. This methodology was originally developed in Italy, (Galimberti) passed on to Sweden, (Bergstein) before being acquired by ABACUS in exchange for another piece of software. Major modifications were made by ABACUS in an effort to provide an inexpensive, robust and useful design tool which could be used interactively on inexpensive graphics terminals.

In essence this category of program operated on a set of data which comprises the co-ordinates of the edges of planar surfaces of a limited number of polygons, in order to produce three point perspective views of objects, as specified by the user.

The program operates by first transforming all the geometry into eye space, that is a co-ordinate system with the eye point at the origin, then all bodies behind the eye are eliminated from the calculation. The core of the algorithm is dedicated to determining the inter-relationship of each edge defining each body with respect to all other edges of all other bodies. Firstly, self obscured body edges are eliminated, then the edge of each body is compared with the edges of every other body and any intersections are flagged. Next, a pass is made to find the surfaces, if any, that obscure each segment defined by the intersections. Finally the visible segments are identified and drawn.

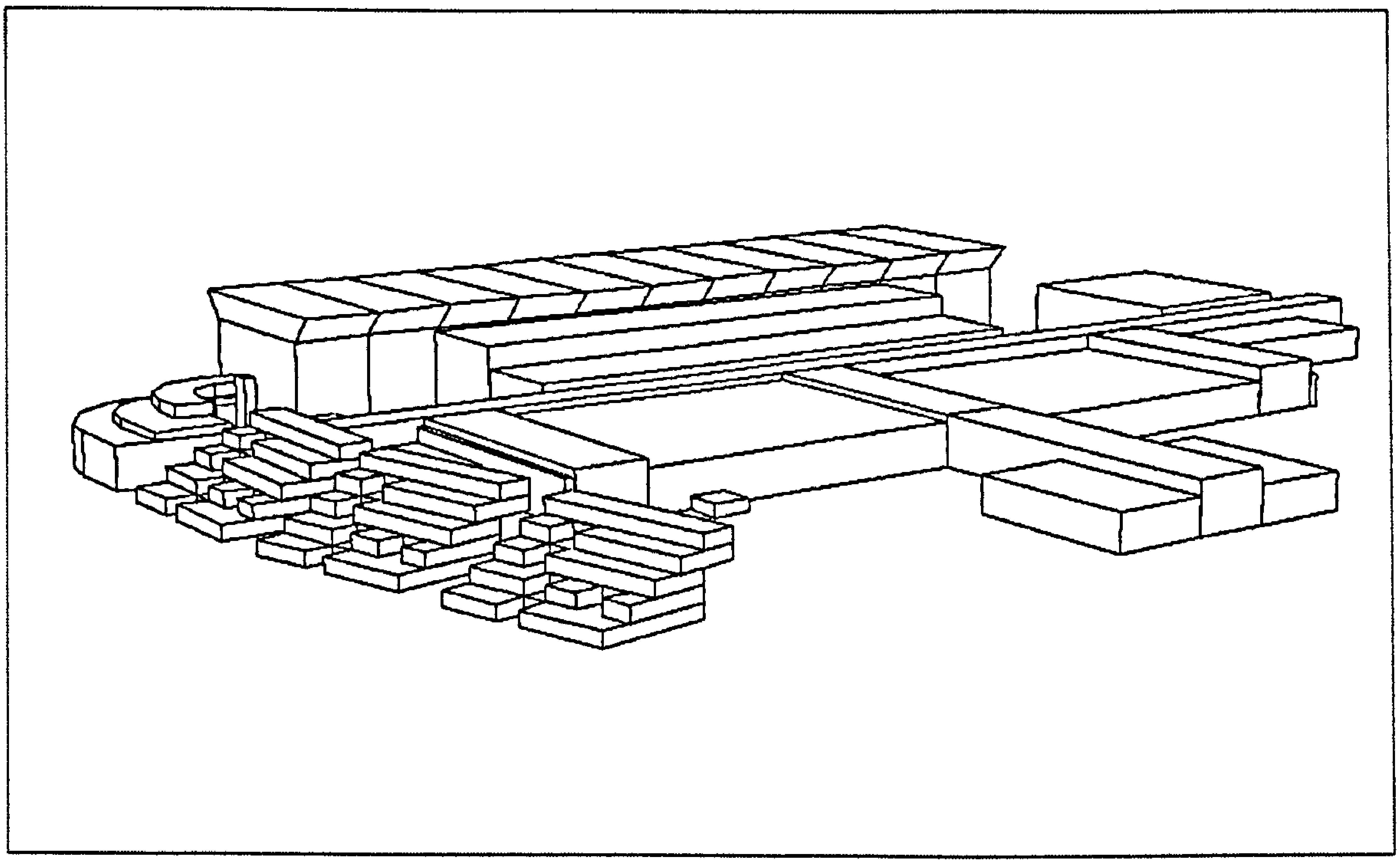


Figure 1.10 Hidden line perspective. The ship factory, student project 1977.

Program operation was very limited, having loaded the program, named a geometry input file and specified the type of terminal and line speed, the computer prompts by asking for four sets of parameters:

```
> EYE POINT/FOCUS POINT/VIEWCONE/DASHED
```

The appropriate response is to type in the `x,y` and `z` co-ordinates of the eye position, the `x,y` and `z` co-ordinates of the focus point, the tan of the half angle of the viewcone and either an integer flag depending on whether the hidden lines are not to be drawn are to be drawn dashed, or are to be drawn solid.

A typical response might be:

```
1.0, 1.0, 50.0,  50.0, 50.0, 0.0, 0.6, 0
```

In the case of the viewcone 45° = 0.4, 60° = 0.6, 90° = 1.0, 120° = 1.7

The view is then drawn on the terminal screen and may be dumped to local hard copy similar to the examples shown in Figure 1.10.

Since the cost of running programs within a time sharing environment was a function of how much core memory was required it was common to compile a number of versions of the same program with different array bounds. At run time the user would select the version with the smallest memory footprint that could accommodate the data. Typical maximum size limitations as applicable to a UNIVAC 1108 in 1973 are tabled below. (Sussock 1974)

Array Dimensions	LINK*L HLE	LINK*L HLE/BIG
Max. points	500	1300
Max. surfaces	350	1000
Max. edges	700	2000
Max. bodies	50	160
Max. vertices/surface	20	20
Max. vertices/body	50	100
Max. surfaces/body	50	150
Max. intersections/edge	50	100
Core Size	24k	44k

Table 1.5 Core size of HLE program

The cost per view of running a scheme using all of the 24k core was approximately 50p. [£3.28] The cost per view of running a scheme using all of the 44k would have been approximately £5.00. [£32.80]

It is telling that these programme remained the principle vehicle for defining and viewing geometry over the next decade! Due to the lack of adoption of visualisation techniques in their own right neither the scope of the primitives, the method in which they were defined or the operations available with which to manipulate them changed over this period. Enhancements that did appear related to global editing functions, implemented as stand alone utilities and operating on entire files, that allowed global translation, scaling and rotation.

1.8 Summary

It is now common to think of the progress of computing technology as being driven be parallel advances in hardware and software. During the 1970s the field of architectural computing was still too young for any evolution to be driven by hardware performance. However, there was a growing awareness of the benefits that could be attained through harnessing the graphical capabilities of the computer but computers of this era were remote, communication was second hand and graphical capabilities were almost non existent. Despite these difficulties progress was made with batch processing giving way to demand mode operation and alphanumeric "art" on a line printer being replaced by vector graphics on a storage tube.

With these capabilities came the desire to interact with a design, to experience its visual qualities all combined with the ability to operate these programs in a timely and efficient manner. These requirements called for revolutionary thinking not just to imagine an architectural application that could be harnessed to the power of a computer but also to devise ways of describing that process in a form that could be understood both by the user and the machine. The developmental problems were much greater when the researcher had not only to create an algorithm for drawing lines but also the logic of describing and defining geometry.

From these limited beginnings progress was made but at a much more limited pace than in today's computing environment. However the foundations laid during this period would serve to inform the future development of all aspects of architectural computer over the next three decades.

Chapter 2 The Glasgow Model

The decade of the 1970s saw the introduction of graphical computing into many aspects of the architectural design process. Not all of the innovations during this period proved to be lasting but within this overall context there would eventually be a finite number of threads that pointed the way to what would become the mainstream practice of the future. The computing environment in terms of processors, terminals and the current communications infrastructure would still be the limiting factors across all channels of endeavour but even by the end of this decade there were clear signs that the rate of technological change was going to increase exponentially.

2.1 The Technological Context

While faster processors and more memory might allow larger and more complex calculations it could also be seen that the move away from batch processing and towards more off-site terminal provision increased the availability of computing resources and served to initiate a self sustaining process where computing applications moved from the preserve of the laboratory or research environment into all corners of modern business practice.

Previously, funded research into architectural computing applications had tended to address issues that were more focused on investigating the general applicability of computing with regard to issues in building design. Now that many of the fundamental computing concepts had been addressed, and with a growing need for applications targeted at specific design questions, development turned to producing packages of programs that would harness these fundamental concepts and apply them to applications within specific areas. One such application area would be centred around the visualisation of the built environment.

By the early 1980s there was growing concern in many industrial nations as to the threat of irreversible change occurring to their traditional townscape and landscape environments. In Britain, despite some of the most sophisticated planning legislation, local communities were still concerned that fundamental changes would occur in the physical and visual quality of their environment and often suspected planning consent as not revealing the full impact of proposals on the community. Additionally the Royal Fine Art Commission had attacked architects for not considering how their buildings might integrate with the visual environment and had called for refusal of planning permission unless the new design is shown in the context of neighbouring buildings and features. Consequently, within the building profession there was a perceived need for a better means of addressing these concerns at an early design stage. (Bridges 1982)

Up till now all that was asked, or expected, from a CAD system was the ability to model and evaluate a single building, yet in a very short space of time it became possible to model ever larger groups of buildings. With this came the realisation that the quality of information content increases with the quantity. Given a group of buildings more evaluations become relevant, studies of blocking, massing, methods of access and egress all become apparent and the ability to view a larger context makes all these evaluations more telling. In order to place this period of application development it is profitable to review progress with regard to the capabilities of the hardware available over this period and the technical and administrative limitations imposed on their use.

2.2 Computing Technology

Although computing technology was continually evolving throughout the decade of the 1970s the rate of change was slow. The UNIVAC 1108, first used by ABACUS in 1972, was still in operation at the end of the decade and the same Tektronix vector storage tube terminals would still continue to provide a service until the mid 1980s. Although the original mini-computers, epitomised by machines like Data General's Nova and the DEC PDP-11 series, had made some inroads to the domain of the mainframe the lack of native processing power and restricted memory availability had still made the larger machines a more useful platform. This is not to say that powerful minis with ample memory were not becoming available but to configure a machine to a specification that would match the capabilities of a mainframe would price their availability beyond the reach of most potential users. The capital cost, administration and maintenance of a large scale multi-user facility meant that time sharing on institutional or national computing facilities still provided the cheapest access to a capable processing platform, but this was about to change.

2.2.1 Mini-Computers

While the mainframe continued through the '80s occupying a niche market by virtue of a range of unique features and the wealth of software targeted at that class of machine, the growing capabilities of the rest of the computing industry inevitably relegated them to a declining role over the longer term.

The dominance of the mini-computer in traditional computing environments was heralded by the arrival of the Digital Equipment Corporations (DEC) VAX - 11/750. This machine was the second evolution of the VAX product line and was to be the industry's first Large Scale Integration (LSI) 32-bit mini-computer. The price performance ratio of the minis now made them affordable at an institutional as opposed to a national level while still providing all the computational capabilities of their predecessors. The reason for this suddenly affordable hardware range lay in the development of microchip technology, bringing with it lower manufacturing costs and increased reliability.

The minis dominated the multi-user market for most of the early to mid 1980s despite being possessed of all the drawbacks of traditional mainframes in terms of the limitations of time sharing and serial line communications between terminals.

2.2.2 Personal Computing

In the late 1970s a new market sector could be seen to be evolving with the advent of the micro computer. During this decade computing began to evidence a popular appeal. While still too expensive for the home user, unless they built their own, more and more micros began to be marketed by a wide range of manufacturers. One of the first machines to have market impact was the Apple II. This was hailed as a landmark in personal computing mainly due to its innovative architecture and:

> *"its colour graphics, its extra high resolution graphics package, extreme portability and high quality construction" (PCW 1978)*

Of these features the capabilities of the colour graphics were hailed as the most impressive, even though being only capable of displaying 16 colours at a resolution of 280 x 160 pixels. This hardware was based around the 4.7MHz MOS 6502 processor and just 4Kb of RAM. The cost in 1978 was £1,250 [£4,339] excluding taxes and without the colour television required as a monitor.

The progression of this level of consumer computing power can be illustrated by tracking the development of the IBM personal computer (PC) series. In 1981 the first IBM PC was launched at a specification based around an Intel 8088 clocked at 4.77KHz, 16Kb of RAM and keyboard - all available at a price of £1265. [£2,998] A mono display cost an additional £345 [£818], and a matrix printer a further £755 [£1,789]

In 1983 IBM followed the success of the PC with an enhanced version designated the XT and retailing for around £5000 [£10,085] depending on specification. The basic machine included 128Kb of RAM, a 10Mb hard drive and 360Kb floppy drive. This was in turn followed by the IBM PC AT, first equipped with a 6MHz Intel 80286 processor, 1.2Mb 5.25 inch floppy and 256Kb of RAM. However completing the system with development software, enhanced graphics, a monitor and 20Mb hard drive pushed the price to over £9000 [£18,153]. IBM's graphics option, the Enhanced Graphics Adapter (EGA), offered up to 640 x 350 pixels in 16 colours. Processor performance was capable of delivering around 0.75 Million Instructions per Second (MIPS).

IBM were quickly established as market leaders in this field, despite competition from the "clones" which had been assimilating market share since 1982. However, IBM's development cycle could not keep pace with the introduction of new technology and by 1986 a purchaser would have to source hardware from a manufacturer like Compaq to obtain the latest components. This heralded a new problem concerning compatibility. The new Intel 80386 16 MHz processors had true 32 bit instruction paths but clone manufacturers were forced to maintain backward compatibility by emulating the 80286 protected mode features and their 16-bit instruction sets. Sadly this ended up by setting the PC standard for years to come and the overhead of providing backward compatibility stifled many innovative features.

In 1986 a consumer level Compaq DeskPro 386 with 16MHz 80386, 1Mb RAM, one 1.2Mb floppy and a 40Mb hard drive cost £5399 [£9350] without a monitor. Just over £2000 [£3460] increased disk capacity to 130Mb while a 4Mb RAM expansion was priced at £2695 [£4662].

2.2.3 Engineering Workstations

The mini computers gave way to a new class of desktop workstation, differentiated from the PC class by superior performance, high resolution graphics and a UNIX operating systems. Not suprisingly they also commanded a much higher entry price. A number of manufacturers competed in this arena notably SUN, Apollo and HP. In the UK domestic market Whitechapel made a brief appearance but were unable to compete over the long term. Engineering workstations such as these were typically driven by the Motorola 68000 processor series. Introduced in a 16-bit format in 1978 this series formed the basis of a number of computers including the Apple product line starting with the Lisa in 1983. Sun, formed in 1982, utilised the successor of this chip, the 68010 in the SUN2 series. Throughout the 1980s future Motorola generations provided the core of the SUN3 with the 68030 and the SUN4 based on the 68040. Within the engineering workstation market the end of the decade saw a move away from the use of processors with a Complex Instruction Set (CISC) towards the faster

Reduced Instruction Set (RISC) processors. Typical of this generation was the SUN SPARCstation, available in 1989. This was an attempt to lower the bottom end of the workstation market and align it with the top end of the PC market.

The SPARCstation 1 was built around a LSI-built SPARC RISC processor running at 20MHz, accompanied by a Weitech 3167 co-processor, 8Mb of RAM, expandable to a total of 16Mb, a 3.5 inch floppy and one 104Mb SCSI hard disk. The display ran at 1152 x 900 pixels and was available with 8-bit greyscale or colour graphics at a much reduced resolution and the option of 16, 17 or 19inch monitors. Configured with the accelerated colour display the SPARCstation cost £17,000 [£25,500] at a time when the average engineering workstation had been priced around £30000 [£45,000]. (All prices referenced from PCW 1998)

2.2.4 Operating Systems

The opportunities manifest in each of these new computing platforms was not apparent in just the provision of more capable and powerful hardware but also in the availability of novel graphical software and window management systems. In this arena the workstations prevailed, offering a much more capable environment throughout the 1980s. Within the personal computing class Apple lead the field from the start while Microsoft only released Windows 1.0 in 1985. Window managers for PCs were largely ignored until Windows 2.0 arrived in 1987 and paved the way for mass adoption in the next decade aided by the introduction of Intel's 486 and Microsoft Windows 3.0.

Within the workstation market development of graphics software was more complex with each manufacturer having their own proprietary standard. SUN initially provided SunView before finding a replacement with News and finally adopting the ubiquitous X-Windows environment. This was obviously problematic for developers who were forced to either adopt a single standard or spread their endeavours over a range of products. One consequence of this situation was the retention of the ability to emulate the existing Tektronix vector graphics standard. This had a number of advantages in that developers did not have to reimplement existing code to take advantage of the newer hardware and co-incidentally the graphics proved to be network transparent as the embedded graphics commands were implemented as ASCII escape codes. The major disadvantage was that the incentive to adopt more modern standards was reduced.

2.3 Strathclyde Hardware Resources

Despite the emergence of a new class of personal computers the inertia of legacy applications and the need to cater for a large and geographically diverse user base ensured that a centralised resource would continue to be the preferred option for the best part of the 1980s. This section details the nature of the computing environment at Strathclyde during that period and looks at the implications of the available machine base, the communications infrastructure and the management policies in place at that time. (Evans 1984).

2.3.1 Hardware Provision

Machine	VAX 11/782 VMS
Location	Strathclyde
Memory	8 Mbytes, Max program 5 Mbytes
Speed	2 (dual processor)

Machine	VAX 11/785 VMS
Location	Strathclyde
Memory	8 Mbytes, Max program 5 Mbytes
Speed	1.5 (single processor)

Machine	VAX 8600 VMS
Location	Strathclyde
Memory	16 Mbytes, Max program 5 Mbytes
Speed	3 (single processor)

Machine	PDP 11/44
Location	Strathclyde
Memory	1.5 Mbytes, Max program 29K words
Speed	0.1 (single processor)

Machine	VAX 11/750 UNIX 4.5
Location	ABACUS Strathclyde
Memory	6 Mbytes, 6 Mbytes (VM)
Speed	0.8 (single processor)

Machine	EMAS 2988
Location	ERCC Edinburgh
Memory	4 Mbytes, Max program 1 Mbytes
Speed	2.4 (dual processor)

Machine	CDC 7600
Location	UMRCC Manchester
Memory	64K SCM, 512K LCM 56K code, 186K data
Speed	10 (single processor)

Machine	Cyber 176
Location	UMRCC Manchester
Memory	128K SCM, 512K LCM 112K code, 186K data
Speed	10 (single processor)

Machine	Cyber 205
Location	UMRCC Manchester
Memory	2 Million Words VM
Speed	10 (vector processor)

Machine	Cray 1S
Location	ULCC London
Memory	1 Million Words 881152 Words
Speed	30 (vector processor)

Table 2.1 Hardware resources, 1985.

The above table describes the diversity of a typical university computing provision of this era. The three VAX machines forming a central cluster belonging to the institutional computer center, research groups operating a single machine and more specialised computers available at Regional Computer Centres.

The standard unit of measurement of speed was taken to be that of a VAX 11/780, about 1 MIP (Million Instructions per Second). A VAX 11/782 was essentially a dual processor system made up of two VAX 11/780 processors, therefor one 11/782 had a speed rating of 2. A VAX 11/785 was a single processor system which was approximately 50% faster than a VAX 11/780. It therefore had a speed rating of 1.5. These speed comparisons are only indicative and many other factors affected the real time performance.

```
VAX 11/782 Specification:

2 x Floating Point Processor
8 MByte memory
1 x 625 bpi Magnetic tape drive
1 x 600 lpm Line printer (with upper and lower case)
32 x Asynchronous 19.2kbps terminal ports (limited to 9.6k bps)
1 x Synchronous 48kbps port
1 x Computer Interconnect 70Mbps
```

Table 2.2 VAX 11/782 specification

All the VAX-VMS processor units were connected to two HSC50 disk controllers on a high speed serial bus, the Computer Interconnect (CI). A total of fourteen 456 Mbyte disc drives were connected to the controllers, of which three were dedicated to the various operating systems. This grouping formed a cluster and allowed the file store to be transparently shared between all machines. At this time, 1985, file space allocations were limited to half a megabyte for students and two megabytes for use by individual members of staff.

2.3.2 Operating System Restrictions

The operating system on VAX computers (VMS) permitted virtual memory addressing, this would allow programs with a large address space to be mapped into an area of physical memory without the programmer having to use more complicated overlay programming techniques. However if the size of the program was large in comparison to the available physical memory then as the size of the working set approached that of the peak page file set the amount of page faults causes a degradation of overall system performance. In order to share resources equitably a number of limits were established, these were set at a system level and could not be varied for individual users. Typically, these would be implemented as below:

```
Interactive CPU limit,        10 [CPU]       minutes.
Number of sub-processes,      2.
Working set default,          200 pages,     100Kbytes.
Working set extent,           1000 pages,    500Kbytes.
Working set quota,            500 pages,     250Kbytes.
Maximum program size,         10000 pages,   5Mbytes.
```

Table 2.3 VAX operating system limits

In practice the maximum program size was restricted to 1 Mbyte while programs above this limit, but below 5 Mbytes, could only be executed at the system operator's discretion. This meant that applications developers had to closely scrutinise the memory requirements of all program developments in order not to exceed the available machine assets. Typically, the OS would report resource usage as outlined below.

```
Buffered I/O count    136               Peak working set size     793
Direct I/O count      8747              Peak page file size       1593
Page faults           10408             Mounted Volumes           0
Charged CPU time      0.00:05:46.44     Elapsed time              0.00:39:23.09
```

Table 2.4 System resource limits

The purpose of monitoring system resources is to enable the programmer to optimise program execution within any given set of run time parameters. The two most influential indicators being memory consumption and CPU expenditure. The peak page file size returns the total size of the program in pages (512 bytes per page) while the peak working set size gives the maximum amount of memory that the program occupied during execution. This latter figure will normally be smaller than the size of the program since not all of the program is normally loaded into memory at any one time and many arrays may not be filled to their declared sizes. A page fault is generated whenever data has to be transferred from the page file to the working set. Large numbers of page faults indicate that the program is working at the limits of the available physical memory allocation and performance will be degraded as the disk is "thrashed" continually swapping pages from the disk store to RAM. Without the benefits of dynamic memory allocation it was common practice to compile different versions of programs, by varying array dimension declarations, in order to generate executables with differing memory requirements. This would give different versions of the same application each with a memory requirement that could be matched against available system resources, Since systems resource allocation varies dependant on whether the program is executed in interactive or batch mode this strategy would give the user the choice of running the "small" version at a terminal or submitting a "large" job to the batch queue.

A further layer of administrative restrictions governed the time that any one user could spend on the system. This limited both the time spent logged on to the machine as well as the amount of CPU time consumed by running jobs. CPU time is a measure of the amount of processing time execution of a program would take if the user had sole use of the system. Elapsed time is the "wall clock" time taken for program execution which in a multi-user environment, would always be greater than the CPU time by a factor dependant on current system load.

Typical system limits are shown below:

```
11:40:39   vaxa  users=27  scarcity=32  pre-emption=35  maximum=40   session=60
11:39:26   vaxb  users=34  scarcity=33  pre-emption=36  maximum=45   session=60
11:06:53   vaxc  down      scarcity=33  pre-emption=36  maximum=45   session=60
11:42:18   vaxd  users=27  scarcity=32  pre-emption=35  maximum=40   session=60
```

Table 2.5 System limits

When the number of interactive users exceeds "pre-emption" those users who have been logged on the longest are given a warning message that that they must log off within five minutes or they will be forcibly disconnected. The system attempts to reduce the numbers of users to "scarcity". Users who have been logged on less than "session" minutes would not be affected but at all times the number of interactive users could not exceed "maximum".

2.3.3 Batch Queues

In practice users could only utilise interactive programs if these programs used a minimal amount of CPU time, less than 10 minutes of CPU per session. Any jobs requiring greater computational resources had to be submitted to a batch queue, the range and availability of which are given below.

Queue Name	CPU Limit	Availability
SYS$BATCH	10mins	24hrs 7days
FAST$Q	3mins	24hrs 7days
SLOW$Q	30mins	17:00 - 08:00 Monday to Friday and all Weekend
LONG$Q	120mins	22:00 - 08:00 Monday to Friday and all Weekend
WEND$Q	600mins	22:00 Friday to 08:00 Monday

Table 2.6 Batch queues

The choice of batch queue was at the users discretion but if a process ran out of CPU time before completion then execution would be automatically terminated and all output lost. The queues were incremented after each job completed, if any specific job had not been processed before the time slot allocated to that queue expired then it would not be submitted until the next time that queue was activated. This could mean that at busy periods a job requiring the weekend queue (WEND) may have taken several weeks before completion. Once a user's job submission is completed, those jobs remaining in the queue submitted by that user would be returned to the back of the queue. This strategy prevented a single user from monopolising all the resources controlled by that queue.

2.3.4 Communications

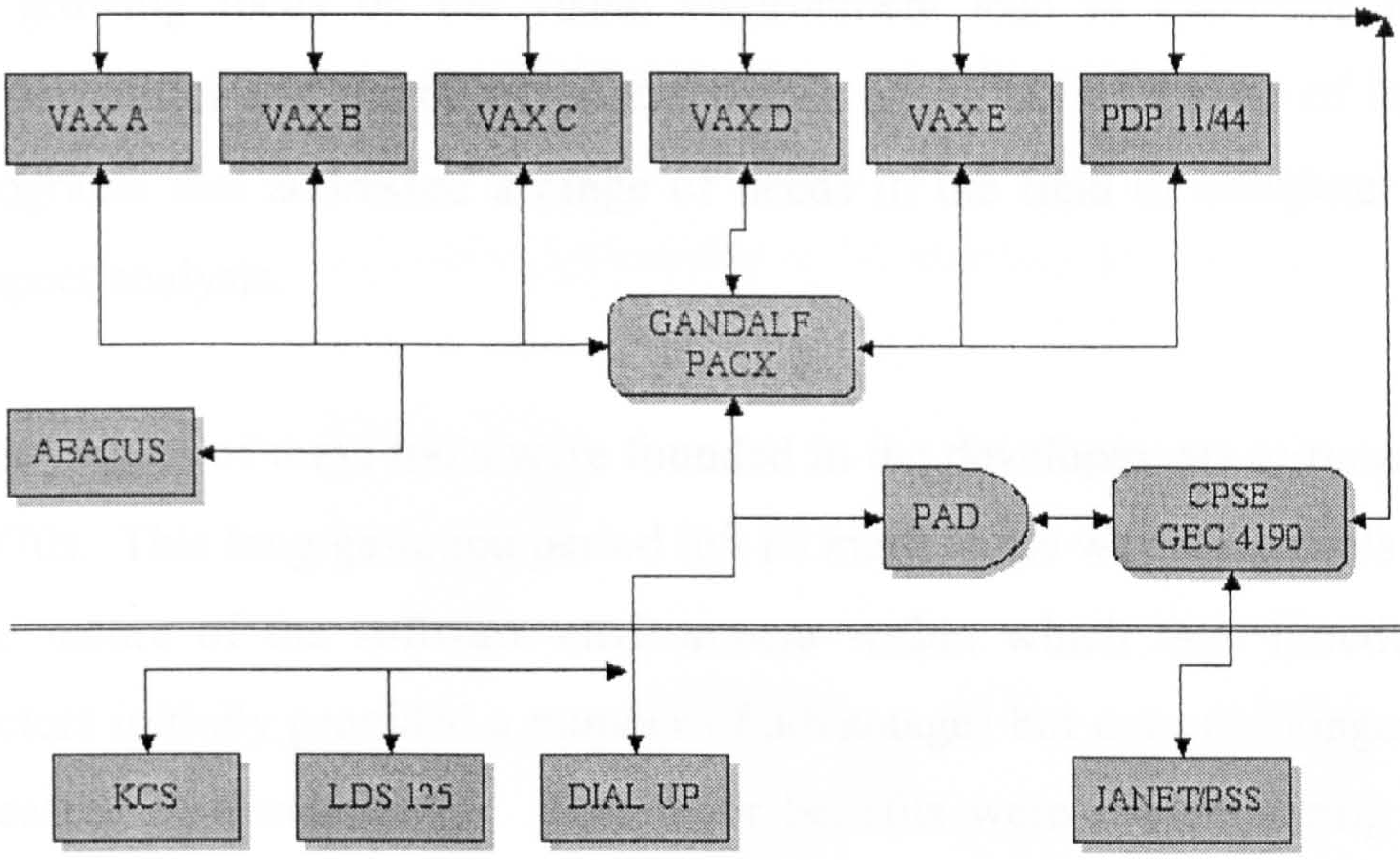


Figure 2.1 Communications layout

In order to link a user's terminals to the centralised computing resources four main methods of access were implemented. These are shown on the diagram above. All methods provided a route through either the Gandalf PACX (Private Automatic Computer Exchange) or the PAD, a terminal concentrator, which fed the CPSE. At a terminal, the user could connect via KCS (Keyboard Class Selection) or LDS 125 (Line Driver System) where the appropriate class connection would be selected by a TDU (Thumbwheel Dial Unit). In either instance the user would select the desired machine by specifying its class number, a two digit class code ie 12 specified the ABACUS 11/750. Access via the dial up modem facility was limited to 900 bps while the serial line connections operated at 9600 bps. A typical terminal configuration might be a VT100 for alphanumeric work or for graphics a Tektronix storage tube. Towards the mid 1980s more advanced terminals became available with the capability of alphanumeric and graphics modes using a minimum of local storage and processing.

2.4 Software Development Trends

A growing focus on the visual environment lead to improved R+D funding opportunities and the development within ABACUS of a suite of interconnected programs that addressed a range of needs in the field of computer aided visual impact analysis.

The origins of these tools were founded in the developments initiated in the early 1970s. This long gestation period left its mark in the way these tools operated and the nature of the software environment within which they functioned. These factors initially provided a number of advantages but over the longer term, also a measure of disadvantage. The major benefits were seeded during development within an era where processing power was finite and core memory was both limited and expensive. This lead to the development of algorithms that were both highly optimised and possessed of a small memory footprint. In turn these factors lead to the ability to handle large and complicated geometry sets in a timely and efficient manner.

On the downside, although the long development cycle had lead to the evolution of a set of robust and efficient tools, it also meant that it was harder to justify the input required to take advantage of new technological developments, such as the introduction of bitmapped displays and the more modern range of window managers with consolidated graphical user interfaces (GUI). However this would only prove to be a disadvantage over the longer term and for the project under discussion, the advantages proved to be a key enabling technology in allowing progress to be made.

2.5 The VIEWER Software Suite

The core of this software system is a computational engine that owes its origins to the HLE program described previously. This application evolved into BIBLE (Buildings with Invisible Back facing Lines Eliminated) (Parkins 1979) which was functionally identical but was enhanced by greater algorithmic optimisation and improvements to the user interface before it was in turn reincarnated with the more politically acceptable name of VIEWER. (Sussock 1985) As the need arose other modules were developed that either fed information to VIEWER or served to post process the output. Within the limits of the current technology, these modules combined to make a powerful and competent visualisation system.

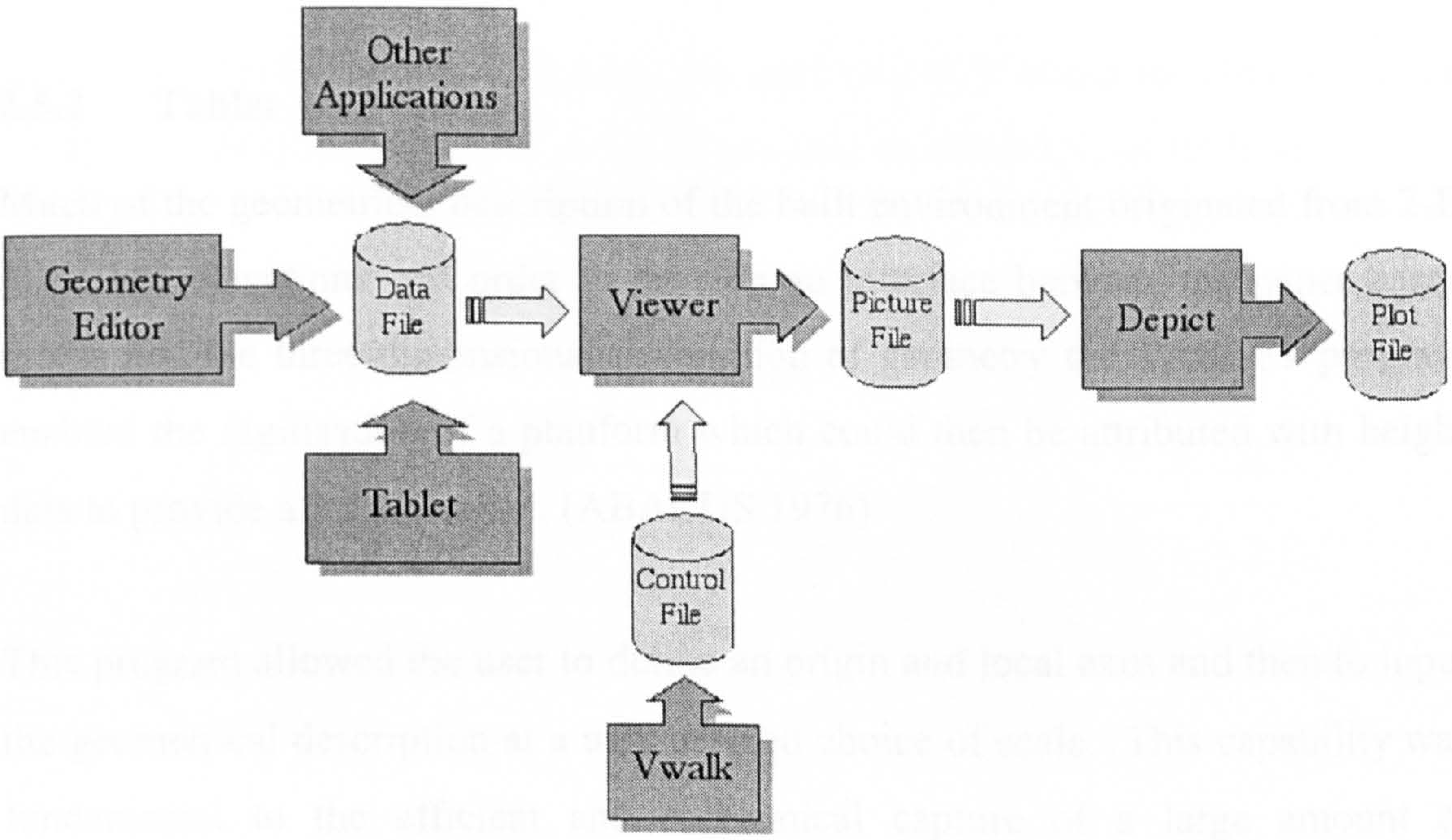


Figure 2.2 VIEWER software suite

2.5.1 Input

The input format for VIEWER was standardised to provide a common geometrical description for all ABACUS software. This input stream comprised a flat scene graph describing geometry in a range of formats although it could also contain a set of control parameters that dictated how the program would process a scene. Most of the input data consists of the geometry specification which is described as the conjunction of a series of bodies. Bodies are by default solid polyhedral shapes bounded by planar surfaces, and may be either convex or concave.

This input stream could be produced interactively using the geometry editor in conjunction with a digitising tablet or could be derived from the output of another application program. Since VIEWER's geometry format provided a ubiquitous lowest common denominator for the description of geometry it became common practice within other applications to provide the ability to write out their geometry and then utilize VIEWER for visualisation purposes. This provided the means to check the geometrical description for any errors and also to produce graphics for report generation and graphical analysis.

2.5.2 Tablet

Much of the geometrical description of the built environment originated from 2-D plans and elevations. In order to provide an interface between the paper based media and the three dimensional description of geometry the TABLET program enabled the digitisation of a planform which could then be attributed with height data to provide a 2.5-D format. (ABACUS 1976)

This program allowed the user to define an origin and local axes and then to input the geometrical description at a user defined choice of scale. This capability was fundamental to the efficient and economical capture of a large amount of geometrical data from cartography and similar paper based sources. The process was only capable of handling a limited range of options these being restricted to planar surfaces, extruded planforms and heighted line strings depicting landform contours.

2.5.3 The VIM Geometry Editor

VIM (Viewer Input Management) was a utility used to create and manipulate three dimensional objects, with the objective of producing geometrical data for VIEWER. (Stearn 1983)

Most 3D artefacts either have a primitive geometrical shape or can be composed by adding primitive shapes together. This program provided the basic tools to help the designer create the desired objects. Often there could be no unique solution as to how the objects may be assembled from some variety of component pieces and the operator had to arrive at a strategy which would combine efficiency with aesthetic merit.

The program options can be considered under four categories

INPUT	geometry creation
MANIPULATE	geometry editing
ATTRIBUTE	geometry attribution
CONTROL	geometry information

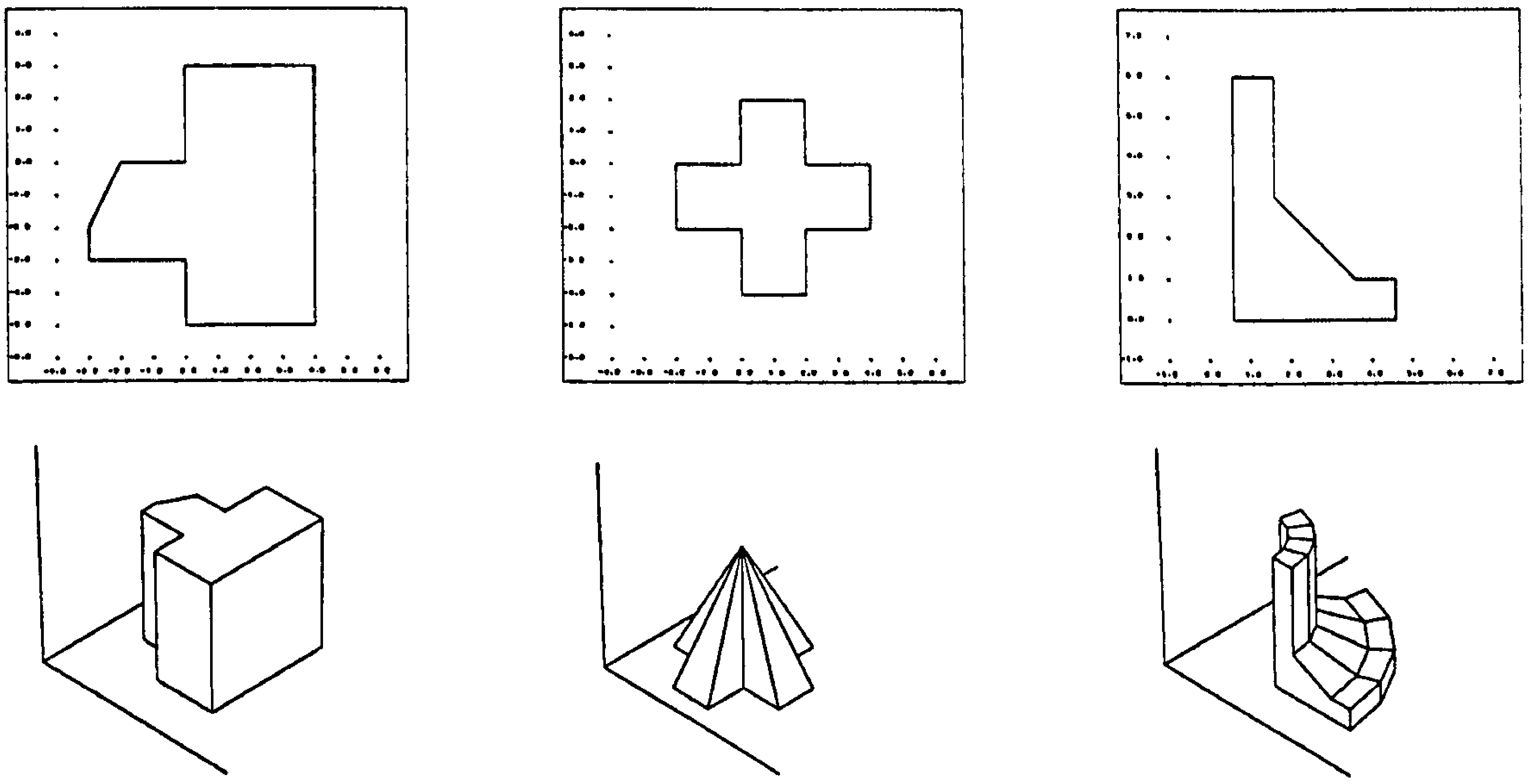
The user selects options by typing the appropriate 3 character mnemonic e.g. `"CYL"`, `"MOV"`, `"DRA"`. A list of possible commands is given below.

TOT	list totals	FIN	end program
NEW	read News file	AXO	define axonometric view
CID	colour identify	FLO	float Lines onto plane
SID	surface identify	REC	create Rectangular Prism
EXT	create Extruded shape	ELN	convert LIN to EXT shape
MEX	create EXT shape manually	REV	create Solid of Revolution
PYR	create Pyramid	SPH	create Sphere
HEM	create Hemisphere	CYL	create Cylinder
BOT	create Bottle	GEN	input GEN shape manually
VEW	input from file	MOV	move body
COP	copy body	ROT	rotate body
SCA	scale body	SPL	split database to file
GRE	graphical edit	EDI	edit body
DEL	delete body	GRO	group body
UNG	ungroup body	LIS	list database
BOD	list body	NAM	list names in database
REN	rename body	DUM	dump database to file
DRA	draw bodies	IDE	identify body
SEL	select bodies from database	GLA	set body to Glass
WIN	define Window	TIL	create Tile
UND	undelete body	GID	identify Group
VER	create vertical planes	CON	convert body
OBO	create Offset Bottle	OEX	create Offset Extruded
COL	set Colour	MAP	define Colour Map
COU	count bodies in file	EED	extruded edit
PED	graphical edit of plane	LIN	create Line
VIS	set Visibility	VOL	volume of body
TEM	define Template	PAT	read path file
PIT	create Pitched roof	PLO	output plot file
INV	invert body	SUR	create SUR object manually
NBO	redefine Bottle shape	LAY	set Layer
FAC	transform facade onto face	GRM	graphical move
GRW	scale object about itself	EXP	move body by scaling factor
HOL	merge Ext as Hole in EXT	ANG	define angle of two points

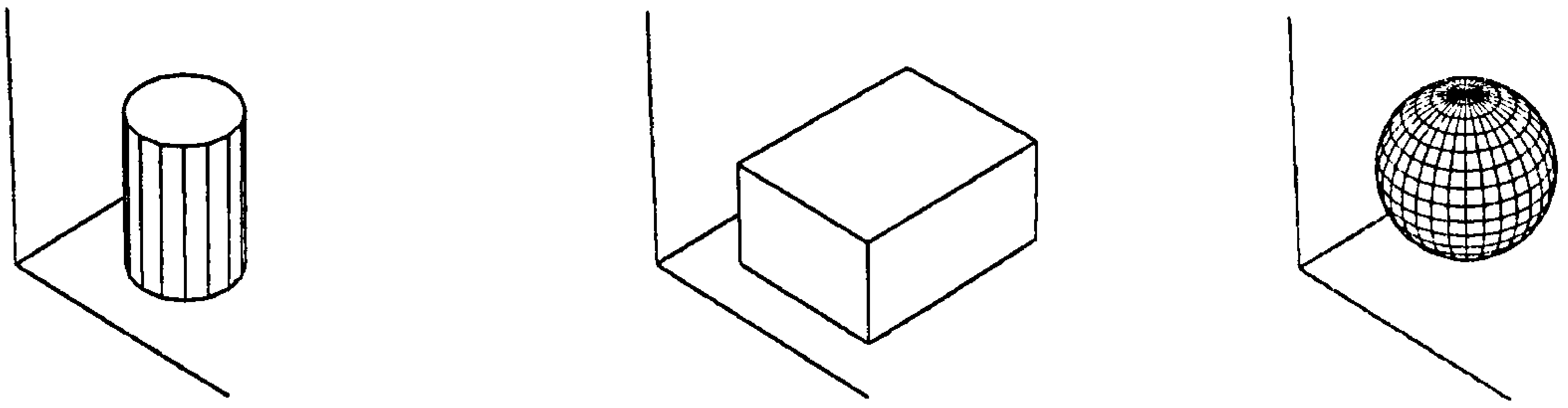
Figure 2.3 VIM geometry editor command set

2.5.3.1 Geometry Creation

Geometry could either be read in from a pre existing file or created directly within the program. The majority of body types were created by defining a section in a two dimensional co-ordinate system and then either extruding parallel to one of the co-ordinate axes or by sweeping the section around a given axis. Other configurations that could not be completed in this manner had their own interface conventions, for example spheres and pyramids. Despite the limitations of this interface there proved to be few objects that could not be formed either by these simple operations or by some combination of separate primitives.



Extrusion Pyramid Solid of Revolution



Cylinder Rectangle Sphere

Figure 2.4 VIM geometry primitives

2.5.3.2 Geometry Manipulation

Due to the limitations of Tektronic's graphics on a storage tube terminal the biggest obstacle to easy editing was the lack of any ability to graphically interact with the program in order to manipulate geometry. This was overcome, in part, by the ability to copy, move, scale and rotate primitives by typing in the relevant commands but the complexity of the procedure meant that the user had always to be aware of the exact consequences of the operation.

2.5.4 VIEWER

VIEWER is the core module of a suite of programs that allow a user to perform visual appraisal of a designed object or some group of objects. These objects may be limited to a single building, modelled in some detail or a larger group of buildings depicted with less data. The program can produce orthogonal, parallel or perspective projections of the scene from any chosen viewpoint. It also provides facilities for generating perspective views that may be used in conjunction with site photographs to form photomontages of proposed developments. The images produced by the program are of a form known by the generic term as "wire-line" in that the edges of objects are drawn as vectors. Lines that should be hidden when the scene is viewed from certain eye points may be left in, drawn dashed or removed altogether.

2.5.4.1 VIEWER Operation

The user interface of VIEWER had been designed to be as easy for an inexperienced user as was possible, without sacrificing the capability to specify arbitrary projections. This was achieved through the integration of four key features.

Fail-soft design : any invalid or wrong input produces an English language error explaining the expected parameters and the nature of the options available to the user at that point. The user could always ask for help which in turn made the program self documenting to an extent. It was intended that the program should be impossible to crash by specifying invalid data.

Graphical specification of the view point : This allows the user to point on a displayed plan view to a place in the vicinity of the object and, in effect, ask "What does it look like from here?" This avoids the often abortive results of having to choose view parameters by specifying co-ordinates.

Use of default parameters : all the view control parameters have default values computed by the program for each data set as it is read in. Therefor it is possible to simply read the data file and immediately call for a view and get a meaningful image.

A command menu on the screen gives the user a visible reminder of what he can do without the need for conversational prompts, more detailed information can be obtained by selecting the help option. A sample menu and the output of the help facility are given below.

VIEWER	
/INPUT	?HELP
EYEP	NEWFIL
FOC&MP	TYPALL
MIDP	4PAVER
HIDDEN	3PAHOR
DASHED	2PERSP
VISIBL	1ORTHO
ANGLEV	SCREEN
LENSFL	OUTFIL
*ENLRG	BYEBYE
INPALL	%STERE
+DEFV	.LAYER

```
?HELP  to list command options available
BYEBYE to exit
/INPUT to TYPE IN numerical values of E F or M
EYEP   to pick the Eye-point with cursor
FOC&MP to pick the Focus and Mid-points together
MIDP   to pick the Mid-point alone
HIDDEN to remove hidden lines
DASHED to draw them dashed
VISIBL to draw them solid
LENSFL to specify Lens focal length (in mm)
*ENLRG to specify Enlargement factor (neg. to print)
ANGLEV to specify viewcone angle (degrees)
          (negative for biggest picture that will fit)
1ORTHO to select orthogonal projection
2PERSP to select perspective projection
3PAHOR to select parallel proj. onto horiz. plane
4PAVER to select parallel proj. onto vert. plane
TYPALL to list all current view parameters
INPALL to type in all view parameters
SCREEN to draw the picture on the terminal screen
CPLOT  to draw/plot it on remote plotter
PLOTL  to draw it on local TEK plotter
OUTFIL to file it in a VIEWER picture-file
NEWFIL to select a new VIEWER input geometry file
+DEFV  to select default views
%STERE to select stereo view option
```

Figure 2.5 VIEWER interface

2.5.4.2 Logical Constraints

As with all descriptions of geometry there are a range of constraints that differentiate between valid and invalid input data. For example any surface description entered with one or more co-incident vertices in a polygon would, by defeating the logic of the algorithm, produce visual anomalies when the scene was processed. In the case of VIEWER it was decided that an anti-clockwise definition of vertex order, in the definition of surfaces, would result in a surface normal that would point at the towards the eye when viewed from the outside of a body. If the definition is reversed each such body would appear to be drawn inside out as shown below.

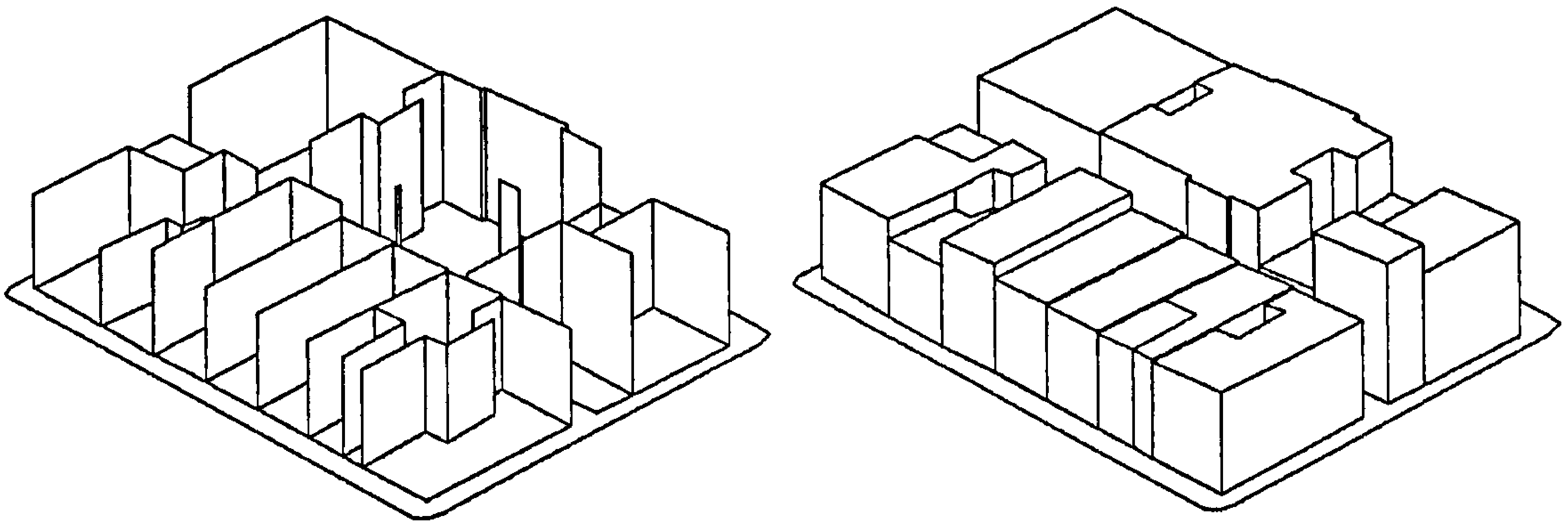


Figure 2.6 Visual consequences of defining geometry in reverse order

By far the most limiting constraint was the need for all bodies to avoid intersecting with each other. This was due to the logic of the hidden line removal algorithm which expected to only test for intersections between pairs of edges and not edge to surface interaction. This is illustrated in the example below where the two intersecting cubes shown on the right are visualised with the result given on the left. The omission of edges which should be visible is the most common outcome although as the orientation of the object is varied and the ordering of occluding surfaces changes then the outcome is undefined.

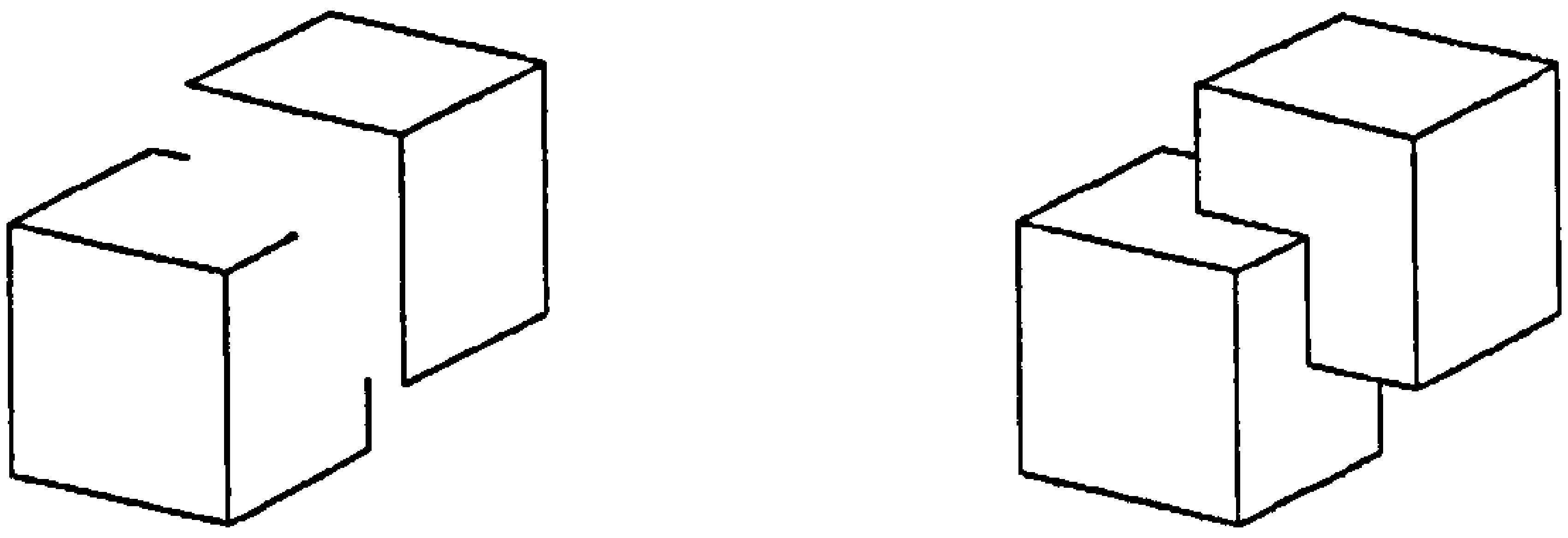


Figure 2.7 Visual consequences of inter-penetrating geometry

Within the composition of an object definition bodies may abut although in some cases this could also defeat the logic as the program attempts to separate edges by applying a small shrinkage factor to each body before calculating hidden line removal. This could produce unexpected results from apparently legal geometry when a concavity in one body is filled by another, considerably shorter, body.

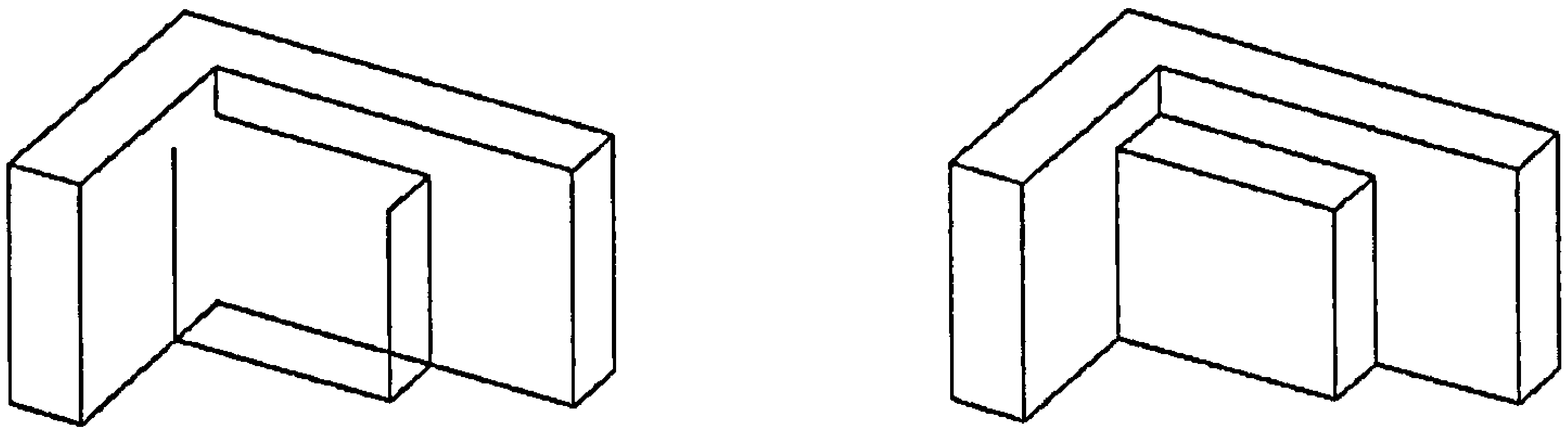


Figure 2.8 Visual consequences of nesting concave geometry

These constraints only affect the determination of hidden lines, when all lines are plotted there are no problems with visualising the geometry of abutting or intersecting bodies.

Failure to conform to any of the above constraints would not necessarily be detected other than in the production of an image with obvious artifacts. However the discipline required to adhere to the rules for avoiding these limitations generally required a greater degree of thought and planning as to how best to model a design which in turn resulted in data-sets which were smaller, cleaner and processed faster. These efficiencies proved fundamental to the success of processing the growing size of data sets.

2.5.4.3 Output Commands

Once the viewing parameters had been set and the display mode defined the program could be commanded to deliver the processed scene in a range of formats.

The SCREEN command tells VIEWER to draw the view on the Tektronix screen. First the screen is erased and then the view is displayed until the user pages the display to return to the menu. If this option is selected in alphanumeric mode i.e. when operated from a teletype, an error message is displayed. Running a graphical program in alphnumeric mode is not as paradoxical as it may first seem. Additional functionality was provided to store the processed view in an intermediate file format such as by writing to a plotfile on the host which may then process it directly or append it to a queue for subsequent plotting. This was essential when running in batch mode as all commands were regulated by a command file and any form of interactive graphical interaction could not be supported. The use of an intermediate file type meant that the procedure of input processing could be separated from output processing giving greater flexibility in the use of CPU time and queue submission.

If immediate results were required then a local plotter, connected in series with the terminal, could be activated by the PLOTL command which sent appropriate escape codes to the device. A final option allows the user to copy the view defined by the currently used geometry with the current parameter settings to a 2-D picture file which could then be used as input to the DEPICT program.

2.5.5 Output

A final utility, DEPICTt, was provided to post process VIEWER's output file format and prepare the data for eventual plotting. (Stream 1984)

The most common hard copy device of the period was still a drum, or flat bed plotter, which transcribed the vector output to a series of lines drawn with an ink pen on a paper based medium. This process incurred the requirement to support a range of devices by providing instructions in a vendor specific format. By storing the output in an intermediate format an additional advantaged was enjoyed through the ability to generate a number of plots at differing scales from a single VIEWER run.

There was also a fundamental need to work within the limitations of the medium. A common problem was manifest in all ink pen and paper combinations. This could be observed in a number of ways. Due to the process of perspective transformation in the viewing frustum reverse depth cueing occurs from the greater density of lines drawn to the back of a scene. This inverts the normal visual clues where an observer would perceive the decrease of detail as the vista retreats into the background. In this instance the opposite is true as the scene becomes darker the further it recedes. Coupled with this anomaly is the pragmatic problem of the paper becoming saturated with ink in areas of high line density, at its limits the paper could become so weakened that the pen might drop through the surface.

In order to counteract these problems DEPICT had the ability to filter the picture file and remove all vectors beyond a user specified tolerance. Not only did this alleviate some of these problems but also it served to reduce the plot time by discarded lines of little consequence.

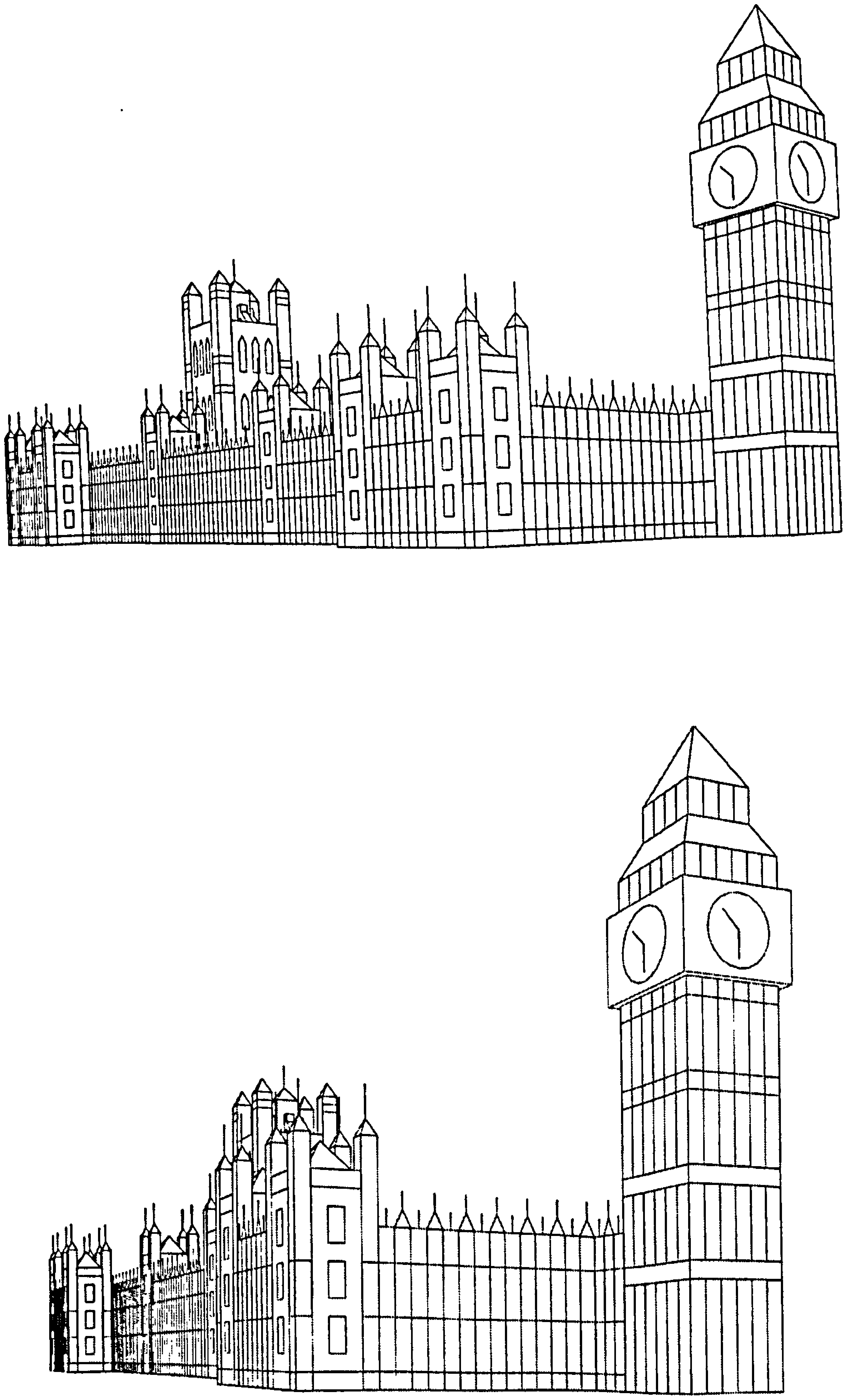


Figure 2.9 Comparison of output showing reverse depth cue effect

2.5.6 Command File Control and Animation

As VIEWER originated in an era when batch processing was the normal mode of operation it maintained the capability to operate in a non graphical and non interactive mode. This required that all input parameters be supplied from a control file and all output be written to an intermediate output file. This operational configuration meant that a series of views could be generated and held in a single output sequence. Once the computationally expensive task of hidden line elimination had been performed the processed views could be quickly regenerated from the 2D data file and either plotted as a sequence or captured to 8mm film.

This rudimentary animation procedure could be used to generate simple walk through sequences. To aid the process a utility called VWALK was provided which took as input a series of control points, these being eye and focus co-ordinates, which coupled with a choice of motion between subsequent points would generate a sequence of parameters as a function of velocity and frame rate that would then be passed as input to VIEWER. The image below shows a diagrammatic representation of the eye point being moved around a model. The square represents the eye position at each frame, the arrow indicating the associated view vector. The photograph shows a stop frame camera capturing sequences of frames from a Pericom graphics terminal. (Shamwana 1986)

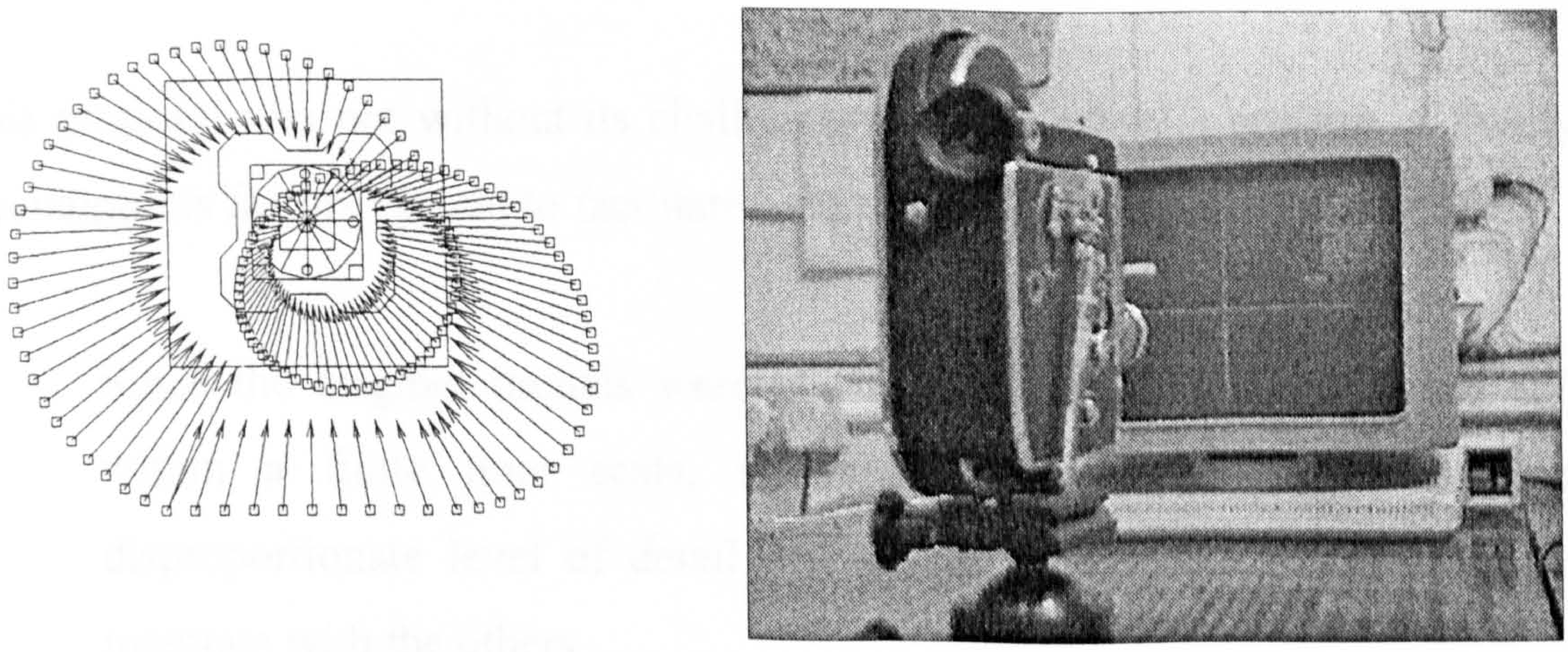


Figure 2.10 Laying off an animation sequence to 8mm film

2.6 The Strathclyde University Campus Model

Initially, the data set that was destined to become the Glasgow Model consisted of a number of individual buildings on the Strathclyde University Campus. These buildings were being modelled as part of an ongoing, annual, student exercise. The choice of this set of buildings had been made largely due to the availability of data and ease of access to the originals. This availability was an important factor because data capture, in the real world, would often be a considerable order of magnitude of difficulty in comparison with modelling "virtual" buildings. After the course described above had been running for some time it was apparent that this existing data set now comprised most of the buildings on Campus. Each of the buildings had been modelled in isolation, this being a limitation of time and the processing power of the then available computer environment. However with the continual upgrade of the University's central computing resource it became possible to consider conglomerating these individual models into a single data set.

As individual buildings these models lacked any sense of place, so while they might be considered worthy representations in their own right, it was also readily apparent that, should it be possible to work with the entire data set, many more evaluations would become possible, such as an appreciation of access, egress and blocking or massing. Potentially this amalgamation would create an entity whose worth would be more than the sum of its parts.

This proposal was not without its challenges but fortunately a number of purely serendipitous factors existed to facilitate the project. These are outlined below:

- Since the original models were constructed with limited resources and within a finite time scale, no one individual building boasted a disproportionate level of detail that would have made it impossible to integrate with the others.

- Most of the on campus building stock consists of individual properties that stand in relative isolation. This made the task of integrating data easier as opposed to the situation that would have existed had, say, the subjects been portions of individual buildings.

- The campus is situated on top of a steep rise, this meant that the buildings naturally took up the land form giving them a position regarding relative elevation as well as a two dimensional location in plan.

- The size and area of the campus is relatively limited, allowing all the campus buildings to be usefully considered without proving too large for the current computing environment or too small to give an impression of a community space.

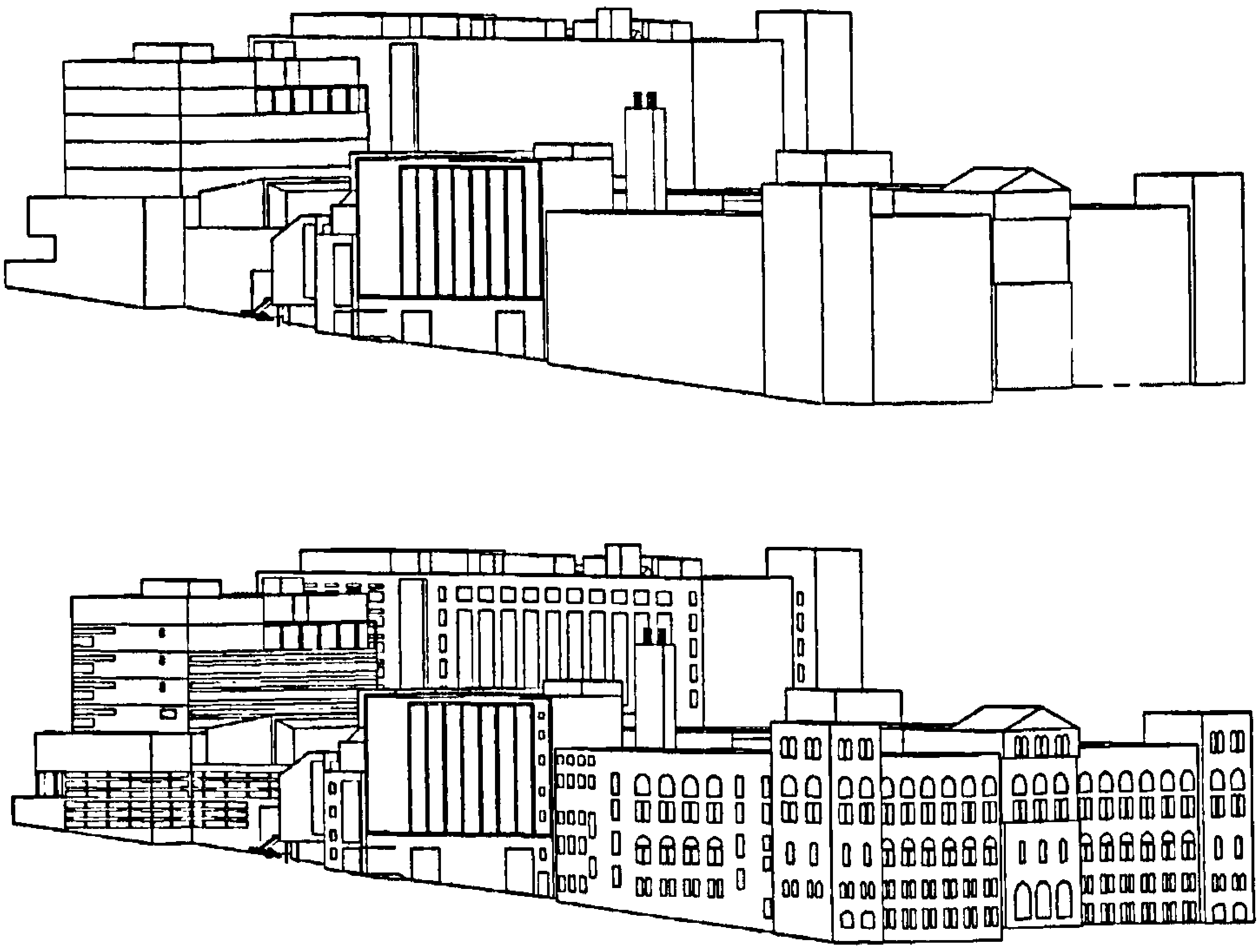


Figure 2.11 Basic and detailed building models.

The development of the Campus Model was not driven by any one predetermined set of circumstances or strategy. Progress was marked by the availability of new hardware, newly introduced features in the suite of software used in its construction or even the addition of new data. This has meant that the structure and format of this area of the model was continually in a fluid state that suited this particular form of evolution. This proved to be a distinct advantage in that the model could exist in different forms, from the most basic of building to more detailed and realistic representations as shown above. With this facet of the project being revisited on an annual basis this iterative design approach gave increased insights into the requirements and potential of a larger scale model.

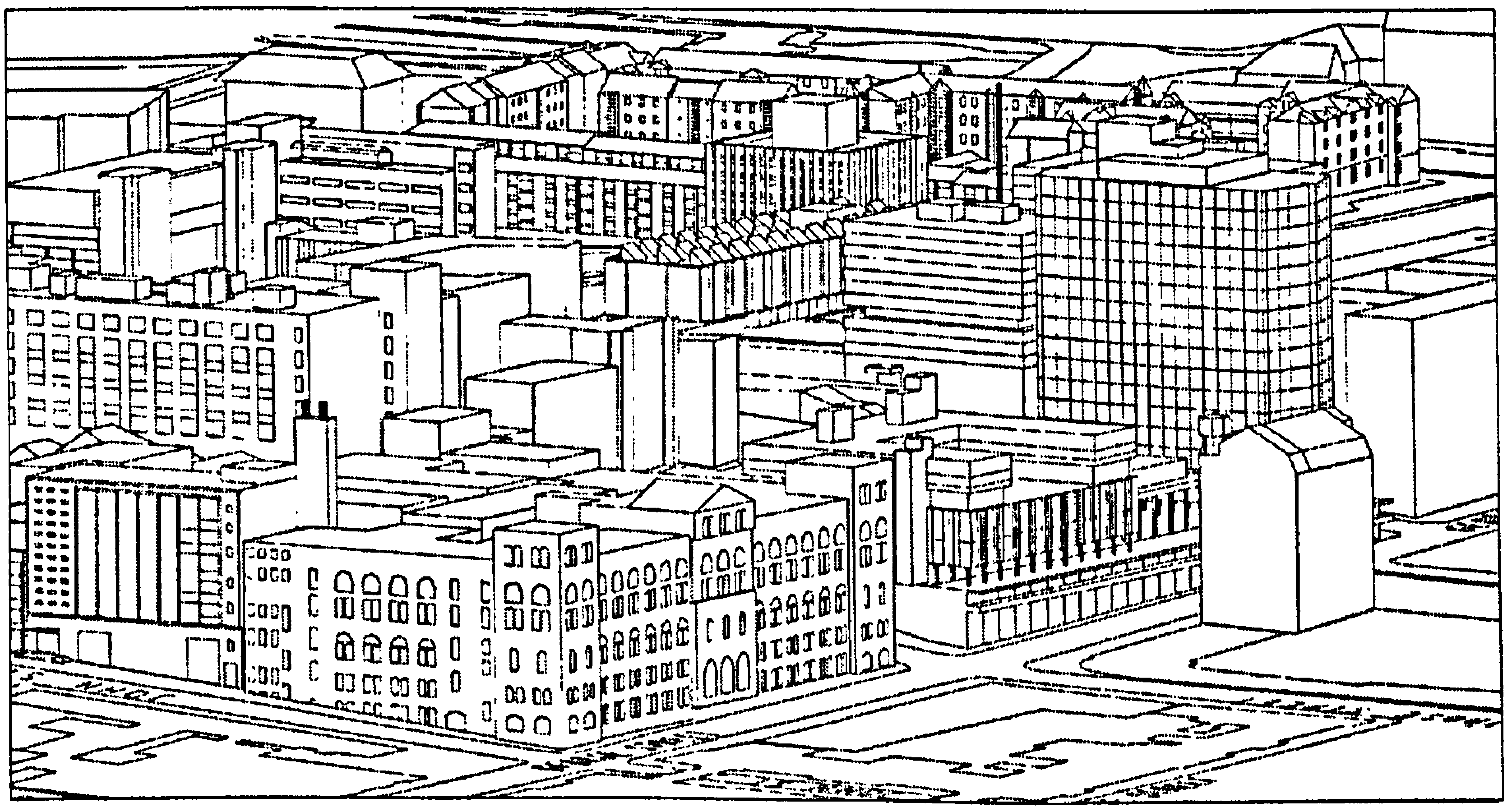


Figure 2.12 Strathclyde campus model

The detailed versions of each building were based around their more primitive counterparts. This allowed for the substitution of the one format for the other as needs arose. Architectural detail is applied by first refining the three dimensional form of the basic building. At this stage the correct choice of the constructional methodology can provide much of the required characterisation, as discussed previously. Most remaining features were then provided by the addition of line primitives. The choice of a geometrical line primitive was made as these were faster and less problematic to process than either planar polygons or additional

]

3-D elements. The examples below contrast the existing buildings with their computer modelled counterparts.

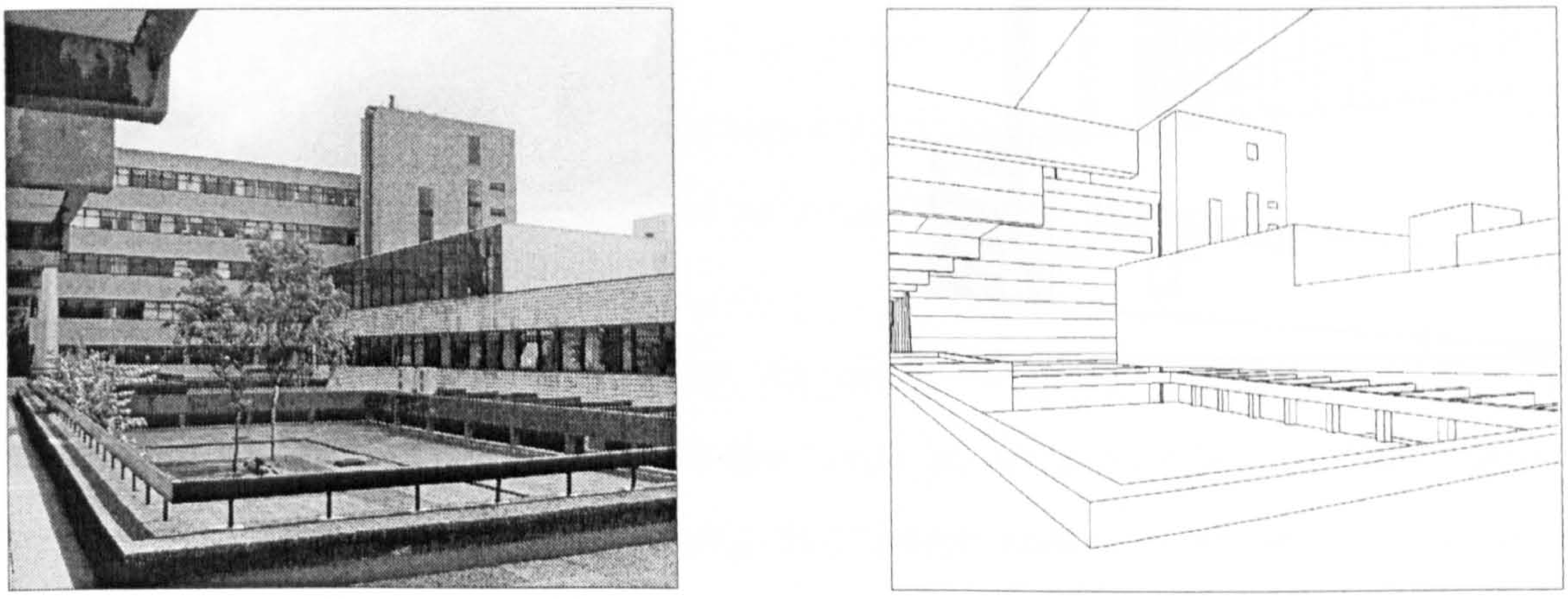


Figure 2.13 The Architecture building

The tight coupling of some buildings with their immediate landscape leads to some instances where the building and their surroundings were modelled as a contiguous units as below. This helped to generate the impression that the discrete buildings were part of a unified conglomeration.

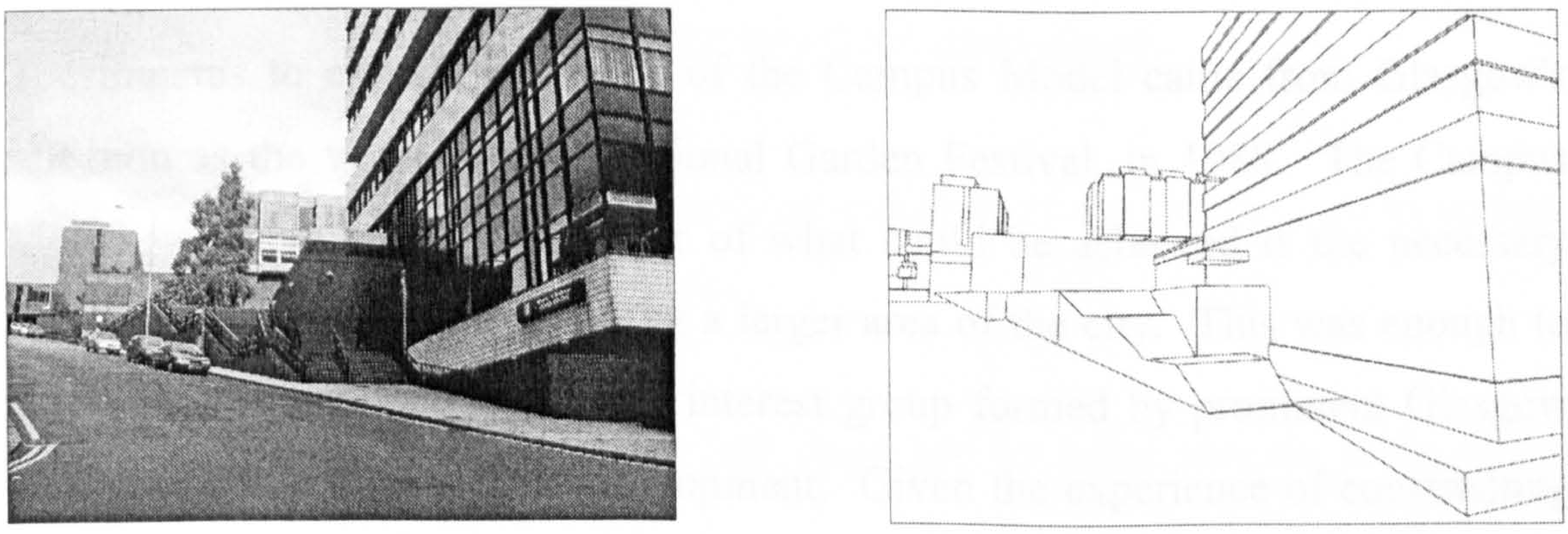


Figure 2.14 The Colville building

The campus model was originally constructed without any form of terrain model to provide a ground plane. Individual buildings were modeled in-situ with enough of the surroundings to suggest the form of the landscape. When visualised in a wire frame mode the "transparency" of the ground was not considered a problem although care in the choice of viewpoint was required

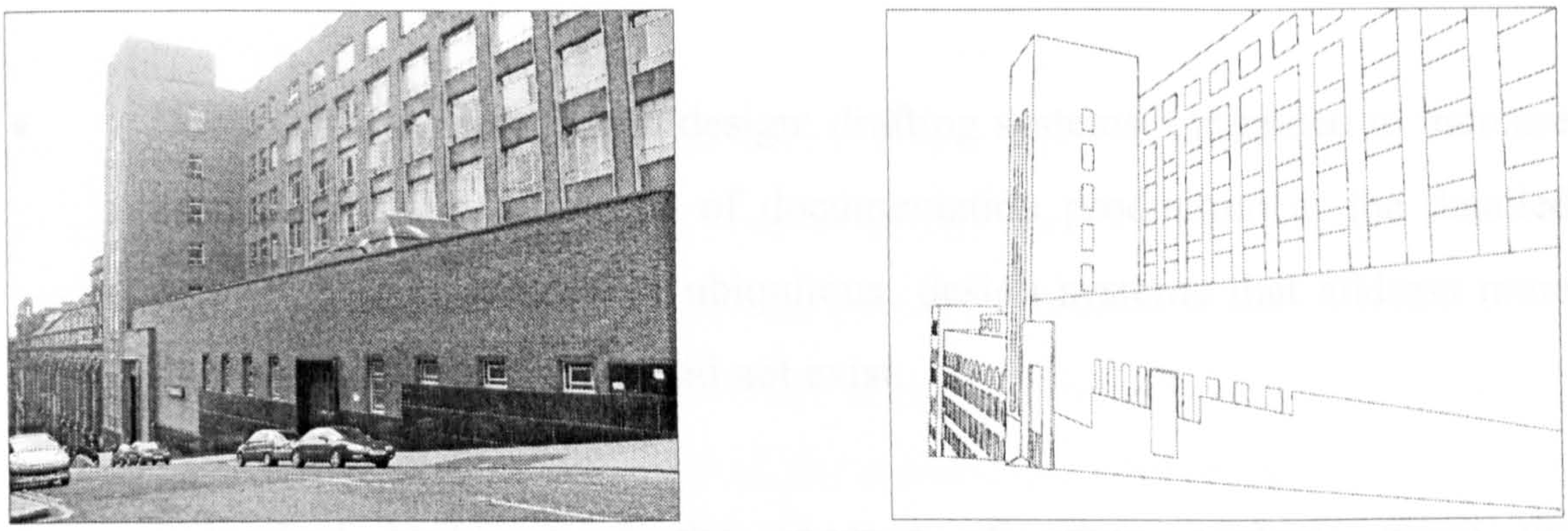


Figure 2.15 The James Weir building

Time has shown that the campus model did prove to be successful, not least as an entity in its own right, but also as proof of concept in that large scale computer models, in geographical terms, could be constructed and could indeed fulfil a useful role. Perhaps the most useful outcome of the construction of the campus model was its use as an exemplar in generating funding that allowed its scope to be extended until it achieved its eventual metamorphosis into the Glasgow Model.

2.7 Glasgow City Model

The impetus to expand the scope of the Campus Model came from Glasgow's selection as the venue for the National Garden Festival. in 1988. The Campus Model was help up as a exemplar of what could be achieved if the necessary resources were applied to capturing a larger area of the city. This was enough to persuade Glasgow Action, a local interest group formed by prominent Glasgow businessmen, to fund further development. Given the experience of constructing the Campus Model it was easy to provide a rational and formulate a simple strategy.

In the 1980s developments in computer aided architectural design had been stimulated not only by technological advances but also by the increasing expectations of architects, planners, developers and owners of building stock.

The major investment in software development over this period had centred on:

- Documentation rather than design: drafting systems - intended to increase productivity and efficiency of documentation production at the detailed design stage have become ubiquitous: design systems that address more than a single development did not exist.

- Individual buildings at the expense of urban context: given the technological limitations and the focus on the production of detail design drawings it was not surprising that design issues at an urban level have been neglected.

- Static as opposed to dynamic treatment of form: the constraints of affordable computer graphics have reinforced the formal static view of the built environment at the expense of addressing the dynamics of an urban environment.

With these observations in mind the strategy initially envisaged for the construction of the model broke down into three main phases:

Stage 1: Construct a simple topographical database of the city centre - to include streets, railways, rivers and city blocks.

Stage 2: Extend the scope of the first stage and to detail specific buildings of interest.

Stage 3: Integration of the topographical model with existing alpha numeric databases of building attributes for example: address; usage; ownership; age; architect; commercial value.

2.7.1 Capturing the Building Plan-Forms

Strathclyde University Campus is situated on the north east edge of what is generally considered to be Glasgow's city centre and as the city model was a continuation of the existing campus model it was considered practical to extend the existing model to encompass the rest of the city centre. At this stage it was impossible to gauge the extent of coverage achievable with the limited resource to hand but the aim was to at least capture the entire city centre, an area of some five square kilometres. This ambition set the scale and scope of the task of data capture.

The construction of the model was in part a commercial exercise, in that a sponsor was looking for a return on their investment, and in part an academic exercise in that data capture and modelling on this scale was a technological first. This in turn meant that some of the methods employed were experimental and also that new research had to be conducted to establish the availability of data sources and the most productive means of data capture.

2.7.2 Data Sources

On investigation the availability of useful data proved to be relatively limited.

- Aerial photography.
- Pre-existing architectural surveys.
- Field surveys.
- Physical models.
- Ordnance survey cartography.

2.7.2.1 Aerial Photography and Photogrametry

Strathclyde Region were able to provide aerial photography of most of the area under investigation. These photographs had originally been commissioned as part of a photogrametric survey at the time of the introduction of Glasgow's urban motorway links.

Photogrametric surveys offer the benefit of combining a consistent treatment of the plan-form over the entire survey area with the added advantage of being able to achieve absolute heighting for all observable points. Film diapositives are located in a precision stereo-plotter which allows the operator to recreate the relative positions and altitudes of objects at the instant of exposure. Measurements of both plan and elevation (altitude) can be made in the model space which exists in the overlap between the stereo pairs. The accuracy of these measurements will vary with the scale of photography, the availability and quality of control points, the type of stereo plotter employed and the experience of the operator. For any one set of photography, accuracy is also dependant on the aspect of the surface under consideration. An experienced operator can achieve a repeatability for the measurement of height of about 0.3 m in model space for a well defined target, but this error will increase in areas of shadow or where control points are harder to define. Errors in plan can be expected to be of a similar magnitude.

In urban areas the context is typified by tall buildings and relatively narrow streets which will cause problems in photogrametric data capture. As buildings are relatively taller than surrounding features they are subject to an increasing "lean back" away from the principle point of exposure which will cause significant areas of detail to be obscured. Even where ground detail is well defined, the "canyon" effect of narrow space surrounded by taller structures will tend to obscure the necessary data points in dense shadow.

The ability to utilise photogrametric survey data is, of course, dependant on access to the necessary stereographic plotting equipment. Only two photogrametric plotters were available - a Kern PG2 and a Wild B8, both situated in the Department of Topographical Science at Glasgow University. The relative inaccessibility of these plotters precluded their use on anything other than an experimental basis but the loan of a parallax bar enabled relative height measurements to be made. This tool was employed extensively during the project and although relatively crude enabled the capture of most of the building height measurements.

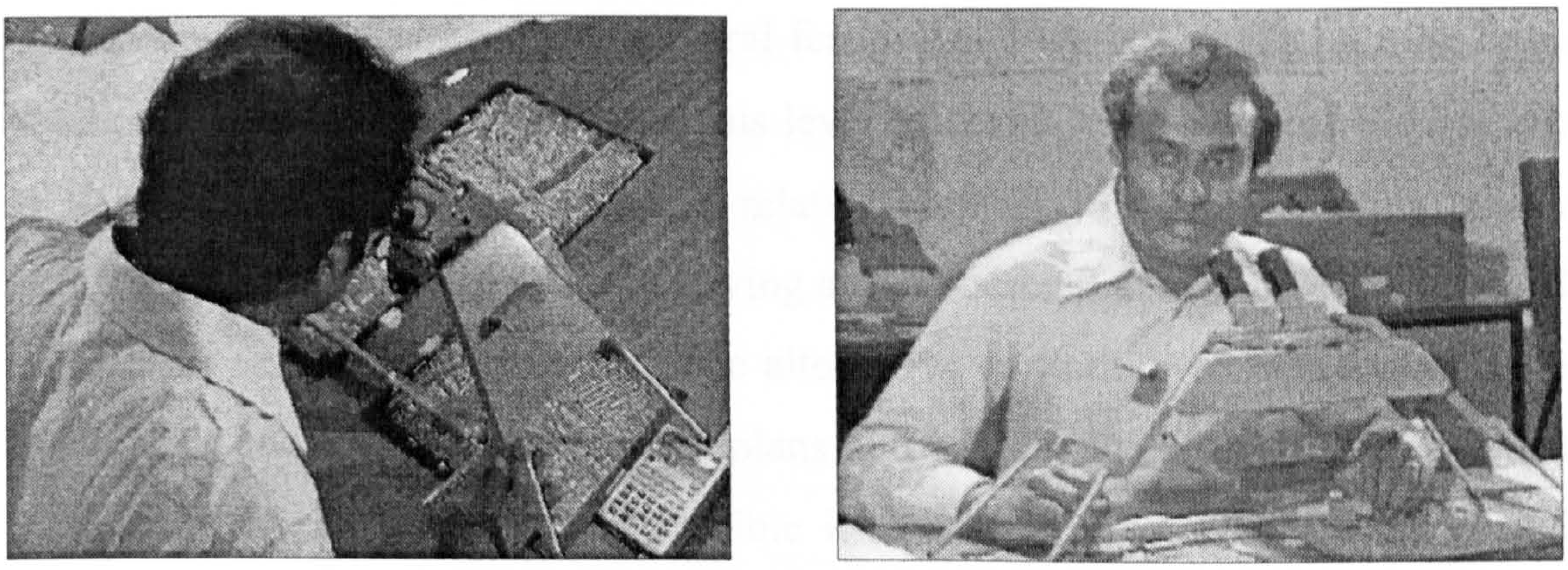


Figure 2.16 Using a parallax bar to capture building heights

Taking into consideration the time scale, the proficiency of the team and the envisaged scale of the project it was decided to represent the buildings as accurate in plan and with a height based apon the dimension of the average of the ground to eaves separation. One feature of the parallax bar method of obtaining height information was the degree of training and experience required to obtain repeatable results. It was readily apparent that as the task was shared within the team there was a difficulty in maintaining consistency not only over what constituted an average eaves location but also in obtaining correct readings from the Vernier scale on the parallax bar. In the event an exact dimension was not critical as the relative proportions of the building were largely unaffected and the texture of the urban fabric was maintained.

2.7.2.2 Existing Architectural Surveys

In the past a number of architectural surveys of Glasgow had been commissioned. One such survey covered an area of current interest. This data proved to be of great help although with hindsight it is apparent that the proposed model was not sophisticated enough to make full use of the information available. The problem endemic with this sort of data in the context of the project was that there was too much information concentrated on too small an area. For example one such source was a survey of an area of Glasgow known as the Merchant City, this consisted of measured and drawn elevations of the street frontages of an area of some six city blocks. The drawings were well detailed revealing the position and dimensions of all pertinent architectural features and while this would have been invaluable had the model addressed this level of detail, in this context the amount of useful information provided was relatively small. Having said this it was important to note the usefulness of having a data source that offered a uniform and contiguous series of dimensions. The alternative to finding an area survey is to track down individual drawings and plans and capture the relevant data of each building in turn. While this is possible it represents a considerable effort and investment in time.

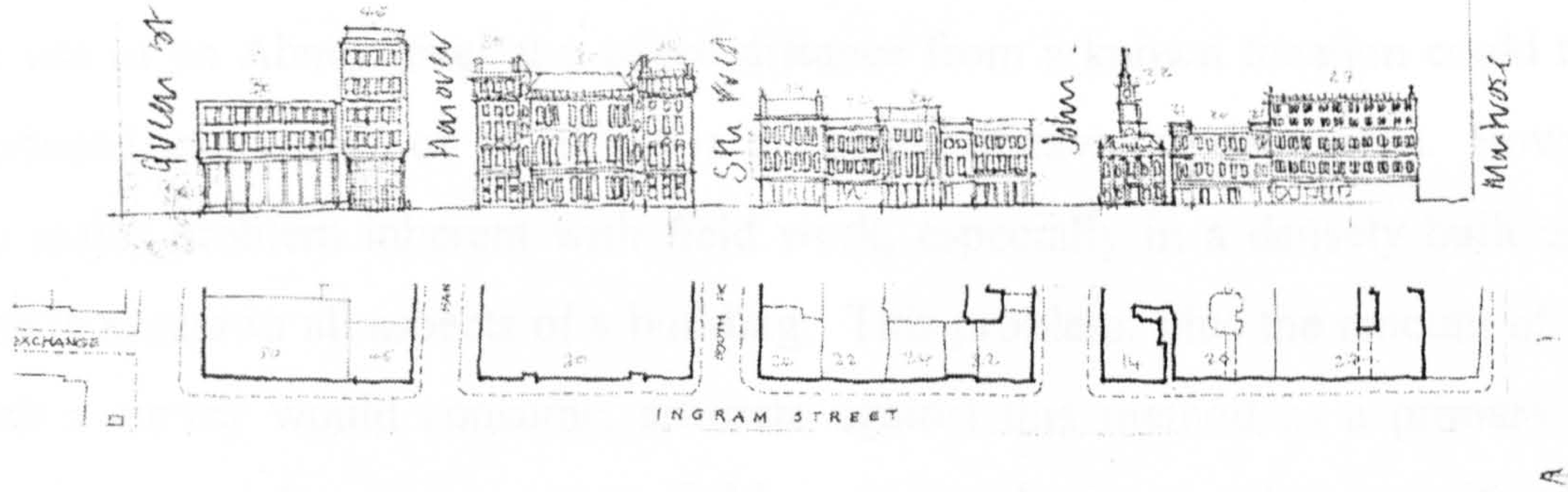


Figure 2.17 Merchant City survey drawings

In theory all plans are available from the city planning department however in a city as old as Glasgow this proved not always to be the case. Some plans and elevations may be available from individual architectural practices. The problems inherent with this method of data capture are numerous:

- Coverage within the project area will be patchy at best.
- Tracking down specific drawings will be very labour intensive.
- Even where drawings exist access may be denied.
- Buildings are often detailed individually, their relationship to their neighbours may be difficult to define.

- Given the level of detail for the model the amount of useful data that can be derived from such sources is limited.
- The most pertinent data values in this context are those of building height but rarely do the drawings indicate an absolute ground level.
- Even where drawings do exist there is no guarantee that the architect's proposals match the final building construction.

2.7.2.3 Field Surveys

In the absence of more traditional media there was often little alternative to performing a field survey. At the time the principle dimension of interest was still the building height which was relatively easy to obtain. This was usually done by direct observation, either by comparison with a neighbour of known proportion or by counting stories and estimating average floor to ceiling heights. If available, the use of an Abney level at a paced distance from a known location could have produced an accuracy of +/- 0.5 m in the height of eaves or ridge lines. However the major problem inherent with field work, especially in a densely built urban area is access to all aspects of a building. This problem, plus the amount of time such a survey would consume, arbitrate against this method as a primary data capture route. In some instances field work may be inescapable even if only to validate data from another source.

2.7.2.4 Physical Models

Glasgow is perhaps unique in that the City Council have a physical model of most of the city centre. This model is a 1:50 scale representation of the buildings. Although somewhat unusual the physical model did prove a useful source of height information, measurements in plan, however were naturally more problematic.

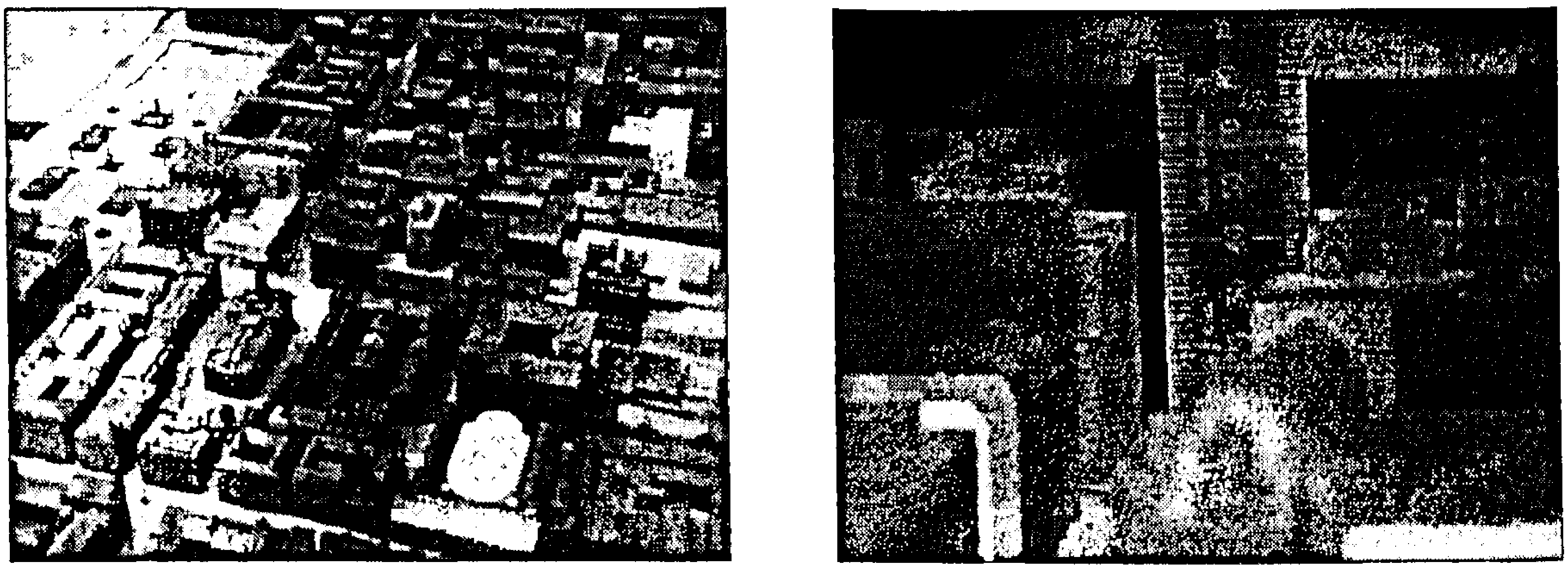


Figure 2.18 Taking measurements off the Glasgow Council physical model

2.7.2.5 Ordnance Survey Cartography

In essence the Glasgow Model database is a three dimensional plan representation of the city so it is not surprising that accessing existing cartographic information proved to be the easiest route to capturing basic building information.

Figure 2.19 Capturing the building planforms

The survey area was captured from the 1:1250 map base. The team was then divided into groups and each allocated an area of responsibility. Within each of these areas the group would estimate the heights of each individual building from the sources identified above and this data would be transferred to the map. The maps were then passed to the digitising team who would use the TABLET program to capture individual buildings. The tablet was operated in serial with a Tektronics terminal as shown above. The terminal operator would initialise the tablet, open the current data file and add the height and address of each building block as it was input.

2.8 Constructing Computer Models Of Buildings

Constructing a computer graphics model of a building is not necessarily a straight forward task, much depends on a number of interrelated factors such as:

- the software used for its construction.
- the software used for its eventual display.
- the hardware that supports the software tools.
- the end use envisaged for the product.

These interrelated technological factors not only determine the methods of constructing and viewing geometry but also tend to characterise the resulting aesthetic nature of the product.

In the most general of terms the task of constructing a successful computer model of a building can be summarised as the process of providing the maximum characterisation of the subject for the minimum of input data. As such, constructing models of buildings is an art in its own right in that the modeler must have both a good knowledge of what architectural features characterise a building and an intimate understanding of how the software will allow a visualisation of these features.

With the software technology under discussion it is possible to construct the same object in a number of different ways. The differing methods of construction, although resulting in an identical object topography, produce an image with marked variations in representation. These differences are due to the witness lines present at the boundaries of the primitives from which the total object is composed. The art in fabricating models using this technology is to choose a construction system that results in a beneficial disposition of these witness lines.

For example, these lines may reinforce the impression of scale by depicting stories on a building or may serve to separate regions of differing materials on a facade. Dependant on the building under consideration, for example a salient feature might be on a relatively "macro" or "micro" scale depending on the 3-dimensionality of the built form. So in one instance a stair tower might be considered a prime feature on a building already rich with other forms yet on its neighbour of equal planform with largely planar facades a useful feature might be as minimal as the boundary between different construction materials or the edge of a portion of the fenestration.

The following modelling procedure, resulting in a diagrammatic representation of an office block bounded by twin towers, illustrates these outcomes.

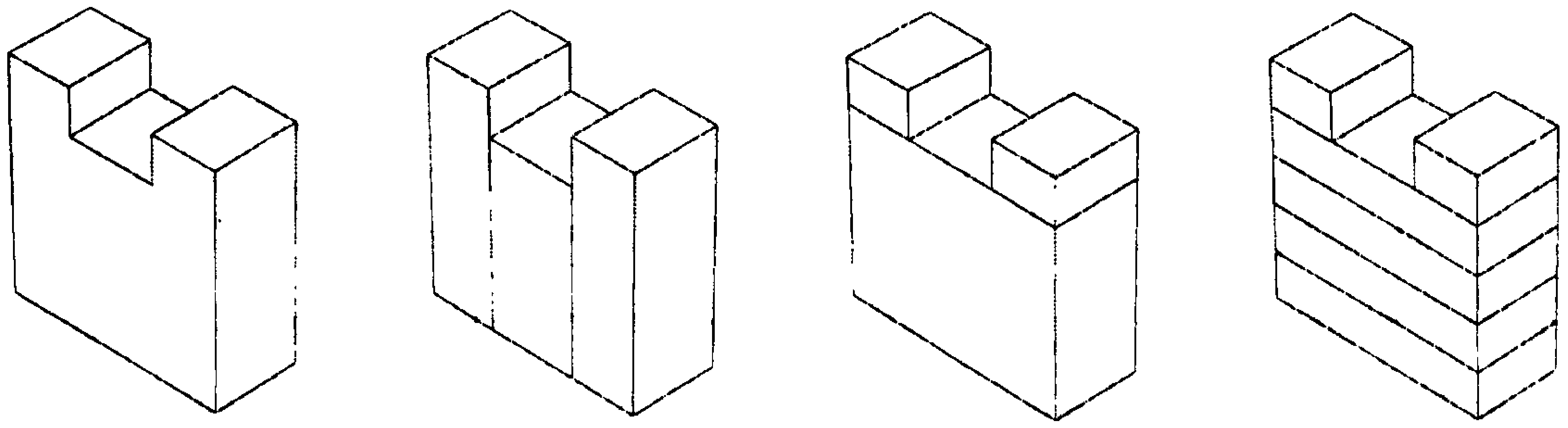


Figure 2.20 Visual consequences of variations in the modelling procedure.

In the first example the building is modeled by extruding an outline along one of the co-ordinate axis. The result is a faithful representation, as long as the cross section is uniform throughout, but results in a bland and featureless presentation with no other apparent detail.

The second technique forms the building by defining three rectangles in plan and then lofting these primitives to the correct height. The resulting image demonstrates the demarcation between the central block and the adjoining towers. This depiction serves to reinforce the formal structure of the building.

The third example is formed from a single block with the addition of two further primitives forming the tops of the towers. In this case the witness lines are distracting as they fail to correspond to a logical construction strategy.

The final example is constructed as an assembly of layers, each vertically dimensioned to correspond to the story height. This produces a faithful representation of the building structure but suffers from greater data redundancy as the inclusion of the abutting floor and ceiling surfaces is inefficient as they serve no practical purpose within the model.

While these considerations are of prime importance at the scale of a single building there are many more factors to be considered when this single building is just one of many representing a area the size of a city.

2.9 City Models

The task of constructing a city model is complicated by the need to account for more technological and practical constraints than would be immediately apparent when addressing a single building. An obvious balance must be weighed between the standard of representation of an individual building and the level of detail applied at a city wide scale. From a technological perspective the capacity of the hardware environment must be considered as this will influence the quantity of data capable of being held in memory and will prove to be a fundamentally limiting factor. Similarly, the time required to process any given amount of data will impact on the degree of ease with which the product can be used. Purely practical consideration must also be given to the facility with which large models can be handled in terms of access, storage and updating. Even given these considerations it is also true that the worth of a city model is in direct proportion to the area covered and the degree of fidelity with which it is modeled. The success of any model must then be sought in the ability to reconcile these requirements against the attendant constraints.

2.9.1 Modelling at Different Levels of Detail

Generally, it can be safely said that there will always be some parts of a model that will purely serve as a context within which the more interesting portions reside. This indicates that a strategy which is capable of addressing variable levels of detail across the model will yield the maximum benefits. In most instances three levels of detail are deemed to satisfy. At the lowest level general building blocks provide context, landmark buildings could be represented by a minimal level of detail and at the top of the hierarchy some building were treated to greater architectural detailing.

2.9.2 Low Level Models of Urban Buildings

In order to provide the width and breadth of the city there was a need to model a great number of buildings at the minimal level of detail. This was achieved by digitising the plan and extruding the building form to a height roughly conforming to the roof level. This was satisfactory from a number of perspectives. In terms of data management this strategy conformed to the ABACUS REGular body format which provided a suitable terse description and was amenable to both easy editing and compact data storage. While still providing only a 2.5-D description this representation proved to be satisfactory as long as care was taken to pick up salient features from with the plan. The figure below compares the rendition of a group of buildings where the planform has been treated purely as a rectangular shape as opposed to capturing a modicum of the character of the building. Even at this primitive level of detail there is a compromise between a simple depiction and expending more resources of a more accurate representation.

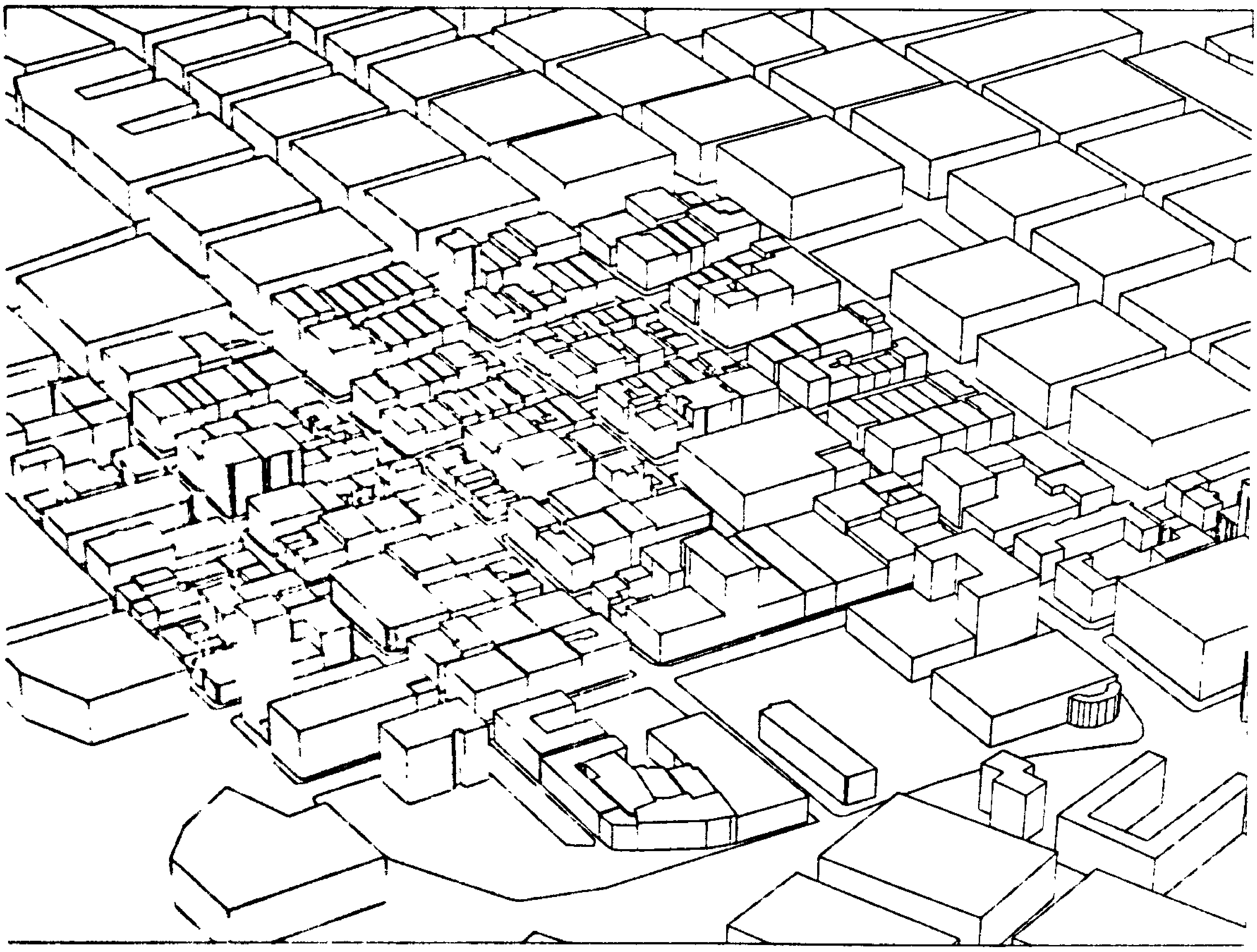


Figure 2.21 Hierarchical levels of detail.

2.9.3 Medium Level Models of Urban Buildings.

It was found that when the entire city was represented at the lowest level of detail users reported difficulties in orienting themselves with regard to their view point within the model. In order to alleviate this problem a number of landmark buildings were identified and provided with just enough generic detail to assist in their identification and subsequent recognition. The techniques employed to provide this level of detail are less formal than would be used to make a truly detailed architectural model as the intent was only to characterise the building. To this end, most of the features are quite coarse and are weighted so as not to burden the model when viewed at a larger geographical scale.

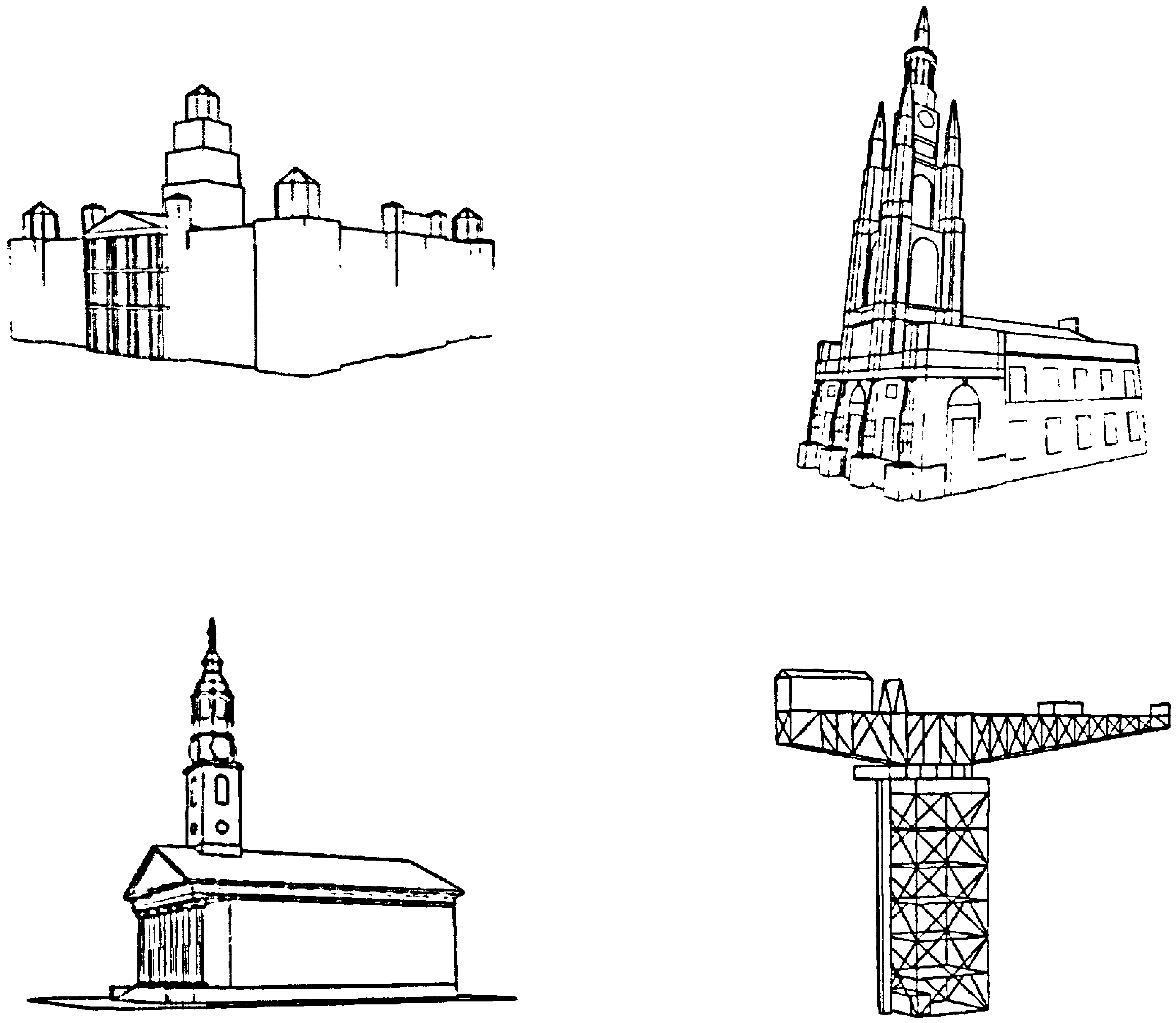


Figure 2.22 Modelling at the medium level of detail

2.9.4 High Level Models of Urban Buildings

As the development of the city model progressed it became apparent that there was a need to address some buildings at a greater level of detail. This was usually because they were the target of some ongoing development strategy and the requirements of the investigation called for a more detailed visualisation over a smaller area.

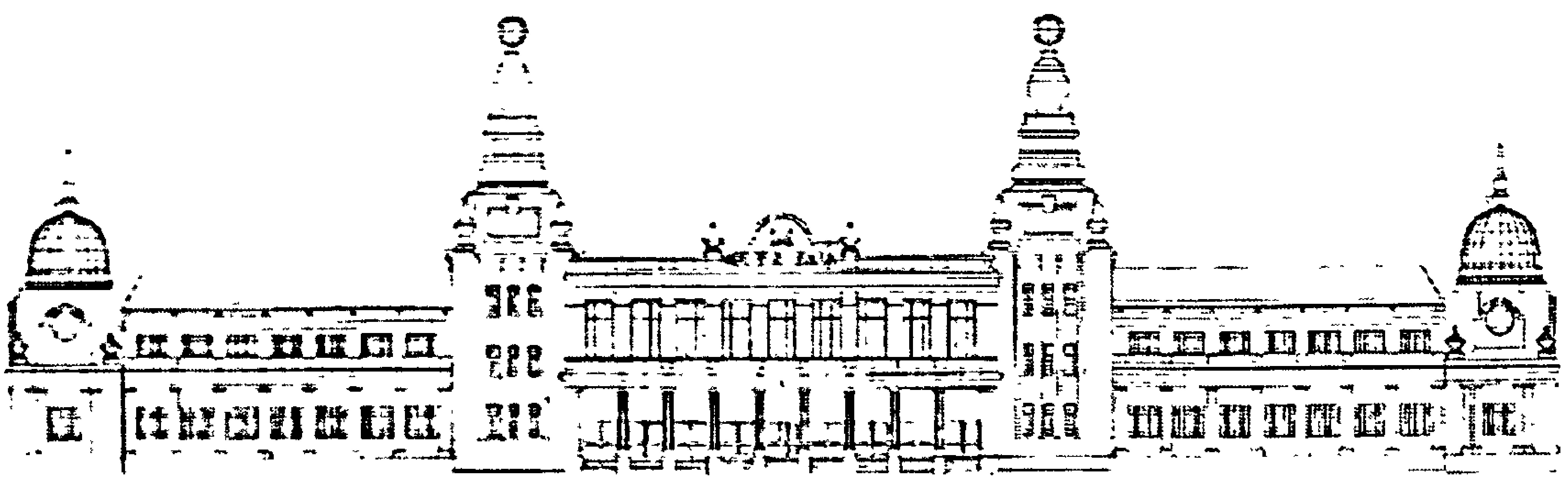


Figure 2.23 Modelling at the highest level of detail

2.10 Transport Networks

Transport networks, including the road, rail and river features, are an integral part of the urban environment and from a visual and aesthetic perspective serve to bind the city fabric together. These features were treated in a number of distinct ways. Individual entities, which had a geometrically distinct character of their own, such as the river Clyde, segments of motorway and railway tracks, were digitised as separate geometry and stored as such. City centre streets were captured in relation to the block that they surrounded. These features were treated as line segments and were assumed to be defined by the pavement edge, where available, and as such were more representative of the land surrounding the buildings rather than the road surfaces themselves. Visualised as a set of lines in space, the distinction is not apparent, yet at a later stage when efforts were made to separate out the salient features of the urban ground-scape this distinction becomes all important.

The intention was to capture those features that comprise "general line detail" in the OS terminology by digitising roads, water features and pathways. The lines so generated were draped over the terrain in a similar manner to the heighting of buildings. As the DTM was not to be used for visualistation the provision of these line networks was seen as being important in reinforcing the illusion of the terrain in areas where the built density was low.

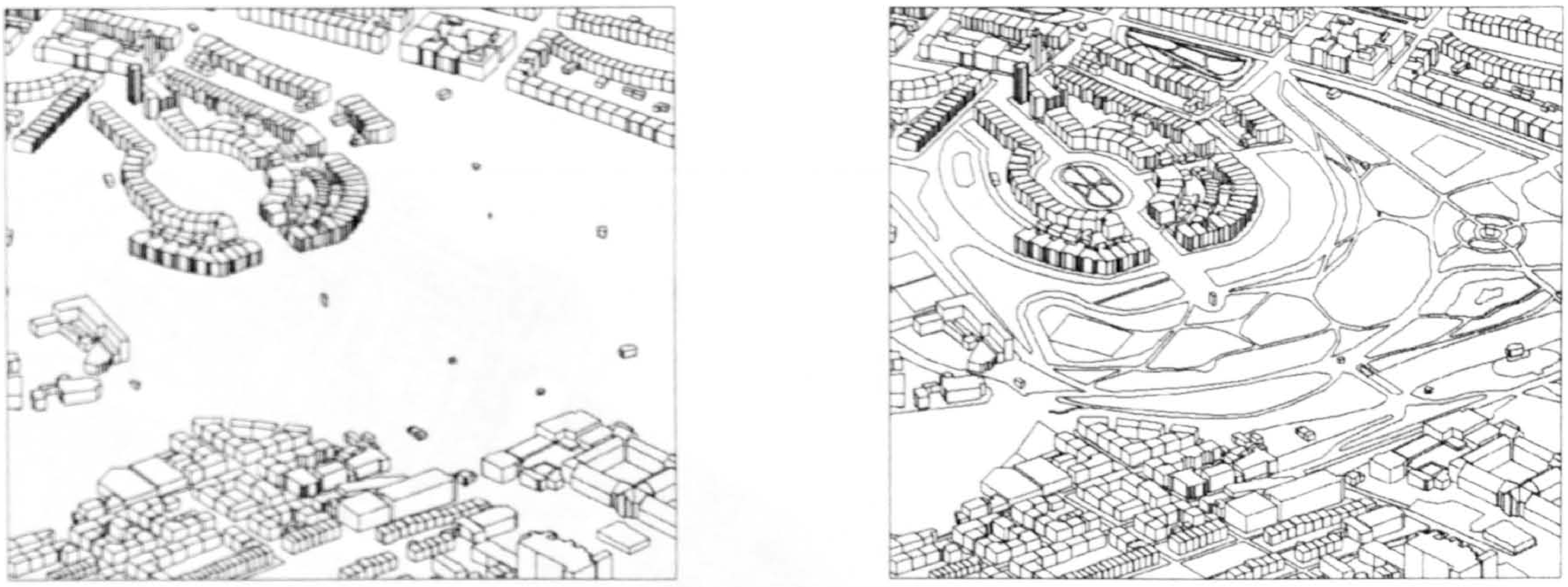


Figure 2.24 Introduction of general line detail

Figure 2.25 Glasgow plan shown without the road networks

Figure 2.26 The complete Glasgow plan

2.11 The Urban Landform

Glasgow is not a "flat earth" city and much of the character of the urban fabric comes from the rolling disposition of the buildings on the terrain. It was evident that to achieve the correct relative elevation of the buildings then some recognition of the underlying landform was essential. This would require finding a relative origin in height for each individual geometric entity that formed the entire model. Other than the fact that manually editing this information into every geometry file would be a tedious and time consuming task the major difficulty would be in locating a suitable source of data. While elevation values were available from a range of cartographic representations no one source could supply a suitably high sample density in order to provide height attributes at all the building locations. As discussed previously some of the missing information could have come from photogrametry or even from commissioned survey information but the real need was for an accurate, economic and perhaps most important an automatic method of applying height information from existing sources.

2.11.1 Terrain Modelling

At this stage it was evident that there was a need for an intermediate representation of the landform – one that could take the localised samples either from contour strings or spot heights – and interpolate these values over the entire action area. The research was then directed towards what did represent a terrain model and what were the precedents for its range of uses.

The ability to model and represent the topography and topology of terrain is of fundamental importance to a range of scientific applications. The exact terminology used to describe the various formats is open to interpretation, opinions differ as to the usage of "terrain" as opposed to "elevation" when describing a digital representation of a portion of the earth's surface. Terrain is now generally preferred when other landscape attributes are included along with the topography in the representation, such as demographic or geological overlays. In this discussion the difference is not germane to the argument and since, historically, this strand of research has always used the nomenclature of Digital Terrain Model (DTM) this naming convention will be adopted for all representations.

The usage of DTMs in the modelling and analysis of spatial and topographic information can be traced back to the 1950s. The range of application domains has grown enormously with the increased usage and competence of computer based systems and it was easy to identify a wide spectrum of fields within which an understanding of the local topology was central to the application. For example in civil engineering road design, cut and fill operations and volumetric calculations are all common problems where the analysis of terrain specific data is required. In geo-scientific applications the ability to generate slope and aspect maps which are required in the interpretation of run off modelling are central to geomorphic simulation. Similarly the growing dependence of the military strategists on the information technologies has fuelled a need for ever more accurate and reliable terrain data not just for simulation and training but also for the development of autonomous guidance systems.

2.11.2 DTM Generation

There are two phases within the process of generating a DTM. Firstly, there is the collection of some, potentially random, set of three dimensional points in space which represent the surface of the terrain. Secondly there is a process which seeks to form an efficient topological structure from the point data set.

The basic point data set can be furnished from a range of sources, principally consisting of existing analogue maps of contour data, photogrametric data capture and surveying. Traditionally digitising contours has been seen as the prime source, largely for historical reasons, this having being the most readily available supply. The adoption of these methods of data capture results in a data set with certain generic characteristics. The means and manner of extraction of some useful subset of this pool of data is dependant on the mode of storage and representation of the final model. There are a number of alternative modelling typologies available, the final choice of which is again determined by the resulting mode of usage.

2.11.3 Contour Models

The traditional representation of elevation is through contour lines either labeled or colour coded to represent change in elevation when represented on 2-D media. Although obviously rooted in two dimensional cartography this format does translate to the third dimension and forms a recognisable abstraction of the input data. The obvious visual disadvantage is the lack of any form of surface model although this can be inferred from the disposition of the contour lines themselves. In terms of data this representations suffers from a lack of any second order continuity so while there is succession along the contour string, areas between height intervals are devoid of data.

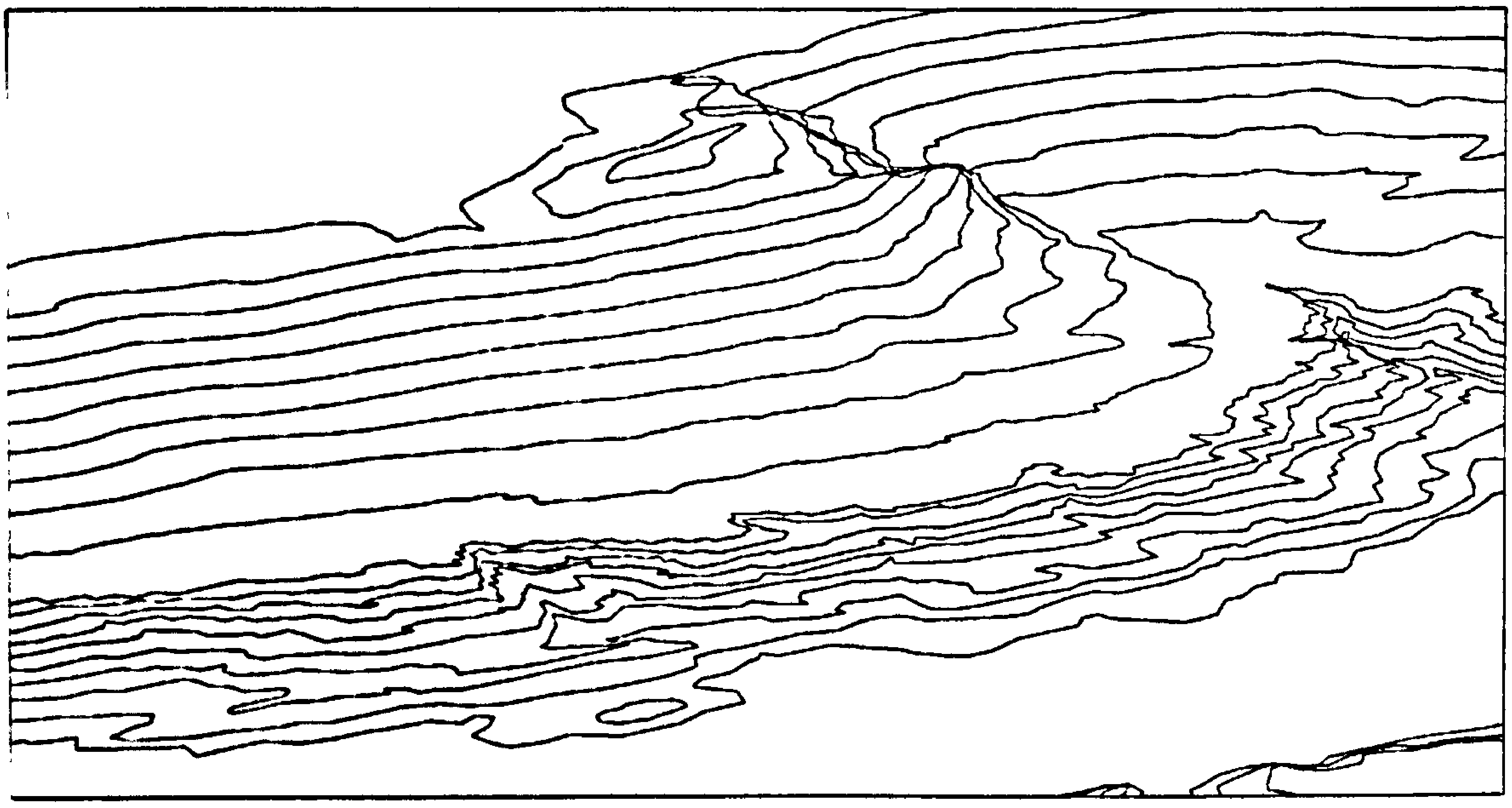


Figure 2.27 Contour model

2.11.4 Slab Models

If the input data is in the form of contours, X Y and Z triples in a linear string, then it is possible to form geometrical entities out of the boundaries defined by the contours themselves. These contours do not have to form closed loops but must be bounded by the extents of the model and must also be "complete" in terms of the description of that iso-height field. The resulting model provides a convincing rendition of the relief but is obviously diagrammatic as opposed to accurate. However the speed of construction and lack of need for any algorithmic pre-processing of the data lends this method to large scale context modelling. The major drawback inherent to this approach is the difficulty of combining the relief with any form of small scale construction that may span contour levels, for instance buildings. Aliasing due to different stepping between increments in the horizontal and vertical domains will always lead to discontinuities but if the sampling frequency in both domains is equalised then the error will tend to zero as the sample rate is increased. Adopting this approach suggests a potential solution.

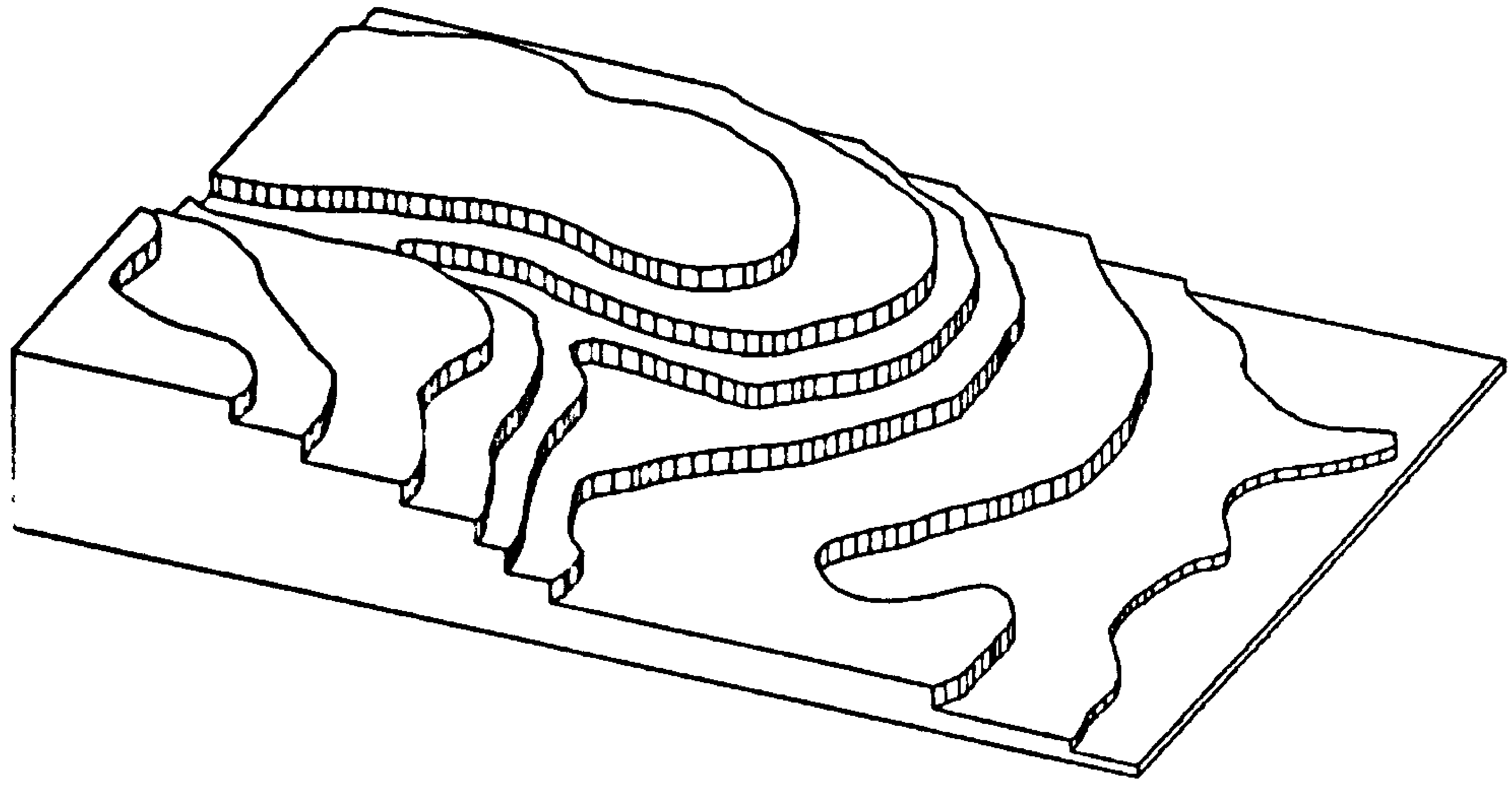


Figure 2.28 Slab model

2.11.5 Grid Models

The grid model, or elevation matrix, is a set of height values stored in a two dimensional array. This model is one that is widely adopted in that it displays certain inherent characteristics that predispose it to computer implementations. The supposition is that only the elevational data is stored, the two dimensional spatial relationships being inferred from the rows and columns formed by the rectangular matrix. Traditionally the columns form the world, or local, Y axis and the rows the corresponding X axis. This structure lends itself ideally to a computer implementation in that storage, processing and algorithmic development are much simplified.

Reinstatement of the implicit X and Y values is performed through the knowledge that the stepping between adjacent cells in the structure is common over the entire data set. There may be a disparity between the incremental X dimension and that assigned to the Y axis but this is uncommon. A non-linear scaling is also possible but even more rarely implemented. The only other data items to be stored with the matrix are an origin, usually representing the bottom left hand extremity and in some cases a rotation, however it is more usual to assume that the positive Y axis points north.

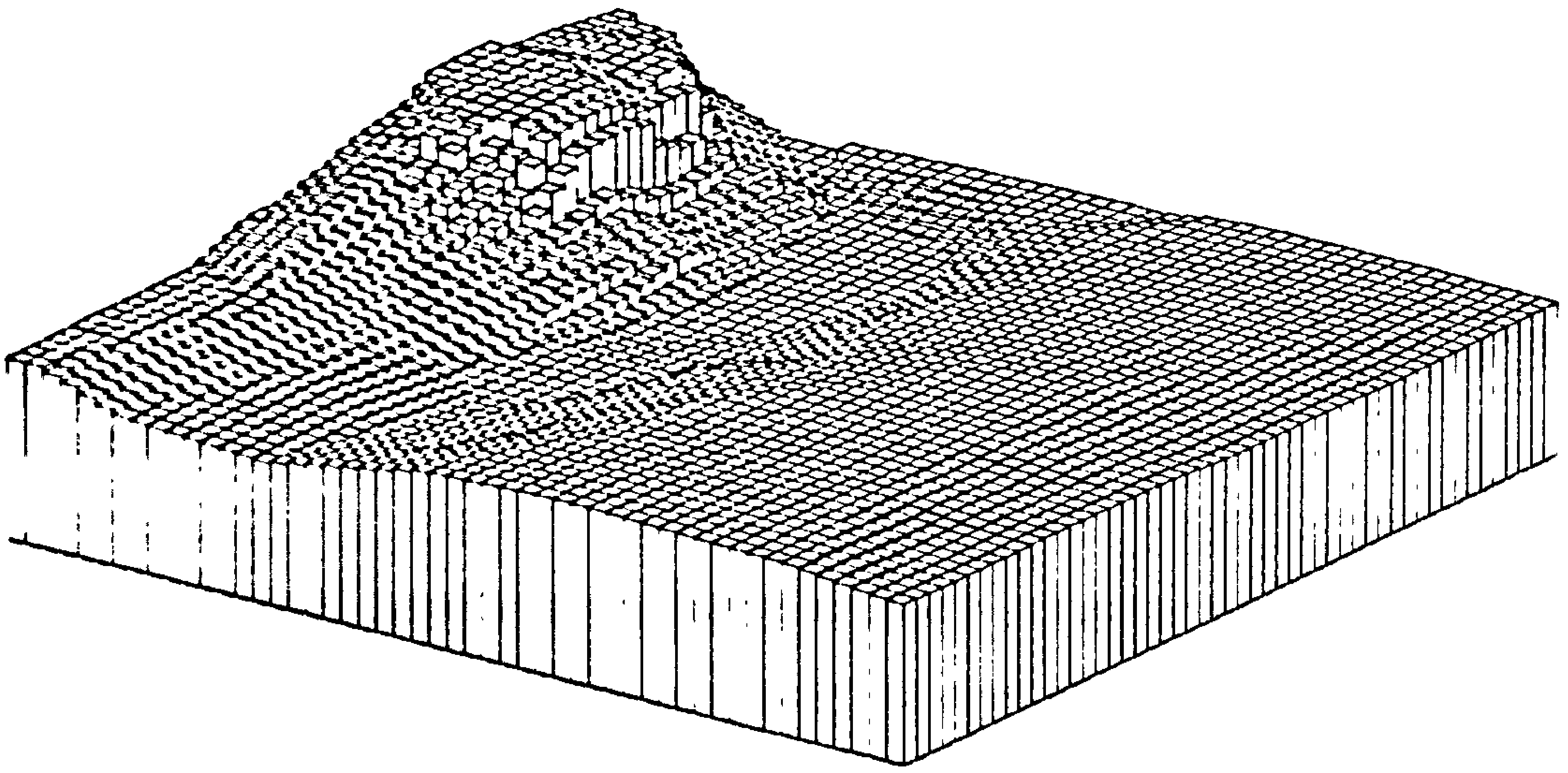


Figure 2.29 Graphical representation of an elevation matrix

A generic characteristic of most methods of data capture is that there will always be regions of high and low sample density. This is inevitable not just because of the need to concentrate data capture on the zone of greatest interest but also because of the nature of the terrain itself. Areas of increased gradient will naturally provide more data points than will planar regions such as lakes or other areas of open water. Since a matrix representation is characterised by a fixed frequency of sample intervals the problem is one of interpolation between adjacent samples. This implies that in regions of rapid change in elevation the will be an under-sampling of the available data and inversely there will be an over-sampling scenario in regions of more gradual gradients. If the data is under sampled then there is a danger of introducing aliasing artifacts, over sampling results in a waste of storage space and an overhead in processing power. The difficulty then is to decide on a sample frequency that will still represent complex relief without providing a disproportional amount of redundant data in other areas.

Interpolation is commonly performed on each grid cell by projecting eight radial vectors from the test point until they intersect eight contour lines. The elevation at the test point is then interpolated through a weighted average. Although there are a variety of interpolation schemes available this method is prone to introducing artifacts such as "terracing", produced by nearest neighbour inverse weighting, or by "ringing" when using higher order derivatives of partial differential equations. These artifacts can be reduced by adding more data, if available, in problem areas but the human visual system is acutely aware of such discontinuities and in a visualisation context they prove hard to hide

Despite its potential drawbacks the elevation matrix remains a useful construct for a number of reasons. As stated previously its representation is predisposed to simplistic algorithms directed towards the generation of topographical relationships between adjacent data points and also the "tiling" of separate matrices to form a mosaic of a larger area. The regularity of the rows and columns of the underlying structure additionally provide a very fast and easy method of access to other data that may be geo-referenced to each data cell.

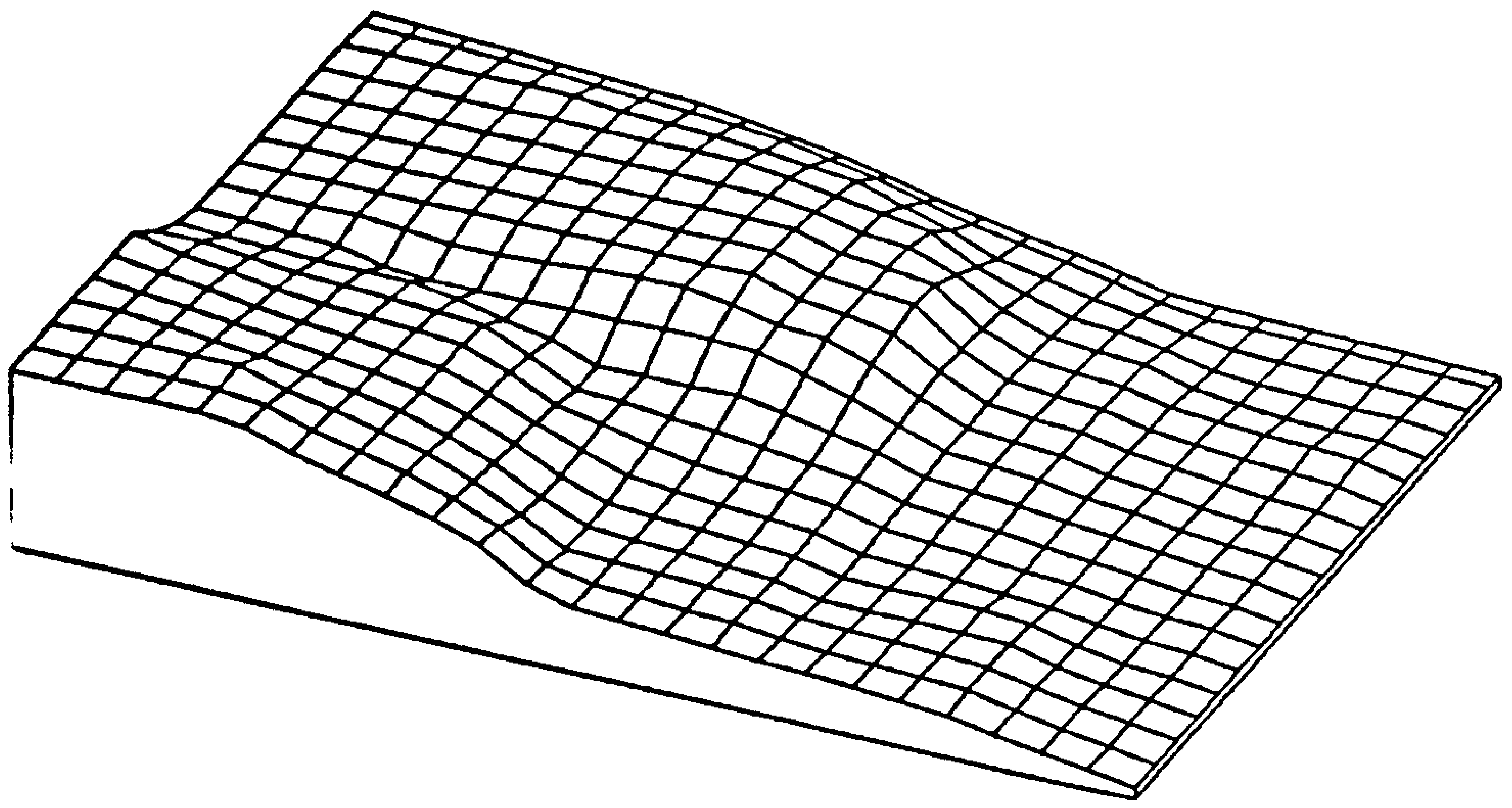


Figure 2.30 Digital terrain model mesh

2.12 Generating the Glasgow Terrain Model

The adopted strategy for generating the terrain model first sought to captured the height information by using the TABLET program to digitise contours from 1:10,000 Ordnance Survey maps. This data was saved to an ASCII file as individual strings of values representing both location and height. Subsequently this file was input to another ABACUS program, SITE, which converted the contour strings to a digital elevation matrix. The computational procedure then implemented for this process utilised nearest neighbour sampling to build a polynomial function which could then be sampled at a frequency relating to the output parameters of the matrix. Unfortunately this approach was characterised by surface overshoot, oscillations and numerical instabilities leading to artifacts in the final model. The insertion of additional data points obtained from spot heights and manual interpolation between contours in areas where problems were apparent eventually generated a satisfactory terrain model.

This data set had a regular cell dimension of 50 meters and covered an area of 64 sq. kilometres. The extent of this model was deliberately generous in order to allow for future expansion of the city model without having to regenerate the terrain information and yet still provide enough detail to represent the texture of the urban landform.

The terrain model was primarily seen as an intermediate stage - most useful for its ability to transfer the height data from the source to the city model. This process had previously been developed as a tool for assessing the visual impact of electricity transmission towers within the landscape. (Petric 1988). Part of this process required that geometry representing pylons and trees should be "floated" down onto the DTM to take up their correct elevational position. This method was adapted to provide the same automatic capabilities for heighting the buildings and roads within the Glasgow model.

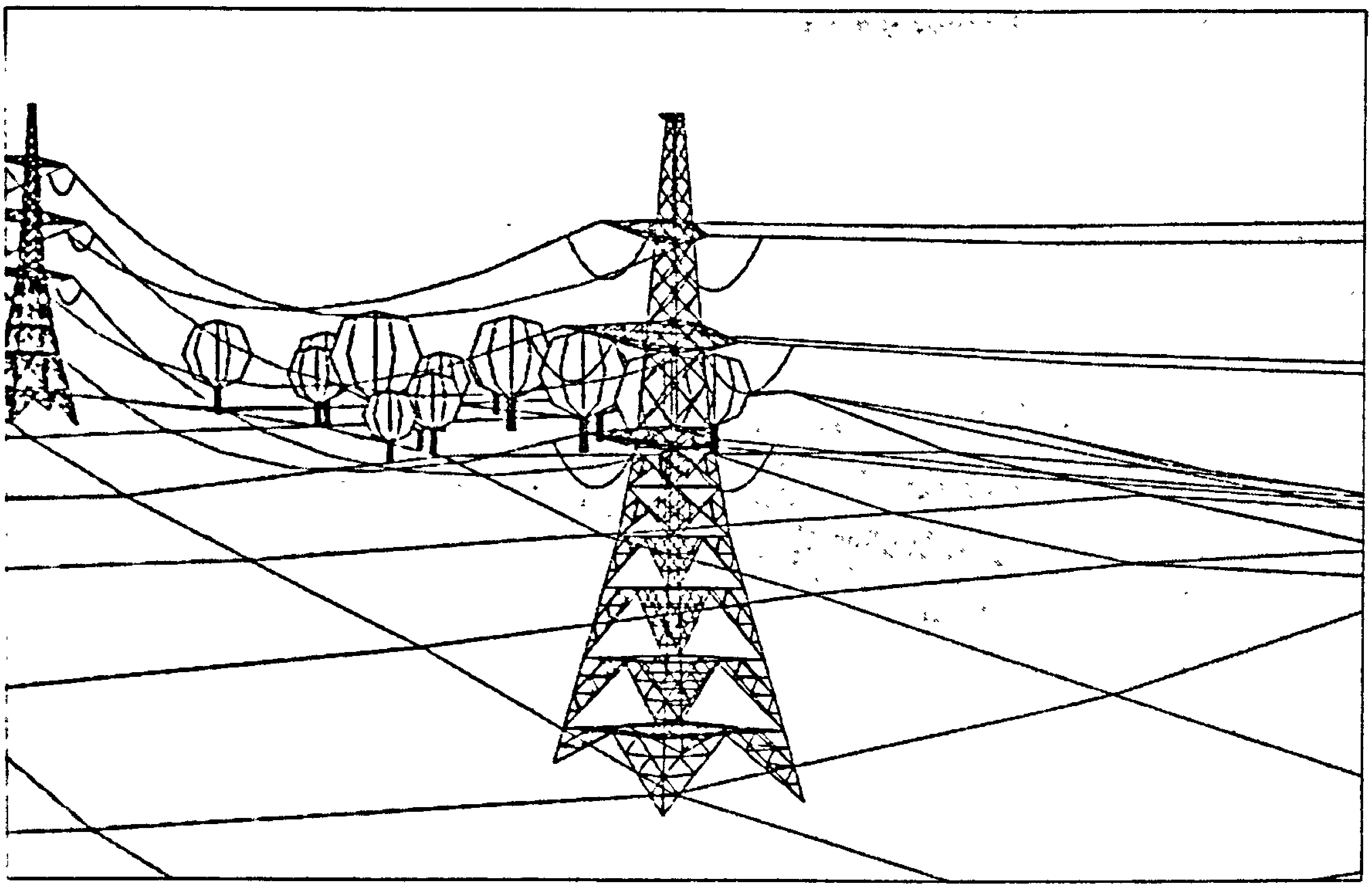


Figure 2.31 Transmission line study

The adoption of a row and column format for the storage of the terrain model exhibits a number of computational advantages. In addition to the ease of recreating the geometry the format also facilitates the retrieval of height information from any arbitrary point within the model. This process proceeds through three stages, firstly there is the identification of which cell within the matrix contains the sample point. This can be found from the equations below.

```
I = (XP - XMIN) / NROW + 1
J = (YP - YMIN) / NCOL + 1
```

Where:

```
I and J are indices pointing to the required row and column
XP and YP are the 2D co-ordinates of the sample point
XMIN and YMIN are the minimum spatial extents of the terrain model
NROW and NCOL are the number of cells in each row and column
```

The ability to retrieve the correct cell with such a simple procedure represents a major computational saving compared with having to potentially search the entire data set with a point in polygon test should an irregular format be employed.

Once the correct cell has been identified the transformation of the data within the elevation matrix into some spatial format is trivial. Each cell in the matrix can be formed by looping on two registers, shown here as `I` for rows and `J` for columns, `DX` and `DY` are the spatial offsets for data cells, `Z(I,J)` is a two dimensional array containing the elevation matrix.

V	X	Y	Z
V1	I*DX	J*DY	Z(I,J)
V2	(I+1)*DX	J*DY	Z(I+1,J)
V3	(I+1)*DX	(J+1)*DY	Z(I+1,J+1)
V4	I*DX	(J+1)*DY	Z(I,J+1)

Table 2.7 Vertex generation from elevation matrix

This representation assumes that the geometry is constructed by using adjacent elevation values as the `Z` co-ordinate for each vertex, alternately matrix values could be assumed as occurring at the centre of each cell and values at the vertices obtained by interpolation.

2.12.1 Interpolation

As mentioned previously the software employed for “floating” roads and buildings onto the landscape was originally developed for the assessment of the visual impact of pylons and forestry within the natural environment. Trees and transmission towers make point contact with the ground and therefore do not require any degree of sophistication to ensure a good “fit”. The dense urban fabric of Glasgow is much more complex and with buildings having a substantial footprint rather than a single point of contact this technique proved to exhibit fundamental limitations.

The illustration below shows the effect of floating the buildings down onto the terrain. The top image shows the original campus model with the correct disposition of buildings in plan but with a single vertex used to position the buildings vertically. This results in the buildings hovering in space and for a long structure may mean that a corner can be as much as 20m off the actual ground level. The lower image represents a later version of the same area from a similar viewpoint and reveals the improvements offered when the bases of the buildings are modified by their interaction with the ground data. The campus is situated at the top of a steep rise and when viewed from ground level, as shown here, the absence of the ground plane allows the viewer to see below and beyond the buildings.

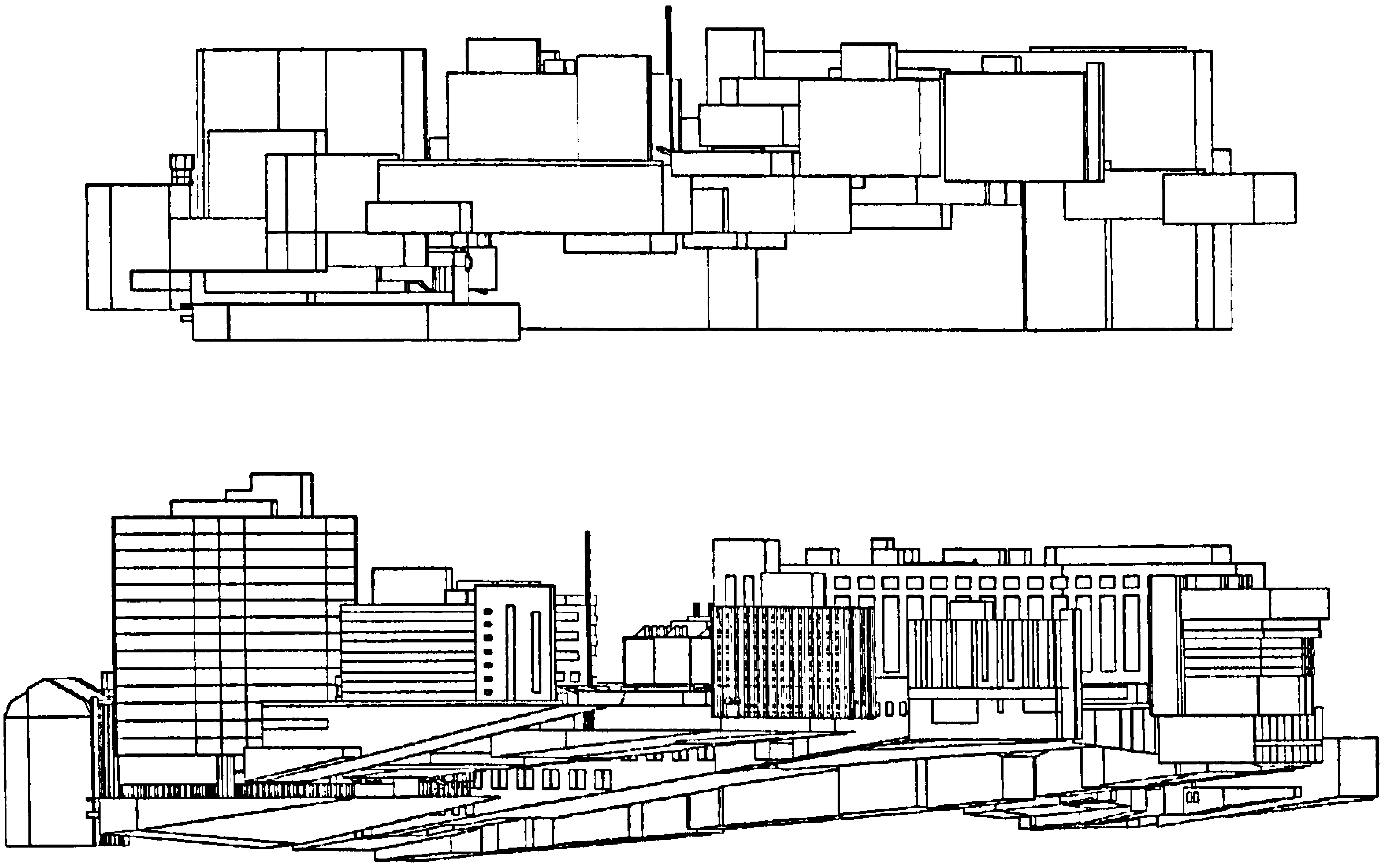


Figure 2.32 Buildings aligned with the digital terrain model

2.13 Model Data Structure

One of the first decisions to be made was how to break down the gross area of the city into logical units that could then be assigned to the teams performing the data capture. Originally this division was executed on an arbitrary basis, each team member being assigned the next street block as the previous area was completed. On completion of any section the new data was amalgamated into the whole within a single data file. This soon proved unworkable as the difficulty in manipulating and visualising the model grew in proportion to the size of the data-set. The system limit on available CPU time within any one session would allow this conglomerate model to be loaded into VIEWER but would not then provide sufficient remaining time to create a plot with hidden lines removed.

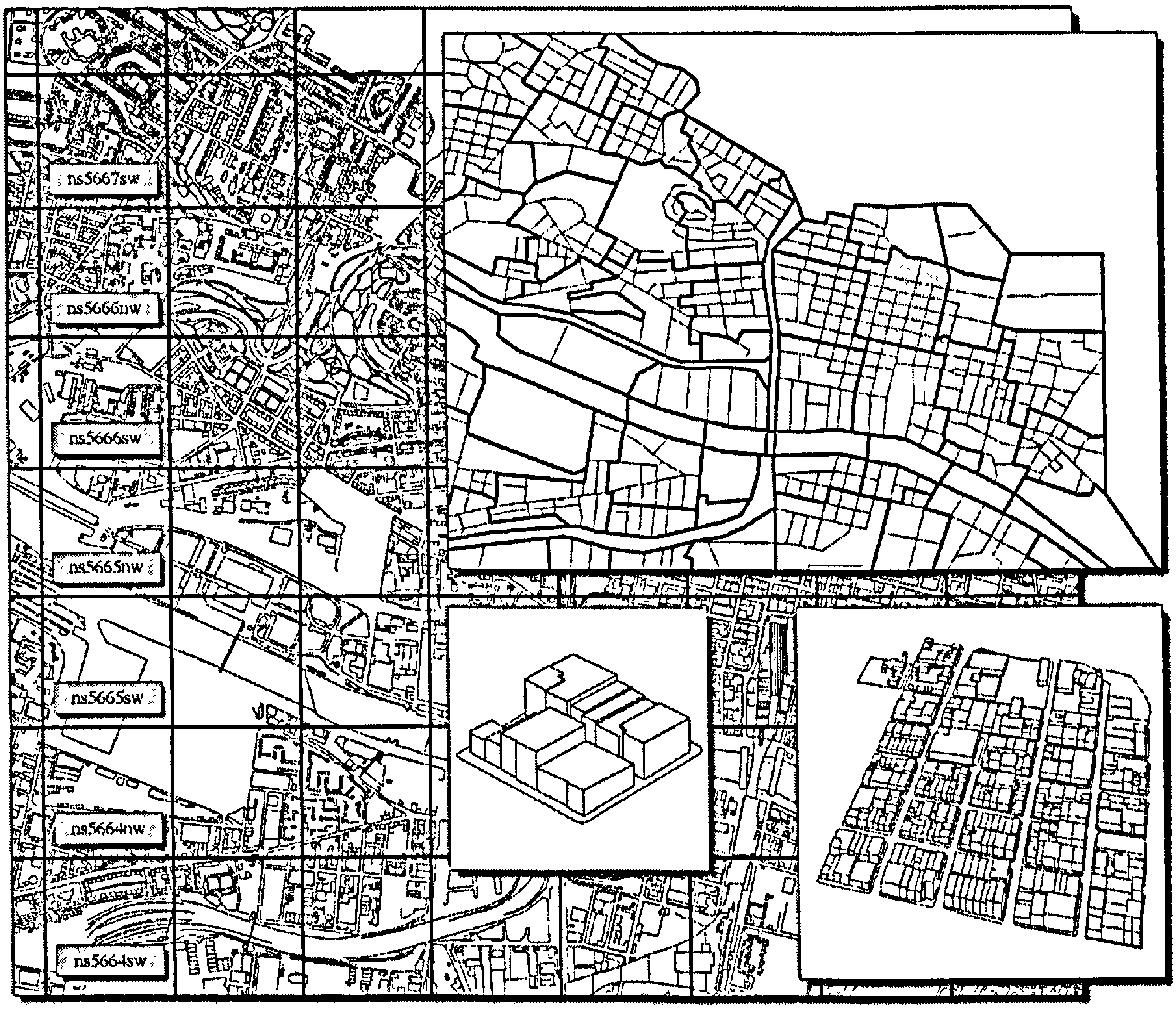


Figure 2.33 Glasgow Model subdivision

To reduce the monolithic structure of the model a decision was made to superimpose a grid, based on the 1:1250 OS map size, and localise an interlocking structure of city blocks within this layout.

Once the city had been subdivided sets of buildings within each region were captured into a single file, numbered sequentially, proceeding in an anti clockwise direction. The intention was to use the hierarchy of the file store on the computer to provide an order for the data. This was achieved by establishing "Glasgow" as a user account on the machine, which then established a "root" for the structure, then each district formed a named directory which in turn contained a number of files containing the building data. Within these files each geometric entity represented a group of buildings captured in a clockwise order from the north west corner of the block. The hierarchy branched to provide a tree for the transport network in parallel to that of the building data.

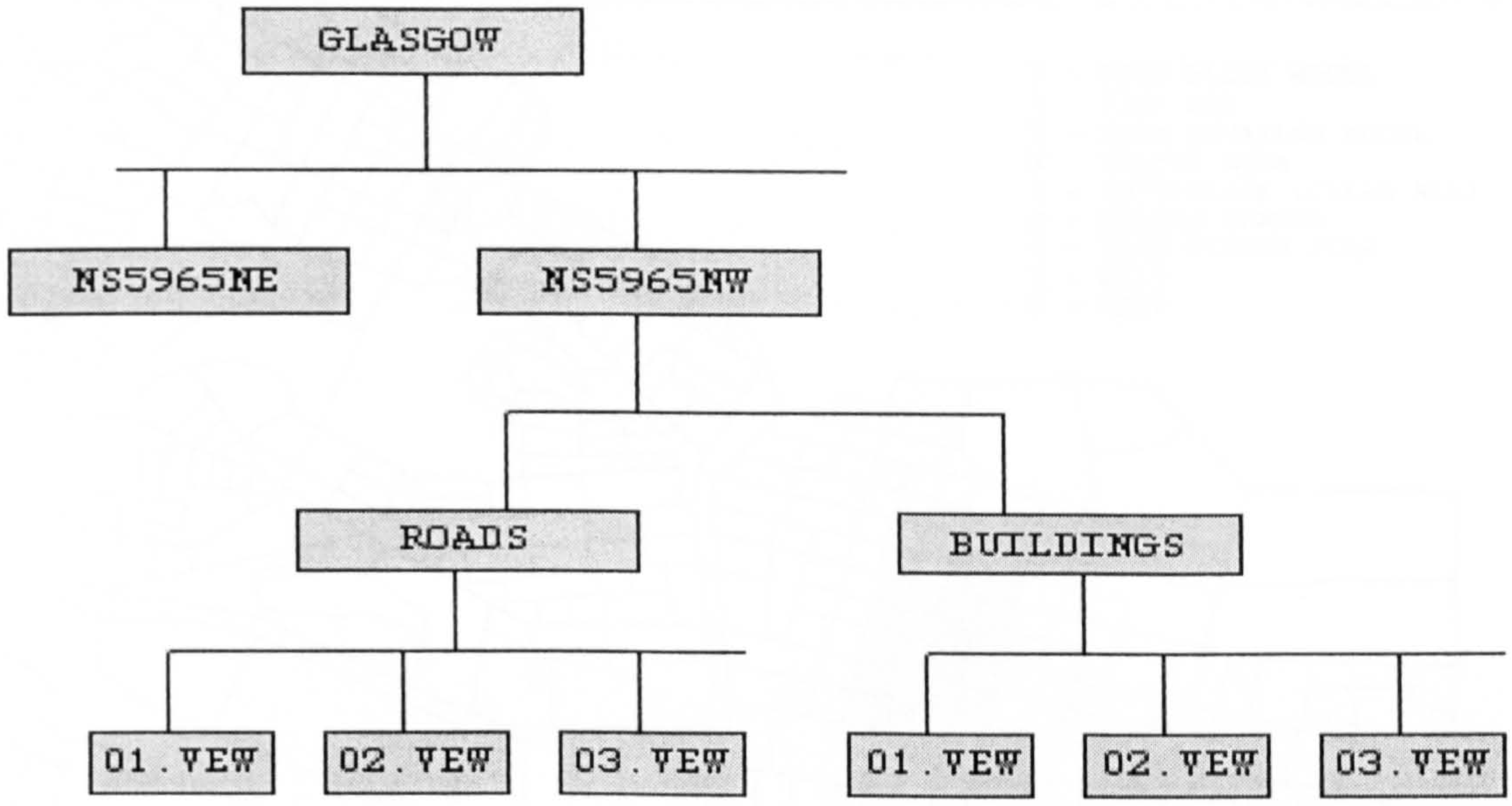


Figure 2.34 Glasgow Model data hierarchy

In order to visualise any desired part of the city the data were concatenated using a script run from the VAX VMS command line. While this was a practical short term solution it required an intimate knowledge of the file naming conventions and a degree of familiarity with the VMS operating system. This restricted the usage of the model to an unacceptable degree. A solution was found through the modification of the VIEWER command set to allow a new command entity that would provide an external recursive reference to data within a separate file structure. This entity was structured as below:

```
FIL
GLASGOW/NS5866SE/BLDNG/01.VEW
0.0, 0.0, 0.0, 1.0, 1.0, 1.0, 0.0
```

Where FIL is the three character token indicating the data type in the input stream. The next line gives the absolute path name for the externally referenced data and the final line supplies a translation, scale factor and rotation for the data. This development provided immense flexibility in the structuring and use of the model not least because it then became possible to develop a graphical interface to the model.

Figure 2.35 The Glasgow Model interface

The contents of the original help file are given below:

```
Welcome to the Glasgow Interface

This program allows you to select an area from the Glasgow Model which can then be
loaded into other ABACUS programs.

The model is mainly stored in a simple block format.  Some blocks contain more
detailed models, these are mainly situated in the Merchant city and the University
Campus areas.  Such blocks are identified on the plan by a small dot in the centre
of the block.

The program is fairly straight forward, keys used are as follows:
P - Pick Block Model (block is hatched)
A - Pick All
D - Pick Detailed Model (Block is hatched if model is available)
X - Deletes Block (Hatching remains until screen is redrawn)
R - Redraws screen
S - Saves VIEWER file (Default name - CITY.FIL)
E - Exits from program
T - A toggle switch

The toggle either includes or excludes buildings from the output file. Roads are
always included and buildings are initially included by default.  Pictures can be
generated interactively using roads only, thus saving time on trial runs using
large areas.
```

2.14 Model Developments

The Glasgow Model was built as a natural progression in the pursuit of a technological means of representing the built environment on a large scale. The preceding description does not do justice to the number of iterations in its format or variations in the method of construction. The developments reported here spanned a period of some four to five years, many people contributed to the project, much of the model was continually reworked – some areas several times – and within this period the model grew, and metamorphosed from an academic exercise into a design tool used by a wide range of Glasgow professionals.

Figure 2.36 The five perceptual landmarks

One key aspect of its success was its transformation from a map to a model. As reported earlier, the first version of the model was largely a 2.5 D extrusion of the cartographic data. As such it was devoid of detail or formal articulation and many users reported difficulty in orientating themselves within the city. As the construction progressed it became apparent that the model needed to account for the way in which people form a mental image of the real world. This image is formed by the conjunction of a set of physical forms which are comprised of five urban elements; paths, edges, districts, nodes and landmarks.. This concept is explored in much greater detail in "The Image of the City" (Lynch 1960). It is still interesting to note that the same theories are applicable to the Digital City.

Different uses and varying viewpoints all tend toward the representation and illustrative media that best suit their purpose. However it is readily apparent that Lynch's classification of visual elements is crucial to the spatial understanding of the urban city-scape, no matter what virtual format it is presented in. The lack of emphasis of these elements was a prime omission from the early versions of the Glasgow city model and when they were introduced the increase in a user's comprehension of the model was greatly enhanced.

A definition and example of these five urban elements in the context of the Glasgow model is given below.

- Paths : Paths may be any form of transport network from which the observer gains his experience of the city. Generally these are taken to be familiar routes along which the observer often travels.

- Edges : Edges are discontinuities between separate areas of the city and are often formed by barriers such as rivers, motorways and railway tracks. These are much in evidence within Glasgow where the city centre is partitioned from the rest of the city by the motorway to the west and the river to the south.

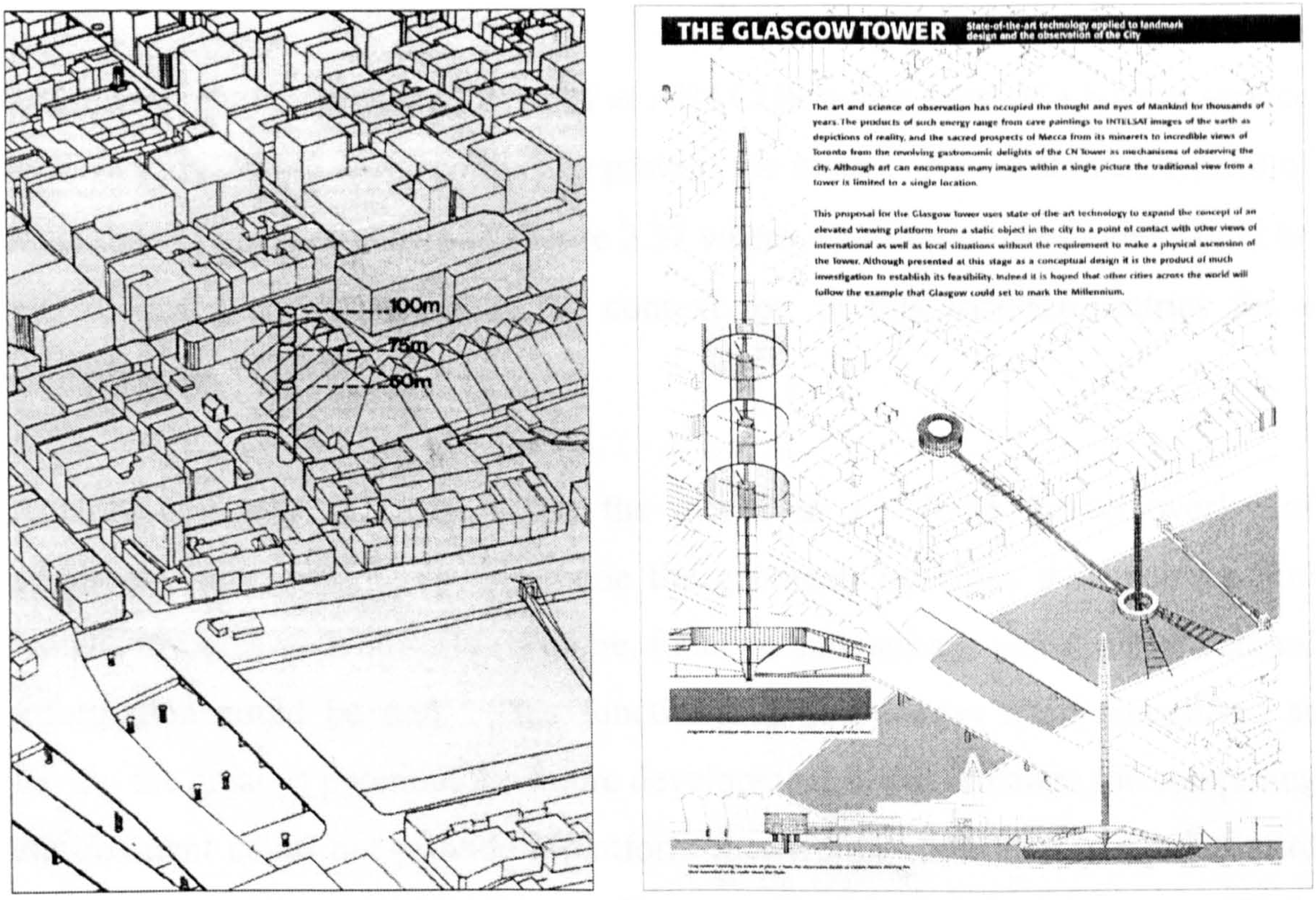
- Districts : Districts are any grouping of buildings within an area which are capable of demonstrating a distinct character which separate them from their immediate neighbours, such as those exhibiting the regular grid pattern of areas within the city centre.

- Nodes : Nodes are a point of reference within the observer's mental image of the city. They can be instanced by the presence of a public square or through the junction of two or more paths that serve as a way point on a journey or even as a focus for the character of a district such as within George Square.

- Landmarks : Landmarks are physical structures that signpost a location. They may be buildings, statues or natural features, but in the case of the Glasgow model it was found that these were most often structures visible within an aerial view like buildings with a characteristic roofscape such as Central Station or the St Enochs Centre or a distinctive grouping of buildings as at Park Circus.

2.15 Using the Model

As the model evolved, becoming a more faithful and graphically interesting representation of the city, interest in exploiting its potential grew proportionally. While there were some technical uses for the model – such as solar availability studies – the most common application was to regard the model as an infinite source of hard copy with designers requesting views onto which they could sketch their proposals.



THE GLASGOW TOWER State-of-the-art technology applied to landmark design and the observation of the City

The art and science of observation has occupied the thought and eyes of Mankind for thousands of years. The products of such energy range from cave paintings to INTELSAT images of the earth as depictions of reality, and the sacred prospects of Mecca from its minarets to incredible views of Toronto from the revolving gastronomic delights of the CN Tower as mechanisms of observing the city. Although art can encompass many images within a single picture the traditional view from a tower is limited to a single location.

This proposal for the Glasgow tower uses state of the art technology to expand the concept of an elevated viewing platform from a static object in the city to a point of contact with other views of international as well as local situations without the requirement to make a physical ascension of the Tower. Although presented at this stage as a conceptual design it is the product of much investigation to establish its feasibility. Indeed it is hoped that other cities across the world will follow the example that Glasgow could set to mark the Millennium.

Figure 2.37 The Glasgow Tower competition.

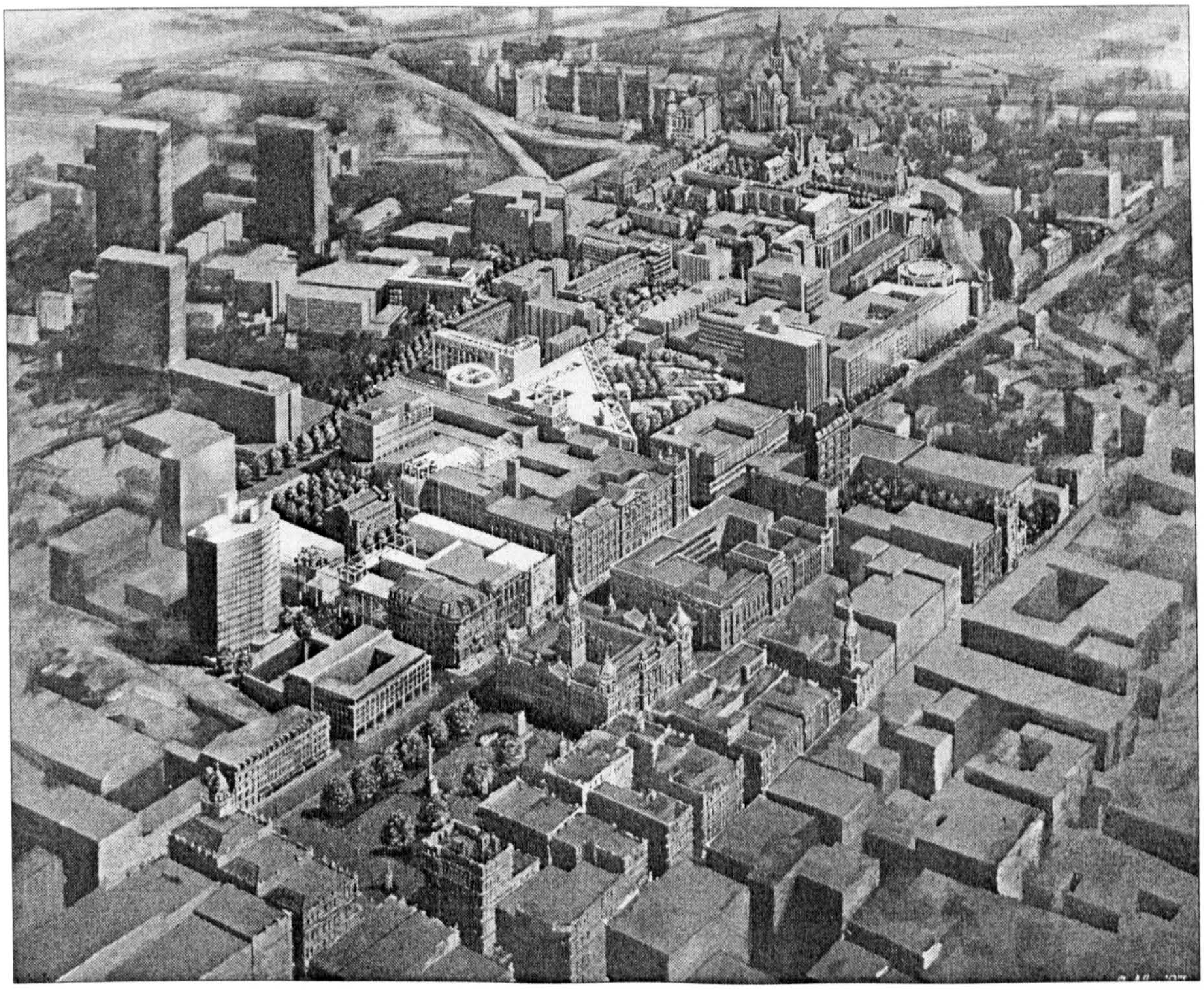


Figure 2.38 Artist's perspective based on the Glasgow Model

These were mostly produced by staff at ABACUS in the form of a bureau service as few practices had the capability to process the model. An illustration typical of this mode of usage is shown in Figure 2.37 where a portion of the model would be plotted and made available as the context for, in this instance, entries for a competition.

Another common use was to use the model as a vehicle for an overlay of additional information. In this mode the geometry becomes the interface into which attributes such as building name, function and perhaps economic or cultural information could be tied. This functionality had always been recognised as having the greatest potential for future development but at this time the computing environment could not provide a platform that would allow the manipulation of this quantity of data with the degree of ease required for interactive access. While a static rendition of the model proved useful it was also limiting and it would take a further generation of hardware and software development to break free from these constraints.

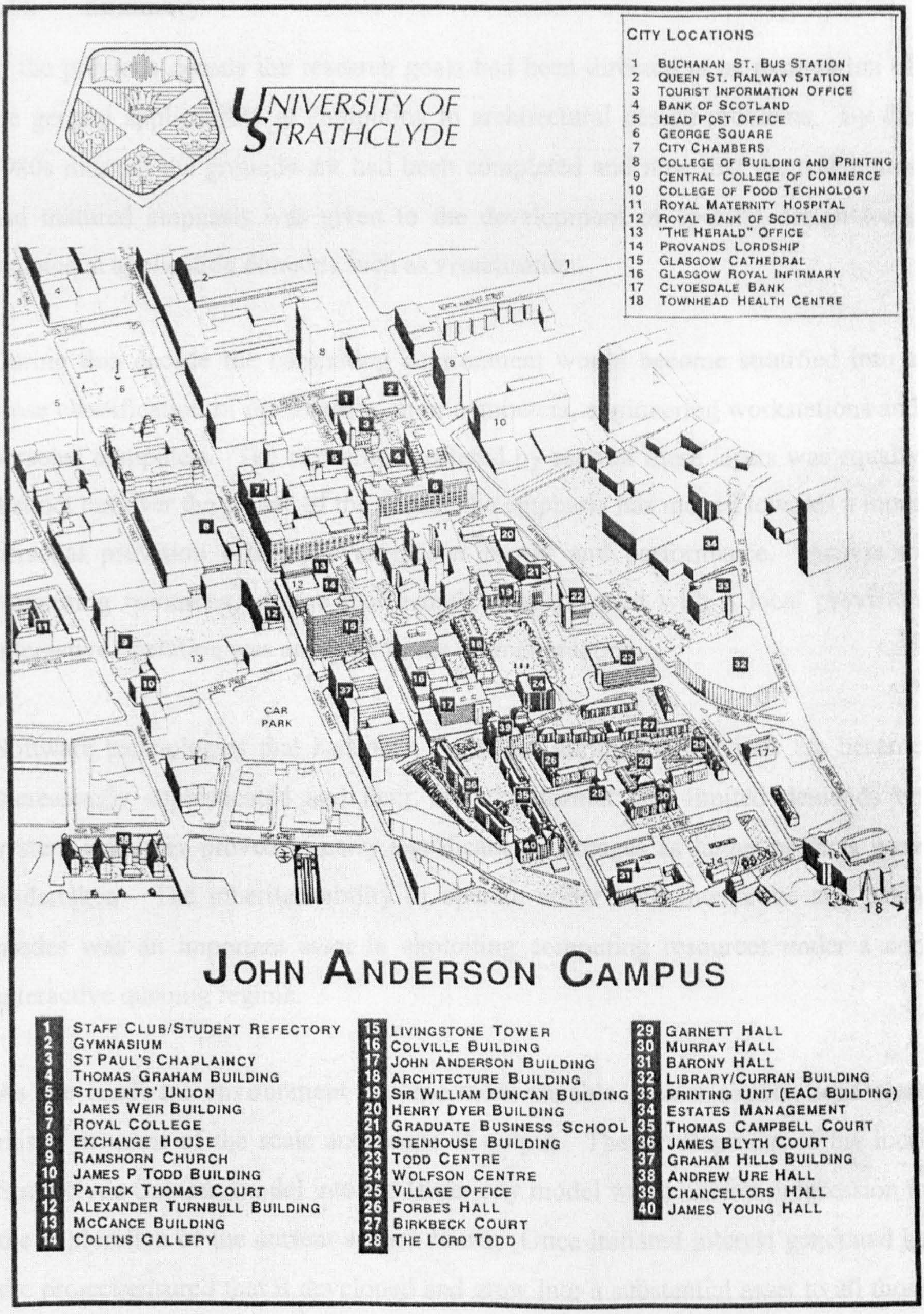


Figure 2.39 Strathclyde University campus guide

2.16 Summary

In the previous decade the research goals had been directed at an exploration of the general applicability of computing to architectural design problems. By the 1980s most of the groundwork had been completed and now that the techniques had matured emphasis was given to the development of specific design tools targeted at application domains such as visualisation.

During this decade the computing environment would become stratified into a loose classification of mainframes, mini computers, engineering workstations and personal computers. The capabilities offered by each of these layers was equally distinct but over the course of the decade the emphasis has moved towards a more personal provision offering ever greater power and performance. Access to computing resources became increasingly available and with a local provision interactive operation was accepted as the normal practice.

Software technologies that had been pioneered during the previous era became increasingly sophisticated and their modular format and limited demands on system resources proved to carry significant advantages as larger projects were undertaken. The inherited ability to operate under both interactive and batch modes was an important asset in exploiting computing resources under a non interactive queuing regime.

As the hardware environment became more capable greater expectations were raised in terms of the scale and scope of output. The development of the local Strathclyde Campus model into the larger city model was a natural progression in the exploration of the current system limits. Once initiated interest generated by the project ensured that it developed and grew into a substantial asset to all those with an interest in the city. This development had to break new ground in techniques relating to data capture, collation and structuring. The growth of the model brought additional problems in data management and forced the development of new methodologies and tools for interacting with large data sets. These interface paradigms have been reused in all subsequent urban models.

While many innovative features were developed at this time the computing environment was not yet mature enough to provide all the capabilities demanded by the vision of an urban information system. The scarcity of colour graphics precluded the inclusion of image data and a static depiction of the city prevented exploration by any dynamic means.

Chapter 3 The Edinburgh Model

The mini computer had previously enjoyed the prime share of what might have been called the scientific and technical market. This was due to the lack of any competition offering the same capabilities or performance. From the mid 1980s this situation began to change with the introduction of the engineering workstations. A number of key technologies helped to gain an increasing market share for the workstations, high on this list was graphics performance. In a multi-user mini computer environment communications between a user's terminal and the processor had always been the limiting factor in interactive graphical use. This was due to the finite bandwidth provided by the 9600 baud serial lines required for a distributed user base. While these data rates were not so restrictive for character based applications the transmission of graphical data was much more problematic.

The technology required for a wireline rendition of geometry offered an economical transfer medium in that little data was required to describe the scene being depicted. As applications technology developed there was a growing interest in generating rendered images that were pixel based as opposed to being constructed from vectors. This new technology could offer colour, shadow and texture, but only at a price. Graphics software that generated colour images was notoriously computationally intensive and so it was not uncommon for a complex scene to take many hours to render. This was not unduly restrictive as large jobs could always be submitted to a batch queue but in order to visualise the output the data had to be downloaded to a local terminal. A high resolution image at a resolution of 1024 by 768 with a colour depth of 24 bits would take over half an hour to transmit at 9600 baud. This was obviously a severe limitation to using interactive imaging within the design process. Some software such as VISTA (Stearn 1982) attempted to circumvent this disadvantage by utilising a polygon and tile based image format thus combining the advantages of a colour display with efficient data trasmission.

The advantage of workstations was realised in employing local processing and a fast internal bus between the host processor and the graphics subsystem giving an instantaneous screen refresh from local graphics memory. While bitmapped displays supporting a multi windowed environment offering the ability to display diverse graphical media within an easily operated user interface could point the way towards future goals there was still one fundamental problem to be overcome.

Despite the advantages of colour graphics, bitmapped displays and the speed offered by the bandwidth of the new architecture there was still a need to process the scene before the output could be displayed. Only when the processes of transformation, rendering and hidden surface removal could be performed at realtime rates could a dynamic experience of architectural space be explored.

3.1 The Visualisation Pipeline

As discussed previously the VIEWER program was the pivotal application for the suite of software tools used to interact with the Glasgow model. Although VIEWER's performance was impressively fast it was still not capable of processing graphics at interactive rates. If these large scale urban models were to be regarded as anything other than a database capable of generating a series of static views there was an obvious requirement to perform real time interaction with the model. The solution to this need appeared in the introduction of a new type of hardware architecture, the graphics engine.

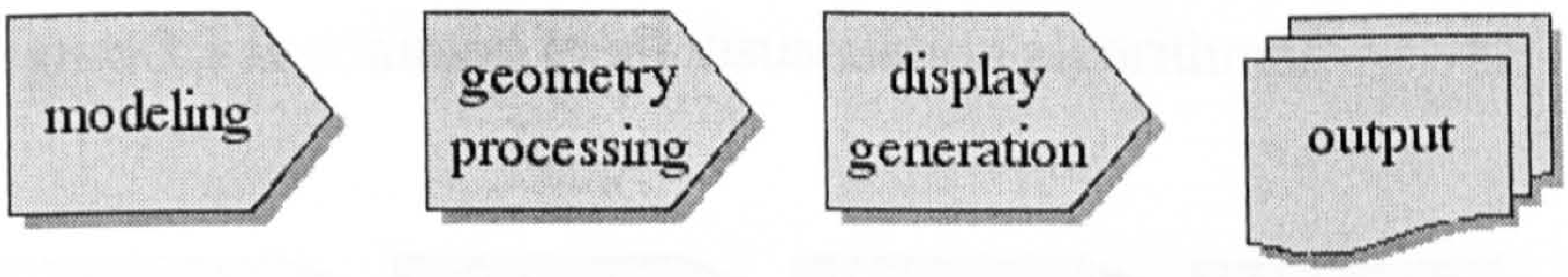


Figure 3.1 Conceptual view of the graphics pipeline

To understand this concept it is worth looking at the methodology employed by the VIEWER implementation as previously described. While not directly analogous the concept is broadly similar as regardless of the approach all these tasks are common and must be executed in the same order. These divisions can be conceptually visualised as components of some pipeline as shown in Figure 3.1.

In the ABACUS software implementation these components were formed by separate software programs connected by intermediate data files, see Figure 3.2.

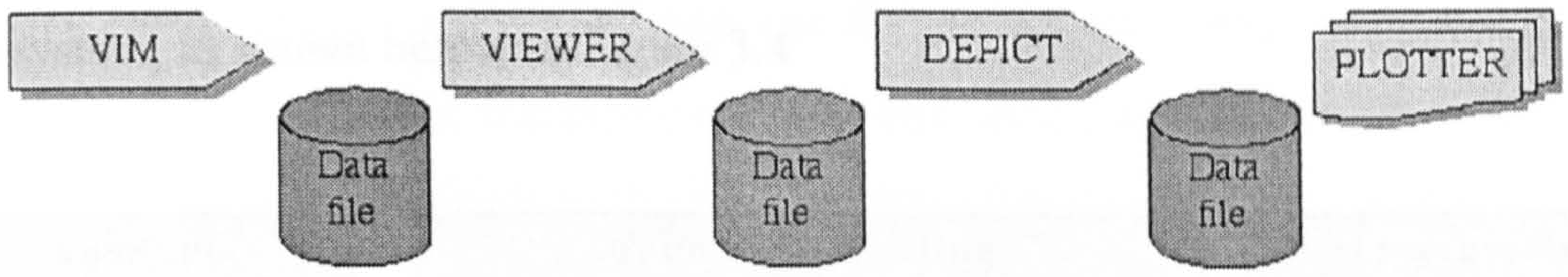


Figure 3.2 Pipelined application modules

This rationalisation of the process was in part dictated by the current computing climate and although an economical solution also suffered from a number of redundancies. Greater efficiency could be gained from integrating much of this functionality into a single module.

Stepping inside the methodology reveals that VIEWER is possessed of the algorithmic components of the pipeline in that it was capable of taking a description of geometry and through a process of transformations and hidden line removal render the output to some device, however extended functionality was offered by the contributing programs. This algorithmic pipeline as shown in Figure 3.3 is common to all visualisation algorithms.

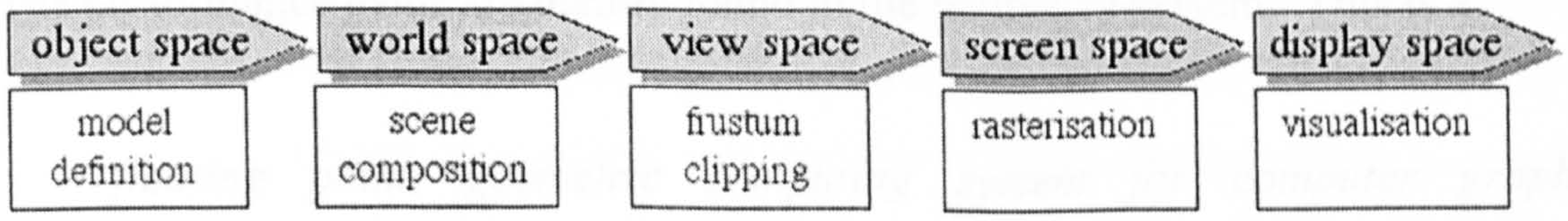


Figure 3.3 Visualisation pipeline

3.1.1 Hardware Acceleration

The visualisation process involves a series of transformations from one coordinate system to the next, the problem is that although the transformation is a simple computation the calculations have to be repeated for every geometric primitive sent down the pipeline. The solution was sought by isolating the subroutines responsible for this series of transformations and implementing their functionality in hardware. This in turn re-orders the concept of the pipeline into three functional components: the application process, geometry pipeline and raster subsystem, as shown below in Figure 3.4

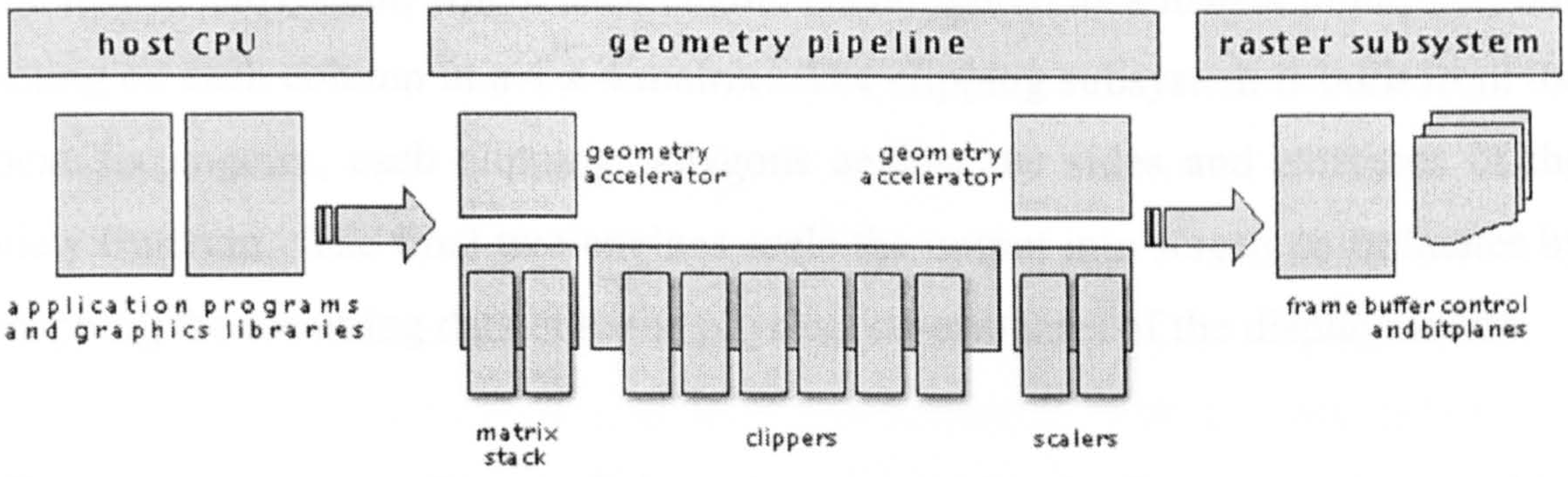


Figure 3.4 Geometry pipeline

The application program controls the current view co-ordinates and executes any scene composition or set-up routines before passing the geometrical data to the next stage. The geometry pipeline performs matrix transformation of the data, clips and normalises the co-ordinates then scales the transformed, clipped co-ordinates to screen space. This output is then passed to the raster subsystem which renders the polygons to the frame buffer.

The performance gains are mainly found in the geometry system. This is a:

> *"floating point, geometric computing system for computer graphics constructed from a basic building block, the Geometry Engine"*

In its original incarnation each geometry engine is a single chip containing around 75000 transistors forming individual arithmetic logic units (ALU). (Clark 1982). Every group of ALUs is capable of performing a custom operation on a four component vector in floating point format and can be configured by software to perform matrix multiplication, clipping and perspective projection.

Taking a general view of a generic geometry pipeline implemented in hardware there are twelve geometry engines set sequentially between a pair of geometry accelerators. These accelerators perform floating point format conversions and provide an i/o buffer for the pipeline. The first four geometry engines comprise the matrix subsystem, where each engine is configured as a matrix multiplier, one acting on each column in a 4 x 4 matrix. The clipping subsystem is built from the next six engines, each clipping polygons against the sides and extremes of the view frustrum. The final two engines scale the output into screen co-ordinates by mapping the incoming data onto the physical co-ordinates of the display device.

The geometry system discussed above is an example of a first generation graphics architecture (Akeley 1993), typical of the technology implemented in the Silicon Graphics IRIS 2400 and the Apollo DN570. At this time floating point hardware was just becoming available, memory for frame buffers was expensive and application specific integrated circuits (ASICs) were rare. Because of the limitations within the frame buffer this style of architecture was predominantly directed towards the acceleration of the geometric portion of the pipeline – such as transformation and clipping - rather than any algorithmic procedures – such as lighting, rendering and hidden surface removal.

High memory prices coupled with the limited processor performance of early engineering workstations had left the traditional mini computers as the platform of choice but the advent of graphical workstations with geometry engines gave notice of a new dynamic outlook to geometrical manipulation. An early example such as the Silicon Graphics 2400 was equipped with 8MHz geometry engines capable of performing 85700 3-dimensional 32 bit floating point transformations a second.

This gave impressive animation capabilities for a wire frame model but transformation rates dropped dramatically when polygons were rendered. The lack of hidden surface removal negated most of the advantage given by the geometry acceleration as this had to be performed in software and it was not until the arrival of the second generation machines that near real time animation with fluid navigation could be attained.

Typical of the low end of these early second generation machines was the Silicon Graphics Personal IRIS 4D25. With lower memory prices and better ASICs the frame buffer capabilities had been improved with the addition of hardware lighting, rendering and hidden surface removal. The claimed performance figures for this machine show a peak throughput of 20,000 lit and rendered independent quadrilaterals a second although the figure of 200,000 3D vectors a second for the same architecture indicates that frame buffer operations were still the limiting factor. These benchmarks were based on a 4 sided 10 pixel polygon and a 10 pixel vector respectively so real world performance was significantly lower. However this still represented a considerable improvement over using software routines not least because each graphics engine was capable of 20 million floating point operations per second while the host's MIPS R2000 CPU could only manage 0.9. (SGI 1988)

3.2 Software Issues

In order to exploit these new workstations ABACUS developed a new software application, ROVE the Realtime Object Visualisation Environment (ASL 1989) This application was designed to occupy the same functionality as VIEWER yet take advantage of the real time animation capabilities of the new hardware. The input data was identical to the existing format with the addition of extra attributes to allow for colour definition and assignment.

The adoption of an interactive as opposed to static interface to the data posed a set of new challenges. Instead of entering co-ordinates designating the eye point and focus point the user could now move through the scene changing the viewpoint at will. This required a new interface paradigm that was capable of giving the user total freedom to go anywhere and look at anything. This is more complex than it might first appear as there is a need to differentiate between the movement of the observer's location and the orientation and direction of their view. It was found to be important to de-couple these two functions since, if they are incapable of independent movement, the constraints on exploring the scene are considerable.

-	-	R	turn right
-	M	-	move forward
L	-	-	turn left
-	M	R	turn down
L	M	-	turn up
L	-	R	move backwards
L	M	R	rotate

Table 3.1 Mouse button movement commands

q	move up
w	move down
z	zoom in
x	zoom out
h	display help message
r/R	reset to default view parameters
N	input new geometry file
A	turn the cursor on/off
a	axes and data on/off
b	back-facing planes on/off
B	switch between single/double buffer mode
d	turn depth-cueing on/off
g	ground and sky on/off
i	filled / outline mode
f/F	faster x1.5 x10
s/S	slower x1.5 x10
t/T	thinner/thicker lines
m/-	continuous rotation on/off, switch rotation direction
p	run the commands in a path file
ctrl-p	re-run the commands in previous path file
.	change the current active layers
+	add some new geometry to the database
X	turn on/off display of the bound box of the model
$	turn on/off shadows cast on the ground
F1 key	change to high resolution monitor
F2 key	change to PAL (video) monitor
ESC	quit

Table 3.2 Keyboard commands

The user controls their movement through the model by moving the mouse as well as by pressing the mouse buttons. These movements are forwards, backwards, left and right and are achieved by moving the mouse in those directions. The offset from the centre of the screen dictates how fast in any particular direction the user will move, the greater the displacement, the faster the resulting motion.

The standard three button UNIX™ mouse offers seven combinatorial selections which are used to orientate the view vector. The exceptions to this rule allow the viewer's location to translate along the direction of the view vector. When combined with keyboard commands the use of the cursor position and mouse buttons allows the user complete freedom of movement in all axes.

3.3 A New Route to Data Capture

The original Glasgow Model was pioneered in order to study the methods of data capture, storage and representation required by large scale urban geometrys. Over the period described in the preceding chapter the capacity and capability of available computer hardware proved equal to providing a platform for these activities. However, whereas the hardware and software environment had matured in line with the scope of the project the single most limiting factor remained unchanged. In all the projects described up to this point, the collection, description and implementation of geometry has remained the one most difficult, time consuming and error prone activity. Unless this restriction could be removed the implication would be that the scale and scope of constructing large scale urban models would always be constrained.

3.3.1 Ordnance Survey Digital Data.

Within the UK, the Ordnance Survey had always been the single body charged with acquiring and maintaining cartographic products at a national level of coverage. Given that this resource is held to be the definitive data set from which all other similar mapping products are either derived or referenced within, then it would appear prudent to adopt this data as the prime source of two dimensional scale and location.

In the late 1980s the Ordnance Survey basic scale plans covered the scale range 1:1250 to 1:10,000 and comprised a total of over 220,000 sheets. 1:1250 was the largest scale of publication and the 55,000 sheets in the series covered the major urban areas. The 1:2500 scale series covered most of the country while areas of limited information content such as the "mountain and moorland" regions were supplied at 1:10,000. Urban areas are covered at 1:1250 the largest scale of publication.

Previously attempts to format building geometry had been made by manually digitising data from the OS 1:1250 series maps. To date this had proved the most productive method of data capture although it was only capable of providing a treatment of planform and could not address the three dimensionality of the building stock. Despite being the most efficient available method this route still proved to be tedious and time consuming. This lead to many errors in the data capture and, for reasons of speed, many of the available and perhaps pertinent, features were ignored.

However a fortuitous meeting between ABACUS and members of the Ordnance Survey at the 1988 annual conference of the British Cartographic Society lead to the establishment of a collaborative project to investigate the suitability of the emerging OS digital data for use in generating three dimensional computer models. (Jones 1990)

3.3.2 Ordnance Survey Digital Data Specification

During the 1980s the Ordnance Survey had been producing an increasing number of its published large scale maps by digital methods and was gradually shifting it's in house archive from film based media to magnetic tape. There was a dual aim to this process in that while both editing and production of conventional products could be streamlined there was also a growing demand from the existing customer base to acquire digital data formats for use on their own computer systems. Although the OS had maintained digital data, for in house use, since the mid 1970s the large scale supply of magnetic media to customers had only commenced in 1983. The initial data format had been known as the Digital Map Base (DMB) but the evolution of the product had resulted in an enhanced format now called the Ordnance Survey Transfer Format (OSTF). In 1989 the OS were unable to supply data on any other media than 2400 ft magnetic tape reels as it was perceived that the lack of physical standardisation of other formats would preclude the economical operation of such a service. (OS 1988)

The OSTF data was supplied in a FORTRAN readable format with a fixed 8 character record length, in blocks of 1800 characters. The data was organised so that all line, point and text detail required to produce a map of an area appropriate to the source map scale are held in one logical unit, but detail features are intermingled. In OS terminology a feature is a point, line or series of lines forming a coherent object on the ground or, in the case of a virtual entity like text labels, on the map. A feature may be considered a subjective entity in that so long as the constituent lines belong to the same feature code then any one feature may not have a one to one correspondence with the object being described in the real world. For example a building may be described in OSTF data as being made of one or more 'features' which may or may not have the same description. The consequence of this unstructured definition will be dealt with later. (OS 1983)

3.3.3 Digital Map Data Accuracy

Due to the nature of the digitising process employed by the OS, which converts draughtsman's lines into their discrete digital equivalents there is an inherent accuracy factor due to the line width of the basic scale resolutions. This is enumerated in Table 3.3. The consequence of this conversion is that, in addition to any rounding errors incurred by floating point accuracy, the lack of any "snapping" functionality in the production process means that junctions may not be totally precise and polygons shaped by the original data may not be well formed.

In the process of converting the traditional media base to the new digital media the data adopts an unstructured format that mirrors the system operators methodology and may appear as disjointed strings of co-ordinates captured in an apparently random sequence. For example an operator might be following a boundary line but then continue to pursue the outline of a wall before returning to the boundary.

Basic Parameters on OS Large Scale Maps			
Scale	1:1250	1:2500	1:10000
Format	Square	Square	Square
Extent	500m	1000m	5000m
Grid Interval	100m	100m	1000m
Grid Interval	50m	100m	500m
Digital Resolution	0.05m	0.1m	0.5m
RMS Error	0.4m	0.8m	3.5m
Minimum line width	0.1mm	0.1mm	0.1mm
Ground	0.125m	0.25m	1.0m
Modal Line Width	0.2mm	0.2mm	0.2mm
Ground	0.25m	0.5m	2.0m

Table 3.3 Basic parameters on OS large scale maps.

While perfectly adequate for the intended purpose of map production this procedure results in a specific series of potential problem areas when it is intended to use the data for geometrical modelling. This characteristic, and the need to parse the OSTF data file to extract only relevant features, lead to the need for an intermediate stage before the data could be usefully employed as the basis for three dimensional building modelling.

3.3.4 LaserScan.

The software tools required to process the OSTF data in order to provide a logical structure to the co-ordinates describing the required features are far from trivial. It was fortunate that ABACUS were able to gain access to this capability through a member of the OS who was studying within the Department of Topographical Science at Glasgow University. In exchange for access to the modelling capabilities within ABACUS a collaboration was constructed that linked the OS, ABACUS and a commercial company LaserScan. LaserScan were among the pioneers of digital data gathering and had a long standing collaborative link to the OS through their work in automatic digitising and photogrametry. (LaserScan 2001).

LaserScan were able to provide a suite of programs known as the LaserScan Automated Map Production System (LAMPS). LAMPS consisted of a number of modules which format, process and display a range of vector digital data. All modules operate on data held in LaserScan's Internal Feature Format (IFF).

3.3.5 Feature Selection

Since the data had to be imported from the OSTF format into LaserScan's own IFF format it was possible to exercise a process of feature selection. In theory all information displayed on OS basic scale plans could be incorporated into a 3-D model but some would be of doubtful relevance.

File	Feature Code	Feature
Buildings	fc1 fc2	Building - outline Building - Pecks
Boundaries	fc7 fc8 fc9 fc10 fc79	Boundary Parish/community Boundary - District Boundary - County/Region Boundary - Electoral Boundary - Parliamentry
Roads	fc15 fc21 fc30 fc59	Railway - standard gauge Road metalling - pecked General Line Detail Water Feature
Seeds	fc321 fc323	Building Seed Glasshouse Seed

Table 3.4 Grouped feature codes

The full specification of cartographic elements contained in the OS map base far exceeds the requirements of the basic data capture required. In keeping with the methodology employed in the original Glasgow project it was determined that only elements relating to buildings, road networks and heighting information should be necessary. In the event it was recognised that all of the general line detail could be concatonated into the generic road network category. In order to cater for possible future uses of the data set it was also decided to extract two further feature sets, those relating to administrative boundaries and also the co-ordinate points of all building seeds. A policy of positive selection was therefor applied - grouping salient features into the four categories as described above. This is enumerated in Table 3.4 above.

- The buildings file contains the building outlines and overhangs. As the desired product must consist of closed features formed by linked co-ordinates these data require be further processed to form structured polygons.

- The boundary file contains all the information on the local and administrative boundaries. In the event this data was not used but could have had value in being able to integrate database information with the model.

- The road file contains data describing roads, railways, water features and other linear details which form important components of the urban landscape. The addition of general line detail provides ground cues and helps to link 3-D objects distributed over the terrain. These co-ordinates describe a series of disjoint vectors and require processing into sets of polylines.

- The seed file contains the co-ordinates of the building seeds and although not used could form the location of georeferenced attributes to the buildings and features to which they refer.

3.3.6 Terrain Data Capture.

Of the available sources of terrain height data, contours, spot heights and bench marks, none were available as a straight conversion from the digital media. Contours are only reproduced on the 1:10000 scale sheets which at that time were unavailable in a digital format. Also while it would have been possible to write a program to extract the co-ordinates of spot heights the actual elevation is contained within a character string and is held at some otherwise unknown location within the file with no structured link between the two. In the event both contours and point height data were manually digitised from traditional media.

At the conventional scale of paper based media contours can prove difficult to digitise. This is because, despite being depicted in a contrasting colour, they tend to merge with building outlines in urban areas and may often terminate against rock features in areas of broken ground. In order to provide a more open document from which to digitise, an enlargement on film of the combined grid / contour repromat was obtained. Once all the extraneous detail is removed contour string are more easily defined. Contour digitising was carried out in point mode resulting in approximately 4000 values over an area of 2 x 2 km.

Point heights were also manually digitised from the 1:1250 scale plans with each height being keyed in at the required co-ordinates. Because spot heights in urban areas tend to be distributed along thoroughfares, care has to be taken to ensure that they are representative of true terrain levels as some points may be substantially elevated, as can happen with flyovers and long bridges.

Once all the data had been collected it was processed through the LAMPS package to produce a DTM matrix on a 10m grid covering a 2 by 2km section. The background to this process has been discussed in Chapter 2, Other than changing the representation from row major to column major the DTM matrix only required the addition of a file header to make it fully compatible with the ABACUS software data format.

3.3.7 Conversion to Structured Geometry

Since the vectors defined in the OSTF standard form a series of disjoint strings it was necessary to post process the two major file types, the roads and the buildings, to attempt to structure the data into a format compatible with the conversion into usable geometry. There are a finite range of operations that are applied to both clean the geometry and to rationalise its final format. These operations were performed using the LaserScan Structure package.

The first pass through the data searches for features or some part of a feature that has been digitised twice. The user specifies a join tolerance within which the software will attempt to precisely align one vector with another. The magnitude of the tolerance is critical to the success of this operation. If the value is too small then some instances of double digitising may be missed, too high and there is a risk of amalgamating parallel but quite separate features. See Figure 3.5.

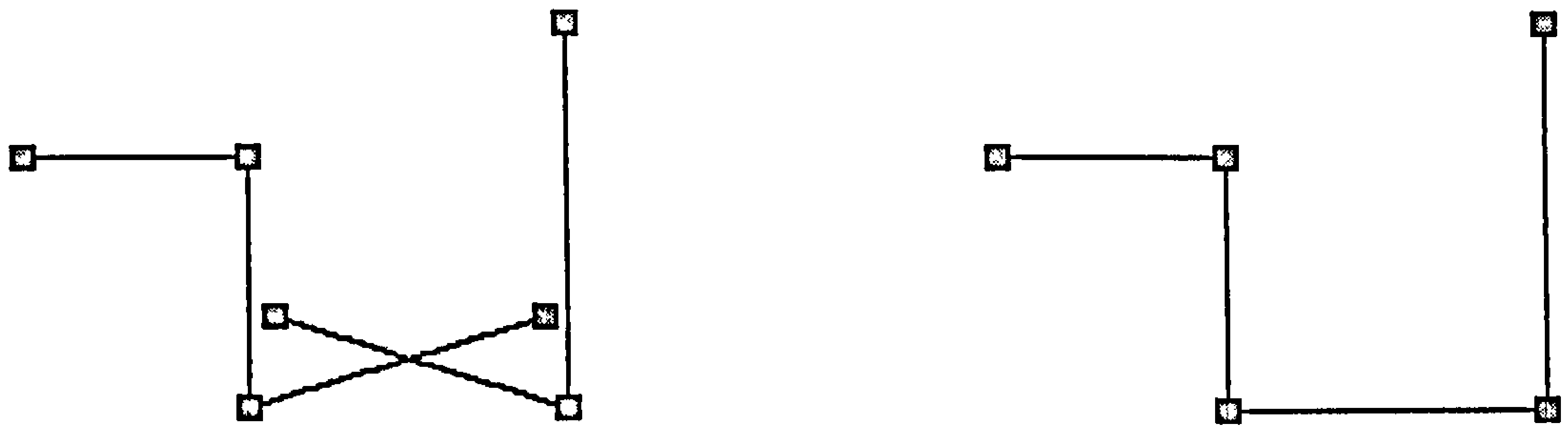


Figure 3.5 Aligning double digitised vectors

The next pass seeks out vector end point co-ordinates and if they are within the given tolerance makes them coincident. Two parameters are specified, the first of which ensures that any movement of the target co-ordinate on the current vector is constrained to move only along the direction defined by that vector. The second parameter allows the vector to move in a sideways direction. A final iteration through this function targets co-ordinates on separate vectors whose endpoints fall within the tolerance and attempts to make them coincident. See Figure 3.6

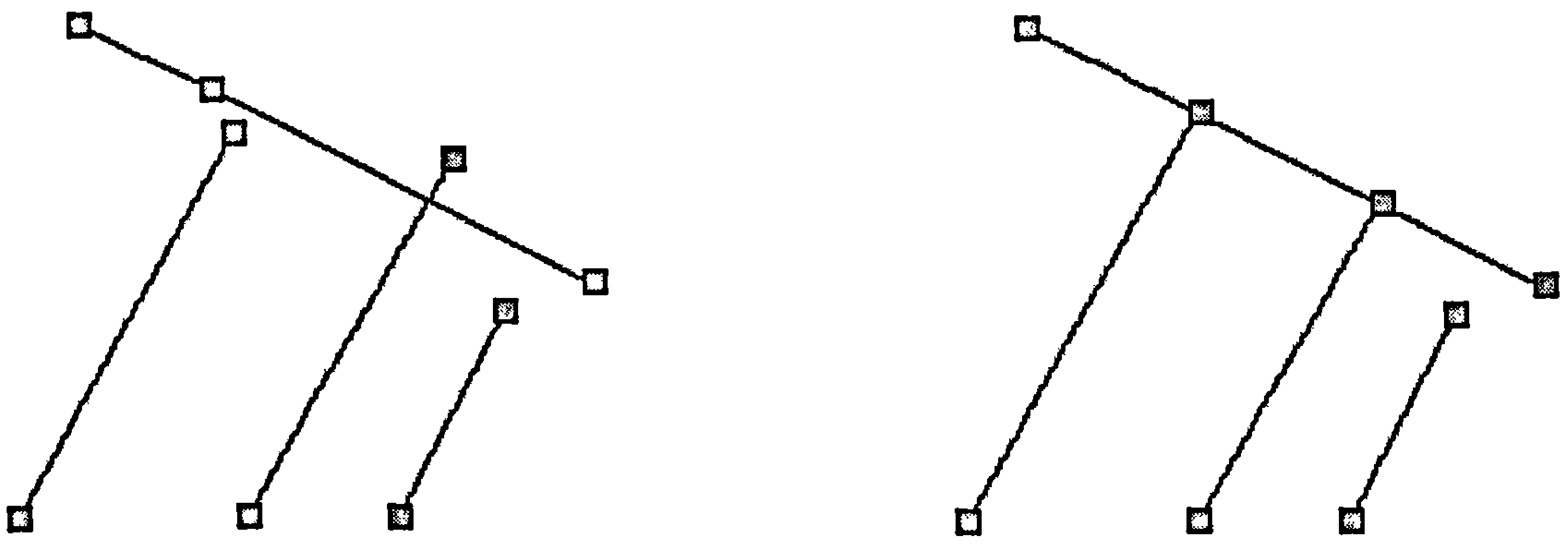


Figure 3.6 Aligning vector vertices within a given tolerance

Once duplicate vertices have been eliminated and all hanging junctions are welded a final process is applied to the data in order to ensure that no feature may cross another other than at a vertex. This also has the effect of eliminating 'T' junctions. See Figure 3.7.

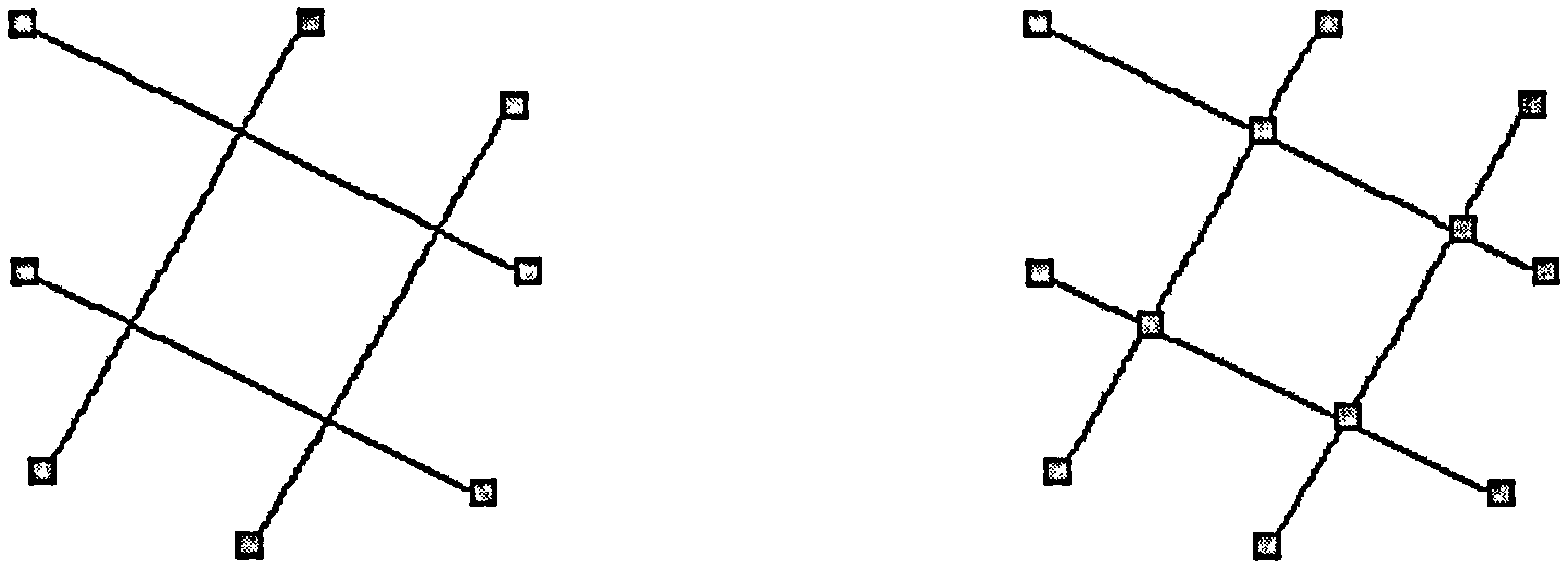


Figure 3.7 Eliminating crossing vectors

Once the geometry has been sanitised an automated polygon building routine is applied. This has the effect of joining individual vectors into closed polygons and, where polygons are nested, as in the case of internal courtyards, ensuring that the correct ordering of external and internal polygon contours is maintained.

3.3.8 Conversion to VIEWER Format

Once the geometry has been structured the format of the IFF data file contains a large volume of header information and a number of inter record fields which are not required in the VIEWER format. In order to parse this file structure and convert valid data into a form compatible with the VIEWER software a number of conversion utilities were written. All conversion routines shared the same logic in that the module would search the file for new record indicators, the characters NF, the next line is skipped and the first numerical field in the next line determines the number of vertices in the polygon or polyline. These are then read and rewritten to the output file in standard VIEWER format. For trivial examples the only modification to the co-ordinate data is due to the last co-ordinate triplet being coincident with the first, as shown in Table 3.5. However this is easy to trap. More complicated logic has to be applied when there are multiple holes representing courtyards with the outline of a building.

```
NF 2 1
FS 1 0 0 0
ZS 7 0
601.35    986.35    65.0
601.75    985.05    65.0
622.50    991.05    65.0
624.25    993.60    65.0
622.30   1000.00    65.0
597.40   1000.00    65.0
601.35    986.35    65.0
EF
```

```
RAG
6 70.0
601.35    986.35    65.0
601.75    985.05    65.0
622.50    991.05    65.0
624.25    993.60    65.0
622.30   1000.00    65.0
597.40   1000.00    65.0
```

Table 3.5 Comparison of simple IFF and VIEWER data formats.

3.4 Generating 3-D Geometry

In line with the strategy adopted for the Glasgow model the adopted modelling hierarchy addressed three sets of elements that in combination described the urban environment. This breakdown divides the topography into three sets - buildings, transport networks and terrain.

As previously discussed, the input data is parsed to provide a set of closed polygons that describe the 2-dimensional plan of each individual building in the data set. Once converted into the ABACUS native format two further operations are performed. The first augments each building plan with additional vertices at the intersections with the cell boundaries of the terrain grid. This serves to provide a better match at the intersection of the building with the ground.

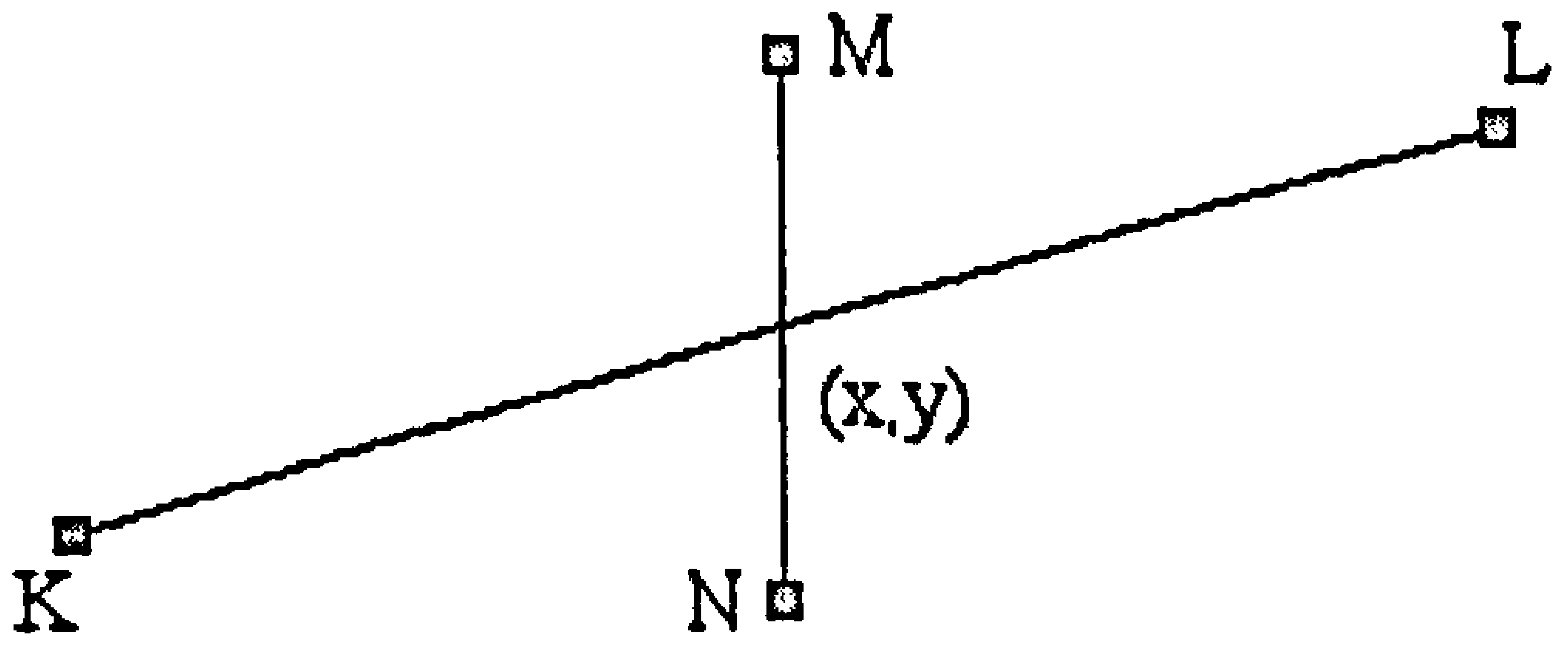


Figure 3.8 Intersection of outlines with terrain grid.

In the case illustrated in Figure 3.8 above, line KL is one segment of the boundary of a building's plan and line MN is the north / south grid line then the intersection point is given by :

$$x = x_m$$

$$y = y_k + ((x_m - x_k) * (y_l - y_k)) / (x_l - x_k)$$

This is a special case of the more general representation of the intersection of straight line segments and as such is simpler and more efficient saving seven multiplications and a division over the general method. (Bower 1983)

The second operation "floats" the building down onto the terrain by attributing each vertex with the height found by projecting the plan down onto the terrain model then interpolating over the grid to obtain the height of the contact point. To provide a 3-Dimensional volume for each building block an arbitrary height is assigned to each building by extruding the roof plane 10 meters above the average of the ground level at the plan. The outcome of this process is shown below and while the effect might be visually subtle the cohesion given to the model is important. The top image shows a building plan before augmentation with additional vertices at the grid crossings, the lower image after the process is complete.

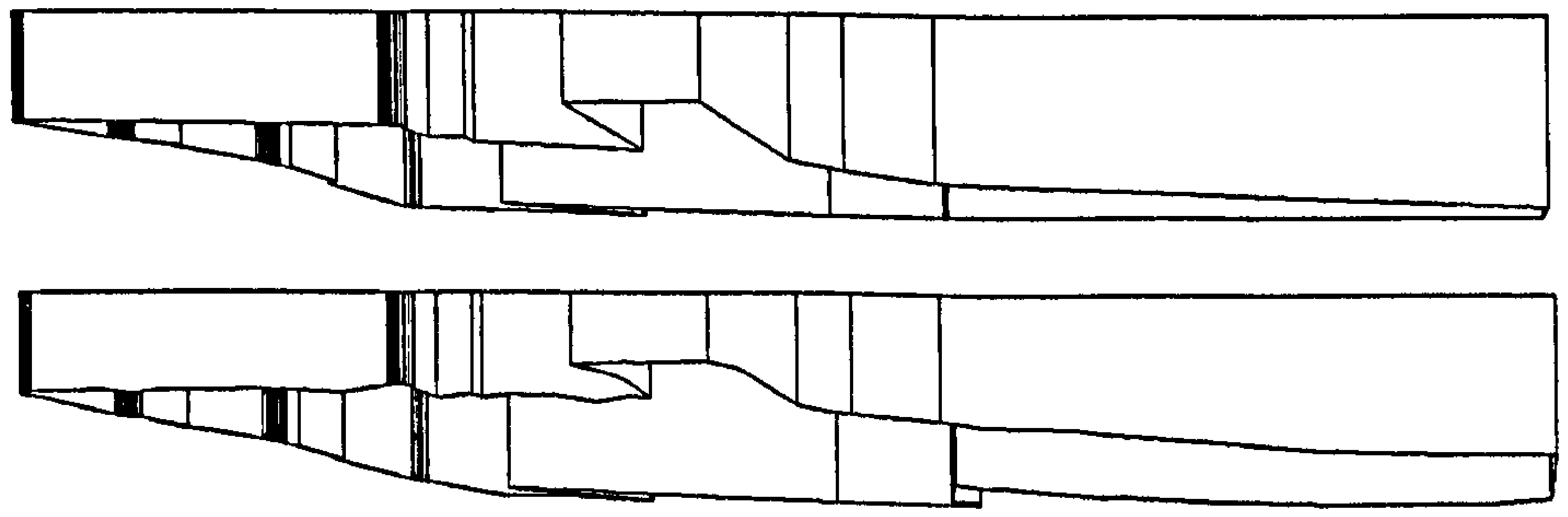


Figure 3.9 Augmenting building geometry at grid intersections

As regards the buildings these operations result in a passable facsimile of the urban context. Certainly, with regard to the man hours expended, this methodology results in an ability to rapidly prototype the first level of detail for large tracts of the urban landscape.

Although the preceding operations do result in a set of volumes that are broadly representative of the matching buildings there are still some intrinsic problems resulting from the adoption of the cartographic 2-D representation as the basis for a 3-D model.

The prime requirement of the data conversion process was that all the vector information should be transformed into closed polygons. To ensure that this process was completed the parameters supplied to the conversion algorithms were incrementally increased until all the data was successfully converted. A side effect of forcing this process to completion was that that less well fitting vectors were coerced into conjunction. This resulted in building outlines where regions of fine detail became distorted and needed manual editing to regain their true form.

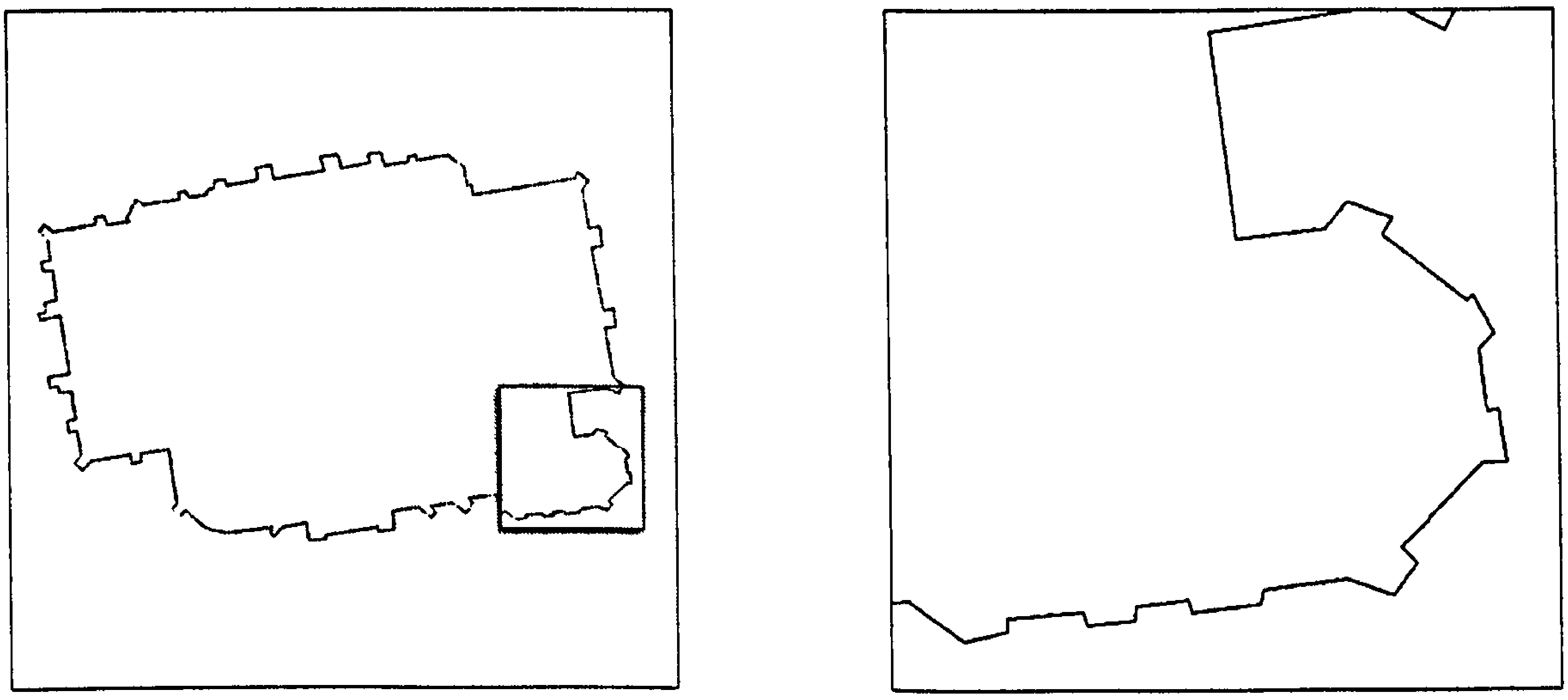


Figure 3.10 Shape distortion due to vector association

The next of these problems results from the manner in which buildings are depicted in a cartographic context. On a map it is common to represent the 2-dimensional plan outline as a series of polygons representing not only the outlines of a building's boundary walls but also inclusive of pecked outlines representing overhangs and in-gos. Similarly, columns and other architectural features may be represented and while all these outlines may serve to aid recognition in a 2-D depiction they tend to add spurious data to the 3-D form.

In order to assimilate the maximum amount of information it was decided to integrate the pecked outlines within the building data set. While this proved advantageous in that a template for future editing was provided it also highlighted the fact that a one to one correspondence between the cartographic representation and that required by volumetric modelling could only be obtained by human intervention.

Figure 3.11 below shows the plan of a building with an internal courtyard and groups of columns at each corner. The entrance is shown as a series of penetrating corridors linking the inner courtyard to the outside. In the extruded form it can be seen that these features are evidenced by the presence of separate bodies. In the extrusion process each body is given an average height of 10 meters plus or minus a small increment which helps to identify each element as an aid to picking within the editing process.

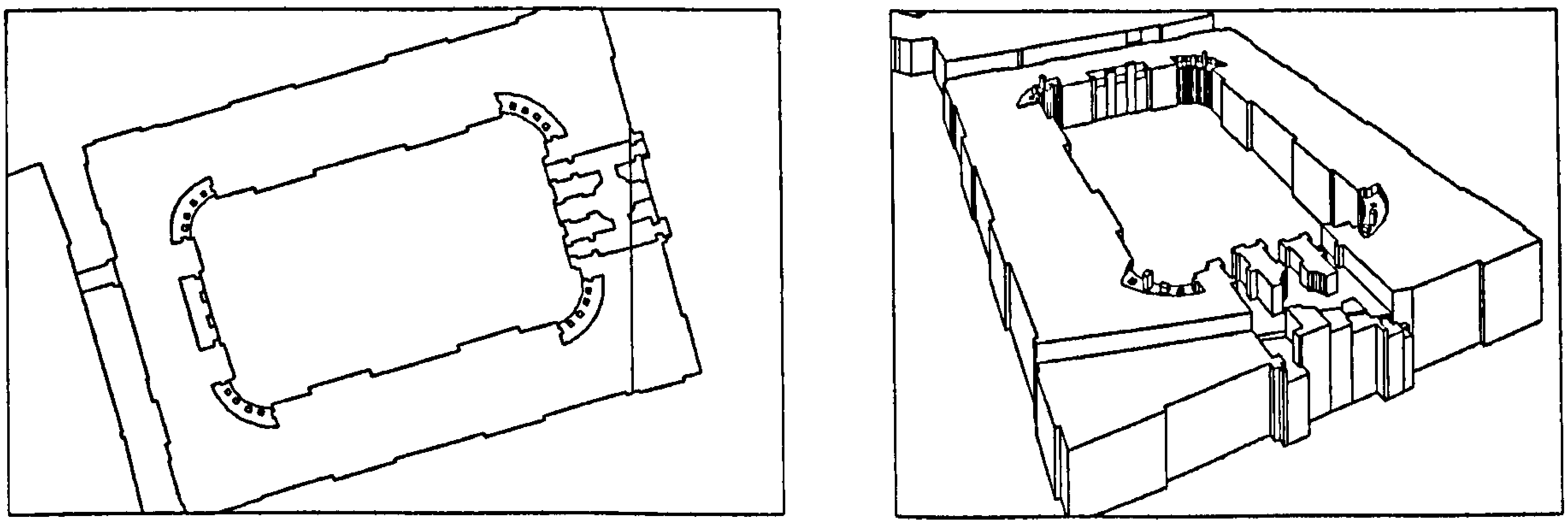


Figure 3.11 Consequences of including pecked outlines

A generic problem when moving between the 2-D and 3-D domains occurs where one feature overlaps another in 3-D space. The 2-D rendition is incapable of portraying this situation accurately. In this scenario the underlying structure is lost with perhaps only the extremities being shown. This situation is relatively rare but may be seen when a large bridge or other structure obscures buildings beneath.

The images in Figure 3.12 show how Waverley Station is bisected by the North Bridge.

Figure 3.12 Plan and axonometric of Waverley Station & the North Bridge

The last discontinuity produced by this conversion process occurs where the data bridges the boundary between two map sheets. In this case the building feature is separated along the sheet edge. In the figure below this horizontal discontinuity has been emphasised by separating the two portions by a small distance. These artifacts are identified during the edit phase and stitched together to provide a single volume.

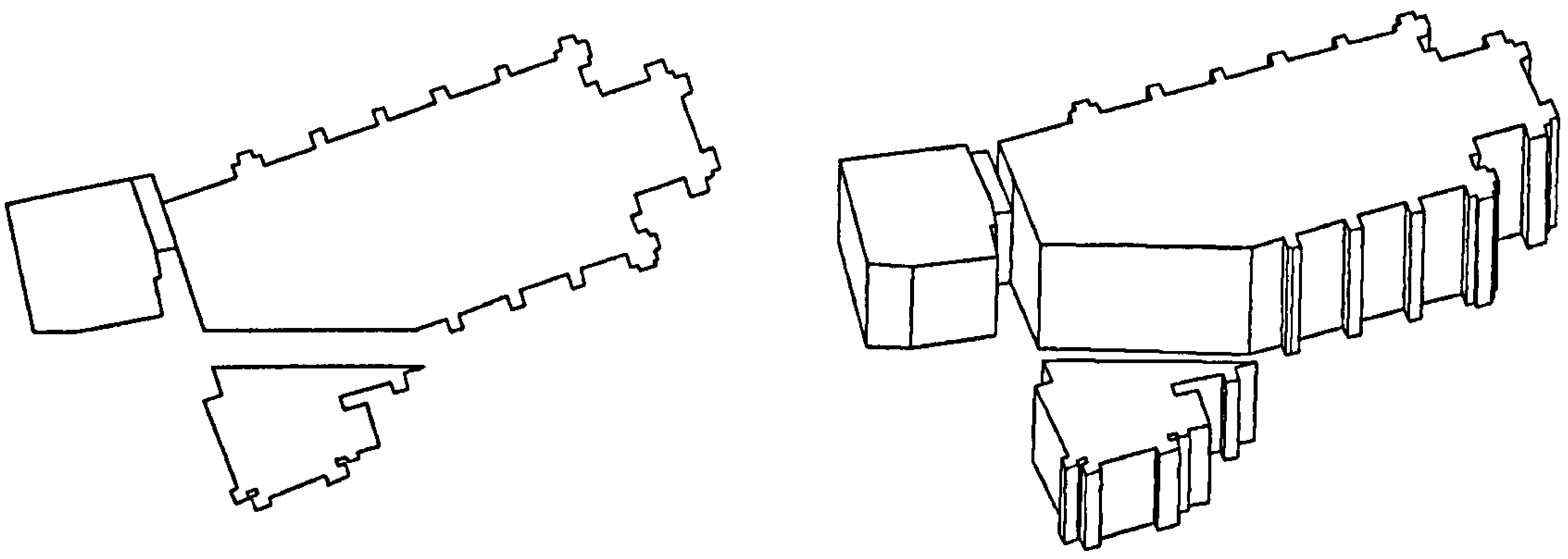


Figure 3.13 The separation of building outlines at map edges

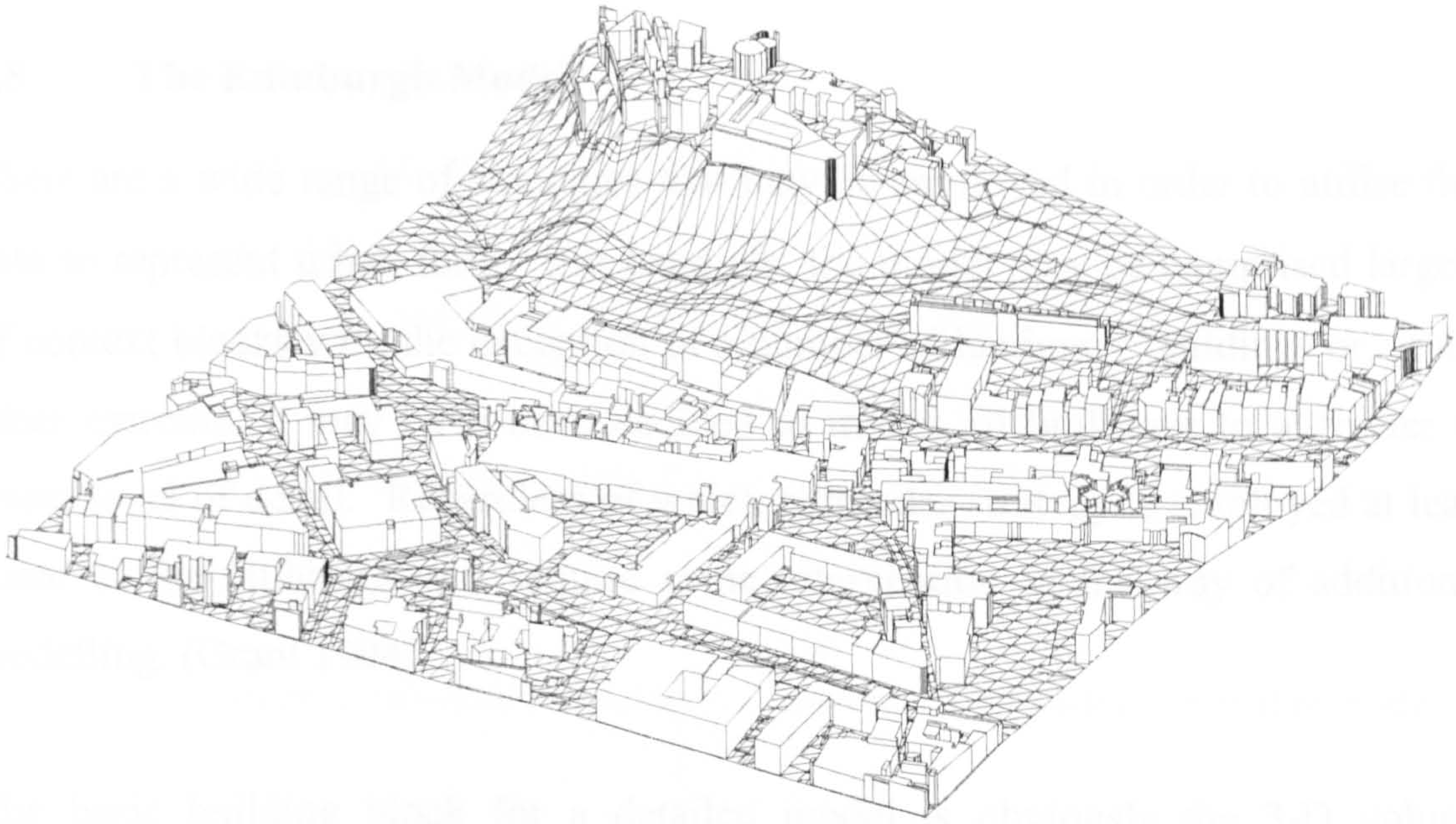


Figure 3.14 A complete map tile visualised by VIEWER

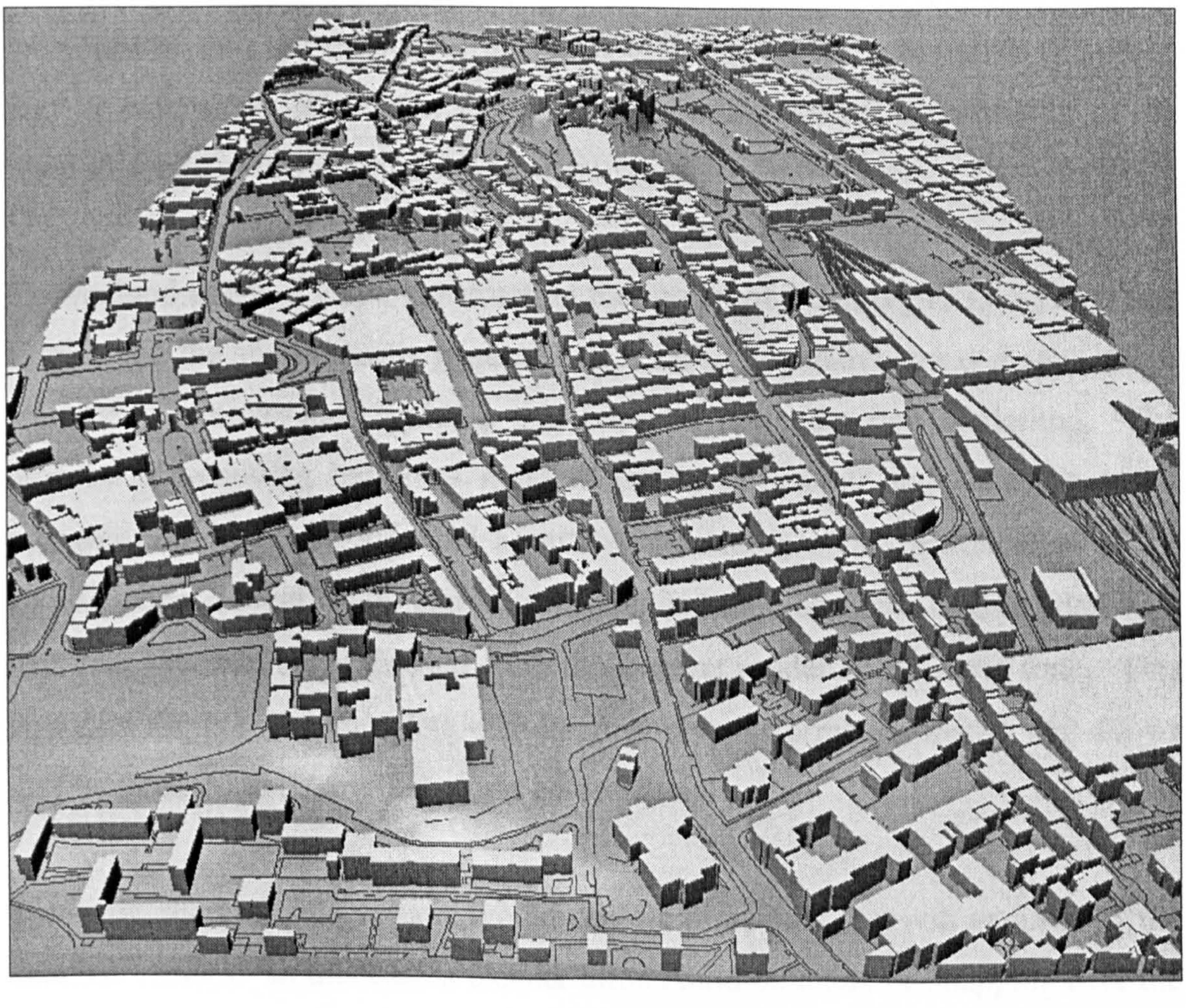


Figure 3.15 Rendered visualisation by ROVE

3.5 The Edinburgh Model

There are a wide range of strategies that may be employed in order to utilise this data to represent urban form. For example, the model may be comprised largely of context blocks with the exception of a number of landmark buildings or, at the other extreme, it may have been decided to model all buildings to a greater or lesser level of detail. Regardless of which particular strategy is employed at least some of the buildings will require some intervention in the way of additional modelling. (Grant 1994).

The basic building block for a detailed model is obviously the 3-D volume obtained from the original data source. However this is purely an extrusion of the 2-D plan and might not bear much relationship to the true 3-D form. Typical problems occur when the 2-D boundary contains spurious representative detail as discussed above but also when the outlines of secondary structures are incorporated into the outline of the main building. These secondary structures may be extensions, lean-tos or utility sheds and may not be representative of the main structure of the building.

The method adopted while constructing the model of Edinburgh's Old Town was to maintain the automatically generated data set as the main context model and to selectively remove each building that required more detailed modelling. The process of removing buildings required the identification and selection of all polygon outlines that comprised a particular building. This action allowed the user to select a local origin and normal axis and then copy the selected data to a new file, the data being deleted from the context model at the same time. This new data file is saved complete with the reverse transformation matrix that allows it to be returned to its original world co-ordinates.

Typically, the base 3-D volume would contain more data that was required by the detailed version. In this case a cleaner basis could be made be snapping over the template to provide the required outline. This also served to normalise the co-ordinates to a known degree of accuracy and to rectify any minor deviation from the co-ordinate axis. This procedure made for easier modelling operations but care had to be taken not to distort the building profile along shared boundaries or party walls otherwise there may be contention when individual buildings, modelled in local co-ordinates, were transformed back into world space. The image sequence if Figure 3.16 below illustrates the process of identifying a building from the map; transposing it to a local co-ordinate system; cleaning the outline and finally applying detail.

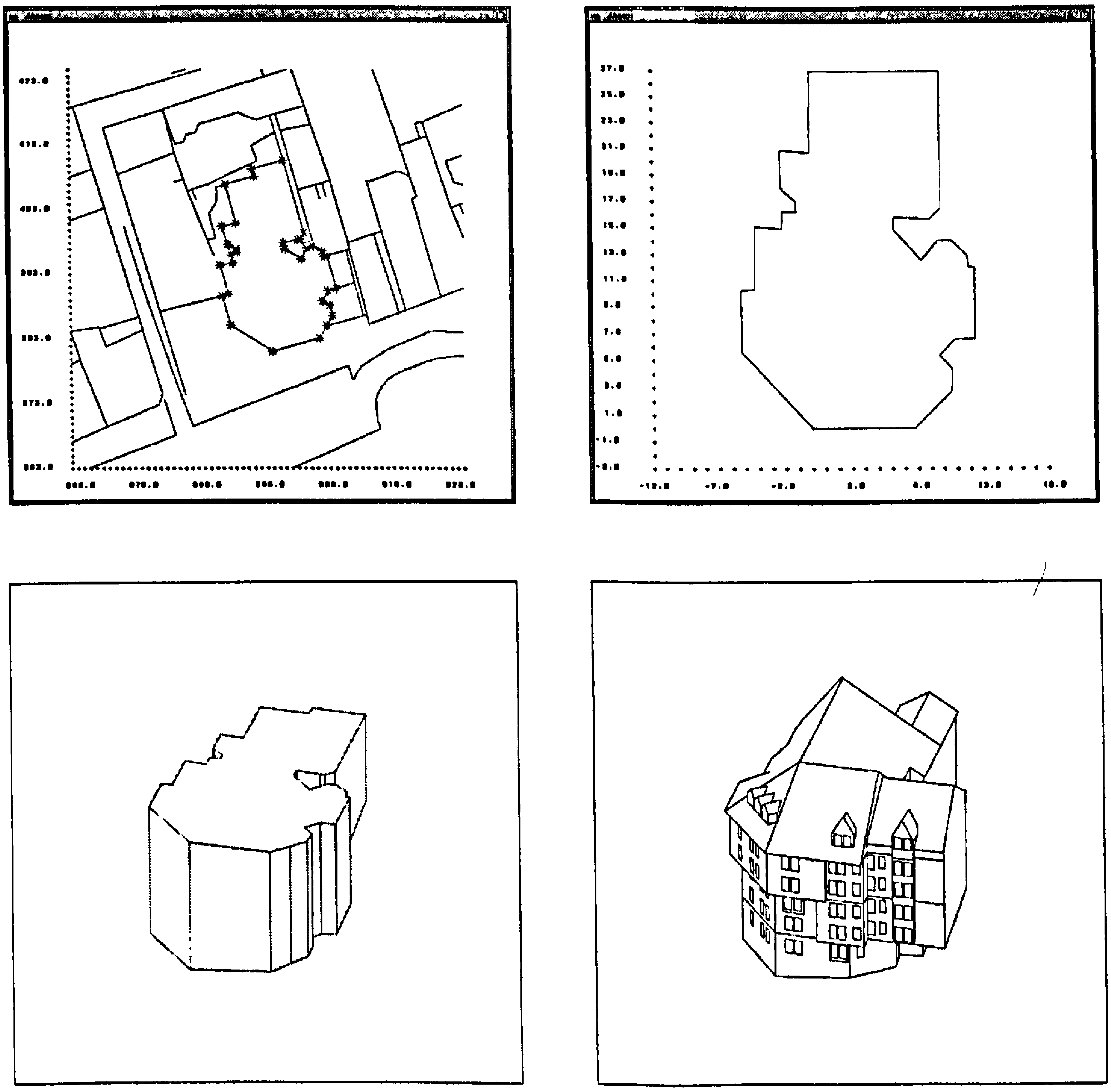


Figure 3.16 The three stages of modelling

At this stage of the process the data represents the base volume as a vertical extrusion within the planform of the building. Detail can be easily added by adding elements to the vertical surfaces of the building but there will be a large number of instances when the correct topology can only be found be removing smaller volumes from the main block. This could be realised by using Boolean operations to sculpt the true form but when dealing with less than perfect geometrical entities in a CAD environment it is generally easier to reconstruct the 3-D volume by using a number of smaller extrusions within the plan outline. One problem with this methodology is illustrated below in Figure 3.17.

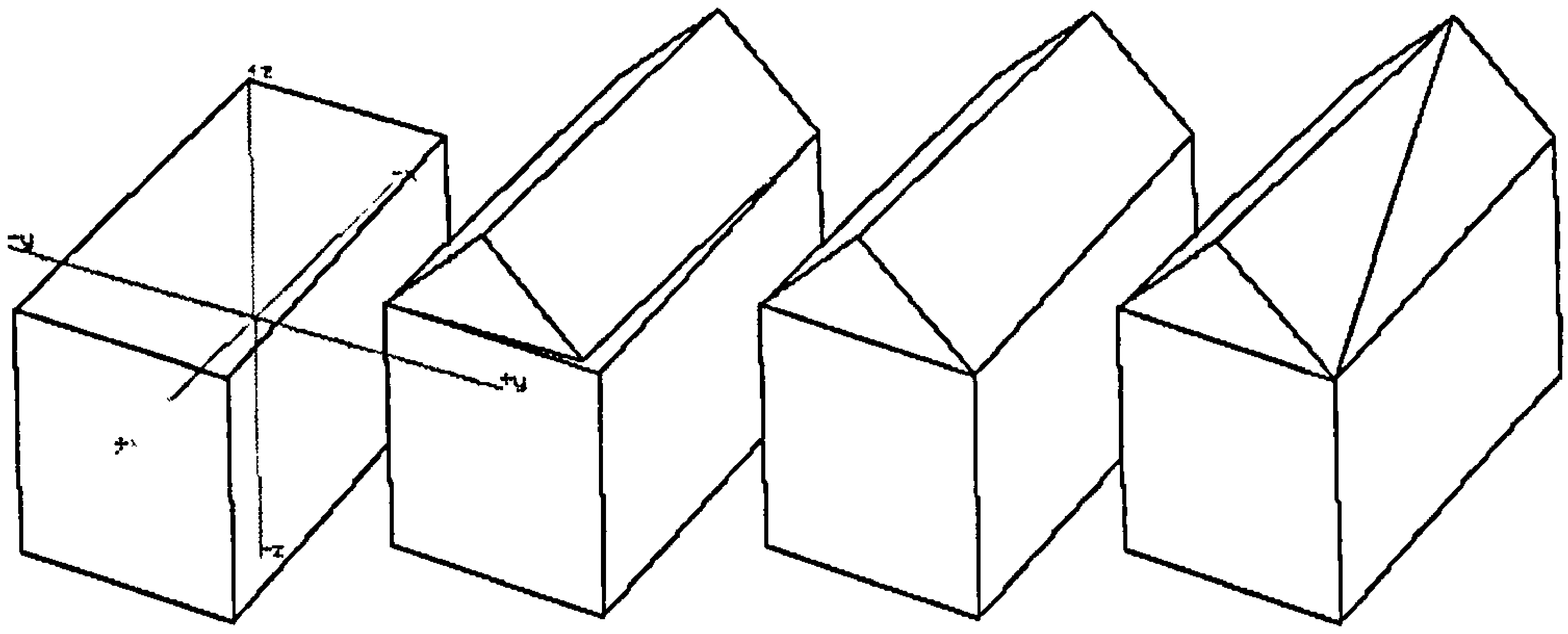


Figure 3.17 A simple building block made from a number of extrusions

A fundamental limitation of the software used to edit the building geometry was that it was constrained to work orthogonal to the world axes. This meant that the creation of new geometrical entities was commonly performed as an extrusion along a given axis. The historical nature of Edinburgh's Old Town means that the buildings are far from regular in plan or elevation and when the geometry is developed there are few instances where the regular shapes dictated by Euclidean geometry match the shape of the structure. Where this occurs the regular geometry produced by the modelling software must be edited to equate with the desired shape. This proves workable but results in surfaces which are non-planar and if grossly distorted can lead to visual anomalies when rendered. Any subsequent geometrical operation requiring accurate plane equations will also produce undefined results.

Depending on the physical nature of the building under consideration the geometrical development may be performed by using vertical or horizontal building blocks. As the rendering process of choice is performed by utilising flat shaded polygons the manner in which conjunct entities are abutted does not affect the final image. If the model were to be rendered as wireline, with hidden line elimination, the position and disposition of witness lines along borders may have been more critical.

Once the true 3-D form has been constructed the model can be embellished with architectural detail in order to more closely represent the building under consideration. In this instance all details are constructed from polygonal data. Areas of different materiality are represented with planar polygons applied to the facades. This approach can be applied to large areas, such as shop fronts or small details such as individual doors and windows. One problem evident with this methodology becomes apparent if the additional detail is applied in the plane of the external façade. Conventional rendering process are not able to resolve the presence of separate surfaces in the same plane. This causes "Z-buffer fighting" in a raster based imaging system and undefined visibility conditions in the wire line images produced by VIEWER. An attempt was made to resolve this conflict by standing off the detail polygons by a small factor. While this did ameliorate the problem with regards to the technical limitations of the system the solution only appeared visually satisfactory when the building was seen in elevation as when viewed parallel to the façade the detail elements appeared to float in front of the wall. A compromise was reached by representing detail as 3-d bodies as opposed to 2-d planes. This gave spatial continuity but at the expense of double the number of vertices and a number of redundant surfaces.

As in real world architecture some elements are common to a number of building types. This allows a limited library of basic architectural forms to be constructed and then applied with a suitable translation and scaling, saving both construction time and maintenance as well as reducing memory requirements and processing time. If all this additional detail is held separately then it can be applied selectively when the entire model is processed for visualisation. Although this does not represent a true implementation of a LOD hierarchy the ability to chose between full and partially detailed representations can be most useful as an aid to off loading processing in areas that are not visually critical.

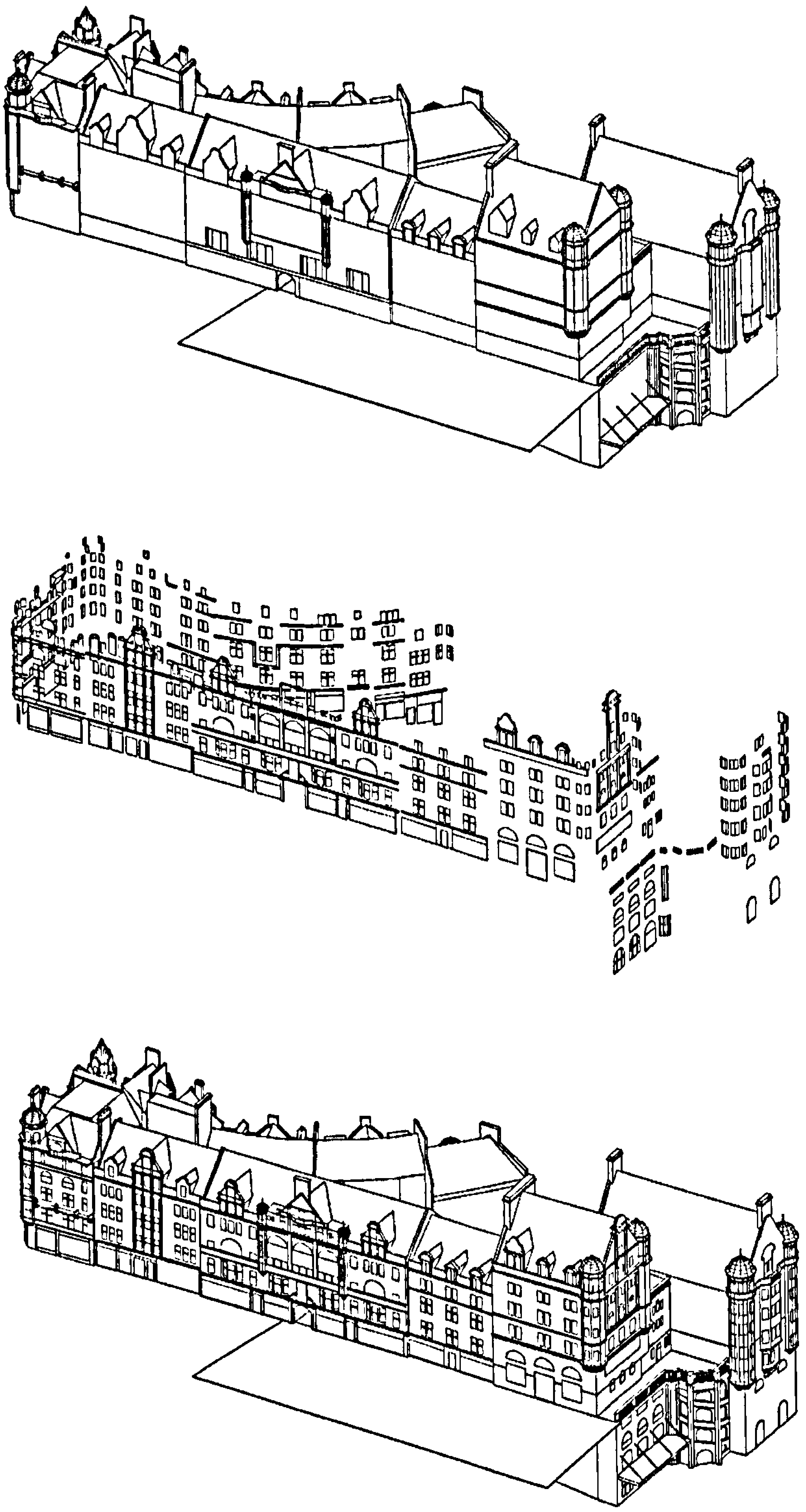


Figure 3.18 The combination of building volumes and details.

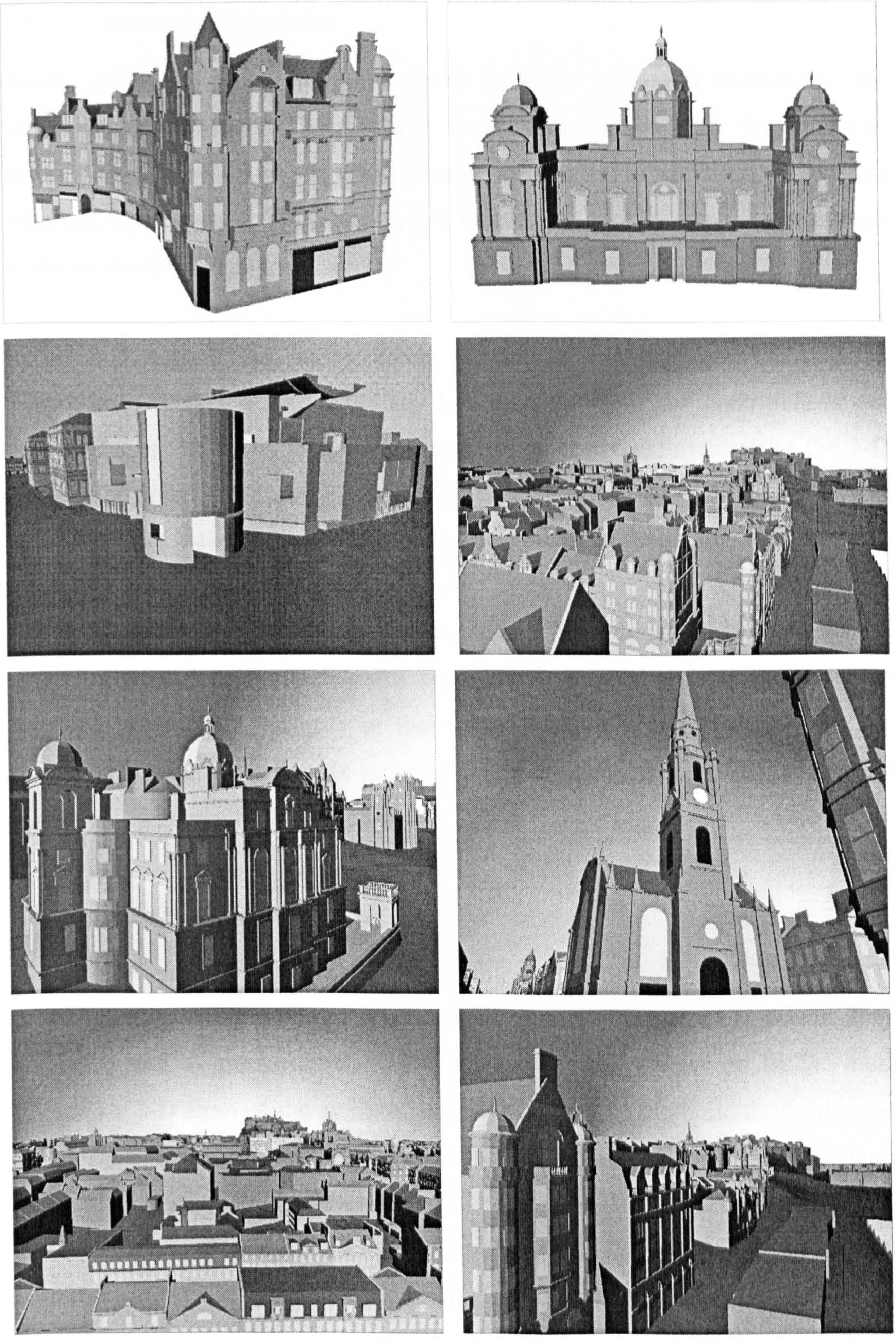


Figure 3.19 Edinburgh landmarks

3.5.1 Transport Networks.

Roads, pathways and general line detail, such as walls and boundaries, are represented as polylines. These entities are pre-processed in much the same way as the building footprints. The final format being a set of polylines, augmented with additional vertices at the terrain cell boundaries, and heighted to correspond to the terrain topography.

In the 2-D domain this representation contains a surprising amount of detail representing as such many of the key Lynchian elements that contribute to our recognition of the urban landscape. However in the transformation to the 3-D domain the potency of these purely linear elements is lost and their representation as 3-D features lacks the ability to define area, volume or materiality.

When the viewed eye point is at altitude, as in a "fly-through" then the generous detail provided by all the features represented as general line detail can be regarded as an asset, quantity in this instance being a ready replacement for quality. However when the eye point is located at street level the lack of veracity in the representation is all too apparent.

3.5.2 Road Construction

If the intention is to provide a realistic representation of the urban landscape then, when tethered to ground level, the viewer is primarily influenced by the more immediate 3-D form of the streetscape and the extent of the building detailing is perhaps of lesser importance. Within the limitations of the 60 degree view frustum common in computer graphics it is evident that when stood in the middle of a street the predominant features with the viewed volume will be street and sky bounded only at the extremes by the building facades. This can make it difficult to compose meaningful views if the vista is not terminated by some other structure, a problem which is compounded by the barren expanse of the roadway.

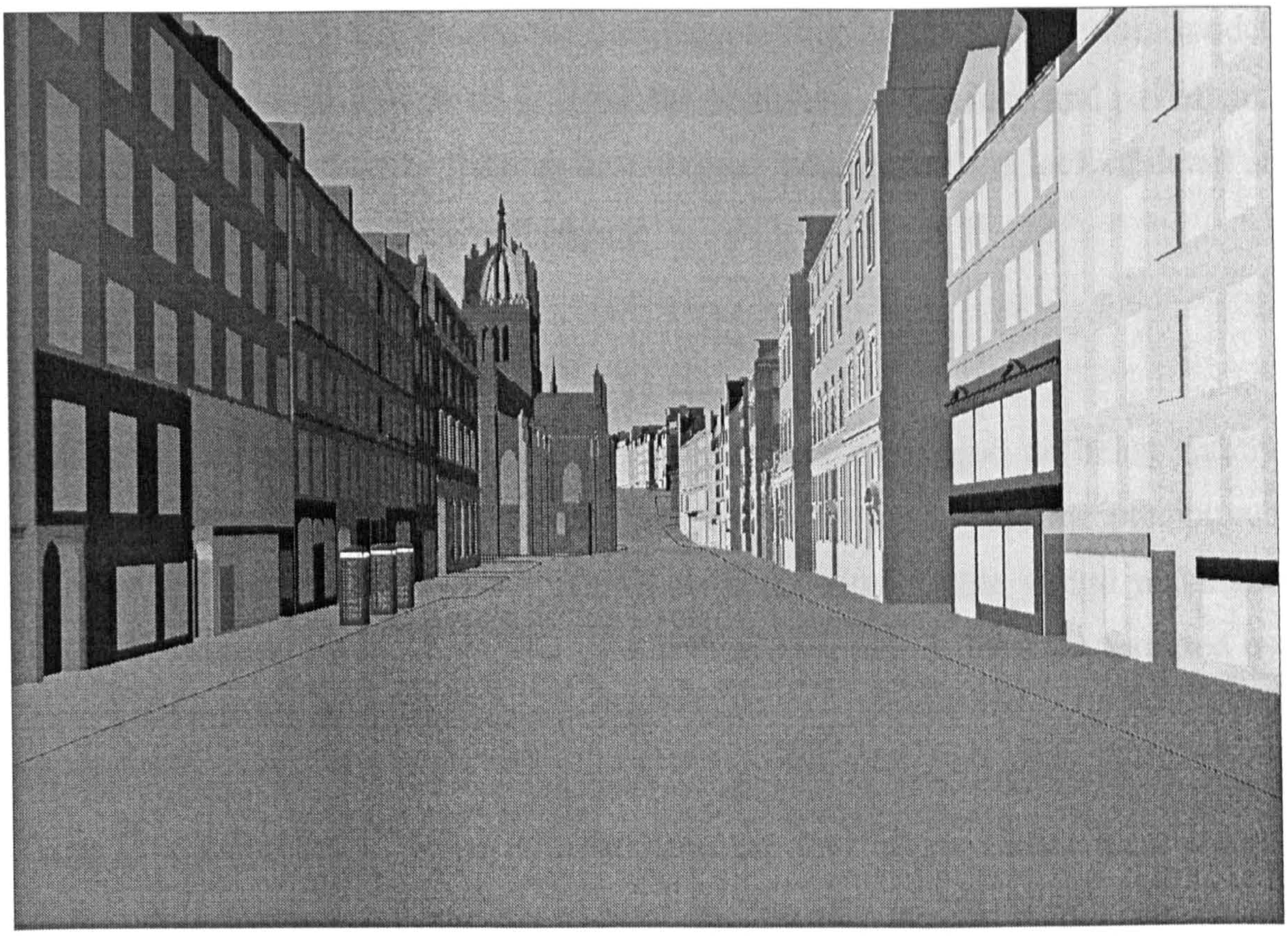


Figure 3.20 The Royal Mile in Edinburgh

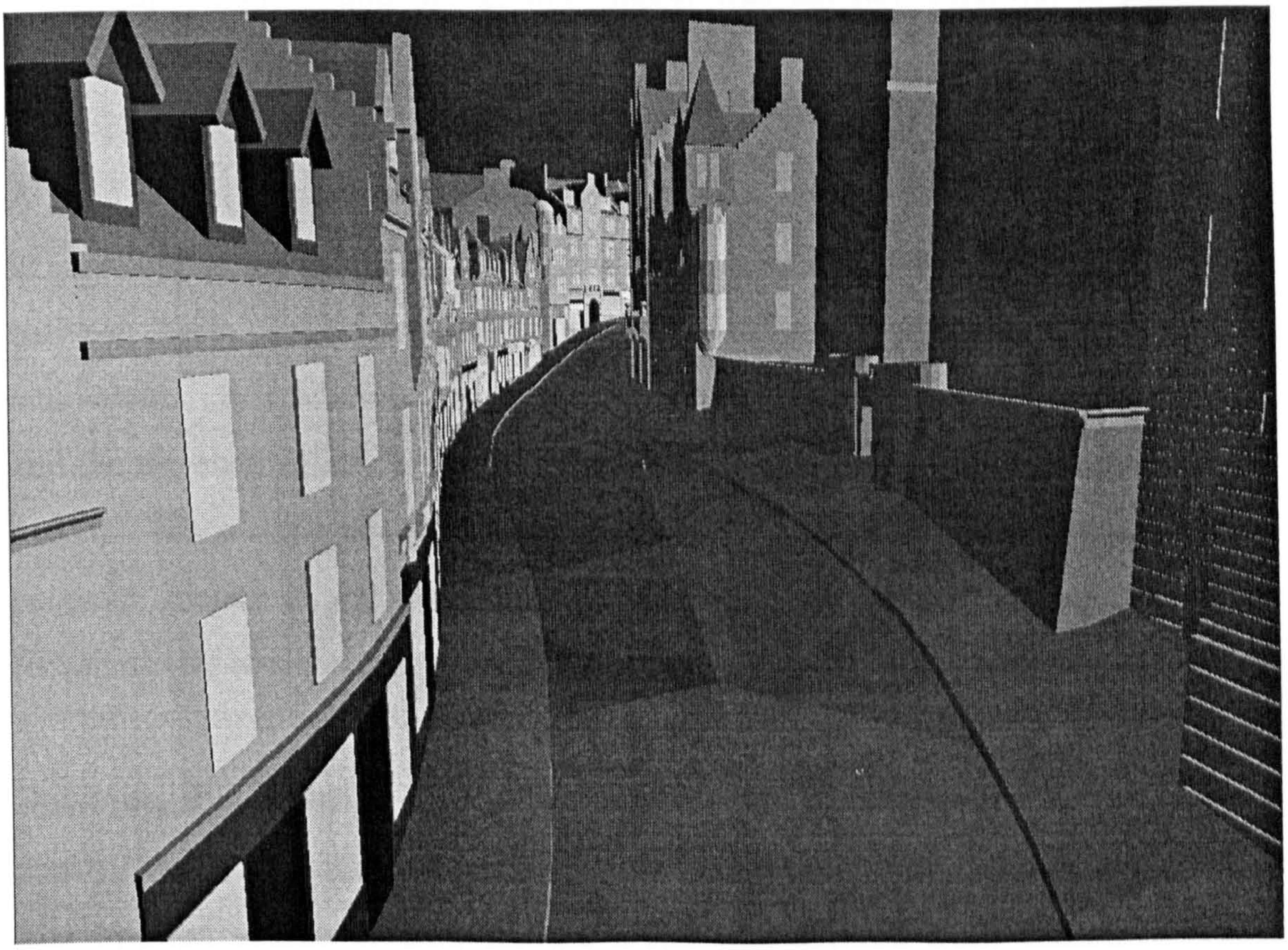


Figure 3.21 Cockburn Street

The image in Figure 3.20 shows the buildings sitting on the basic terrain model with only the general line detail to mark the boundary of the road and pavement. While the scene is accurate from an architectural perspective the lack of detail at ground level diminishes the visual impact

3.5.3 The Terrain Model

Unlike the Glasgow model Edinburgh was always envisaged as being tightly integrated with the terrain. This was due to two factors in that the precipitous cliffs of the Castle rock and the steep tail of the Royal Mile would make the choice of viewing parameters difficult and the perceptual differences incurred by using solid models as opposed to wireline.

There is a difference of some 70m between the foot of the Castle rock and its peak, when viewed in wireline without the benefit of the terrain the castle buildings would appear to float unsupported over the town. The close coupling of the building models and the ground model was deemed essential to provide a sense of continuity between the two.

A solid model was chosen in order to capitalise on new hardware opportunities. In developing the model it was found that there was a fundamental difference between how people perceive a rendered image as opposed to its wireline equivalent. In a wireline model the depiction is that of boundaries not surfaces and the observer is left to decide whether the intervening spaces are occupied or vacant. When surfaces are rendered there is no such dubiety and the model must be more explicit in describing the scene. For these reasons the inclusion of the terrain model was seen as essential.

However, the inclusion of the terrain was not without its problems. In the construction of the model the DTM is used as the datum for all relative elevations. The original source of data for the DTM was derived from contours. These contours were designed to convey height information of landscape not cityscape and fail to take into account the civil engineering interventions that are so characteristic of urban terrain. This has meant that while the generated terrain is broadly accurate of the underlying ground section the local differences due to cut and fill operations are not present. The data required to account for the local, as opposed to global, variations in surface are not readily available but could be collected by a commissioned survey at extra cost. This lack of local detail resulted in a number of discontinuities that became apparent as the model was constructed. For instance there were a number of occasions where buildings, detailed using drawn elevations, would not match the immediate ground profile. This results in the ground surface overlapping features on the building's façade. Similarly, the surface of roads would follow the terrain instead of remaining horizontal in section. This is shown in Figure 3.21 above, where the road surface can be seen to incline across as well as down the hill

A final concern is in regard to the redundancy of data within this representation. While the DTM matrix is a compact format it is still bound by the regular sample interval embodied within the concept. If the area under consideration contains regions that are flat or are of constant gradient then the format is less than optimal. As the sample density increases then this situation is exacerbated and if there is a need to represent the terrain as a series of independent triangles then the worst case condition is soon reached.

The diagrams in Figure 3.22 show the effect of one strategy used to mitigate this situation. Each figure shows the wire line representation to the left of its rendered counterpart. The first figure represents the verbose format with no decimation applied, this is followed by the result of joining any co-planar triangle along their shared edges. The final two diagrams show the effect of varying the joining tolerance. This tolerance is applied by recursively calculating the dihedral angle between the planes of adjacent triangles then if they fall within the correct range merging them into polygons. If the triangles are exactly co-planar there is no loss of accuracy but as the tolerance is increased the margin of error grows proportionally. Table 3.6 shows the effect on the number of faces, edges and vertices within the model

Editing the terrain is difficult as any changes tend to be pervasive and since all buildings rely on the initial DTM for their vertical displacement small variations applied to correct a discrepancy in one area will affect all the surrounding structures.

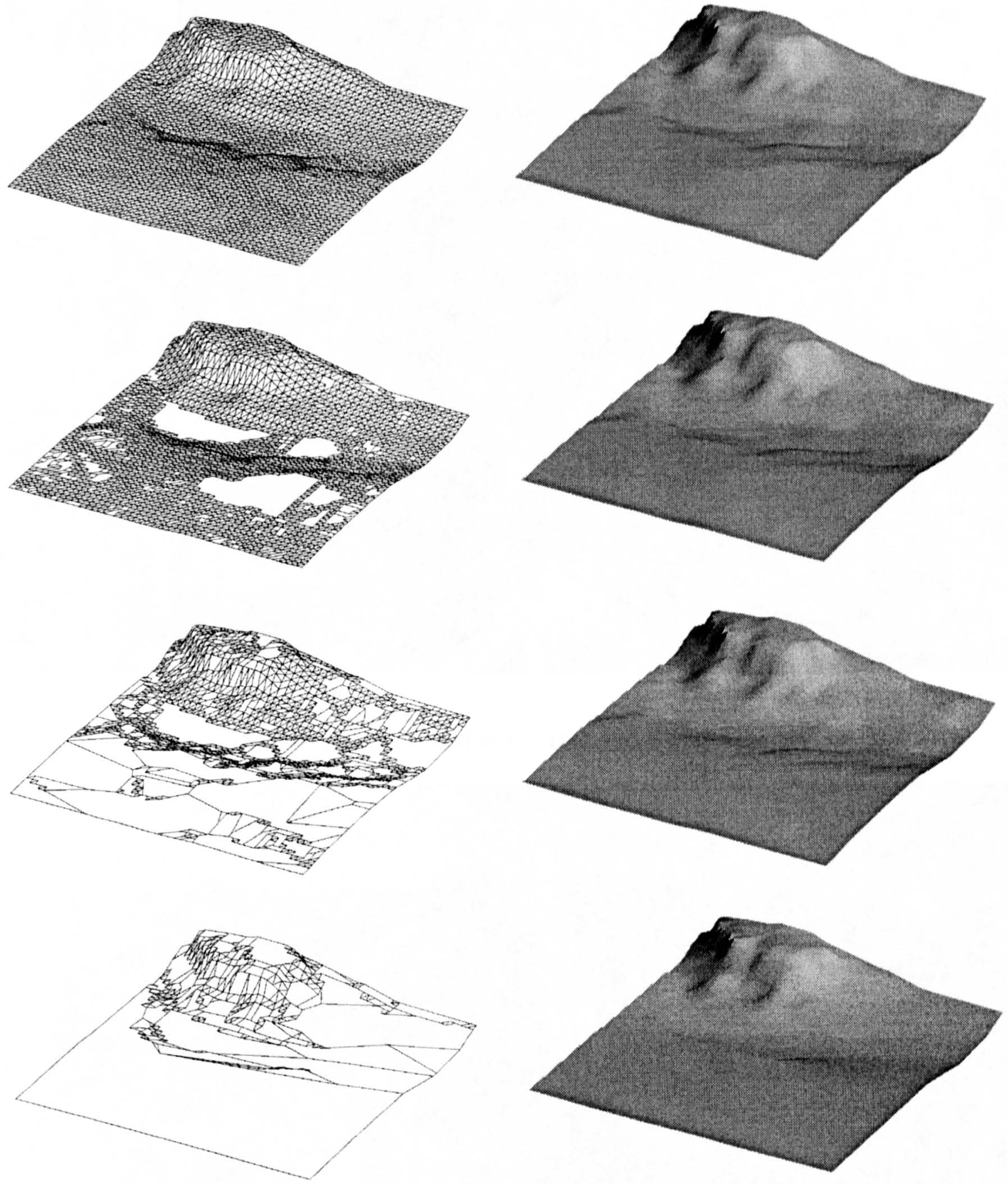


Figure 3.22 Terrain mesh decimation

Verbose format	5000 faces	7600 edges	2601 points
Co planar surfaces	6349 faces	4025 edges	2325 points
1 degree tolerance	1569 faces	2844 edges	1277 points
6 degree tolerance	731 faces	1119 edges	389 points

Table 3.6 Effects of changing tolerances

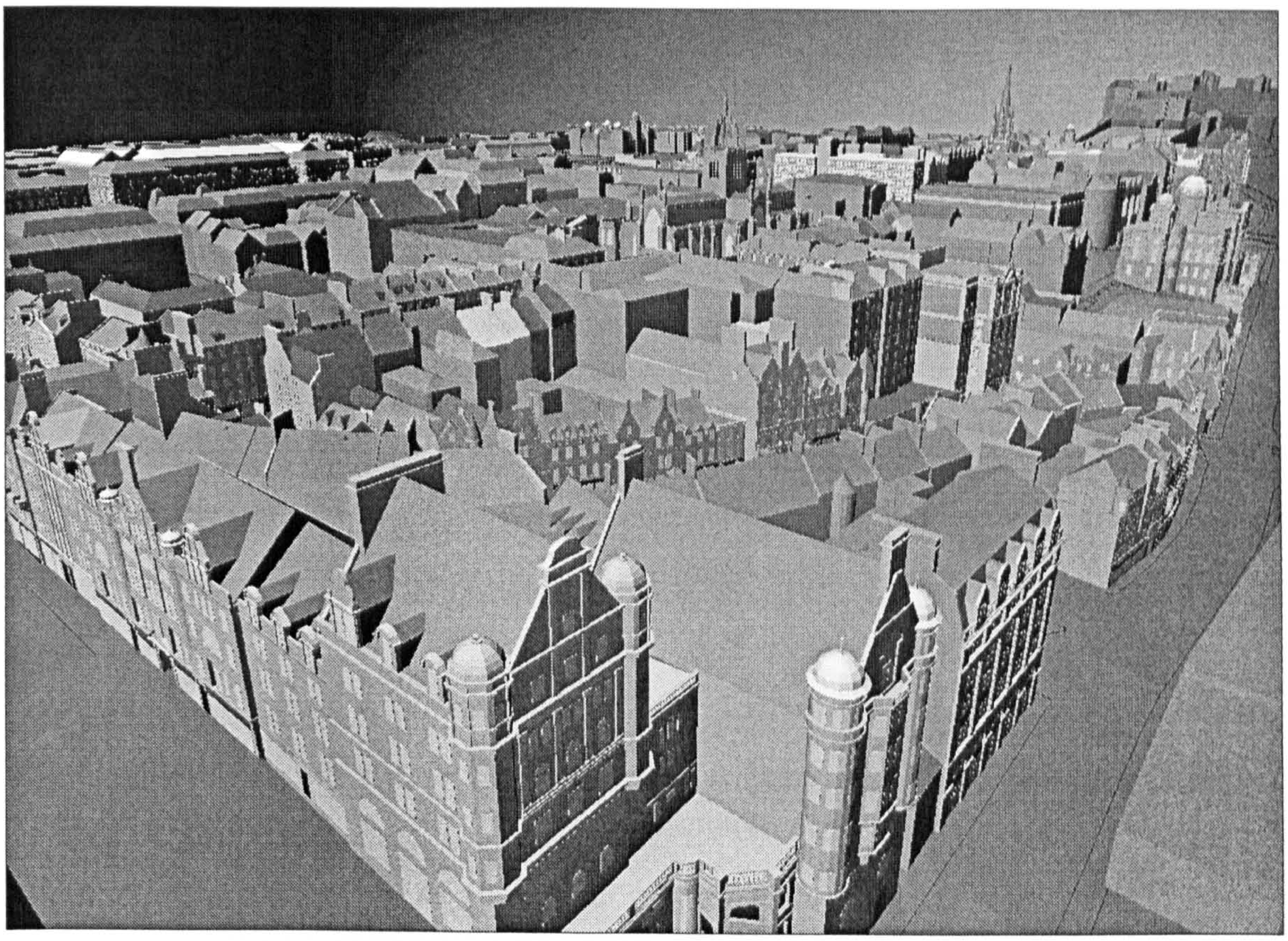


Figure 3.23 From North Bridge with Cockburn Street and Royal Mile beyond

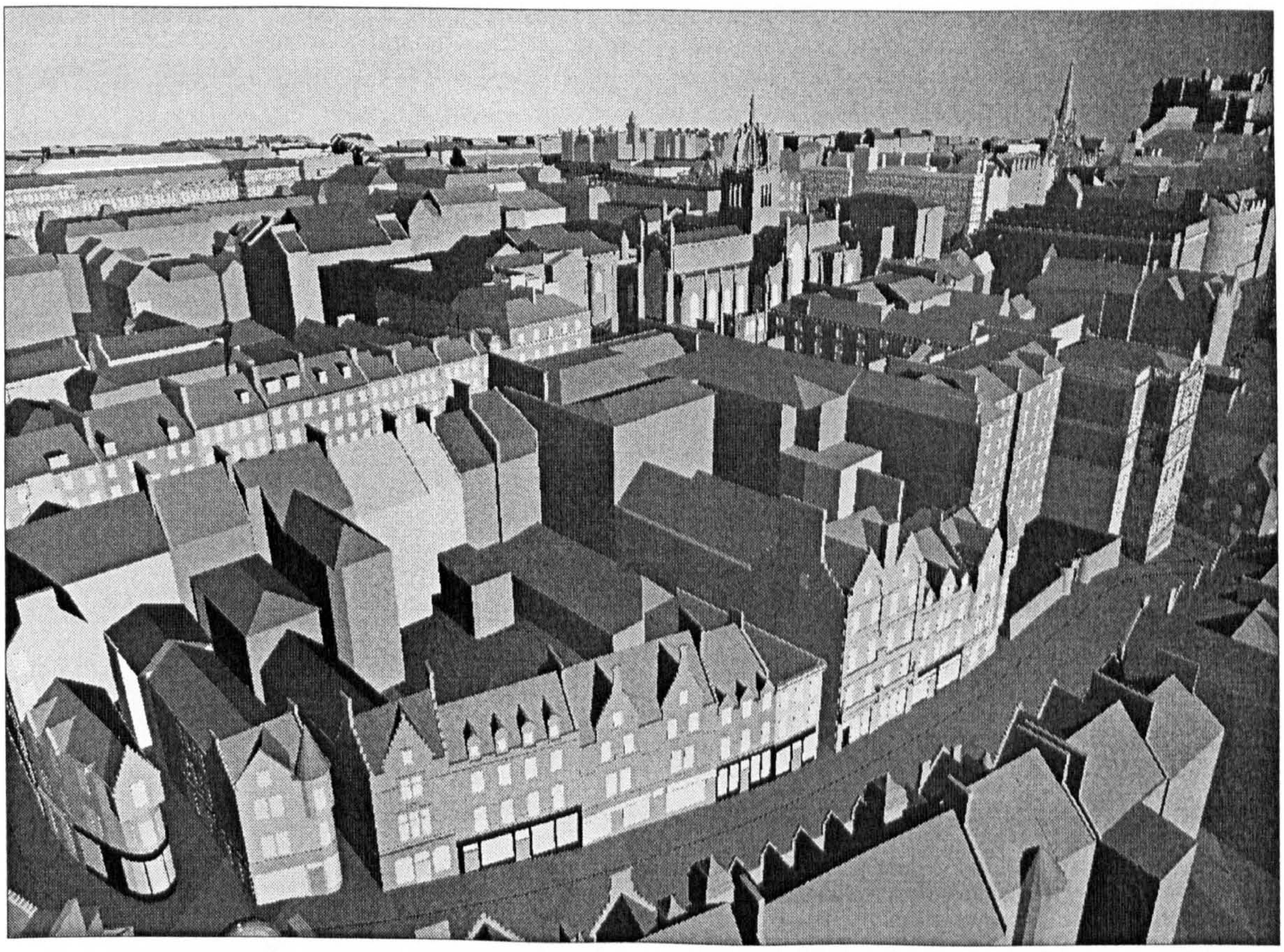


Figure 3.24 From Cockburn Street with Royal Mile beyond

3.6 Edinburgh Old Town

The Old Town originally developed as two separate burghs - that of Edinburgh spreading downhill from the Castle and that of the Canongate climbing uphill from the Abbey of Holyrood. By the early seventeenth century, the Royal Mile had become established as the link between these two districts.

Prior to the development of the New Town, pressure for space within the Old Town's restricted physical boundaries resulted in a dense pattern of tall buildings stretching downhill to the north and south of the Royal Mile and separated only by narrow closes. During the eighteenth and nineteenth centuries, significant improvements were carried out in the Old Town in a manner which ensured the continuance of its original function as both the core of the City and as a definable community in its own right. (EOTRT 1994).

3.6.1 The Edinburgh Old Town Renewal Trust

The Edinburgh Old Town Renewal Trust (EOTRT) was formed in 1991 to promote the environmental, social and economic renewal of the Old Town, It aims to improve the environment and economy of the area by sustaining residential growth and pursuing a creative balance between the incorporation of new developments and the preservation of a unique urban environment.

The Trust received funded for running and management costs from Edinburgh District Council (EDC), Lothian and Edinburgh Enterprise Limited (LEEL) and Lothian Regional Council (LRC). In addition it received a capital budget for conservation grants provided by Historic Scotland and EDC

Recent years had seen major sites within the area become available for redevelopment while other existing buildings were in need of restoration and upgrading. As a result the Old Town faced a period of sustained redevelopment. The overall objective for the future development and improvement was given in the Action Plan for the Old Town, Second Review, 1994 as :

> *"To achieve long term sustainable improvement in the environment and economy of Edinburgh's Old Town, by promoting a productive balance between the interests of residents, businesses and visitors. This balance is to be established through active partnership between the local community and developmental bodies." (EOTRT 1994)*

The Trust had recognised the need for tools which would both assist the process of communicating intended strategy to the local community as well as promoting the work of the Trust over its wide remit. As recognised in the statement from the action plan outlined above, a key task would be to bridge the divide between the local community and the development bodies. The development of a computer model, within which strategy could be described, was seen as one possible method of approaching this aim.

3.6.2 The Edinburgh Interface

One stated aim of the Edinburgh Old Town Modelling project was to construct the model in such a fashion that it would be possible to attribute the geometrical model with other, alpha numeric, properties relating to the building stock. This was proposed in order to address shortcomings discovered to be inherent in the construction of the prior model, that of Glasgow. Previously, attempts to link geometry and information had been thwarted by a mismatch in the granularity of the two sets of data. This was due to the capture of Glasgow's buildings in a block format as opposed to structuring the data at the level of individual buildings. Retrospective action to correct these problems in the Glasgow model were considered to be too close to the effort and manpower required to remodel the whole city and unfortunately any corrective action was never implemented. Since the basis of the Edinburgh Model was the source material for the OS mapping data there would be a one-to-one relationship between the captured geometry and the main source of addressing information. It was not difficult to recognise that the closer the structure of the Edinburgh Model adhered to that of the OS data then the easier it would be to identify and cross reference buildings with their spatial reference.

While it was realised that the resource required to attribute all of the buildings, within even the limited area under consideration, was beyond the scope of the main project it was still possible to view the exercise as a worthwhile investigation into what could be achieved over a larger area. The case for a generalised urban information system is made later in this chapter, at this stage it was recognised that even a limited subset of information linked to the geometrical model would proved significantly useful to the activities of the Old Town Renewal Trust (OTRT). Originally the worth of the model had been seen in its ability to help describe and communicate interventions in the fabric of the Old Town, however if the model also became a repository, and the vehicle, of an information database describing the building stock then its worth would be more than the sum of its parts.

The need to revitalise the Old Town of Edinburgh was recognised in the early 1980's and in 1984 the Old Town Advisory Group (OTAG), the forerunner of the EOTRT, set up a survey steering committee to programme a fabric survey of the Old Town. The survey was carried out within the designated conservation area with the objective of establishing a paper based archive relating to each existing building and, where possible, information pertaining to past buildings on that same site. The goal was to establish a resource that could be readily used to inform decisions on the need for future action. The outcome of the survey was a set of records, differentiated at the building level, containing details of ownership, use, fabric condition and other relevant facts.

EDINBURGH OLD TOWN STUDY FABRIC SURVEY	REF: grm31.fil
ADDRESS	31 - 35 GRASSMARKET
description	Large university building of concrete slab construction. 5 to 7 storeys with flat roof. Ground floor car park with large entrances at either end of the building. 14 - windows across.
listing	
date built	1968
architect / client	George Walls and Partners
GENERAL CONDITION	Good
present use	Educational
vacancy	
owners	Heriot Watt University
TOWNSCAPE	This building dominates the south side of the Grassmarket and disrupts the continuity formed by the prevailing character of tenements and ground floor shops. Totally out of scale with adjoining buildings, and use of concrete, and metal framed windows is out of sympathy with these also. Pend and doorway alienate pedestrians by their sheer size and view onto the rear car park, and also by their failure to define an actual entrance to the building and only lead to spaces behind it.
REFERENCES	Presscuttings: colleges Heriot Watt 1873 - 1964 Vol 4 1964 - 1977 p.175/6

Table 3.7 Old Town Action Committee fabric survey

Urban environments are naturally subject to constant change due to their dynamic nature and any form of collated data can be quickly outdated. The survey data collected by the OTAC had not been comprehensively reviewed since its original capture date and many of the records were now inappropriate or outdated. Despite this, the information provided in the archive was still pertinent to the functions of the EOTRT and as such was considered worthy of use as a proof of concept exercise. At the time of this project commercial database programs were relatively limited, costly and difficult to maintain by the uninitiated while applications which addressed multi-media content just did not exist.

The design and implementation of an interface that would link this survey information with the geometrical model sought to address three objectives.

- The development of an interface that would allow interactive management of geo-referenced data. The principle objective was to investigate the means of storing information held in a variety of formats while allowing unskilled users to search and access the data in a flexible and useful manner.

- Since most information was held in a paper based or non-digital format it was important to gain some understanding of the cost in time, specialist data capture equipment and storage requirements for a multimedia archive.

- If the geometrical model was to be capable of responding to single building queries then the lowest common denominator of the entire geometrical set must relate to the geometry of a single building. Given this scenario an interface must be provided that would allow individual access to single buildings as well as the ability to rapidly define and select city blocks.

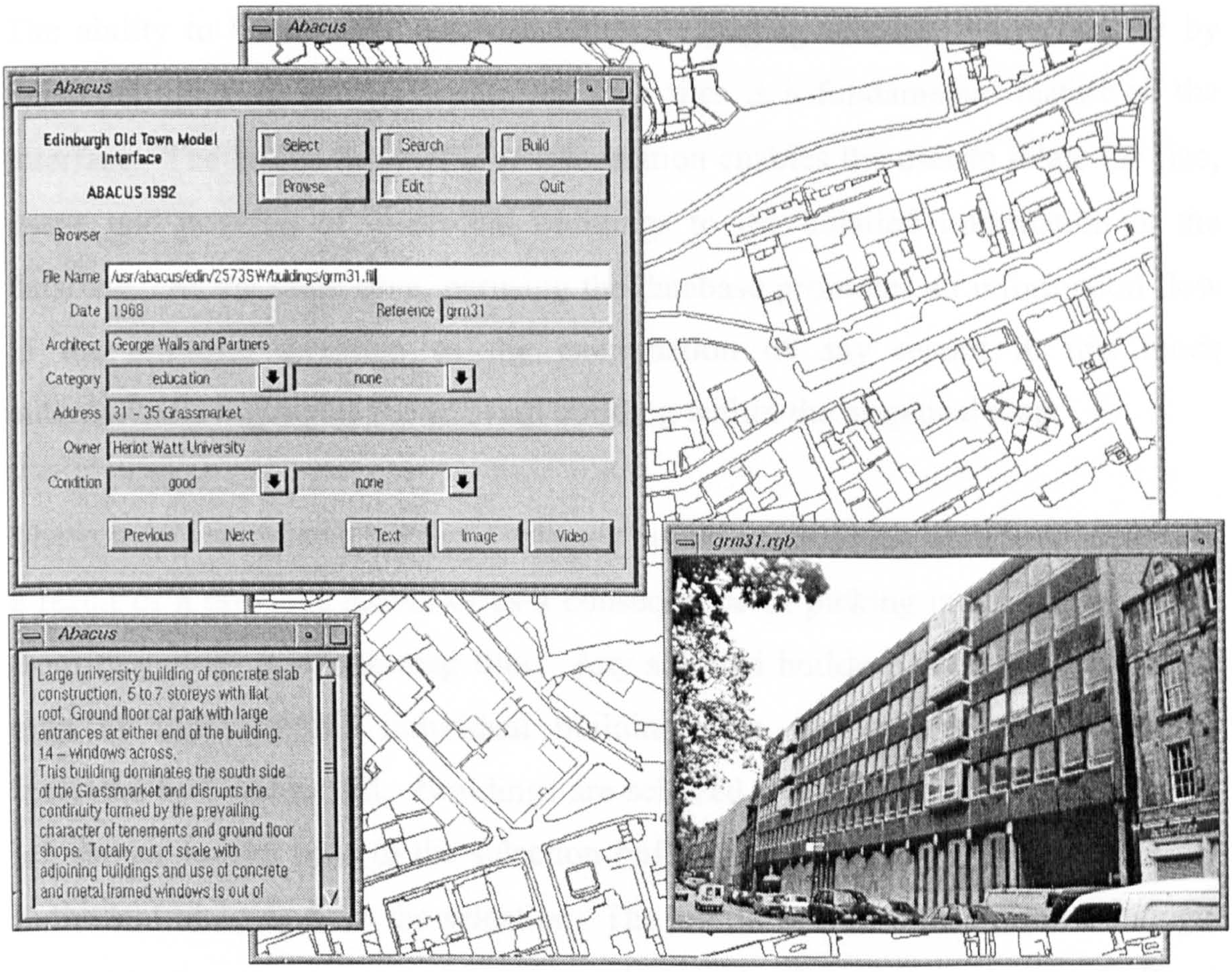
The methodology adopted was to create an interface that allowed the user to move around the database and graphically browse the data using 2-D maps to select buildings and their associated records. Unlike a conventional database where each record might point to a building, this implementation takes each building as a graphical pointer to a record containing further data. Also provided is a query interface that enables a limited "search" functionality. This enables a user to search for a record on the basis of context combined with a selection of Boolean, arithmetical or lexical operators. This will allow for a search pattern of, say, a mixture of use category, ownership, date, condition and location. The added value inherent in this approach is the bi-directional flow of information between the graphical content and the information content.

For ease of implementation as well as access efficiency the entire data set was subdivided into tiles dimensioned 500m by 500m. This mirrored the format of the

1:1250 series of OS urban scale maps. The advantage of being able to restrict queries to discrete areas of the data-set was most obvious when it became necessary to address localised strategies within the greater area. As no commercial database management software was available the decision was made to construct a suite of software tools that would provide the limited functionality required. A set of requirements was drawn up that would allow the users to import, edit and manipulate the data that represented each field in the set of records. This was originally implemented in FORTRAN, latterly in 'C'. The hierarchy of the data base is structured utilising the same file system strategy as was employed for the geometrical data. The spatial layout of the real world geography is mirrored by the naming conventions adopted for the storage of the geometry and data on the hard disk. Therefore the entire model has a root directory under which a number of leaf directories accommodate groups of node directories containing geometry, text, images, data fields and other attributes. Every building is identified by a unique name generated from a combination of its particular house number and a common street name. This convention was adopted for all other data items with differentiation being performed by the addition of a generic suffix. Basic fields for each building record are held in a direct access file.

The interface is comprised of two separate windows, the first contains a graphical menu of the application's functionality, the second displays the map tiles of the area, or areas, currently under consideration. The menu section, depending on which option is chosen, gives access to a further set of palettes allowing interaction with the database

3.6.2.1 Selection



File Name /usr/abacus/edin/2573SW/buildings/grm31.fil

Date 1968

Reference grm31

Architect George Walls and Partners

Category education none

Address 31 - 35 Grassmarket

Owner Heriot Watt University

Condition good none

Large university building of concrete slab construction. 5 to 7 storeys with flat roof. Ground floor car park with large entrances at either end of the building. 14 – windows across.
This building dominates the south side of the Grassmarket and disrupts the continuity formed by the prevailing character of tenements and ground floor shops. Totally out of scale with adjoining buildings and use of concrete and metal framed windows is out of

Figure 3.25 The Edinburgh interface

The selection option presents a choice of the tiles that comprise the area of the basic model. As one or more tiles are selected the application will load a 2-D representation of the chosen area into the graphics window. These 2-D representations of the model are formed from a set of polygons representing each road, building and differentiated area of soft or hard landscape. As each map is loaded these tokens are stored in a lookup table and form the index for the database. The polygons themselves not only depict the model entities that comprise the data set but form pickable primitives allowing selection from the map via a point in polygon test. The purpose of this selection mechanism is to preserve the hierarchy of the database and to provide a means of limiting enquiries to specific geographical regions within the model

3.6.2.2 Browsing

The ability to browse the database either by paging through the records or by selectively picking entities within the map tiles is a fundamental feature of the interface. The bi-direction flow of information enables the user to relate the size, shape and position of individual buildings to the detailed information in the database. At the same time, perusing the database promotes an information flow in the opposite direction as the examination of any record in this mode automatically highlights the relevant building within the graphical map.

The system maintains a buffer of selected buildings, this may have been created as a result of a previous search or as a consequence of picking individual buildings from one or more of the map tiles. Any selected building is highlighted on the map by outlining the individual building plan, the current selection being indicated by a colour fill. Buildings are selected by pointing with the cursor, the left mouse button adds to the selection buffer while the right button deletes that individual building from the selection. The middle mouse button posts a dialogue to enable the user to zero the entire selection buffer. If the selection buffer is current then the user will be constrained to browsing only those records that are indexed within the buffer. If the user needs to browse the database sequentially then the buffer must be reset as previously described. As the user pages through the existing selection, or adds to it, the data base fields relating to the current record are displayed within the interface. The information within the database was limited to that provided by the OTAC survey as shown below. Where more than one entry per field was applicable these were highlighted within the accompanying text.

file name	Relating to the data file within the model
date	Date of construction (where available)
reference	OTAC reference number
architect	The name of the architect or practice
condition	The fabric condition of the building
address	The street name and number
owner	The individual or company in current possession
use	The building use category

Table 3.8 Field structure for database records

Supplementary information was provided in the form of scanned images, video sequences and a textual description. These items were accessed within each record from the interface buttons and presented in separate pop-up windows as shown in Figure 3.25. Once these features were activated they were automatically updated as subsequent records were accessed.

3.6.2.3 Searching

The interface provides a rudimentary search facility. On selection the user is presented with a form filling layout relating to the fields incorporated in each record. The scope of the search operation is limited to the current tile selection which provides a mechanism for either addressing the entire database or the ability to focus on individual sets of tiles. This allows the geographical scope of any query to be limited.

A search may be made by lexical matching on any text field or combination of fields, the user may enter a full or partial string within the query field and the database will return the matching records if that string is found in the corresponding field of any record. All strings are converted to uppercase to avoid the additional complexity of allowing for mixed case data.

Further numerical matching may be performed within a user specified margin on any numerical field. This would allow the identification of buildings dates within a given age range, for instance any constructions within 50 years of 1870. The user supplied tolerance is applied to both sides of the given date so a given tolerance of 10 years provides a 20 year margin.

Certain keyword options are prone to subjective definition, for example the description of building condition or use category. In this instance a pre-defined vocabulary of terms can be accessed via a pull down menu. Since many buildings are multi-occupancy and hence fall under more than one use category multiple criteria can be placed in one entry. If one or more of these criteria is matched then the corresponding record is returned.

Each field can be included or excluded from the search criteria by selecting the activation function to the left of each field. On completing a search records are returned as matched only if all the participating criteria are found. A set of indicators show which criteria were matched during the search, these can be used to refine the search for subsequent queries.

If the search is successful then a series of pointers to all matching records are placed into a selection buffer, unless this list is reset subsequent operations search within this list allowing a hierarchical search option.

3.6.2.4 Edit

The edit function provides a mechanism for both entering and updating record fields. New records can only be added by introducing additional geometry within the 2-D maps, each new addition introduces an additional index that provides a unique pointer into the database file structure. The current record can be chosen by paging back or forwards through the database or specific buildings can be picked from the selected maps. Once the fields have been filled out the data can be written to file.

3.6.2.5 Build

This function refers to the interface's ability to choose between different representations of the geometrical data. On completion of a successful search or as the result of choosing buildings by picking from the maps the selection can be viewed in a variety of ways. Depending on the application that the data had to illustrate not all elements of the model might be relevant. This palette allows the user to choose the level of representation by selecting from a series of options that would toggle the required level of detail. Options include the inclusion or exclusion of terrain, roads and the ability to select the level of detail on the buildings.

3.7 Urban Information

Many cities in the UK, and indeed throughout the developed world, are characterised by the all too familiar symptoms of urban blight caused by insensitive intervention in the environment. The common denominator within this class of problem is the lack of a coordinated, integrated approach to the planning, design and maintenance of our cities. The cycle of development and redevelopment involves input from a diverse range of disciplines relating to architecture, civil engineering, transport engineering and the management of city utilities. This lack of a common, updatable, information base renders access to a global view of the city difficult if not impossible. (Grant 1993).

A possible solution could be sought through the use of a model of the city. Utilising the geometry as a common denominator, the attribution of the physical form of the buildings, transport networks and landscape, with information relevant to the formal and functional evaluation of all aspects of urban management would make for a flexible and powerful tool.

Naturally the introduction of this strand of the information technologies into local authority usage is not a straightforward matter due to the constraints involved. These include funding, UK central government's enthusiasm and willingness to encourage uptake, information availability and quality, localised initiatives and the influence of EC policy. These factors are currently more pressing than technological advancements, although progress in the relevant technology is also vital in order that systems can continue to meet user requirements and expectations.

In an urban context, the wealth of spatially attributable data means that the scope for potential applications is limited only by the imagination. Proprietary systems are already utilised in numerous fields including planning, property management and appraisal, transport planning and routing, policing, health and environmental monitoring, marketing, land use management and in many more related areas. These data sets are often collated independently, either by the bodies concerned or acquired from external sources, notably the applications relate to information handling in both public and private sectors.

UrbanGIS, in this context means a networked system which contains disparate yet standardised data sets linked to what is essentially 3D cartography. This would enable an indefinite number of data sets relating to the same urban environment to be linked and cross referenced by their unique spatial coordinates, giving added value to each piece of data through the spatial manipulations allowed and their interrelationship with other relevant data sets. Each data set could be maintained by the principal user of that data but the important factor would be that, during any particular query, data from separate data bases relating to the same geographical location could also be accessed and merged.

To consider urbanGIS implementation and maintenance on this scale, the incorporation of data relating to all major aspects of urban management is undoubtably a huge undertaking, implying the incorporation of massive amounts of data relating to each of the applications mentioned above and also the centralisation of effort. This would most likely be held by one principal coordinating body. Regional and local authorities, as the major collators, processors and disseminators and users of urban information are the natural accommodators and users of urbanGIS. It is a major objective to illustrate that urbanGIS implementations on this scale carry significant benefits for local government planning and policy making processes and therefore to society as a whole.

3.7.1 Local Government and Urban Management

Although the general economic climate is primarily the responsibility of central government policies, it is evident during periods of economic growth or recession that variations occur in the extent to which regions are affected by swings between either extreme. It can be assumed that the ability of individual local authorities to successfully manage and coordinate the resources under their control will impact on the relative wealth created and hence the opportunities available to inhabitants as producers and consumers. Consequently the social well being and quality of life for the population as a whole, at a localised level, may be related to the performance of local planning and policy. Moreover, this wealth naturally contributes to overall national prosperity and economic growth.

Aside from the complexities of urban management, which naturally place increasing demands on management and planning practices, there are other influences which will affect changes in structure, role and orientation of policies. These changes are themselves instrumental in pushing policy makers to use their information resources to greater effect in an attempt to produce better and speedier judgements.

Pressure from central government to be cost effective and efficient in the ability to produce solutions to infrastructure, management and social problems in urban centres is passed down to local authority level along with an increasing burden to police various legislative changes, notable domestic rates, but also in education, social services and housing. These changes, with accompanying financial constraints are part of a shift in emphasis from service provision towards to some form of "enabling" function, where the principal role is that of management within a framework in which services would be provided from the public sector. In order to support such policy it is obvious that information management will play an increasingly important role and, to an extent, will become the essence of good management (particularly the management of change) and decision making practices.

3.7.2 Applications and Users

The principal objectives of an urbanGIS targeted at local government can be tabulated in order to define the extent of information requirements, internal and external users and desired system capabilities. (after Gault and Peutherer)

Information systems should not be application specific but should be designed to meet corporate goals. The needs of all levels, and parts, of an organisation must be met, identifyable key groups fall into the following catagories:

- Senior management and strategic advisors
- Service managers and function planners
- Service deliverers
- Elected representatives
- The public
- Special interest groups

The folowing have been identified as key issues.

- The system should be able to provide each part of the organisation with the information required, when it is required, but must avoid both information overload and information starvation.

- Information has to be available to be used selectively and used to measure progress towards planning objectives

- The system should be integrated, interactive and networked to allow mobility of data between source and user.

- Such systems should be able to provide information on the entire gamut of spatial units of interest to the organisation.

- The system should be capable of holding or generating, past and current information in order to allow time series analysis. The system should also contain simulation modules to enable trend extrapolation and policy testing.

- The data must be of good, or at least of defined quality.

3.7.3 Spatial Data Concepts

Geographical Information Systems have been a major advance in land management. Their success is in no small part due to the ability to communicate the information via associated mapping packages. Given that the urban model can be considered an extension of conventional cartography, enriched by the addition of the third dimension, then this spatial data set must be seen as an intrinsic part of the system. The urban fabric is the single common denominator that associates all the disparate data that may lie underneath and the ability to spatially reference attributes is a prerequisite. The geometric information affords a number of advantages; firstly since it is a potentially accurate digital representation it can itself be be focus of evaluations, in terms of area, height, volume, orientation etc. Secondly, because it is a "natural language" representation it can be easily used as both an interface to the layered information beneath and, at the other end of the process, a vehicle for conveying more abstract information.

Local administrations have the most to offer and the most to gain by moving towards an integrated philosophy regarding information collection, collation and dissemination. The motivation to move towards this goal comes from a number of sources, primarily through the increasing complexity of urban management but also due to central government policy to progress towards the decentralisation of services. Fiscal pressure to increase efficiency, lower manpower resources and arrive at speedier judgements all point towards an increasing reliance on the information technologies.

The quality of planning and decision making processes can be substantially improved when valid data is appropriately and efficiently handled.

Relevant information systems which support the activities of planning and decision making must be based on a thorough and clear analysis of the planning processes adopted. The use of an urbanGIS allows information to be assembled and applied in new ways. It offers practical means to manage large and diverse spatial databases and provides effective tools to understand relationships between disparate phenomena. An increasing number of decision makers and managers have recognised that these technologies will be essential if they are to address the expanding mandates and complex decisions they now face.

The use of a graphically orientated urbanGIS system has a threefold advantage. Firstly, since 80% of urban information can be spatially referenced it is a logical format under which disparate data sets can be collated. Secondly, communication is of paramount importance. The ability to use a readily understandable and recognisable medium with which to communicate proposed strategies and outcomes will make the task of disseminating policy open to all. Finally the use of a digitally accurate urban model will open up greater roles for simulation based tools.

An increasing knowledgeable and aware user base will be a significant factor in the uptake of urbanGIS. Users will demand systems that accept data in diverse formats from existing databases, that they be easier to use and present higher levels of performance. Market forces will provide the incentive for the incorporation of advanced graphics, simulation and animation facilities, faster processors, more competent data exchange and query functions and expert system shells to reduce areas of complexity identified by the users.

3.7.4 Utilising Geometrical Data

Currently the system does not provide for the evaluation of existing or proposed changes to the urban fabric or public spaces. To use the system as an instrument for evaluating the qualities and quantities of the building stock within the context of any city and as a support mechanism for resource management, it is then necessary to attribute further values to the data. To utilise the existing geometrical data, the base level of detail can be used to allow the calculations of floor areas, volumes, densities and distances. As the underlying model is in an accurate, digital format it make sense to utilise this fact to gather data that would otherwise only be available from costly onsite surveys. Increasing the capacity of the system to perform more complex arithmetical calculations would enable the user to obtain the following evaluations of the urban form and structure.

Given this ability it would then become plausible to utilise the model in this new mode and allow the evaluation of, say, built density as a function of other parameters. Similarly the results of a design intervention could quickly be assessed by evaluating existing and/or potential floor areas to be used for new or redevelopment. The potential extends beyond the building stock since the ability to differentiate, and calculate, the fractions of hard and soft landscape as well as that proportion given over to transport networks. This will allow the evaluation of the road network structure, including parking areas, individual and accumulative distances between specific points in the city and even the ability to evaluate the extent of infrastructure given to road networks and the overall length of systems.

The categorisation of building usage e.g. housing, commercial, education, retail, carparking, vehicle routes would enable complex evaluations of urban development projects on varying levels of scales: single buildings, whole blocks or within the overall context of the city. Such evaluations could review: the mixture of categories of use within a specified area showing the distribution of different building categories within an urban district. Similarly the same technique could be applied to the transportation densities in an area.

Specifically the size of distribution networks and distances from supply points of goods to receiver points could be calculated, along with the individual, average or accumulative distances between places of residence and areas of work, retail and recreational areas and social and cultural points. This type of evaluation could immediately pinpoint potential problems in a proposed development such as the ability of the transport networks to meet expected demand, the availability of critical services (fire, police) within a designated area and even the trivial, such as the relative proportion of car parking facilities to expected residents. The location and frequency of specific services such as public transport, police stations, fire stations and schools could be defined for a specific density of development with a particular use profile. This could be aided by applying simple algorithmic planning rules.

Linking cost information to the building use categories would provide a means of evaluating the economic feasibility of proposed developments which in turn could provide an estimation of project costs and investment modelling. Similarly, the viability of transport networks can be evaluated by attributing cost information to different modes of transport. As a result it would be possible to assess: the accumulative distances between places of residence, areas of work and other targeted locations. Given this ability it would then be possible to test proposals against the optimum capacity of the system allied with the existing population density as a function of the desired frequency of the system and the cost per mile.

3.8 Summary

The arrival of a new generation of engineering workstations served to mark a clear distinction between the era of the previous generation's static wireframe view of the world and a dynamic colourful experience of the future. Colour bitmapped displays and geometry acceleration promised a vibrant and dynamic mode of interaction with the city but performance was always in short supply and expectations would still be curtailed by the limitations of an affordable system.

The single most expensive aspect of building the Glasgow model had been the task of collecting geometrical data to describe the city. The introduction of computing to all aspects of information provision brought the opportunity to transfer data, in a digital format, from a cartographic domain into a new model. While this promised greater accuracy and efficiency the lack of a fundamental structure to the imported data brought additional problems and the process was not as expedient as first hoped. The undeniably useful ability to generate large tracts of urban landscape at a contextual level with little human intervention was compromised by the basic limitations of a cartographic format when interventions at the level of individual buildings were made. The requirement for greater realism and detail moved the burden of data capture from the broad brush of the urban context to the domain of individual structures. The ability to model in greater detail and the expectations that this raised in terms of the overall visual quality of the model proved limiting in both the scope of the project and the ease of interacting with the geometry.

The goal of attributing the geometry with urban information was still as elusive as in the previous era. The limitations of the hardware platform restricted the view of the city to the level of individual buildings and not the global extent of the entire model. This served to concentrate access to information at the level of the interface and relegated the buildings to just another attribute of the city rather than as the key component. The availability of information relating to the city was noticeably more restricted not just in legal terms of copyright and the data protection act but also with potential supplier's recognition of the worth of data sources. This lead to some reluctance to share assets as it was perceived as diminishing individual departments control over their own information currency. The concept of gathering diverse data types within a geometrical model of the city was still valid but the fundamental problems were shifting from technical issues of hardware and software towards a more political arena.

Chapter 4 MultiMedia Technology

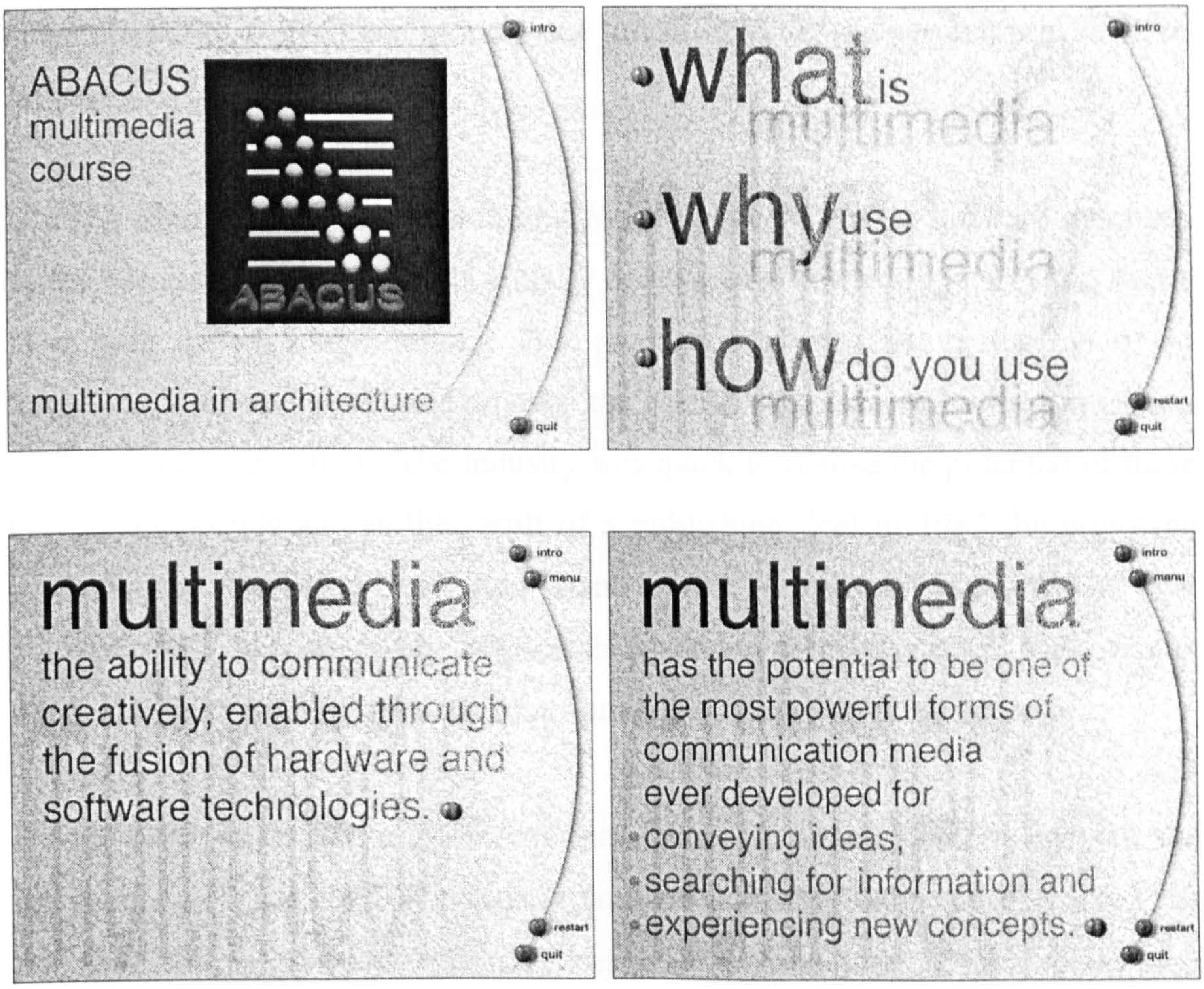
At the close of the previous section it could be seen that the evolution of the Edinburgh Model and an attendant desire to develop a close coupled system of graphics and associated information content could potentially yield a powerful and genuinely useful tool with wide ranging applications. However it was equally obvious that there were a number of impediments to the realisation of this ambition.

The worth of the concept of an Urban Information System was embodied in the conjunction of the geometrical description, its allied information content and the ability to interact with, and manipulate, the data. The disadvantages were mainly found in the particular hardware and software platform that only provided a limited development solution. This wasn't to say that the potential was not there but that the lack of any high level interface to the required functionality meant that the technical skills and attendant time scale required for the development of authoring and presentation tools was a negative factor. Equally, it had been proven that the most immediate audience for such a system would be mainly comprised of essentially non technical users and that the high cost graphical workstation with it's complex UNIX™ operating system would prove unsuitable.

It was fortunate that at this stage of development there was a parallel evolution in the capabilities of ubiquitous, low cost machines at the bottom end of the computational hardware scale. This step formed the basis of a new generation of IT products that encompass all aspects of diverse media types and hence fall into the generic category of Multimedia. This was a major advance in that for the first time developers had access to a platform that was essentially "good at everything" and that would encourage not just progress but also diversity.

This chapter is illustrated with a number of screen shots from a range of multimedia applications realised within ABACUS since 1993. Some of the applications were commissioned but most images are extracted from a suite of

CEC funded international CPD courses directed at developing authoring skills for both practicing architects and lecturing staff in undergraduate education.



ABACUS
multimedia
course

multimedia in architecture

what is multimedia

why use multimedia

how do you use multimedia

multimedia
the ability to communicate creatively, enabled through the fusion of hardware and software technologies.

multimedia
has the potential to be one of the most powerful forms of communication media ever developed for
- conveying ideas,
- searching for information and
- experiencing new concepts.

Figure 4.1 Multimedia courseware, using multimedia to teach multimedia

4.1 The Origins of the Product

Books were cheap, durable and accessible to anyone that could read, music and video were both widely available in industry standard formats, yet in the early to mid 1990s the potential niche for multimedia products was far from certain. Not only did the technical challenges of the proposed media stress even the state of the art in consumer computer hardware but a wide range of competing platforms vied to become the standard. This dilemma gave rise to a vicious circle within which developers were reluctant to target a delivery system with little installed base and end users were equally reluctant to invest in hardware without a good choice of desirable software. Paradoxically this provided a huge opportunity for Apple and the PC clone manufactures who energetically promoted their products signalling an extension of their target audience out from the office and education and into the home. An analysis of this frantic period of development could be hard to constrain but it is profitable to trace one thread of development that proved to be lasting.

In 1984 Marc Cantor, Jay Fenton and Mark Pierce formed a software company called Macromind. Macromind initially developed a product for the 128k Apple Macintosh called SoundVision. This program was the first realisation of an integrated multimedia tool and allowed the user to sequence audio in conjunction with real time animation. The industry was quick to realise the potential of these new developments and as the result of a publishing deal in 1985 the prototype product was separated into two stand alone programs, MusicWorks and VideoWorks. In 1986 every Macintosh produced shipped with a VideoWorks animation that provided a guided tour illustrating the machine's features.

During 1987 VideoWorks Interactive was launched, a development that saw the integration of Tiny BASIC, a tightly coupled scripting language that freed the developer from the purely linear sequence of a time line. This product was subsequently used in authoring a wide range of applications ranging from training material through marketing presentations to television commercials. This rapid take up was the result of a product which was then unique in it's ability to both quickly and cheaply develop and display interactive, integrated multimedia productions finding a niche in media markets which were previously the preserve of comparatively heavily capitalised and skilled technology like audio and video editing.

In the intervening period VideoWorks has evolved into MacroMind Director (latterly MacroMedia Director) the de facto industry standard for multimedia authoring and the platform on which the rest of the discussion within this chapter is based. (Macromedia 2001) In terms of hardware, despite the wishful predictions of market surveys, the wide range of platforms – CDTV, 3DO, Sega

MegaCD, Atari Jaguar, VIS and Cdi – have all failed to retain their market and only the Mackintosh and PC formats have survived.

4.2 Multimedia Foundations

The word "mutlimedia" is used in a variety of contexts, often to describe a wide range of products. It is almost impossible to provide any final and unambiguous definition of the concept. Perhaps the best way to approach an understanding of this concept is to look at how over the past decade there has been an obvious and exponential growth of that fusion of hardware and software products that collectively comprise the technology that has come to be known as multimedia. Taken individually these products are not exceptional, however what can be deemed revolutionary is the way that, when they are applied in some conjunction, this combination can open up new horizons.

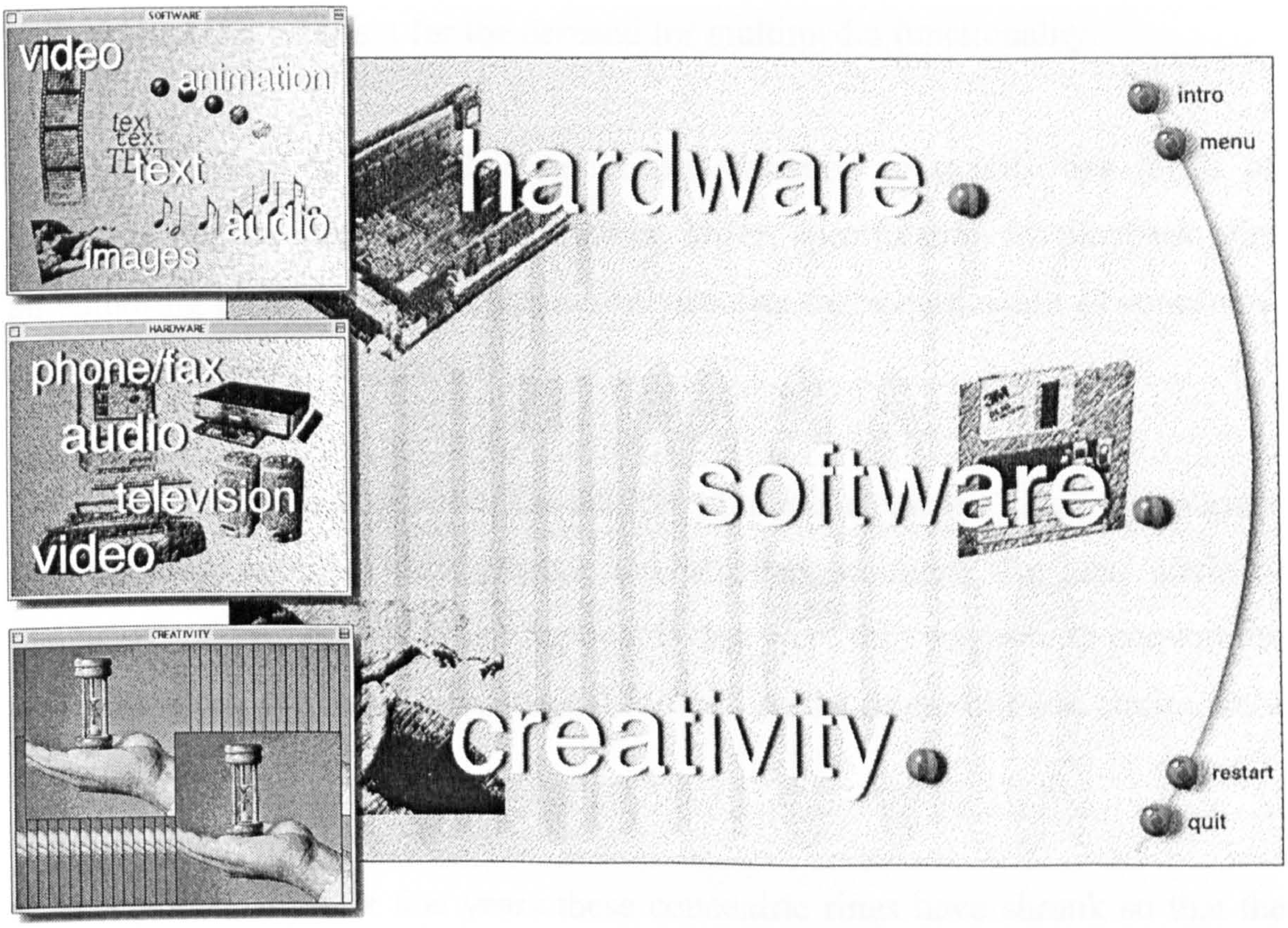


Figure 4.2 The fusion of hardware, software and the human element of design.

Approached at a simplistic level multimedia can be considered as being compartmentalised into three categories, hardware, software and the human element of design.

4.2.1 Hardware Developments

The very nature of multimedia implies that the hardware should be capable of processing, storing and performing input/output on a wide range of media. It is a moot point as to whether it was the advent of multimedia itself that promoted the adoption of all the diverse hardware functionality that is now integrated into a computer or whether the reverse is true and it has been the growth in hardware that laid the foundations for multimedia.

Whichever view point might be adopted it is readily apparent that market forces have driven the capabilities of consumer level computing. These forces would not have existed if it were not for the demand for multimedia functionality.

In the early days of the technology it was common to specify two levels of hardware, one for authoring and the other, lower, specification for playback. On an authoring platform it was common to visualise the specification as concentric rings of technology. Figure 4.3

The outer circle is comprised of the sources of media elements, video tape players and cameras, compact disc players and assorted scanners, the next circle is comprised of the range of digitising boards that were then required to convert the analogue audio and video into a digital format. At the centre of these circles sits a conventional PC.

Over a period of just a few years these concentric rings have shrunk so that the average specification of a consumer level computer will now incorporate most, if not all, of these features. The purpose of this illustration is to demonstrate that the ramifications of media integration go far beyond just the software product and have in fact generated fundamental changes in basic computer capabilities.

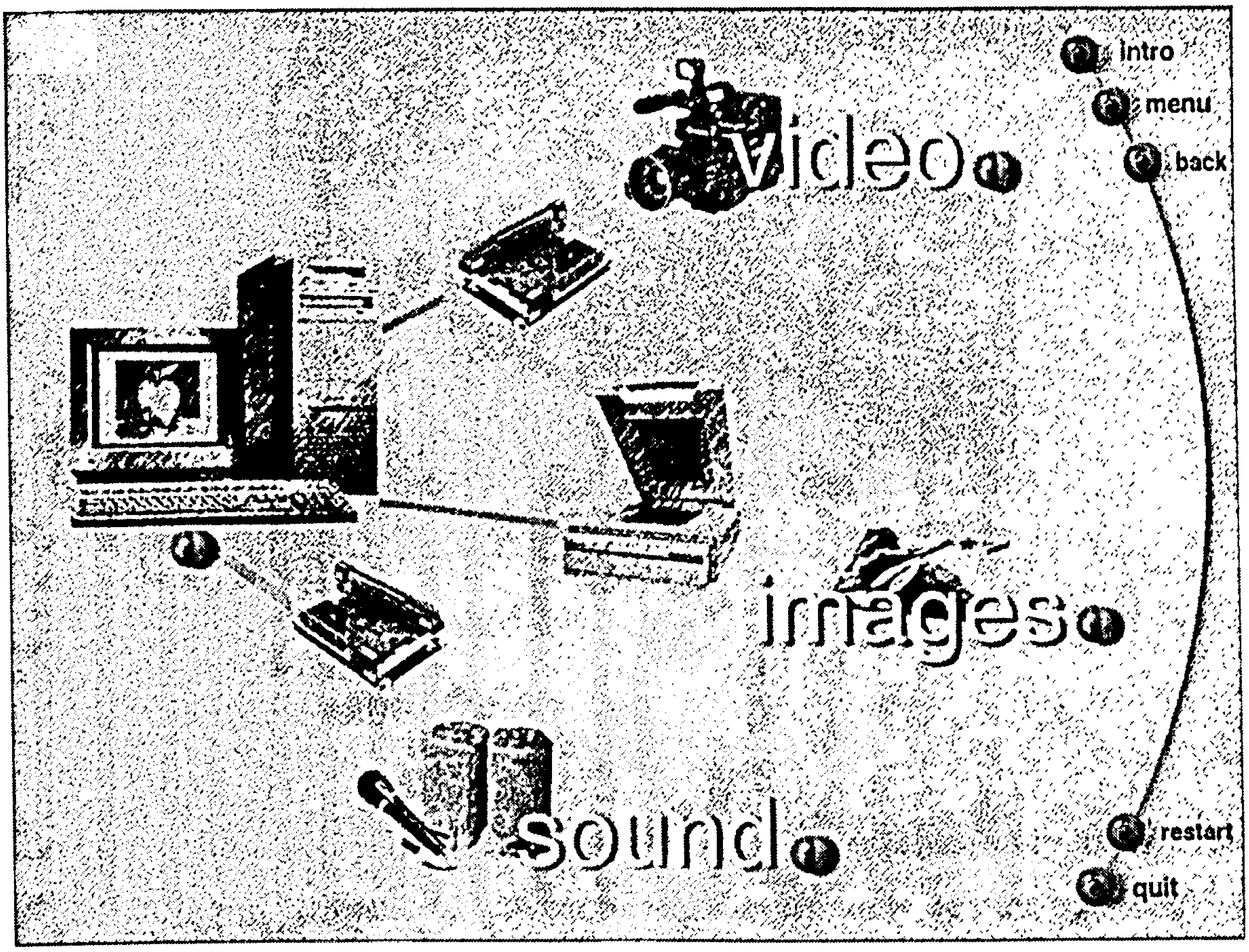


Figure 4.3 Concentric rings of technology.

The capabilities of a modern machine go far beyond mere data processing and it is not uncommon to find integrated hardware addressing functionality such as phone and fax services, hi-fidelity audio, television and radio reception, video processing and a range of other features that were once the preserve of separate and disunited pieces of equipment.

4.2.2 Software Developments

The hardware may provide the platform on which to integrate different media, but there is still a need for software to capture, create or edit the range of media types. Due to the open ended potential of multimedia development this in itself creates a dilemma. As illustrated in Figure 4.4 there are essentially three skill sets required by the multimedia developer.

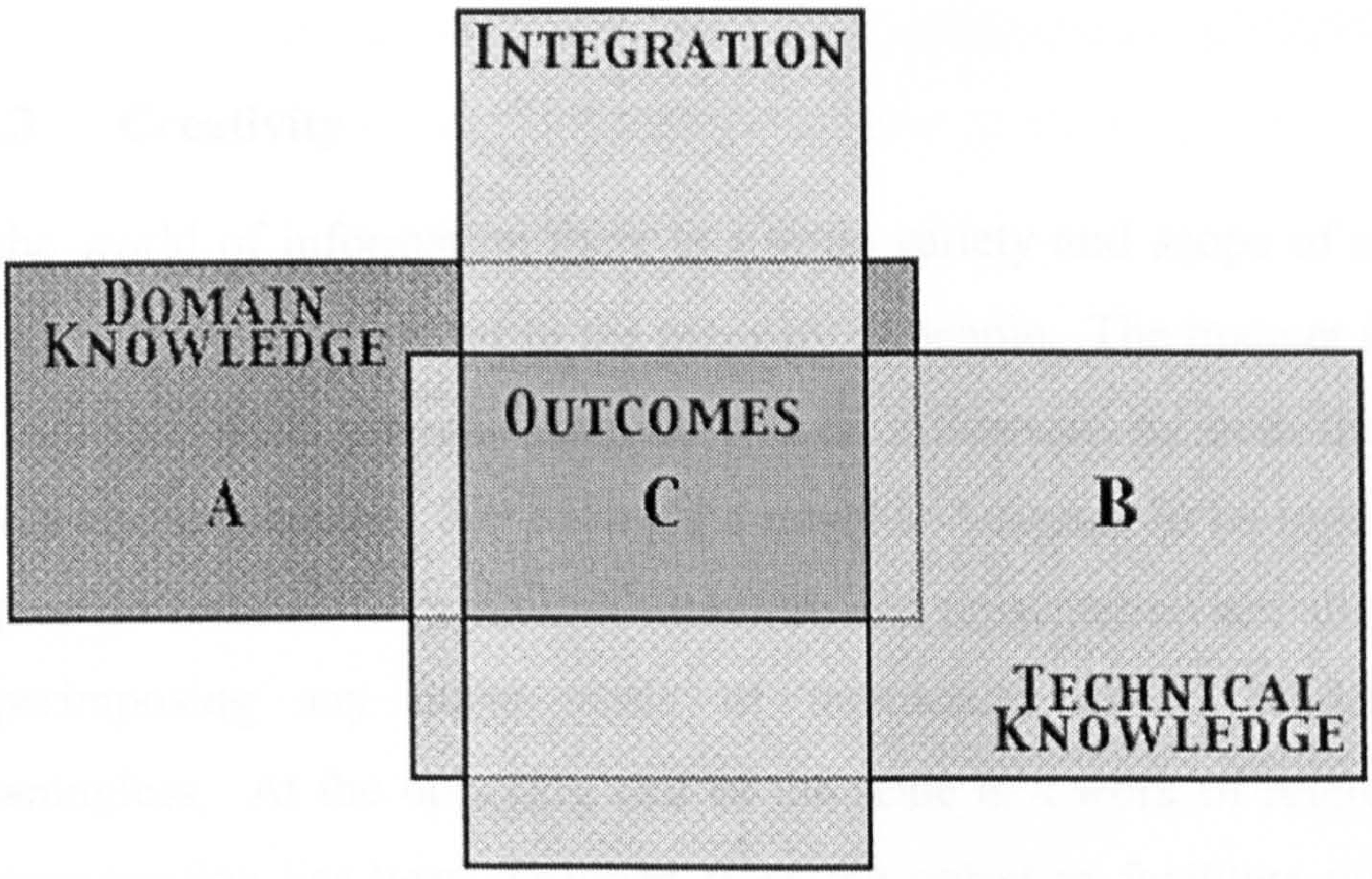


Figure 4.4 Multimedia domain integration

The required outcome (C) is a marriage of domain knowledge (A) in terms of an understanding of the topic of interest which then can only be communicated by means of technical knowledge (B) which is the set of skills necessary to design the structure of the application, apply software tools to transform media elements and program the integration of narrative, user interaction and content to produce the end result. (Anderson 1994).

This technical skill set is represented by the ability to work with a spread of applications that enable the capture, edit and transformation of all the required media types. This implies that the author, whether an individual or a team, must not only know how to control the hardware and operate the software but also must possess an understanding of the media itself. For example in the case of video footage knowledge of that individual domain involves camerawork, lighting and perhaps sound recording, then when the material is digitised further expertise is required to judge frame rates, colour depth and data storage requirements. The production of a successful multimedia application must appreciate the fit both between the higher level domains and between the sub groupings within each domain.

4.2.3 Creativity

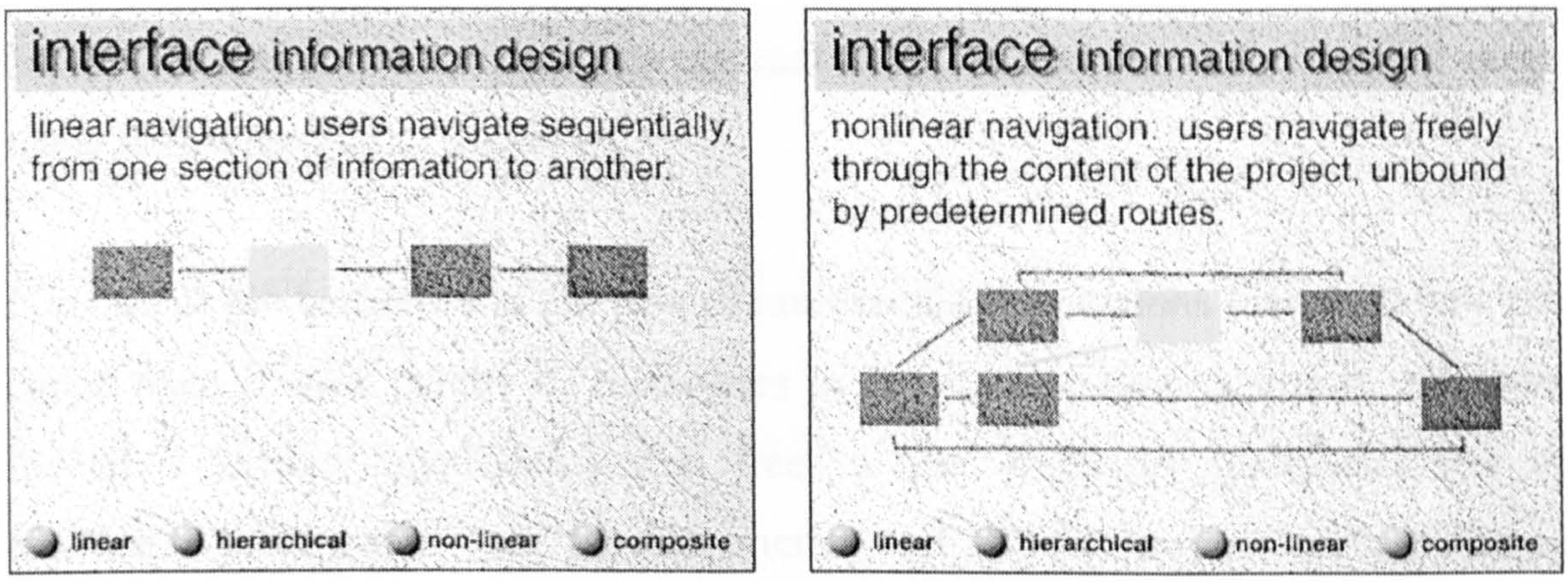
In the world of information there is a wide variety and scope of media, most of which will be very familiar to the majority of people. The manner in which a user will interact with any one example will be influenced by both the nature of the media and its content. For example a novel is designed to be ingested in a linear sequential fashion, the narrative, structure and presentation are all predetermined. Superimposing any other mode of interaction would render the content meaningless. At the opposing end of the scale is a work of reference where the systematisation has been designed from the outset to facilitate piecewise access and cross-referencing.

The restrictions of a traditional paper based media fail to provide a vehicle for any form of dynamic content, the ability to search on arbitrary keys and to offer any form of interactive conjunction with the topic other than through the user's imagination.

The core strength of multimedia is that it can answer each of these shortcomings and so provide the potential to become one of the most powerful forms of communication media ever developed. Multimedia is characterised by combining a number of expressive forms such as text and numerals, graphics, images, animation, sound and video. This is what serves to provide such a rich medium with which to convey ideas, but if a user is to experience the "message" as a coherent whole then the design and implementation of the structure and principles of navigation must be carefully considered. There is no one technological solution that can address the structure of a multimedia application, a fact which perhaps is responsible for the wide and varied nature of the genre.

4.3 Navigation

The structure of a multimedia programme serves as a framework within which the media is sequenced. The challenge is to define a structure that will both offer a wide range of choices and at the same time appear logical and easily understood. The coherence of the structure may be express in a number of ways but is most likely to fall under one of four categories.



interface information design

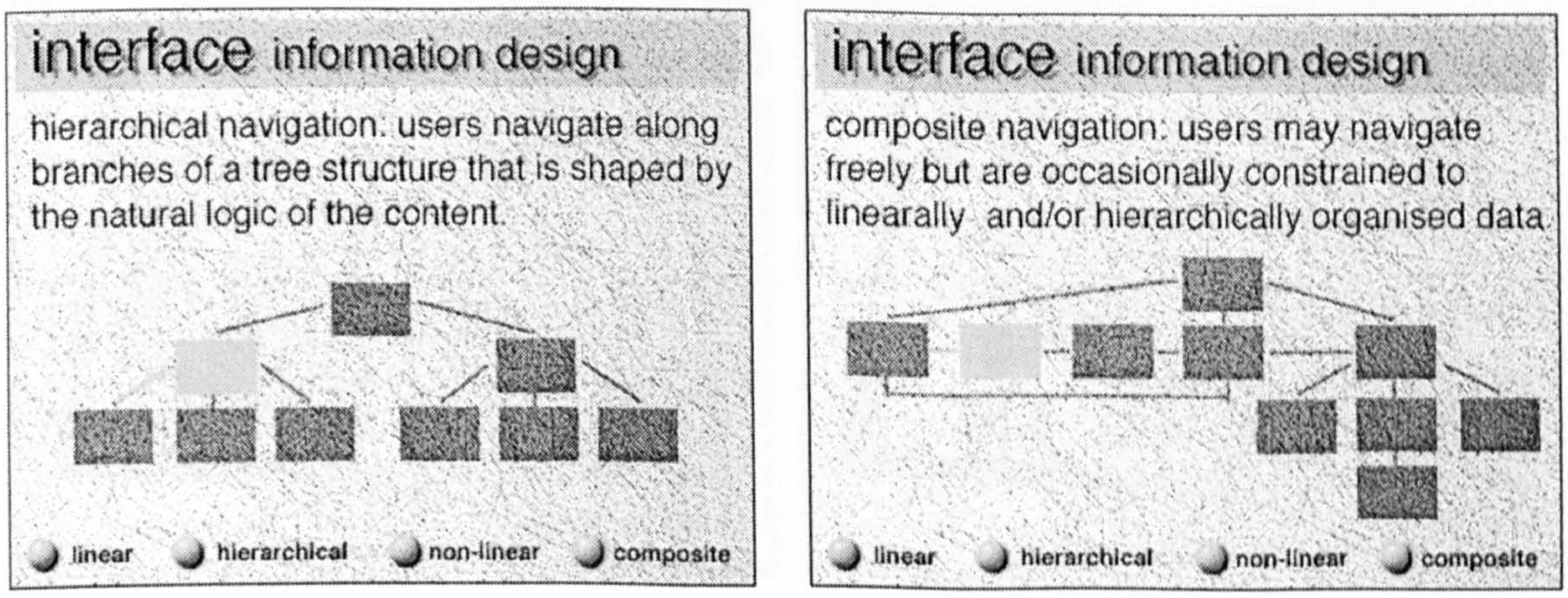
linear navigation: users navigate sequentially, from one section of infomation to another.

interface information design

nonlinear navigation: users navigate freely through the content of the project, unbound by predetermined routes.

Figure 4.5 Navigation strategies, linear and nonlinear navigation.

While the structure might define the layout of the media and to a certain extent predetermine the ease of access to information it is the formalisation and construction of the interface that offers the most scope to the designer. If multimedia is to offer more than the functionality of a database or the slick imagery of a catalogue then it will be through the provision of an experiential interface between user and data.



interface information design

hierarchical navigation: users navigate along branches of a tree structure that is shaped by the natural logic of the content.

interface information design

composite navigation: users may navigate freely but are occasionally constrained to linearally and/or hierarchically organised data.

Figure 4.6 Navigation strategies, hierarchical and composite navigation.

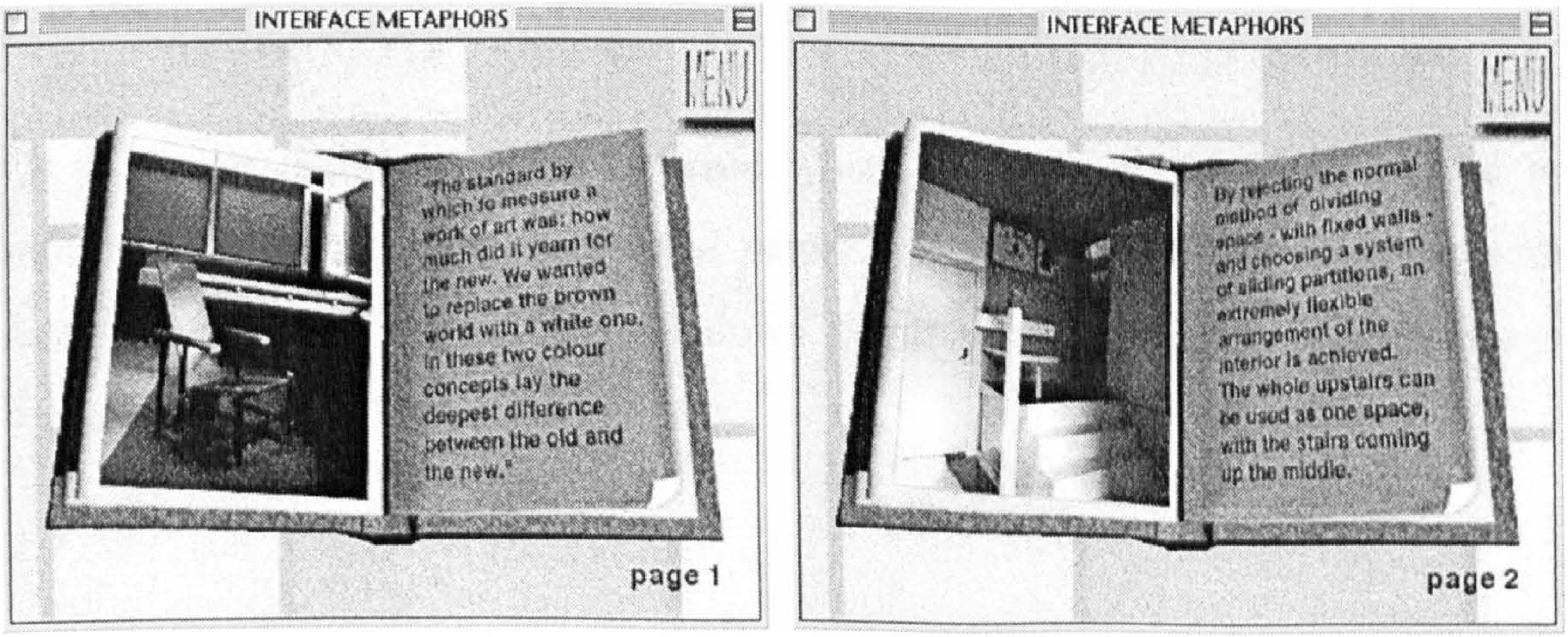
4.3.1 Interface Design

The most engaging multimedia productions are those that have adopted the design paradigm of interfacing through a metaphor of the topic under discussion. Since all users will be familiar with the metaphor then the ability to produce an experiential interface reduces the abstraction of the media and fosters a more ready understanding of the topic.

It is one of the benefits and the joys of multimedia applications that an author can chose from a wide palette of metaphors in the construction of a user interface. Indeed it is this functionality that frees a user from the traditional sets of rectilinear buttons and over familiar menus and allows the possibility of being able to interact with the content in a more natural and exciting way.

4.3.2 Building on a Metaphor

The simplest form of interaction can be addressed using a metaphor that both looks and acts like the real world object. As discussed previously a book can be an example of the simplest linear structure for content. If the nature of the presentation falls into this category then it is possible to structure the interface metaphor on similar lines as in Figure 4.7.



"The standard by which to measure a work of art was: how much did it yearn for the new. We wanted to replace the brown world with a white one. In these two colour concepts lay the deepest difference between the old and the new."

By rejecting the normal method of dividing space - with fixed walls - and choosing a system of sliding partitions, an extremely flexible arrangement of the interior is achieved. The whole upstairs can be used as one space, with the stairs coming up the middle.

Figure 4.7 A simplistic interface metaphor

The layout is designed so that text is displayed on one page, images the next and clicking on the relevant page curl moves the presentation either forward or back. The worth of this metaphor is mainly in its use as an illustration, patently the use of this technology is best suited in providing an alternative to existing products and not as a direct replacement.

The main drawback of the preceding concept is experienced when the user needs to be exposed to information within some superset of conceptual groupings. It is not feasible to expect a user to page back and forward through an entire volume just to be exposed to certain items of information that might have been segregated by several intervening chapters.

4.3.3 Buttons and Menus

The use of buttons and menus will be familiar to anyone who has experienced a graphical user interface (GUI) but there is an added opportunity to give some extra meaning to such an interface by relating it to the subject in hand. An example of such a metaphor is shown in Figure 4.8a. This shows a menu based on a Mondrian work and might serve as an interface leading to selections within a collection of further artwork. Although in this instance utilising the graphic does form a conceptual link to the content this scarcely stretches the metaphorical concept.

On a similar theme consider the case where the content might be related to information of a computer generated nature. In this instance an interface with more deterministic buttons and windows might serve to reinforce the message behind the content. Figure 4.8 b.

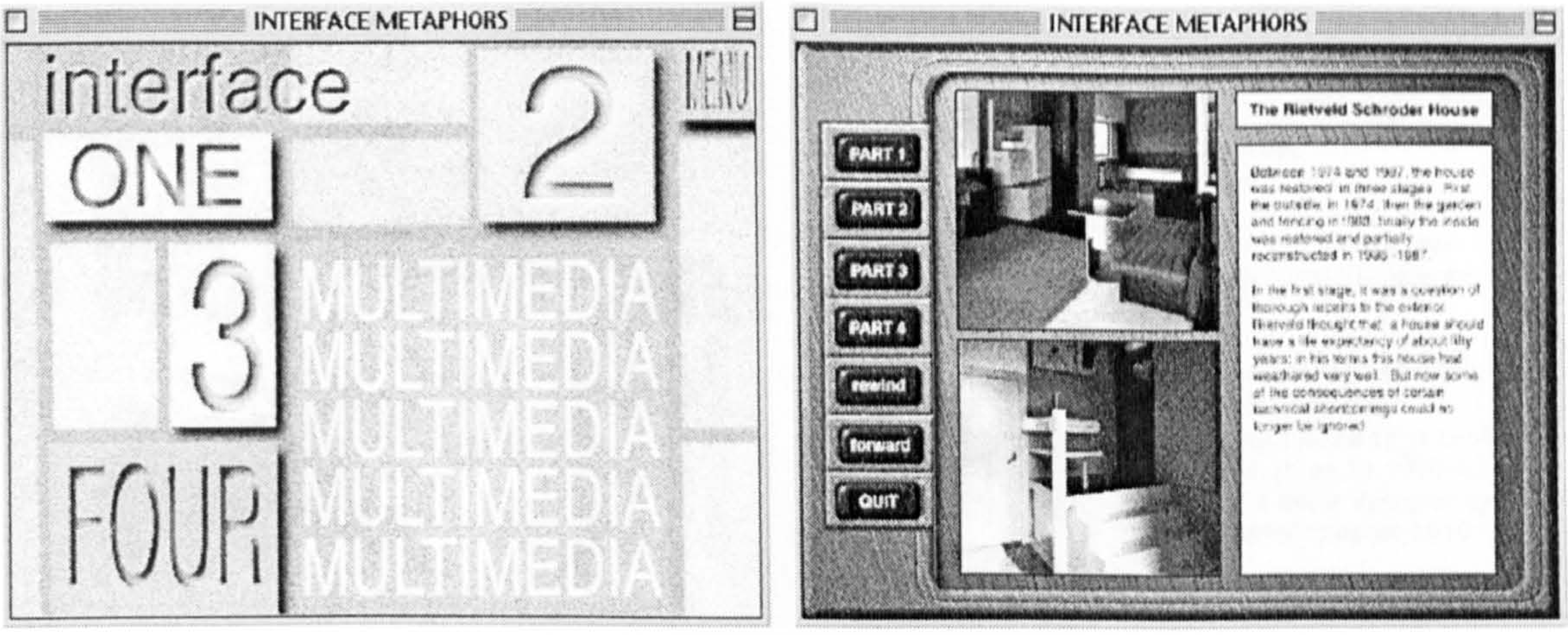
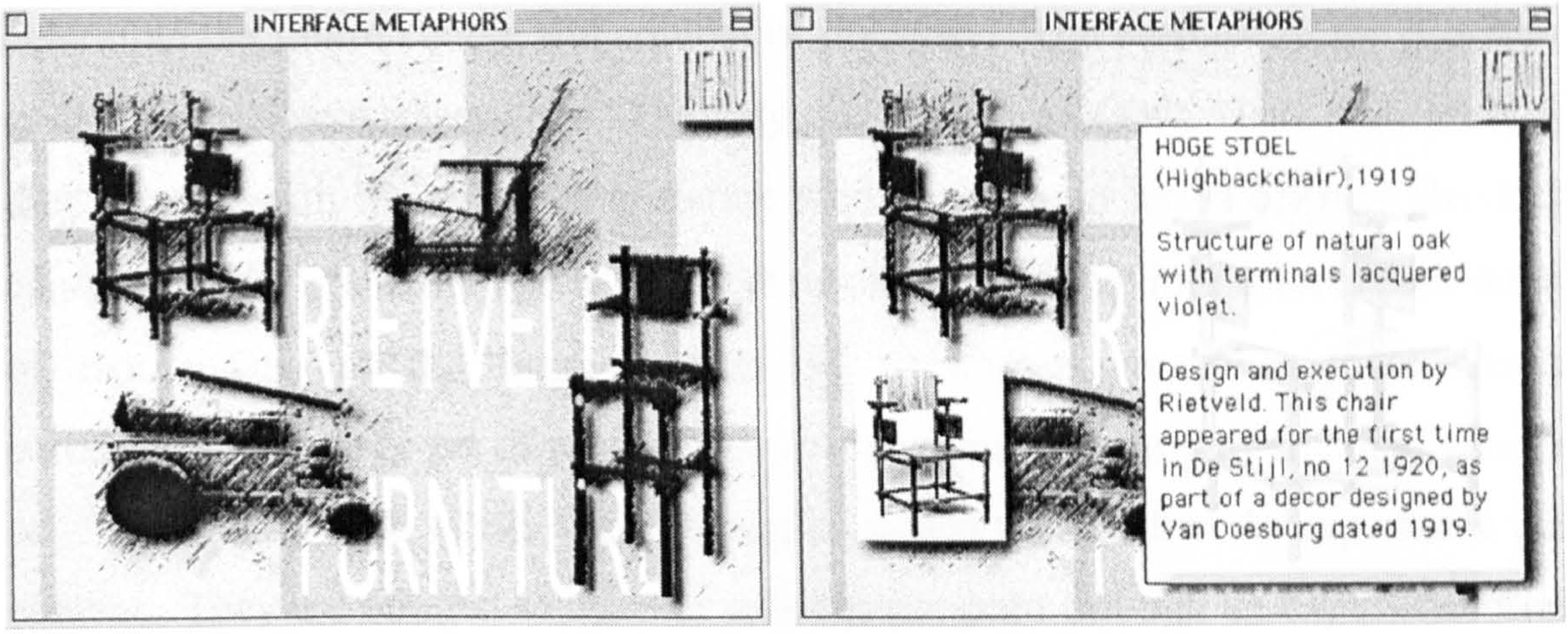


Figure 4.8 Interface metaphors

While this particular approach does serve to address the paradigm of the metaphor it also illustrates a further advance from the simplistic slide show scenario. The number and disposition of buttons hints at the quantity and grouping of information accessible from that page. This is hardly surprising as the "windows icons mouse pointer" (WIMP) interface is the standard for all GUI interfaces. However by exaggerating the concept a very mechanical and regimented interface can be built. This does have a downside in that if a very formalised scheme is adopted then there is a collateral reduction in the flexibility of the layout. This can be manifest, for example, as the difficulty in providing descriptive titles for buttons that possess a common word length or the integration of graphics that are provided in both landscape and portrait format.

4.3.4 The Iconic Interface

In response to the limitations of the above it is possible to construct an interface that is more overt in suggesting its content, Figure 4.9. In this instance there are four icons which, when selected, lead to some corresponding set of information.



HOGE STOEL
(Highbackchair),1919

Structure of natural oak with terminals lacquered violet.

Design and execution by Rietveld. This chair appeared for the first time in De Stijl, no 12 1920, as part of a decor designed by Van Doesburg dated 1919.

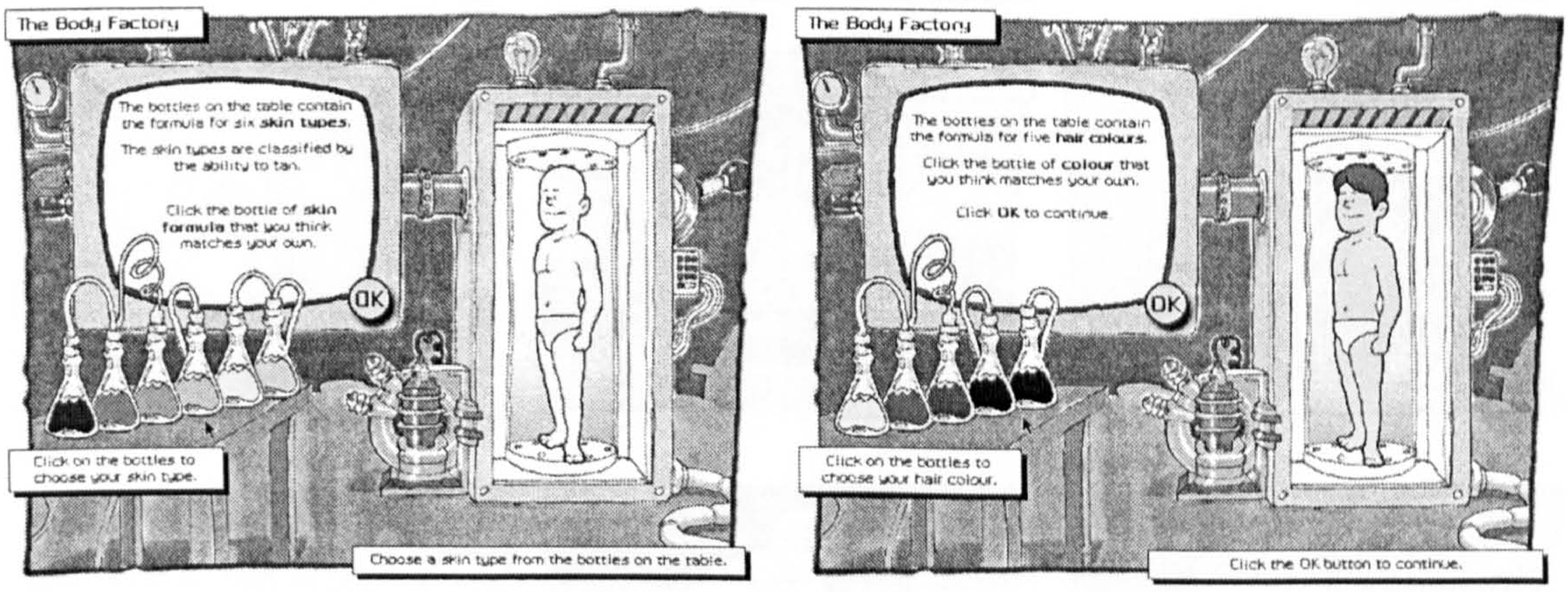
Figure 4.9 The iconic interface.

The advantage of this strategy lies in its transparent nature. A user may arrive at this section and by visually browsing the interface gain a measure of information before even arriving at the true content. In this example the tactic has been to replace the interface with a pseudo hierarchy of abbreviated content that serves as a gateway to deeper levels of data. Because there is no requirement for a strictly ordered set of interface controls these design principles can adopt a more freeform approach to the disposition of on-screen media which serves to accommodate a more diverse range of media types and sizes.

4.3.5 The Experiential Interface

The previous examples serve to illustrate the extraordinarily diverse range of possibilities available to a multimedia designer. However our goal lies in the ability to create and use a truly experiential interface to the content. At its extreme this concept can be categorised as the ability to interact directly with the content while bypassing the need for an interface altogether.

This production was developed by the author as part of a package for primary school pupils illustrating the dangers of over exposure to sunlight. At a conceptual level navigation through the program may be perceived as being purely linear in that there are a set sequence of scenarios that must be visited in a set order. However, within each set piece the user must select certain options that determine the future course and content of the story. The premise of the program is that the user will experience "The Day the Sun Came Out". This requires that they interact with the software to define who they are, in terms of their physical characteristics, where they go and how their dress and behaviour affects their level of risk. Due to the large number of parameters available there are a correspondingly large set of routes through the content leading to an equally large set of outcomes. The worth of this metaphor is that the children do experience the content. They are obliged to interact with the program through an experience that moves them both spatially and temporally in pursuit of the data.


The Body Factory

The bottles on the table contain the formula for six **skin types**.

The skin types are classified by the ability to tan.

Click the bottle of **skin formula** that you think matches your own.

Click on the bottles to choose your skin type.

Choose a skin type from the bottles on the table.

The Body Factory

The bottles on the table contain the formula for five **hair colours**.

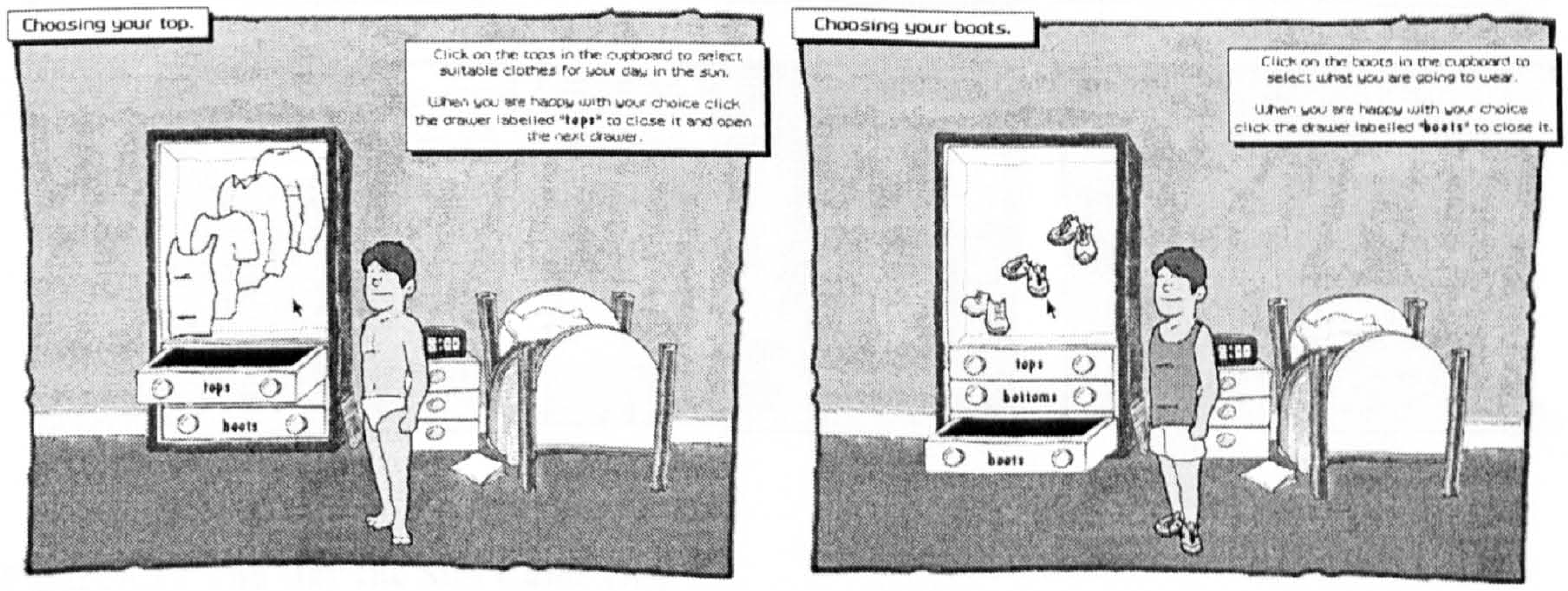
Click the bottle of **colour** that you think matches your own.

Click **OK** to continue

Click on the bottles to choose your hair colour.

Click the OK button to continue.

Figure 4.10 Defining gender, skin type, hair style and colouring.


Choosing your top.

Click on the tops in the cupboard to select suitable clothes for your day in the sun.

When you are happy with your choice click the drawer labelled "**tops**" to close it and open the next drawer.

Choosing your boots.

Click on the boots in the cupboard to select what you are going to wear.

When you are happy with your choice click the drawer labelled "**boots**" to close it.

Figure 4.11 Controlling exposure through choice of clothing.

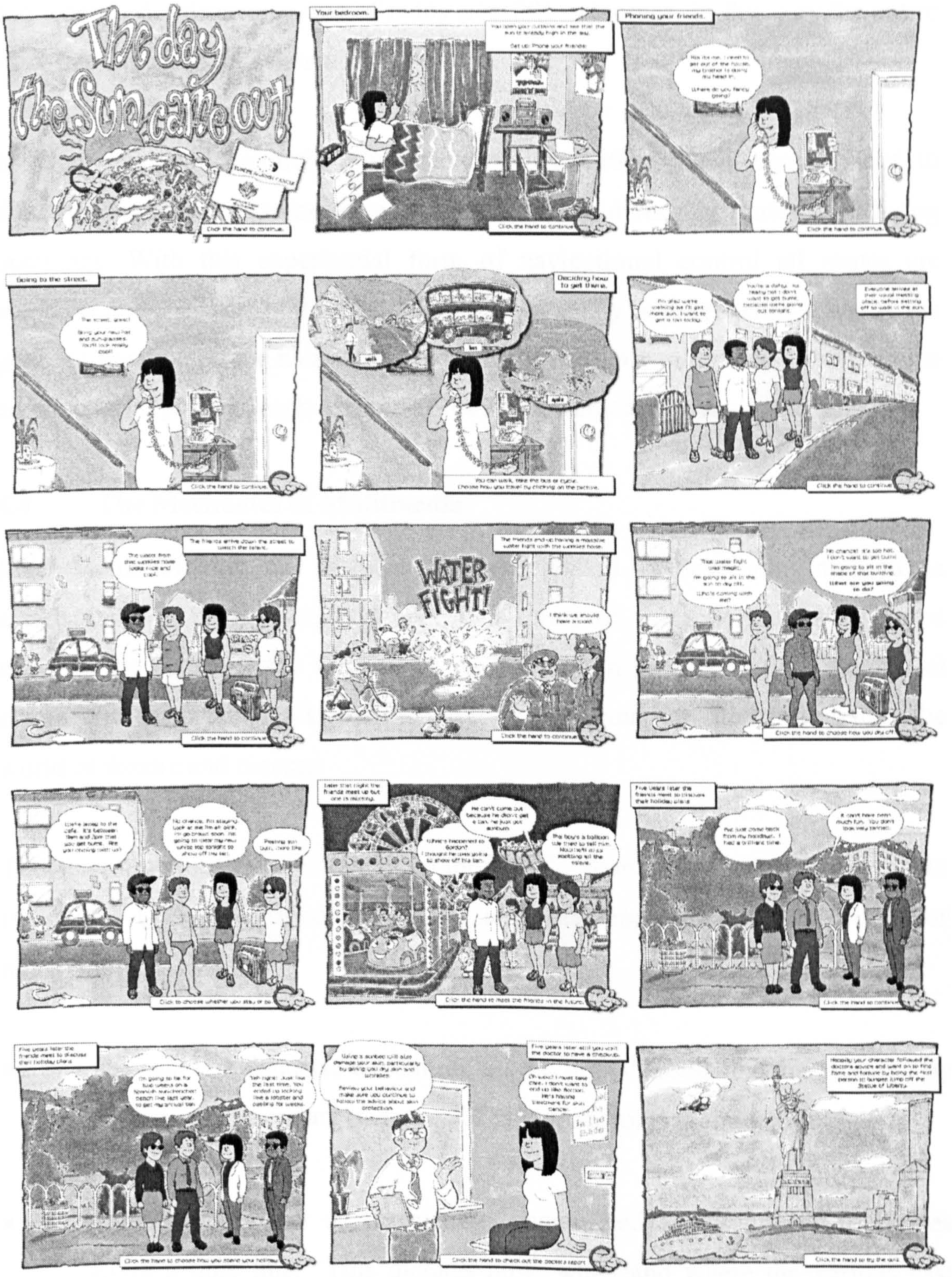


Figure 4.12 The Day the Sun Came Out

Once the user has personalised the central character they can then experience "The Day the Sun Came Out". Further choices throughout the day, such as where to go and what to do, determine one of many possible outcomes, in this way the entire story line functions as a menu.

In a conventional, hierarchically structured application, menu choices tend to direct the user from a common starting point along diverging routes to differing outcomes. With this experiential form of navigational control all routes are common and the user is directed along one of many parallel paths towards the goal. Depending on the nature of choices made there are more than 5000 possible combinatorial pairings each of which provides a weighting towards the outcome.

4.4 The Mechanics of Multimedia

MacroMedia Director, the de facto industry standard for developing multimedia applications, is itself structured around a series of metaphors. As its name suggests the author is the director of the production and the various functional elements of the program take their name from analogous activities from within the world of theatre and cinema.

The main functionality is embodied in the "cast", "score" and the "stage" as shown in Figure 4.13 and described below. The completed application is stored as one or more "movies".

- The cast is a database containing all the digital assets that are required by the program, these are instances of video, audio, text and graphics.

- The score is a set on instructions constructed as a sequence of frames along a time line. This dictates how, when and where the various cast members arrive on the stage.

- The stage is effectively that area of the screen apon which the movie actually plays. Cast members are drawn and animated on the stage under the control of the score and scripts.

Directors default behaviour is to play through the score in a frame sequential order, from first to last. This does not provide much in the way of interactivity, so Director also includes a scripting language "Lingo" which can be used to modify the movie's behaviour.

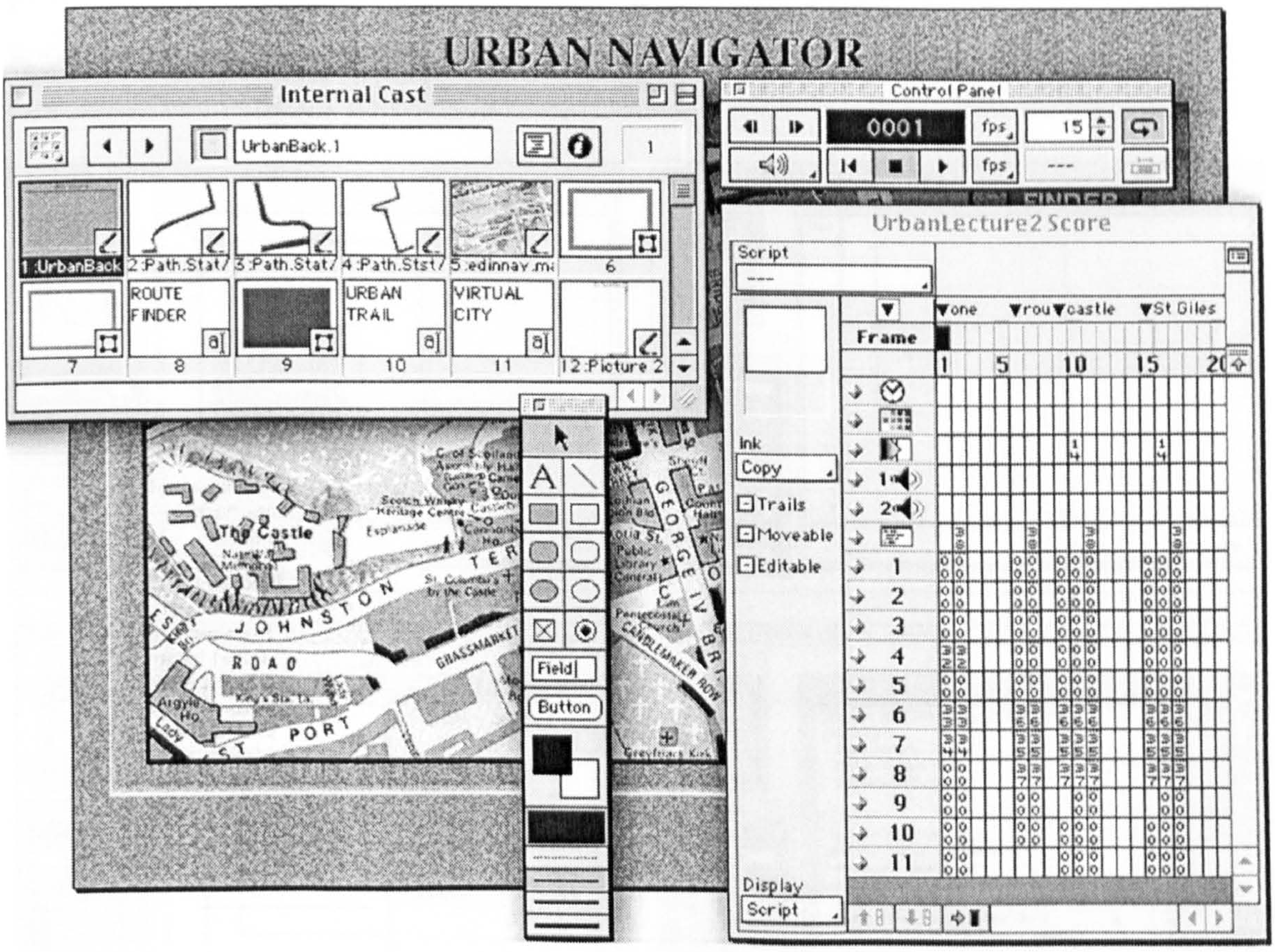
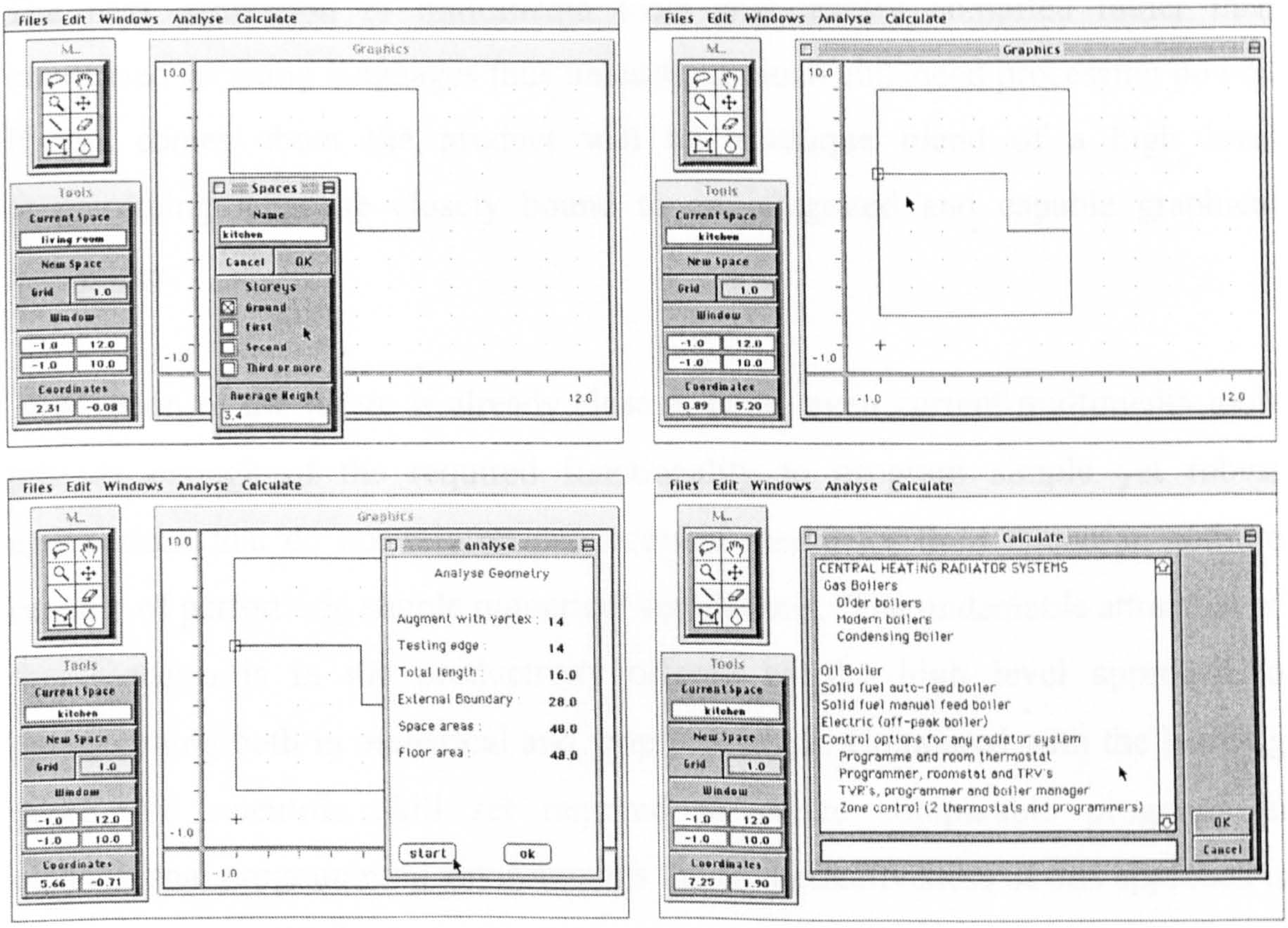


Figure 4.13 Director structure, illustrating elements of code, data and IO.

According to traditional computer science models, the basic elements of a program are its code, its data and its input & output. Within the Director metaphor the score, cast and stage correspond directly to these elements. As applications authored in this environment become more complex emphasis will shift away from the score and towards an increased use of scripting as the main site of a movies control.

4.4.1 The Future of Authoring Tools

The analogy between the computer science model and the components of the Director program is quite powerful. It is all too easy to be dismissive of the potential of multimedia especially in comparison to the power of low level programming languages. However bearing in mind the roots of this analogy it is possible to position multimedia authoring tools not just as a separate and distinct applications but as an entire development environment. (Grant 1995)



```
Analyse Geometry

Augment with vertex :  14
Testing edge :         14
Total length :         16.0
External Boundary :    28.0
Space areas :          48.0
Floor area :           48.0
```

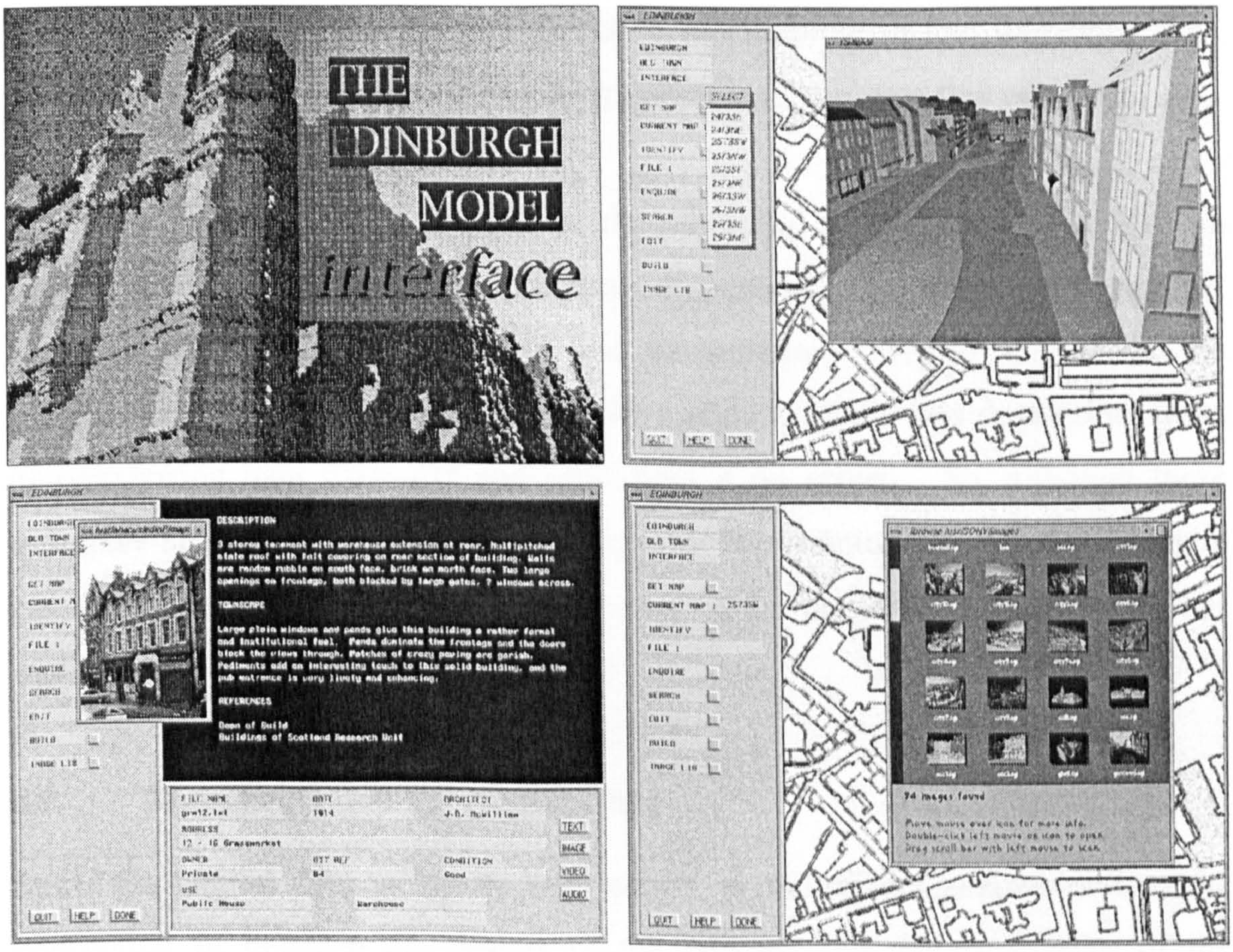
```
CENTRAL HEATING RADIATOR SYSTEMS
 Gas Boilers
   Older boilers
   Modern boilers
   Condensing Boiler

Oil Boiler
Solid fuel auto-feed boiler
Solid fuel manual feed boiler
Electric (off-peak boiler)
Control options for any radiator system
    Programme and room thermostat
    Programmer, roomstat and TRV's
    TVR's, programmer and boiler manager
    Zone control (2 thermostats and programmers)
```

Figure 4.14 Multimedia and CAD

Future stages in this progressive development might see the dissolution of the barrier between conventional programming and multimedia scripting. Given the capabilities of current technology there is no longer such a wide divide between mainstream programming languages and the scripting capabilities of the top end multimedia packages. With the computing power offered by the average desktop platform it is not inconceivable that future multimedia tools may not be regarded as just an authoring package but as a fully fledged programming environment.

explore in safety. However in an expensive hardware environment the cost of providing a machine base purely for training may be prohibitive. A solution that is increasingly being applied is to represent this software environment by recreating the application using those tools commonly used for authoring multimedia. The visual characteristics of the application can very quickly be captured by utilising screen dumps from the original and interactivity can be added by using the functionality inherent in the multimedia authoring package. In this manner the target program can be "cloned" with little effort. The overall functionality of this clone will be restricted, yet with careful attention to the composition of the "media" full functionality for a limited number of routes through the program can be simulated. This strategy can result in the creation of applications that totally mimic the look and feel of the real thing and yet in terms of time, effort and productivity cost much less. Similar economies can be obviously be found in the ability to rapidly prototype and test the interface design for new applications.



DESCRIPTION

3 storey tenement with warehouse extension at rear. Multipitched slate roof with felt covering on rear section of building. Walls are random rubble on south face, brick on north face. Two large openings on frontage, both blocked by large gates. 7 windows across.

TOWNSCAPE

Large plain windows and pends give this building a rather formal and institutional feel. Pends dominate the frontage and the doors block the views through. Patches of crazy paving are garish. Pediments add an interesting touch to this solid building, and the pub entrance is very lively and enhancing.

REFERENCES

Dean of Guild
Buildings of Scotland Research Unit

FILE NAME: gm12.txt
DATE: 1914
ARCHITECT: J.D. McWilliam
ADDRESS: 12 - 16 Grassmarket
OWNER: Private
OS REF: 84
CONDITION: Good
USE: Public House / Warehouse

94 images found

Move mouse over icon for more info.
Double-click left mouse on icon to open.
Drag scroll bar with left mouse to scan.

Figure 4.15 Simulating Unix ™ applications on a PC.

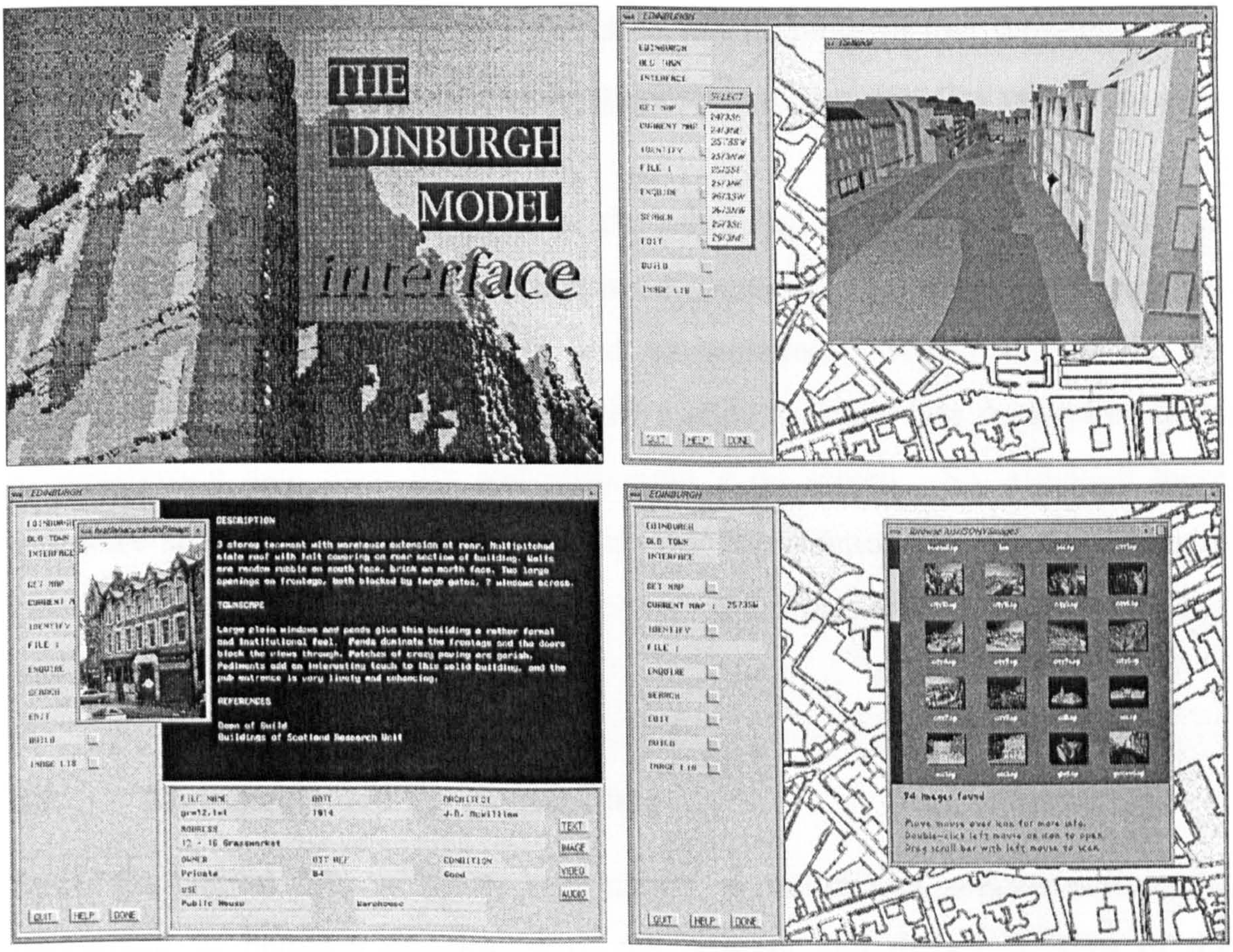
The example shown above, Figure 4.14, was constructed to illustrate this potential mode of development. The application functions as a limited evaluation package allowing the user to digitise floor plans and attribute spaces with functional data. The program then goes on to calculate the formal relationship between spaces evaluating quantities such as volume, internal areas and external wall surface areas before performing steady state heat loss calculations.

The next generation of multimedia tools will feature compiled rather than interpreted scripting languages thus unleashing much enhanced processing power. If this comes about the product will be a unique blend of a high level programming language closely bound to an integrated and capable graphical subsystem.

This vision of the future is already close at hand, even current multimedia tools provide enough of the required functionality to program simple yet robust applications that do not rely on pre-recorded media for their operation and are capable of performing simple numerical simulations. The undeniable attraction of this scenario is in the productivity offered by the high level approach to programming both in numerical and graphical areas. Compared with the learning curve and scientific skill set required to create comparable programs in conventional programming environments the cost effectiveness of this approach is very attractive. While there are many examples that illustrate the progression of this theme it may be useful to consider, as a potential milestone, an application where multimedia tools are used to recreate existing software applications.

4.4.2 In the Role of Simulation

As an example consider the need to provide user training on a complex, multi-user system containing critical data. It may be inappropriate to give a novice users access to such a system yet hands on experience is vital in order to train staff in it's operation and maintenance. One solution is obviously to provide access on a duplicate system where mistakes would not corrupt data and the trainee could explore in safety. However in an expensive hardware environment the cost of providing a machine base purely for training may be prohibitive. A solution that is increasingly being applied is to represent this software environment by recreating the application using those tools commonly used for authoring multimedia. The visual characteristics of the application can very quickly be captured by utilising screen dumps from the original and interactivity can be added by using the functionality inherent in the multimedia authoring package. In this manner the target program can be "cloned" with little effort. The overall functionality of this clone will be restricted, yet with careful attention to the composition of the "media" full functionality for a limited number of routes through the program can be simulated. This strategy can result in the creation of applications that totally mimic the look and feel of the real thing and yet in terms of time, effort and productivity cost much less. Similar economies can be obviously be found in the ability to rapidly prototype and test the interface design for new applications.



DESCRIPTION

3 storey tenement with warehouse extension at rear. Multipitched slate roof with felt covering on rear section of building. Walls are random rubble on south face, brick on north face. Two large openings on frontage, both blocked by large gates. 7 windows across.

TOWNSCAPE

Large plain windows and pends give this building a rather formal and institutional feel. Pends dominate the frontage and the doors block the views through. Patches of crazy paving are garish. Pediments add an interesting touch to this solid building, and the pub entrance is very lively and enhancing.

REFERENCES

Dean of Guild
Buildings of Scotland Research Unit

FILE NAME: grm12.txt
DATE: 1914
ARCHITECT: J.D. McWilliam
ADDRESS: 12 - 16 Grassmarket
OWNER: Private
CONDITION: Good
USE: Public House Warehouse

94 images found

Move mouse over icon for more info.
Double-click left mouse on icon to open.
Drag scroll bar with left mouse to scan.

Figure 4.15 Simulating Unix ™ applications on a PC.

In the previous example the multimedia application was used as a vehicle for presenting pre-recorded material, in response to user interaction, to simulate the operation of some existing program. The next developmental milestone might be an application that is not restricted to the play back of pre-ordained material but functions, in a limited capacity, in it's own right. A typical example might be the provision of a graphical front end to an existing simulation program. Many simulation programs, particularly from an engineering background, still lack the user interfaces that are now taken for granted in most other disciplines. These programs operate either in batch mode or at a very low level of user interaction. The ease with which multimedia tools can create such a graphical interface is their greatest strength and it is easy to see the possibilities inherent in using multimedia applications to perform data preparation to assemble an input stream that can then be used as input to older, less friendly applications. In this instance the user's prime task is to select parameters and data sets that will be used to control the eventual simulation. The capabilities of even the most modest multimedia authoring are tailor made for this aspect of functionality in that they have been designed with the prime aim of providing the tools to facilitate this very task.

The same argument is equally true at the other end of the process, in that the capabilities of programs created in multimedia authoring packages can be applied to tasks related to data recovery and presentation. Again this allies the functionality of the user interface with the computationally less demanding tasks of transferring data from file and presenting it on screen. So, although these examples are not ground breaking in terms of the functionality offered, they are genuinely useful and represent a major step in the progression from the earlier examples.

4.5 Multimedia and the Digital City.

Since the Digital City is, by its very nature, a multimedia conglomeration the conceptual fit between the desires of the concept and the capabilities of the technology potentially provides a direct mapping.

Progress in the development of multimedia applications directed towards describing the built environment can be mapped out through a series of landmarks. The first of which has its origins in sequencing media and is essentially a digital slide show.

4.5.1 The Electronic Slide Show

In most instances of the electronic slide show the incumbent functionality is similar to the traditional use of 35mm photographic transparencies. The added benefits are apparent in a number of areas especially where there is a need to communicate using a heterogeneous set of media. The ability to reduce photographic transparencies and prints, computer generated images and a whole diverse range of traditionally printed materials all to a single, common homogeneous set is perhaps one of the greatest assets of the technology. This means that a presentation can be enriched, to the benefit of the audience, from a wider range of sources. The author has the advantage of creating a permanent asset that can be edited, annotated and updated and which reduces a significant set of materials to a single readily transportable medium which can be duplicated, copied and disseminated with ease. While the above is still true of the photographic process, the availability of cheap, easy to use computer based multimedia editing tools has the advantage of putting this process into the hands of the author certainly allowing greater freedom and perhaps a greater degree of creativity.

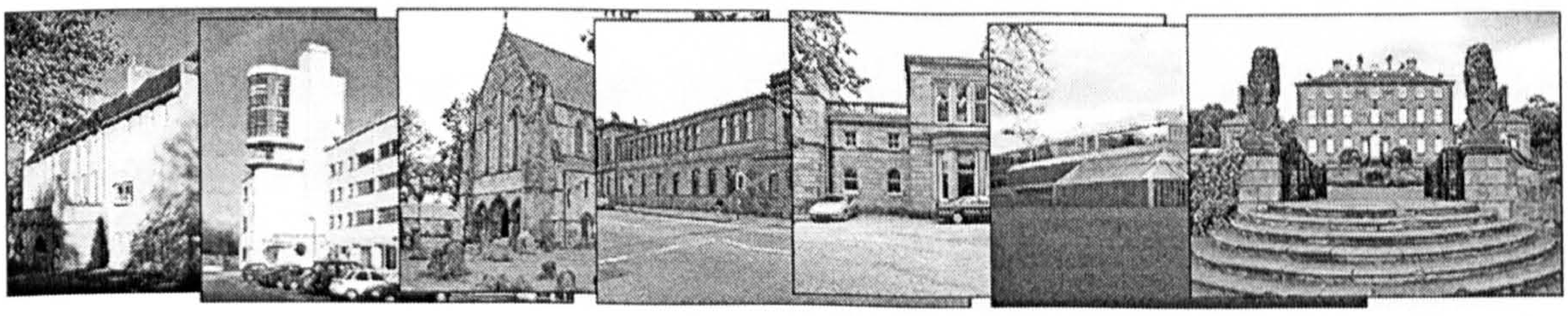
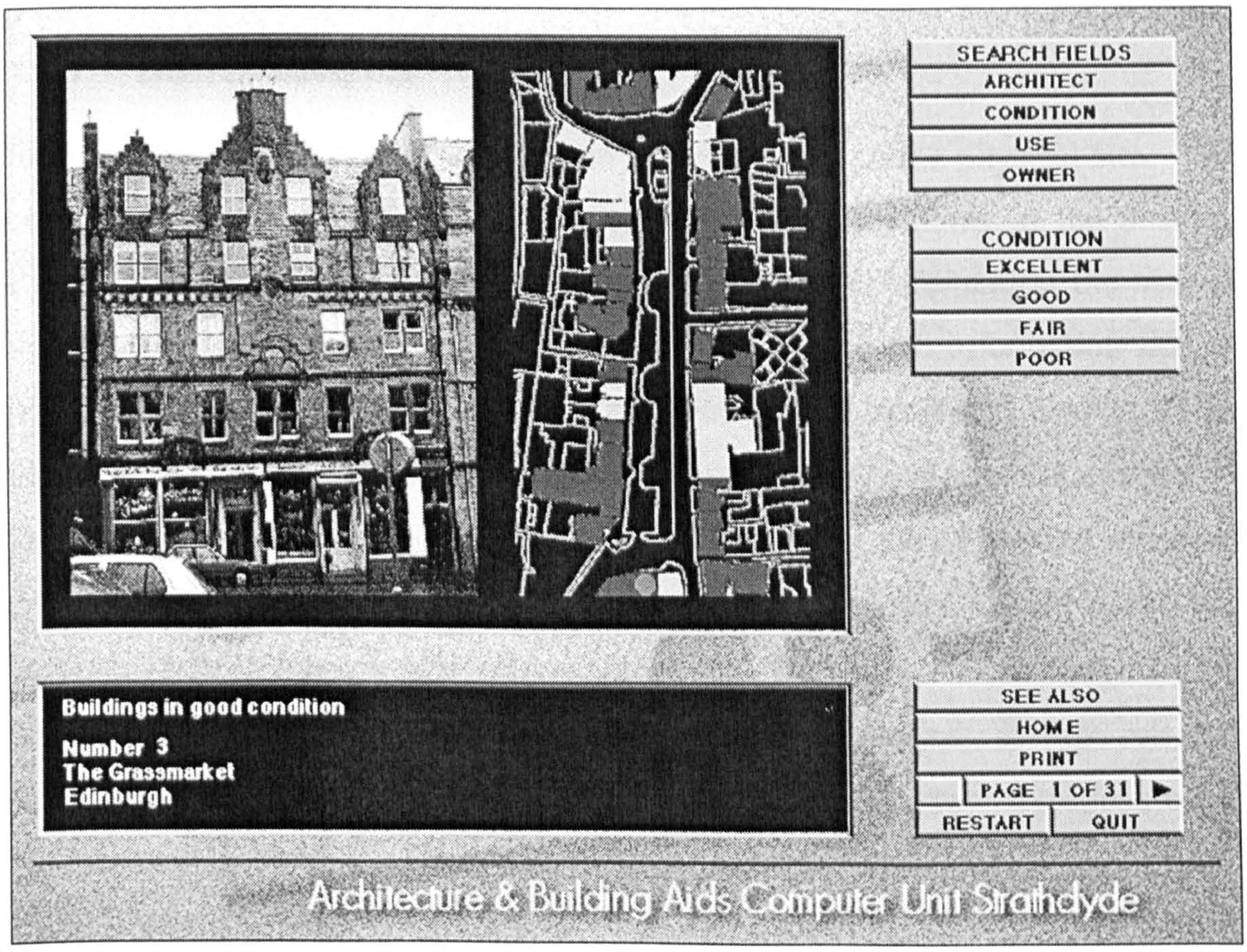


Figure 4.16 Sequencing slides in the electronic slide show.

4.5.2 Hierarchical Navigation

The natural evolution of the product from this point is evidenced in the development of the ability to navigate, first in a linear sequential manner, then bi-directionally, leading on to hierarchical structures and their more complex composite offspring.



Buildings in good condition

Number 3
The Grassmarket
Edinburgh

Figure 4.17 Hierarchical menu implementation

Given that the contents of the slide show might remain unchanged during the above evolution it is apparent that the increased ability to navigate adds functionality, not least in the user's ability to sequence data as the user desires but also through the insights that the un-sequenced combinations might bring. The addition of a hierarchy to the sequence brings an additional element to the presentation, that of an interface through which the user may choose the path through the data. This then forms the foundations of all applications that might lay claim to being an instance of the Digital City. The data may remain unchanged yet the media employed and the nature of the interface provide the distinction between a mere presentation and some more involved form of experiential interaction. Liberation from the constraints of the slide show might come with the provision of a more competent interface but the essential ingredient of the ability to interact with the formal fabric of the city is first evidenced by a metaphor that might be described as discretised space

4.5.3 Discretised Space

In this instance the interface is provided by the context of the physical site. The context can be discretised to any degree within the constraints of the display. Rather than descending a hierarchy when browsing for information the user is unrestrained within the limits of the displayed context. Each information site, or node, can be delineated by an arrow or some similar graphical device. This accrues a number of immediate benefits. The user can instantly see the quantity of information available, by just counting nodes. Similarly the position of the node denotes its position within the real world and the direction of the arrow, not surprisingly, corresponds to a vector within the real world. Each node may be selected in any sequence, there being no sequential constraints, and the relevant information or view is presented on top of the display. Further selections cycle the information content in relation to the choice made.

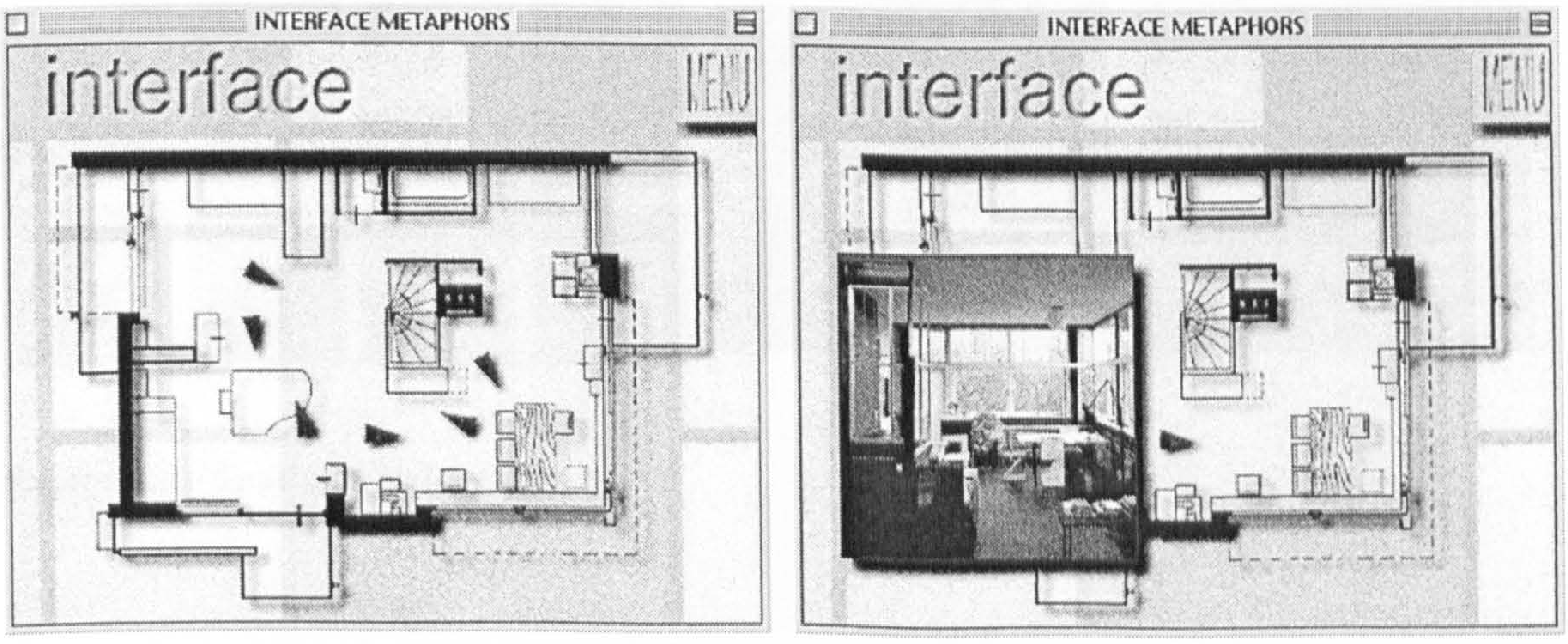


Figure 4.18 Interfacing through the plan and images of 3-D space

This method of depicting architectural space is directly analogous to utilising cartography, or in the instance of a single building, two dimensional plans. The benefits are visible in the ability to layer information within the context of the problem using the strengths of the multimedia technologies to provide a suitable range of media types. While the method might transcend the limitations of the two dimension depiction of the context by reinforcing the scene with three dimension images it does still present one serious limitation.

The major deficiency in this methodology is that while the interface provides a limited context for this scenario there is no natural coherence between nodes and their attached information content. The spatial translation and rotation between one selection and the next has no relevance to past or future selections. This is a limitation in that the user is constrained to sampling the space as against being able to progress through it in a more natural manner.

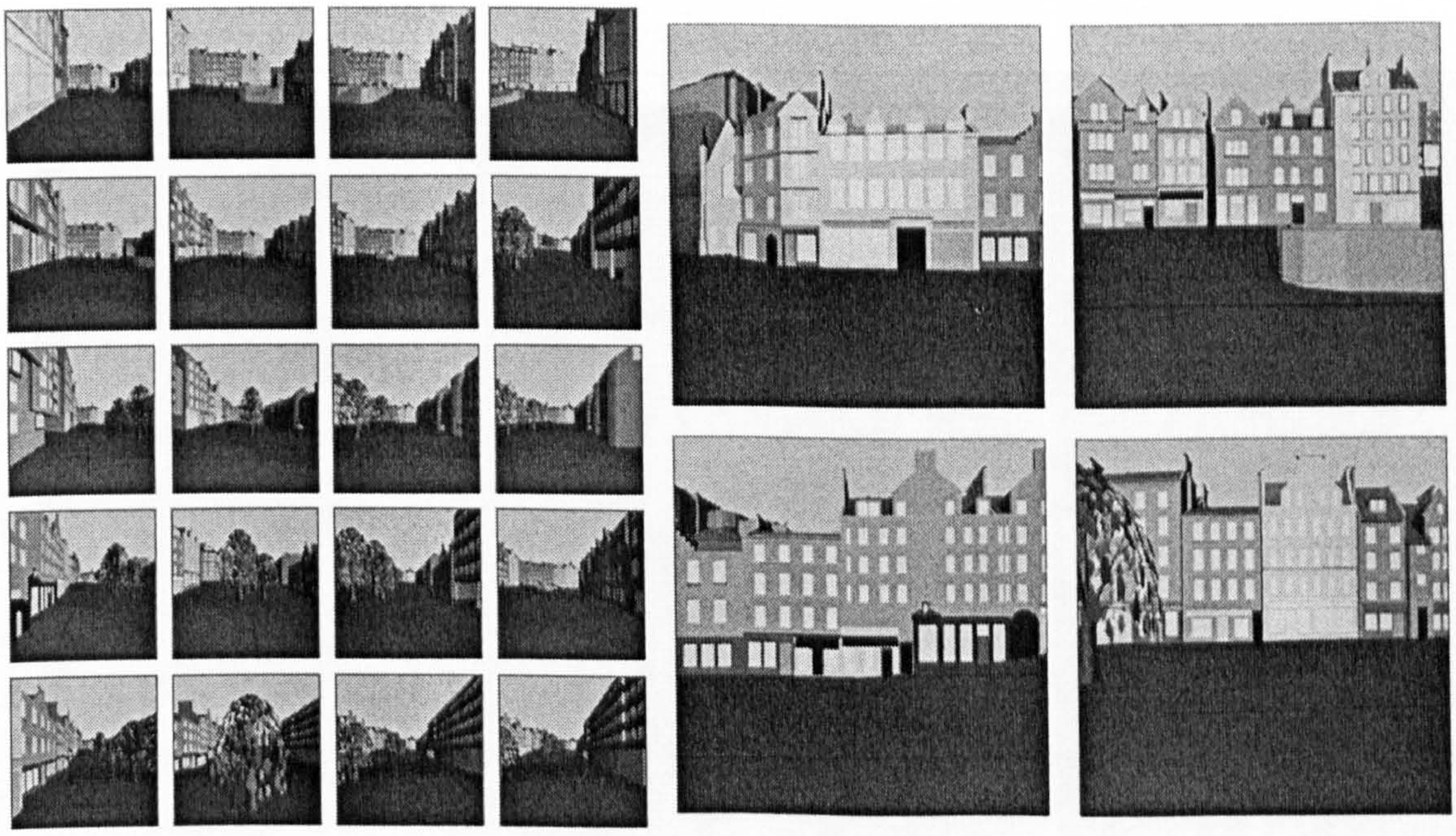


Figure 4.19 Spatial matrices.

The technical implementation of the above is relatively simplistic, any selection sends the application to another location on the time line where the relevant supplementary information is stored. In this instance there will be as many locations on the time line as there are items of information to be displayed. In order to progress beyond the technical limitations of this scenario there is a requirement to progress the computational interaction between the interface and the content. To understand this point consider the classic problem of creating a virtual space where the objective is to allow the user to roam the environment by moving through the space, changing position and direction of view at will. Traditionally this is achieved by creating two matrices of images. The first is the positional matrix which may contain 20 positions, say five in the East - West axis and four in the North - South axis. (Shown above on the LHS) Associated with each index pointing into the positional matrix is a direction matrix allowing four views, say North, South, East and West. (Shown above on the RHS). This gives a sum of eighty images which represent the total degree of freedom allowed to the viewer. In order to interact with this virtual space it would be necessary to provide two sets of controls allowing the user to move forward and back, left and right and to view North, South, East and West.

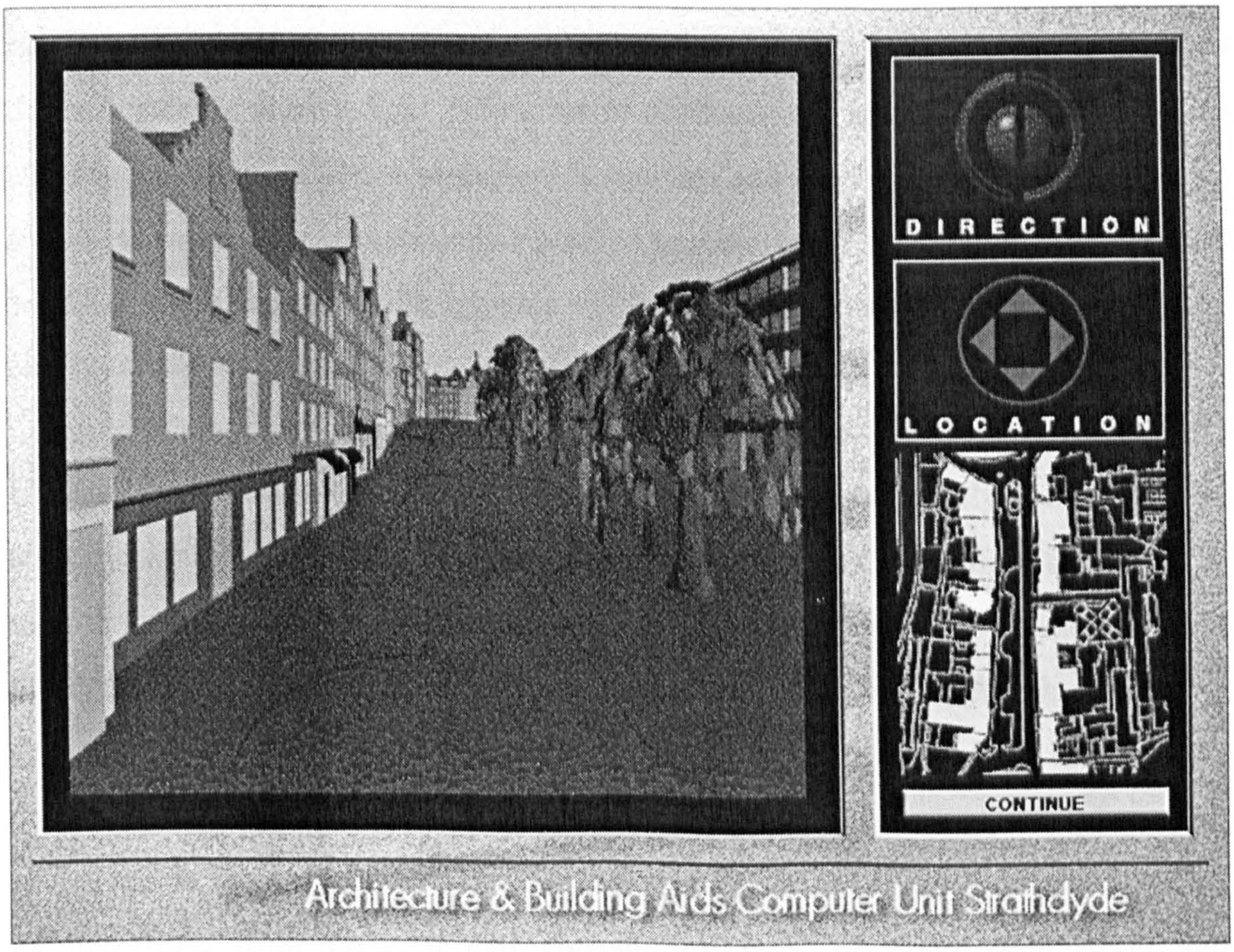


Figure 4.20 The interface to discretised space

Logically there are potential 80x4x2 individual interactions that define the navigational structure of the problem. To sequence this interaction by tying explicit actions to 640 buttons imposes obvious limits to the scope of the technique. However if the authoring environment supports a scripting language then the entire navigational problem is reduced to four conditional case statements and a simple mathematical formula.

This then opens up a whole new horizon in the functionality of multimedia applications in that the constraints of purely sequencing media are overcome and the tools can be applied in a more creative manner.

4.5.4 The Route Finder

Given that the above has suggested some solutions for navigating in and through a virtual space and to retrieving information in the process it is timely to consider further enhancements to the application. One set of features that are always nominated for inclusion is the ability to overlay the application with some set of domain specific knowledge. This class of production is usually aimed at a public audience and has particular relevance to tourism and other casual users of the city. In this instance there is a particular need to communicate some level of familiarity with the spatial layout of the city to a visitor. This is most often evidenced by the ability to use the Digital City as a form of route finder.

If it can be assumed that the visitor will arrive at a conveniently local transport terminus and that there is a limited set of destinations that they might then want to visit then the "route finder" functionality is trivial to implement. Assuming that the application can be localised to the departure point, or maybe this could be selected by the user, then selection from a menu of destinations will result in the correct route being indicated. This can be more advantageous than it might appear in that while similar functionality might be gained from identifying the end points of the route from conventional cartography it is only the application and integration of local knowledge that can suggest the most advantageous route due to local conditions that might not be immediately apparent.

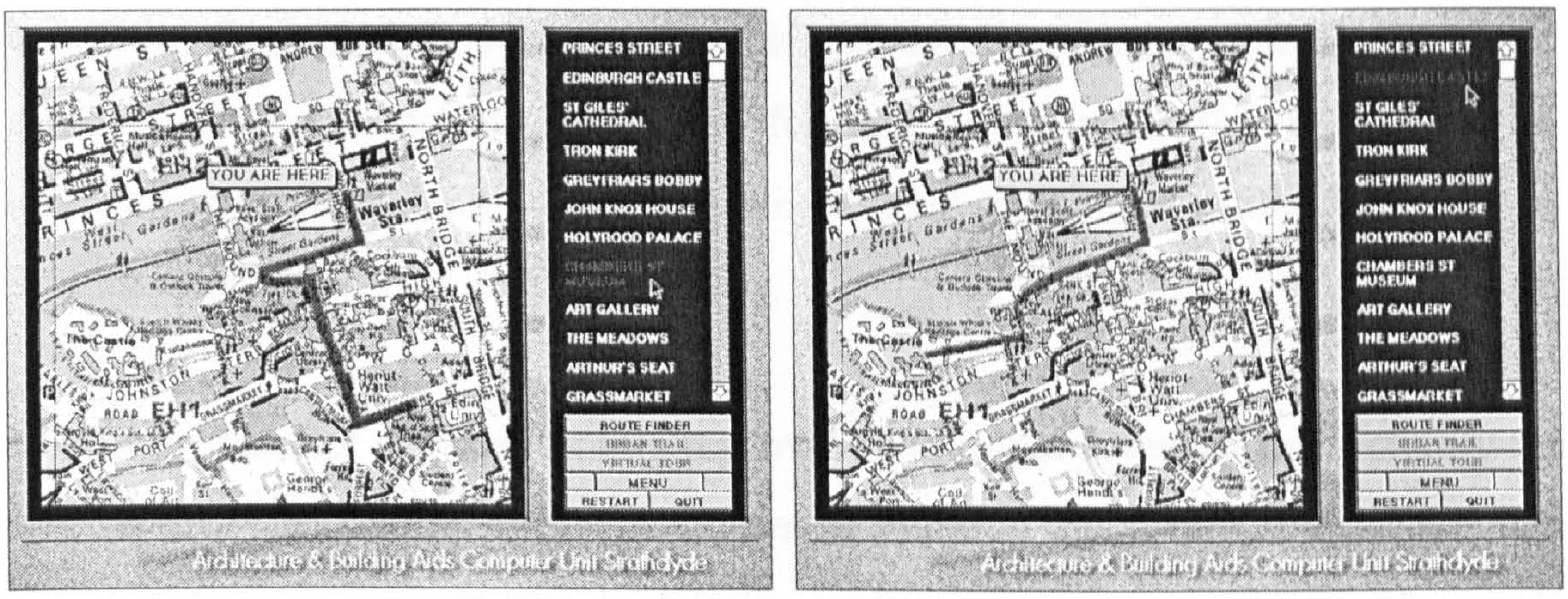
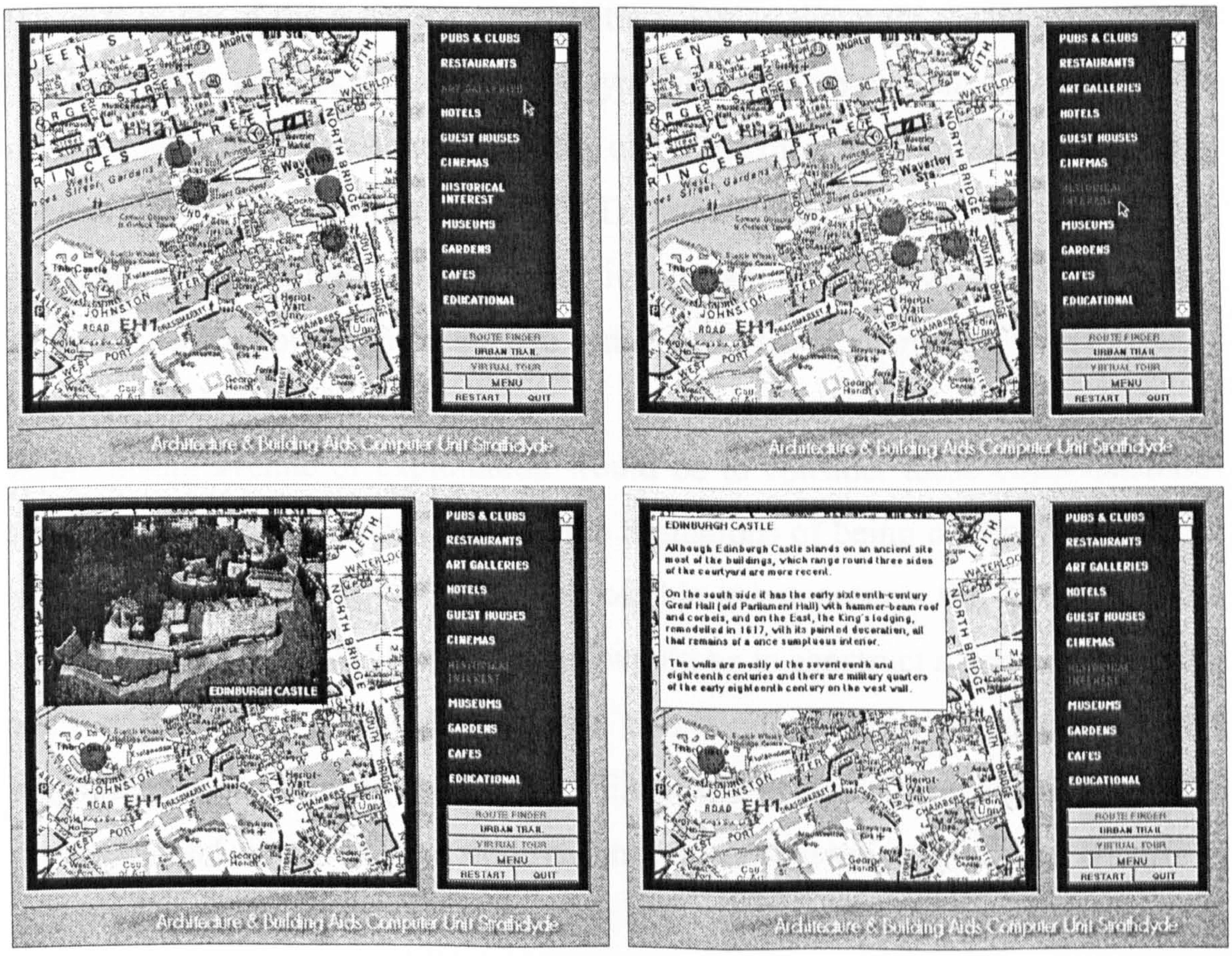


Figure 4.21 Alternate destinations on the route finder

4.5.5 The Urban Trail



EDINBURGH CASTLE

Although Edinburgh Castle stands on an ancient site most of the buildings, which range round three sides of the courtyard are more recent.

On the south side it has the early sixteenth-century Great Hall (old Parliament Hall) with hammer-beam roof and corbels, and on the East, the King's lodging, remodelled in 1617, with its painted decoration, all that remains of a once sumptuous interior.

The walls are mostly of the seventeenth and eighteenth centuries and there are military quarters of the early eighteenth century on the west wall.

Figure 4.22 On the urban trail

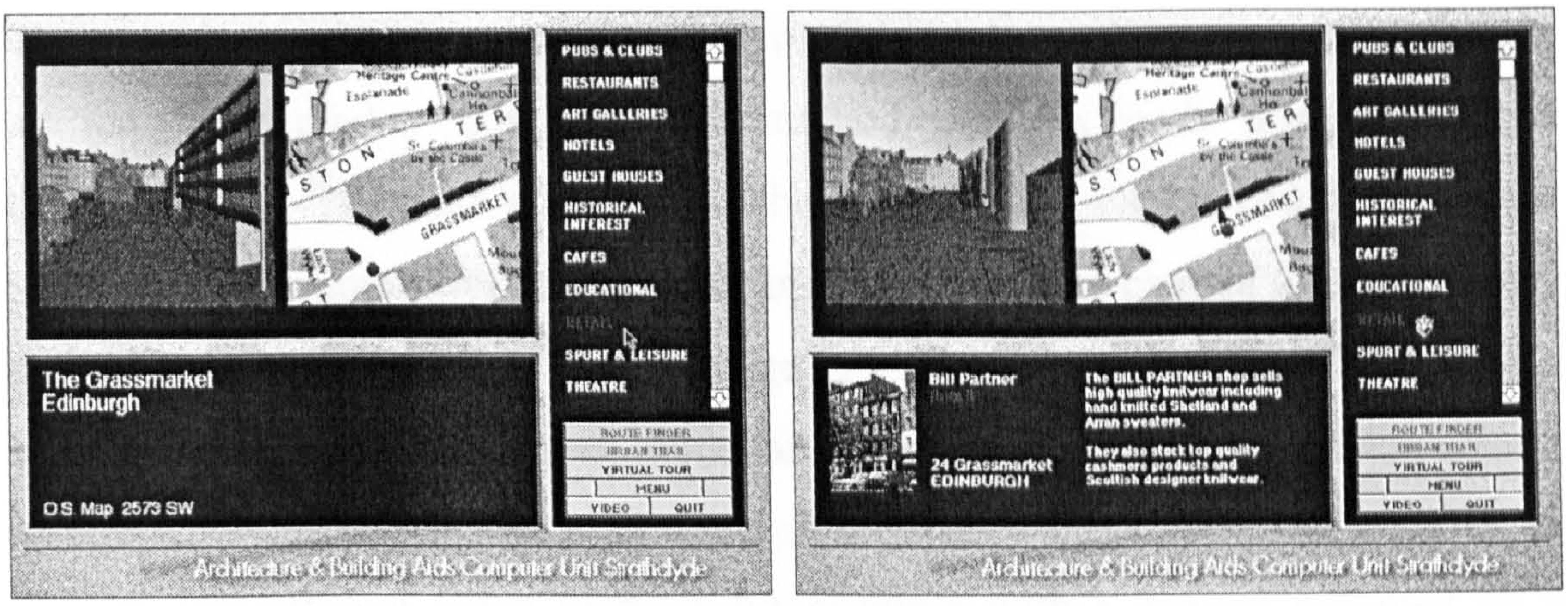
This small amount of "intelligence" is something that is easily embodied in such an application and may be evidenced by the ability to correlate similar buildings or activities into convenient groupings. This then allows the user to access further levels of functionality such as the identification of geographically localised destinations within certain categories forming an urban trail. In addition to indicating the whereabouts of such buildings some measure of information may be also provided, as shown in Figure 4.22.

While this class of application might provide additional layers of functionality the realisation of the Digital City will always be found in the ability to move through the context in a manner that is directly analogous to the way in which people would experience the real world.

4.5.6 A Virtual Tour

Given that one of the main advantages of employing the multi media technologies is the ability to harness differing media types it is then possible to experiment with diverse methods of representation of the city. The most significant impediment to capitalising on the previously cited CAD models was the lack of a suitable low cost computing platform with the graphical capabilities that would allow the use of the model to provide an interactive context for virtual interaction with the city.

A promising avenue of exploitation would be to consider these CAD databases not as models but as media. Within the limitations of being confined to a subset of routes through the city it would be possible to pre process the data and record the sequences as digital video. The aim would be to construct a network of routes, each evidenced by a digital video file, connecting a series of nodes at logical waypoints such as road junctions. Each route would be recorded in both directions and on arrival at a node the subsequent direction would be chosen by selecting the appropriate link to the next file. A wider insight into the immediate context is provided by conventional 2-D cartography with a moving icon giving feed back as to the current location and direction of travel.



The Grassmarket
Edinburgh

O.S. Map 2573 SW

Bill Partner

24 Grassmarket
EDINBURGH

The BILL PARTNER shop sells high quality knitwear including hand knitted Shetland and Arran sweaters.

They also stock top quality cashmere products and Scottish designer knitwear.

Figure 4.23 VR tour

Digital video is a time based medium so it is possible to seek to a known location within the file and display the resulting frame. This property can then be used to index the sequence so that spatial locations in the real world can be related to a temporal location within the video sequence. This relationship is bi-directional in that the sequence can be commanded to display the image of a building that is known to be at a certain number of seconds from the beginning of the clip or if the clip is played linearly some flag could be raised when that same time interval has passed, this then could trigger some other event.

Using this methodology it is possible to transform the purely passive video media into an interactive interface. For example not only can the user navigate freely within the limitations of the prerecorded network but they can also be informed of the location of places or items of interest as they pass by, as shown in Figure 4.23. Similarly the can be transported to some set of destinations as the result of a menu pick or some other user interaction. Pursuing this train of thought to its logical conclusion it becomes apparent that this methodology would offer enormous flexibility. Utilising the same media base the system could be programmed to take the user on such diverse missions as a retail tour, flagging shops within some predetermined set of parameters, or an architectural tour differentiating between buildings of the Georgian or Victorian eras, or maybe just allowing the visitor to browse by wandering at will throughout the network.

The preceding example considers using rendered images from computer graphics as the base media, however this method exhibits both advantages and disadvantages. The main advantage is that, being based on the virtual world, the author has full control over the content and may direct the camera to go anywhere and look at anything. The disadvantage is evidenced in the need for adequate territorial coverage in the source model and the intensive effort required in setting up and rendering all the required routes. A more productive approach would be to acquire conventional video footage, for example with a camera mounted on a vehicle, and then subsequently digitise the analogue output. The problems and advantages incurred by this approach are the inverse of the above in that there is no need to recreate the existing world but that real world physical constraints still exist. However the quantity and quality of the data captured would be significantly greater allowing wider geographical coverage and greater realism.

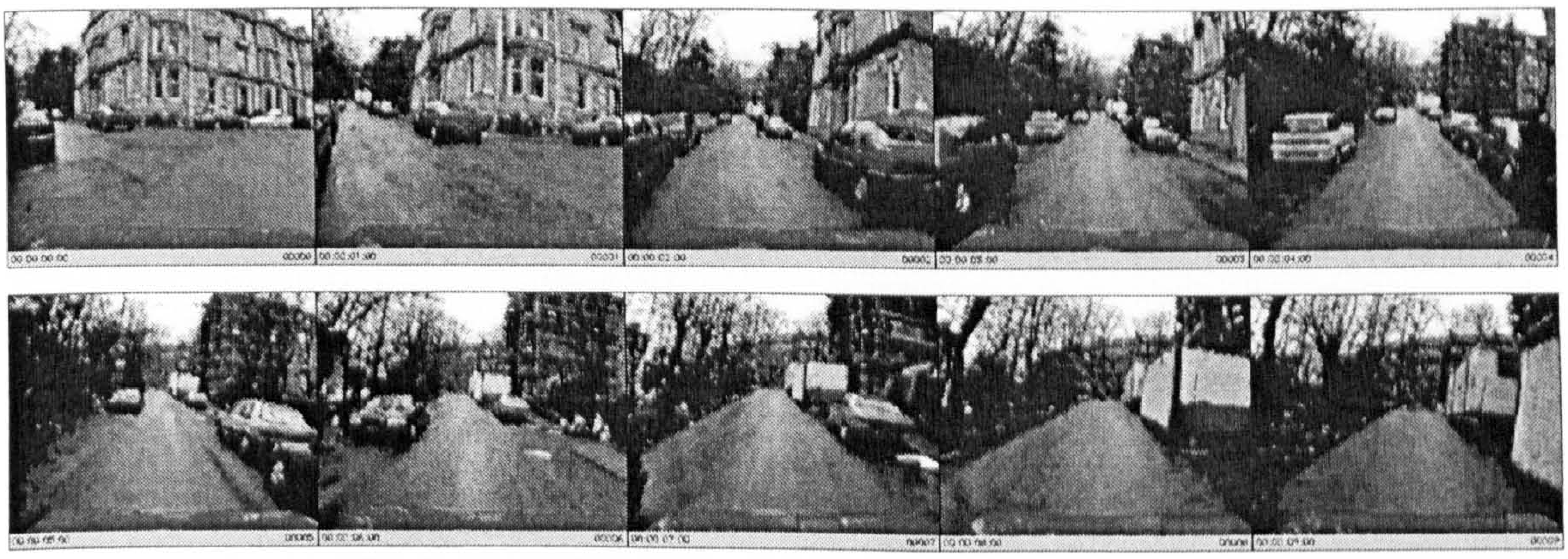


Figure 4.24 Digital video footage

It has been shown that the inherent advantages of the new media technologies is in the ability to harness a wide range of media types in the representation of the Digital City and also to provide an experiential or at least exploratory interface to the context and its associated information content. The flexibility and ease of constructing simple yet competent interfaces has positioned these technologies at the "casual user" end of the technological scale and this has served to relegate their exploitation to archival rather that technical information storage and retrieval. This distinction is based on differentiating between information archetypes that tend towards the graphical rather than the predominantly textual and numerical. Equally a multimedia application will typical tend to augment the purely factual content with some element of narrative that serves to position the production within some domain relative to the information being conveyed. This targeted outcome is due in no small part to the nature of the technology. Multimedia documents are generically more time consuming to author and therefore tend to be more focussed on limited domains. In turn this leads to the adoption of media elements bound within a specific style of interface that tend to serve a particular application area.

4.6 Case Studies

This section gives a brief overview of a small range of multimedia productions that, in their own way, provide a unique example of one of more of the strengths of the technology. These factors encompass the treatment of time – in an historical context – detail, scope and media and knowledge integration. They are all typically representative of the genre and yet are markedly different in their "look and feel" and modes of interaction.

4.6.1 Skara Brae

In 1996, ABACUS was approached by Historic Scotland to assess the suitability of new media in offering an alternative to the traditional presentation formats then in use. This test case focused on Skara Brae, northern Europe's most extensive and elaborate Neolithic Village. Revealed in 1850 as the result of a massive storm, the site exhibits most, if not all of the problems associated with heritage sites. It is remote, situated near Stromness in the Orkney Islands, off the northernmost coast of mainland Scotland, conservation is a strong issue, as visitors are already being restricted in terms of access to some areas of the site, and there is also a need for explanation and interpretation as much of the original building fabric has disappeared over time.

The brief called for an exemplar that could be used as a vehicle for debate as to the focus and scope of multimedia applications in providing interpretative material for a proposed visitor centre at the site. Historic Scotland cited three main goals that required to be addressed, these being; navigation, reconstruction and interpretation.

A central concern over the conservation of the Skara Brae site had been raised by the attrition on the remaining fabric due to erosion caused by visitors. This had lead to restricted access to much of the area with the consequent draw-back of separating the visitors from the objects of interest. This paradox is typified by the images shown in Figure 4.25 which show on the one hand the barrier to access and on the other the freedom provided by replacing the reality with an alternative route provided by captured digital video sequences.

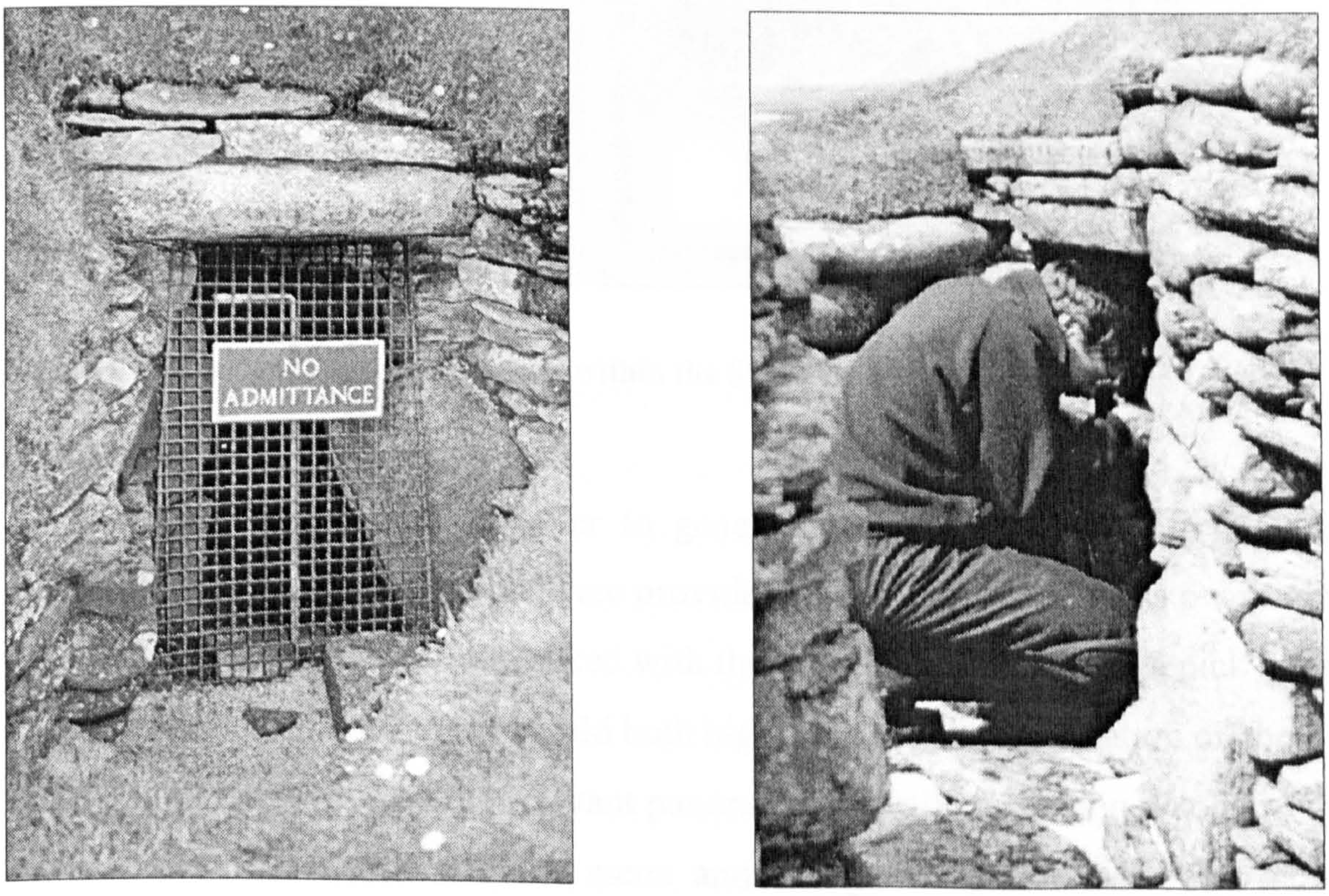
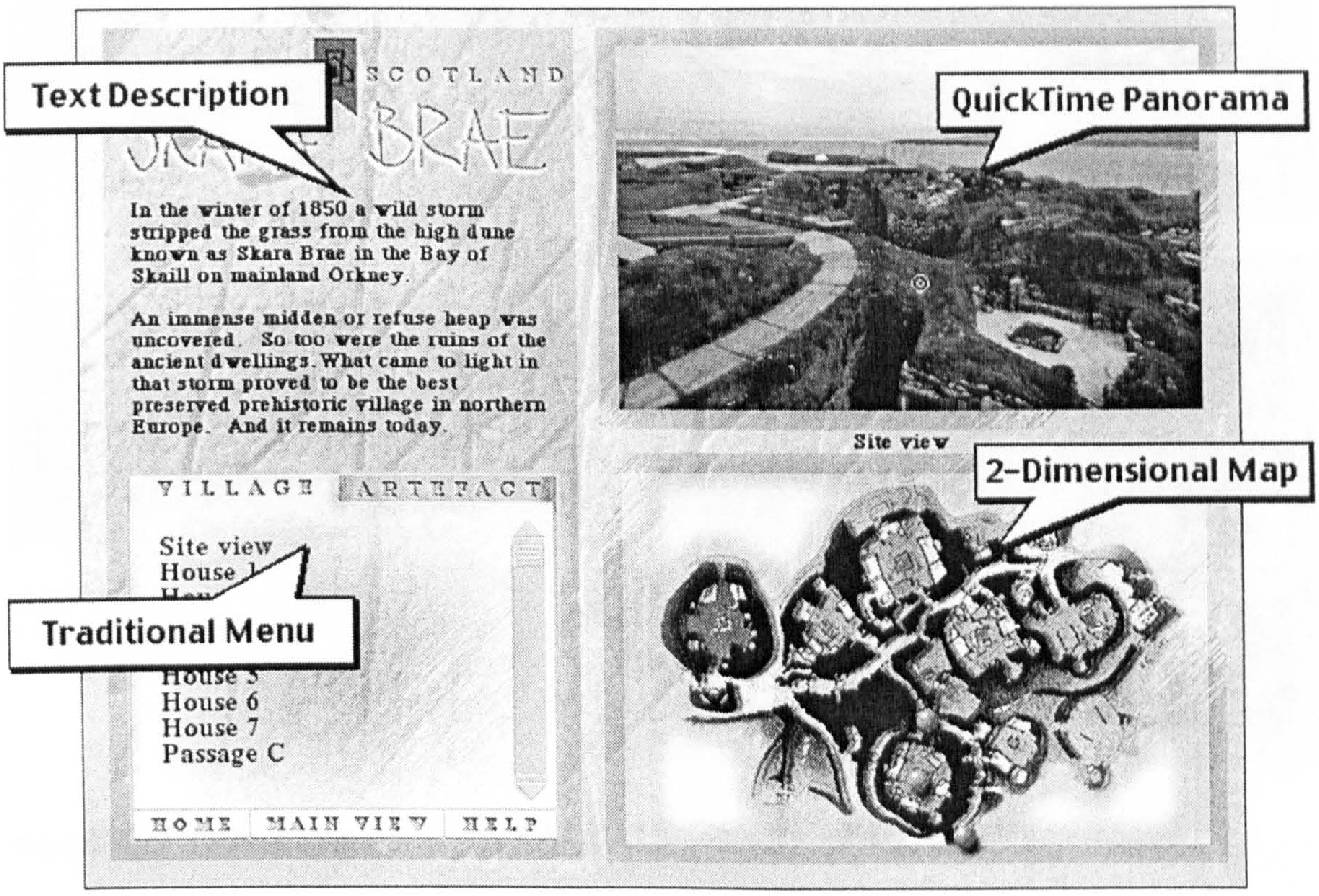


Figure 4.25 Replacing physical access with the virtual equivalent.

The goal was then to replace this experience by creating a virtual environment through which visitors could then explore the site. The technical challenge would be to formulate a low cost, short time scale approach that would be both easy to operate and still provide an experience commensurate with a visitor's expectations. In order to explore the available options the exemplar was constructed around the three most logical strategies, as shown in Figure 4.26.

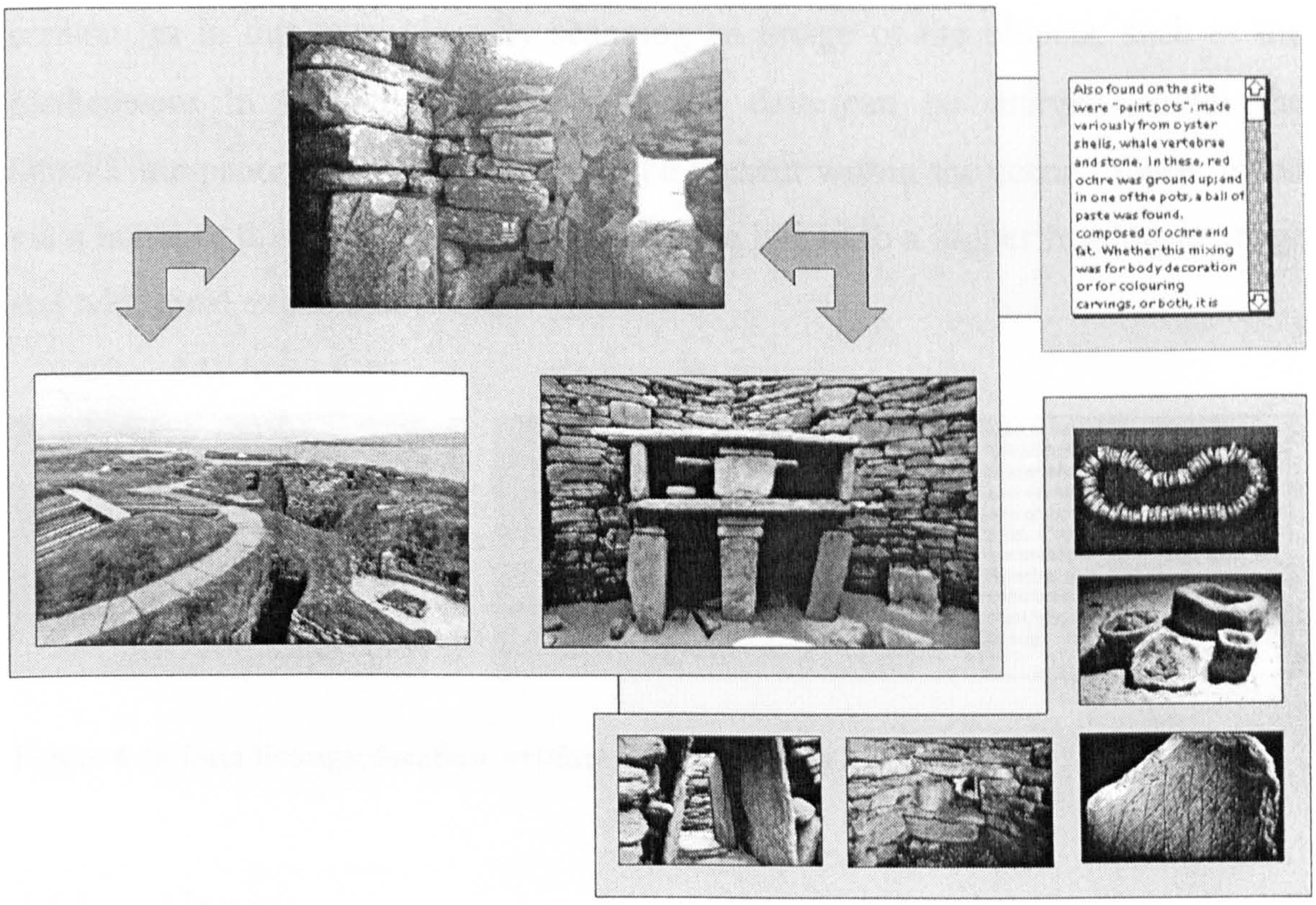


In the winter of 1850 a wild storm stripped the grass from the high dune known as Skara Brae in the Bay of Skaill on mainland Orkney.

An immense midden or refuse heap was uncovered. So too were the ruins of the ancient dwellings. What came to light in that storm proved to be the best preserved prehistoric village in northern Europe. And it remains today.

Figure 4.26 Navigational opportunities within the Skara Brae exemplar

As previously stated, and in order to generate an informed debate, all three potential navigational metaphors were provided. Each functioned in its own right and each was functionally inter-linked with the others. In this manner a pick from the traditional text based menu would both high-light the chosen location on the 2-dimensional map and load the relevant panorama. Equally, choosing a node from the map would activate both the menu and change the panorama. The most natural and immersive mode of access proved to lie in the ability to utilise the panoramas to browse the site, allowing the user to find the inter-linked passages and doorways, and by so doing traverse the site much as they would if on foot. This mechanism was enabled by linking each node by a sequence of digital video that when activated would transport the user from one node to the next, as shown in Figure 4.27.

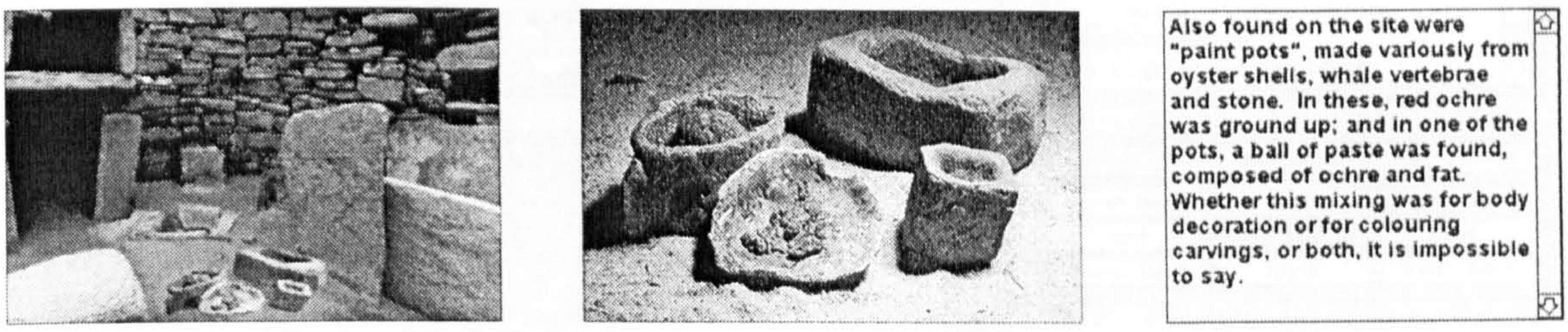


Also found on the site were "paint pots", made variously from oyster shells, whale vertebrae and stone. In these, red ochre was ground up; and in one of the pots, a ball of paste was found, composed of ochre and fat. Whether this mixing was for body decoration or for colouring carvings, or both, it is

Figure 4.27 Linked nodes to access graphics and textual descriptions.

Figure 4.27 shows the relationship between three of the nodes, each interconnected by sequences of digital video, linked between and activated by, hot-spots. Each video sequence is recorded bi-directionally so that any path can be traversed in either direction. A user can thus browse any node in the scene and on finding a doorway or passage entrance potentially select any one of a number of possible routes. In this manner the entire site could be explored and while the user may be constrained within the pre-recorded set of routes the order in which they can be explored is of their choice.

The panorama at bottom right of Figure 4.27 has additional hot-spots linked to coupled graphic and textual information that can be referenced from locations within the scene. These items are usually those that, once found during the original archaeological investigation, are then removed for display at national museums. The displacement of an object from its context is a much criticised but practical necessity, when considering the security and conservation of historical artefacts. However it is now possible to "virtually" reunite objects with their context, as in this example. By scanning an image of the objects, such as the earthenware in Figure 4.28, the resulting data can be merged within the QuickTime panorama and so appear to be present within the scene. Once linked via a hot-spot this portion of the image can be linked to a higher resolution image and additional explanatory text.



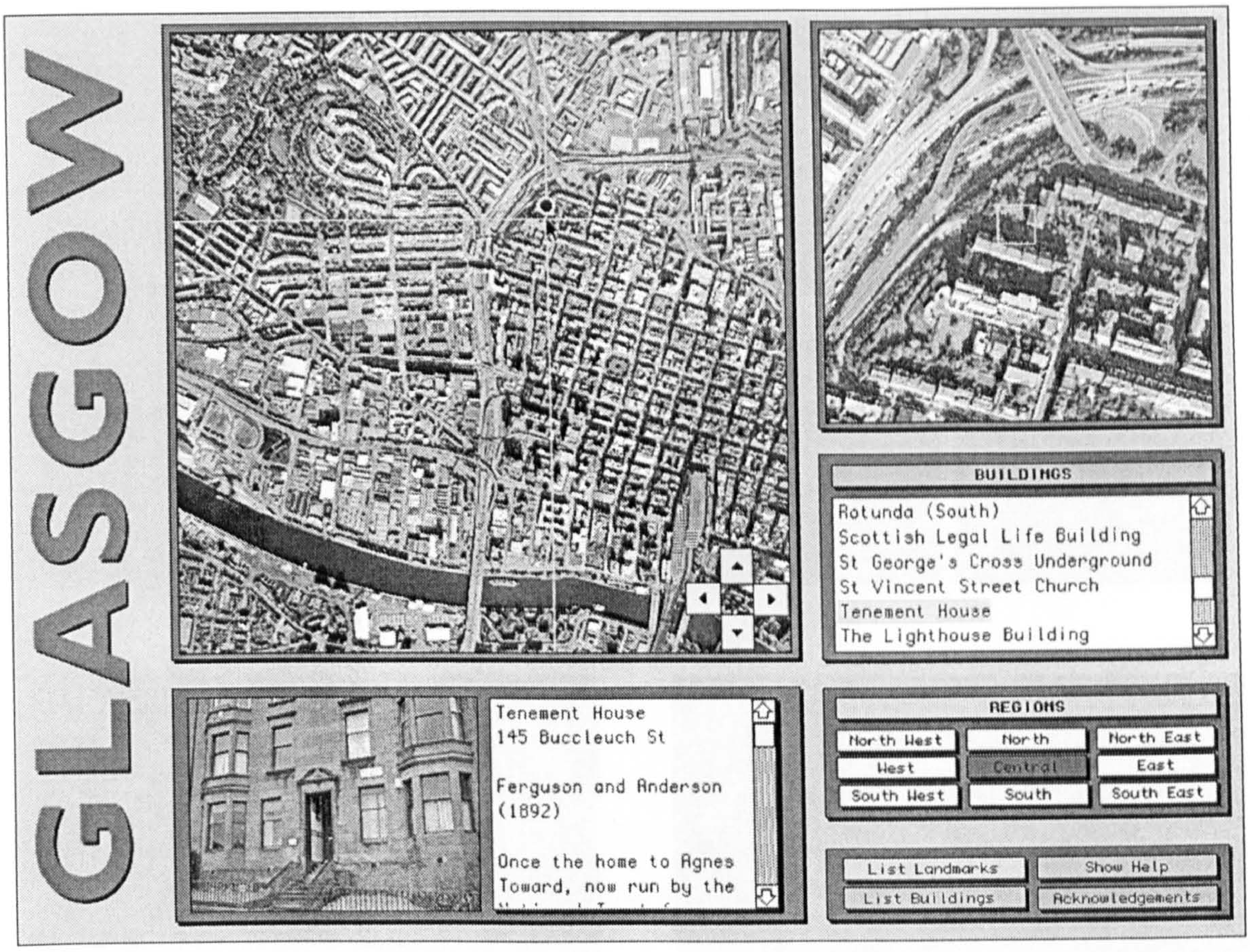
Also found on the site were "paint pots", made variously from oyster shells, whale vertebrae and stone. In these, red ochre was ground up; and in one of the pots, a ball of paste was found, composed of ochre and fat. Whether this mixing was for body decoration or for colouring carvings, or both, it is impossible to say.

Figure 4.28 Data linkage; location, artifact and description

4.6.2 Airmaps

As previously stated one of the key advantages of the multimedia technologies is their ability to transcend time and space through the ability to handle diverse media in an efficient manner. Common to all information applications with a wide coverage, be it geographical or otherwise, there is always a difficulty in providing an access paradigm that facilitates browsing, on a plane on which the user is not subject to information overload, while still allowing ingress to the data below. This is particularly relevant where the object of interest is contained within a context that in itself provides a measure of the information sought.

In the specific case of an interface to urban information there will always be the desire to access information on a large scale while still preserving the ability to tunnel down to specific material. Conceptually this format forms an inverted pyramid with a broad base representing the context and progressively smaller layers providing information until the apex, the topic of interest, is reached.



Rotunda (South)
Scottish Legal Life Building
St George's Cross Underground
St Vincent Street Church
Tenement House
The Lighthouse Building

Tenement House
145 Buccleuch St

Ferguson and Anderson
(1892)

Once the home to Agnes
Toward, now run by the

Figure 4.29 Glasgow Airmaps

An application of the inverted information pyramid is shown in Figure 4.30. The block of nine images is a mosaic of aerial photography covering most of the conurbation of Glasgow and the outlying region. Each of the nine regional maps is composed of four elements, as shown on the top RHS of Figure 4.30, allowing for thirty-six individually accessible area maps in total. Each of the area maps can be browsed at low resolution and zoomed to a high resolution.

The user can navigate over the region by accessing each of the nine maps through an interface that relates the maps to the points of the compass arrayed around the centre of Glasgow.

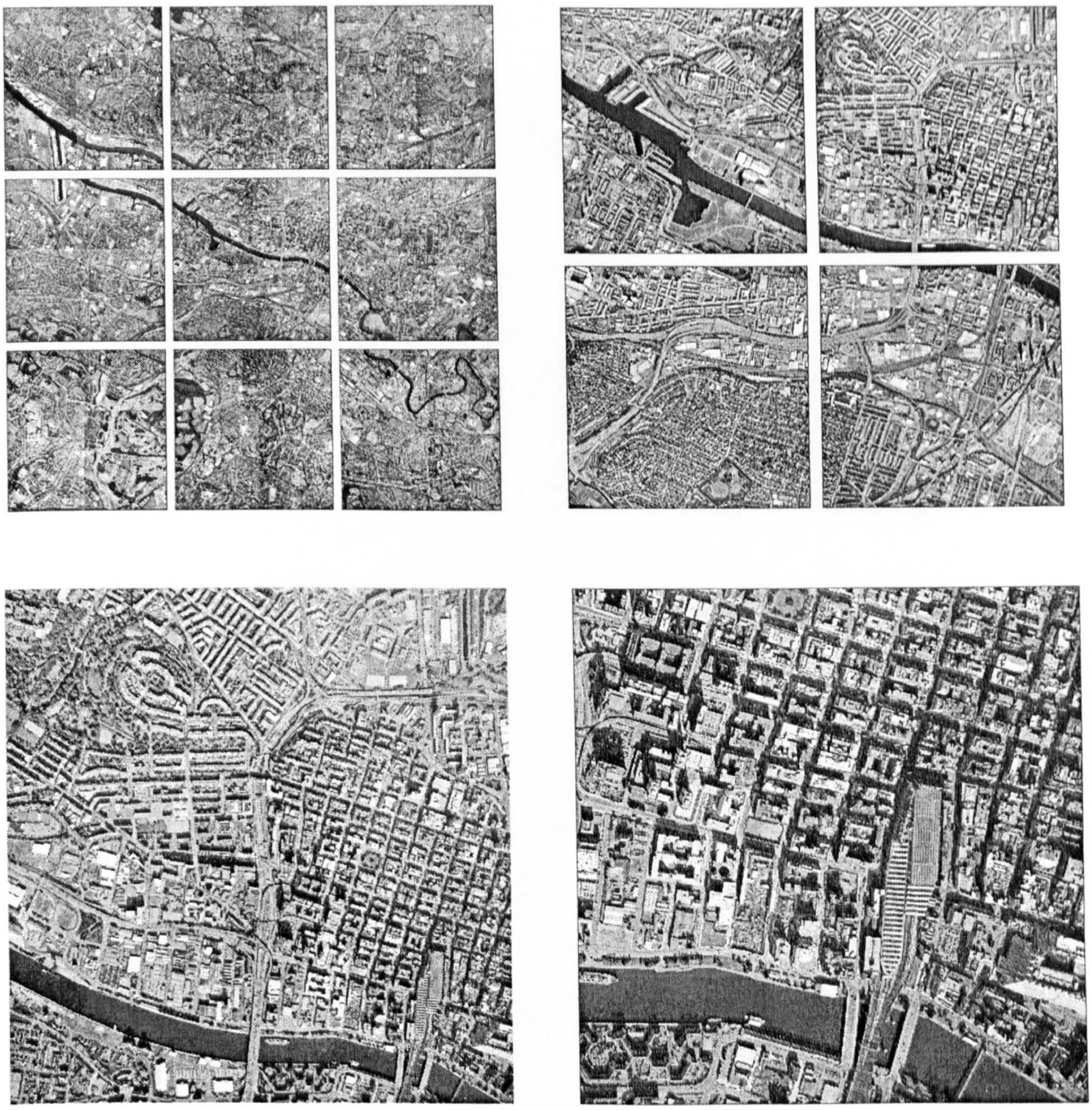


Figure 4.30 Heirarchy of detail within maps

Once the region is chosen, each of its constituent parts can be accessed in turn using the four arrow buttons located to the lower right of the current map, as shown in Figure 4.31. These serially provide the next left, top, right or lower map as required. Within each map landmarks are accessible from a scrolling list, these flags are provided to let users orientate themselves in unfamiliar surroundings.

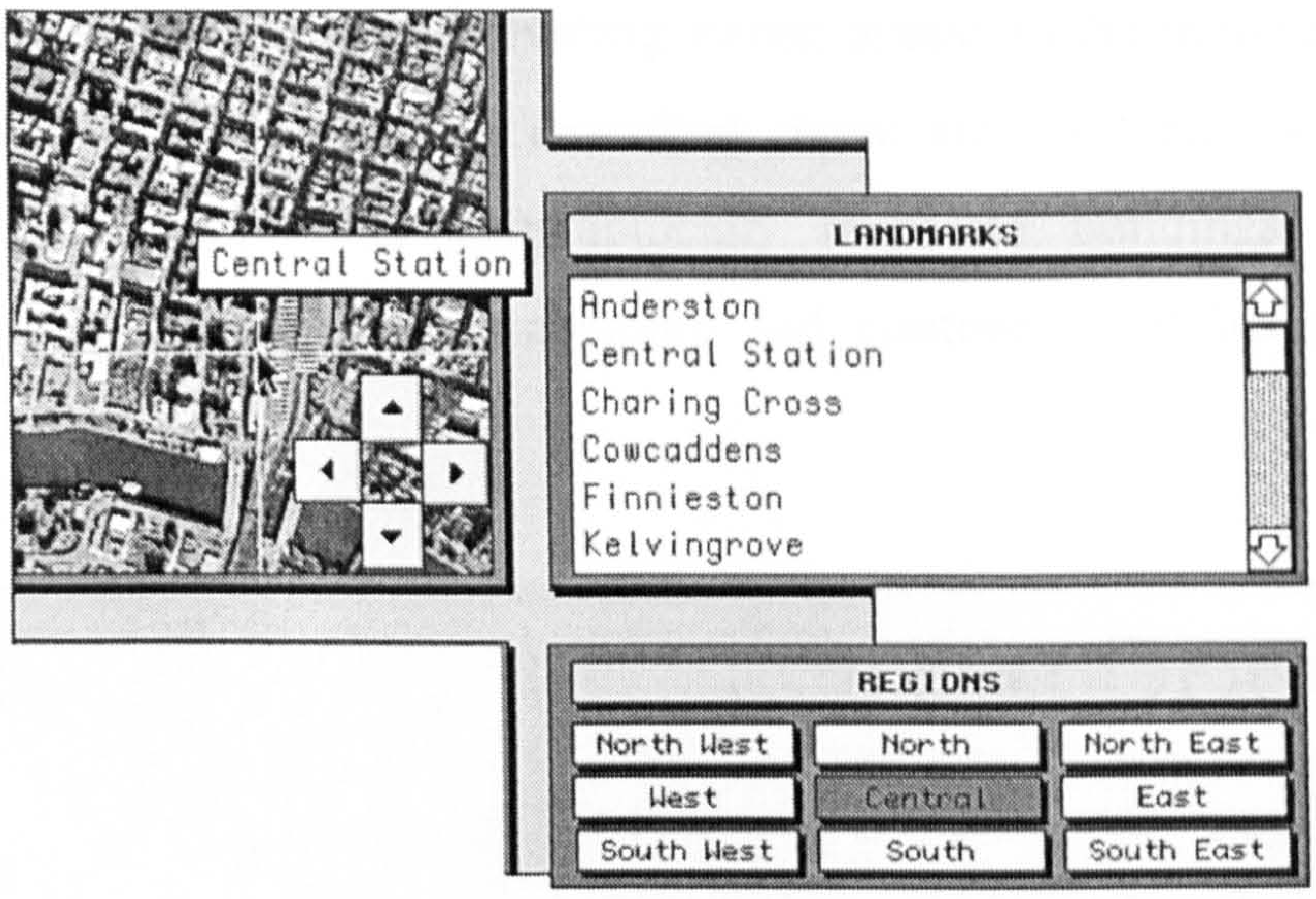


Figure 4.31 Navigation within the hierarchy of maps

The cross-hairs provide feedback as the mouse is moved over the area map. This position is reflected in the panel at the top right of the interface which shows the same view at four times magnification. This level of enlargement reveals details of individual buildings, houses and streets and is an information source in its own right. At a final level of detail, information is accessed from a text menu – geographically linked to the area maps – which when selected indicates the chosen location on the map and presents an image of the building along with a textual description.

4.6.3 Glasgow: Virtual Open Doors

Glasgow's Doors Open Day is an annual event allowing access to a selection of some of the cities most magnificent, colourful and unusual interiors, many of which are normally closed to the general public. This event is enormously popular and although a great success, the lack of time (one weekend) in which to visit the ninety-three available venues often results in a lost opportunity or rushed experience. (Grant 1998)

ABACUS produced a CD featuring a selection of the more interesting buildings, utilising photographic panoramas accessed through an index linking descriptive text to an image database. By providing virtual access to the highlights of these buildings most of the constraints described above are resolved. The narrative supplies context, access to geographically disparate buildings is next to instantaneous and the ability to compare and contrast amplifies the learning experience.

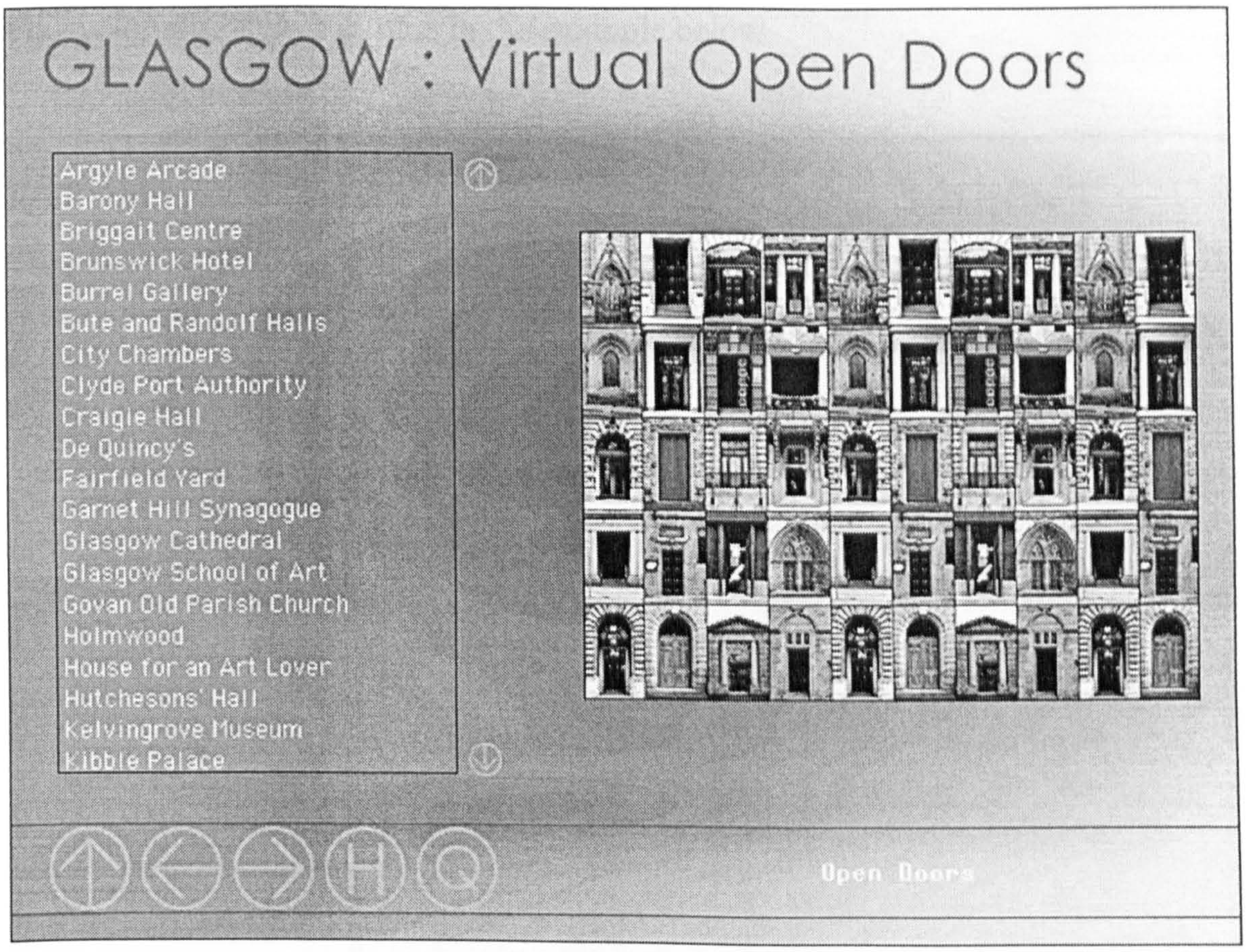
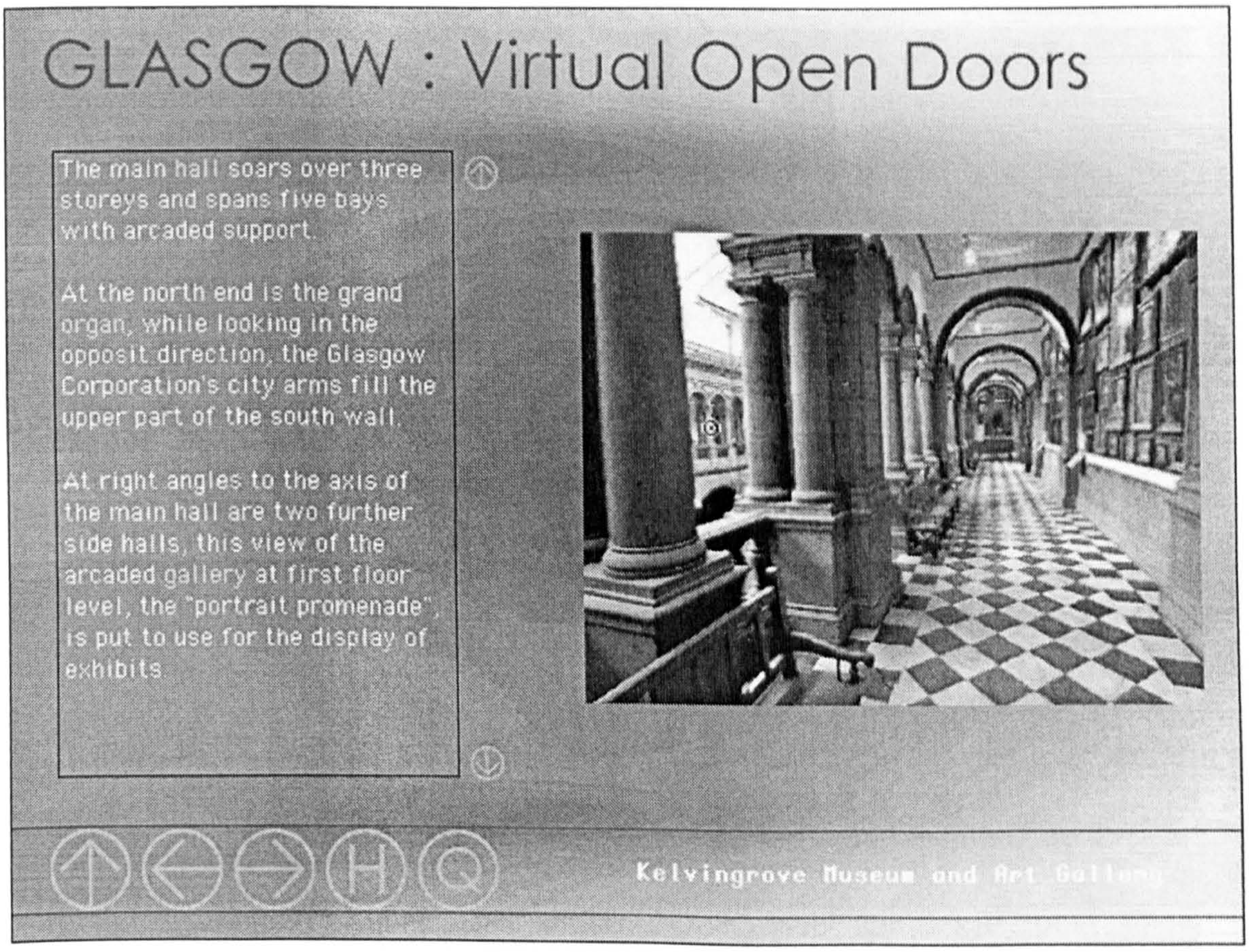


Figure 4.32 Glasgow Open Doors CDRom

The visual format of the Open Doors CD is designed to be graphically simply. All the information is presented on a single page within which one third is given over to the textual description of the building and the remaining two thirds to the imagery. Navigational controls are provided at the bottom left of the screen and are depicted as arrows, for moving through the content and two further icons accessing the help function and also the ubiquitous “quit”.

A scrolling index allows each building to be selected inat random or the user can move sequentially through the list. On selection the index is replaced by the description of the chosen building. This description follows a common format, giving first the address, date of construction and architect followed in turn by a general description of the building before detailing the interiors. Each description is supported by up to ten illustrations, including interactive panoramas, accessed via hyperlinks, shown in blue in the example below.



The main hall soars over three storeys and spans five bays with arcaded support.

At the north end is the grand organ, while looking in the opposit direction, the Glasgow Corporation's city arms fill the upper part of the south wall.

At right angles to the axis of the main hall are two further side halls, this view of the arcaded gallery at first floor level, the "portrait promenade", is put to use for the display of exhibits.

Figure 4.33 Format of the Open Doors CD.

Figure 4.34 The kitchen in the Tenement House.

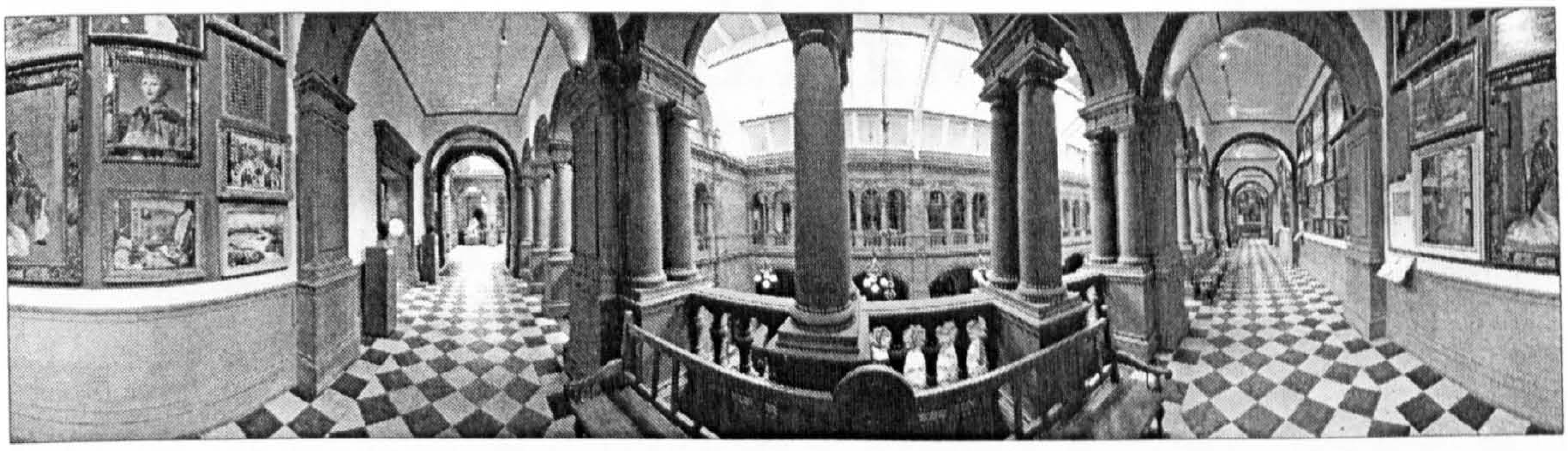


Figure 4.35 The Portrait Promenade, Kelvingrove Museum.

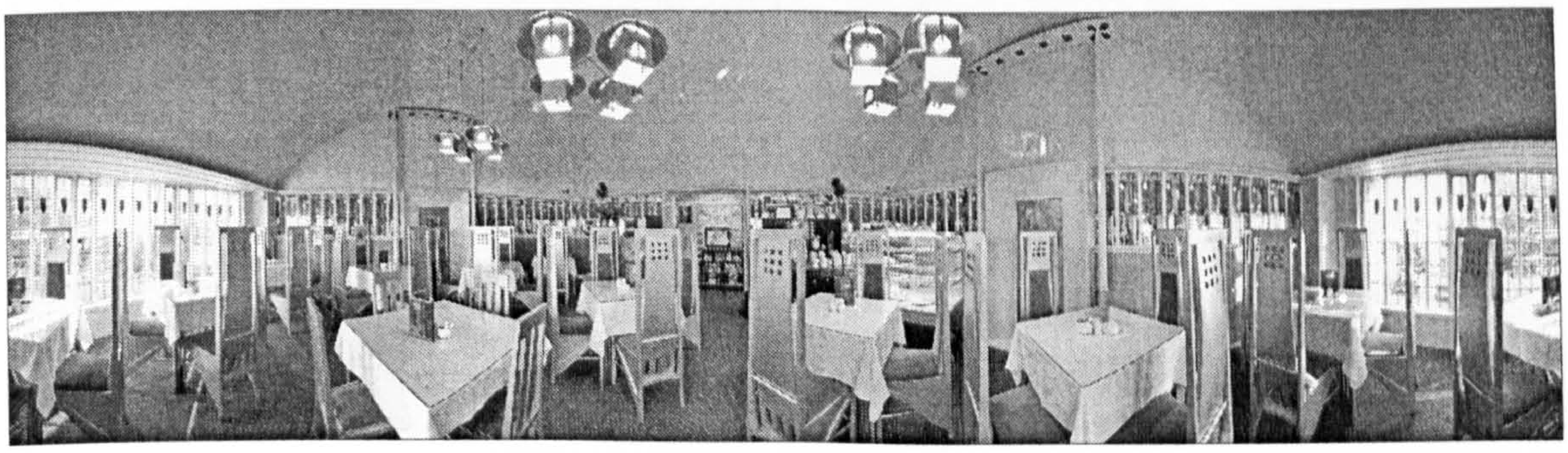


Figure 4.36 The Willow Tea Rooms, Sauchiehall Street

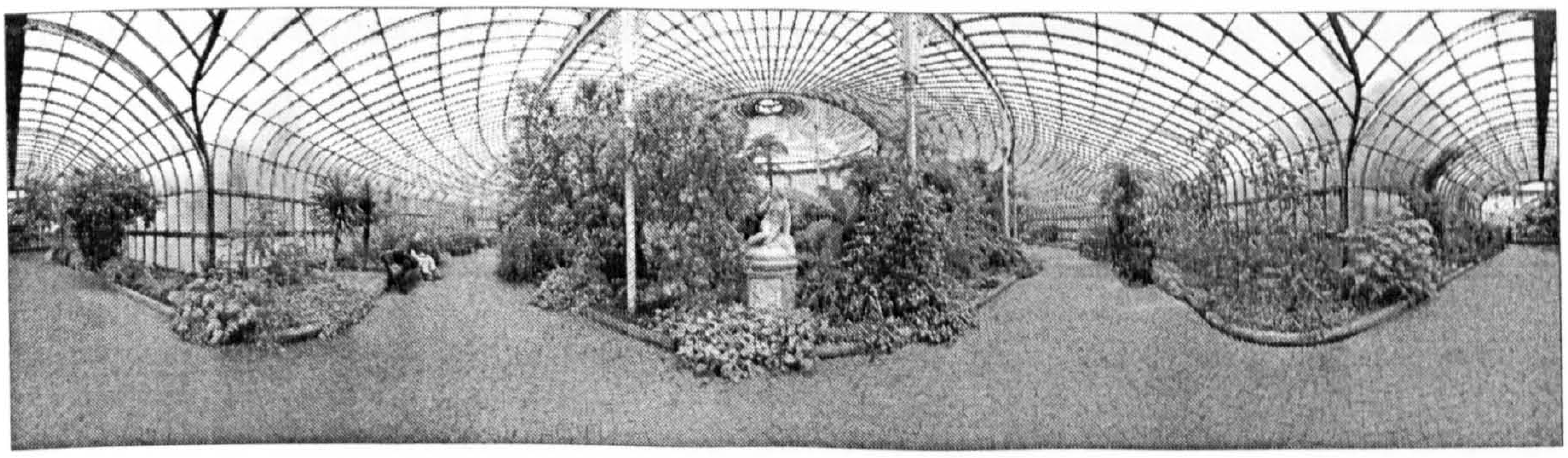


Figure 4.37 Kibble Palace, Botanic Gardens

4.6.4 The Tenement Story

The Tenement Story is about Scotland's flats, urban and rural, from their medieval beginnings to the present day, with the buildings as a starting point. It is a story of grim contrasts and pleasant surprises. It is about the place of Scotland in a mainstream European tradition. Above all it is about how people live.

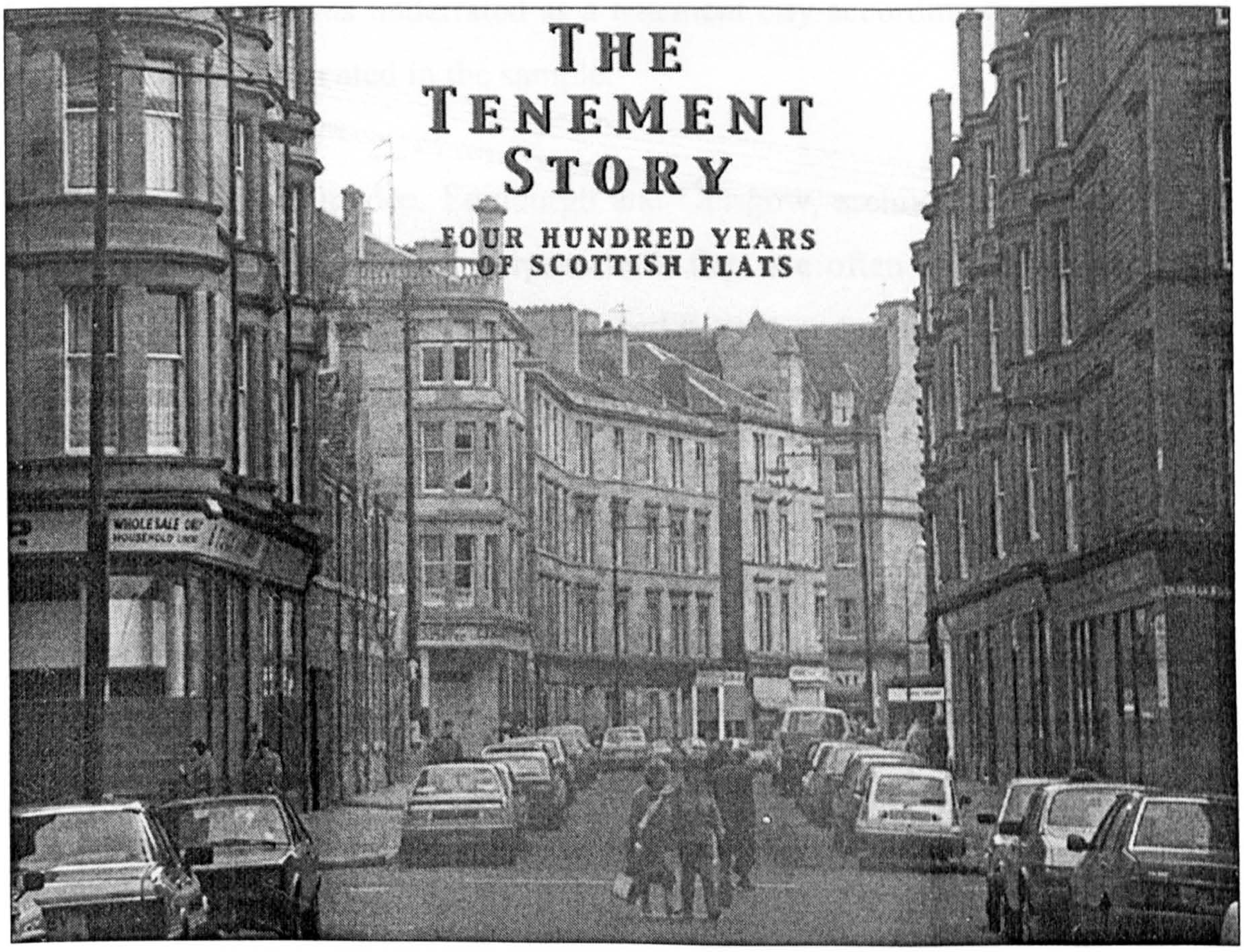


Figure 4.38 Four hundred years of Scottish flats

The archive comprises 140 examples of flats taken from all over Scotland. It is based on a collection of more than 500 flatted buildings assembled between 1982 and 1986 by Peter Robinson as part of a PhD thesis on flats, "Aspects of a Scottish Flat Tradition", at the University of Strathclyde. (Robinson 1986) The intention was to trace the development of the Scottish flat from the first hand evidence of the buildings themselves. The record has been updated by ABACUS for SCRAN with the help of Scottish Homes.

The core of the research study was a survey of a representative sample of flatted buildings, primarily from Scotland, assembled from scratch to create a permanent archive. The intention was to study the form of purpose-built flats by going back to the evidence of the buildings themselves.

The study showed that tenement developments tended to be largest in Glasgow, but that Edinburgh was underrated as a tenement city according to the range and quality of flats represented in the sample.

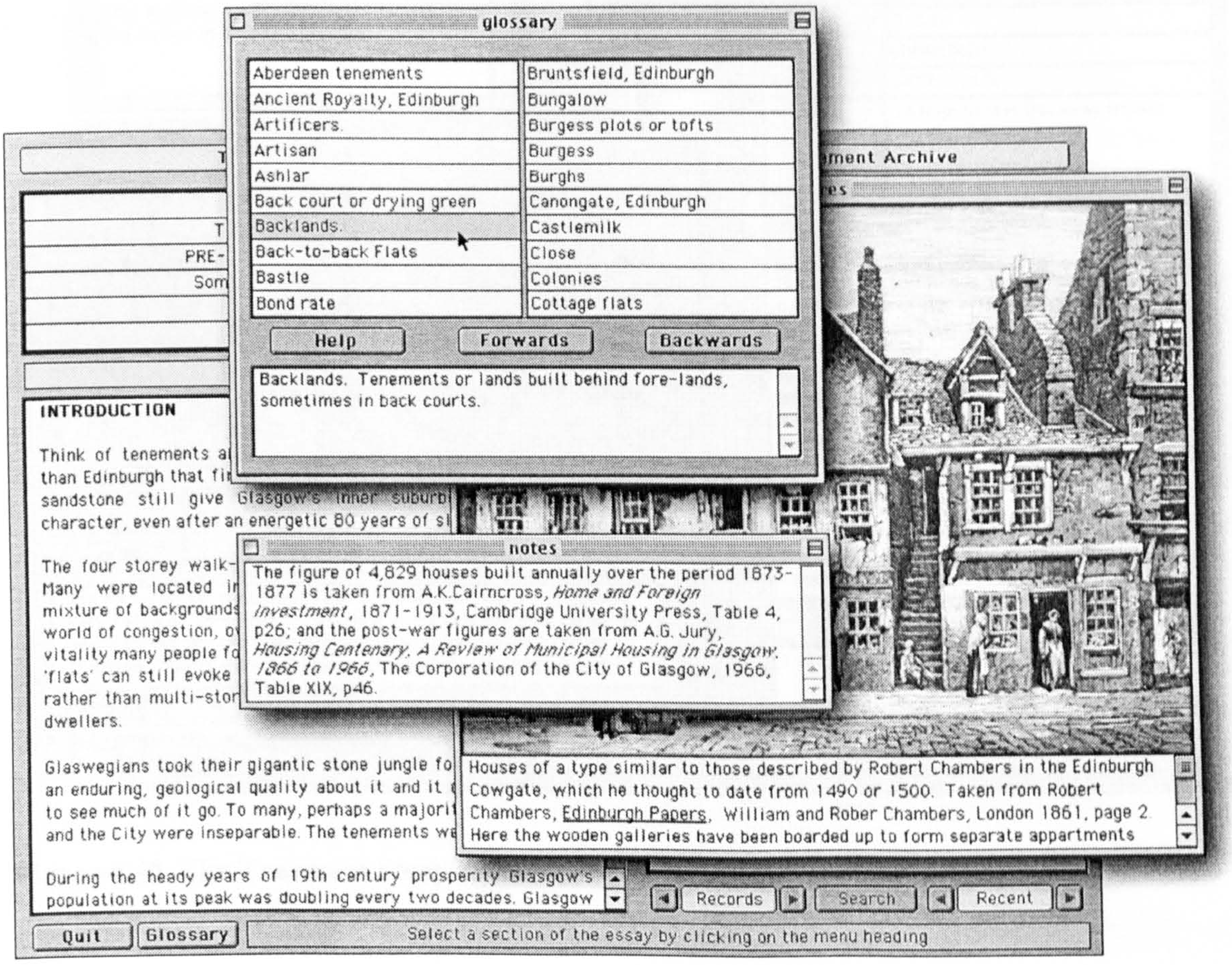
Outside Aberdeen, Dundee, Edinburgh and Glasgow, archive entries were more varied in size and treatment. Purpose built flats are often hidden from view in what look like two-storey cottages, terraced houses and villas. In the mid 1980s survivors tended to be later in date and more artisan in quality outside the cities and particularly so in the islands.

Original material was supplied from many sources; building control records, plans prepared by architects for improvement schemes, archives and first-hand observation. The study was based on single structures, that is a single building constructed at one time that may contain many flats. Each building was photographed where possible and some were measured.

The multimedia essay differs from the norm in that the bulk of the contextual material is textual, only being supported by imagery. The multimedia document is split into two functional halves, each cross referenced and mutually supportive. The tenement story occupies the left hand portion of the screen and the archive records, the right. Supporting information is presented within pop-up windows activated from hyperlinks or button actions.

The Tenement Story is a substantial body of text, equating to some forty pages of A4 script, this is broken down into thirty-one topics each accessed from a scrolling menu allowing a style of non linear navigation. Within each section the content is supported by subsidiary information access from hyperlinks within the text. As seen in Figure 4.39 this gives access to captioned images, notes and a glossary explaining commonly used terms.

Where specific examples are cited within the text based component of the essay these are cross-referenced to the archive.



- Aberdeen tenements
- Ancient Royalty, Edinburgh
- Artificers.
- Artisan
- Ashlar
- Back court or drying green
- Backlands.
- Back-to-back Flats
- Bastle
- Bond rate
- Bruntsfield, Edinburgh
- Bungalow
- Burgess plots or tofts
- Burgess
- Burghs
- Canongate, Edinburgh
- Castlemilk
- Close
- Colonies
- Cottage flats

Backlands. Tenements or lands built behind fore-lands, sometimes in back courts.

INTRODUCTION

Think of tenements a
than Edinburgh that fi
sandstone still give Glasgow's inner suburb
character, even after an energetic 80 years of sl

The four storey walk-
Many were located i
mixture of backgrounds
world of congestion, o
vitality many people f
'flats' can still evoke
rather than multi-stor
dwellers.

Glaswegians took their gigantic stone jungle fo
an enduring, geological quality about it and it
to see much of it go. To many, perhaps a majori
and the City were inseparable. The tenements w

During the heady years of 19th century prosperity Glasgow's population at its peak was doubling every two decades. Glasgow

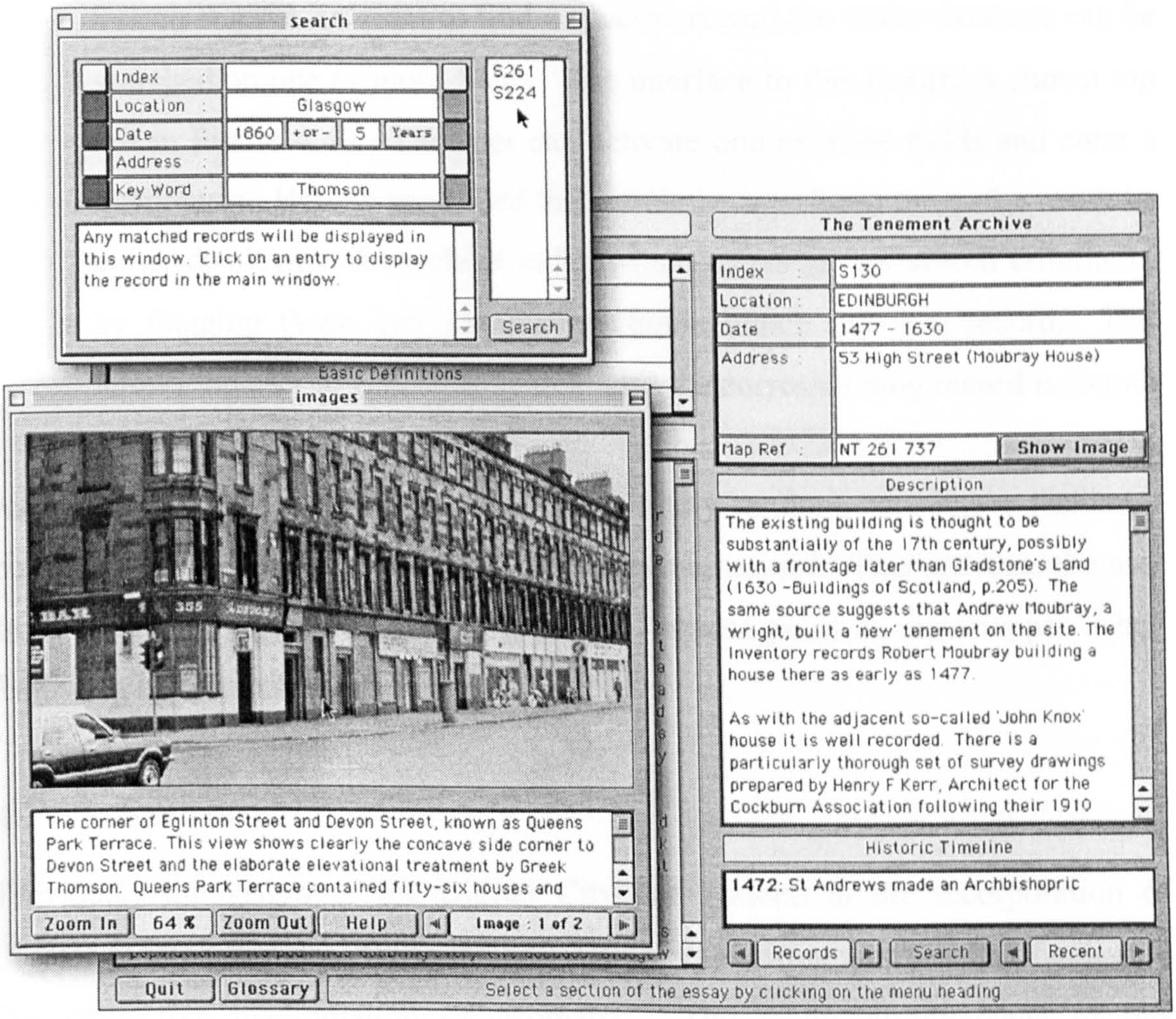
The figure of 4,829 houses built annually over the period 1873-1877 is taken from A.K.Cairncross, *Home and Foreign Investment*, 1871-1913, Cambridge University Press, Table 4, p26; and the post-war figures are taken from A.G. Jury, *Housing Centenary, A Review of Municipal Housing in Glasgow, 1866 to 1966*, The Corporation of the City of Glasgow, 1966, Table XIX, p46.

Houses of a type similar to those described by Robert Chambers in the Edinburgh Cowgate, which he thought to date from 1490 or 1500. Taken from Robert Chambers, Edinburgh Papers, William and Rober Chambers, London 1861, page 2. Here the wooden galleries have been boarded up to form separate appartments

Select a section of the essay by clicking on the menu heading

Figure 4.39 The Tenement Story

The archive contains key examples of 140 tenements from throughout Scotland, the data being formatted into three sections. The first part is comprised of a general architectural description of the building, this in turn is supported by a series of fields containing the archive index, geographical location, construction date, address and a map reference. Supporting the descriptive information are a series of captioned images displaying a photograph and where available, plans, sections and elevations. Where more than one image is used to illustrate a record, forward and back buttons are provided to page through the selection.



search

Index :
Location : Glasgow
Date : 1860 + or - 5 Years
Address :
Key Word : Thomson

Any matched records will be displayed in this window. Click on an entry to display the record in the main window.

S261
S224

images

The corner of Eglinton Street and Devon Street, known as Queens Park Terrace. This view shows clearly the concave side corner to Devon Street and the elaborate elevational treatment by Greek Thomson. Queens Park Terrace contained fifty-six houses and

Image : 1 of 2

The Tenement Archive

Index : S130
Location : EDINBURGH
Date : 1477 - 1630
Address : 53 High Street (Moubray House)
Map Ref : NT 261 737

Description

The existing building is thought to be substantially of the 17th century, possibly with a frontage later than Gladstone's Land (1630 -Buildings of Scotland, p.205). The same source suggests that Andrew Moubray, a wright, built a 'new' tenement on the site. The Inventory records Robert Moubray building a house there as early as 1477.

As with the adjacent so-called 'John Knox' house it is well recorded. There is a particularly thorough set of survey drawings prepared by Henry F Kerr, Architect for the Cockburn Association following their 1910

Historic Timeline

1472: St Andrews made an Archbishopric

Select a section of the essay by clicking on the menu heading

Figure 4.40 The Tenement Story archive

Access to this data is obtained in by a number of methods:

- Sequential access: all records are ordered in chronological order. The database can be browsed linearly by calling up each record in turn from the interface

- Cross referencing: where a particular building is cited from within the essay a hypertext link reveals the relevant archive entry. If any particular archive record relates to a further entry then these too are cross-referenced via a hyperlink.

- Search engine: in order to find a specific record the entire database can be searched on one or more fields. The interface to this facility is shown top left in Figure 4.40. The user can activate one or more fields and enter a query term. Having processed the search the interface returns the result by either displaying the archive index which meets all the search criteria, or by flagging those individual query terms which match a record. This allows the user to refine the search until the corresponding record is found.

As each record may be accessed in a potentially random, non-linear fashion a buffer of the ten most recent records visited is maintained. This prevents the user from becoming separated from their original access point as may happen when following hyperlinks.

4.7 Summary

Previously the vision of the Digital City had resided in the incorporation of diverse media types describing the city within an environment that had facilitated the interactive manipulation of geometry. The major disadvantage of this approach was with the computing infrastructure within which it had been conceived. The lack of any inbuilt high level functionality meant that every feature of the system had to be laboriously constructed from primitive elements. In turn this made for slow development cycles and a limited audience restricted to a similar computing system. The arrival of that strand of technology broadly labelled multimedia has the potential to provide a very different view of the problem and proved fundamental to the accelerated pace of development in terms of consumer hardware and software products.

The key asset introduced by this new technology was the fact that it provided a platform that was essentially “good at everything”. Commercial software developed to exploit these features provided the developer with unparalleled freedom to explore the creative opportunities offered by an interface that was limited only by their imagination. The adoption of metaphors out with the constraints of standard interface idioms made for a product that could be directed at a different audience, one which also possessed suitable hardware.

The lack of traditional 3-D functionality was offset by the ability to replace geometry rendering with digital video and 2-D animation techniques. Far from proving restrictive this offered diversity in the character and format of the applications and provided new ways of capturing the context of the city. The inclusion of different graphical formats and the ease with which a homogeneous package could be built allowed greater insights through the conjunction and cross referencing of information which would have been difficult if not impossible within a conventional computing system.

The case studies reported here are a small sample within of a range of application that can all lay claim to reflecting some aspects of the desired format and functionality of the Digital City. This could take the form of the diversionary narrative structure within the Day the Sun Came Out or the inclusion of interactive elements as shown in the Open Doors CD or the historical context of the Tenement Story with its integrated illustrated database.

Chapter 5 Back to the Future

An intriguing facet of the development of the Digital City is the manner in which applicable technologies get re-discovered or re-invented. The growth of the Internet has provided many opportunities not just as a medium but also in terms of access to knowledge, techniques and software products. The ready availability of these tools has allowed the rapid development of applications that would otherwise have taken many man-years to realise. This economy has made it possible to re-visit the Glasgow model, the dissemination of knowledge has improved the tools used to construct the data-set and the availability of a common denominator in terms of the delivery medium has made the product available to all. The final factor that has enabled the uptake of 3-D applications has been the increased availability of consumer grade graphics systems bringing affordable high level performance to almost every machine specification. It is interesting to observe that in the space of ten years what was once the pinnacle of graphical development has become the lowest common denominator in an era of new technology.

5.1 The Internet Context

In the late 1980s CERN, the European Organisation for Nuclear Research, encountered a logistical problem concerning the correlation and dissemination of information between geographically disparate members of the research community. An increasing emphasis on collaboration between laboratories throughout the world and the rapid recruitment of new staff meant that access to existing documentation and the ability to integrate new material quickly gained a high priority. At CERN Tim Berners-Lee created a proposal for a hypertext document system that could be used to enhance communications within the High Energy Physics community. This proposal laid the foundations for what was to become the World Wide Web. By 1990 the prototype was in service and although implemented on a single platform it incorporated several key components on which all future developments would be based.

Berners-Lee defines

> *"several important components necessary to realise the vision and which, in a nutshell, define the nature of the WWW today" as:*

- It must be cross-platform.
- Must be able to use existing informational resource systems while also allowing new information to be easily added.
- A transport mechanism was necessary to move documents across networks. [evolved into http]
- An identification scheme for addressing both local and remote hypertext documents. [evolved into URL addressing]
- A formatting language for the hypertext documents. This was not explicitly mentioned but was part and parcel of presenting the information received. [evolved into HTML]

In 1991 the specification was published in a number of Usenet special interest groups and was readily adopted by the Internet community. With a growing pool of common code and the availability of browsers for a variety of hardware and software platforms the amount of information published on the Internet started to grow exponentially as did the number of people wishing to access it.

5.1.1 Standardisation

With the rapid adoption of the core functionality of the WWW there was an inevitable degree of deviation from the original specifications. This was in part due to the commercial interests adding "extensions" to the mainstream functionality in order to differentiate themselves from competitors. In order to prevent the emergence of proprietary systems the decision was made to establish an open standard for all to use. This is the prime goal of the International World-Wide Web consortium, "W3C", a body of institutes and companies from all over the world. The W3C is run jointly by INRIA, the Institute National Pour la

Recherche en Informatique et en Automatique (for Europe), MIT, the Massachusetts Institute of Technology (for the USA) and Keio University (for Asia).

5.1.2 Browsers

The original browser was named Mosaic, produced by the National Centre for Supercomputing Applications, NCSA. Originally released in 1993 this software did much to promote the popularity of the early WWW. Development continued until 1997 although many features of the later HTML specifications were never implemented. Once the commercial potential for web browsing software was realised the process of innovation proceeded at an increasing pace. The early developments at the University of Illinois can lay claim to being the basis for both the future market leaders in this area. In 1994 the software engineers responsible for the Mosaic project at NSCA set up an external company, Mosaic Communications, later to become Netscape which was to hold, for a time, the largest share of the browser market.

The next year Microsoft launch their own proprietary browser, Internet Explorer, based on code licensed from NCSA. In the following years intense competition between Explorer and Netscape resulted in a stream of innovations as each competed for market share. Among these features were the integration of server and client side scripting and the inclusion of plug-in architectures. These features provide key functionality in the process of implementing the Digital City on the Internet. The race to occupy market dominance has undoubtedly forced the pace of development yet it has also introduced deviations from the standard, a situation which tends to negate many of the advantages.

5.1.3 Scripting

Both of the main stream browsers implement scripting using JavaScript. This is a Netscape technology that implements a cross-platform, object oriented scripting language for both client and server side applications. JavaScript is a simple interpreted programming language that can be used to enhance the functionality of HTML and not only produce more dynamic output and process user input but also provide the means of communicating between the extensible capabilities of the plug in architecture and the core capabilities of the browser. (Netscape 2002)

5.1.4 VRML

At the first international conference on the WWW Berners-Lee invited Mark Pesce to present a paper on the concept of a three dimensional interface to the Web. Following the presentation the conference delegates concurred that there was a need for a standardised language for specifying 3-D worlds. The ongoing discussion as to the specification was facilitated by an electronic mailing list which attracted wide spread support. After much deliberation a sub-set of the Open Inventor ASCII file format was chosen as the basis for VRML. Inventor was a CAD oriented scene graph description with support for geometry, materials, lighting and viewing parameters. (VRML 2002a).

The preliminary specification was deliberately designed to meet the following requirements:

- Platform independence
- extendibility
- suitability for low band width connections
- independence from HTML
- streamlined design and implementation

The VRML Architecture Group, who were charged with overseeing the initial implementation, managed to avoid a specification that appeared to be "designed by a committee". The initial goal was to take the concept to the open market without burdening future developers with an overly complicated syntax or features that were deemed unnecessary in a first incarnation. To achieve these aims development was concentrated on defining the structure and content of the scene graph and issues such as behaviours were set aside for future consideration. The key functionality embodied with VRML out with that of a normal scene graph, was the ability to define a set of "inline" objects that may exist on separate servers and the addition of hyperlinks allowing an interactive dialogue from within the 3-D world.

The first draft of VRML 1.0 was published in November of 1994 and after review was formalised as a specification in May of 1995. However it was always the intention that VRML should be continually developed and over the next year an enhance specification was raised and officially published in August 1996. This was followed by its homologation as an ISO/IEC international standard later that same year. (VRML 2002b).

VRML 2.0 or VRML 1997 differs from its predecessor by the addition of a range of enhancements and additional features. To this end the specification was extended to nodes which represented not only aesthetic features but also relate to physical attributes, such as gravity and collision detection, and communications routing. The main difference to core functionality is through the addition of sensors which generate events corresponding to user interaction and scripts which can process incoming events and generate outgoing ones. These and other additions were intended to progress the VRML standard from its initial static representation of the "world" into one capable of dynamic interaction with the user. (VRML 2002c)

5.2 The Glasgow Model Revisited

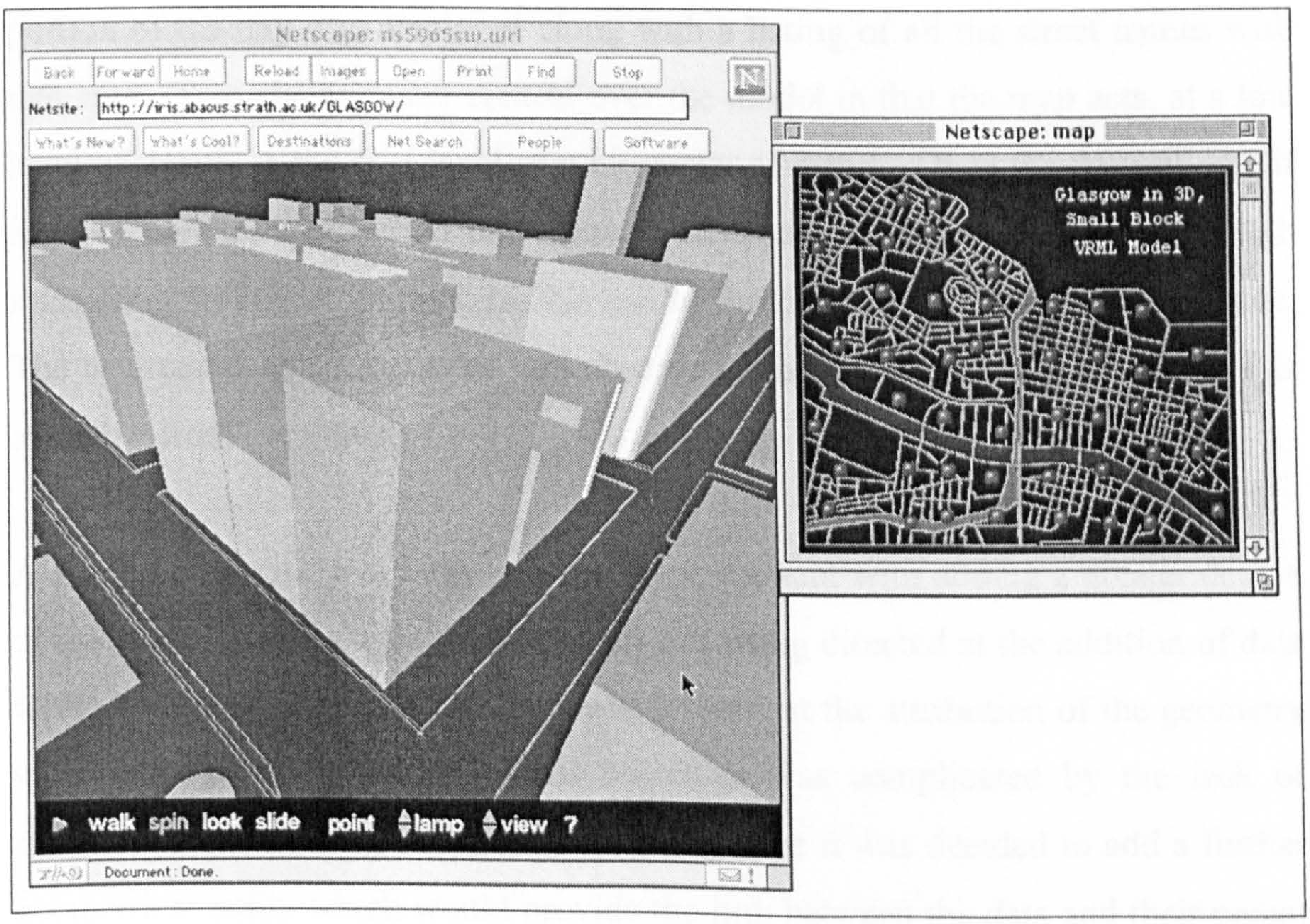
Despite the passage of over 10 years the original Glasgow model was still in existence although no further development had taken place. Despite this the model still proved useful and was continually in demand serving the same sort of functionality as it always had. Unfortunately it also still suffer all the limitations that were present over a decade ago. These shortcomings were well understood and can be enumerated as below.

- The hardware and software combination required to exploit the model was proprietary to ABACUS.
- Even with the scale of technology at the disposal of a research group the level of interactivity with the model was limited
- The interface to the model had been developed but never to the extent that made it useful to a casual user
- The ability to attribute detailed data within the model to increase its legibility had never been performed
- The integration of disparate databases, originally a prime goal, had always proved too demanding.

5.2.1 Initial Developments

With the introduction of VRML1 in November 1995 the sharing of 3-D models over the Internet became a reality. The arrival of VRML as a standard made the issue of application specific hardware and software a thing of the past. The entire philosophy behind the WWW is directed at platform independence by positioning browsers to form an abstraction layer over the native hardware.

In order to experiment with this new technology ABACUS looked for an existing source of data that was suitable for conversion to the VRML format. The data had to simple because despite the advantages of the concept the model still has to be downloaded over a relatively slow connection and the graphical processing power of the average client was still very limited. The VRML geometry format was not dissimilar to the original VIEWER syntax and a conversion program was easily written. The first exemplar of this incarnation of the model was a direct re-implementation of the original model interface – even to reusing the original map graphic. The example shown below was running in the original version of Netscape using the "Live3D" plugin on an 80-MHz Macintosh but the browser software was available for PCs and workstations. This version was constrained by the standard screen resolution of 640 x 480 pixels that allowed little opportunity to create an integrate layout. Ideally there should have been some integration between the 2-D and 3-D components but the limited screen area dictated that the model was displayed in a separate window from the interface.



Netsite: http://iris.abacus.strath.ac.uk/GLASGOW/

Document: Done.

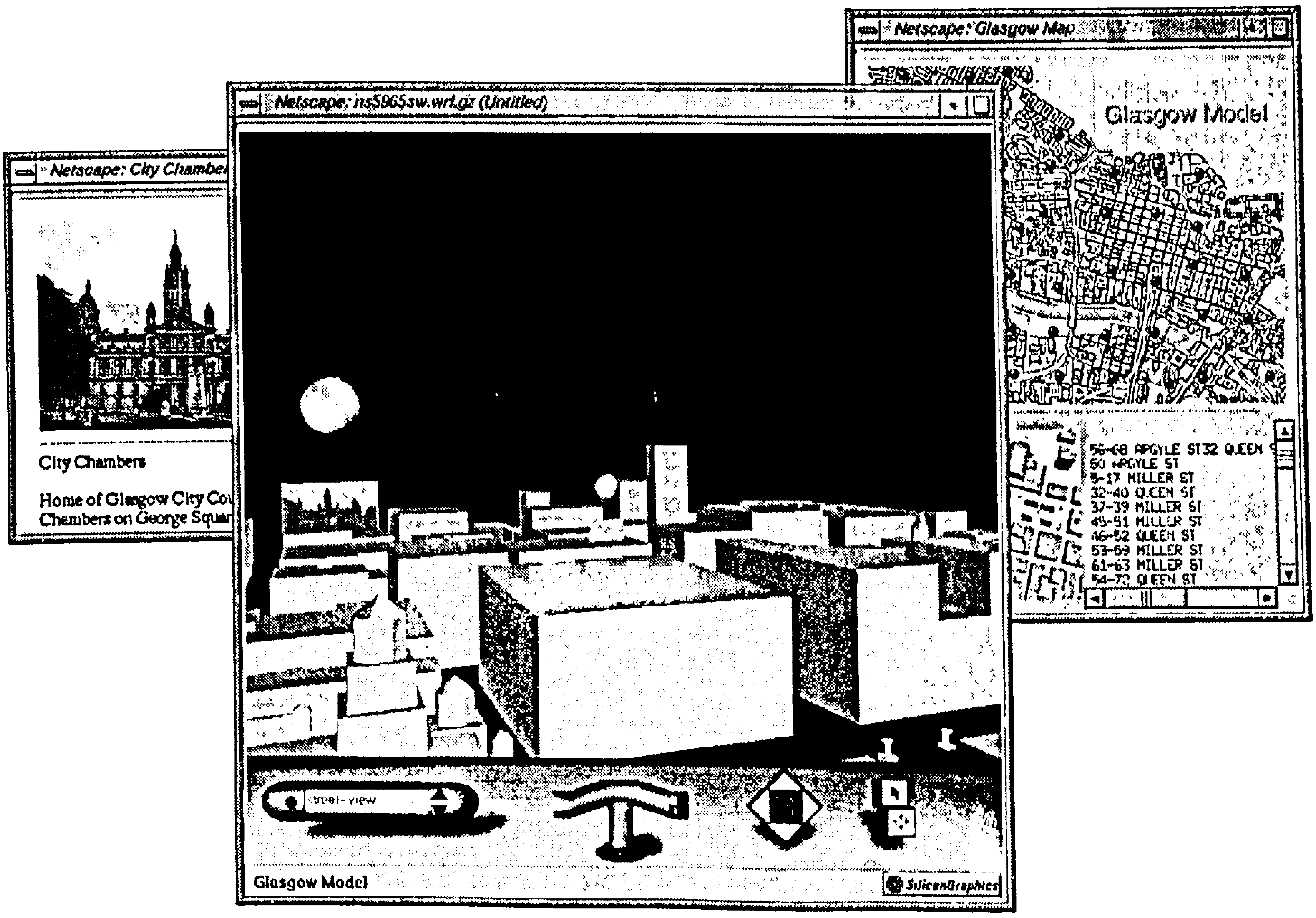
Figure 5.1 The first on-line Glasgow Interface

Despite these limitations the performance and utility of this new direction was unparalleled, anyone with access to the internet could download the model and interact with it to the same degree of functionality that had previously required the power of a £15,000 graphical workstation. The graphical performance might not have been quite the same but since the major bottleneck was related to the speed of the Internet this no longer mattered to the same extent.

The choice of which area of the model to download was made be selecting the relevant portion from the bullet points on the map. While this served to break the model down into sections that were suitably sized with regard to the time taken to complete the download it was difficult for a user unfamiliar with the layout of the real city to decide which area was required. A second version of the map interface was constructed that provided more feedback. Figure 5.2 This was made so as to allow the user to browse the map by creating a layer of indirection before the download of data commenced. Once an area was chosen a plan view of that portion of the city was displayed along with a listing of all the street names with that area. This gives greater control over the model in that the map acts, at a low level of detail, as the first level of a hierarchy of detail. Once the desired area if found then a further button press downloads that part of the model. Although interesting within its limitations, this use of the model was a passive experience. The user could download areas and observe them but no further information was available from the system.

A second prototype was built in order to experiment with adding a greater degree of interactivity and interest to the model this being directed at the addition of data to the buildings. Previously it had been found that the attribution of the geometry with data regarding the individual buildings was complicated by the lack of structure inherent within the model. At this stage it was decided to add a further geometrical entity which would provide the link between the data and their parent building. This took the form of an iconic marker – a balloon – pinned to the relevant building. The implementation of this functionality had been seen as a significant hurdle to realising the use of the model as the underlying structure of a more integrated information system. However VRML brought two of the required elements in the form of the node types WWWinline and WWWanchor. An inline node allowed the integration of geometry within a separate structure from the rest of the model, this additional geometry being identified and accessed by a specified

URL. The anchor provided an encapsulated mechanism for activating a picking function that in turn loads the contents of the specified URL into the designated target. These two features made it possible to create a series of icons identifying buildings that, when activated, would load a description of the building into a further window where it was identified by an image and accompanying text.. Each building of interest was indexed by storing a viewpoint in the VRML file structure that would orientate the view with that building readily visible. A similar construct was used to add detail to the street views. A viewpoint would locate the user within the streetscape and a hyperlink anchored onto a street sign would display a photographic view and pertinent details.


City Chambers

Home of Glasgow City Co
Chambers on George Squar

Figure 5.2 The second on-line Glasgow Interface

A notable deficiency with this version of the city was the lack any of physical constraint it provided to a user's navigation. Without the confinement imposed by gravity and collision detection there can be no sense of the "right way up" and buildings are transparent to the user's progress. This diminishes the recognition of the model as a representation of a city and can lead to it being perceived as just an object in space.

The release of the VRML2 standard in 1997 provided the solution to both these problems as well as offering many more enhancements that could potentially be exploited. However, in order to capitalise on these new features it was imperative to provide the city with a ground model but over the period since the original model had been constructed this part of the data had been lost. The task was then to recreate the DTM to the original standard. This was complicated by the fact that the buildings, which had all been matched to the terrain as described previously, had to align with the recreated DTM. In order to guarantee a match the new DTM had to be derived from not only the standard contour information but also the heighted footprint of all the buildings and general line detail. Without the original software a replacement algorithm had to be found. This illustrated another benefit of the Internet in that global access to freely available information was unprecedented. In the preceding era ownership of key algorithms such as this would have been the treasured currency of established research groups, now this information is willingly shared. A search on the Internet reveals over 7,500 references to "triangular irregular networks" the basis for terrain construction and further investigation reveals thousands more instances of references to algorithms, complete software packages and numerous published methodologies. The result of this pooled knowledge provides all the required information in order to implement a DTM solution as described below.

5.3 Terrain Model Construction

The terrain modelling software first used in the early 1980s was functional but renowned for its instability and required much human intervention and modification of the input data to achieved the required result. The Edinburgh model was based on pre-constructed data obtained from the Ordnance Survey and required little modification to adapt to ABACUS formats. Now that a new requirement had arisen the task of selecting a suitable approach was made much easier by the wealth of published information available via the Internet.

5.3.1 Triangular Irregular Networks

The number of methods of forming a triangular irregular network are legion. Most of the common implementations fall under the banner headline of Delaunay triangulation. Delaunay triangulation does not itself refer to any one algorithmic approach but indicates a criterion for forming a well structured mesh. The history of this methodology goes back over a hundred years and owes its existence to a range of scientists who have enhanced and rediscovered its benefits over this period. The earliest reference to this widely used geometrical construction is accorded to G.L.Dirichlet but it is also commonly referred to as a Voronoi diagram and as a derivative of Thiessen polygons. A short list of scientists who have lent their name to the construct is shown below.

G.L.Dirichlet	mathematics	1850
M.G. Voronoi	mathematics	1908
Thiessen	meteorology	1911
Wigner Seitz	metallurgy	1933
Blum	bio-science	1967

Figure 5.3 Scientists associated with triangular constructs

The property in common with all these methods is that they split the 2-dimensional plane into a set of tiles each surrounding a single generating site. All other points within the tile are closer to the generating site than to any other. The Delaunay triangulation is then a superimposed construct that is created by connecting all generating sites that share a common edge, the triangle edges forming perpendicular bisectors of the tile edges. Such a triangulation exhibits many desirable features. It can be shown that the convex quadrilateral formed by two adjacent triangles with a common edge has a greater minimum internal angle than if the quadrilateral were formed in any other way. This leads on to the observation that the triangulation is unique, for non degenerate data, and within the network the triangles are as well formed as possible.

5.3.2 Radial Sweep Algorithm.

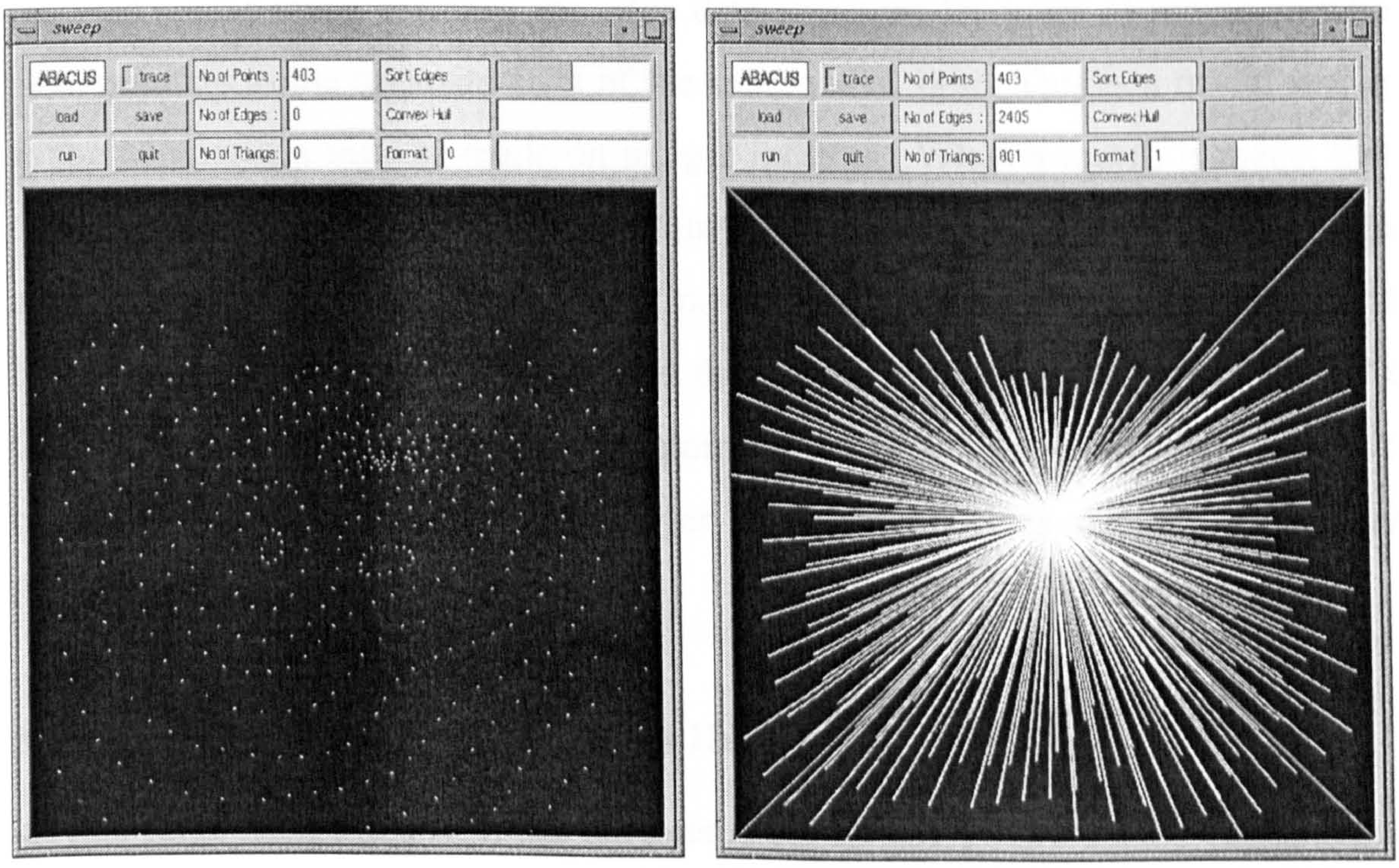
The Radial Sweep algorithm, also known as the sweepline algorithm, is one of many methods of constructing a triangular network from a set of arbitrary distributed points. While this approach is certainly not the fastest or the most fully featured solution to the problem it does possess a unique advantage in that the data structure and computational processes are so simple that for a small scale example the solution can be found by inspection. This factor makes explanation, implementation and de-bugging a much more simple task than some of the more esoteric methodologies, while still producing identical results. It does suffer from some disadvantages in that the execution time is relatively long, there being few opportunities for aggressive optimisation, and in large scale data sets a lack of floating point precision can lead to errors. The ease of implementation and the ability to take unstructured data as input lead to the choice of the radial sweep algorithm as the prefered option.

5.3.2.1 Input Data

The input required by the algorithm is any set of measurement sites. These may have originated as contour strings, randomly distributed survey data or as heighted building footprints or feature boundaries. The only requirement at this stage is that all input sites should belong to a similar co-ordinate system. As the data is read, each new site is compared to the preceding values. If a coincident point is found it will not be loaded as duplicate co-ordinates will defeat the logic of the algorithm. Once all the valid data is loaded, each site is tested to obtain the bounding box of the data set, from this, the geometrical centre of the point distribution can be calculated. Once the centre is found then the distance between each of the data items and the centre is compared, the shortest distance indicating the data point to be nominated as the hub of the radial sweep. This point is then shifted to the end of the input data array and the running total of participating sites is decremented. Each distance calculated is saved as a member of the structure containing the input data. This value is required later on to differentiate between input sites that have an equal bearing from the centre point.

5.3.2.2 Determination of Bearings

Radial line segments are constructed between the central point and all other data points. If the central point is considered as a local origin then the magnitude of the relative offsets in the x-axis and y-axis of each of the data points can be used to calculate the bearing of each of these line segments. The point structure array containing all the input sites can now be sorted in ascending order on the basis of the magnitude of their bearing. If two or more radial lines posses an identical bearing then a further pass is made to sort these sites on the basis of ascending distance from the central point.



No of Points : 403
No of Edges : 0
No of Triangs: 0
Format 0

No of Points : 403
No of Edges : 2405
No of Triangs: 801
Format 1

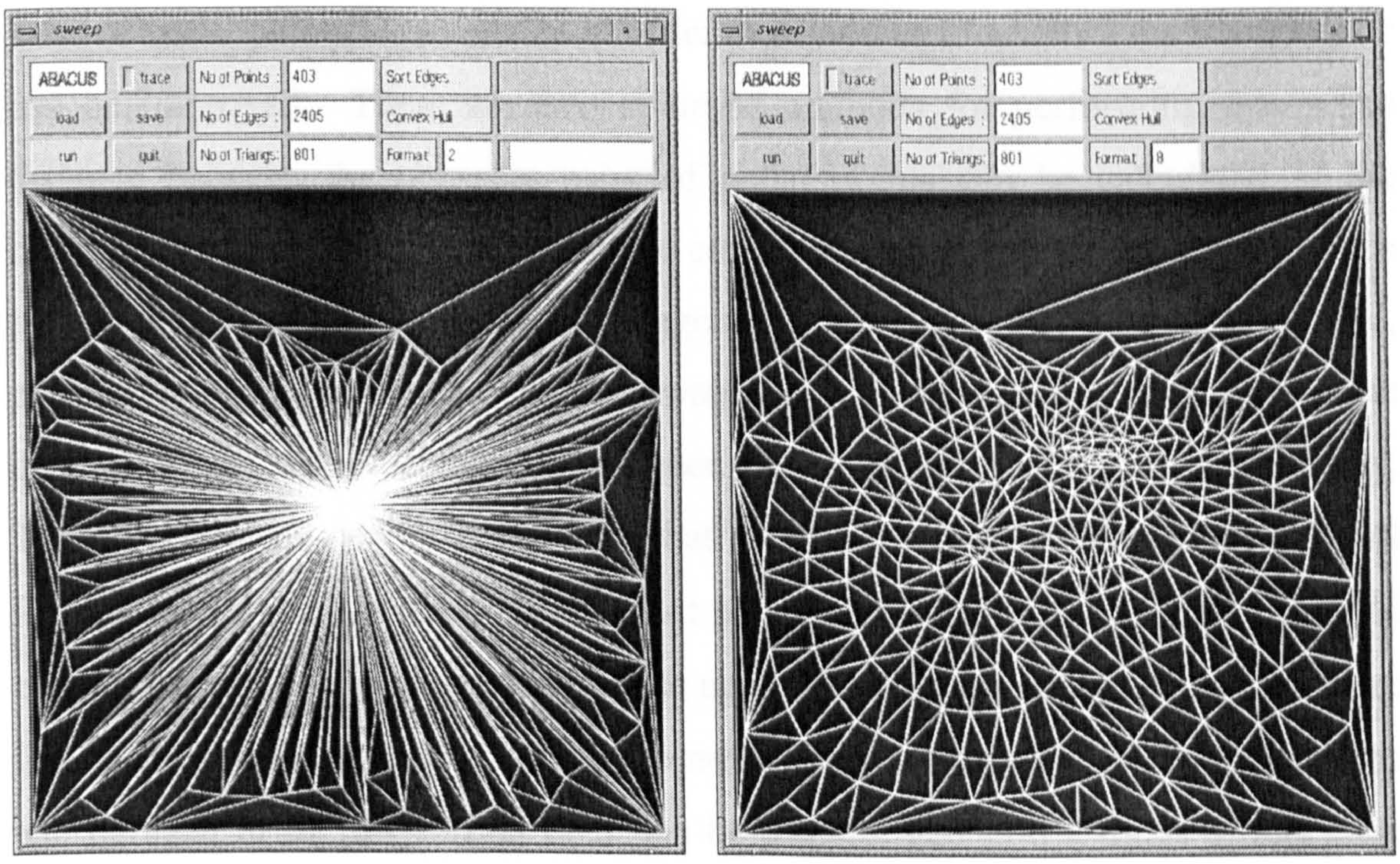
Figure 5.4 Point input data and calculation of initial bearings

5.3.2.3 Building the Initial Edge Data Structure.

At this stage of the calculation procedure only the array of structures containing the input data and relative bearing and distance exists. The task is now to populate the edge data structure and hence form the initial triangular network. This is achieved by looping through each participating member of the point data structure. A radiating line is projected from the central point to each point under consideration. A long thin triangle is formed by connecting this line, at its extremity, to the next point in the array. This point is then connected back to the central point, forming the initial triangle. Complications ensue when there are more than one data point sharing an identical bearing. When this condition is detected the index of each point on the identical bearing is pushed onto a stack. When the bearing eventually does change, each of these indexes can be popped off the stack, and since they were sorted in order of ascending distance from the centre, valid triangles can be made by connecting multiple points on the same bearing to the next point possessing a unique bearing. As each point is accessed pointers are added to the edge data structure that then describe a set of points, edges and unique triangles. To establish a triangular network a further pass is required. This stage searches the edge data structure for pairs of members who have pointers to the opposite ends of the same radial line. This is a trivial search as the comparison may be made on integer indexes. When a match is made, that member of the edge structure that contains a pointer to adjacency is updated with a pointer to the corresponding edge. If no match can be found then the edge must lie on the boundary of the network and the pointer is negated to reflect this. At the end of the primary sweep all data points have been triangulated by a network of non-overlapping triangles. However this network does not yet adequately describe the complete surface topology

5.3.2.4 Generating the Convex Hull

At this stage each triangle in the base set is comprised of two shared edges and a boundary edge. While each site is a member of the triangular network, the network itself is incomplete. Some sites might lie at a distance from the centre while its neighbouring triangle might be relatively closer. This forms deep concavities in the boundary of the network and represents one or more "missing" triangles which must now be formed. The edge data structure is search for pairs of edges that are both flagged as being on the boundary and who both point to a common site. These edge pairs are then tested to ascertain whether they will form a concave or convex angle. If the angle is concave then they will form a valid triangle with a positive signed area. If convex then the triangle so formed would be invalid as it would potentially overlap existing members of the network. Inclusion can be based on the sign of the planar area of the triangle to be formed or on the basis of the sign of the cross product of the vectors of the adjunct edges. If the new triangle is deemed to be valid then the edge array is grown and all relevant pointers updated to incorporate the new member. This procedure is iterated over until no new members are produced.



No of Points : 403
No of Edges : 2405
No of Triangs: 801
Format 2

No of Points : 403
No of Edges : 2405
No of Triangs: 801
Format 8

Figure 5.5 Building the convex hull and final data structure

5.3.2.5 Improving the Surface Topology

The network is now formed by a sheet of unique, non overlapping triangles that describe the convex hull of the data set. The greater majority of these triangle exhibit a poor aspect ratio. The aspect ratio of a triangle is represented by the ratio of its inner circle to its circumcircle. This is of relevance in a number of application areas. If the triangle is overly long and thin then it is unlikely to conform to desired surface features such as summits, ridges or breaklines as well as smaller scale surface discontinuities. Ideally the formation of near equilateral triangles will produce the most consistent results for all applications. This can be achieved by understanding that if two neighbouring triangles form a quadrilateral the topology of the network can be improved by recursively swapping the common, diagonal, edges until no further iterations can be performed.

Taking each edge in turn the quadrilateral formed by its neighbouring triangles can be tested to determine if they are a suitable candidate for recombining. If the quadrilateral so formed possesses a pair of edges that form a concave angle it is rejected, however if it passes this test then the relationship between the opposite pairs of nodal points is calculated in order to decide as to whether the triangle pair can be improved. There are two common methods for performing this. The distance between the opposite pairs of nodal points can be calculated, if the distance between the two points on the common edge is greater than the distance between the remaining nodes the diagonal is swapped. The other, preferred, method is to base the benefit of swapping the diagonal by calculating the difference in the included angle between the same pairs of sites as previously. The later method yields a network that provides a greater proportion of well formed triangle and reduces the amount of terracing artefacts along the edges of contour strings. This procedure is once again performed iteratively until no more improvements are possible. Once the network is complete the triangular mesh can be used in its raw format or it can be transformed into a mesh of any edge dimension by overlaying a grid. The elevation values at each grid intersection being determined by analysis of the plane equation of the underlying triangle.

5.4 Road networks

A fundamental problem, already discussed in relation to the Edinburgh model, relates to the differentiation of the terrain into areas of hard and soft landscape and differentiating between road surfaces and the generic terrain. This is particularly pertinent where there is little overall detail to provide perceptual cues as to the nature of the groundscape. If the user is to travel at a height consistent with a normal viewer's elevation then it is imperative to relieve the otherwise nondescript surroundings of grey walls and ground.

One solution can be sought by subdividing the ground into areas of differing material and colour. This is more easily said than done if the entire terrain is composed of a single entity in the form of a DTM matrix. This format is extremely efficient in terms of storage but does not offer any means of differentiation across its surface. In order to subdivide the terrain it must first be decomposed into an array of independent triangles from which roads can be subtracted via a series of Boolean operations. This allows the different surfaces to be attributed with distinct colour attributes. A further enhancement can be found through differentiating between levels by adding an upstand but as each of these interventions substantially increases the polygon count there is a finite limit to the amount of detail added in this manner.

A final concern is in regard to the redundancy of data within this representation. While the DTM matrix is a compact format it is still bound by the regular sample interval embodied within the concept. If the area under consideration contains regions that are flat or are of constant gradient then the format is less than optimal. As the sample density increases then this situation is exacerbated and if there is a need to represent the terrain as a series of independent triangles then the worst case condition is soon reached. This is a tradeoff between the requirements of the visual representation and the need for a compact data format. Given the gains in legibility offered by this technique the additional cost incurred by losing the advantage of economical file size is slight.

Figure 5.6 illustrates the advantages of differentiating the terrain. At top right is the image generated using just the buildings, leaving the viewer with no horizon or feedback as to spatial orientation. The inclusion of the ground model anchors the buildings to the ground but provides few visual cues. Delineating the transport network with line detail differentiates between roads and pavements but cannot contrast between hard and soft landscape. The final image shows the effect of separating these features into different geometry which solves most of these problems while generating little overhead.

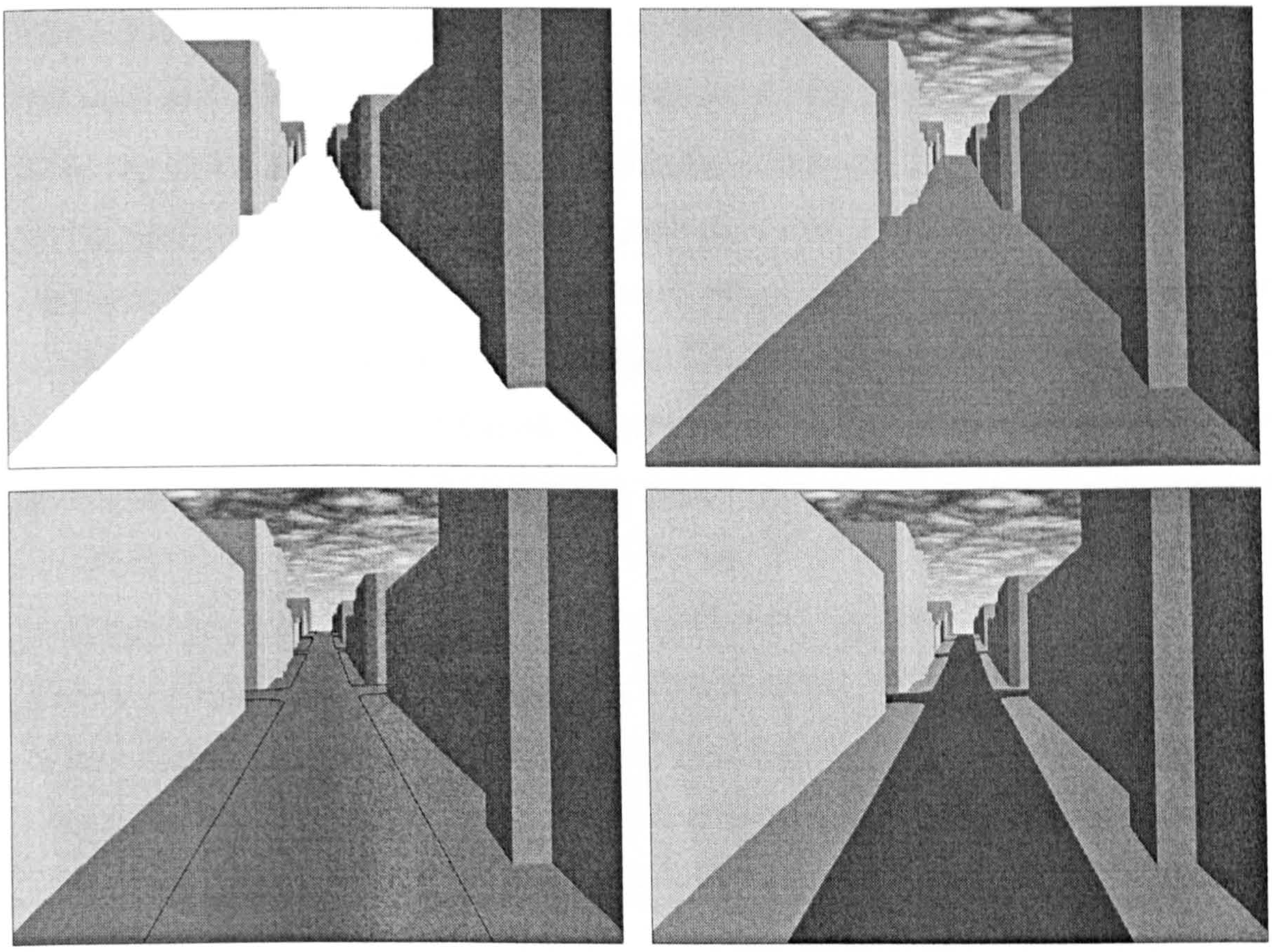


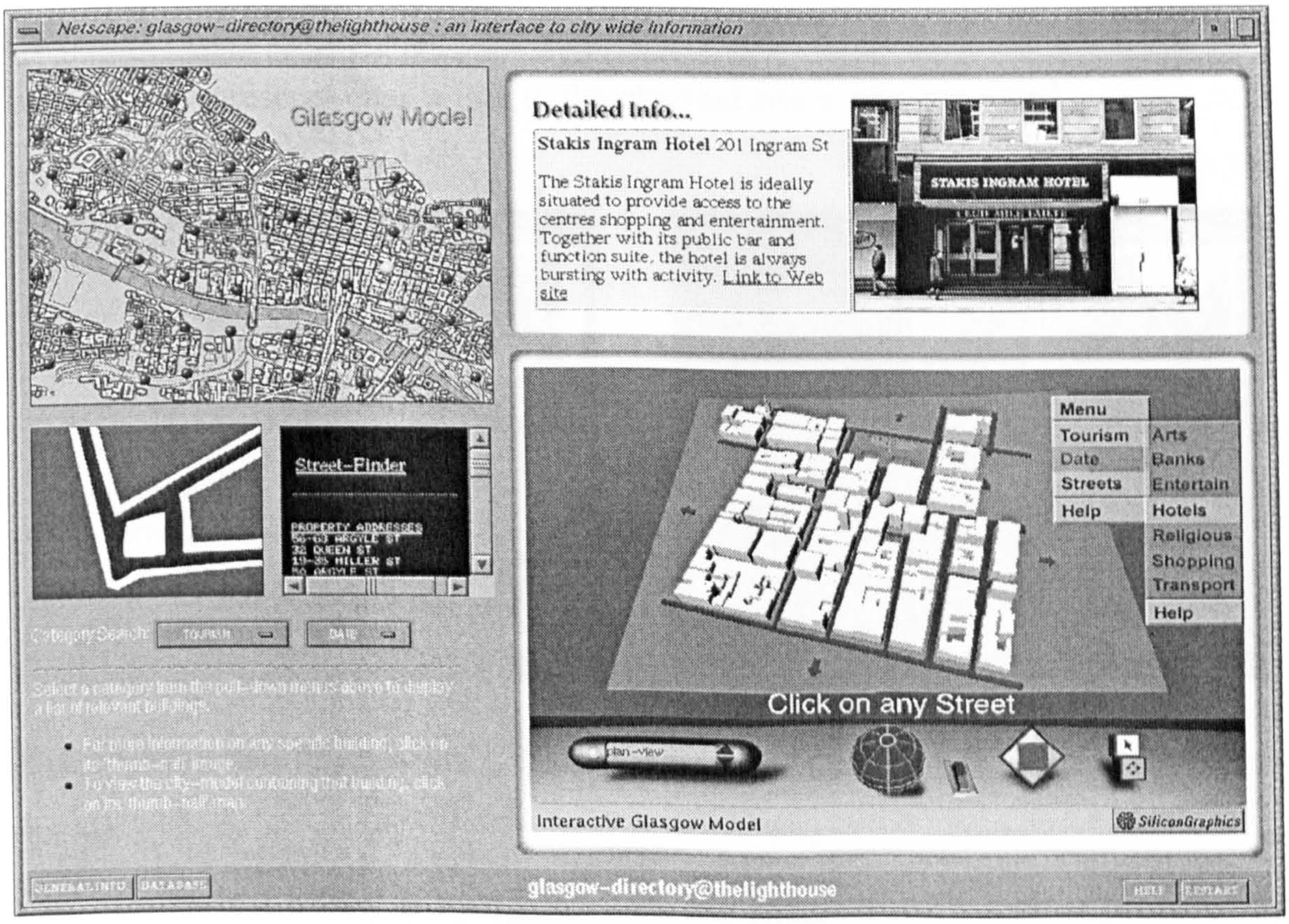
Figure 5.6 Terrain differentiation

5.5 The Glasgow Directory

The concept of developing an on line information system based around the original Glasgow city model was clearly valid and when the release of the VRML2 standard coincided with a request from Glasgow's Lighthouse corporation for an innovative project to associate with the 1999 Festival of Architecture it was decided to rework the model to take advantage of all the latest developments. (ABACUS 2002)

The city model was brought up to date by the addition of several of the more prominent developments that had taken place since the 1980s. The terrain model was also rebuilt and using the experience gained from the Edinburgh model was subdivided to differentiate between streets, rivers and areas of hard and soft landscape as described above. The enhanced standard incorporated within

VRML2 allowed for the definition of "in-line" nodes. This made it possible to define a standard scene graph for every model section and then populate this structure with the required data. This made for a very flexible format where data could be added or subtracted programaticaly. This ability allowed the user to change the content of the graphical representation and so relate the visual scene to the nature of their interest. To provide a unifying structure and to make room for all the proposed functionality the interface was sized to fit within a 1024 x 768 screen dimension. The intention was to unify the interface and generate the appearance of an integrated application rather than the fragmented collection of windows as seen in the prototype. The adoption of this format allowed the program to be divided into two columns, each with two windows, those on the left dealing with queries, those on the right displaying the results. The "query" aspect of the interface represents input to the system the "result" being the information returned from the system. Within the four quadrants of the interface any user interaction results in a cyclic flow of information between each of the other functions.



Detailed Info...

Stakis Ingram Hotel 201 Ingram St

The Stakis Ingram Hotel is ideally situated to provide access to the centres shopping and entertainment. Together with its public bar and function suite, the hotel is always bursting with activity. Link to Web site

PROPERTY ADDRESSES
56-63 ARGYLE ST
32 QUEEN ST
19-35 MILLER ST
50 ARGYLE ST

Figure 5.7 The Glasgow Directory Interface

The map section gives an overview of the entire area covered by the model. Individual sections are denoted by buttons which divide the model into regions that are sized according to a compromise between what is easily downloaded and visually interesting. Feedback is provided by highlighting the current selection. Once a user selects a portion of the model the geometry is downloaded into the graphical window along with a listing of street names contained within the area.

The query section is contained at the bottom left of the interface. This section allows access in two ways. Firstly the information attributed to the model is subdivided into categories under pull-down menus. Selection of any category of building generates a listing of all matching buildings. This listing, Figure 5.8, also provides access to detailed information and will also load the correct section of the model with the viewpoint orientated to the selected building.

Detailed information is proved for around 400 buildings within the area of the model. This information is comprised of an image and a short textual description. Where relevant access to other media elements, such as a detailed model or QTVR movie, is accessible through a further link. If the building or company has their own Internet presence then a link to the external site is also provided.

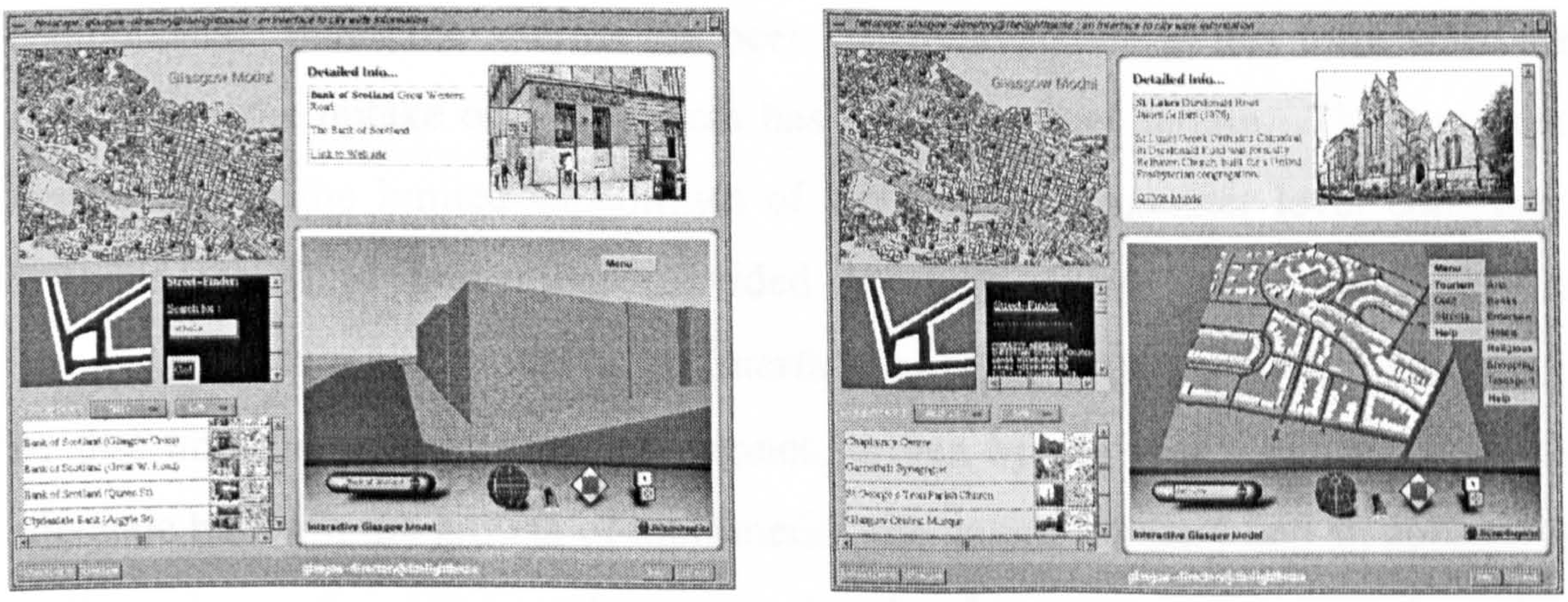


Figure 5.8 Interface functionality

The geometrical model is navigated with the standardised interface provided by VRML. This is enhanced by the addition of features to make it easier to navigate both within the individual model sections and throughout the city as a whole. In order to access the building attributes each building in the database is identified by an icon as described previously. In order to decrease the visual clutter produced by marking a large number of buildings a menu is provided that switches between each category within the information database. When a building category is chosen the icons are revealed, picking an icon loads additional information into the detail window above. User orientation is reinforced by providing a mode that allows the graphical selection of individual streets, a mouse pick returning the relevant street name. Navigating between regions of the city from within the model is facilitated by the provision of arrows situated on the boundary of the model. Clicking on the arrow loads the designatedt region of the model that lies on that boundary. Additionally invisible portals are situated at the ends of major streets that end at a boundary. If the user is about the "walk" off the edge of the model the portal is activated and the next appropriate region is loaded thus providing a spatially seamless transition between regions.

This incarnation of the Glasgow model has achieved all of the original aims set out some ten years previously. The ability to meet these criteria has largely been enabled by the collective wisdom available from the global Internet in the form of ubiquitous standards, shared knowledge and freely available software components. While the Internet has been widely available in this form since the early 1990s the uptake of 3-D content has been less than certain. This has been largely due to the limited capabilities of the average consumer level computer. Without the fluid movement provided by high frame rates the graphical manipulation of a city model as an interface to other data is less than satisfying. At that time the current graphics engines, driven by the host CPU, were barely adequate but with the growth of multimedia, the games industry and an apparently endless desire for more performance the development of graphics hardware targeted at the consumer rather than the professional was about to transform the market.

5.6 Consumer Graphics Cards

As computers have become faster and more capable the tasks that they are programmed to complete have become more demanding. This in turn means that the task of interpreting the outcomes of the process are also proving more onerous. Since the bulk of our information regarding the outside world comes through the human visual system it is only natural that a graphical depiction is the most meaningful and easily assimilated. Any improvement on the capabilities of a graphical display to convey meaning can be derived from the quality, veracity and quantity of the graphics themselves. This has been understood for as long as there have been mechanical aides to calculation and the progression of this technology has been one of the threads throughout this thesis.

As the information technologies have advanced, the capabilities of graphics systems have become ever more important and can now be seen to be evolving at a greater rate than the any other subsystem within the conventional computer. Taking the Personal Computer market as a standpoint it is possible to trace a strand of the development of graphics hardware from the days of basic text displays to the powerful geometry processing units of today.

5.6.1 A Graphics Taxonomy

Keith Cok et al (Cok 1999) offer a taxonomy of graphics hardware which classifies graphics subsystem as one of four different types. This classification is based on the representation of the five basic stages of the pipeline and can be used to position technologies both with regard to their introduction into the market and the performance offered.

Each member of the proposed taxonomy can be described using **G,T,X,R** and **D** as above. A dash represents the division between the stages performed in dedicated graphics hardware and those stages performed in software on the host CPU.

Pipeline Stage		Process
Generation	G	The process of generating a graphics data structure from the geometrical primitives that comprise the scene. This is performed by software on the host CPU.
Traversal	T	The process of scanning through this data structure and performing any user supplied or temporal interventions before calling those functions that supply the appropriate data to the next stage.
Transformation	X	The process of matrix transformation from one co-ordinate system to another i.e. world space to eye space and the calculation of lighting and shading before clipping the primitives and projecting them into screen space.
Rasterisation	R	The process of calculating the depth colour texture and blending of pixel fragments before drawing the fragments into the framebuffer.
Display	D	The process that outputs the contents of the framebuffer to the display.

Table 5.1 Representation of the graphics rendering pipeline

GTXR-D The sole function of a GTXR-D graphics subsystem is to regularly update the screen by scanning pixel values from video memory to a display monitor. All other graphics functions are executed on the CPU.

GTX-RD This configuration has a rendering engine that implements the scan conversion of screen space objects into video memory and performs screen space shading and other pixel operations. Transformation and lighting are still executed on the host CPU.

GT-XRD These subsystems add transform engines that implement in hardware the transformation from world space to eye space, perform lighting calculations and finally transform the output to screen space. Most of the execution overhead is offloaded from the CPU which then has more cycles to expend on traversing the scene graph.

G-TXRD This configuration would prove to be too inflexible for a general purpose system and while useful for an embedded system has little relevance to interactive graphics.

The performance of these subsystems can be equated with the movement of the dash towards the left. This represents the implementation of more functionality in dedicated hardware and, to a certain extent, also serves to mark developmental stages of the technology. While some manufactures have always offered specialised hardware outwith this time scale the consumer market can be seen to closely conform to this view. The development of graphics hardware is closely linked to the introduction of standards and their acceptance by both the manufactures and software developers. From 1982 all PC graphics applications that required any level of performance used the ROM BIOS chip to access graphics functionality. While there were limited hardware offerings from other manufactures, notably Hercules, this meant that graphical functions were restricted to the standards that IBM provided. Without an independent and thus non standard interface to developers were required to adhere to the standards as enumerated below. (IBM 1984)

Monochrome Display Adapter (MDA)	monitor adapter	£186 £181
Colour Display Adapter (CDA)	monitor adapter	£552 £194
Enhanced Graphics Adapter (EGA)	monitor adapter	£736 £482
Professional Graphics Adapter PGA)	monitor adapter	£3060 £2829

Table 5.2 IBM graphic standards in 1985.

The MDA standard was the first dedicated graphics card. Designed to display text only it offered a resolution of 720x250 pixels and could display upper case characters in character cells of 9x14 pixel blocks giving 18 rows of 80 column text. While the MDA standard could only address on screen character positions, CDA supported pixel addressing and offered 16 foreground colours and 8 background colours. The choice of resolutions was comprehensive offering a choice of horizontal resolutions, at 640, 320 and 160 all at a vertical size of 200. This would support an 8x8 “dot box” giving 25 lines of 40 characters or 25 lines of 80 characters. Pixel addressing also supported the display of filled shapes and line vectors. The Enhanced Graphics Adapter offered the same capabilities but at a higher screen resolution of 640x350 pixels. (IBM 1996) The Professional Graphics Adapter was a complete subsystem in its own right and consisted of a board set that occupied three of the four vacant slots on the bus. This system offered;

> *"its own built in processor and graphics software , as well as 320Kb of display storage. These combine to produce a display controller with many built in functions such as rotation, clipping and translation. This meant that neither the user nor the IBM PC need be concerned with the complexities of, for example, rotating a three dimensional figure."*

This was offered with the proposed ISO/ANSI standard Graphical Kernel System.

The next standard emerged in 1987 and became the base configuration for all subsequent PCs. The Video Graphics Array [VGA] supported 16 colours at 640 x 480 or 256 colours at 320x200. By this stage IBM were trying to diverge from the standard PC line-up with the PS 2, microchannel bus and OS2 operating system. This also introduced a new graphics adapter called the 8514/A which was compatible with VGA but offered greater functionality. IBM were reluctant to release engineering information on the 8514/A specifications and it soon became eclipsed as rival manufacturers bypassed this standard and developed their own version based on the Texas Instruments 3410 series processor. (Hall 1989)

The pace of development was beginning to accelerate and the next standard emerged in 1990. This was termed the Extended Graphics Array XGA and with 1Mb of frame buffer was capable of 1024x788 in 8 bit colour or 800x600 in 16bit. The final standard was termed SuperVGA and supported all resolutions up to 1600x1200 at 32 bit colour.

These standards only addressed the **GTXR-D** version of the taxonomy but with the advent of new technologies like multimedia and a growing demand from consumers for more performance regarding 3-D applications in computer games then market forces encouraged the development of greater functionality. The first graphics card manufacturer to exploit this opportunity was ATI with the Mach8 in 1990. This was an attempt to leapfrog the current standards with the development of a co-processor that added accelerated functionality at the rasterisation stage. The exploitation of the **GTX-RD** taxonomy specifically for 3-D was left until 1994 with the 3Dfx Voodoo series of subsystems. The problem with the introduction of enhanced functionality lies in the ability of software to exploit the new features if they deviate from the current standard. The lack of an industry wide API for new features culminated in the adoption of the GLide library as a defacto standard. Without the use of standard libraries a graphics card would only provide enhanced functionality to applications that had been specifically developed for that hardware. This resulted in the sale of graphics cards complete with bundled software titles which limited market possibilities. In the intervening period development concentrated on adding features to the rasterisation stage of the pipeline and it was not until 1999 that the **GT-XRD** taxonomy was exploited with the introduction of the transform and lighting engine of the nVidia Gforce series of hardware in 1999. This final development heralds the start of the law of diminishing returns for pure processing power. The bottleneck within the pipeline has been moved from the rate at which the system can feed the pipeline with primitives to the rate at which the physical limitations of memory and bus bandwidth can function. A six month development cycle between products and the huge complexity of the graphics processing units (GPU) have taken GPUs to the limits of current technology. nVidia's current GPU boasts more transistors that Intel's Pentium4 and Moore's Law governing performance increments has been bypassed by introducing multiple parallel pipelines. (nVidia 2002)

The pursuit of performance through the ability to draw faster will always show some returns but there are also performance gains to be won by drawing less.

5.7 Optimisation Strategies

The pursuit of performance through the ability to draw faster will always show some returns but there are also performance gains to be won by drawing less. The major advantage with VRML, other than its ubiquitous format, was that it imposed a regimented ordering on the geometrical data. This structuring could be used to implement a scene graph that offered substantial efficiencies to the rendering process. The VRML format considers all members of the input file to be nodes within the data stream. By grouping these nodes within a nested hierarchy it is possible to imply a degree of spatial cohesion within the graph. The addition of this minimal degree of intelligence to the input stream allows a number of optimisation strategies to be applied that could reduce the amount of data that had to be drawn.

The ability to draw less has obvious performance advantages. Within any scene, and particularly an urban landscape, there will always be substantial areas of the city that are not at any one time within view.

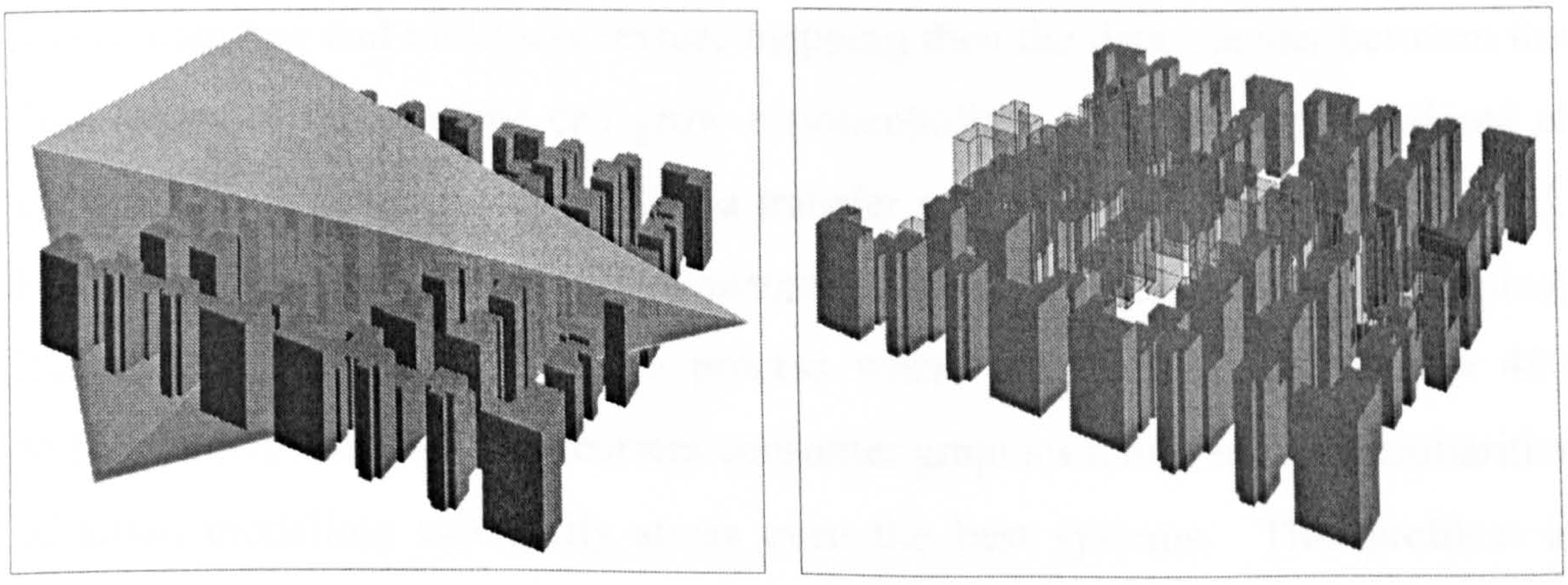


Figure 5.9 The intersection of the view frustum and the object database

Figure 5.9 shows a diagrammatic representation of the view frustum intersecting with a group of buildings. This shows that even within one block only a small number of the buildings contribute to the scene. In the VRML scene graph each grouped set of geometry nodes is provided with a bounding volume, this volume is tested against the clipping planes of the frustum and if outside the boundary is pruned from the tree. This suggests that the spatial organisation of the original data is of prime importance. The ideal scenario is that the world should be recursively subdivided into squares so that each square would occupy the smallest bounding volume. By recursively subdividing the geometry a number of smaller volumes can be contained within a single larger group, then if this parent group is culled the traversal need not descend the tree any further. Applying this technique to the geometry of the Glasgow model resulted in a substantial improvement in the average frame rate. Further optimisation becomes considerably less trivial and although a number of techniques exist for culling occluded geometry they are not universally applicable or require specialised interventions during the database traversal which were not implemented within the VRML standard.

As previously discussed the used of dedicated hardware will improve the system rendering performance to a significant degree yet there is a finite limit to gains that can be realised. As the target speed in frames per second is increased and as the scene complexity rises then the amount of data sent to the frame buffer increases in proportion. With advanced rendering techniques requiring deeper pixels, blending and multipass texture mapping then the data transfer between the final stages of the pipeline can grow exponentially. A 32bit image rendered at 1600 x 1200 resolution will require a transfer rate of over 5.8 Gbytes a second. Further up the pipeline similar constraints are met when accumulating pixel fragment data in the rasterisation process where despite fill rates of over 480 Million pixels a second from current consumer graphics hardware the peculiarities of urban modelling can easily stress even the best systems. This problem is largely due to the depth complexity of an urban scene.

As shown in Figure 5.10, a typical urban scene is composed of numerous buildings receding into the distance. Each polygon in the scene must be sent down the pipeline to the frame buffer before its contribution to the final image can be assessed. Because of the degree of mutual shadowing when seen from the eye point there may be tens or hundreds of faces that have to be tested and rejected before the scene is complete. Each pixel that is processed occupies some portion of the system's time then when a lot of processing is wasted on geometry that is never seen the overhead can be considerable. VRML offers the ability to toggle the Level of Detail (LOD) on the basis of the geometry's perceived contribution.

This scheme manages the display of different levels of detail on the basis of predetermined switching distances that correspond to elements of the model with successively diminishing detail. The application of this functionality was not so dramatically successful as in the previous case because the level of detail on the Glasgow model buildings was already minimal and the areas visualised at any one time were small. The image in Figure 5.10 illustrates the extent of depth complexity by blending semi-transparent surfaces – the greater the opacity the more pixels are overwritten in the frame buffer.

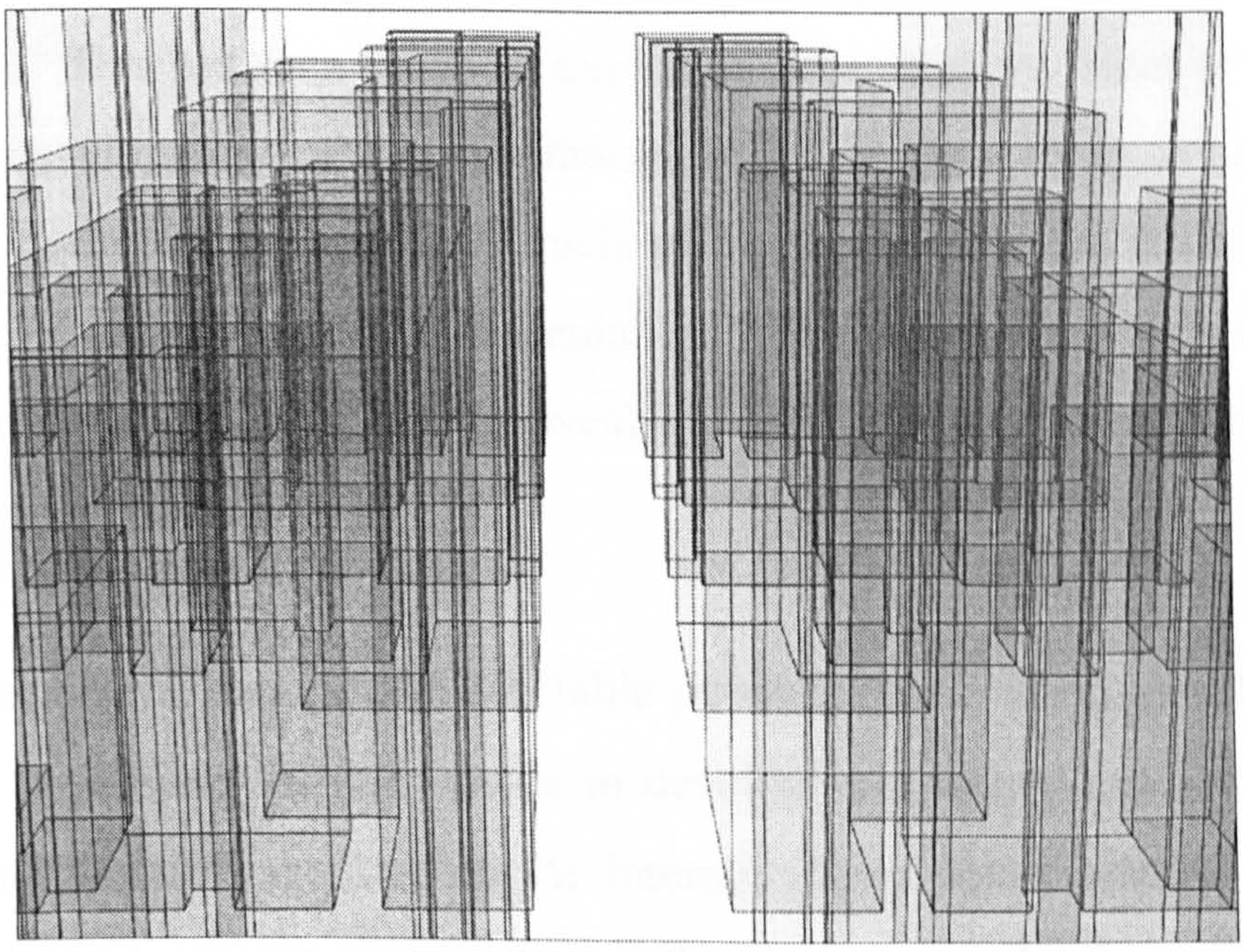


Figure 5.10 Depth complexity within an urban scene

This is a fundamental limitation to performance and as models become more sophisticated and hence more detailed can only prove to become more limiting. There are a range of culling techniques that seek to ameliorate the problem but most rely on constrained navigation or complicated pre-processing. However it is interesting to note that when a fundamental limitation is reached it is rarely that the whole solution is found in improved and optimised software and that the resolution is more likely to be delivered by hardware. In a software solution the only way forward is through the aggressive optimisation of the scene graph, asking for more expertise of the model creator and requiring more features on the modelling and visualisation applications. In a hardware solution the optimisation is transparent and the performance is available to all applications. In the case in question, two technologies have surfaced, one which reduces the image plane to a set of tiles small enough to reside in the raster processors cache where occlusion culling can take place without expensive memory transfers and another which automatically hashes the scene into bounding boxes which when rendered front to back are culled on the basis of their occlusion and contribution to the final scene.

5.8 Summary

The prime motivation during this era has been clearly drawn from the evolution of the Internet. This has seen the most explosive growth of any facet of technology within the development of the information technologies. Despite having many parallels with the "boom and bust" cycle of the multimedia era the lasting worth of the Internet and its content is undeniable. The real legacy of this era is in the standardisation of products and the wealth of freely available knowledge that is willingly disseminated.

These developments have made available powerful tools which would otherwise have been impossible for individuals to develop in a timely manner. The cross platform compatibility required by the Internet allows developments to be shared without hindrance and the continual quest for market share by the providers of end user technology ensure a steady stream of new features that can be leveraged.

The success of the Glasgow Directory has been built on these developments allowing the inclusion of functionality that would have otherwise been unavailable. The requirement for graphical performance in order to enjoy these benefits has produces hardware technology that has eclipsed Moore's Law governing processor development and consumers can now access capabilities far in excess of even the most optimistic prediction of just a few years ago.

Chapter 6 Into The Future

The ultimate Digital City would be an environment that embodied as much of reality as was possible. This would have to mean that an observer should be able to experience all aspects of the computer model in as natural a format as possible. In turn this sets the requirement that all the separate processes and technologies discussed so far should come together to reinforce this concept through mutual enhancement. Technology has now progressed to such an extent that the application of any one advance, no matter how revolutionary, will show less overall benefit than the combination of several techniques. This chapter seeks to bring together a range of processes that are all just incremental advances in their field yet when taken in combination point the way towards how the Digital City may be implemented in the future.

So far each chapter has dealt with a particular era represented by the defining technology of that period - from the discovery of human computer interaction through vector graphics, the introduction of solid shading techniques and on to the commercial explosion of all things multimedia and the promise of the Internet. Now that the exponential pace of development has increased to such a dramatic extent there is no one defining technology that can label the present time period. Rather the worth of any one particular advance is diminished if taken out of the context of all other technologies.

This means that no matter how advanced the modelling and software methodologies might be then without the graphics hardware to process it, the projection system to display it and an haptic interface to control it the net benefit may never be enjoyed to the full.

6.1 Data Sources

The fundamental problem with constructing a model of any urban situation is the shear size, scale and complexity of the subject matter. A typical inner city sector may be comprised of upwards of 300 buildings within the footprint of a single 1:1250 map, an area of just 0.25 sq. kilometres. If we budget for a maximum expenditure of no more than a day for each building this then assumes a time scale of around one entire man year to research, initiate and complete a minimal model of each principle structure and would even preclude any attention to the groundscape, street scenery or other contextual media.

Some of the answers can be found in the prospect of the traditional agencies broadening their outlook to capture and hold relevant data in a format that is more compatible with the requirements of 3-D modelling. The Ordnance Survey has traditionally been the prime source of data for projects involving data capture with an urban area. This source had been exploited by manually digitising the relevant features, as in the description of the Glasgow model, or by parsing existing digital data as with the Edinburgh project. While this route to data acquisition was undoubtedly efficacious it was still subject to a number of shortcomings in regard to this particular purpose due to its lack of a native structure. In the intervening period the information technologies had enjoyed a much higher profile and the field of digital cartography had continued to evolve at an equal pace.

6.1.1 Ordnance survey

The OS encapsulates their remit within four functional statements:

- Providing reliable positioning and reference information
- Collecting, holding and maintaining information
- Linking the above with information held by others
- Marketing cartographic information to meet customer needs

Historically the OS have been positioned to maintain the only source capable of providing a complete mapping of the country. This fact coupled with the de-facto national referencing system has enabled the OS to support a wide range of users and applications.

The OS made the move into the electronic age simply by digitising their existing paper based map stock. This conversion process was driven by the need to provide a new distribution medium. However while this filled temporary market needs the process did not address the wider requirements of users in the digital information age. During the 1990s the dominance of the OS was challenged by the emergence of a number of technologies that made it easier for external agencies to gather, process and distribute geographical information. As a consequence there was a very real danger that the fundamental asset of the OS, as the national depository of cartographic information, might be devalued by the fragmentation incurred by diverging standards and technologies. This has been addressed by a change in strategy in that the concept of cartography has been replaced and augmented by a broader vision of a Digital National Framework. [DNF] (OS 2001)

The DNF is a re-engineered version of the existing National Topographical Database, (NTD), where the purpose of enhancing the existing provision is to move towards helping the dataset meet modern information requirements. This is being performed by changing from an abstract, cartographic metaphor concerning the representation of features to the definition of data depicting real world objects in a structured format. Each of these features is to be given a unique identifier that will allow the superimposition of layers of metadata so that other forms of information can be spatially referenced and assimilated within the database.

In theory this should make the task of building city models all the more practical and efficient. The polygons representing buildings and other area features can be easily extracted from the database and augmented data can be reinserted. This would solve many of the problems encountered in utilising digital cartographic data as described earlier. However the basic concept still lies in the two dimensional domain and there are still fundamental problems regarding the description of the city.

Since the original collaboration between ABACUS and the OS was based around the premise that additional 3-D information could be integrated within the existing cartographic dataset then the future still holds the prospect of better data sources emerging from some impending release of the Digital National Framework. Unless we look an extraordinarily long way into the future it is very unlikely that the fundamental 2-D nature of the map will be usurped by a 3-D alternative and until then there will always have to be some compromise to the transition between the two domains.

6.2 Modelling

The increasing availability of digital cartography has done much to alleviate the base-line time scale and attendant effort for fundamental data capture. However, experience has shown that this approach is potentially flawed in that the methodology tends to result in models that are less than amenable to the proposed application processes due to shortcomings in the structuring of the data. The mapping of cartography to model tends to result in a dataset that is monolithic in concept if not in structure. This is due to the inherent accuracy of the cartographic abstraction. Given a one-to-one mapping between the seed of a building, it's 2-D footprint as originated from a cartographic source, and the derived 3-D model then there is an inevitability as to the uniqueness of the eventual product. This is in direct opposition to the key strengths of a computer generated system, the ability to copy, clone and reuse geometrical components.

6.2.1 The Economics of Urban Modelling

If we pursue this line of reasoning to it's inevitable conclusion it would become obvious that to construct an accurate yet basic model of even a small urban area to a reasonable degree of completeness would be a task destined to economic failure. New technologies, such as large scale photogrametric data capture, can undoubtedly speed parts of the process yet are destined to trade off problems of data quantity, redundancy and overall usefulness against apparent speed and accuracy.

The core precept that make this methodology unworkable is the user's inability to separate the regimes of draughting, where dimensional accuracy is at a premium, from the less numerically taxing arenat of modelling.

Ideally, we would wish to see a one-to-one mapping of reality onto the model as it would then prove equal to any evaluative task. However even with the most sophisticated technology envisioned within the near future this will not be achieved. The only reason that some prior models have come into existence is through the use of some measure of abstraction in the process of converting real world artefacts into their digital equivalent. If this is to be the route to an economical process it then begs the question as to what level of abstraction is appropriate for the tasks envisioned for the resulting model.

6.2.2 A Choice of Representation

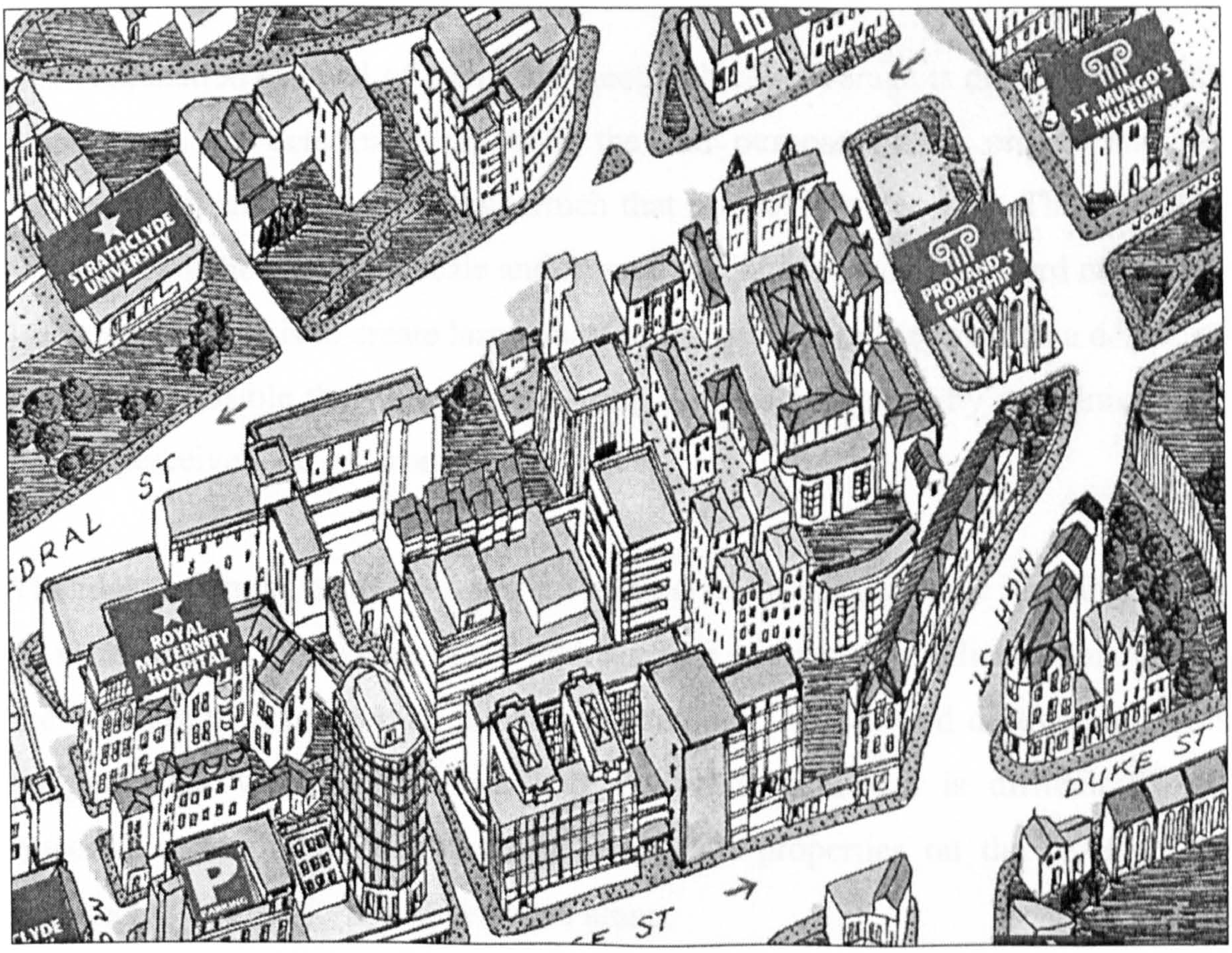


Figure 6.1 An abstract representation of urban form

Cartography has developed a standardised language for the depiction of all those features that are commonly found within our landscapes and in our cities. On reflection there are remarkably few symbols that are required to describe a very varied environment and yet maps are renowned for their accuracy and clarity of content. The legibility of a map is in direct contrast to the abstract nature of the symbolism employed especially when these symbols bear so little resemblance to their real world subjects.

This phenomena can be successfully exploited in a range of media where a diagram, as long as it bears representative and recognisable features, can form an immediately recognisable rendition of reality. This is illustrated in Figure 6.1 where a diagrammatic map, occupying a position midway between cartography and the image, can readily convey sufficient information for a user to both locate and orientate themselves. This is in spite of gross distortions in scale and structural accuracy.

The conjunction of level of detail and geographical coverage is on a sliding scale whose axes are determined by both the end purpose of the project and the technological environment within which that project is conceived. This is amply illustrated by the variety of scale and format available within a standard map base. If the project goal is to create large tracts of urban landscape to as high a degree of realism as possible then a suitable starting point is suggested by examining how people perceive their environment.

The most common, useful and proven method of locating a single building within an urban area is through its postal address. An address will indicate which street the building is in and which individual building is the desired destination. This suggests that within substantial portions of our cities it is difficult, if not impossible, to differentiate between individual properties on the basis of the uniqueness of their formal appearance alone.

As an example consider Glasgow's West End which, despite it's large geographical area, boasts many buildings of a similar design constructed with similar materials. Within this area buildings are more readily identified by their situation rather than any individual distinctive characterisation. This is illustrated in Figure 6.2 which depicts a range of addresses spread throughout a wide area. Taken outwith their context, individual buildings are superficially indistinguishable from one another. While these patricular buildings were constructed within the same period, from similar materials and to much the same basic design it would still be possible to visit any other area of the city and reduce the vernacular building stock to very few characteristic examples.

This fact suggests that the depiction of groups of building types may be reduced to a more singular or iconic representation. The method of depiction is dependant on the proposed end use. If a high degree of verisimilitude is demanded then there may be no alternative but to model each building as an individual unique entity, yet at the other end of the scale a purely iconic representation may be enough to augment a primarily cartographic representation and serve the needs of displaying a wider area at a lower level of detail. Midway between these two extremes lies the proposition that we could take the strengths of each of these strategies and allow an iconic representation but formed from accurately modeled archetypes of representative building configurations.

Figure 6.2 The Victorian archetype in Glasgow's West End.

6.2.3 The Choice of an Archetype.

There are a wide variety of purposes that may be asked of the Digital City and depending on the required outcome the degree of abstraction may be more or less. The original Glasgow model was less than realistic in its depiction of individual buildings yet the forms adopted were suitable for communicating the entire city at a computationally economical level. If the goal is to represent rather than recreate the city then there is much more freedom in the expression of architectural form. Taking this to a logical conclusion suggests that if the goal is to present *some city* rather than *a city* then the model only has to adhere to accepted urban design constructs in order to be entirely accurate.

This has been recognised for some time since Frey used the VIM and VIEWER combination to construct a library of generic building components in order to explore the role of continuity and variety in the urban environment. (Frey 1987)

> *"It consists of a library of facades, building forms and other urban space elements, even trees, cars etc. which can be assembled with a simple allocation program. This library is still limited and somewhat unsophisticated but allows already at this stage rather interesting insights into the operational possibilities and environmental consequences of collages of such elements; the possibilities and dangers of mixing formal languages, the unifying principle behind harmonic space, the compatibility or incompatibility of urban elements , but also operational opportunities of the display of space manipulation and the participation of those interested in the process".*

This rational breaks the city down into its most basic components and allows the formation of a descriptive language that can be used to reassemble these elements in order to depict any chosen design scenario. Figure 6.1 shows two possible urban blocks and in this instance might illustrate the effectiveness of height articulation in adding visual interest to the sky-line.

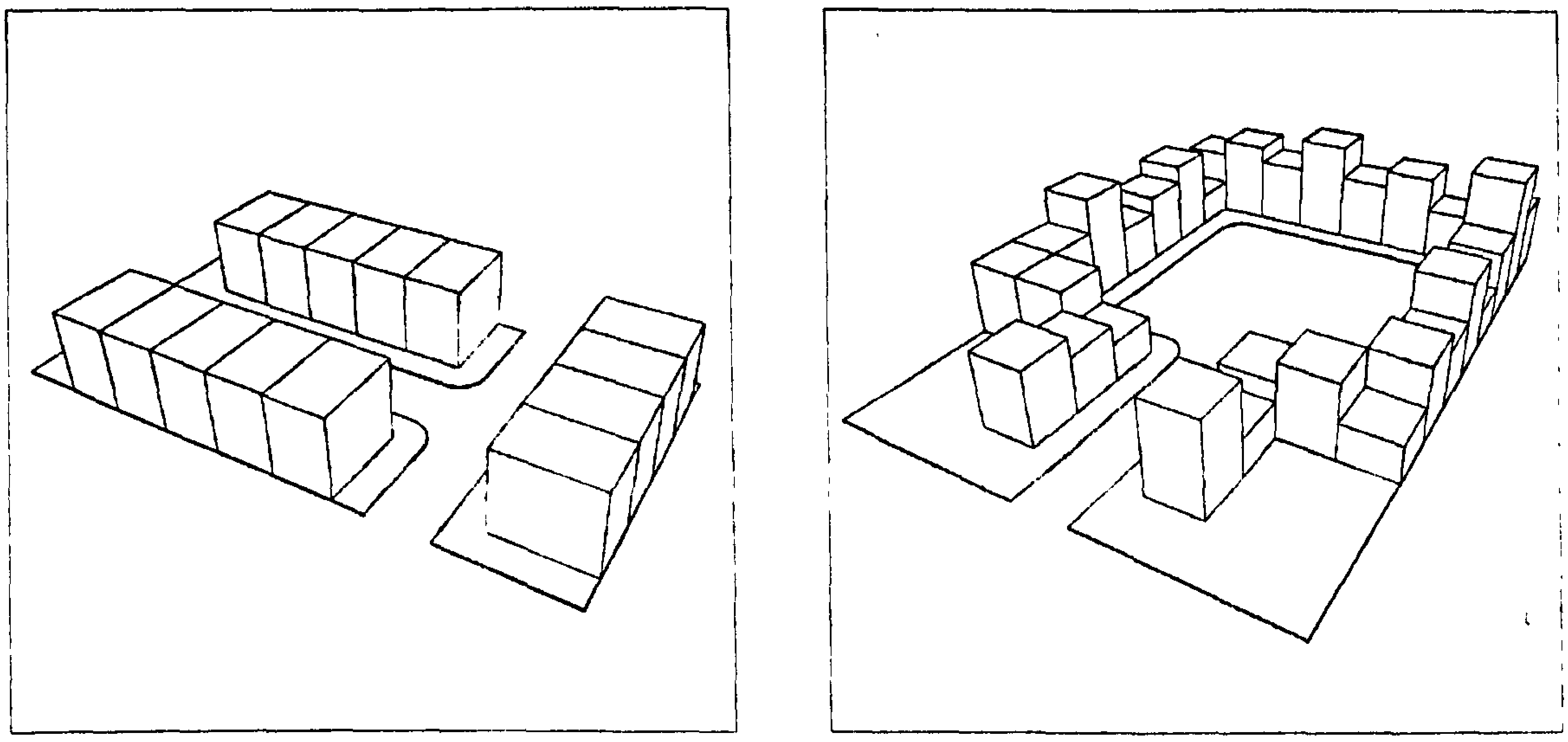


Figure 6.3 Urban blocks

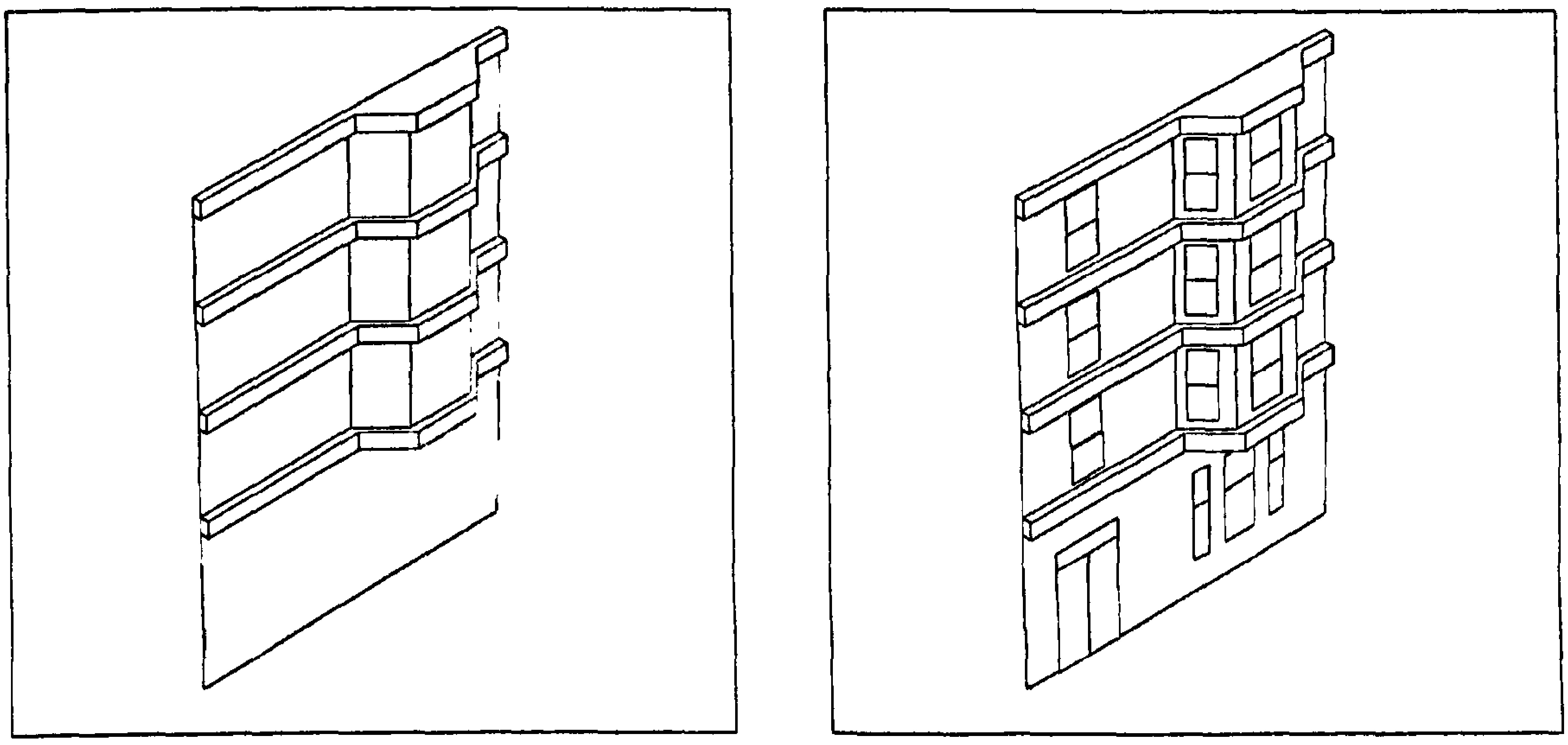


Figure 6.4 Elevations- base level detail

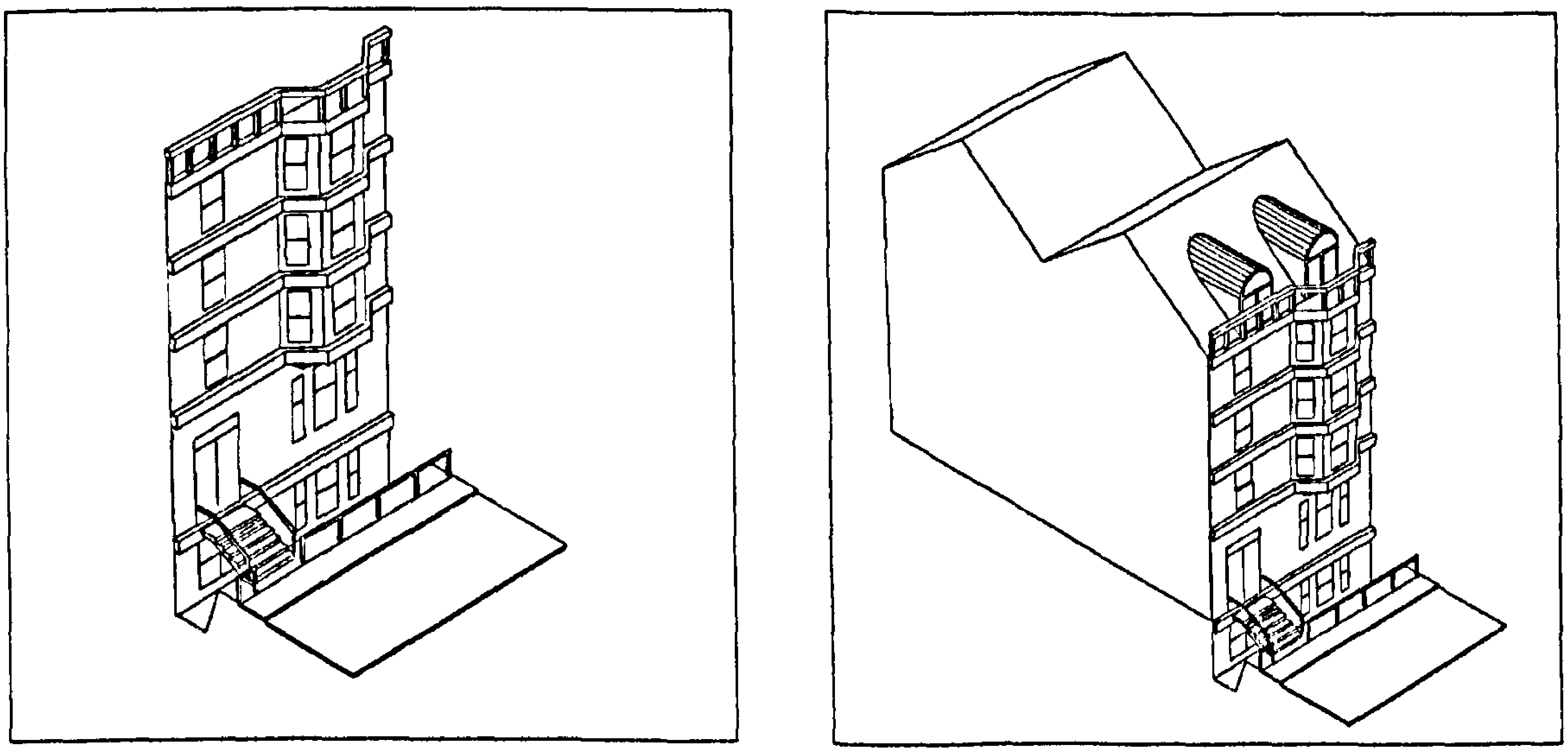


Figure 6.5 Elevations- assembling the component library

This architectural language is embodied in a grammar which combines a set of architectural elements – the vocabulary – with a systematic statement of rules governing the placement and use of these symbols - the syntax.

A hierarchical composition is formed from a common vocabulary of minor elements, like windows and doors, which when combined within the structure of a façade form an individual building's identity. These facades can then be attributed to the building blocks, shown in Figure 6.3, and become part of a tool capable of describing a rich and varied architectural landscape. The variety that such an approach is capable of generating is shown in Figure 6.6. These are just a few of the examples in the library and demonstrate the possibilities for quickly and economically describing generic but very recognisable and representative design scenarios.

The extent and classification of potential scenarios is limited only by the range of archetypes held within the library but might include some of the design issues described below.

In a typical street scene the aggregation of identical facades does not necessarily result in monotony as long as the facades have interesting and meaningful details. The composition shown in Figure 6.7 distinguishing features and when repeated generates a dull and meaningless space. If the facades are multilayered the result is usually a lively and interesting scene. The richer the detail then the more people will find it stimulating. Figure 6.8 also shows the consequence of emphasising classical symmetry which can becomes overpowering and may even acquire an undesirable political overtone.

Variety in the urban environment is required to avoid dull and meaningless spaces but too much variety destroys continuity. The aggregation of facades with no common visual denominator, as in Figure 6.9, can result in a formally disintegrative and chaotic space. If there is commonality with regard to scale and architectural language, then the composition is likely to be read as an integrated whole and provide a harmonic and acceptable form. Figure 6.10

Buildings of civic importance usually have a richer architectural language are may be larger in scale with different materials and colours in order to contrast with their context. Those that form the space in which such monuments are located should therefore have a simpler language and should be smaller and simpler in scale to form a neutral background in order not to compete. Figure 6.11

If private buildings with little civic importance acquire a much larger scale than other buildings in their immediate surroundings then the meaning and legibility of the urban environment is destroyed. Very often such large-scale intrusions express little more than the economic power of private companies and in some cases the dictatorial power of systems. Figure 6.12

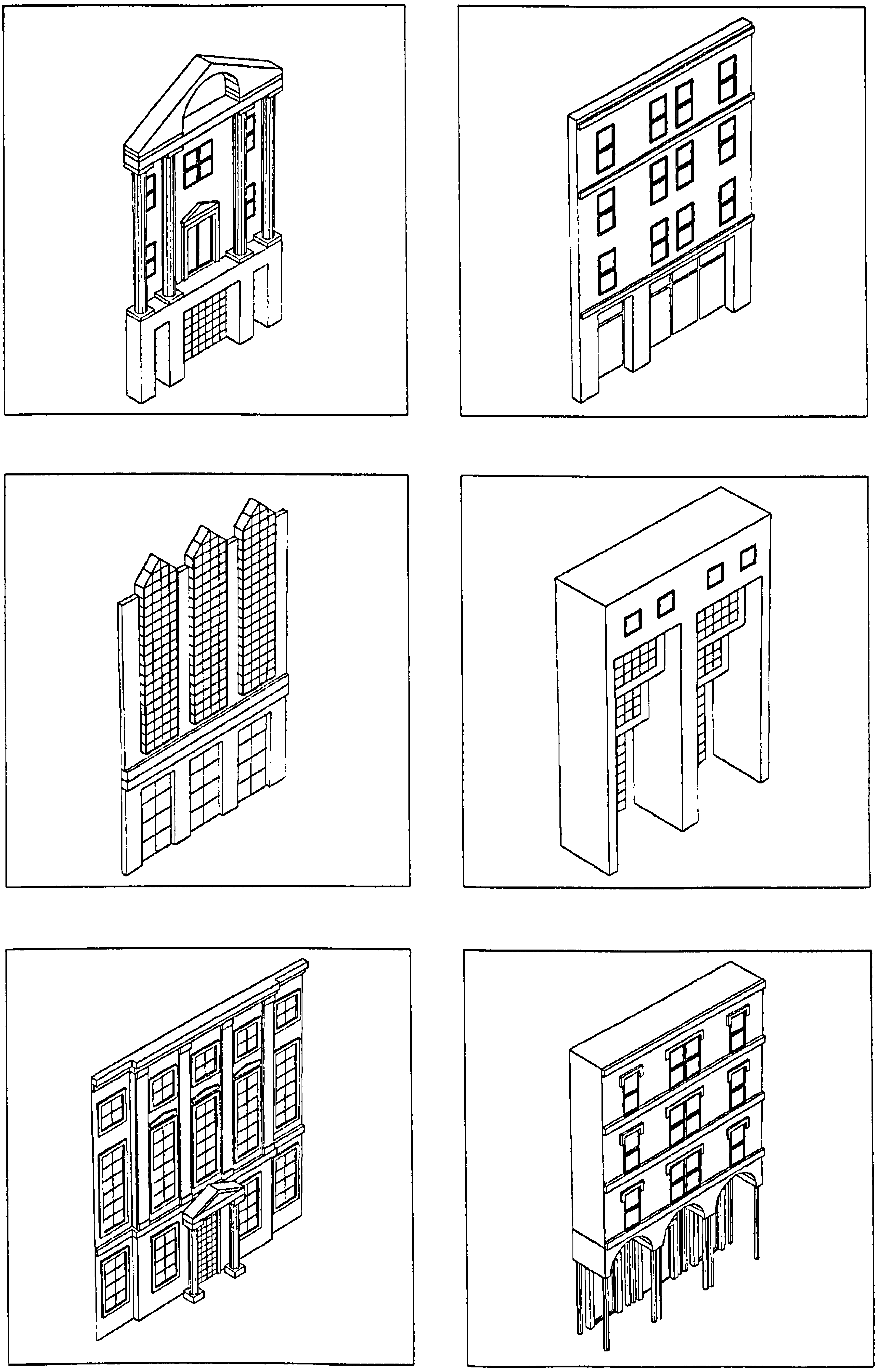


Figure 6.6 Elevation - hierarchy

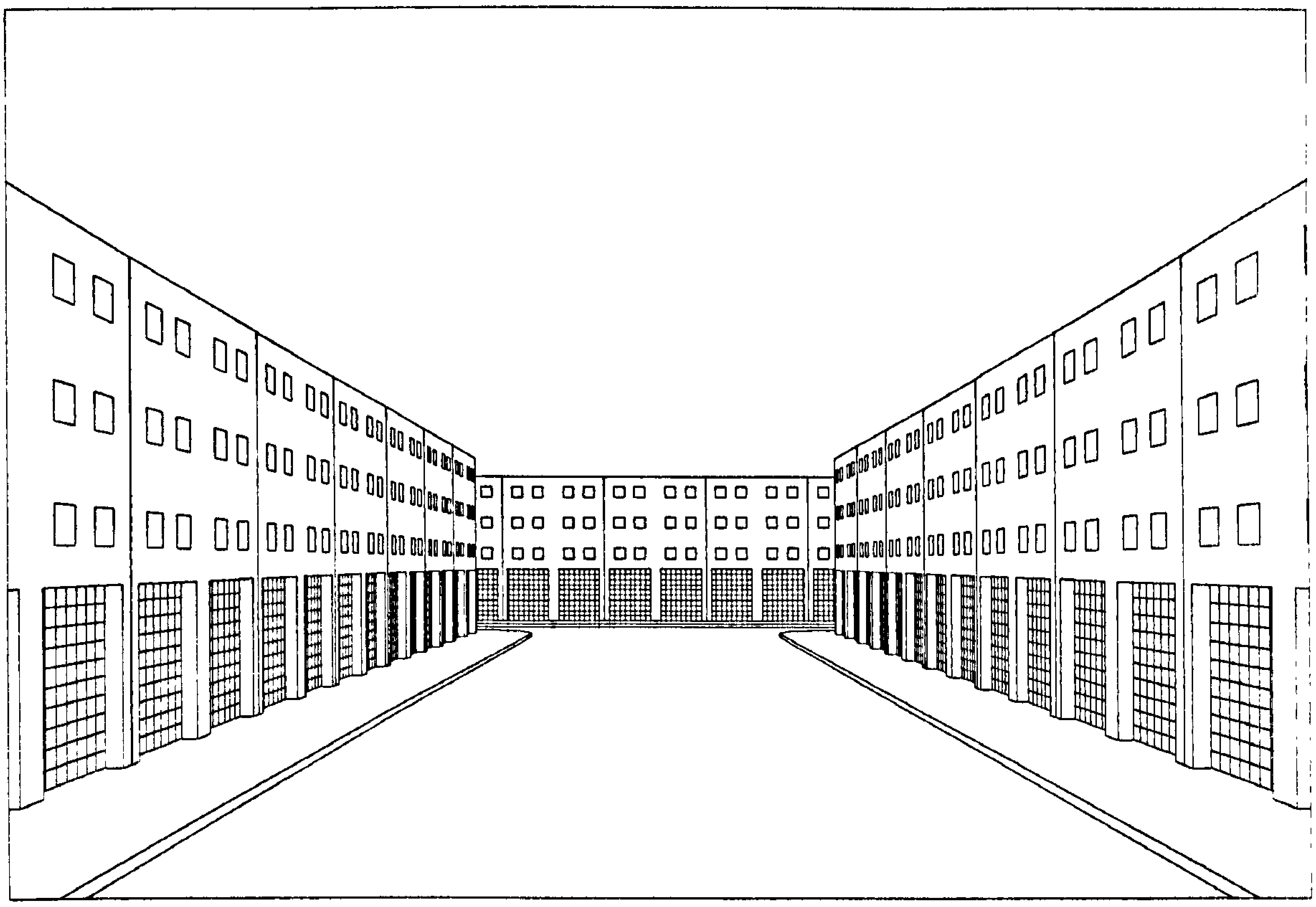


Figure 6.7 Continuity- an aggregation of identical single layered facades

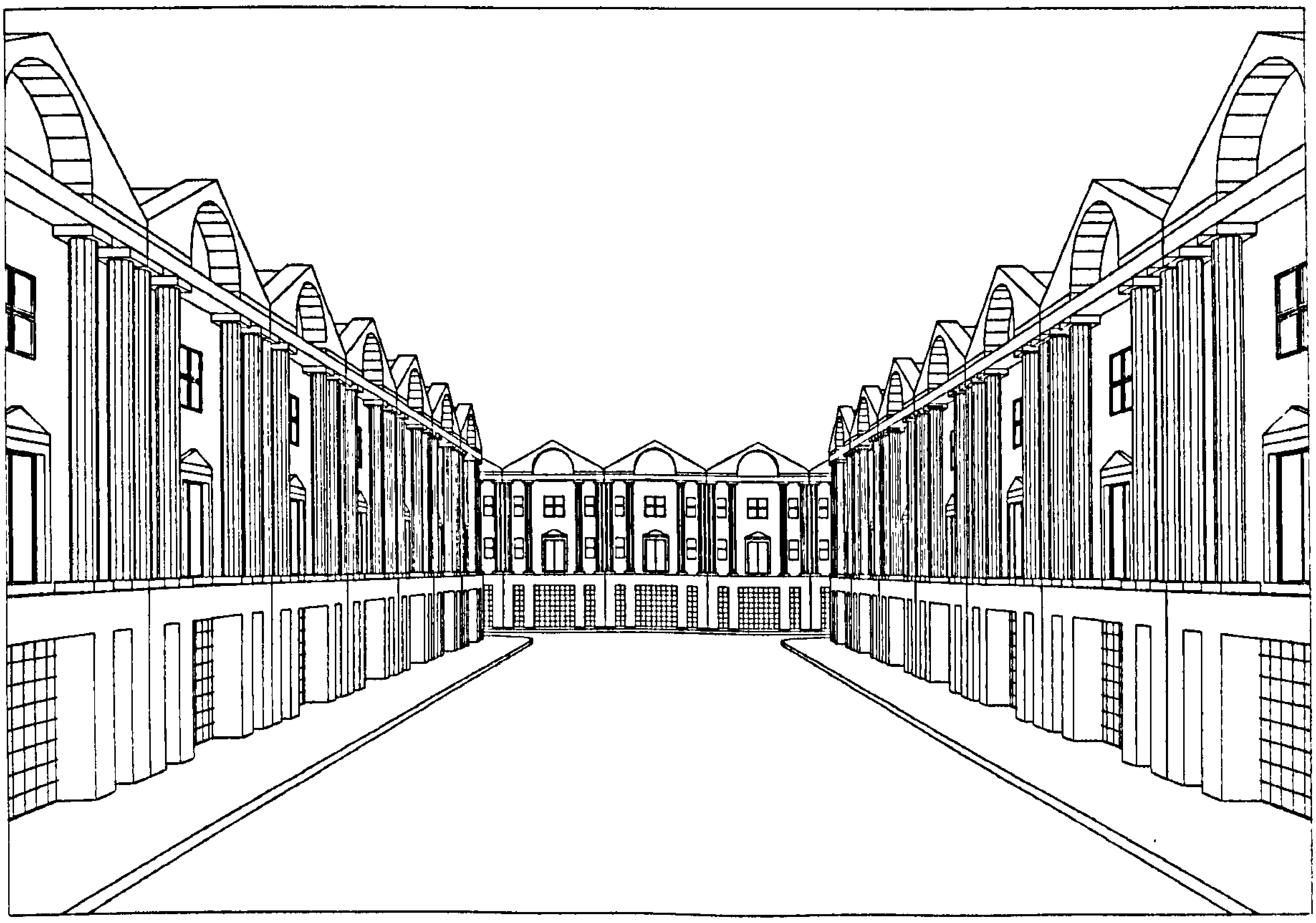


Figure 6.8 Continuity- an aggregation of identical multi-layered facades

Figure 6.9 Variety- an aggregation of entirely different facades

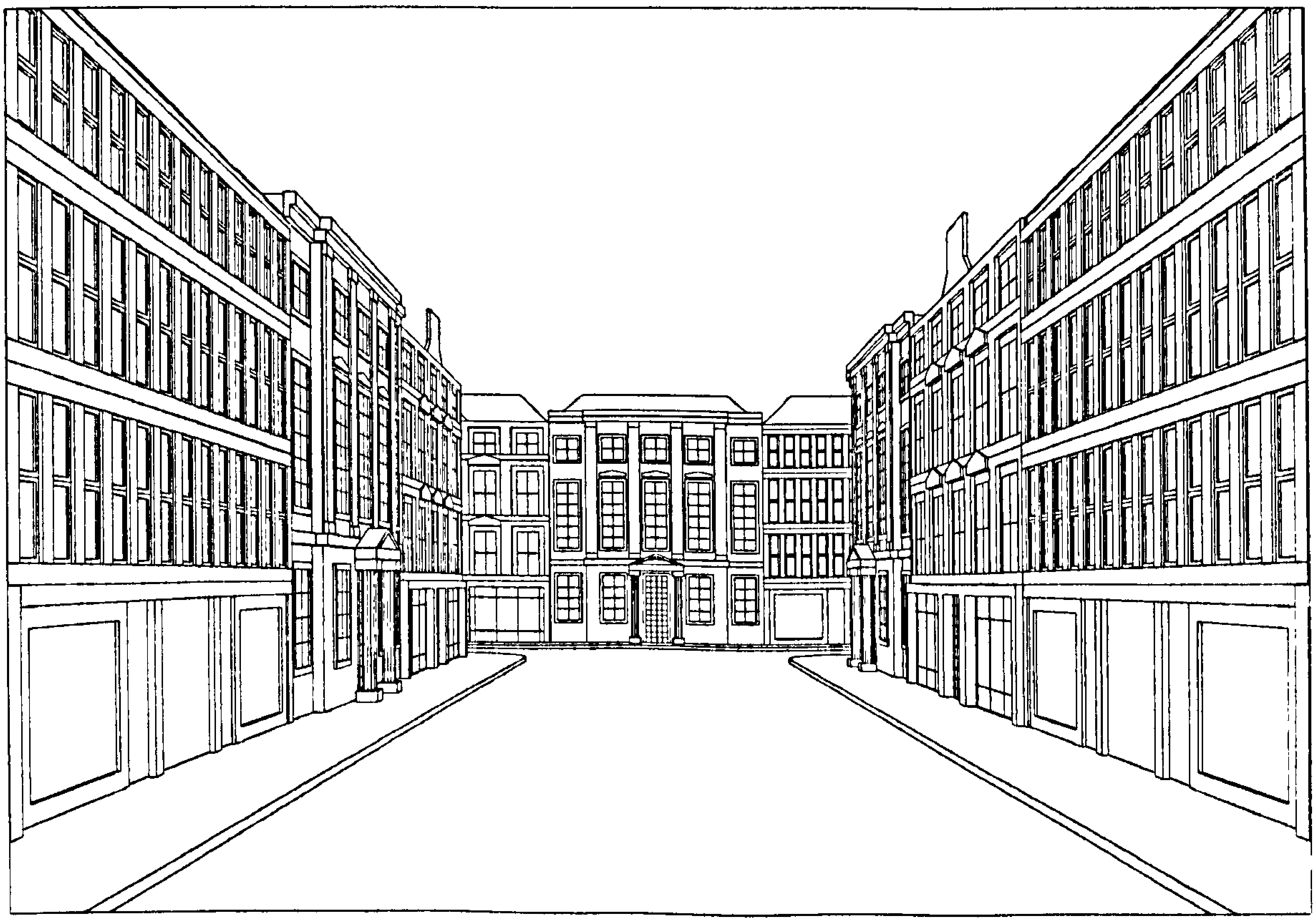


Figure 6.10 Variety- an aggregation of different but similar facades

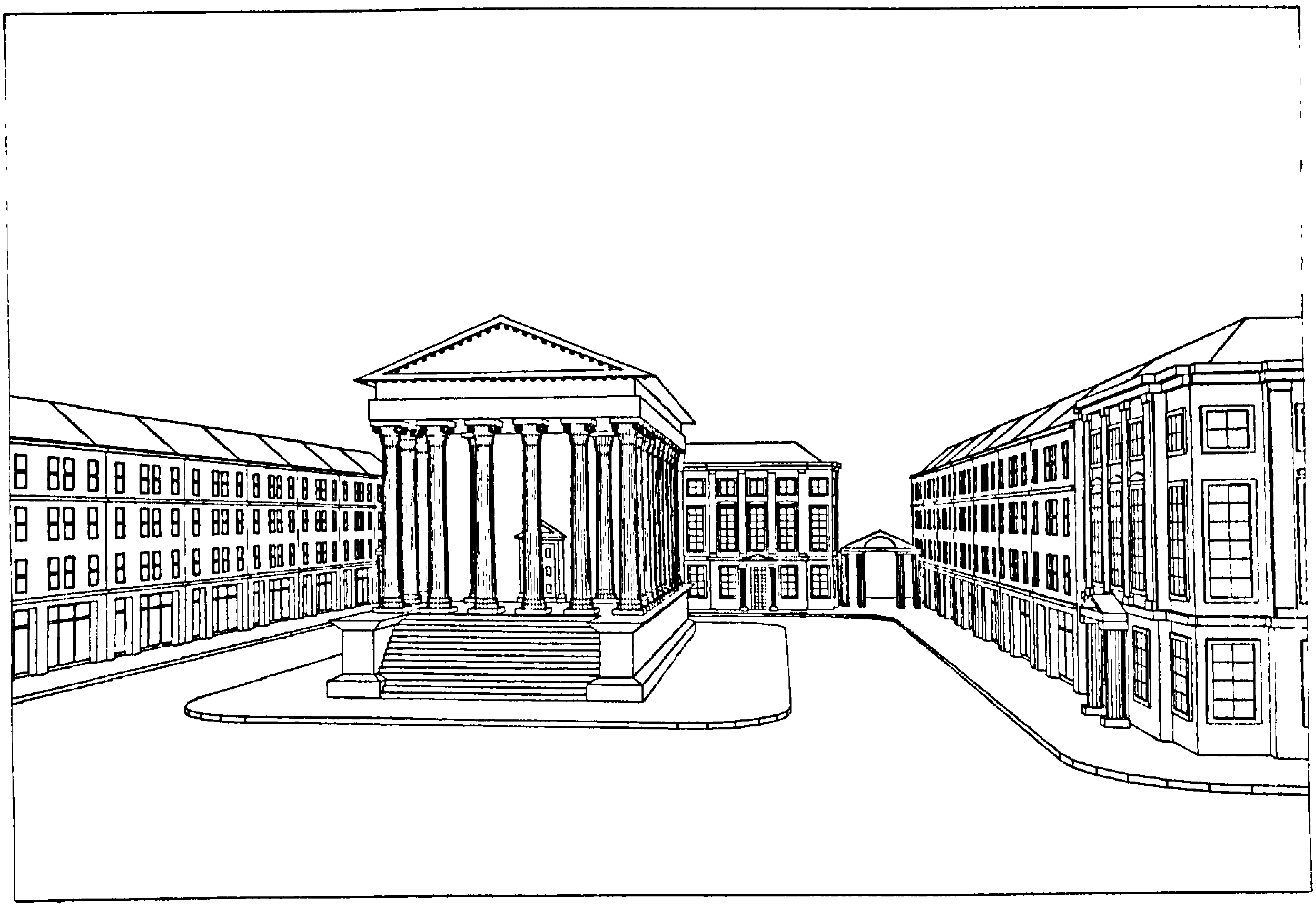


Figure 6.11 Hierarchy- facades framing important public buildings

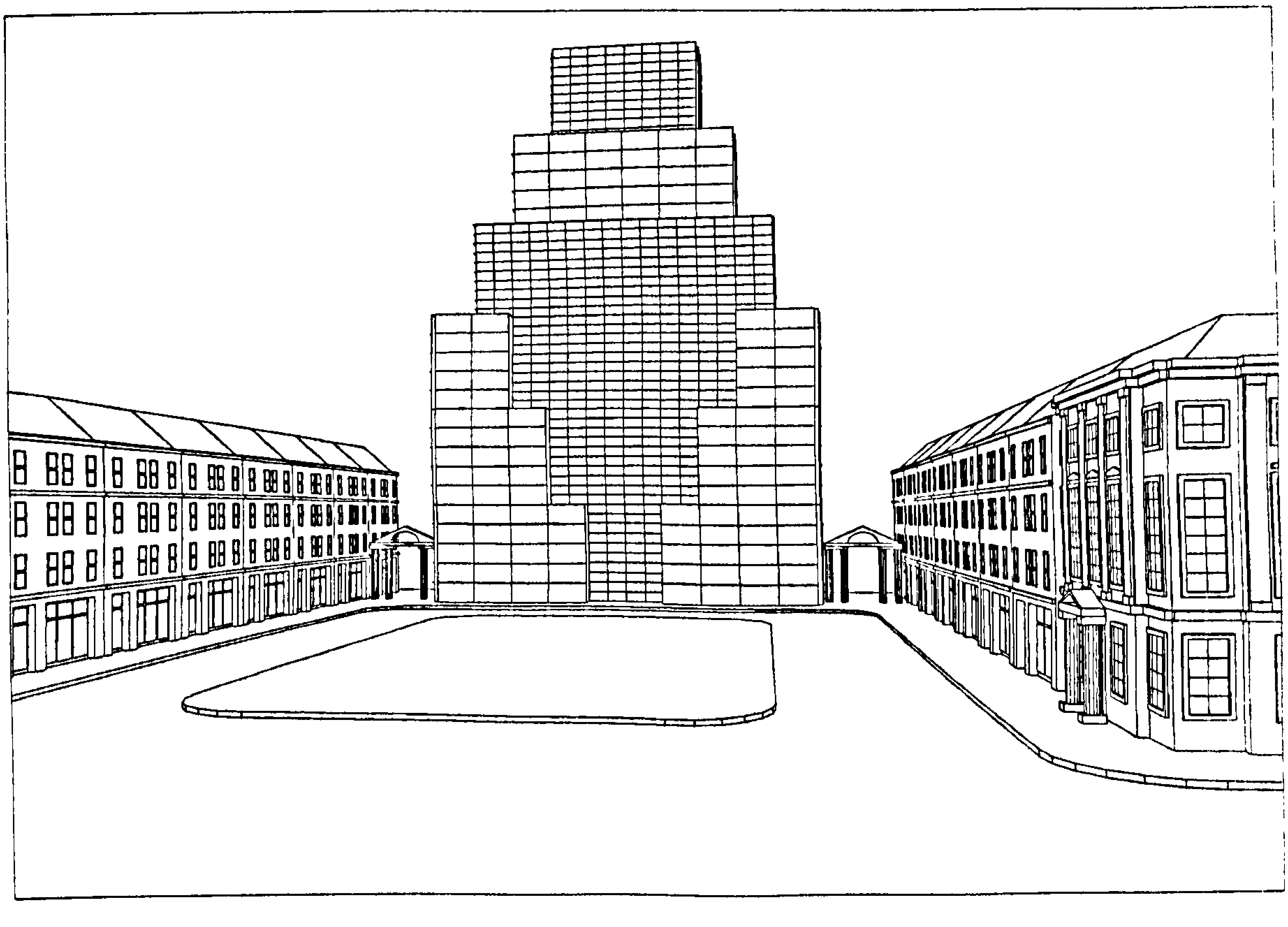


Figure 6.12 Hierarchy- facades framing unimportant private buildings

The experiments previously described were first conducted in 1987, 15 years before the time of writing. It is perhaps unique that not only has the data survived intact but also the software tools required to utilise it. Other than writing a simple utility to convert the VIEWER picture format to generic level 1 postscript in order that it could be rasterised into a TIF image file the software is as useful today as it was then. However, due to the passage of time and despite the continuing utility of this approach this particular form of graphical representation has become dated

Renewed interest has lead to the rediscovery of this methodology as the basis for a series of student exercises directed at exploring visual design issues at the level of the façade, individual buildings and finally at an urban scale. The briefs for these exercises are reproduced below.

6.2.4 In the Service of Urban Design

These exercises comprise a body of work that builds on itself to form a methodology for the efficient and timely design and construction of architectural computer models.

The aim of thisexercise is three fold:

- The capture of architectural form
- The recreation of architectural form
- The exploration of rules of composition

To accomplish these aims a computer based methodology must be designed and implemented. This method will be based on the formulation of a grammar composed of a vocabulary of architectural elements and a syntax that describes how these elements are combined to create the formal composition of the buildings

Working in groups, each team member will nominate a building, each within the near vicinity – a mix of old and new over a range of functional types. The task is to capture all of the formal attributes of the building to an appropriate level of detail so that it will be possible to reconstruct each building as a computer model. These attributes will consist of characteristic dimensions, formal elements, material properties and all other data that is required to formulate a grammar. The collected data will be collated in the form of first hand photography, sketches, and measured quantities as well as derived data from cartography, archive drawings and reference sources.

Figure 6.13 Data capture through digital photography & freehand sketches.

The outcome of this first exercise will be evidence of an exploration of the capture and expression of architectural form utilising the methodology outlined above. The objective is to gain an understanding of the atomic composition of a building at a component level and how rules of representation are used to define the formal appearance of a façade. The syntax derived from this exercise will be combined with a vocabulary of elements, created as a library of components, that can be used and re-used to create different architectures. A secondary outcome is an appreciation of the role of abstraction in computer modelling, especially as related to the level and depth of detail required to represent a building.

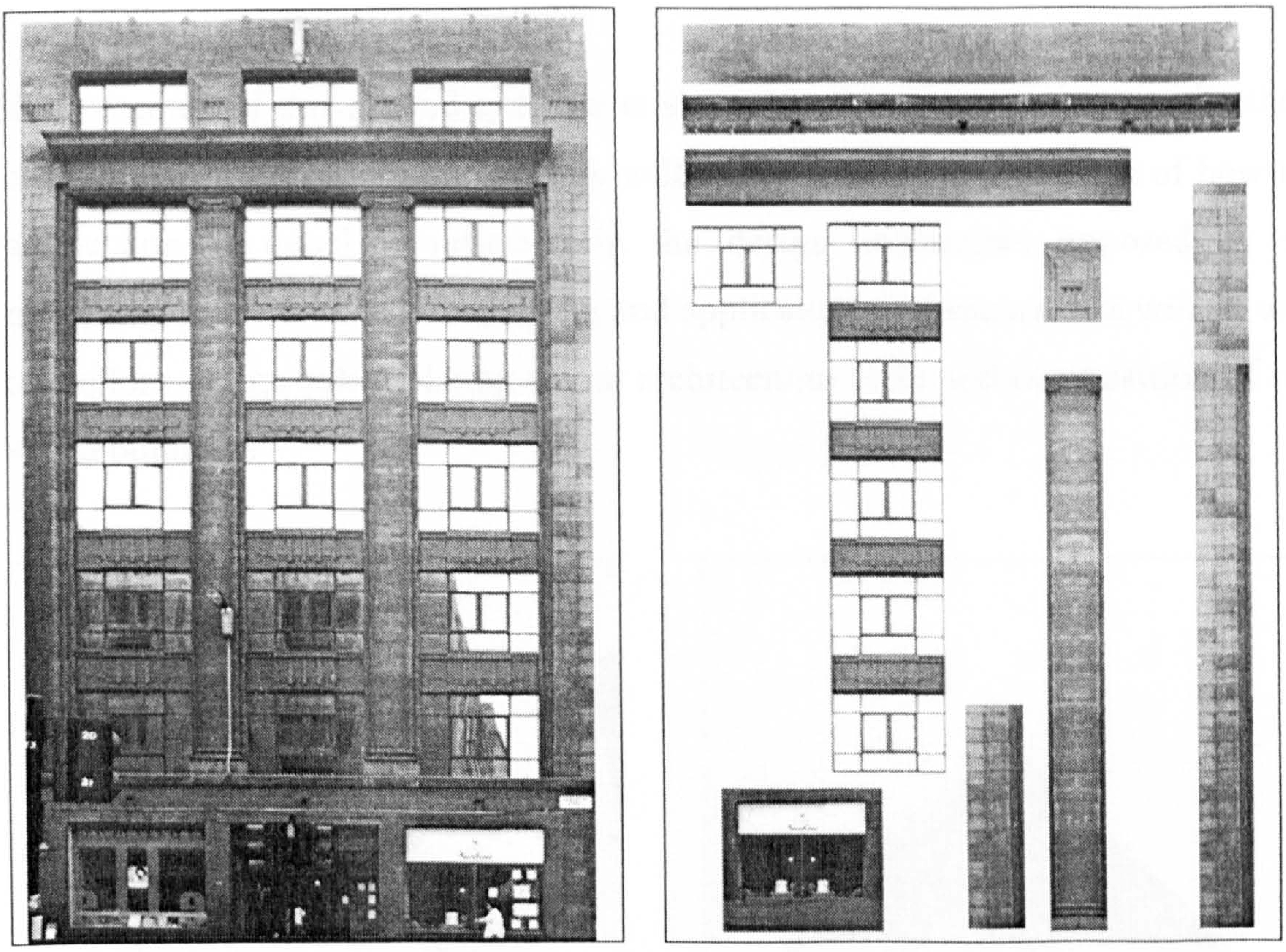


Figure 6.14 The reduction of a facade into a library of components.

6.2.4.1 3D representations

Previously the representation of a building was limited to the two dimensional domain, this exercise seeks to explore the third dimension. The "best" computer models are those that succeed in representing the object by using the minimum of data to the maximum effect. This exercise seeks to derive a protocol where computer models can be designed so as to provide a quick and accurate method of prototyping typical building structures.

In the previous exercise the teams undertook the process of data capture in order to provide that set of formal and material properties which, in conjunction with a description of the stylistic syntax, allowed the recreation of a representation of those buildings in 2-D. This next stage is to explore the route to the third dimension and maps the boundaries that define the trade off between level of detail and computational effort. The vehicle for this exercise is an exploration of the composition of buildings both in isolation and in conjunction.

The outcome of this second exercise is shown in the ability to capture the three dimensional form of buildings. This will encompass an appreciation of how the application of detail is reflected in the design constraints imposed by the modelling process. The construction and application of these models will be used to explore design rules relating to the architectural style and composition of real world buildings

Figure 6.15 Comparison between model & photograph of subject building

6.2.4.2 Urban scenarios

Previously the representation of a building was limited to a single instance of a structure, this exercise seeks to explore the creation of models on an urban scale. Computer modelling has the potential to provide new and exciting means of investigating design principles relating to our urban environment. However the challenge of modelling the city-scape requires a careful approach if the constraints of scale and complexity are to be satisfactorily addressed. This module seeks to explore a methodology within which urban scale computer models can be usefully constructed and employed. The objective is to develop a methodology that will allow the designer to quickly and easily create design scenarios at an urban scale. The key to this process is the ability to reduce the structure of the city to a minimal set of archetypes. – Streets, corners, junctions, spaces. – This is directly analogous to the process of deconstructing and reconstructing buildings in that a library of instances can be developed and deployed to model these situations.

Figure 6.16 The Streetscape – terminated vista

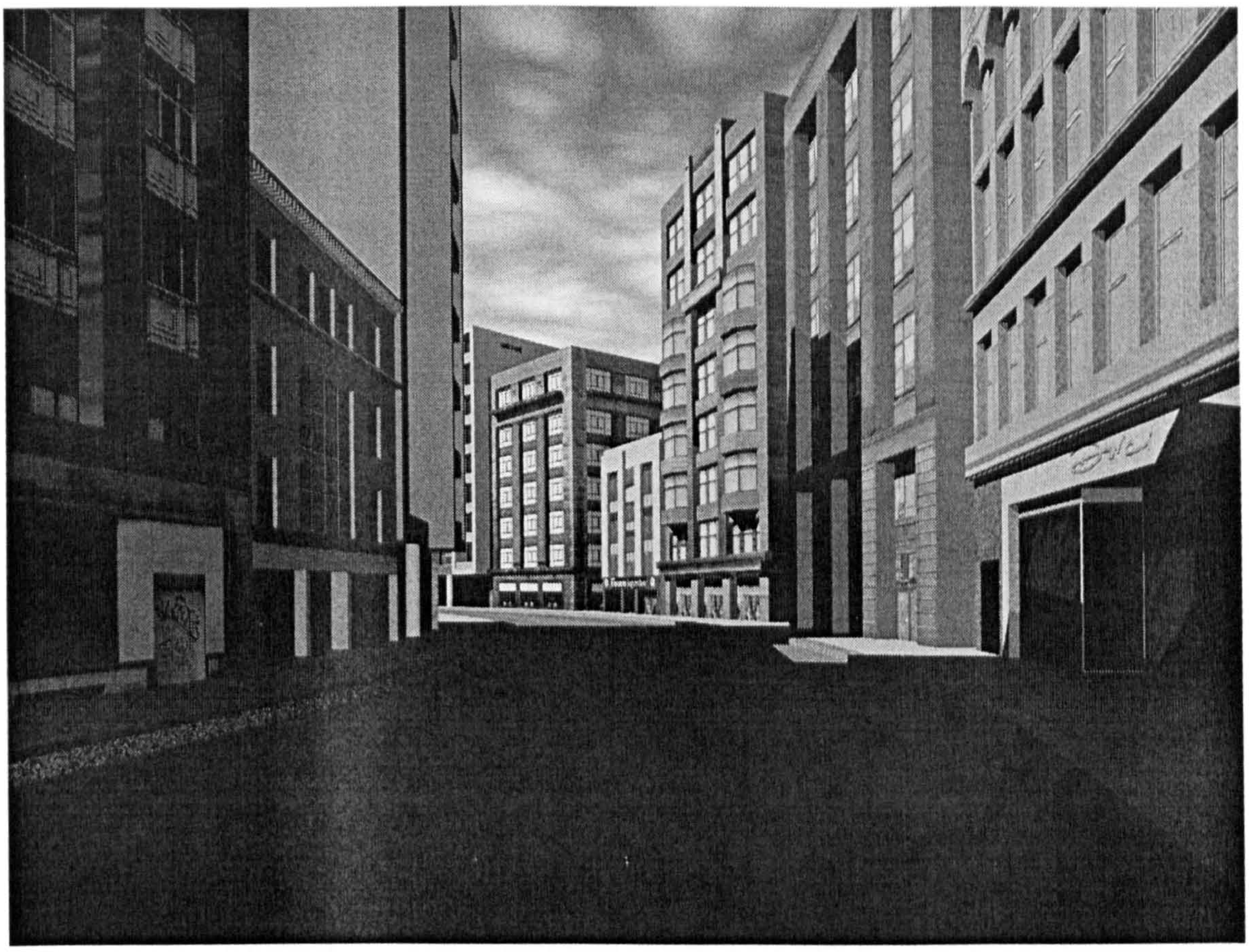


Figure 6.17 The Streetscape – closed views

6.3 Display Formats

Over the period covered by the research reported here, there has been an exponential increase in the performance and capabilities of all types of computer systems. System memory, disk storage and processing power all look likely to continue to multiply at a steady rate well into the future. One aspect of system performance that has only recently reached an accelerated stage of development relates to the graphics subsystem. This is currently driven by the popularity of graphical games software where pleasure can be directly equated with on screen performance. However while the games market might well be the fastest evolving industry sector it is not alone in wanting to benefit from a more complete graphical experience. Any graphical application where the operator seeks meaning or greater value from the display will find an advantage in technology which enhances the experience. This is not a new phenomena and it is perhaps surprising to see how many times this knowledge has been rediscovered and reinvented over the years.

6.3.1 The Panorama

While the scenic image has had a long and varied history in the passage of the arts the credit for the invention of the Panorama is given to Robert Barker who patented the concept in 1787. Barker's patent application referred to the:

> *"invention of an entire new contrivance or apparatus, called by him La Nature a Coupe d' Oeil, for the purpose of displaying Views of Nature at large by Oil painting, Fresco, Water-colours, Crayons or any other Mode of painting or drawing". (EDVEC 2002)*

This 18th Century vision consisted of a 360 degree painting, usually of a landscape or similar "out to the horizon" image, covering the internal wall of a circular building or "rotunda". The spectacle soon became a popular form of mass entertainment and was acclaimed as being able to provide *"not a pale reflection of a distant scene but offering an almost palpable sense of reality".* (Roberts 2002)

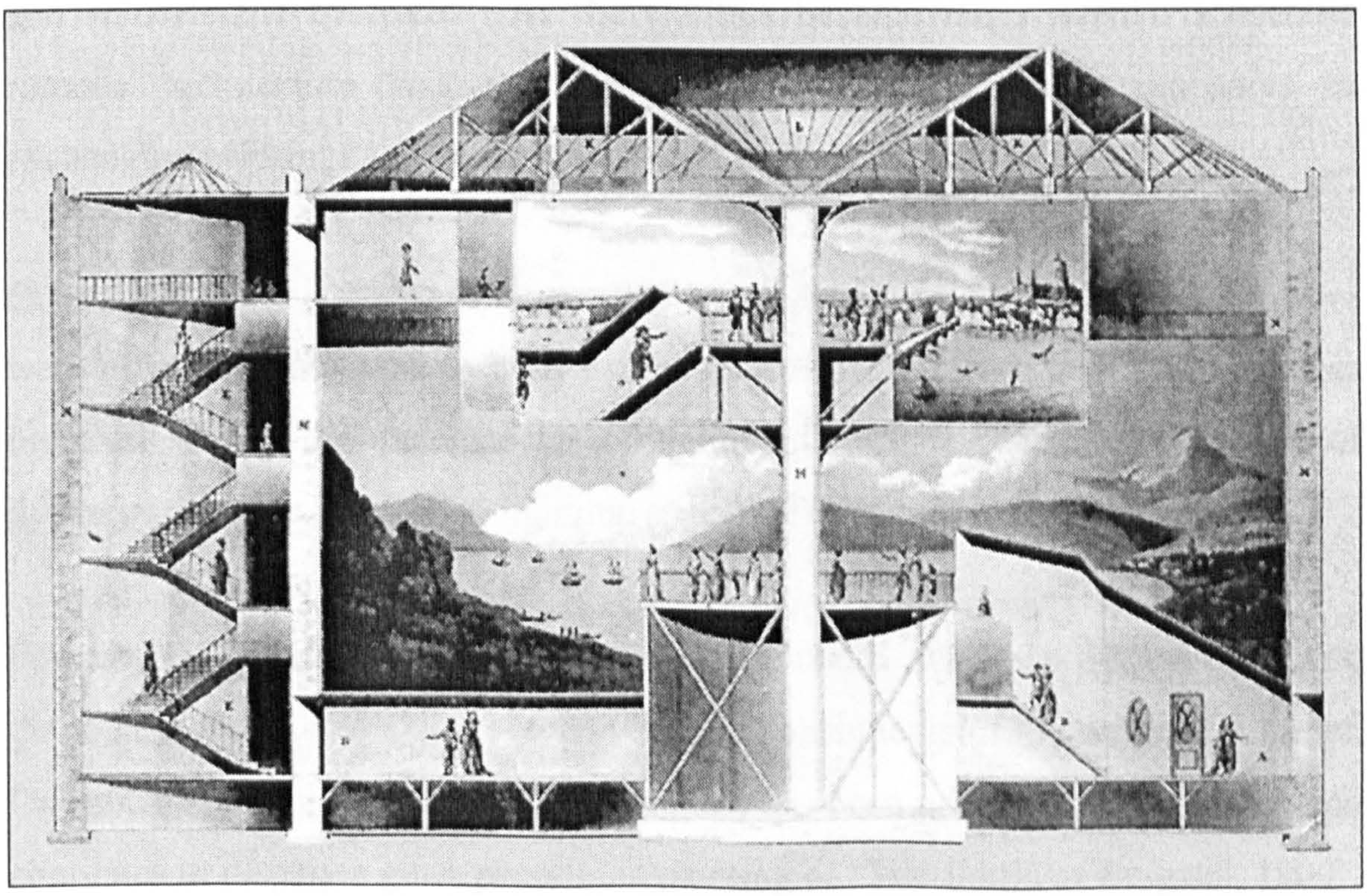


Figure 6.18 Section of Barker's Panorama, Leicester Square, London, 1789

Dating from a time before photography the panorama offered a surrogate reality that was then unattainable in any other form of media and proved to be capable of educating and thrilling in equal measure. For a naïve audience the result was claimed as *"being shocking, the illusion leading to unmanageable visual distortions and in certain instances hysteria."* The educational value was evident in the ability to make the world of distant, exotic and often dangerous cultures visible to those unable to travel, a fact that the advertising was quick to claim was preferable to the real thing.

It is interesting to note that the execution of the concept was not just confined to the image but the space in which it was displayed was also an integral part of the experience. The Panorama was also called the "all view" or "picture without boundaries" and it was found the suspension of disbelief was dependant on removing all extraneous visual cues other than the image itself. The top edge was hidden by the velum a top lit umbrella like roof and at the bottom by an artificial foreground. Access to the raised viewing area was provided through a darkened corridor leading to a spiral staircase which was intended to disorientate the visitor and enhance the spectacle. As the exhibits became more popular a specific rotunda architecture was even standardised thus allowing companies to exchanging paintings and broaden the scope of their listings.

Once the Panorama began to be rivalled by the new science of photography there were a range of attempts to increase the appeal of existing installations. These included seating the audience on a rotating platform and the incorporation of dynamic elements, controlled lighting, special effects and other theatricals.

Few historical panoramas and fewer original rotundas survive today but those that do endure allow us to see the technical accomplishments of the original pioneers. One such example, termed the "Cyclorama", has been preserved in Innsbruck and continues to display a giant panoramic painting of "The Battle of Bergisel, 1809". (Innsbruk 2002)

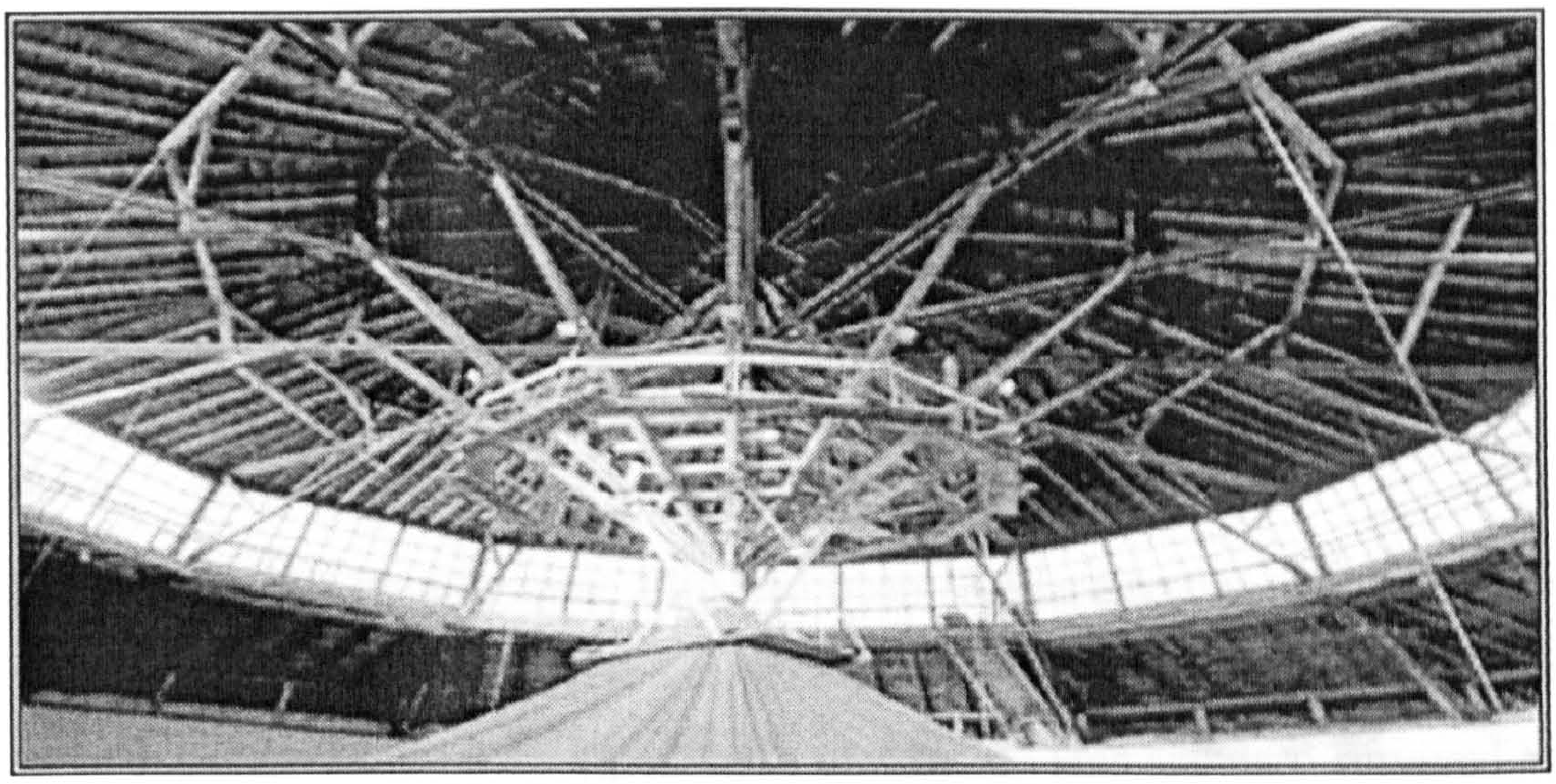


Figure 6.19 A view of the velum at the Cyclorama, Innsbruck.

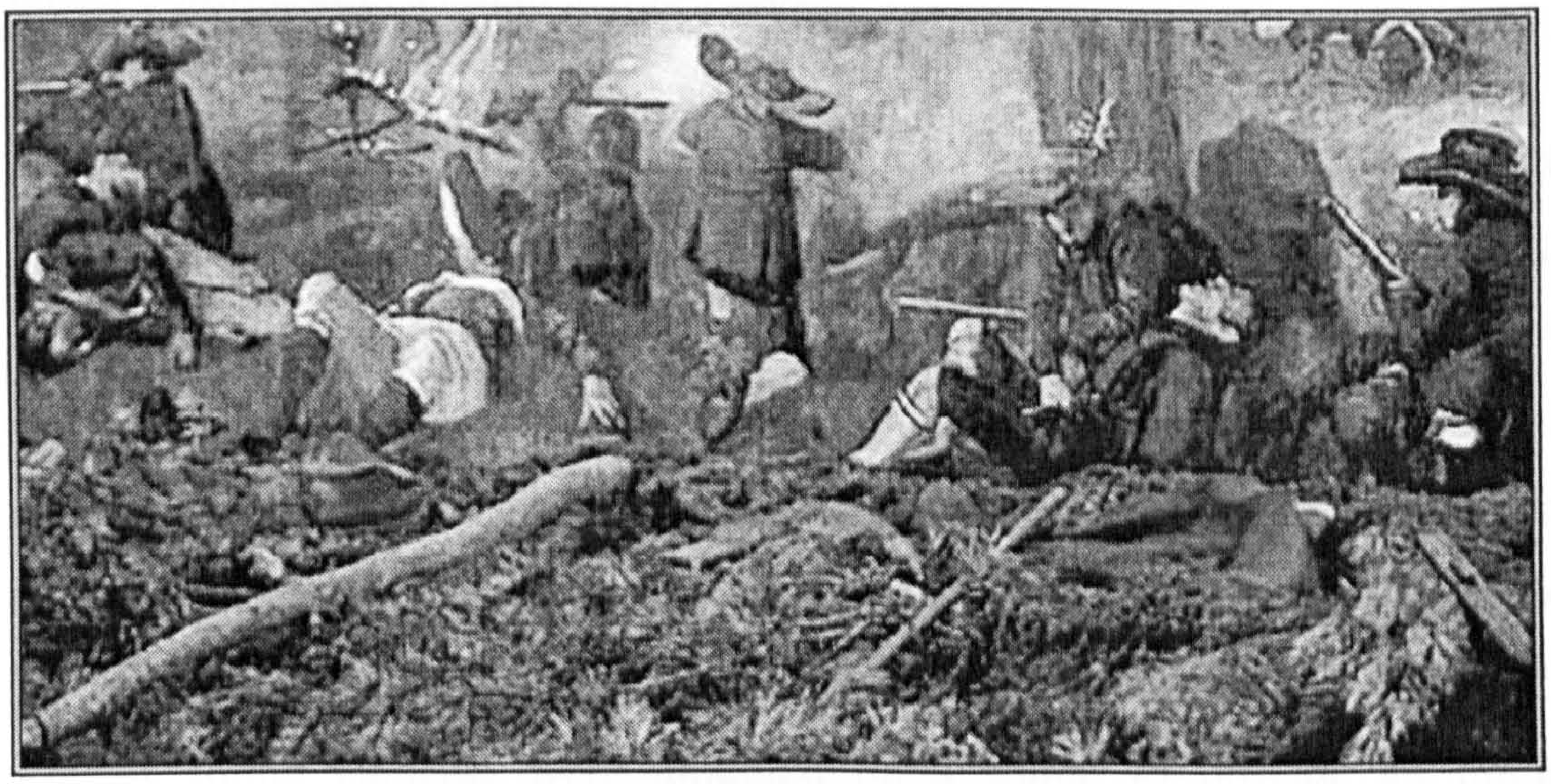


Figure 6.20 A close view showing the false terrain in front of the panorama

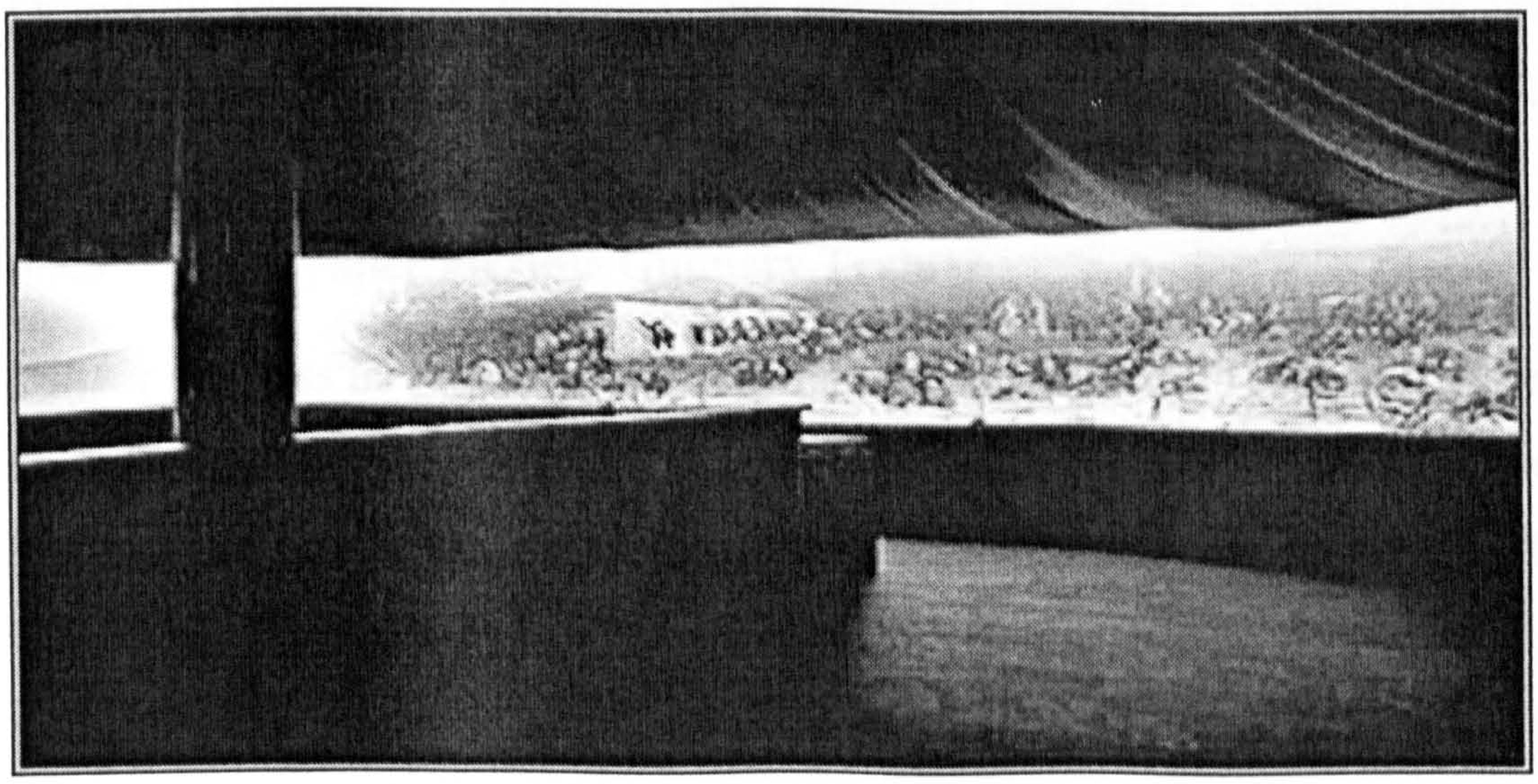


Figure 6.21 The viewing gallery

Eventually the ubiquitous nature of photography and the exposure afforded by new printing methods meant that despite all efforts the Panorama's 100 year reign was over. This however was just one of the concepts many technical incarnations

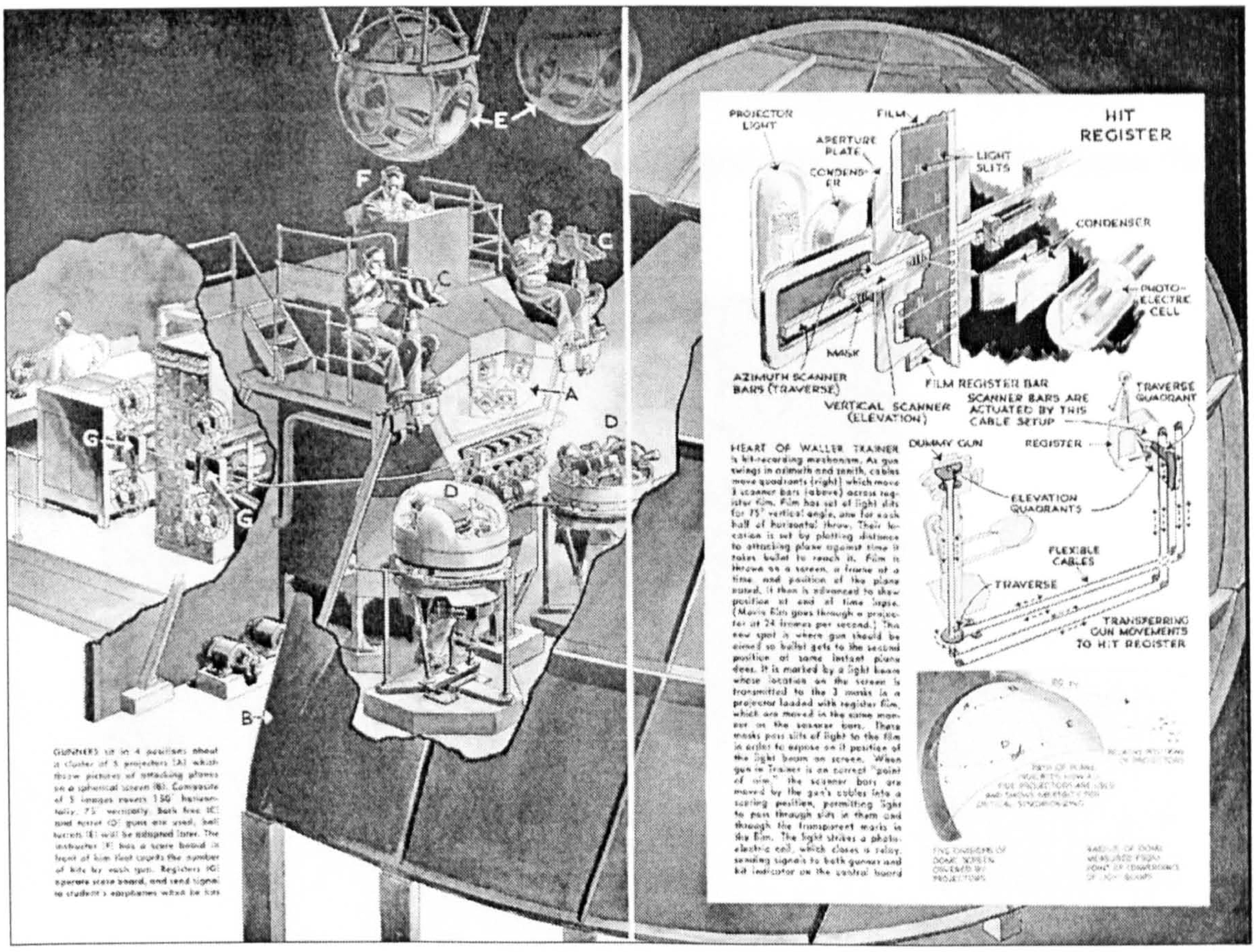
6.3.2 Widescreen Cinema

Since the 1890s photographic film had been standardised on 35mm stock which was drawn vertically through the camera 4 sprocket holes per frame. This resulted in an exposed area with an aspect ratio of 4 by 3. This standard survived into the 1950s and still persists today in that most common television formats and frame buffer configurations retain a 1.33 ratio.

While there were a number of experiments with larger formats it was not until the introduction of television and its influence on declining cinema audiences that the film production companies started to look for a means of revitalising interest in film theatres. The first evidence that the cinema going public would respond to a new experience was seen by their reaction to Cinerama. Cinerama was born out of a series of experiments dating from 1938 when Fred Waller had developed an exhibit for the Kodak Eastman pavilion during the 1939 World Fair in New York. Eleven 16mm cameras were mounted on a frame to provide an immersive environment when the resulting footage was projected on the inside of a dome. Originally patented as Vitarama this spherical projection system was used to display both moving and still images to an appreciative audience. (Hart 2002)

6.3.3 The Waller Flexible Gunnery Trainer

Further progress in Viterama was halted by the Second World War but Waller was commissioned by the US Defence Department to develop the system for use as a gunnery trainer. A five camera assembly was mounted in a target aircraft and filmed successive passes by attacking fighters. Interlocking projectors then projected the film on a hemispherical screen. Three projectors tiled the image over the horizontal area above the horizon while the remaining two provided the overhead fill. All the projectors were electronically synchronised and provided an accurate realtime visual environment.



GUNNERS sit in 4 positions about a cluster of 5 projectors (A) which throw pictures of attacking planes on a spherical screen (B). Composite of 5 images covers 150° horizontally, 75° vertically. Both free (C) and turret (D) guns are used; ball turrets (E) will be adapted later. The instructor (F) has a score board in front of him that counts the number of hits by each gun. Registers (G) operate score board, and send signal to student's earphones when he hits

HEART OF WALLER TRAINER is hit-recording mechanism. As gun swings in azimuth and zenith, cables move quadrants (right) which move 3 scanner bars (above) across register film. Film has set of light slits for 75° vertical angle, one for each half of horizontal throw. Their location is set by plotting distance to attacking plane against time it takes bullet to reach it. Film is thrown on a screen, a frame at a time, and position of the plane noted. It then is advanced to show position at end of time lapse. (Movie film goes through a projector at 24 frames per second.) The new spot is where gun should be aimed so bullet gets to the second position at same instant plane does. It is marked by a light beam whose location on the screen is transmitted to the 3 masks in a projector loaded with register film, which are moved in the same manner as the scanner bars. These masks pass slits of light to the film in order to expose on it position of the light beam on screen. When gun in Trainer is on correct "point of aim," the scanner bars are moved by the gun's cables into a scoring position, permitting light to pass through slits in them and through the transparent marks in the film. The light strikes a photoelectric cell, which closes a relay, sending signals to both gunner and hit indicator on the central board

Figure 6.22 Waller Flexible Gunnery Trainer

Trainee gunners – two manning free mounted weapons and two in powered turrets – were able to track and shoot at attacking planes. The gun mounts were connected by cables to a photoelectric system that matched the altitude and azimuth of the gunner's aim to the position of the aircraft's image, hits being recorded electronically.

6.3.4 Cinerama

After the war the development of Vitarama was restarted but by 1946 additional investment prompted a reorientation of the company to form Cinerama. In the Cinerama process the spherical screen was abandoned and the camera / projector combination reduced to three. A high aspect ratio cylindrical screen subtending

144 degrees in the horizontal and 55 degrees in the vertical was adopted as a compromise between retaining the experience of immersion while also catering for larger audiences. The reduction in the number of cameras / projectors was a concession to the economics and practicality of operating multi-camera systems. The production camera system consisted of a single body with three offset lenses and a central rotating shutter exposing the images onto three separate 35mm films.

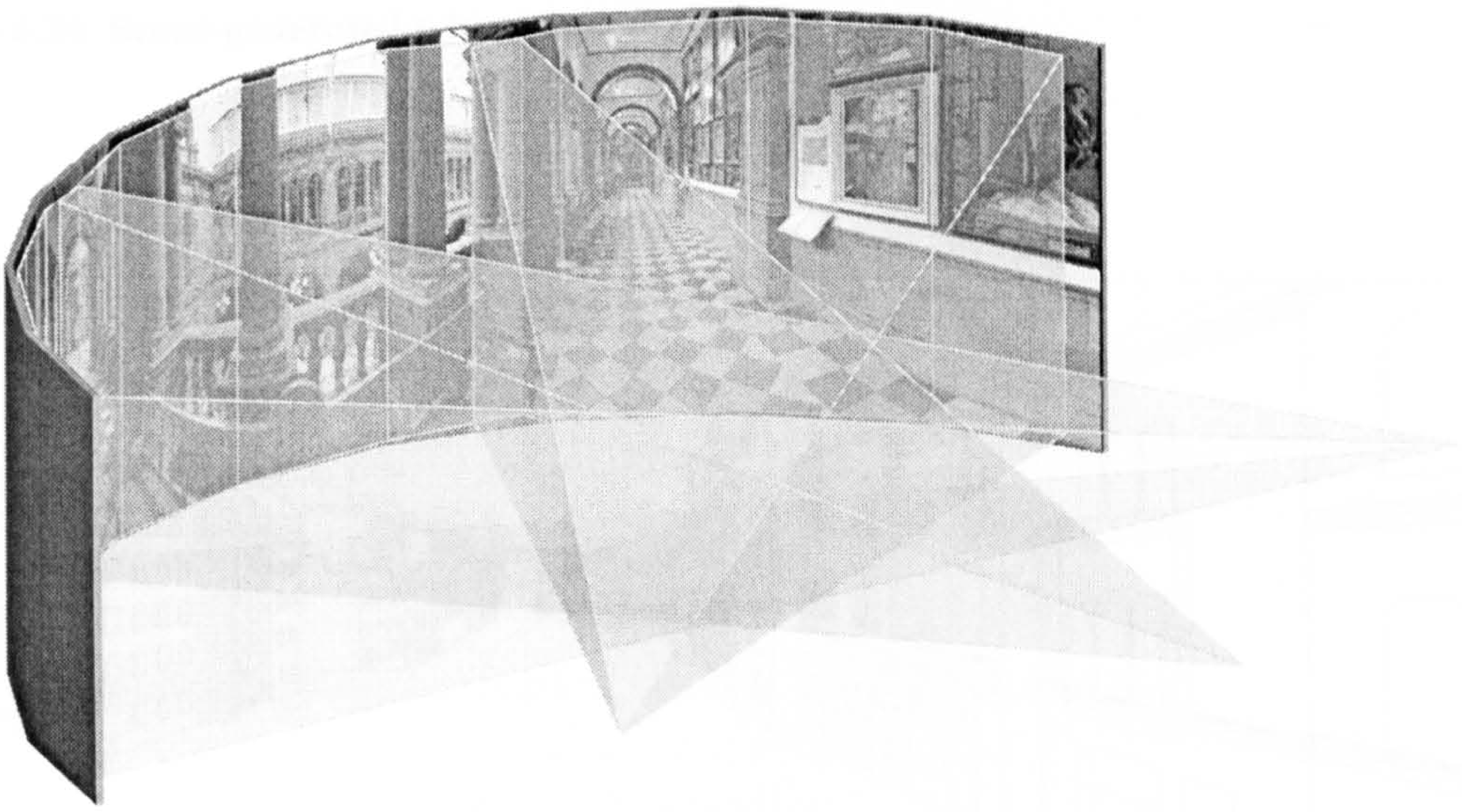


Figure 6.23 Diagram of the Cinerama three projector system.

The use of a wide, curved screen provided the audience with an enhanced viewing experience but introduced a number of technical problems in its implementation. These challenges related to the geometrical distortion of the image and the difficulty of seamlessly merging the three separate image panels.

When a camera captures a scene with a lens offering a wide field of view [FOV] then the angular field of view decreases towards the edges of the image. This is illustrated in Figure 6.24 to Figure 6.26 and shows the growth in distortion as the FOV is increased.

Figure 6.24 Scene generated with 50 degree FOV

Figure 6.25 Scene generated with 100 degree FOV

Figure 6.26 Scene generated with 150 degree FOV

A further problem is manifest when projecting a rectangular image onto a curved surface. Since the image has to travel a greater distance to reach the centre of curvature then it will appear larger at the centre. This effect can be seen in figure 6.xx. These optical problems are significantly reduced by using the three camera / projector system as each image is projected normal to its own third of the screen. The remaining geometrical distortion can be removed using an anamorphic lens in each projector to compensate for the "pin cushion" effect.

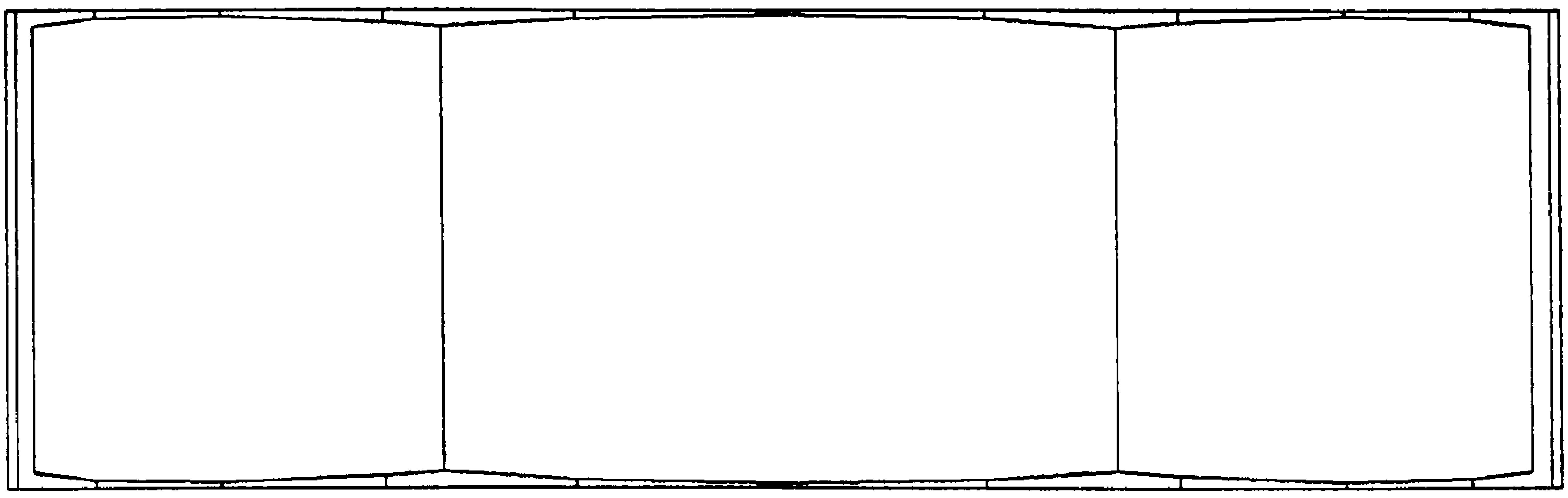


Figure 6.27 Diagram of the geometrical distortion of projected images

The figure above also demonstrates the unequal projected areas of the three panels when observed from a single central viewpoint. The central panel maintains much of its original aspect ration but the two outer zones are severely reduced in width.

The creation of a seamless single image from the three projectors was realised by the addition of a device that provided primitive edge blending between the adjoining images. This was a mechanical device mounted on the aperture frame and operated by a cam that caused a comb shaped shutter to oscillate rapidly, blurring the edge of the image. Without some method of diffusing the common edges between adjacent images the additive effect of the overlapping projection beams causes bright bands between the panels. Inevitably there were problems with registration and variable film stock that lead to less than perfected joins and variations in brightness between the three panels yet audiences seem to have been remarkably tolerant and the system enjoyed great success.

Despite much critical acclaim Cinerama was always burdened by the financial overheads of the multi-camera system. In addition to triple the film production and processing costs the main overhead was incurred by the number of staff required to project the film in a cinema. Each of the six projectors – two for each channel to allow for reel changes – required at least two staff in addition to a sound engineer to manually pan the soundtrack as required and a front of house engineer to continually fine tune the image balance and registration between image panels. This made it an expensive format for the cinema operator to run even after the $70,000 dollar capital costs of refitting the cinema in the first instance.

6.3.5 Single Lens Formats

While the benefits to an audience of the wide screen system were undeniable the crippling cost of the process lead the production companies to look for a more economical method. This economy was sought in a compromise process where much of the technical quality of the three strip original was replaced by a single lens format.

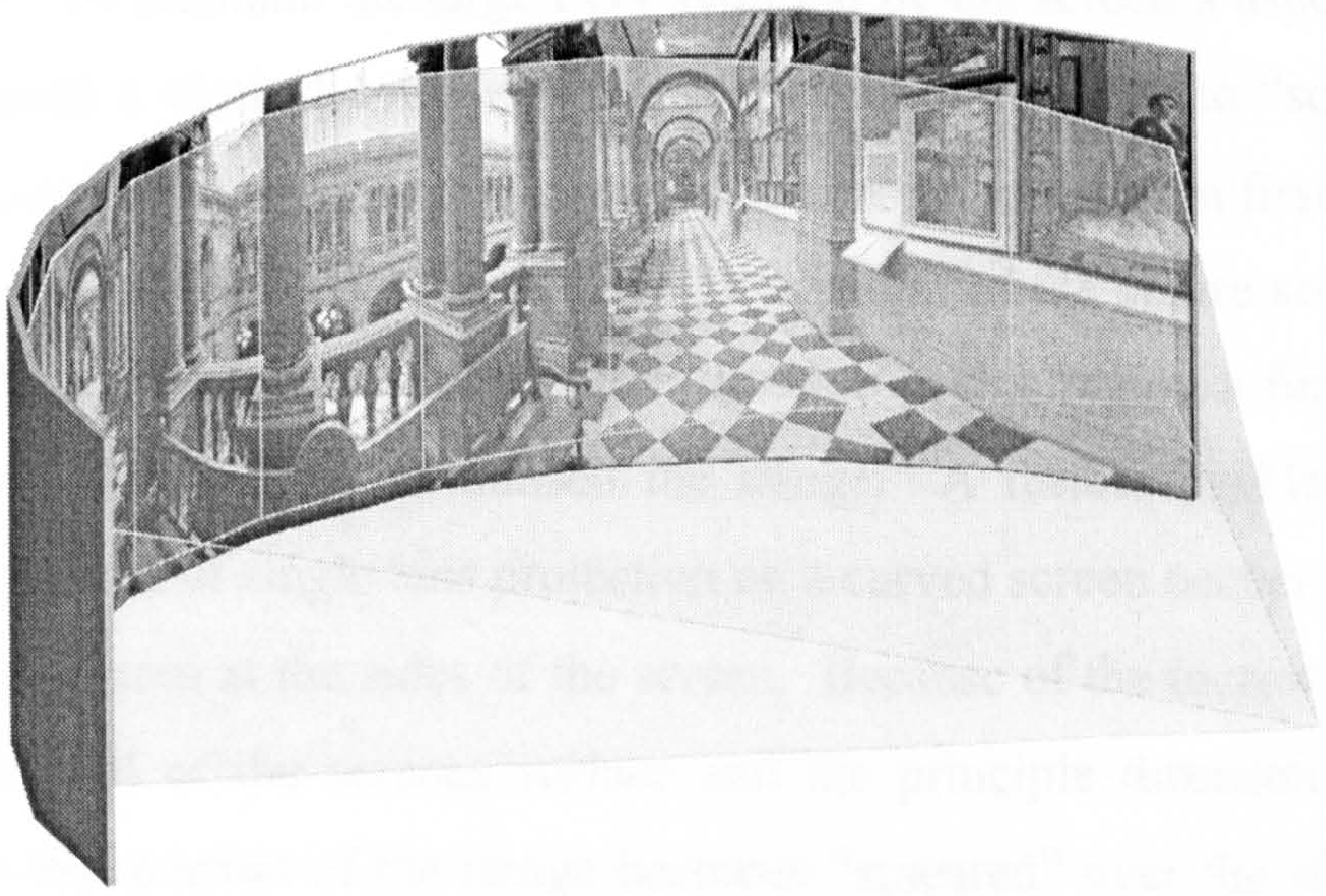


Figure 6.28 Diagram of the single lens widescreen process

Using a single lens exacerbated many of the geometrical problems relating to the display system. Figure 6.29 shows the magnified image distortion that results from keeping the projector normal to just the centre of the curved screen.

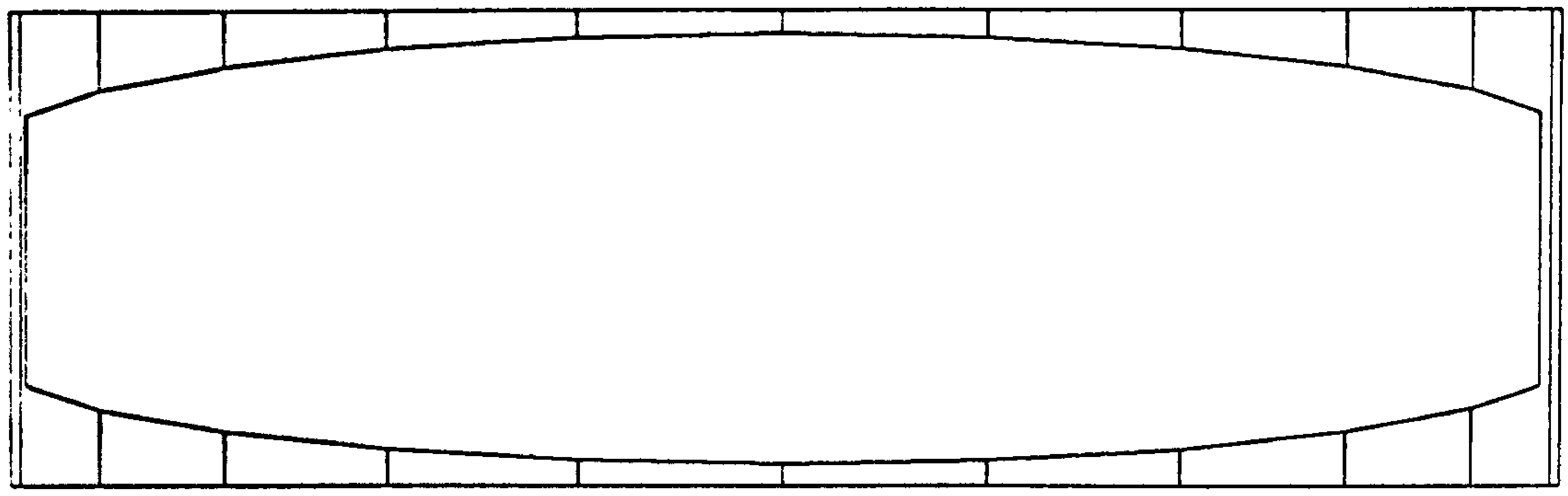


Figure 6.29 Geometrical distortion due to single lens projection

This distortion is inevitable when projecting a straight edged fustrum onto a curved display screen. The solution was found in the use of more radical anamorphic lenses that sought to apply an inverse distortion that would counteract this problem. A normal camera lens in spherical in form but an anamorphic lens is a composite that makes use of a series of cylindrical optics to give the desired effect. To maintain the large FOV required by the screen's aspect ratio the taking lens used a vertical cylinder within the anamorphic lens to "squeeze" the image onto a standard negative. The lens on the projection system first sought to remove the "pin cushion" effect by reducing the height of the centre section of the image and then to compensate for the minification in this region a further barrel lens in series with the first normalised the image. A further problem peculiar to the combination of single lens projection on a curved screen occurs due to the reduced projected area at the sides of the screen. Because of the increasing angle between the normal of the screens surface and the principle direction of the projection beam that portion of the image becomes "smeared" over the display surface. To counteract this a further non-linear "squeeze" has to be applied to the image at either side.

Figure 6.30 Single frame showing anamorphic non-linear squeeze

The success of the widescreen format is undeniable and while the original and superior Cinerama concept became marginalised due to excessive production costs the anamorphic single lens process lives on today under a wide range of formats.

6.4 Digital Display Technology

Despite all these advances it is inexplicable that the technology commonly used to interface with the computer has been subject to such a minimal development over the recent period. Ever since the replacement of storage tube terminals by raster scanned CRT devices in the mid 1980s there has been no real progress in this field. Resolution, scanning frequency and the wide spread availability of deeper bit-depth frame buffers might have become common place but the viewer is essentially still restricted to visualising output in the same configuration as was available at the outset of computing technology.

Recent developments suggest that there is a growing move away from the conventional viewing paradigm and as techniques, technologies and applications become more wide spread the proliferation of alternatives will become more widely available. This is largely due to the decreasing cost/performance ratio of all the technologies that must be brought into conjunction in order to realise a more advanced concept.

6.4.1 Projector Technologies

Large screen video projection has been pioneered in two original application areas, entertainment and military simulators. Since these initial developments the technology has gained wide spread acceptance in training applications, business displays, conference presentations and corporate boardrooms. The primary medium for these displays was video but more recently projectors have become capable of displaying real time computer images with all their attendant benefits.

There are two main design philosophies employed in projector construction, cathode ray tubes (CRT) and liquid crystal diodes (LCD). A CRT projector generates the image by scanning an electron beam onto the face of a phosphor coated tube. The excited phosphor then glows forming an image much like on a conventional monitor. Typically three tubes are employed, one for each primary colour, each beam being converged on the image plane to form a single image in the additive colour combination. An LCD projector forms the image by creating each individual RGB component on separate panels and then by reassembling the image through an optical system emitting a single beam.

Digital Light Processing (DLP) is the newest technology which has yet to make an impact on the market. Images are created by millions of tiny mirrors each mapped to an individual pixel. The mirror is capable of being modulated at high speed reflecting light through the optics and onto the screen. The benefits promised include significantly greater light output and better colour rendering. Although LCD projectors are currently both cheaper and brighter than the more expensive and complicated CRT projectors they suffer from a lack of image contrast and poorer geometry control characteristics.

6.4.2 Reinventing the Widescreen

The distinctive feature about these new display modes is that the image is available to an audience, previously computer graphics could only be experienced through a single monitor or perhaps a VR headset. With the advent of a change in philosophy that extends the concept of using the computer purely as a production tool and allows its use as a presentation tool then the ability for more than one individual to share in the experience had added advantages.

To create an environment within which an audience can access the computer generated output requires the image to be scaled to some extent. As the capabilities of conventional monitor construction preclude their use in this context, other than by constructing a wall of multiple units, then the use of a projection system becomes the natural choice. While there are many available configurations of physical environment the design and implementation of the most appropriate configuration is dependant on both the content to be viewed and the manner in which the audience is expected to perceived it.

As no one configuration is able to optimally address all the requirements of all the possible modes of interaction there are a range of solutions available. The choice between these techniques is generally made on the basis of whether the virtual content is best visualised between the viewer and the horizon or whether it is at or near arms length distance. Naturally architecture, certainly when it reaches the scale of the city, falls into the first category while engineering components and similar objects meet the later criterion.

The next choice to be made relates to the surface onto which the image will be projected. When specifying a display system the decision is between flat, cylindrical or spherical screen types. Flat screen configurations lend themselves to pixel mapped data where the projected image is a rectilinear, two dimensional projection within which all edges and boundaries are held to be parallel and perpendicular. If the image is generated through a perspective projection of a three dimensional data set then a flat screen will exhibit severe limitations where large fields of view are displayed. This is due to the both the increasing angular separation of parallel horizontal and vertical lines towards the boundaries of the image due to the increasing separation of the viewer and the image plane. To compensate the screen can be divided into elements and wrapped around the eye point, however the technical difficulties of achieving a satisfactory edge blend incurred by this method mitigate against its use other than where there is a natural discontinuity between the screens such as glazing bars in an "out the window" view.

The use of a cylindrical screen will solve these problems but again there is a problem with generating a geometrically true scene from a frustum with a large field of view. Similar optical problems come from trying to project an orthogonal image onto a curved surface. Both these limitations can be overcome by dividing the scene into multiple channels, usually three, and generating the image from a single eye point while incorporating a suitable offset in azimuth between each channel. Display surfaces subtending an azimuth angle of 150 degrees and an altitude angle of 50 degree can be made utilizing this method. This is directly analogous to the original Cinerama concept.

The fact that the screen and its image are wrapped around the audience allows a further benefit regarding the user'perception of immersion within the scene. Because the image fills a large proportion of the viewer's own field of view the senses are tricked into providing matching perceptual cues as if the environment were real. This greatly enhances the audiences feeling of connection with the content on display.

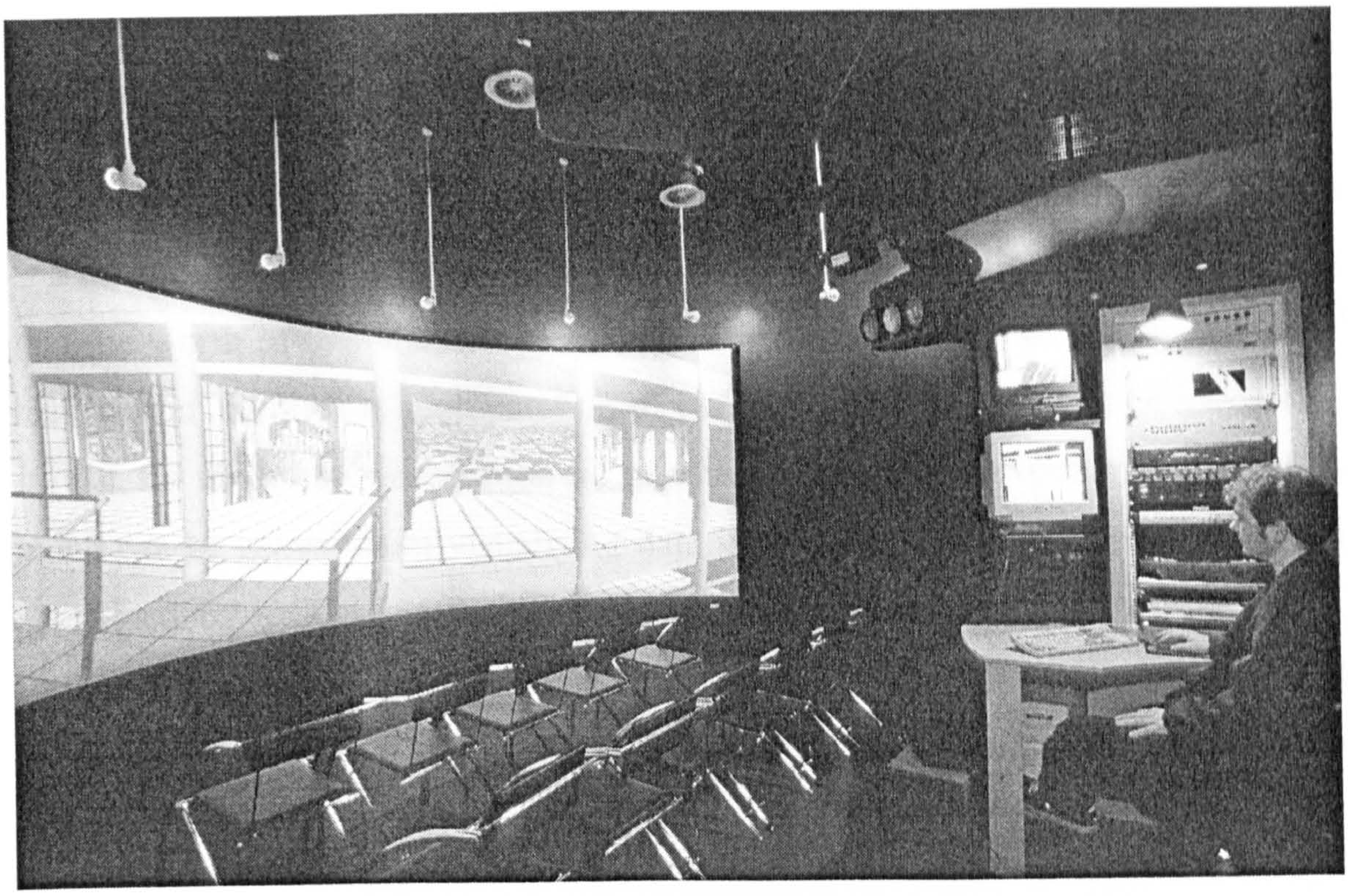


Figure 6.31 The Virtual Environment Laboratory

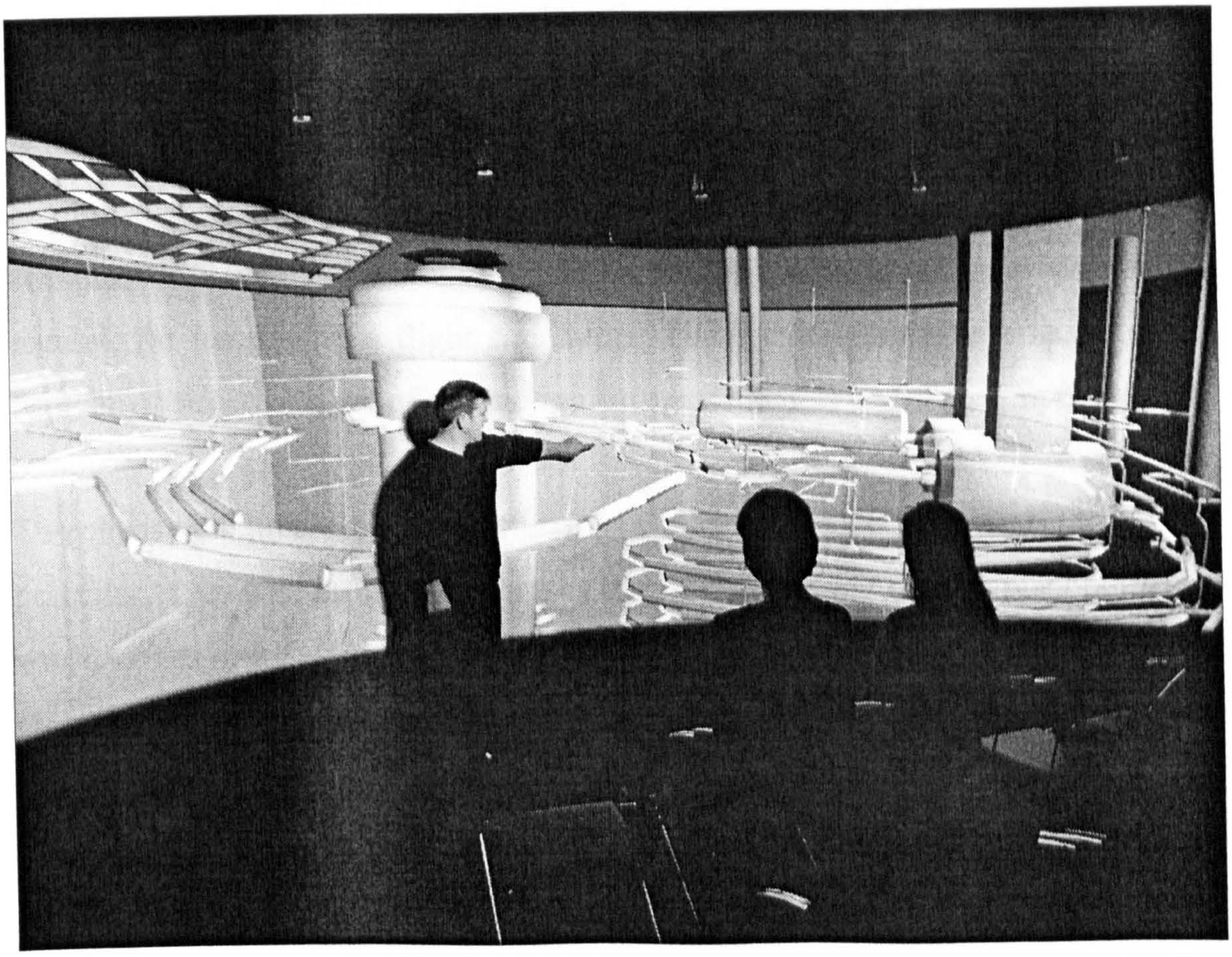


Figure 6.32 Sharing graphical immersion with an audience

6.5 User Interaction

In normal circumstances the creation of a sense of immersion within a VR system is a function of the perceptual cues generated solely by the display system. Given that the worth of this technology is embodied in the pursuit of a better sense of reality then any further augmentation of the user's experience must be harnessed. In a basic mode the observation of graphics is a passive experience and while much can be gained from visualising data in an enhanced environment more can be gained by seeking to improve the control and manipulation of the scene. Research has shown that if the observer's interaction with the scene is subjugated to the constraints of an artificial metaphor then much of the perceived realism is lost.

A mouse may be a convenient control device but only because every computer has one and it is a mechanism with which most users will be familiar. Since the mouse has only two degrees of freedom the consequence of its movement cannot map onto a literal translation in the virtual world. This then requires a layer of abstraction where some set of software algorithms takes the cursor position and in combination with key or button presses transforms these actions into motion. Joysticks make for a better approximation but only where that form of control is natural, for instance in a flight simulator. These examples of interfaces tend to form an obstructive layer of abstraction between the user and their sense of reality.

To enhance the perceived impression of reality an interface must mimic real world practice. The most important factor is that it must recognise and respond to real world constraints. If the means of locomotion is some form of vehicle then it must "drive" as it would in reality but if the metaphor is that of some means of human locomotion then the opportunity for kinesthetic feedback is much greater. The disadvantage of a passive interface is that there is no energy input into the system, with a human locomotion interface effort is rewarded by progress. This results in more realistic navigation and an enhanced appreciation of the spatial layout. It has also be demonstrated that locomotion helps to calibrate vision and can improve any decision making processes that rely on the judgement of distances or spatial qualities. (Rieser 1995)

A system that utilises this philosophy has been developed at ABACUS. Although directed towards assessing issues of access for the disabled this projected employed the notion of utilising an energy extractive interface to provide the greatest sense of realism and engagement with the scene. (Grant 2000)

6.5.1 System Overview

The entire assembly is comprised of a number of mutually inter-linked components that combine to form the two systems on either side of the divide of reality. The locomotion interface, of which the user is an integral part, remains on the real side of the boundary while the graphical simulation complete with data and algorithms remains virtual.

Each of these components are situated diametrically opposite on a feed back loop. They communicate action and reaction only over one half of the circuit, the other half is connected by the users optical and proprioceptive senses. By closing this part of the feedback loop with a human rather than a further pair of sensor connections, it was found that any minimal latency or hysteresis inherent within the rest of the communications path can be compensated for by the user. Both interface and simulation have a dual task to perform in that they must simultaneously provide the input for the other while responding to the current condition.

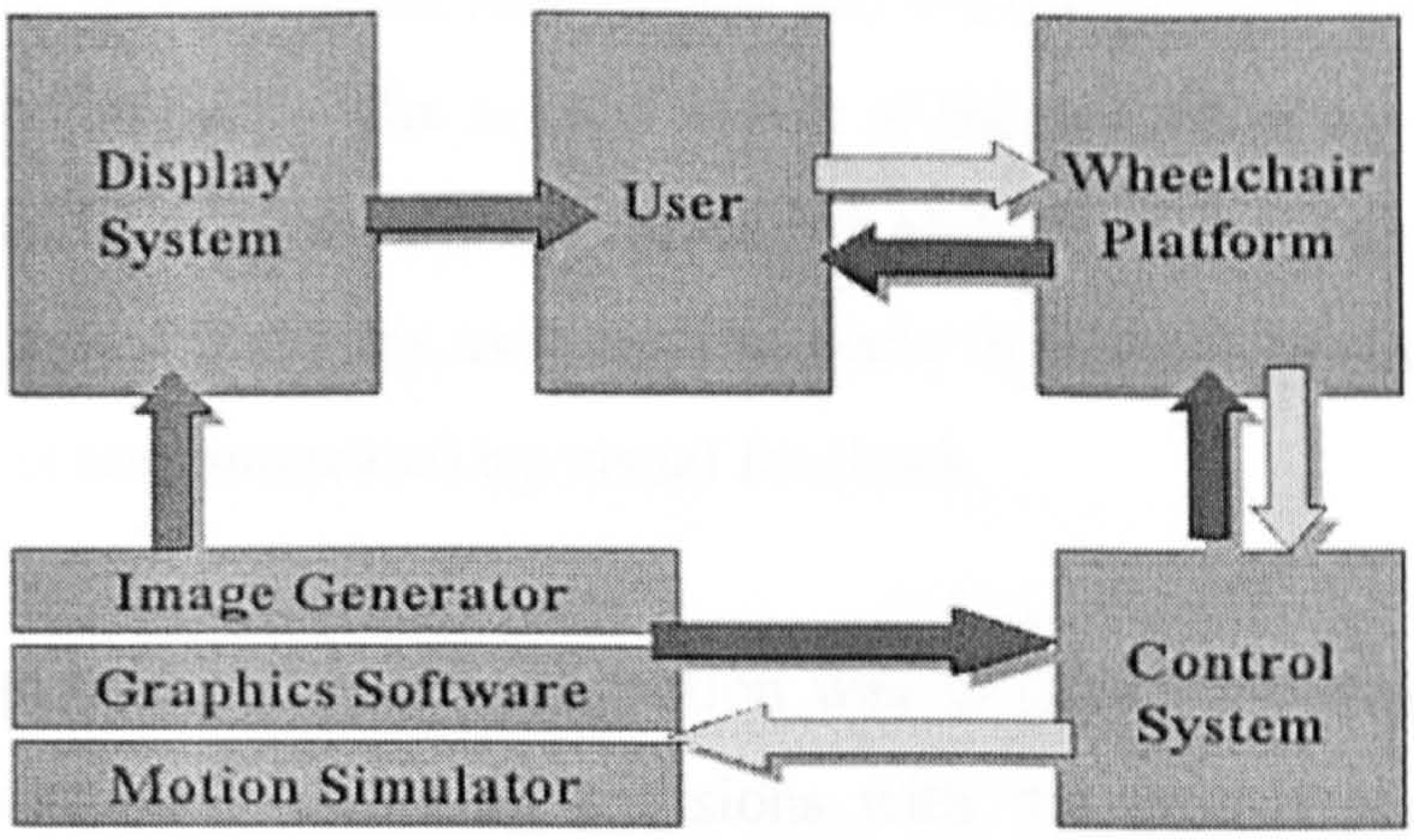


Figure 6.33 System overview

6.5.2 Physical Interaction

The wheelchair is statically mounted on a platform with rollers under each driving wheel. In this way the user is not limited within the virtual environment by the constraints of their physical environment. The normal manual input to each wheel turns either roller, the rotation of which is sensed by rotary encoders. At a set sample interval the individual incremental angular displacement of both wheels on the motion platform is measured. This data is passed to a control algorithm where the current data values are compared to the previous increment, to determine whether either wheel is rotated forward or backward. The basis of the motion control algorithm is the determination, through an analysis of similar triangles, of any translation and also, using the location of the centre of rotation along the rear axle of the virtual wheelchair, the angle through which it is turned. These values are then passed to the graphics system where the transformation of the eye point and rotation of the view vector can be determined.

6.5.2.1 Virtual Interaction

The graphics application requires the Cartesian co-ordinates of the eye point, plus the yaw, pitch and roll angles of the direction of view. Given the yaw angle the remaining two angular parameters can be calculated based on the wheelchairs attitude on the ground plane. In the database traversal three rays corresponding to the contact patch of each of the rear wheels and the midpoint of the front axle are intersected with the floor. The normal vector of the ground plane at these points can then be used to calculate the roll, pitch and altitude of the chair and hence the corresponding view. Since the user remains static in the real world the attitude of the wheelchair is communicated by visual feedback.

Another vital part of real world interaction was to detect collisions between the user and obstructions. Detecting collisions with the ground plane provides a semblance of gravity, collisions with objects reinforces reality by not allowing the viewpoint to pass through buildings. A separate class of collision detection was implemented to perform this task where a matrix of randomly "jittered" rays was aligned with the current direction of motion. This proved to be capable of detecting collisions between the wheelchair and even relatively narrow vertical or horizontal obstructions. The length of these rays was varied according to the distance traversed by the wheelchair between frame updates.

The same intersection procedure can also be used to identify the surface under each wheel, this information then being used to index material properties, such as rolling resistance and surface texture. This information is passed back to the platform control system where it is used to calibrate active input into the system. Varying floor surfaces are simulated by altering the resistance to motion of the rollers, using the brakes. Simulation of the wheelchair moving down a slope requires active input into the system, providing a torque against which the user can control their movement down the gradient. Similarly a resistive torque is supplied when ascending a gradient. A variable torque motor is used to provide this input for different grades of slope. When the graphics system detects a collision event a flag is passed back to the control system which then resets the users position to the last "non collided" location and cycles the brakes on the rollers. This temporarily locks the wheels on the wheelchair and coincidentally provides convincing audio feedback.

6.5.3 User Evaluation

In evaluating the system volunteers were asked to navigate around the VR model and on completion were asked to rate how realistic various features of the simulated environment compared with 'real life' experiences/expectations. In total thirteen features, Table 6.1, relating to the performance of the haptic interface and the simulation of wheelchair mobility were rated by each user. This was achieved by use of a questionnaire in which each user was required to rate how realistic each feature was. A simple rating scale was used to achieve this.

Interface Feature	Rating Scale Response
A. Negotiation of kerbs	1. Unable to judge
B. Wheelchair coasting/freewheeling	2. Very unrealistic
C. Negotiation of doorways	3. Unrealistic
D. Turning	4. Moderately realistic
E. Negotiation of corridors	5. Realistic
F. Effort for propulsion	6. Very realistic
G. Changes in floor surface	
H. Ascending slopes	
I. Descending slopes	
J. Perception of motion	
K. Quality of VR image	
L. Collisions with objects	

Table 6.1 List of features rated by the users and scoring used

Data from the completed questionnaires were then collated and the results are summarised in the histogram shown in Figure 6.34. Out of the 15 subjects, a single user reported sensations of nausea due to motion sickness and this subject was unable to complete the evaluation. The data from this subject has been included in Figure 6.34 but the majority of responses were in the 'unable to judge' category. Importantly, when asked about susceptibility to motion sickness the subject indicated that he often became ill during car, train and boat journeys. However, the remaining 14 subjects reported no incidences of motion sickness and their feedback from the questionnaire provides an important pointer to the performance of the system.

The features rated are listed within the histogram on the x-axis in ascending (from left to right) order of 'realism'. Thus on average the negotiation of kerbs was the least realistic feature of the interface with collisions rated as the most realistic feature. The method of depicting a wheelchair traversing a kerb in our system relies exclusively on a transient visual stimulation that correlates with the change in wheelchair tilt that would accompany rolling off or over a kerb. There is no additional haptic feedback provided as the platform itself is stationary and does not have a tilting mechanism incorporated in its design. It is because of this lack of non-visual feedback that features such as kerbs are poorly rated. However, note that the average response for this feature is 'moderately realistic' . Features that utilise haptics such as changes in floor surface, changes in slope and collisions are all considered as realistic and it is the combination of the sense of effort experienced by the user together with an accurate visual representation of expected motion that provides the perception of reality and is a vital component in making the overall VR simulation truly immersive.

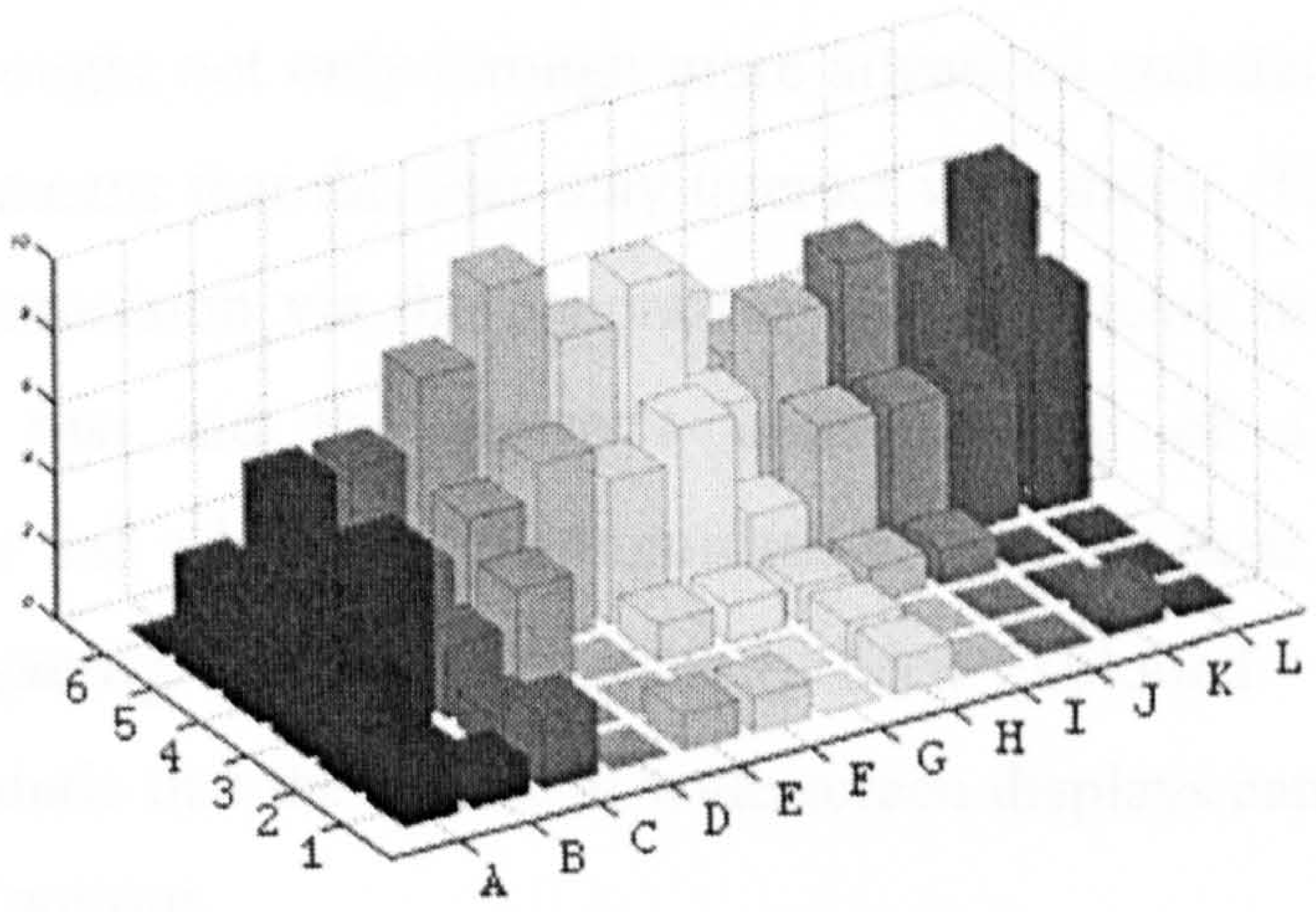


Figure 6.34. Histogram charting user's perceived realism of features

6.6 Summary

The pursuit of reality through an increasing more detailed and accurate representation of the city is inevitably self defeating as the goal can only be approached asymptotically and beyond a certain stage the gains made are disproportionate to the effort expended. It has been profitable to reconsider the form of expression of the urban environment as investigation has shown that there are parallel assertions which can provide similar benefits at greater economies. The adoption of cartography as the prime source of all urban data has come at the price of an inflexible adherence to the accuracy of the model.

The realisation that this route can only lead to diminishing returns suggests the adoption of a more abstract representation using a wide range of generic archetypes to express urban form. The success of this format is demonstrated by the ease with which a limited library of components has been used to great effect in the exploration of urban design issues. Modern methodologies can result in models with a greater visual appeal while still providing a fast and accurate process. The derivation of a grammar composed of a vocabulary of elements capable of being deployed within a range of architectural syntaxes promises great flexibility and at the limit ever greater accuracy.

Realism can be sought not only through more advanced and detailed models but also through the means that the user may interact with them. The rediscovery of the benefits of immersion via the panoramic display allows the introduction of perceptual cues that aid the viewer's understanding of a scene. These technologies have had a long and distinguished history yet their introduction into mainstream computing use has only recently been exploited. The gains to be found are so dramatic that the uptake of wide screen displays can only continue to become more ubiquitous.

In the quest for a better interface and more natural interaction with a scene the ability to provide an energy extractive interface will augment the user's experience by providing feedback from the physical characteristics attributed to the model. This also provides relief from the artificial metaphors usually employed to control the simulation and brings additional perceptual benefits in the enhanced ability to judge scale and distance.

As technology progresses the ultimate reality will not be found in the application of any one technique or the use of any particular piece of hardware, rather the conjunction of a number of innovations, as described here, will make for a system whose worth is greater than the sum of the parts.

Conclusions

> *"In conclusion it is clear that, generally, computer graphics related to architecture are as yet in their early days. The hardware is available but software development has only been possible for organisations which have sufficient skill and resources to research effectively. General uses of graphical techniques will most probably develop with the advance of multi-access systems and appropriate software." (Gutteridge 1973)*

In retrospect, the above quote seems to be a fair summary of the state of the art in the early 1970s. The perceived impression is one of a young branch of technology that offered great promise but had yet to fulfil its potential due to a number of shortcomings. The greatest problem was the embryonic state of the infrastructure. In the early 1970s there were only five time sharing machines available for commercial use, those others that were available were limited to academic and scientific domains. The slow adoption of interactive usage was in part due to the expense of using computers in an interactive mode. A more subtle advantage comes from breaking down the barrier between man and machine in that the user of an interactive terminal may develop a "feel" for the system, an experience that is absent in the batch mode, and may be able to sense when something is going wrong and take immediate action accordingly. This leads to a greater sense of control over the process and allows the computer to become more of a tool in the designer's hands than would be possible in other modes of usage. This factor, although hard to quantify in economic terms, has proven to be a key goal in the development of computer systems for processes in the design and analysis cycle.

In architectural practice justification for the use of computers was sought by attempting to find economic leverage in augmenting the productivity of the existing design staff. An analysis of the design office showed that the greatest number of man hours went into the preparation and production of working drawings and schedules for use on site. It followed that any attempt to reduce design costs by introducing technology into the design office was most likely to be successful if it could addresses the production of working drawings as opposed to a purely numerical appraisal or analysis. This line of thought lead to a proportionally greater emphasis on developing graphics applications for drafting as opposed to modelling and hence documentation rather than design.

In terms of the evolution of computer graphics and associated hardware, once the initial problems were overcome, there were few major landmarks in terms of the evolution of hardware throughout the decade. Vector displays on storage tube terminals would remain the standard until half way through the next decade allowing little innovation within the underlying technology. The general impression is of a slow but steady improvement of all aspects of computer technology, each development progressing the state of the art and allowing, in turn, progress in associated areas. What is important about this era is that each new innovation formed the foundation for much of what was to come in future generations of graphics applications.

The Glasgow model was created in an period of rapidly evolving technology and with hindsight it can be seen that no single innovation made this research possible rather it was movement on all fronts of the key technologies – hardware software and communications – that provided the momentum.

Despite the emergence of a new tier of hardware in the form of the Personal Computer these new offerings never had any major impact on the research. In the early stages of their development PCs were targeted on the business community and had little to offer to memory hungry, graphically intensive processes. For example the much acclaimed APPLE II could only provide an on screen resolution of 280 x 160 pixels while a Tektronix storage tube had an addressable display of 4096 x 4096 and plotter output had an almost infinite resolution.

One of the more liberating characteristics of this era was the freedom from financial constraints in the form of AUs and the ease of access to the University's computing provision. The VAX cluster with in house plotting facilities and the associated communications infrastructure were all available on demand and free of the financial burden incumbent with the use of the Regional Computing Centres. Interactive operation of programs was facilitated by what was, for its day, a high speed network certainly faster and more reliable than the previous leased line GPO modem communications. The administrative limits were restraining rather than repressive and served to reinforce the need for maximising performance and use of resources. This was a major benefit within the process of software development.

In the 1970s ABACUS had received a rolling grant from the Science and Engineering Research Council, this continuity of funding made it possible to develop long standing themes of research, the consequence of which was the development of practical software tools rather than a series of research prototypes. This focused research enabled the group to develop, refine and improve the software over a number of iterations. One example of this process is clearly evident in the development of the visualisation software with its beginnings as HLE, refinement as BIBLE and eventual delivery as VIEWER. With this long gestation period, other researchers within the group were able to add the utility functions around this core, for example VIM, VWALK, TABLET and DEPICT. This software was all built in a time of technological poverty, by today's standards, yet this proved to be a important benefit. The meagre availability of memory discriminated for modular as opposed to monolithic software design which in turn allowed the system to be flexible and amenable to adaptation to new computing platforms. The fact that the period of software development also spanned the eras of off-site computing and subsequently interactive operation brought a useful legacy in that batch processing capabilities were preserved and the attendant subdivision of the data pipeline made for a mode of usage that was fundamental to the exploitation of the current computing environment.

The development of the model was facilitated by the hardware and software factors as outlined above but the main impetus for development came from a wider range of considerations. The existence of the Campus Model had proven invaluable in demonstrating the concept and provided an exemplar with which to attract future funding but the most telling asset was the unique nature of the model itself. At this time there were no other comparable geometrical computer databases and interest in the project enabled it to be self sustaining over a period of years. This continuity provided an almost continuous input enabling an iterative approach to problem solving, the periodical injection of new technology and the refinement of the model's structure and detail. It is also interesting to look back and note the ease of forming a relationship with information providers before data became a valuable commodity in its own right. Without the assistance of the various departments of the District Council, architectural practices and all those who provided invaluable information the formation of the model would have been potentially much more expensive and time consuming.

This area of application was always easy to bring to the attention of the public and commercial community as all those who lived in, worked in or were involved in the management of the city could find a use for the model. Initially the model was very busy as architects, developers and illustrators would commission views of the city which they could then embellish with their own proposals. While some of this work was used to show new developments within their context a large proportion of users utilised the model as a vehicle for adding their own information. This might consist of zoning for redevelopment opportunities or as a location map showing the adjacency of hotels, transport routes and parking facilities.

It became obvious that with the prime mode of usage for the model being as a medium for a further layer of information then future development would have to take this into account. It had always been recognised that the attribution of the geometry with alphanumeric data would create a powerful tool but with the primitive vector graphics initially employed the difficulty of interfacing with this information was always a limiting factor. Once engineering workstations with powerful processors and sufficient memory became available this factor was no longer so restrictive but the structure of the model – with many buildings grouped in the same file – proved difficult to manipulate. Given the lack of structural coherence in the model the ability to utilise any inbuilt intelligence within the geometry would always be compromised. As the usage of the model became more sophisticated queries that hinged on these properties became more common. Requests for information as to the cumulative lengths of pavement within an area, the minimum radius of street corners and printouts of the address of buildings within an ownership category could not readily be answered and served to indicate that without substantial investment directed towards a fundamental upgrade of the model then its future would be limited.

The decline of the multi-user mini computer was mirrored by the ascent of the personal workstation. This brought the benefits of new graphics technologies and increased performance. In turn this was translated into a broadening of the horizons of what was, and would be possible in terms of modelling and visualisation.

The final developmental goal of the Glasgow model had always been targeted on the attribution of the geometry with alpha numeric data. This desire had been limited by the lack of graphical capability in a predominantly vector based wire line world. As the workstation market became established a number of investigations were made seeking to capitalise on the strengths of window managers and bit mapped displays. While these projects were successful to a limited degree the development was centred around the interface to the model rather than towards the enhancement of the model itself. As such the limited additional information offered was more decorative than informative.

The additional capabilities of the graphical workstations allowed the use of colour, giving greater realism and the ability to define subtle detail but also introduced new constraints in the way that an observer perceives the scene. The hardware assisted transformation and rendering not only gives increased performance but also proved to be more forgiving of less than perfect geometry. A major constraint with VIEWER had been the software's inability to cope with inter-penetrating bodies and non-planar surfaces. Now that building construction could be less inhibited it became easier to describe the built form through the conjunction of a number of bodies. Since the Old Town was possessed of a much more organic nature than Glasgow the simplification of this process imparted substantial gains in modelling efficiency.

The dynamic interaction supported by advanced hardware heralded a totally new era in computer graphics and while the ability to interactively visualise a single building proved invaluable during the construction phase the performance offered by these early machines was incapable of handling even a small proportion of the entire city model. As the amount of geometry expanded the frame rate decreased and below about 10 frames per second it proved too difficult to exercise enough control for precise navigation.

A supplementary benefit of the new technology was the provision of advanced software libraries that added new features and capabilities to existing software. These included tools for image manipulation and processing as well as functional elements for graphics

The opportunity to collaborate on the development of a new urban model rekindled interest in this area and the availability of a new generation of hardware and software offered the opportunity to refine the methodology and improve on the original concept. The availability of accurate digital cartography immediately made it possible to generate large 3-D models with the minimum of human intervention. This process showed great potential for constructing wide areas of urban landscape at a contextual level of detail. Once the procedure was in place, and with suitable tools for data translation, there were few boundaries to the potential scale of a model.

Once the emphasis is turned to the detail of the model the worth of this phase of the data capture is proportionally smaller than the remaining effort required. Most of the time and effort is then spent on research and data gathering of those attributes that are unavailable in a 2-D format. The limiting factor then becomes tied to data that is not yet available in the digital domain and must be collected by manual surveys or transformed from conventional media.

The form of the building stock is relatively easy to both identify and capture but when an observer is situated at ground level then other aspects of the urban environment become equally if not more important. The translation of general line detail can provide much of the immediate context for a building but there is a greater need to transform this data into walls, fences, hedgerows and areas of hard and soft landscape. These characteristics are every bit as important as the architecture but are much harder to identify and execute. Without these details the streetscape is largely visually desolate and barren but to include just some of the required features would be a task equal in effort and complexity to capturing the buildings alone.

Although the Edinburgh model was more comprehensive in terms of detail, enjoying a better rendition of the terrain and building stock, it was not necessarily a better urban model. It has previously been stated that, in terms of the data requirements of an urban model, quantity can be directly analogous to quality. In practice this is not always the case. When working with large amounts of data there will always come a point where the practical difficulties of handling large quantities will start to offset the benefits of having access to that data. As the quantity of data increases so does the net benefit. This holds true until the practical difficulties start to increase at a greater rate than the benefits. As this occurs the graph will approach a maximal turning point before the benefits start to actually decrease. The shape and turning point on this graph will be dependant on the ease with which the data can be handled, a more logical structure or an improved interface to the model may make it possible to extract significantly greater benefits from the same physical quantities of data.

Working with the Edinburgh Old Town Trust as a client brought another level of complexity to the development of the model. The original strategy had been contrived so as to work up detail over the entire area to much the same extent as had been applied the Glasgow model. A second pass would have then picked out the landmark buildings with further effort directed towards the facades of the major streets. However, since the Trust was involved in strategic interventions at a number of locations they were naturally anxious that effort should be concentrated around these sites. This tended to make the application of detail more uneven than desired and lead to some areas of the model receiving limited attention. This made the model less globally useful and restricted animated and static scene generation to only those parts that had been previously considered.

The development of the interface proved to be one of the more interesting aspects of the model. While it was both restricted in geographical scope and limited in content it proved to be a most useful addition to the methodology. As urban models get larger the management of data will become an ever increasing overhead. With only a cryptic filename to differentiate data files the ability to select buildings from a map proved to both facilitate the task and save much time and frustration. Despite adopting a strictly hierarchical file naming policy based on street names and house numbers the sheer volume of files associated with this form of model can prove difficult to handle. The adoption of a graphical interface should be implemented from the start. Once attributed with textual information concerning age, ownership and use the database can be browsed, queried or searched with the results presented in a graphical manner. Not only is this a useful piece of functionality but it also adds interest for the casual user.

Developments from all corners of the information technologies will aid the uptake of urban information systems, emerging standards, the move towards open systems, increased networking infrastructures and the every greater availability of data in digital formats will increase the mobility and sharing of databases. These factors coupled with a growing user community will create an environment that will foster an increasing growth rate.

The issue of financial justification will be important, especially in a local authority context. Although hardware costs are on a downward trend, many of the costs associated with the system remain high. Budget expenditure will be greatest at early stages while many benefits will only accrue once the system is in widespread usage. A complex issue is one of asset management. The better the existing internal asset management then the more difficult is the justification for new technology. For example, the authority which has recently invested in high quality OS paper map coverage of it's land coverage may find it difficult to justify recapture in a digital format.

Given all the factors that influence the success of an implementation such as that discussed here, the technical aspects of authoring the software and targeting a suitable hardware platform, are undoubtedly in the minority. The real challenge is in the legal, ethical, political and economical barriers to promoting the sharing of data and encouraging centralised management of assets. Even within the tight boundaries of local and regional authorities the duplication of effort, the adoption of mutually incompatible formats and the lack of willingness to pool resources all tend to negate the potential benefits. In short, the only remaining barriers to this concept are essentially human ones. Society is increasingly dependent on information, in all forms. It is easy to see the move towards a critical mass where enough interested parties hold enough information to make the move towards the philosophy of centralisation of information assets self perpetuating. Once this occurs then the potential benefits may be realised.

Multimedia has always been hard to delimit. As a ubiquitous concept it could be applied to a wide spectrum of ideas, as a reference to hardware it labelled a broad range of platforms, in terms of software it underpinned an entire industry and as a product it was both a rousing success and an abject failure. What multimedia best represented was the collision of a number of enterprises spanning communications, entertainment and the traditional computer industries. Each of these industry giants saw multimedia as a chance to gain some leverage on the market, an opportunity which they exploited to the full. This in turn lead to success and failure being opposing sides of the same coin.

Failure is represented by the inability of the industry to live up to the claims made by overzealous marketing departments. No one was quite sure whether multimedia was a software product, an item of consumer electronics or an extension of current computer hardware yet this did not prevent market analysts forecasting the death of the traditional media formats and the unlimited potential of the new. Industry giants from sectors such as cinema (Walt Disney), publishing (Time Warner) and consumer electronics (Philips, Sony) invested billions of dollars and yet the mass market never quite materialised.

On the other hand, common sense dictates that the conflation of a variety of descriptive media must make for a tool that exhibits the potential for being "One of the most powerful forms of communications media ever developed". This fact is not in dispute but the niche for a product that allowed the user to "convey ideas, search for information and experience new concepts" was always somewhat smaller than the imagination of the market might have forecast especially when this opportunity only became available with the purchase of an enhanced, and expensive, combination of hardware and software.

From this perspective multimedia did not live up to expectations yet it was the driving force behind a revolution in the computer industry that took computing out of the commercial environment and into the home. This expanding market was energetically pursued by the manufacturers who first offered the "Multimedia PC" as an enhanced product before time and market forces closed the market to any hardware that did not offer these now fundamental capabilities. Multimedia is a "hungry" format, where quality only ever comes at the expense of greater storage, processor power, memory capacity and hardware functionality, once these features were deemed necessary then the "best" hardware was only ever going to be the offering that provided greater quantity and, by association, quality.

The mass market might have not have lived up to initial expectations yet there is still a niche where the traditional vision of multimedia continues to fulfil its promise. These examples occur whenever there is a need to conflate a variety of media and where the benefits of a homogeneous collection of heterogeneous data types can be arranged in a systematic format designed from the outset to facilitate piecewise access and cross-referencing.

Exemplars of this type can be found in education particularly in the realm of architecture and design, much of which is rooted in the past. Architectural learning can be seen to be retrospective, with modern design methods and practices based on their precedents. Students are encouraged to look at examples from history for both inspiration and evidence of good practice. The provision of exemplary subjects has long been the goal of many educational architectural texts and it is plain to see that multimedia has a role to play in a move away from traditional paper based formats to more modern electronic archives.

Similarly, access to information, artefacts or buildings is just one aspect of the problem associated with learning from our heritage. To gain full advantage from these initiatives there must also be an attempt to explain their significance, to reunite disparate items with their context and interpret their cultural value. This is perhaps where the new media can show the most benefits via its flexible approach to amalgamating media and through this composite medium offer new contexts, interpretations and viewpoints.

These opinions, and the variety of case studies presented previously, show just some of the possible applications of multimedia. The benefits are evident not just in the ability to develop applications that use the technology to integrate media in new and innovative ways but also how, from a greater perspective, multimedia has been subsumed by the computing industry as a whole and now exists as an integral part of the technology and not as a separate entity.

In theory we can no longer differentiate between multimedia applications and any other flavour of computer program. The capabilities exist to incorporate multimedia design ideals into all manner of traditional programs. Even word processing packages boast voice annotation and the ability to incorporate hyperlinks amongst text, graphics and digital video.

The transitional boundary between multimedia and the Internet continues to be eroded. Many of the applications that were initially location based have been repositioned to capitalise on the interest shown in a new medium. Much of the promise associated with the early multimedia productions has been inherited by this new technology and although re-implemented most remain faithful to the original ideals. The areas where the Internet does offer great advantage is through its influence on breaking down barriers to the dissemination and sharing of knowledge. The adoption of open source policies has done much to enhance software development and the platform independent delivery of content serves to make this knowledge available to all. The definition of standards and the availability of ubiquitous tools to exploit the same have proved a strong incentive to adopt the medium as the platform of choice. The ability to define a model of a city on the Internet as a "site" allows the developer to subjugate "the Web" by integrating links to external content within the model. This is perhaps the best opportunity to realise the epitome of the Digital City as a "database of databases" first mooted a decade before. The addition of graphical interactivity adds a natural form of interface and a more interesting mode of interaction than the common WIMP paradigm. While the concept laboured under less than adequate graphical processing power for much of the period the growing market for experiential computer games has provided a strong stimulus for technical development. The worth of all interactive content increases as the ease with which it can be manipulated grows. This is just the start on an infinitely long path towards a virtual reality which will require the embodiment of a wider range of technologies than has been previously discussed.

The pursuit of reality has been the quest for all researchers and users of graphics systems ever since their first introduction. It is only recently that enough of the critical technologies have matured to the stage where this dream might become fact. It is interesting to note how many of these seemingly innovative technologies are neither new nor novel. The existence of almost every aspect of the ultimate graphics system has been pioneered in the distant past, over 100 years in the case of immersive display technology. Primitive motion platforms were in use at the same time and also coincidentally resurfaced at the same World Fair as saw the reintroduction of panoramic cinema.

Despite these advances the concept of the Digital City cannot exist without content. The daunting quantities of data involved come with all the classic problems of capture, manipulation and storage and the prospect of modelling large tracts of urban landscape come complete with a seemingly geological time scale. Even if the resources are available the creation of an accurate facsimile of all the building stock can only ever be given a limited treatment. Experience has shown that relaxing the desire for 100% accuracy can result in major savings with minimal loss in realism or utility. In fact through the acceptance of a degree of abstraction not only can large savings in time and effort be found but also greater detail can be employed through the reuse on components. It makes economical sense to apply greater consideration to individual components if they are to be reused multiple times.

This has lead to another rediscovery, that of a methodology commensurate with the ability to economically manipulate and represent urban space. This frees a designer of many of the constraints concerned with traditional modelling methods and supports the generation of more numerous and telling appraisals. The continued development of this tool set has enabled a more visually accurate and powerful design instrument by adopting a more modern gamut of graphical techniques. The application of a library of abstract or archtypical representations of vernacular buildings can be used to generate startlingly realistic urban scenes. As the breadth and depth of this vocabulary is increased then not only will the representation gain visual realism but will also become more accurate. In the real world, where every different conjunction of forms may produce a unique structure and where the organic process of development and redevelopment shades the regular disposition of the original buildings, then this methodology may fall far short of total accuracy. However experience has shown that humans are remarkable efficient at inferring meaning from the most minimal or iconic archetype and may not need greater accuracy for the majority of visual applications.

Wide screen display techniques have had a long and distinguished history in partnership with a range of visual media. While their inherent benefits have been enjoyed by well resourced organisations like the military it is only recently that price and availability have seen their use pass into the remit of education and commerce. The visual realism offered by these systems can provide all the necessary visual cues required for surface recognition and spatial navigation. The wide angle of view is beneficial as this allows the user's direction of view to be decoupled from the direction of motion. This enables a user to look around within the environment rather than be constrained to the narrow view frustum common to conventional graphics displays. The downside is that this form of display provides an "out of the window" view which separates a user from objects, which would otherwise be within arms length.

The graphical display gains its advantage from enhancing the perceptual cues required by the user in order to experience immersion within the environment. These cues can be reinforced by interacting with the system through a more natural, if mechanical, interface. The key concept has been found to reside in the ability to equate energy expenditure with locomotion as this serves to calibrate the visual system in relation to distance and scale. When operating the system through such an interface a physical simulation can be based on the interaction of the user's presence with the geometry within the virtual world. As such, refinement or extension of these capabilities would usually be just a matter of writing additional software. However there are some fundamental limitations with this concept in that the system was primarily related to the interaction of the interface mechanism with the environment as opposed to modelling the users personal interactions. With reference to the wheelchair project described previously this is exemplified by the manner in which a user might negotiate swing doors. In this event a wheelchair occupant might tend to use their knees, or the chair itself, to wedge the door open while manoeuvring for a favourable position from which to exert additional leverage. This would be beyond the scope of the current implementation. While we may never be free of the physical constraints of our environment the promise of the Digital City is to deliver freedom of access to available data and information as an attribute to the familiar, if virtual, structure of our built environment.

Although much was accomplished during this period many avenues of research did not achieve any lasting impact. Among the ground breaking technologies forecast were "natural language interaction" and "artificial intelligence" both of which have yet to mature. However some predictions that have come of age and have also proved to be remarkably accurate are reproduced below, it just took a little longer than expected.

> *"The building designer of the future will have direct on line links from his office, not only to massive computing power and data resources, but also to people with other skills and interests. The building design held in the computer will be accessible to others who will be able to participate in its use." (Marcus 1969)*

> *"Development of interactive graphics will provide more realistic and useful visual representations than the current line drawings on cathode ray tubes. It will be technically possible to "walk" through full scale buildings at the design stage, the computer providing 3D images automatically adjusted to the user's eye point as he turns his head." (Reiners 1969)*

References

ABACUS (1976) SITE: Site Analysis and Appraisal Routines. ABACUS Occasional Paper No 55.

ABACUS (2002) Website: The Glasgow Directory, Last Accessed January 2002, http://vrglasgow.co.uk

Akeley K. (1993) Reality Engine Graphics. Proceedings of SIGGRAPH. ACM.

Anderson P. (1994) A Practical Guide to MultiMedia, PLS Consult, Denmark.

ASL (1989) ROVE User Manual – Version 4.1 ABACUS Simulations Ltd, University of Strathclyde.

Bergstein L.A. Perspectives in Analysis and Design, College of Architecture, University of Lund, Sweden.

Bowyer A, Woodwark J. (1983) A Programmer's Geometry, Butterworths, London.

Bridges A. (1982) Showing IT how it is. Computer Aided Visual Impact Analysis. CAD/CAM International. No 13 November 1982, Special Supplement, CAD in Construction, pp. x-xii

Clark James H. (1982) The Geometry engine: a VLSI Geometry System for Graphics. Computer Graphics, Vol 16. No 3 1982

Cok (1999) Developing efficient Graphics Software: the Yin and Yang of Graphics. A SIGGRAPH 99 Course. In SIGGRAPH 99 Conference Course Notes, ACM SIGGRAPH, August 1999.

EDVEC (2002) Website: (Last Accessed January 2002). http://www.edvec.ed.ac.uk

EOTRT (1994) Action Plan For The Old Town, Edinburgh Old Town Renewal Trust and the City of Edinburgh District Council Planning Department, Second Review.

Frey H. (1987) Continuity and Variety in the Urban Environment. In the Construction of the Urban Place: Methods and Instruments. Ed Antonella Cartezi, Alinea, Firenza.

Galimberti R. An Algorithm for Hidden Line Elimination, University of Montanari, Italy.

IBM (1994) IBM Monochrome Display and IBM Colour Display, IBM Marketing Information, GU20-4019, IBM.

Innsbruk (2002) Website: The Cyclorama, Last Accessed January 2002, http://www.panorama-innsbruck.at/e/com/info.htm

Jones A. (1990). Three Dimensional Urban Terrain Perspectives, An Investigation into Alternative Sources of Data for the VIEWER/VISTA Architectural Modelling Packages, MSc Thesis, Glasgow University

Kraus K, Myer S, Danielson S, Lewis R. (1968) Architects' Specification for a Computer System, Centre for Building Research, Massachusetts Institute of Technology.

LaserScan (2002) Website. Last accessed January 2002. http://www.laser-scan.com/papers/lamps2_and_3d.htm

Lynch K. (1960) The Image of the City. MIT Press. Cambridge,Mass

MacroMedia (2001) Website. Last accessed January 2002. http://www.macromedia.com

Marcus T.A. (1969) Chairman of Group 3, reporting in Computer Aided Architectural Design, Part 1: Reports of three working groups, Ministry of Public Buildings and Works, Directorate of Research and Information, London 1969.

Marcus T.A. (1967) The Role of Building Performance Measurement and Appraisal in Design Methods, Architects Journal. Dec 1967.

Maver T.W. (1982) Visual Impact Analysis by Computer Graphics. Proceedings of Eurographics '82. North Holland Publishing Co.

Maver T.W. (1972) PACE - An On-Line Design Facility, Proceedings of the EDRA 3 / AR8 Conference, Los Angeles. 1972

NetScape (2002) Website: Last Accessed January 2002, http://www.netscape.com/eng/mozilla/3.0/handbook/javascript

nVidia (2002) Website: Last Accessed January 2002, http://www.nvidia.com

OS (1983) Digital Map Data, DO9 User Guide, Documentation on the Plot Program, Ordnance Survey, Document No D28AAA.

OS (1988) Digital Map Data, Ordnance Survey Transfer Format, Incorporating OS1988 specification, Ordnance Survey, Document No I/3/1/30/Nov,88

OS (2001) Website: Last Accessed January 2002. Ordnance Survey, Digital National Framework, Consultation Paper 1/2000, www.ordnancesurvey.gov.uk

Parkins R.P. (1979) BIBLE: A Computer Program for Generating Perspective Views of Buildings. ABACUS Occasional Paper. No. 75

PCW (1998) 20th Anniversary Special Feature, Personal Computer World, Volume 21, Issue No 5, May 1998.

PCW (1978) Personal Computer World, Issue 4, August 1978

Petric J. (1988) Computer Modelling of Landscapes. PhD Thesis. University of Strathclyde.

Reiners W.J. (1969) Chairman of the Committee on the Application of Computers in the Construction Industry, reporting in Computer Aided Architectural Design, Part 1: Reports of three working groups, Ministry of Public Buildings and Works, Directorate of Research and Information, London 1969.

RIBA (1965) Architectural Management and Practice, Royal Incorporation of British Architects., London.

Rieser J, Pick H, Ashmead D, A.Garing A. (1995) Calibration of Human Locomotion and models of Perceptual-motor Organisation. Journal of Experimental Psychology : Human Perception and Performance, Issue 21, 1995.

Roberts M. (2002) Website: Last Accessed January 2002. http://web.mit.edu/mroberts/WWW/Baraka/panoramas.html

Robinson P. (1986) Aspects of a Scottish Flat Tradition, PhD Thesis, Strathclyde University.

SGI (1988) Personal Iris 4D20, Silicon Graphics Technical Specification Sheet esd18/88

Shamwana M. (1986) VWALK: Viewer Input Path Generation. ABACUS User Guide.

Stearn D. (1982). VISTA: Visual Impact Technical Aid. Proceedings of Eurographics '82. North Holland Publishing Co.

Stearn D. (1983) VIM: Viewer Input Management. ABACUS User Manual. No M41.

Stearn D. (1984) DEPICT: VIEWER Picture File Postprocessor. ABACUS User Guide.

Sussock H. (1973) Graphics facilities for Computer Aided Architectural Design, ABACUS Occasional paper No 25

Sussock H. (1974). Computer Generated Perspective Views. ABACUS Occasional Paper No. 32

Sussock H. (1981) GOAL, General Outline Appraisal of Layouts, ABACUS Occasional Paper, No. 62.

Sussock H. (1982). GOAL : Graphics User Manual. ABACUS User Manual No M27.

Sussock H. (1985) VIEWER User Manual Version 4.2. ABACUS User Manual No M11.

Teller S. (1991) Visibility Preprocessing for Interactive Walkthroughs, Computer Graphics Volume 25, Number 4 July 1991, ACM.

VRML (1995) Website Last accessed January 2002. www.vrml.org/technicalinfo/specifications/vrml1.0.htm

VRML (2002a) Website: Last Accessed, January 2002. http://www.vrml.org

VRML (2002b) Website: Last Accessed, January 2002. http://www.vrml.org/fs_technicalinfo.htm

VRML (2002c) Website: Last Accessed, January 2002. http://www.vrml.org/technicalinfo/specifications/vrml97/index.htm

Walker J. (2000?) UNIVAC Memories, Website: Last Accessed January 2000, http://fourmilab.ch/documents/univac/index.html

APPENDIX

1. 1989 Grant M. "ISSUE an Interactive Software System for the Urban Environment" **Development Of an Evaluation Tool for Urban Interventions - A Case For Support to** the Science and Engineering Research Council.
2. 1992 Grant M. "ISSUE an Interactive Software System for the Urban Environment" Final Report for the Science and Engineering Research Council, Grant GR/F/91919. Subsequently published in: CAAD Futures 1993, Pittsburgh. Published by North Holland. Edited by Flemming and Wyk. ISBN: 0-444 899227
3. 1994, Grant M. "Urban GIS - The Application Of The Information Technologies To Urban Management" The Virtual Studio, ECAADE, Strathclyde. ISBN 0-952368706
4. 1995, Grant M. "Multimedia, A Multi Purpose Programming Environment", Conference Proceedings, ECAADE, Palermo. ISBN 0 9523687 1 4
5. 1996, Grant M. "Virtual Heritage", Conference Proceedings. ECAADE Lunde (Late Paper)
6. 1999 Grant M. Ennis G. Lindsay M, "VRML Possibilities", Proceedings of Milestones for the New Virtual Millennium, Dundee. ISBN 1899 796 053
7. 2000 Grant M. et al, "Development of a Wheelchair Virtual Reality Platform for Use in Evaluating Wheelchair Access", 3rd International Conference on Disability, VR and Associated Technologies, Sardinia, Edited by P.Sharkey, ICDVRAT ISBN 0704911426
8. 2001 Grant M. "Haptic Interfaces for Wheelchair Navigation in the Built Environment", Unpublished Paper, Extracted from the Final Report for EPSRC Grant GR/MO5416/01

ABACUS
Department of Architecture Strathclyde University
Glasgow
Scotland

DEVELOPMENT OF AN EVALUATION TOOL FOR URBAN INTERVENTIONS

- A CASE FOR SUPPORT

M.Grant

1.1 SUMMARY

Many cities in the UK, and indeed throughout the developed world, are characterised by the all too familiar symptoms of urban blight caused by insensitive intervention in the environment. The common denominator within this class of problem is the lack of a coordinated, integrated approach to the planning, design and maintenance of our cities. The cycle of development and redevelopment involves input from a diverse range of disciplines relating to architecture, civil engineering, transport engineering and the management of city utilities. This lack of a common, updatable, information base renders access to a global view of the city difficult if not impossible.

ABACUS, at Strathclyde University's Department of Architecture, has amassed a geometrical database describing the natural topography, buildings and transport networks covering much of Glasgow's inner city. This database, now consisting of some 10,000 buildings spread over 20 square kilometres has been widely used to describe the visual context for a number of interventions in the urban fabric.

The objective of this proposal is to seek support for a prototype system where, by attributing the existing geometrical framework with those physical quantities that are relevant to the formal and functional evaluation of the urban environment, the means of evaluating the qualities and quantities of the buildings within the context of any city may be realised. The outcome will be a prototype of an Integrated Software System for the Urban Environment (ISSUE) allowing interactive access

to a global dynamic view of the city with the ability to interrogate, evaluate and communicate any proposed development strategy relating to the fabric of the city, its utility infrastructure or its inhabitants.

1.2 BACKGROUND

In 1986 ABACUS produced a computer based model of the city centre in Glasgow, the result being the first comprehensive computer simulation of a city in the UK. At this stage the area covered extended over some 7 square kilometres and although the detail was fairly coarse the model was seen as a major advance into the field of computer graphics at an urban scale. This project was seen as a pilot study to assess what may be achieved in modelling the urban environment and as a test case for the development of specialised software tools for data capture, data management and evaluation.

The data came from disparate sources, mainly ordinance survey maps, but for more recent developments the city's planning department provided willing and enthusiastic cooperation. Currently new developments within Glasgow are tested against a wooden model, once planning approval is granted a block is carved and added to the existing buildings. Of guaranteed accuracy this unusual source provided information on buildings for which details were unavailable or simply did not exist. Added to this the highway department's stereoscopic aerial photographs gave useful information on building heights and orientation.

Phase one of the project consisted of capturing the physical terrain, roads and buildings to a simple level of detail, in the case of the majority of buildings just basic shape and height. Phase two resulted in a structured hierarchy of detail, the university campus, the merchant city and areas of the Clyde corridor all being built up in greater detail. The third phase concentrated on consolidating the structure of the model, expanding its geographical scope and increasing the amount of geometrical data by entering facade details and roof profiles.

The database is held in a machine independent format and can be accessed by a variety of programs. The graphics available include wire-line, shade and colour simulation and also the ability to animate sections of the city allowing the user to move through the streets. Output from the database can take a variety of forms and, besides pen plots, there are film, video tape and colour slide facilities.

During the intervening period the city model has developed from an academic exercise in urban modelling, to a substantial asset to the city of Glasgow, the software having progressed from a trial implementation to a robust and user friendly tool. Widespread reportage in press and journals has encouraged its use in projects as diverse as a study of the disposition of festive street lighting, to feasibility studies on the re-introduction of city trams, through to day lighting studies, visual impact analysis of new developments and as an interactive city guide for tourists. With the present professional concern with urban context, and the EEC formalising the need for environmental impact appraisal, this case appears well timed.

ABACUS have been working in the field of visual assessment since their establishment as a research arm of Strathclyde University in 1970. Being part of the department of Architecture and Building Science, much of the work has related to building design. The research and development side of ABACUS parallels its teaching input, and it has been much involved in power line and pylon assessment related to countryside protection. Visual simulation of the city is a natural extension and now allows prospective developers to test their projects within the context of the city. Each new aspect of usage has enabled ABACUS to expand and refine the model until it now encompasses some 10,000 buildings within 20 square kilometres of the city of Glasgow, supporting a hierarchy of detail from basic dimensions to fully detailed architectural modelling.

1.3 SYSTEM OVERVIEW

Although much of the software and expertise required for the process of urban modelling was already in existence, the constraints peculiar to this scale of operation resulted in the development of a suite of software tools and a methodology that are orientated towards facilitating the data capture. In effect, this has allowed the construction of three data-sets: the terrain, the buildings and the transport networks.

A terrain model is made by digitising, from topographical maps, the land contours at a scale of 1;10,000, interpolation between the resulting spot heights results in a solid model that forms the context for the city buildings. The transport networks, roads, rail links and rivers are all entered in 2-D. Custom software then takes this plan representation and moulds it to the terrain by "floating" the data down on to the digital terrain model, similarly, once the geometry of the buildings is prepared using the geometry editor, they too are repositioned in 3-D space in relation to the terrain.

This basic model has evolved into a hierarchically structured data-set, each building block being referenced by its map reference and held, for ease of modification, orthogonal to the axes and originated at the model origin. Each file is then referenced symbolically and translated and rotated into its true "world" location. Parts of the city exist in different stages of detail, from basic blocks through to fully detailed architectural models.

The city geometry is accessed via an interface analogous to a city tourist atlas. The user is presented with a 2-D road map showing the city streets, squares and outlines of the city blocks. This plan can be interrogated by pointing with the mouse and cursor, a "pick" returning the address and identity of the object of interest. A further mode enables picking of individual buildings or whole sections for interactive viewing in 3-D. This operation creates a list of the geometry files that are then passed to the visualisation program. At this level, interaction with the dataset is, even for a novice user, easily achieved. ABACUS software allows editing and manipulation of the geometry, evaluation by visualisation or photomontage, but access to more telling evaluations is hindered by the lack of suitable attributes describing the physical, social and economic properties of the city.

1.4 THE PROPOSAL

In the past, CAD systems have concentrated on documentation rather than design, draughting systems intended to increase the efficiency of documentation production at the detailed stages of design are now ubiquitous, decent design systems intended to improve the quality of the built environment at an urban level simply do not exist. This situation has lead to the promotion of individual buildings at the expense of those issues relating to urban design; similarly the constraints of affordable computer graphics have reinforced the formally static view of the built environment at the expense of a dynamically changing experience in urban development.

More recently we have seen the advent of advanced modelling software, capable of simulating the appearance and performance of all manner of buildings. The application of this technology has been limited almost entirely to the scale of individual buildings, however with new hardware technology becoming available the scale of these endeavours have grown until we now see, in the USA, the Skidmore, Owings and Merrill model of Chicago and in Japan, impressive developments by Inter-Lab on urban scale models of Yokohama, Kobe, Osaka and Paris. Nearer to home, ABACUS at the University of Strathclyde, have produced a computer based model of the city of Glasgow.

The worth of the concept behind this proposal has already been proven as an aid to assessing the visual impact of interventions in the urban fabric, the challenge is now to develop a "value added" system that enables varied evaluations and representations of the physical qualities and quantities required by urban management strategies.

The first task is to review the potential range of usage of such a system, this can be enabled by visualising the concept of urban management as being arranged on two axes. Looking along the horizontal, we can visualise the potential user groups; ranging from the professional urban planners through to managers of city utilities and services and even to the casual users of information, the inhabitants themselves and visitors. From each identifiable user group, the vertical axis can be associated with the information each group holds and the use they find for it. The first stage will then be to explore first the horizontal extent of this matrix, identifying the users, rationalising their role in this schematic view of the city and establishing the connections between groups. Secondly, it is proposed to then select a suitable user group and make a detailed study of the vertical axis, identifying the nature and format of the data already held, the methods of collection, storage and usage and the eventual end use in terms of evaluating qualities and quantities of interest. A huge information set of urban data exists in all cities, the relevance of this stage is in the identification and classification of the useful data required, the knowledge of how it is held and the likelihood, or otherwise, of other cities holding the same information in a similar way.

The second stage, in the light of the information gathered previously, is to review the existing tools, methods and data that the proposed system requires. A range of tools already exist for the purpose of digitising and manipulating geometry, but there is a need for a means of attributing the existing geometrical framework with those physical quantities that are relevant to the formal and functional evaluations envisaged, the aim is to identify those areas of the current data set that are deficient and to introduce more of the relevant attributes. Having identified these areas of interest, software tools can be created that can utilise existing data, if held in an electronic form, or capture the relevant details from other media. The hardest phase of the process of urban modelling is finding suitable sources of data. A number of sources have been explored including existing archives, historical surveys, aerial stereoscopic photographs and even a physical model of the city centre owned by Glasgow District Council. New technologies may facilitate the capture of the relevant details from other media, specifically, cartographic photogrammetry and digital scanning have the potential to release much more information at a significantly lower cost.

The review of this vertical theme will provide information on the nature and purpose of those evaluations commonly applied to buildings at an urban scale.

By providing the users of this data with the means to attribute the urban geometry with the physical characteristics of the buildings it will be possible to evaluate individual properties, city blocks or whole areas of the urban environment within their natural context. The level of detail envisaged will be limited, in the first instance, to a relatively simplistic level of representation, but the potential is limited only by the imagination. The next task is to provide a means of relating the data. in a meaningful manner, back to the user. This requires the development of a front-end, a user friendly interface that allows interactive management of the data, and a similar-back end that presents the data in an easily assimilated and relevant form.

Since the geometry data is held in a digital format the structure of the model itself can provide a degree of information, for instance, quantities such as volume, orientation, surface area, floor area and height can all be easily calculated and returned. This lends itself to simple technical evaluations such as global heat loss, solar availability and built density calculations. The results of the evaluations could be displayed as an attribute to the geometry, say colour or height of an extrusion of the ground plan, providing an effective means of communicating large scale evaluations.

1.5 RELATED WORK

Recent years have seen the advent of Geographical Information Systems, software dedicated to the description and evaluation of the natural environment. Such tools showed great potential in the planning, management and preservation of natural resources, and the rapid take-up by commercial concerns has been justification for the investment involved. Attention has now turned to the provision of similar tools dedicated to the urban environment, recent conferences relating to this theme demonstrate an international interest with a number of research projects working towards common goals. In the UK interest is also growing, ABACUS have commenced informal cooperation with the department of Photogrammetry and Surveying at the University College London and are involved with the Centre for Configurational Studies at the Open University in a European Interest Group on the subject.

It is becoming apparent that information systems of this genus are evolving on separate media, namely engineering workstations with graphical accelerators, Hypertext media on low cost micros and optical disk systems on the lines of the BBC Doomsday project. Within the University of Strathclyde the Department of Information Science has developed a HyperCard based system, Glasgow Online, that contains social, political and economic data relating to city life. The database contents, structure and techniques should all prove relevant to our proposal. Similarly ABACUS is collaborating with a team from Heriot Watt University who are engaged in developing the use of optical disks as a storage and recovery medium for data relating to commercial property in Scotland.

Contact shall be maintained with these organisations and cooperation extended to all other interested parties.

1.6 RESOURCES

Briefly, funding is requested for:

One RA for a period of two years.

Travel funding to keep abreast of parallel developments within the UK, and attendance at major conferences, both to maintain contact with other researchers in the field and to disseminate the results arising from the research.

No major capital expenditure is envisaged as the existing workstations are adequate, however a sum has been budgeted for the acquisition of an optical storage device. The large scale data storage requirements inherent in this project call for a means of storage on a media that is erasable, rewritable and removable. Products such as the SONY NWP 5393 offer these facilities, 596Mbyte capacity with reasonable access times and capable of interfacing to existing hardware through a SCSI port. There is also a requirement for machine maintenance overheads and the provision of consumables.

1.7 MANAGEMENT

The project will be coordinated by the ABACUS management group headed by the applicant, with progress meetings held weekly, progress reviews monthly and interim reports quarterly.

Close contact will be maintained with those agencies that have championed the model during the early stages of development. (Glasgow Action, Glasgow Region and District Council).

Most important of all will be the continuation of those links with the architects, developers and city planners that have shown such enthusiasm for the concept in the past.

1.8 THE OUTCOME

The product will include the establishment of a suit of software tools dedicated to the capture and transformation of data relating to the formal and functional representation of the urban environment. These tools will be implemented on a common machine base along with a representative set of data describing the urban topography and it's physical, social and economic attributes. This data will be collated as a result of a survey of the users of such a product and will form the prototype of a system giving the means be which architects, planners or managers of city development can interactively evaluate and communicate the results of urban intervention in the city fabric. Although proven in the context of Glasgow the concept would be portable to any city, the knowledge, data and tools being available as a template for any future ventures in this area.

ABACUS
Department of Architecture Strathclyde University
Glasgow
Scotland

Final Report
for Grant GR/F/91919

AN INTERACTIVE SOFTWARE SYSTEM FOR THE URBAN ENVIRONMENT

M.Grant

1.1 INTRODUCTION

Many cities in the UK, and indeed throughout the developed world, are characterised by the all too familiar symptoms of urban blight caused by insensitive intervention in the environment. The common denominator within this class of problem is the lack of a coordinated, integrated approach to the planning, design and maintenance of our cities. The cycle of development and redevelopment involves input from a diverse range of disciplines relating to architecture, civil engineering, transport engineering and the management of city utilities. This lack of a common, updateable, information base renders access to a global view of the city difficult if not impossible.

The research reported here sought to examine the prospect for linking data sets which are currently isolated within specific departments or held externally by utility companies or businesses each of whom are likely to benefit from pooling resources. These data sets are potentially capable of being merged into one comprehensive system with the prospect that the sum of the parts would be worth considerable more than their individual worth suggests.

Naturally the introduction of this strand of the information technologies into local authority usage is not a straightforward matter due to the constraints involved. These include funding, UK central government's enthusiasm and willingness to encourage uptake, information availability and quality, localised initiatives and the influence of EC policy. These factors are currently more pressing than technological advancements, although progress in the relevant technology is also vital in order that systems can continue to meet user requirements and expectations.

Urban Information Systems are certainly a beneficial prospect for the future and indeed are already beginning to be realised within many authorities and councils, but difficulties regarding centralised support, standardisation of data formats and other organisational, ethical and legal questions remain to be resolved before significant uptake and true benefits are to be realised.

1.2 GRANT INTENTIONS

The worth of the concept behind this research had already been proven as an aid to assessing the visual impact of interventions in the urban fabric, the challenge was to investigate the potential for a "value added" system that would enable varied evaluations and representations of the physical qualities and quantities required by urban management strategies. To achieve these aims it was necessary to identify those who would use and benefit from such a system and then to prototype and implement a demonstrator to prove the concept.

1.3 BACKGROUND

The developmental thread of this project is a natural progression brought about by the escalating potential of hardware technologies linked to expanding expectations and desires of the users. In the mid 1980s all that was asked, or expected, from a CAD system was the ability to model a single building, yet in a very short space of time it became possible to create and model ever larger groups of buildings. With this came the realisation that the quality of the information content increased with the quantity. Given a group of buildings more evaluations become relevant, studies of blocking, massing, methods of access and egress all become apparent and the ability to view a larger context makes all these evaluations more telling.

Our initial activities in urban modelling stemmed from a purely academic interest in how large scale urban data-sets could be captured, manipulated and stored. This resulted in the capture, at varying levels of detail, of much of Glasgow's city centre. However interest was such that the model became self perpetuating and now extends to some twenty square kilometres encompassing around ten thousand buildings. The format of the model allows the logical data-structure of three interlinked sets of entities, the underlying terrain, the transport networks and the building stock. Within the limitations of this format (which is discussed later) the model proved adequate for our then current needs which were largely limited to the evaluation of new developments in their surrounding context.

As the model was applied to an ever greater range of problems we became frustrated by it's innate inability to respond to a growing number of queries. While it was ideal as a tool to evaluate and communicate the spatial and aesthetic design issues relating to building development we were increasingly having to turn away requests for other information. These requests related to data that the model did contain, and could potentially provide, yet because it lacked that basic level of intelligence the answers remained locked inside the data structure. The nature of these requests was in essence quite simple; "Can you show the relative proportions of commercial, retail and residential accommodation within a given area?", "Can you calculate the total length of footpath between here and there?". This is all information that is intrinsically held in such a model yet the lack of basic attributes to the geometry required its extraction to be performed by hand, if at all.

The level of interest and usage of the model was a clear indication that a unified urban information tool would find a ready audience in all who were involved in all forms of urban management or development. This tool would provide a virtual abstraction of the city, containing not only the spatial information contained in the geometry but also the ability to harness the diversity of other, existing, urban databases. The goal of this research was to develop the, admittedly Utopian, concept of ISSUE, an Interactive Software System for the Urban Environment.

1.4 URBAN INFORMATION AND MANAGEMENT

In the original plan it was proposed that the schema of urban information could be envisaged as a matrix whose horizontal axis related to potential user groups, from each identifiable user group, the vertical axis could then be associated with the information each group holds and the use they find for it. Although a convenient means of conceptualising the boundaries of the problem, as the research progressed it became apparent that the extent of this horizontal axis was almost limitless. Everyone, from central government through regional and local authorities to even the most casual visitor of a city has some requirement on an information system.

To focus the research it was decided to concentrate on the user group who had both the most to offer to and the most to gain from the proposed system. In the urban arena clearly the local authority is the leading figure. To explore the role and needs of the local authority we formed a collaboration with the Information Science Department at Strathclyde and, through interview and correspondence, concluded the following.

1.5 THE CONCEPT OF URBAN GIS

In an urban context, the wealth of spatially attributable data means that the scope for potential applications is limited only by the imagination. Proprietary systems are already utilised in numerous fields including planning, property management and appraisal, transport planning and routing, policing, health and environmental monitoring, marketing, land use management and in many more related areas. These data sets are often collated independently, either by the bodies concerned or acquired from external sources, notably the applications relate to information handling in both public and private sectors.

UrbanGIS, in this context means a networked system which contains disparate yet standardised data sets linked to what is essentially 3-D cartography. This would enable an indefinite number of data sets relating to the same urban environment to be linked and cross-referenced by their unique spatial co-ordinates, giving added value to each piece of data through the spatial manipulations allowed and their interrelationship with other relevant data sets. Each data set could be maintained by the principle user of that data but the important factor would be that, during any particular query, data from separate data bases relating to the same geographical location could also be accessed and merged.

To consider urbanGIS implementation and maintenance on this scale, the incorporation of data relating to all major aspects of urban management is undoubtedly a huge undertaking, implying the incorporation of massive amounts of data relating to each of the applications mentioned above and also the centralisation of effort. This would most likely be held by one principle coordinating body. Regional and local authorities, as the major collators, processors and disseminators and users of urban information are the natural accommodators and users of urbanGIS. It is a major objective to illustrate that urbanGIS implementations on this scale carry significant benefits for local government planning and policy making processes and therefore to society as a whole.

1.5.1 LOCAL GOVERNMENT AND URBAN MANAGEMENT

Although the general economic climate is primarily the responsibility of central government policies, it is evident during periods of economic growth or recession that variations occur in the extent to which regions are affected by swings between either extreme. It can be assumed that the ability of individual local authorities to successfully manage and coordinate the resources under their control will impact on the relative wealth created and hence the opportunities available to inhabitants as producers and consumers. Consequently the social well being and quality of life for the population as a whole, at a localised level, may be related to the performance of local planning and policy. Moreover, this wealth naturally contributes to overall national prosperity and economic growth.

1.5.2 CHANGING ROLES IN LOCAL GOVERNMENT

Aside from the complexities of urban management, which naturally place increasing demands on management and planning practices, there are other influences which will affect changes in structure, role and orientation of policies. These changes are themselves instrumental in pushing policy makers to use their information resources to greater effect in an attempt to produce better and speedier judgements.

Pressure from central government to be cost effective and efficient in the ability to produce solutions to infrastructure, management and social problems in urban centres is passed down to local authority level along with an increasing burden to police various legislative changes, notable domestic rates, but also in education, social services and housing. These changes, with accompanying financial constraints are part of a shift in emphasis from service provision towards to some form of "enabling" function, where the principle role is that of management within a framework in which services would be provided from the public sector. In order to support such policy it is obvious that information management will play an increasingly important role and, to an extent, will become the essence of good management (particularly the management of change) and decision making practices.

1.6 APPLICATIONS AND USERS

A detailed study of two English local authorities carried out by the Local Authorities management Services and Computer Committee (LAMSAC), evaluating GIS strategy benefits for local authorities, produced the list of potential application areas given in table 1.

The list is extensive but not exhaustive and many more application areas would potentially benefit. It does, however, serve to illustrate the diversity of services and applications which might be integrated into and served by urbanGIS. However this would bring it's own technical and managerial problems that must be addressed. Whether local authorities continue to provide services from within their own departments or to assume an organisational role for outside contractors, in either case they must possess the relevant information.

The principle objectives of an urbanGIS targeted at local government can be tabulated in order to define the extent of information requirements, internal and external users and desired system capabilities. (after Gault and Peutherer) Information systems should not be application specific but should be designed to meet corporate goals. The needs of all levels, and parts, of an organisation must be met, identifiable key groups fall into the following categories:

1/. Senior management and strategic advisors

2/. Service managers and function planners

3/. Service deliverers

4/. Elected representatives

5/. The public

6/. Special interest groups

The following have been identified as key issues.

- The system should be able to provide each part of the organisation with the information required, when it is required, but must avoid both information overload and information starvation.
- Information has to be available to be used selectively and used to measure progress towards planning objectives
- The system should be integrated, interactive and networked to allow mobility of data between source and user.
- Such systems should be able to provide information on the entire gamut of spatial units of interest to the organisation.
- The system should be capable of holding or generating, past and current information in order to allow time series analysis. The system should also contain simulation modules to enable trend extrapolation and policy testing.
- The data must be of good or at least defined quality.

1.6.1 SPATIAL DATA CONCEPTS

Geographical Information Systems have been a major advance in land management. Their success is in no small part due to the ability to communicate the information via associated mapping packages. Given that the urban model can be considered an extension of conventional cartography, enriched by the addition of the third dimension, then this spatial data set must be seen as an intrinsic part of the system. The urban fabric is the single common denominator that associates all the disparate data that may lie underneath and the ability to spatially reference attributes is a prerequisite. The geometric information affords a number of advantages; firstly since it is a potentially accurate digital representation it can itself be focus of evaluations, in terms of area, height, volume, orientation etc. Secondly, because it is a "natural language" representation it can be easily used as both an interface to the layered information beneath and, at the other end of the process, a vehicle for conveying more abstract information.

1.6.2 GEOMETRICAL DATA CAPTURE

The original Glasgow model was pioneered in order to study the methods of data capture and representation required by such large-scale endeavours. This activity is potentially the most time consuming as no single source can provide all the data. The methods employed covered source material ranging from cartography, aerial photography and street surveys through to historical archives. However, an increasing amount of data is available in a digital format. The Ordinance Survey (OS) basic scale plans cover the scale range 1:1250 to 1:10,000 and comprise a total of over 220,000 sheets. 1:1250 is the largest scale of publication and the 55,000 sheets in the series cover the major urban areas.

A collaboration with the OS and LaserScan, a commercial company with a long history of digital cartography, lead to an investigation as to how the OS digital 1:1250 maps could be used as source material for three-dimensional urban geometry. The OS maps are available as raster images or vector datasets. Increasingly the vector data is being made available in a structured format. Currently much of the data is unstructured; for example several records may separate parts of the same feature. However, demand from CAD systems means that the provision of structured data will become more available. Structuring the data requires that separate vectors defining features are "structured" into polylines or, if applicable, closed polygons.

In order to exploit this source we developed a filter that was capable of taking structured output from OS data or associated packages, stripping unnecessary information, and adding inter-record headers where required. The output being a series of geometrical entities formed by the base polygon, where each vertex is given a 'Z' co-ordinate at ground level and a mean height inserted into the roof level field of the header information. When this output is accessed by the system it results in a scene that is surprisingly realistic, considering the effort involved. The use of OS data ensures that the building footprint is highly detailed and each building is accurately positioned with reference to the terrain. Obviously roof profiles are generated by flat planes, which is seldom the case in reality, yet the model gives a recognisable impression of the distribution of buildings in the given area. Further development requires an increasing investment in fieldwork, especially to gather accurate information on roof heights and formation. It is estimated that one working day would have to be invested to gather the heights of all building blocks contained on a 1:1250 scale sheet. The use of an Abney level at a paced distance from a known location would produce an accuracy of +/- 0.5m in the height of eaves or ridges relative to the building base. A further man-day would be required to process the geometry and patch the data into the file listing. The resulting data set would then be adequate to represent the urban context for most needs Further refinement is increasing expensive in terms of manpower, to provide architectural embellishment calls for a greater width of source information and, in general terms, the effort involved is in direct proportion to the quality and quantity of this provision.

1.7 SOFTWARE IMPLEMENTATION

At the outset of the project the Glasgow urban database did exist, yet it was in a format that required significant knowledge and skill to utilise. The prime aim in assembling a demonstrator was to increase the ease with which it could be assembled and accessed to those without the experience of proprietary tools and operating procedures. With the prospect of urban geometry becoming increasing available from sources like the OS the methods of creation, manipulation and storage of such quantities of data are of paramount importance.

1.7.1 INTERFACING

Through experience an obvious method of control suggested itself. This was to maintain a hierarchy of detail, accessing the geometry through a series of interfaces similar to the familiar atlas concept. At the highest level a user is presented with a graphic representing the city in broad outline. Glasgow's gridiron road layout within the city centre lends itself well to this concept. A user can then descend through this layer, being presented with ever increasing detail until the base level of building entities is reached. The user can browse the interface at any stage, accessing further details, building names or addresses, with a mouse pick. In this manner the user can ascend and descend the hierarchy accessing data at different levels of definition.

1.7.2 REPRESENTATION

In order to interface with such extensive geometrical data sets there has to be a hierarchy of representation depending on the spatial scale of the application. This has meant that such a model should exist on a number of levels of representation. If the intervention was on a citywide scale then a basic, simplistic model is required. If however interest is focused on a specific area then a greater level of detail must be adopted. Determining what level of representation to adopt is a major factor in the usability of the system. Too much and response is slow, too little and not enough information is present. Very early on it was realised that even if the model was totally accurate in plan, users still had great difficulty in orientating themselves and so minimal detail was added to "landmark buildings" the stations, churches and civic buildings. This led to a further level of detail in which buildings were just caricatures of themselves again accurate but minimalist in terms of representation. In a typical application the basic levels of detail would form the context for the intervention, minimal detailing would be applied to the immediate surroundings and the object of attention would be represented by the fully detailed architectural model. By maintaining this layering of detail large areas can still be manipulated without incurring unwelcome overheads.

1.7.3 ATTRIBUTION

The geometry is attributed with spatial co-ordinates relating to the OS grid reference system. Each basic unit carries a reference to the 1:1250 map sheet which the may be broken down as the urban layout best suggests. Once the level of an urban block is reached each building is individually identified. This layout uses the file system as a structured database, a hierarchy of directories maintaining a unique path and filename for each building. For every database field one record will point to this filename, equally every file has a field relating it's spatial reference to common data. By this means a query can originate from the graphics, a "what is that?" style pick of a building, or may originate from a query relating to age, usage or other such data, and traverse the databases in both directions. This was the prime aim of the implementation, to be able to respond to the "show me" type of question. Here the graphics are indispensable as a means of communication.

1.7.4 DATABASE FUNCTIONALITY.

The application was built on a range of machines to prove that the concept could work at the low, as well as the high end of technology. Although these implementations were "stand alone" in that they did not access a central resource, it did serve to prove that the concept was portable. The low-end implementation was directed at IBM compatible PCs running Dbase4 and our own graphics front end. Although the graphics were limited to 2-D maps the system proved that it could provide the necessary functionality. On UNIX workstations the system was built around an ORACLE database and could accept much greater functionality. An SQL interface complemented the graphics and while the attribute data was limited, the limitations were due to resources available and not to the limitations of the system. The image in figure 3. represents one stage in the evolution of the system. The features available demonstrate that the implementation can access graphics from the 3-D dataset, attribute information from the relational database and multi-media data sets in terms of scanned raster images and even audio data. In a multi-tasking environment the interface can co-ordinate a range of applications, invoking geometry editors to manipulate the data, filters to provide import/export functionality and access to associated peripherals. Given a suitable programming resource the technology that is readily available today would be well matched to the requirements of the proposed system and would not, as is often the case, limit the application.

1.8 ACHIEVEMENTS

A review and classification of the users of urban information systems, the nature and format of the data held and the range of uses to which it is put.

The establishment of a suite of software tools to enable the capture of, and access to, data relating to the formal and functional properties of the urban environment.

This project continues to attract considerable interest from a varied selection of potential users and commercial developers.

A number of trial studies have been made relating to potential usage of the proposed system.

Community Architecture. This was a collaboration between ASSIST Architects, Scottish Holmes and the local community residents association. The objective was to use the Urban Database as a means of communicating the proposed redevelopment strategy (Hutchesontown phase three) to the community and to provide a focal point for informed discussion on the feedback obtained.

Data Capture. A successful collaboration with the Ordnance Survey and LaserScan (the UK's largest digital mapping enterprise) has resulted in the ability to take OS digital data for 2-D maps and to reconstruct the three dimensional terrain, road networks and buildings. This is a major advance in that the ability to capture the urban topology is now faster and more accurate as well as both easier and cheaper. Future collaboration is planned to extend LaserScan's GIS system to include an urban element and to give the OS a visualisation capability regarding their digital data.

Urban Renewal. We are also collaborating with the Edinburgh Old Town Trust to implement an urban database describing the Old Town area of Edinburgh. This area is sensitive to urban interventions and the proposed database will be used to monitor new proposals. Graphical material from the topology will be used in a permanent exhibition relating the historical development of Edinburgh to the present and future cityscapes.

Evaluations. A collaboration has been proposed between ourselves and the Radio Communication Research Group at RAL. The objective being to use the formal representation of the built environment to simulate the propagation of radio waves from local transmitters. An exchange of data and software has already taken place and interest has been shown by INTELSAT on a future development project.

Development. Development of the software continues in response to the feedback from the test-cases studies.

Dissemination. Having identified research groups with similar interests we have hosted a seminar with local and international participants. Contact is being maintained with all parties who have expressed an interest. A presentation, and demonstration, on the current work has been made at the Society of Cartographers conference at Glasgow University in September 1991, and similarly at the ITA theme meeting at Abingdon in October 1992.

1.9 CONCLUSIONS

Local administration has the most to offer and the most to gain by moving towards an integrated philosophy regarding information collection, collation and dissemination. The motivation to move towards this goal comes from a number of sources, primarily through the increasing complexity of urban management but also due to central government policy to progress towards the decentralisation of services. Fiscal pressure to increase efficiency, lower manpower resources and arrive at speedier judgements all point towards an increasing reliance on the information technologies.

The quality of planning and decision making processes can be substantially improved when valid data is appropriately and efficiently handled. Relevant information systems which support the activities of planning and decision making must be based on a thorough and clear analysis of the planning processes adopted. The use of an urbanGIS allows information to be assembled and applied in new ways. It offers practical means to manage large and diverse spatial databases and provides effective tools to understand relationships between disparate phenomena. An increasing number of decision makers and managers have recognised that these technologies will be essential if they are to address the expanding mandates and complex decisions they now face.

The use of a graphically orientated urbanGIS system has a threefold advantage. Firstly, since 80% of urban information can be spatially referenced it is a logical format under which disparate data sets can be collated. Secondly, communication is of paramount importance. The ability to use a readily understandable and recognisable medium with which to communicate proposed strategies and outcomes will make the task of disseminating policy open to all. Finally the use of a digitally accurate urban model will open up greater roles for simulation based tools.

An increasing knowledgeable and aware user base will be a significant factor in the uptake of urbanGIS. Users will demand systems that accept data in diverse formats from existing databases, that they be easier to use and present higher levels of performance. Market forces will provide the incentive for the incorporation of advanced graphics, simulation and animation facilities, faster processors, more competent data exchange and query functions and expert system shells to reduce areas of complexity identified by the users.

Developments from all corners of the information technologies will aid the uptake of urbanGIS, emerging standards, the move towards open systems, increased networking infrastructures and the every greater availability of data in digital formats will increase the mobility and sharing of databases, These factors coupled with a growing user community will create an environment that will foster an increasing growth rate.

The issue of financial justification will be important, especially in a local authority context. Although hardware cost are on a downward trend, many of the costs associated with the system remain high. Budget expenditure will be greatest at early stages while many benefits will only accrue once the system is in widespread usage. A complex issue is one of asset management. The better the existing internal asset management then the more difficult is the justification for new technology. For example, the authority which has recently invested in high quality OS paper map coverage of it's land coverage may find it difficult to justify recapture in a digital format.

Given all the factors that influence the success of an implementation such as that discussed here, the technical aspects of authoring the software and targeting a suitable hardware platform, are undoubtedly in the minority. The real challenge is in the legal, ethical, political and economical barriers to promoting the sharing of data and encouraging centralised management of assets. Even within the tight boundaries of local and regional authorities the duplication of effort, the adoption of mutually incompatible formats and the lack of willingness to pool resources all tend to negate the potential benefits. In short, the only remaining barriers to ISSUE are essentially human ones.

Society is increasing dependent on information, in all forms, it is easy to see the move towards a critical mass where enough interested parties hold enough information to make the move towards the philosophy of centralisation of information assets self perpetuating. Once this occurs then the potential benefits may be realised.

ABACUS
Department of Architecture and Building Science,
Strathclyde University,
Glasgow

URBAN GIS - THE APPLICATION OF THE INFORMATION TECHNOLOGIES TO URBAN MANAGEMENT

M Grant

ABSTRACT

This paper describes a current theme of research whose objective is to identify, and then prototype, a relevant urban information system. It is proposed that by attributing a geometrical framework with those physical quantities that are relevant to the formal and functional evaluation of the urban environment, the means of evaluating the qualities and quantities of the buildings as well as the social and economic prospects may be realised.

Keywords: Urban, Information, Databases, Geometrical models, CAD

1.0 INTRODUCTION

Many cities in the UK and indeed throughout the developed world are characterised by the all too familiar symptoms of urban blight caused by insensitive intervention in the environment. The common denominator within this class of problem is the lack of a coordinated, integrated approach to the planning, design and maintenance of our cities. The cycle of development and redevelopment calls for input from a diverse range of disciplines relating to architecture, civil engineering, transport engineering, and the management of city utilities. This lack of a common up datable information base renders access to a global view of the city difficult, if not impossible. This problem has provided the motivation to move towards an integrated philosophy regarding information collection, collation and dissemination. The impetus is provided primarily through the increasing complexity of urban management but also through central governments policy to progress towards decentralisation of services. Fiscal pressure to increase efficiency, lower manpower resources and arrive at speedier judgements all point to an increasing reliance on the information technologies.

1.1 URBAN INFORMATION SYSTEMS

Modern cities are part of a competitive environment. They must compete for inward investment both within their own national boundaries and abroad. To be competitive they must provide the infrastructure, amenities and the aesthetic environment that is attractive to both commerce and the additional human resource required by relocation. In times of economic recession this competitive pressure is at it's greatest, yet it is obvious that some regions prosper more than others. It can be assumed that those regions who are more capable in the management and co-ordination of the resources under their control can influence the relative prosperity and opportunity available within their jurisdiction. Consequently we can state that local government performance in terms of planning and policy relates directly to social and economic satisfaction within that area.

Aside from the increasing complexities of urban management, which naturally places an escalating burden on management and planning practice, there are other factors which will affect changes in structure, role, and orientation of policies. These changes are themselves instrumental in pushing policy makers to use their information resources to greater effect in an attempt to produce better and speedier judgements.

In the UK it has been central government's policy to decentralise the responsibility for policing legislative changes down to local authority levels. These changes are part of a shift in emphasis from service provision towards some sort of an enabling function where the principle role is that of management within a framework in which services would be provided from the private sector. The aim is to be cost effective and efficient in the ability to produce solutions to infrastructure, management and social problems in urban centres. However in order to support such policy, it is obvious that information management will play an increasingly important role and to an extent will become the essence of good management and decision making practices. These factors all point towards an increasing reliance on the information technologies (Gault and Peutherer 1990).

A Utopian view of the ideal urban information system would be an open system which would contain disparate yet standardised databases linked to what is essentially 3-D cartography. This would enable an indefinite number of databases relating to the same urban environment to be linked and cross referenced by their unique spatial coordinates, giving added value to each piece of data through the spatial manipulations allowed and their interrelationship with other relevant data sets. Each data base could be maintained by the principle user but the important factor would be that during any particular query, data from disparate sources relating to the same geographical location could be accessed and merged. In an urban situation the wealth of spatially attributable data means that the scope for potential applications is almost unlimited. Proprietary systems are already utilised in numerous fields including planning, property management and appraisal, transport planning and routing, policing, health and environmental monitoring, marketing, land use management and in a great many other areas. These data sets are often collated independently, either by the bodies concerned or by external sources.

The quality of the planning and decision making processes can be substantially improved where valid data is appropriately and efficiently handled. The use of such an urban information system allows data to be assembled and applied in new ways. It offers a practical means to manage large and diverse spatial databases and provides effective tools to understand relationships between disparate phenomena. An increasing number of decision makers and planners have recognised that these technologies will be essential if they are to address the expanding mandates and complex decisions that they now face (McLennan 1992).

The use of a graphically oriented information system has a threefold advantage. Firstly since 80 percent of urban information can be spatially referenced, it is a logical format under which disparate data sets can be collated. Secondly, communication is of paramount importance. The ability to use a readily understandable and recognisable medium with which to communicate proposed strategies and outcomes will make the task of disseminating policy open to all. Finally the use of a digitally accurate urban model will open up greater roles for simulation based tools.

1.2 AN APPLICATION

The Edinburgh Old Town Trust is a subsidiary of Lothian and Edinburgh Enterprise Limited. The Trust was formed in order to monitor change and to help analyse the environmental impact of redevelopment The main aim of the Trust is to champion the residential and local business community and offset their needs against those of commerce and tourism.. The Old Town is a conservation area which infers both special status and special needs, the Trust represents these needs and serves to co-ordinate the delivery of services and the progress of long term planning. Over the years the Old Town has maintained it's distinctive character but there is now an urgent need for renewal and the area faces a period of sustained redevelopment. A prime requirement is now for a tool that will aid the Trust communicate proposals, co-ordinate planning and assess their impact on the local environment and population.

The commission given to ABACUS by the Trust was to build a 3-D model of the Old Town for use in visual impact studies and then to elaborate this data base with relevant information in order to integrate the system with the Trust's current working practice.

1.3 GEOMETRY

In order to provide the geometrical basis for the system, a collaboration between ABACUS, LaserScan and the Ordnance Survey lead to the development of a set of software tools capable of extracting 3-D geometry from digital format 2-D cartography (Jones 1991). Once the filters were developed this gave the facility to generate in block form, the basic structure of the urban topography. By using aerial photography, street surveys, and existing architectural drawings the relevant building details were added until a faithful replica of each building had been modelled.

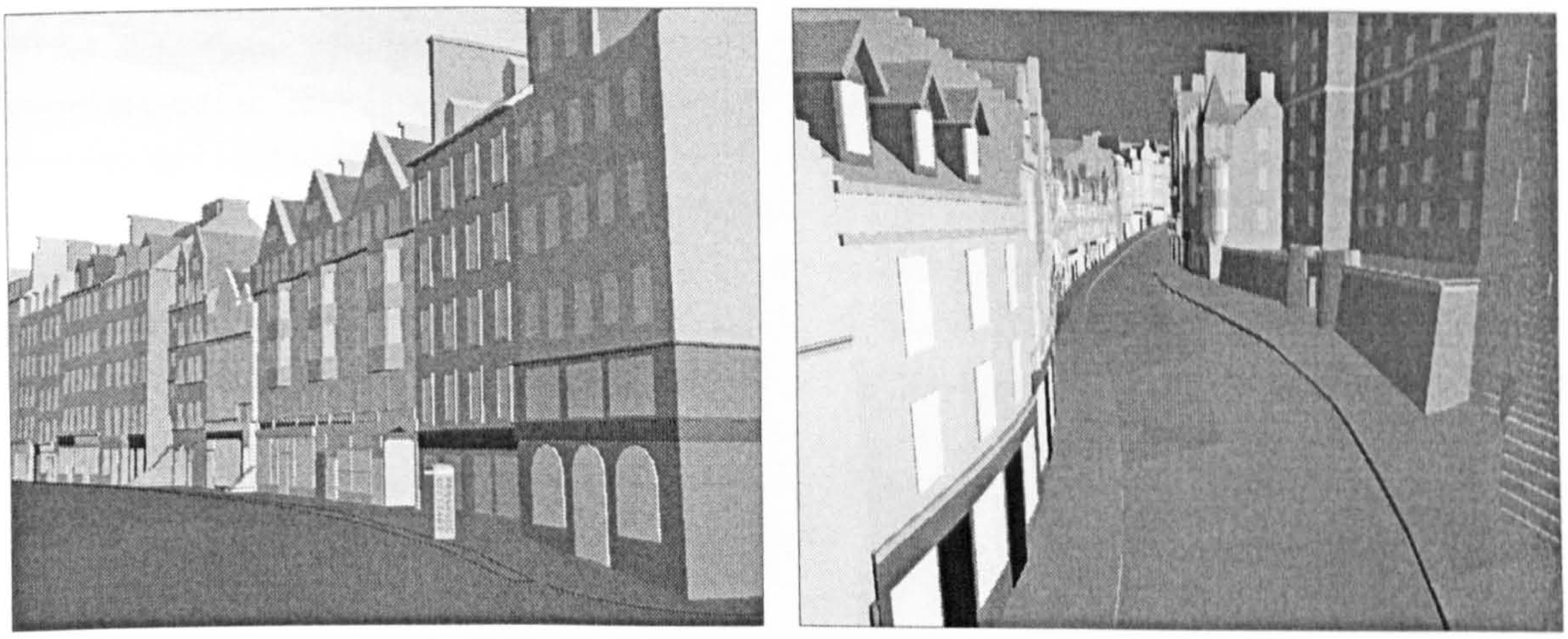


Figure 1 Streetscape showing granularity of data.

1.4 ACCESS

Even within the limited boundaries of the Old Town conservation area the wealth of data in enormous. This leads to a necessity of providing a vertical hierarchy of detail in order to simplify access [Figure 2]. Each building exists in three levels of representation: as a basic block accurate in plan and elevation; the same block with roof detail; and as a fully detailed architectural model. The level of detail [LOD] can be toggled to allow access to a contextual area of the urban environment without the burden of carrying unwanted or unnecessary data. The base LOD is used to automatically generate a 2-D representation which forms the basis of the interface and allows a global view of the entire data set.

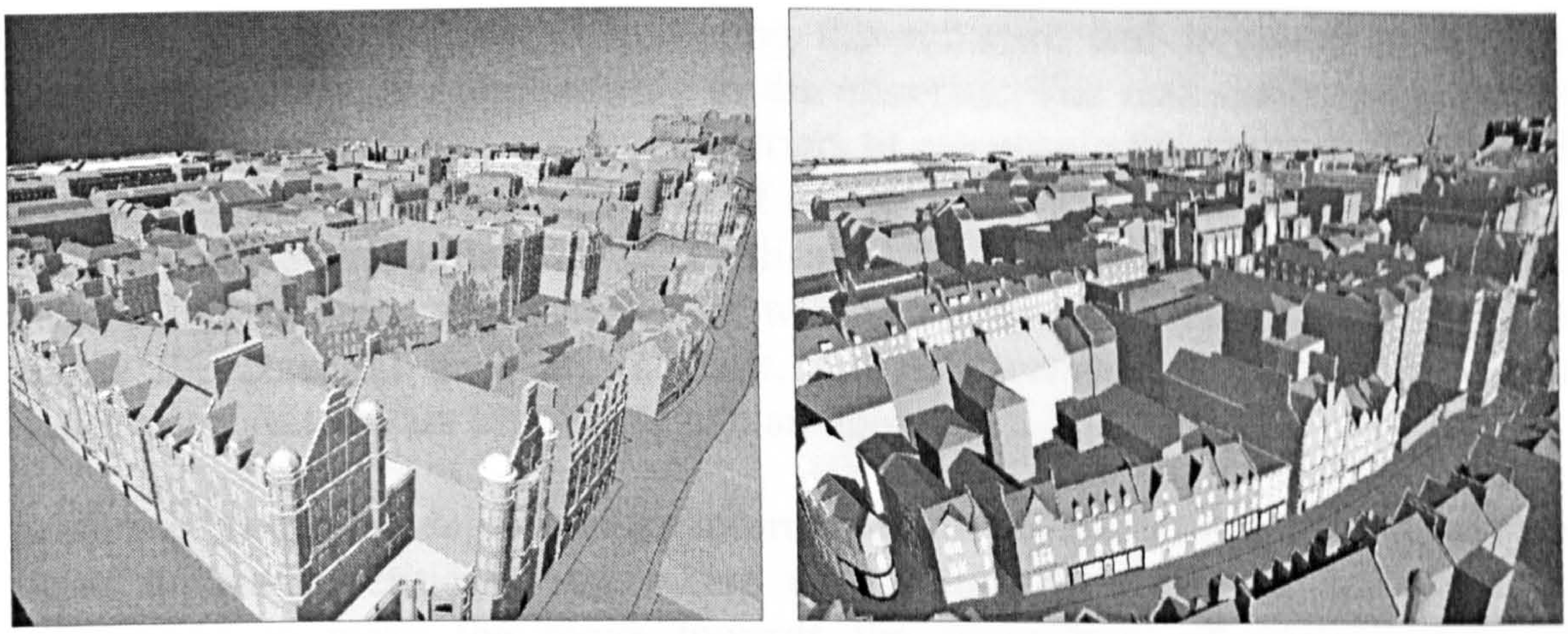


Figure 2 Edinburgh Old Town Conservation Area. A view of half the database.

1.5 DATABASE

This interface allows each building to be accessed independently. The 2-D representation forms a map which can be easily manipulated, allowing picking via mouse and cursor to access further levels of geometrical and related data. Unlike a standard data base implementation where each record would refer to a building, using the geometry as an interface allows each building to be a record, therefore a pick on a building returns a pointer to the database fields relating to that location. Database functionality is implemented in software allowing a standard access paradigm of fields collated on Boolean, lexical and numerical operators. Access to different subject data bases is enabled by making the interface self configuring in that a template file can be loaded to set out the structure of the fields for each set of records. The data sets can contain graphical images, voice information or standard alpha-numeric text.

1.6 USAGE

The structure described above facilitates access to all types of information. A user can browse the data sets looking for objects of interest or can ask pointed questions through the query interface. The added attraction is that the subject of investigation can be depicted graphically allowing the outcome to be readily identified, evaluated and communicated.

1.7 CONCLUSIONS

Developments from all corners of the information technologies will aid in the uptake of similar systems, emerging standards, the move towards open systems, increasing network infrastructures and greater availability of data in digital formats will increase the mobility and the potential to share data bases. Given all the factors that influence the success of an implementation such as that discussed here, the technical aspects of authoring the software and targeting a suitable hardware platform, are undoubtedly in the minority. The real challenge is in the legal, ethical political and economic barriers to promoting the sharing of data and encouraging centralised management of assets. Even within the tight boundaries of local and regional authorities the duplication of effort, the adoption of mutually incompatible formats and the lack of willingness to pool resources all tend to negate the potential benefits. In short, the only remaining barriers to urban information systems are essentially human ones.

Society is increasing dependant on information, in all forms. It is easy to see the move towards a critical mass where enough interested parties hold enough information to make the move towards the philosophy of centralisation of information assets self perpetuating. Once this occurs, then the potential benefits may be realised.

REFERENCES

Gault, I and Peutherer, D. 1990. "Developing Geographical Information Systems in local Government in the UK: Case studies from Birmingham City Council and Strathclyde Regional Council", in L. Worrall (ed) Geographical Information Systems: Developments and Applications. London. Belhaven Press.

McLennan, A. 1992. "Urban Information Systems". MSc Thesis. Strathclyde University.

Jones, A. 1991. "Urban Geometry Development". MSc Thesis. Glasgow University.

ABACUS
Department of Architecture and Building Science,
Strathclyde University,
Glasgow

MULTIMEDIA
-A MULTI PURPOSE
PROGRAMMING ENVIRONMENT

M Grant

ABSTRACT

In the few short years since the emergence of multimedia programming tools this activity has moved from the periphery of the Information Technologies to the mainstream of computing applications. This is due not only to the progressive development in hardware and software technologies but also to the escalating set of desires of authors and users of multimedia products. Perhaps the most interesting theme within this strand of development is in the progression of the capabilities of the scripting and programming capabilities now on offer. The purpose of this paper is to trace the development of this aspect and speculate on the future of multimedia authoring tools as a new generation programming environment where the distinction between multimedia and CAD becomes less well defined.

1.0 INTRODUCTION

Over the past five years we have seen the exponential growth of that fusion of hardware and software products that comprise the technology known as multimedia. Taken individually these products are not exceptional however what is revolutionary is the way that when they are applied in conjunction the combination can open up new horizons. Approached at a simplistic level multimedia can be compartmentalised into three categories, hardware, software and the human element of design. The hardware provides the platform on which to integrate different media, the software the tools to control it, and functionality and purpose come from within the design. As with all aspects of the Information Technologies (IT) nothing stands in the way of progress and it is all too easy to see the rapid advances in hardware and software. However it is in the progress in the design and application of these new generation products that perhaps the most interesting advances are seen and hence form the focus of this paper.

2.0 MULTIMEDIA

Looking at the scope of applications that come under the heading of multimedia it is apparent that this terminology means "all things to all people". It is commonplace to find the most humble electronic slide show alongside a production featuring fully integrated video, audio and animation, both residing quite comfortably under the title of "multimedia". Obviously each serves it's own purpose to good effect and if we look at some of the interim stages between these extremes it is easy to chart the development of all three of the essential ingredients of multimedia, the hardware, software and the element of design, then use this progression to extrapolate beyond the current state of the art and forecast what the future holds.

3.0 PROGRESS

The progress in multimedia applications can be mapped out through a series of landmarks. The first of which has its origins in sequencing media and is essentially the electronic slide show. The natural evolution of the product from this point is evidenced in the development of the ability to navigate, first in a linear sequential manner, then bi-directionally, leading on to hierarchical structures and their more complex composite offspring.

Given that the contents of the slide show remain unchanged during the above evolution it is apparent that the increased ability to navigate adds functionality, not least in the users ability to sequence data as the user desires. It can be seen that the development of this theme leads to the next milestone, one that is perhaps typified by the "Architectural Archive". Essentially this category is just the components of the slide show allied to a navigational structure that allows functional groupings of data which, given these groupings, gives added value to the content.

This strand of development is finite as there is a limit on what can be achieved by purely sequencing media on some time line. To understand this point consider the classic problem of creating a virtual space where the objective is to allow the user to roam the environment by moving through the space, changing position and direction of view at will. Traditionally this is achieved by creating two matrices of images. The first is the positional matrix which may contain 20 positions, say five in the East/West axis and four in the North/South axis. Associated with each index pointing into the positional matrix is a direction matrix allowing four views, say North, South, East and West. This gives a sum of eighty images which represent the total degree of freedom allowed to the viewer. In order to interact with this virtual space it would be necessary to provide two sets of controls allowing the user to move forward and back, left and right and to view North, South, East and West.

Logically there are potential 80x4x2 individual interactions that define the navigational structure of the problem. To sequence this interaction by tying explicit actions to 640 buttons imposes obvious limits to the scope of the technique. However if the authoring environment supports a scripting language then the entire navigational problem is reduced to four conditional case statements and a simple mathematical formula.

This then opens up a whole new horizon in the functionality of multimedia applications in that the constraints of purely sequencing media are overcome and the tools can be applied in a more creative manner.

4.0 A NEW GENERATION

While there are many examples that illustrate the progression of this theme it may be useful to consider, as the next milestone, an application where multimedia tools are used to recreate existing software applications. As an example consider the need to provide user training on a complex, multi-user system containing critical data. It may be inappropriate to give a novice users access to such a system yet hands on experience is vital in order to train staff in it's operation and maintenance. One solution is obviously to provide access on a duplicate system where mistakes would not corrupt data and the trainee could explore in safety. However in an expensive hardware environment the cost of providing a machine base purely for training may be prohibitive. A solution that is increasingly being applied is to represent this software environment by recreating the application using those tools commonly used for authoring multimedia. The visual characteristics of the application can very quickly be captured by utilising screen dumps from the original and interactivity can be added by using the functionality inherent in the multimedia authoring package. In this manner the target program can be "cloned" with little effort. The overall functionality of this clone will be restricted, yet with careful attention to the composition of the "media" full functionality for a limited number of routes through the program can be simulated. This strategy can result in the creation of applications that totally mimic the look and feel of the real thing and yet in terms of time, effort and productivity cost much less. Similar economies can be obviously be found in the ability to rapidly prototype and test the interface design for new applications.

In the above example the multimedia application was used as a vehicle for presenting pre-recorded material, in response to user interaction, to simulate the operation of some existing program. The next developmental milestone is an application that is not restricted to the play back of pre-ordained material but functions, in a limited capacity, in it's own right. A typical example might be the provision of a graphical front end to an existing simulation program. Many simulation programs, particularly from an engineering background, still lack the user interfaces that are now taken for granted in most other disciplines. These programs operate either in batch mode or at a very low level of user interaction. The ease with which multimedia tools can create such a graphical interface is their greatest strength and it is easy to see the possibilities inherent in using multimedia applications to perform data preparation to assemble an input stream that can then be used as input to older, less friendly applications. In this instance the user's prime task is to select parameters and data sets that will be used to control the eventual simulation. The capabilities of even the most modest multimedia authoring are tailor made for this aspect of functionality in that they have been designed with the prime aim of providing the tools to facilitate this very task.

The same argument is equally true at the other end of the process, in that the capabilities of programs created in multimedia authoring packages can be applied to tasks related to data recovery and presentation. Again this allies the functionality of the user interface with the computationally less demanding tasks of transferring data from file and presenting it on screen. So, although these examples are not ground breaking in terms of the functionality offered, they are genuinely useful and represent a major step in the progression from the earlier instances.

5.0 THE FUTURE

The final stage in this progressive development is to dissolve the barrier between conventional programming and multimedia scripting. There is no longer such a wide divide between mainstream programming languages and the scripting capabilities of the top end multimedia packages and with the computing power offered by the average PC or Macintosh it is not inconceivable that future multimedia tools may not be regarded as just an authoring package but as a fully fledged programming environment. The next generation of multimedia tools will feature compiled rather than interpreted scripting languages thus unleashing much enhanced processing power. If this comes about the product will be a unique blend of a high level programming language closely bound to an integrated and capable graphical subsystem. This vision of the future is already close at hand, even current multimedia tools provide enough of the required functionality to program simple yet robust applications that do not rely on pre-recorded media for their operation and are capable of performing simple numerical simulations. The undeniable attraction of this scenario is in the productivity offered by the high level approach to programming both in numerical and graphical areas. Compared with the learning curve and skills required to create comparable programs in conventional programming environments, such as X Windows, the cost effectiveness of this approach is very attractive.

6.0 CONCLUSION

Several years ago, when users were first introduced to the then revolutionary capabilities of multimedia titles, even the most simple capabilities were impressive. Yet as the technology has moved from the sidelines and into the mainstream of IT it is apparent that in today's environment titles must be faster, more capable and offer much more functionality to capture the same level of interest. This leads to an ever greater reliance on the programmability of multimedia tools as designers look for greater control over the media that they employ and will undoubtedly force the technology towards the future as suggested above. Similarly we can expect advances in hardware to influence the future of multimedia production, already it is possible to see the trend towards networking capabilities and as these aspects of functionality come on stream they will then influence the capabilities of the multimedia environment.

It is not difficult to see the future of multimedia authoring tools as a new generation programming environment where the distinction between multimedia and conventional CAD becomes less well defined. This will come about because of progress in hardware technologies, integrating functionality into a single platform, the ever increasing ability of software tools to access this functionality and the creativity of designers to utilise these developments in the production of increasingly more sophisticated and capable multimedia products.

Indeed, it is perhaps true to say that the future of CAD lies in multimedia.

ABACUS
Department of Architecture and Building Science
Strathclyde University
Glasgow

VIRTUAL HERITAGE

M.Grant

ABSTRACT

This paper gives an account of a productive collaboration between the ABACUS research group at the University of Strathclyde and Historic Scotland - custodians of around 2000 magnificent examples of Scotland's architectural heritage - in the application of information technology to the interpretation, and appreciation, of our stunningly sophisticated Scottish ancestry.

1.0 INTRODUCTION

The ever expanding growth in world tourism is raising new issues within the tourism and leisure sectors. In particular, there is a growing concern about the conservation of our heritage and an increasing focus on the need to control the physical impact of tourism on heritage sites. A further goal is to integrate the need for conservation with finding an alternative option to meet the demands of future tourism.

As the life span of the general population increases there is a commensurate increase in the proportion of leisure time, this combined with, on average, a greater level of disposable income has made tourism the fastest growing industry in the western world. Some reports predict that the tourism sector of the leisure market will become the largest global industry not long after the millennium. This does pose new and increasing acute problems for those charged with the twin tasks of both exploiting and conserving our architectural heritage. While these roles have gone hand in hand for many years there is a very real danger that as a critical mass on the demand side is exceeded, these two functions will become mutually exclusive. Heritage operators are now looking to the information technologies as a way of addressing these needs.

1.1 THE CASE FOR A VIRTUAL HERITAGE

Evidence suggests that much of this concept is indeed already with us. A major swing towards new media can be gauged by the presence of almost one thousand museum/heritage sites on the World Wide Web. These statistics are drawn from 1995 figures, there are doubtless far more now. Although a high percentage may be considered as purely advertising, the necessity for this IT based approach will continue to ensure that this activity is demand driven. The provision of new media directed towards this domain serves to address a number of problems.

Education, particularly in architecture and especially in design, is rooted in the past. Much of architectural learning is retrospective, even modern design methods and practices are based on precedents, students are encouraged to look to the past for both inspiration and evidence of good practice. The provision of exemplary subjects has been the goal of many educational architectural texts and it is plain to see that IT has a role to play in the move away from traditional paper based formats to more modern electronic architectural/historical archives. While this format has been exploited to provide newer and better methods of indexing this information this is clearly not enough and there is still a need to address the subject matter in a more complete, experiential fashion.

Archiving, every year many important building are lost through warfare, fire, natural disasters or simply by neglect. In many instances little remains to inform future generations of past glories. Proposals are being discussed, certainly within the United Kingdom and probably in other countries, to introduce new legislation that would seek to ensure that sufficient data is collected to be able to recreate buildings, if not materially then at least virtually.

Remote Access, there are many more interesting historic sites and buildings than any one individual could hope to visit. The ability to gain remote access through a telematic network such as the World Wide Web is as important to the researcher as it is to the casual browser. This concept is materialising under a number of initiatives such as the International Council of Museums Computer Interchange of Museum Information project and the activities of the European Remote Access to Museum Archives consortium. The inverse of this problem is also addressed in that since many historical sites get relatively few visitors remote access could provide tangible benefits to the curators, by raising their popularity in the virtual, if not the real, world.

Interpretation, access to information, artefacts or buildings is just one aspect of the problem associated with learning from our heritage. To gain full value from these initiatives there must be an attempt to explain their significance, to reunite disparate items with their context and interpret their cultural value. This is perhaps where the new media can show the most benefits via it's flexible approach to amalgamating media and through this composite medium offer new contexts, interpretations and viewpoints.

Conservation, one of the most pressing needs, certainly as regards the architecture heritage sites, is that of conservation. Even with today's levels of visitors many popular sites are recording annual visitor numbers in the hundreds of thousands. While this is very gratifying for the curators they recognise that this level of throughput is unsustainable due to the "erosion" on the fabric of buildings. Construction materials, especially those that are hundreds, if not thousands, of years old cannot physically withstand the passage of so many people. It is not uncommon for many important sites to be forced to exclude visitors for much of the year purely in the interests of preservation. This is obviously in total contradiction to the main purpose of maintaining heritage sites and as such becomes a problem to which a solution is sought with increasing urgency.

An Example:

Skara Brae, situated near Stromness in the Orkney Islands, off the Northernmost coast of mainland Scotland, is Northern Europe's most extensive and elaborate Neolithic Village - an extraordinary complex of seven houses leading off a covered passageway, revealed for the first time in 1850 as a result of a massive (and fortuitous) storm. The completeness of the houses - with stone furniture - offers an extraordinary, if incomplete, insight into the life of Scots some 5000 years ago. The site is one of many that are under the custodian ship of Historic Scotland. The objective of this project was to assess the suitability of the new media in offering an alternative to the traditional presentation formats in offering an interpretation of the site for the visiting public. The site itself exhibits most if not all of the problems associated with heritage sites as discussed above. It is remote, Orkney being a small island in the extreme north of Scotland, conservation is a strong issue, as visitors are already being restricted in terms of access to some areas of the site and there is need for explanation and interpretation as much of the original building fabric has disappeared over time.

1.2 OBJECTIVES

- The objective of the project is the development and preparation of computer based presentational material which can be used to enhance the visitors' understanding of the Skara Brae settlement.

- Due to a growing concern about the conservation of our heritage and an increasing focus on the need to manage the physical impact of tourism on heritage sites it is not possible to allow public access to all areas of the settlement. The presentation is designed to take full advantage of the introduction of Interactive Media and especially the role of Virtual Reality in mitigating these constraints.

- The presentation sought to place Skara Brae in the context of Orkney's pre-history as well as dealing with the site's more recent history in terms of it's discovery and subsequent excavation.

- Visitors could be able to explore the village in the company of an expert "interpreter" as well as in an interactive mode. A choice of "interpreter" would allow visitors to gain an overview of the site from different perspectives such as that of an archaeologist, historian or perhaps an original villager.

- Features of special interest are highlighted and explained. These features can be approached on two levels, differentiating between what experts think is unquestionably true of the village and the life of it's people and secondly offering a best guess as to what was thought most probable.

- The interactive nature of the presentation should seek to encourage a visitor's curiosity and help answer the questions that this remarkable village may have provoked.

- Parts of the village that have succumbed to the ravages of time could be reconstructed. This aspect could explore the methods of construction and materials used. It should be possible to experience a recreation of a part of the village as it would have looked 5000 years ago.

- A rich collection of objects have been recovered from the site, these should be reintroduced to their original locations and a mechanism should be provided to encourage visitors to "virtually" touch and handle artefacts.

- Visitors should be left with a greater understanding of the culture and way of life experienced by the original villagers.

1.3 THE TECHNOLOGY

Because of the perceived need for an experiential approach to interpreting the site it was decided to investigate the potential of QuickTime Virtual Reality, a recent addition to the multimedia developers tool kit from Apple. QTVR is a means of creating virtual environments from photographic source. A typical virtual reality environment is comprised of three elements:

- An object is an interactive item that the user can pick up, turn and view from all angles. Objects can be embedded in panoramas.

- A panorama is a 360 degree image that allows the user to pan and zoom within the confines of the space. A single panoramic scene is termed a node. Hot spots can be defined within each node to provide hypertext links that index associated information or provide access to other nodes within a scene

- A scene is a collection of panoramas linked together by hypertext hot spots. In a multi node scene a user can navigate between nodes to move throughout the scene.

This technology was well suited to the project for a number of reasons.

- The high level of interactivity that is expected by today's more technologically literate audience (the SEGA generation) is well catered for as is the requirement for an experiential virtual reality interface.

- The quasi organic nature of ruined structures is not conducive to traditional 3-D modelling CAD tools, which function best in a more rectilinear environment, so the photographic approach combines a high level of realism with a relatively quick and easy mode of construction

- The hardware requirements are at the low end of the technological scale, far bellow those required by VR in it's conventional guise.

- The use of a VR interface simplifies many of the navigational problems encountered in mainstream interactive media. In the traditional format interactive media is structured as a series of hyperlinked pages each containing interactive elements. With a VR interface the context is always apparent, thus reducing the structural depth of the media while interpretative media area always presented at the top level.

- As with many sites, recovered artefacts tend to have redistributed to national museums, virtual media enables the reinstatement of items into their original context.

1.4 CONCLUSIONS

Although only implemented to a "proof of concept level" this approach promises to deliver much of what was envisaged for relatively few drawbacks. The technical challenge of generating the virtual representations proves to be of several orders of magnitude less than the more conventional VR representations and the assemble of the supporting media is no more arduous than in the case of normal multimedia presentations. An inherent advantages is that navigation is through a natural format in that the user progresses through the information by traversing the site. Additionally since all descriptive or discursive material is found along with the item to which it refers it is always presented in it's natural context.

ABACUS
Department of Architecture and Building Science,
Strathclyde University,
Glasgow

VRML POSSIBILITIES:

THE EVOLUTION OF THE GLASGOW MODEL

G. Ennis; M. Lindsay; M. Grant, VSMM '99 Dundee, Scotland

ABSTRACT

During the 1980's, ABACUS, a research unit at the University of Strathclyde developed an interest in the ability to model and manipulate large geometrical databases of urban topography. Initially, this interest lay solely in the ability to source, capture and store the relevant data. However, once constructed, these models proved genuinely useful to a wide range of users and there was soon a demand for more functionality relating to the manipulation not just of the graphics, but also the range of urban attributes. Although a number of improvements were implemented there were drawbacks to the wide adoption of the software produced.
The problems were almost all due to deficiencies in the then current hardware and software system available to the professions, and although this strand of research continued to be pursued, most of the development had to be focused on research applications and deployment.

However, the recent advent of the Virtual Reality Modelling Language (VRML) standards have rekindled interest in this field since this language enables many of the issues that have proved problematic in the past to be addressed and solved. The potential now exists to provide wide access to large scale urban models. This paper focuses on the application of VRML as applied to the 'Glasgow Model'.

1.0 THE ABACUS GLASGOW MODEL

In the 30 years since ABACUS was formed, the group has explored many research directions, the most enduring of which has been centred on the visualisation of the built environment. ABACUS has designed and built many software applications for this purpose, the first towards the end of the 1970's was called VIEWER.

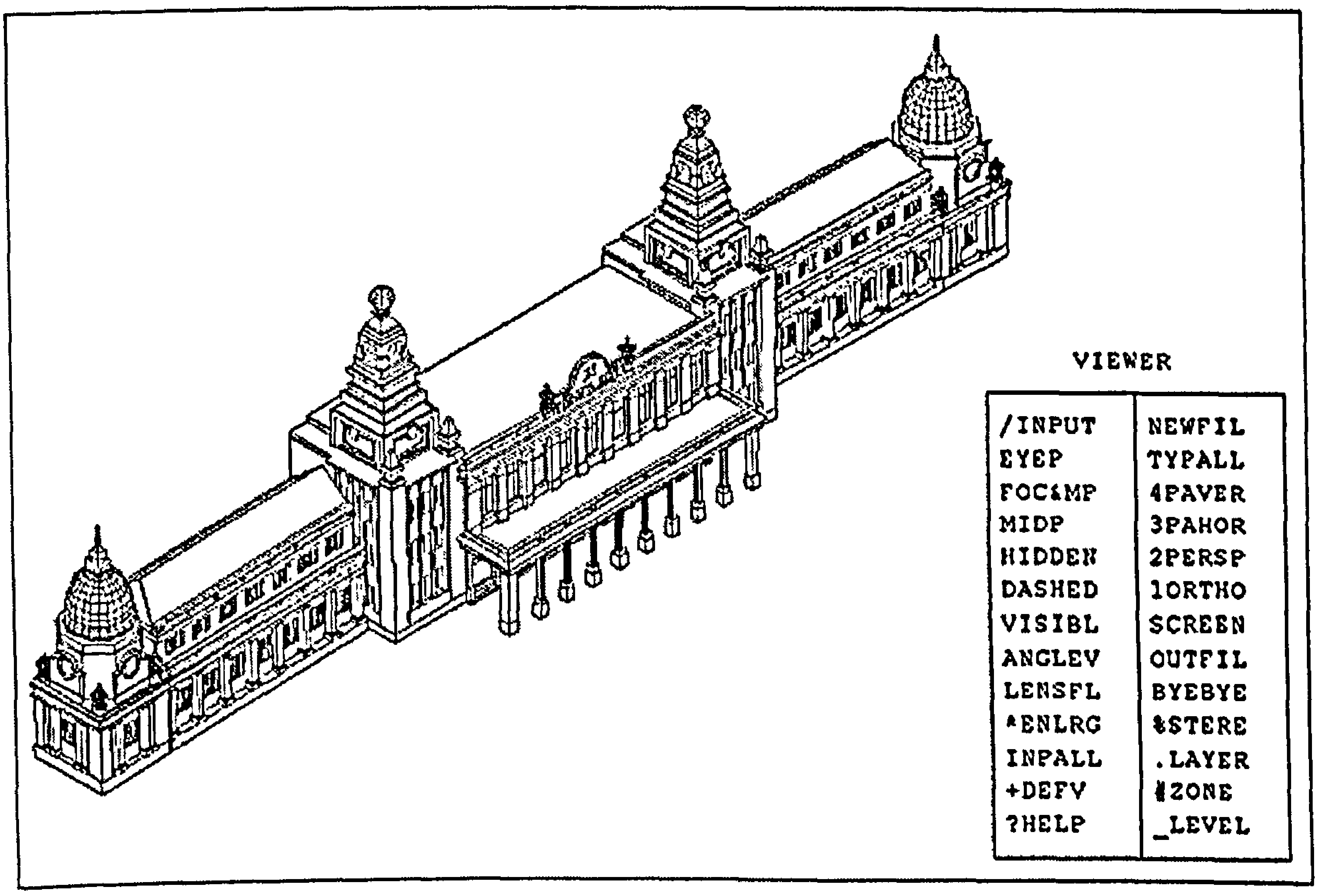


Figure 1 Abacus Software - Viewer

Using wire-frame line images, VIEWER allowed the visual appraisal of an object (e.g. a building) or a group of objects (e.g. an urban scene). Since the program could produce orthogonal, parallel or perspective projections of a scene, from any chosen viewpoint, the resulting images could be combined with photographs to form photomontages of proposed developments in situ. The VIEWER software was closely followed by VISTA, software designed to take the wire-frame model to a more advanced state, overlaying its framework of surfaces with textures to create a believable full colour shaded 'real-world' object.

In the spring of 1986, the Rutherford and Appleton Laboratory approached ABACUS, inviting the group to evaluate the capability of a new generation colour graphics workstation namely the Iris, produced by Silicon Graphics. The Iris was revolutionary in that it had a dedicated graphics engine, specific circuitry devoted to the task of undertaking large geometric transformations at speed, giving the illusion of real-time animation.

In order to stretch the capabilities of this new technology to its limits, ABACUS embarked on an ambitious method of evaluation. Using their own software, the team built a virtual model of the city of Glasgow representing an area of some 25 square kilometres, and attempted to produce interactive real-time 'fly-throughs' of the massive urban geometry data-set.

Backed with funding from 'Glasgow Action', a team of students were employed over the summer period to help with the mammoth task of capturing the necessary data.
Head of ABACUS Professor Tom Maver explains:
"The strategy we decided to adopt was to build a model in three levels as it were... First we had to capture the terrain of the city. Secondly we digitised the road network and floated that down onto the terrain. Thirdly of course, there was the challenge of capturing the geometry of the buildings."(1)

1.1 RECOMMENDATIONS

Building the Glasgow model soon became an academic exercise in how to create, store, access and manipulate large geometric databases. Its appeal was instant, proving popular with architects and developers, who could use the model to display their design proposals within an accurate city context. Widespread use of the data-set was however limited to those organisations (mostly research) who owned similar expensive hardware and who were also prepared to struggle with the less than user-friendly software. Despite being frequently deployed by ABACUS as a design and planning aid, the model failed to directly benefit the wider community.

In April 1987, Steven King, a research student at the University of Strathclyde, conducted a critical analysis of the Glasgow Model(2). The following is a summary of his main recommendations:

The data set is too large to enable realistic animation, use only what is needed in each scene.

Different levels of detail need to be introduced i.e. the nearest block should have the highest detail and those blocks in the background should have very little detail.
Labelling - the data set requires labelling with street names, building names, owners, uses etc. being retrievable interactively.

More powerful hardware is needed.
The user-interface needs to be developed to provide more intuitive navigation.
Software is needed to allow the user's eye-point to remain at a constant height above ground level to aid the sense of 'walking' in the model, yet allowing a 'manual over-ride' for investigative manoeuvring.

Integrated databases containing building information should be linked directly to the model, i.e. it should be possible to stop outside a building and then call up information on that building including a detailed picture. This could be displayed in a separate segment of the VDU.

Without these features, King states, "the full potential of the Glasgow Model will not be realised. When these facilities become available, all manner of information will be available for reference and display, e.g. the ability to display where particular businesses are situated, could be useful in finding a bank either for a member of the public or for a rival bank examining optimum possible locations for a new branch."

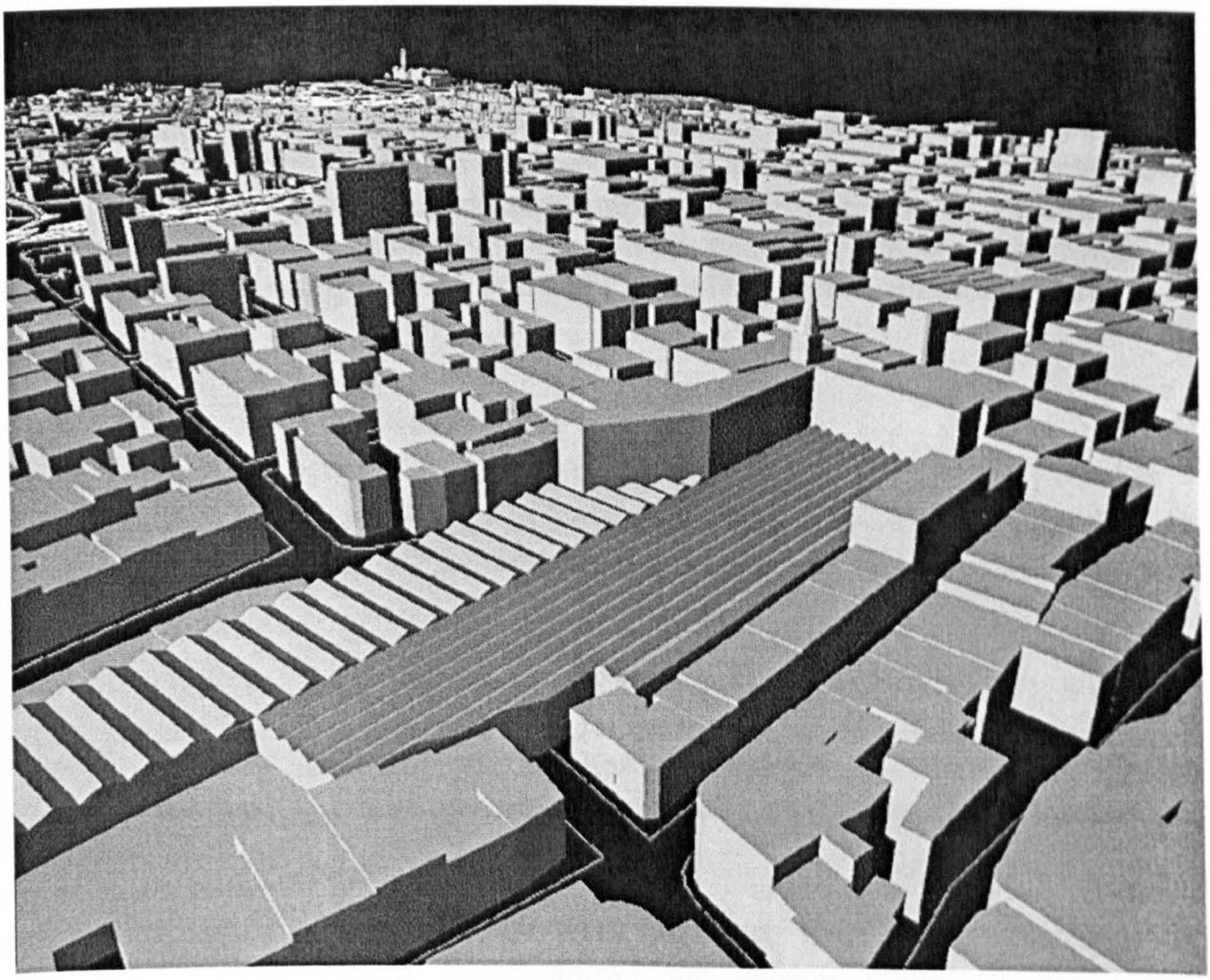


Figure 2 The Glasgow Model in 1987

At the time, ABACUS decided that the implementation of some or all of these features, although feasible, was a futile task, in terms of cost, software development, and time-scale.

Today however, it seems uncanny, that a young architectural research student should predict certain 'key' modelling standards that would emerge and be deployed almost ten years later with the advent of VRML.

1.2 VIRTUAL REALITY MODELLING LANGUAGE

In comparison to other programming languages VRML is in its infancy, yet its growth rate has been exceptional. VRML is not just a 3D display language, and it is unlike all other conventional VR systems:

"The fact that VRML was designed to run within a Internet browser, such as Netscape, offers some significant advantages over more conventional VR software. The browser is available at minimal cost, the software is platform independent and models can be transferred across the Internet."(3) The need for expensive software is eliminated, and "a 100 Megahertz Pentium is the minimum CPU recommended to view 3-D worlds and a 3-D accelerator board will make the rendering smoother"(4), a specification acceptable to most users.

ABACUS have chosen to use VRML in the re-development of the original Glasgow Model. The language's flexibility enables development at both ends of the viewing spectrum. From a web-based system accessible by the general public over the Internet, to a highly detailed immersive environment as viewed in the Virtual Environment Laboratory(5), a new facility recently established by ABACUS within the University of Strathclyde's Department of Architecture.

1.3 THE GLASGOW DIRECTORY

With the introduction of VRML1 in November 1995, ABACUS were for the first time able to incorporate 3D models of selected areas of Glasgow into their web pages.

As the VRML standards progressed (VRML2 followed by VRML'97), it became apparent that this language could offer far more than just the ability to view 'static' models. Extended research began into how the Glasgow Model could be adapted to suit these facets, and be developed into a fully interactive Urban Information System.

In 1998, funded by 'Glasgow 1999', work began on "The Glasgow Directory"(6), an Internet-based system which allows the general public to access city information through an interactive 3D model of central Glasgow.

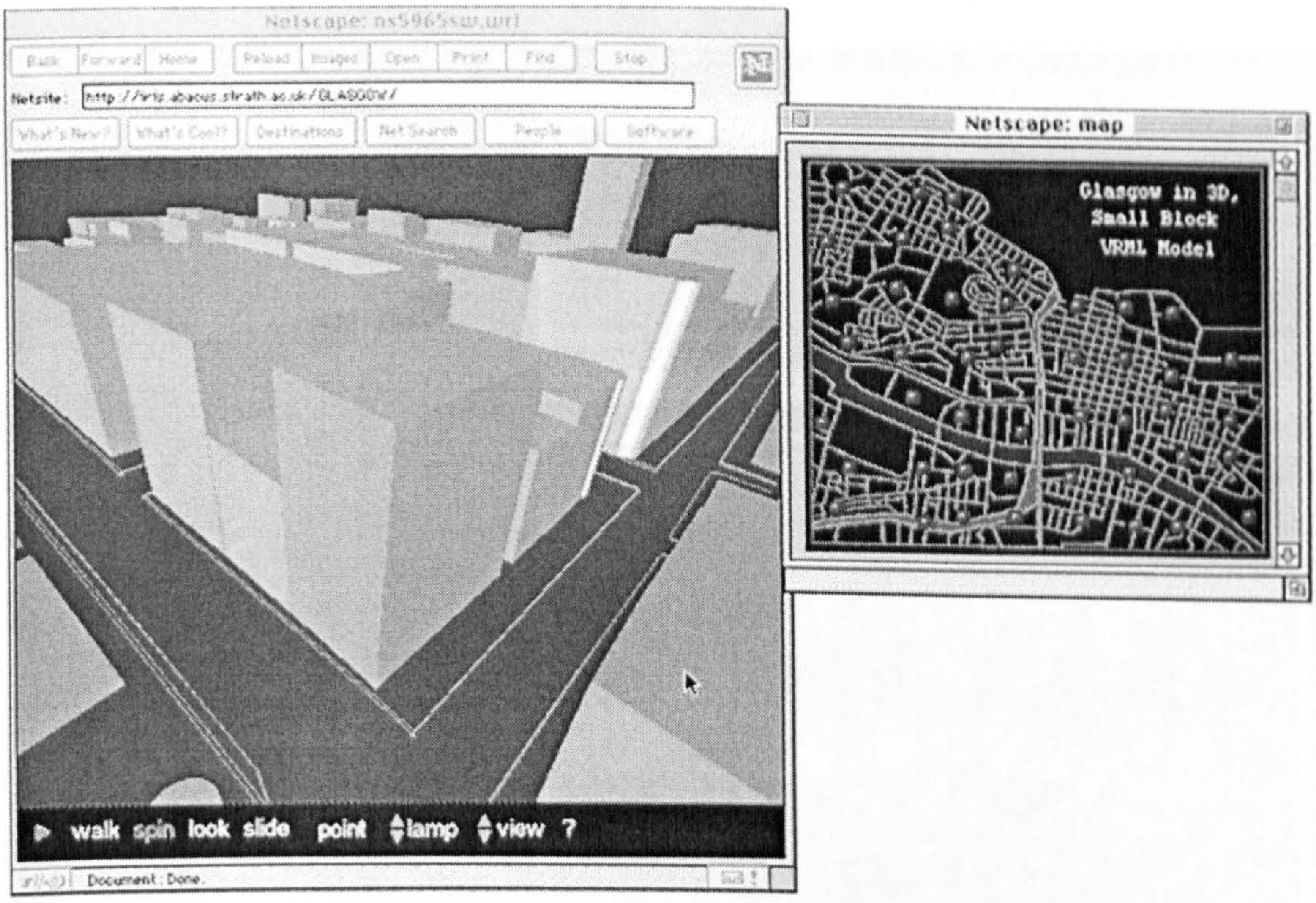


Figure 3 Glasgow VRML1 System viewed on Netscape's 'Live 3D' 1996

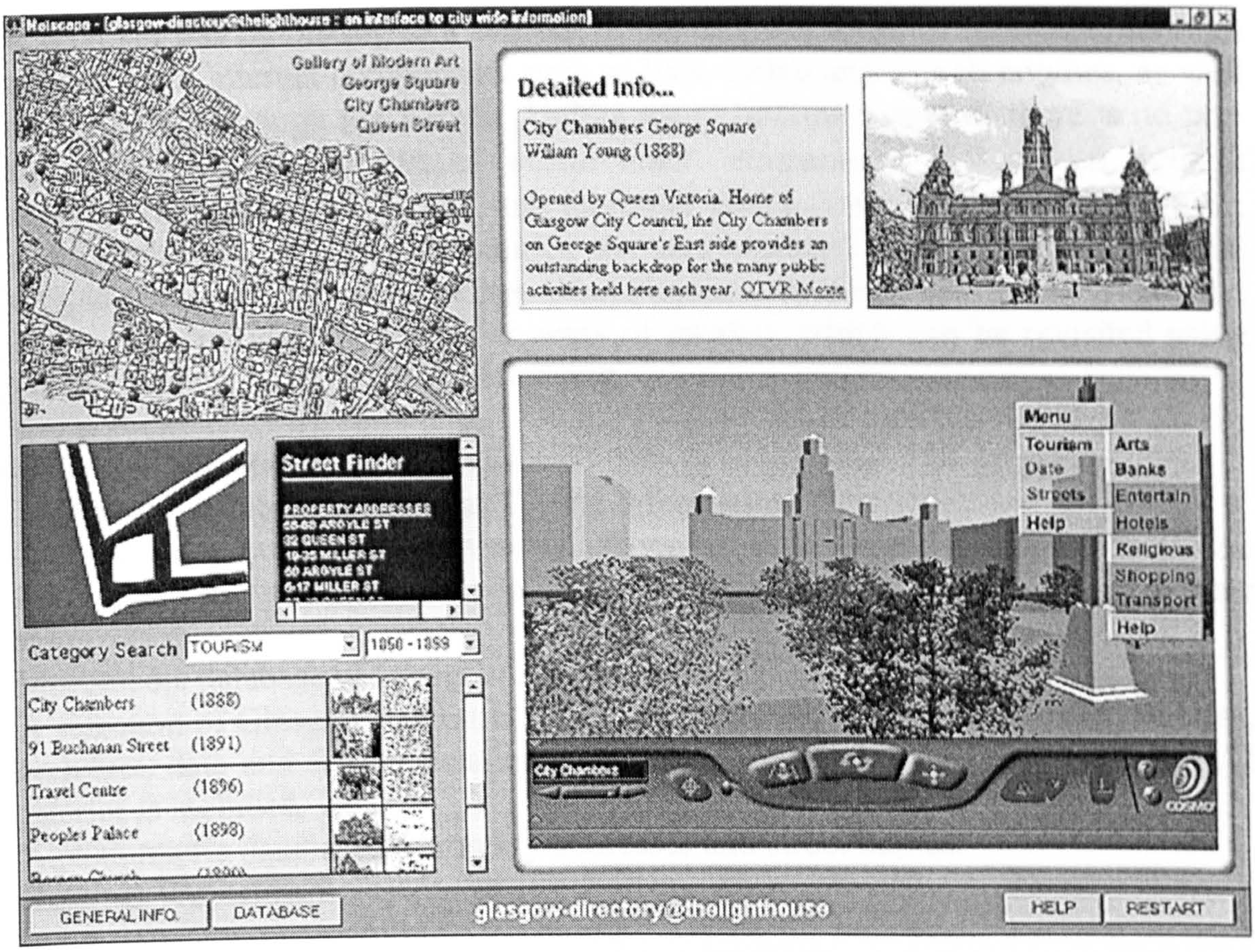
1.4 CONSTRUCTION SUMMARY

The following summarises the construction process of the Glasgow Directory and indicates each steps relevance to King's recommendations.

[Recommendation 1 and 4]
The original Glasgow Model was converted from its native file format into VRML1 using in-house software. It was then converted into VRML2 using a freely available SGI converter. Once the model was in VRML format, it was split up into 47 different model segments allowing for faster download time and manipulation.

[Recommendation 2]
It was decided to keep the level of detail on these models to a minimum, except where individual buildings were considered 'landmarks'. These would be given more detail to help navigation within the scene.

[Recommendation 5]
An intuitive interface was designed to house the VRML model, allowing easy integration of 'frames' hosting different data types. This was achieved through the use of JavaScript, added to both the HTML files and the VRML model. This enabled the automatic display of certain content in one frame as the user explored certain aspects of the VRML model.



Detailed Info...

City Chambers George Square
William Young (1888)

Opened by Queen Victoria. Home of Glasgow City Council, the City Chambers on George Square's East side provides an outstanding backdrop for the many public activities held here each year. QTVR Movie

Category Search TOURISM 1850 - 1859

City Chambers	(1888)
91 Buchanan Street	(1891)
Travel Centre	(1896)
Peoples Palace	(1898)

Figure 4 The Glasgow Directory Interface

[Recommendation 6]
A number of defined 'Viewpoints' were programmed into the VRML code, allowing quick access to places of interest within any model segment. Using the 'collision' and 'gravity' option within the VRML file, the user may walk around the environment while remaining at eye level height above the terrain.

[Recommendation 3 and 7]
An Interactive Menu Bar was programmed into the VRML code allowing a range of functionality. This included the ability to activate street labelling, as well as identify a range of building types within the city by hyper-linking to an external database of property information.

1.4.1 Content

The Glasgow Directory allows users to explore the 'virtual city', search under a range of headings for items of interest, and discover some of Glasgow's finest architecture.

It achieves this by linking to a number of information sources, accessible through conventional Internet components such as lists, tables and search engines, as well as in-directly through the VRML. Unlike many Internet systems, there is no pre-set route to follow, or list of 'useful links'. Information is accessed through intuitive exploration of the site, and therefore varies depending on both the user, and the chosen route. In this manner, the user becomes familiar with the virtual city, in much the same way as they would become familiar with the physical city. He or she may identify particular areas of interest, which can be revisited using familiar routes, or accessed via browsing.

Current information sources include:
Multimedia database of General Tourist Information.
This includes entertainment venues, shopping centres and transport facilities. Each item in the database contains a photograph, an address, and further information.
Multimedia database of Glasgow Architecture.
A selection of Glasgow's finest buildings are archived with a photograph, address, architect, date and further detailed information. In many cases a QuickTime VR interior is included.
Alphanumeric database of property addresses and street names.
External Web-sites.

The system can take advantage of the many Glasgow based web sites that exist on the Internet such as those offering local news and weather.

A special feature within the system allows architects and designers to download appropriate sections of the 3D model for use in proposed developments. Future additions to the system will permit design proposals to be integrated into the system allowing the audience to explore a number of possible urban design alternatives.

1.4.2 Recognition

Recognition of the virtual city is aided by the use of certain key 'Lynchian Elements'(7). As with the physical location, landmarks, streets, edges and nodes play an important part in comprehending the virtual environment.

Extra detail is given to buildings in the model when they are perceived to be landmarks.

The streets can be easily identified by clicking on them.

Each model segment is terminated at a distinct edge such as the river or a main street.
Specified viewpoints (nodes) at junctions etc., help users understand their location in the virtual city.

It achieves this by linking to a number of information sources, accessible through conventional Internet components such as lists, tables and search engines, as well as in-directly through the VRML. Unlike many Internet systems, there is no pre-set route to follow, or list of 'useful links'. Information is accessed through intuitive exploration of the site, and therefore varies depending on both the user, and the chosen route. In this manner, the user becomes familiar with the virtual city, in much the same way as they would become familiar with the physical city. He or she may identify particular areas of interest, which can be revisited using familiar routes, or accessed via browsing.

Current information sources include:
Multimedia database of General Tourist Information.
This includes entertainment venues, shopping centres and transport facilities. Each item in the database contains a photograph, an address, and further information.
Multimedia database of Glasgow Architecture.
A selection of Glasgow's finest buildings are archived with a photograph, address, architect, date and further detailed information. In many cases a QuickTime VR interior is included.
Alphanumeric database of property addresses and street names.
External Web-sites.

The system can take advantage of the many Glasgow based web sites that exist on the Internet such as those offering local news and weather.

A special feature within the system allows architects and designers to download appropriate sections of the 3D model for use in proposed developments. Future additions to the system will permit design proposals to be integrated into the system allowing the audience to explore a number of possible urban design alternatives.

1.4.2 Recognition

Recognition of the virtual city is aided by the use of certain key 'Lynchian Elements'(7). As with the physical location, landmarks, streets, edges and nodes play an important part in comprehending the virtual environment.

Extra detail is given to buildings in the model when they are perceived to be landmarks.

The streets can be easily identified by clicking on them.

Each model segment is terminated at a distinct edge such as the river or a main street.
Specified viewpoints (nodes) at junctions etc., help users understand their location in the virtual city.

The subtle use of basic colours - blue for river, green for parks etc., enhance the urban fabric.

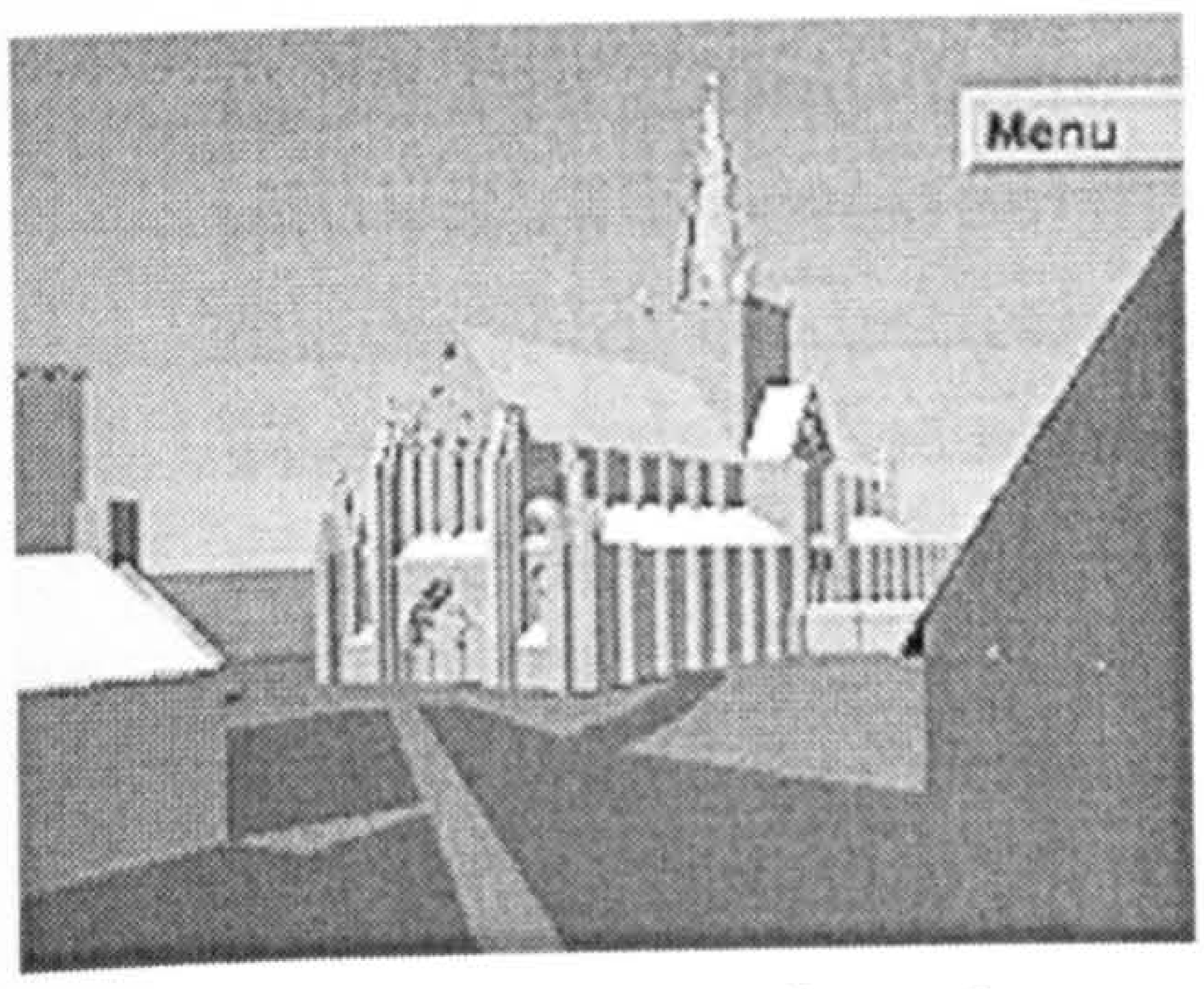


Fig.5 Detailed Landmarks

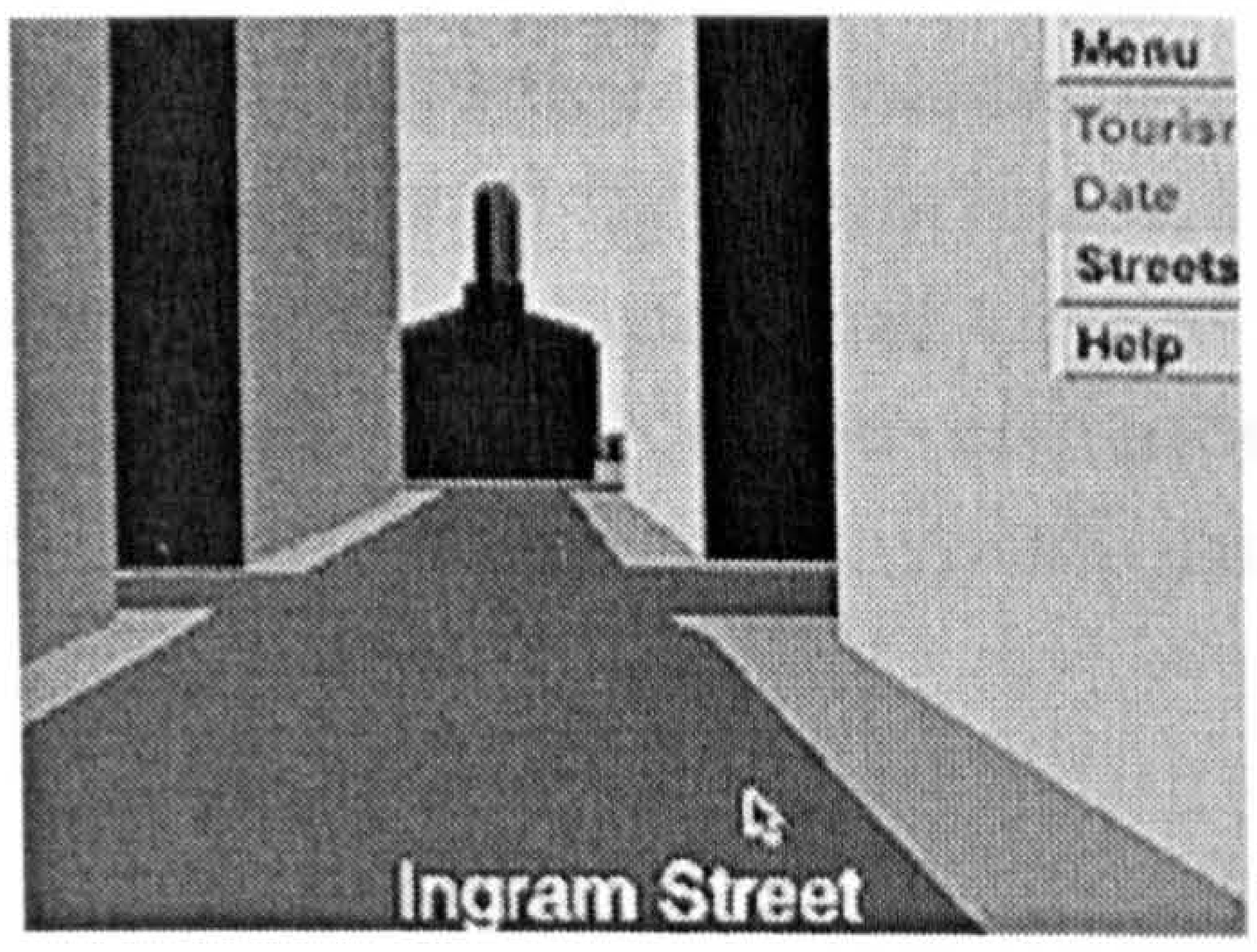


Fig.6 Street Names

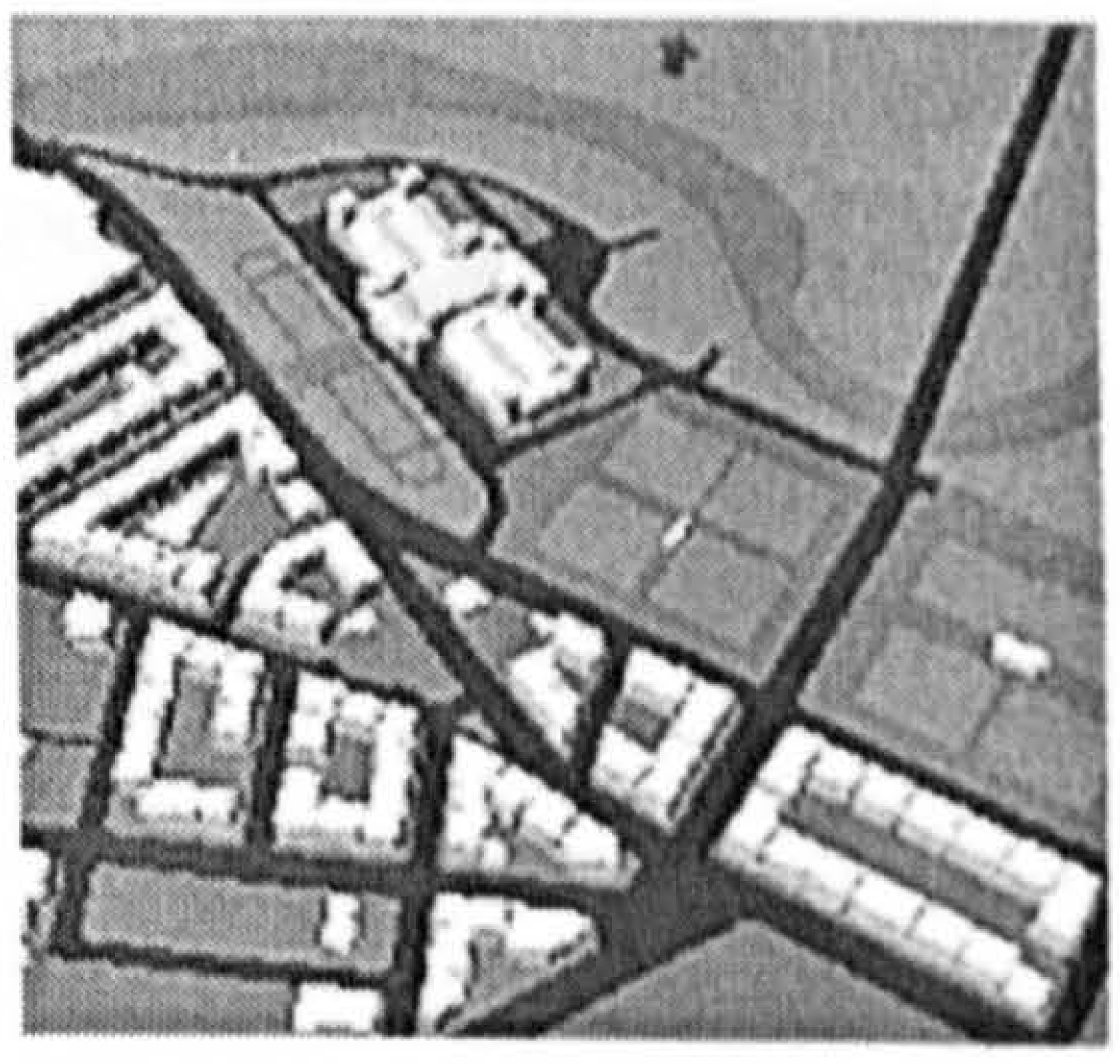


Fig.7 Use of colour

Recognition of individual buildings is achieved through the use of 'layered' data. Since the model of the building itself may be nothing more than a basic block within the city context, the facade is expressed as a photograph displayed in another frame. This frame also displays some introductory text about the building, as well as linking to more detailed information such as QTVR interiors. In this way, the user can recognise the building by collating all the relevant information that is automatically displayed on screen as the building is accessed within the VRML model.

1.4.3 Uses

There are many potential applications for the Glasgow Directory, covering a range of interest and user requirement. Its flexible user interface means that the information gained from each access session will be unique to each user. Here are three examples of system application:

1.4.4 Tourism:

A prospective visitor to the city of Glasgow can explore the virtual city before arrival in the physical, resulting in a greater knowledge of where to go and what to see. Browsing the virtual city can influence choice of hotel, entertainment or shopping area, and transport networks may be examined for choice of optimum route. Places of interest worth a 'real' visit can be identified and details printed out for use during that visit.

Subsequently the virtual city can be revisited and attractions that were missed in the physical city can be enjoyed at leisure through the VRML city model and imaging technology.

1.4.5 Education:

The Glasgow Directory can be used on various levels in an educational context. From primary school pupils, who may be given access and tasks such as, 'explore their city', 'find their school' or 'count and locate the museums', through to university or college students who could by exploring the model be introduced to the concept of computer modelling application and virtual reality.
The models educational potential is of course enhanced because it is easily accessible via the World Wide Web, using platform independent Internet Browsers.

1.4.6 Architects:

Using the Glasgow Directory, architects may access and download parts of the city model which relate to design projects. These sections may be downloaded in a variety of file formats and used to enhance the presentation of new building or urban designs. By displaying design proposals within the context of the city, the architect and their clients, can understand and therefore assess a scheme more clearly.

Once the design is complete, the architect can submit it for integration into the virtual city model, providing public access to the new proposal, as well as contributing to the evolution of the system.

1.4.7 City Planning:

Proposed change to the urban fabric can first be evaluated by simulating the change within the model. Any number of proposals can be analysed in parallel, with the public being given the opportunity to vote for their favourite, a process similar to that used in the Netherlands at Ljburg(8).

This feature would be particularly beneficial during the initial stages of a project, where numerous members of the design team must collaborate on design issues before advancing to the construction phase. This method of increased access to design alternatives would reduce the need for group viewing of physical models.

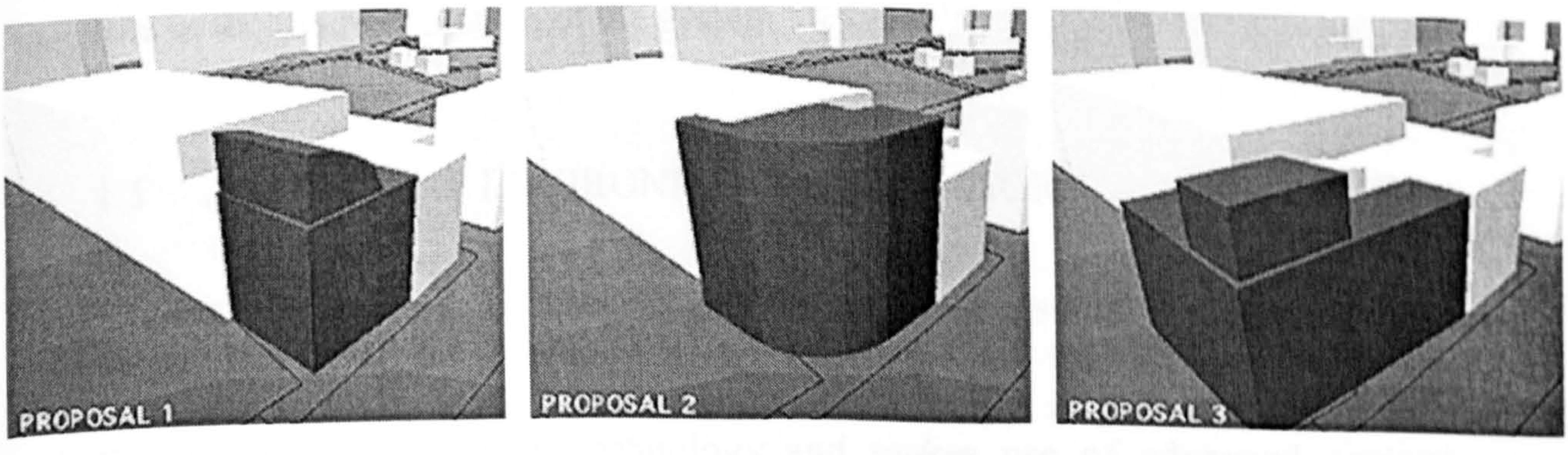


Figure 8 Planning Proposals

1.4.8 Maintenance and Updates

In the initial development stage of the Glasgow Directory, the addition or upgrade of building information and urban elements was a laborious task. Consequently, a partially automated update module was developed, which automated the editing or creation of 6 separate files of various content and output the data in the required four different programming languages.

This method has been recently improved by the development of a fully automated update module, requiring only that the user input certain key details such as building name, architect and date of construction. Model viewpoints can also be automatically created using this module, and user does not need any knowledge of the programming languages deployed.

This module is written entirely in PERL, enabling additions to be made by remote contributors through a standard Internet browser.

The ABACUS "Viewpoint Creator"(9), which is now on-line for public use, is based on the elements used in this automation system, and allows the creation of VRML viewpoint nodes from any VRML model. A task that usually relies on trial and error. User feedback from this and other "ABACUS VRML Tools '98" has been positive.

1.4.9 User Feedback

Throughout its design stages, ABACUS has welcomed evaluation from the general public as to the development of the Glasgow Directory. This has been achieved through posting 'Requests for Comments' (RFC's) on many appropriate Internet Newsgroups(10), asking users to 'test-drive' the project and return their ideas and recommendations by way of a simple on-line 'Feedback' form. To date, over 100 forms have been returned, with users offering from the simple "congratulations on the project", to a complete list of suggested improvements and samples of programming code.

Collecting feedback however is simply not enough. Suggestions must be analysed, evaluated, and where appropriate put into practice. This process will play a critical role in the future evolution of the system.

1.5 THE VIRTUAL ENVIRONMENT LABORATORY

The most recent stage of research and development associated with the Glasgow model has been the construction by ABACUS of the Virtual Environment Laboratory (VEL). This visually immersive facility is powered by advanced Silicon Graphic computer technology and makes use of advanced projection systems. The laboratory seats around 15 people and the 160 degree curved screen fills the viewers' cone of vision, thus giving them the impression of being immersed within the projected environment. The processing power of the Onxy2 Silicon Graphics machine allows interactive manipulation of large computer models, thus enabling real-time navigation of detailed building interiors or urban landscapes with relative ease.

This facility enables two aspects of virtual reality to be addressed, namely interaction and immersion. Development in these areas in the past has been hampered because of a lack of computer processing power. However the graphics processor at the heart of the VEL can interactively display the Glasgow City model which not only represents around 10,000 properties, but also the appropriate topographical data covering an area of 25 square kilometre. Currently the machine displays this model at up to 25 frames a second, compared with one frame every few minutes 10 years ago, and a rate of processing equivalent to a frame ever day or two 20 years ago.

Visual immersion is achieved as a result not only of the screen size, but also due to its curvature. It therefore feeds the viewers peripheral vision with the correct visual clues and thus invokes the sensation of being surrounded by the projected environment, city scene or architecture. So although it is a 2D moving image that is being projected onto the curved screen, viewers report the feeling of being immersed within a 3D environment.

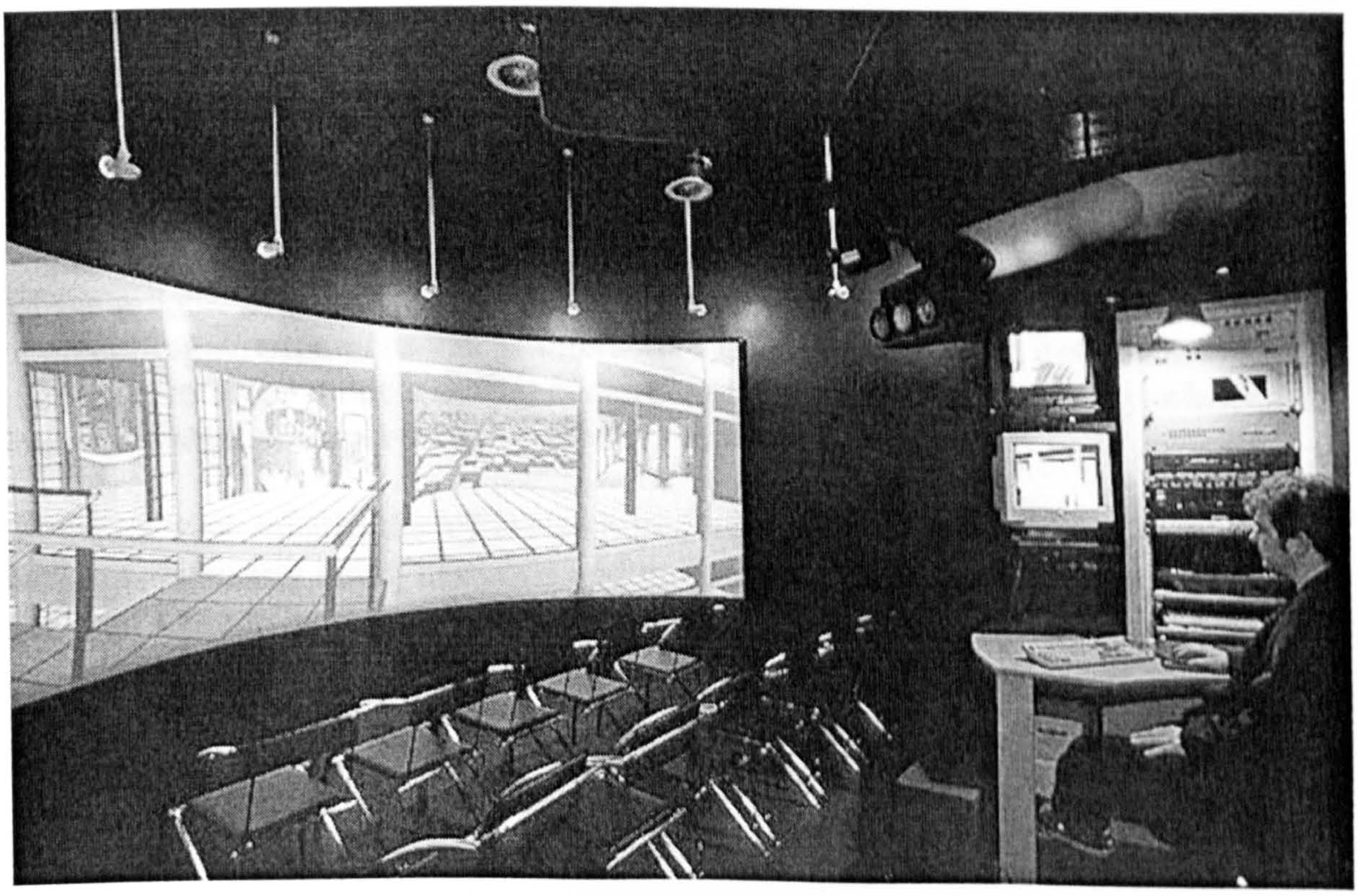


Figure 9 The Virtual Environment Laboratory

1.6 APPLICATIONS

1.6.1 Architecture

The VEL is currently used by architects and designers to test designs prior to the construction phase of a project. Architects can 'experience' their designs in a new way and clients can 'walk though' a building at an relatively early stage of the design process in order to gain an insight into the attributes of the environment that they are commissioning. One of the pleasing aspects of experiencing the industrial use of the facility is that virtually no explanation needs to be given to the design teams during these sessions regarding how to use the VEL, since they are immediately able to see their models clearly on the screen and to navigate them with ease.

For example, one French based designer has commissioned some modelling on an urban scale, in order to assess their design in situ. In this case the Glasgow model was revisited, and two streets identified for detailed modelling. This was achieved by the addition of modelled street furniture, and the inclusion of facade detail by texture mapping. The client's design was converted from ArchiCAD to VRML using 3D Studio Max, and integrated into the scene. During sessions in the VEL, the client's design could be seen either on its own or in context, and design changes were suggested as a result. These changes were subsequently built into a new version of the model, and approved during the next design appraisal session.

The VEL development team are currently modelling the new parliament buildings and surrounding area for viewing in the facility. Again this Scottish Office project is on an urban scale and will be constructed using a new version of the ABACUS modelling package, specially designed for VR projects. This package makes use of the graphic processing capabilities of the Silicon Graphics machines to provide a modelling environment that is interactive in nature, enabling ease of movement through prototype models during construction. It is hoped that this 'fly-while-you-build' software will speed up the development of large-scale VR environments.

In the next six months the VEL also hopes to test the wheelchair motion platform, which is currently being designed and built by the Department of Bioengineering (at the University of Strathclyde). This platform will be able to interpret wheelchair movement and thus control motion through virtual environments. Using this system, architects will be able to test wheelchair access to buildings and urban spaces prior to their construction or redevelopment.

1.6.2 Other Applications

The VEL currently hosts several different types of marine related models, which address areas such as, vessel design, ship evacuation and accident investigation. The ship evacuation model uses the facilities ability to control the inclusion of 'fog' or 'smoke' in the environment, where as the accident investigation work depends on clear visualisation of specially constructed engineering models.

Other areas of interest or research include:
Civil Engineering
Prosthetics & Orthotics
Rapid Prototyping
Tele-presence
Biology
Psychology

Future research issues include:
optimisation of VR modelling and data transfer
impact of visual immersion on the design process
role of stereo visualisation in immersive VR
role of 3D sound in immersive VR.

1.7 CONCLUSIONS

The Glasgow Directory is due to be officially launched during June 1999. The combination of its functionality linked to a massive 3D city scale urban data set has only been possible due to the emergence and versatility of VRML, and the parallel improvements in computer processing power and communications infrastructure of the Internet.

Where as this type of data was once confined to the realm of state-of-the-art research laboratories, it can now be viewed remotely from a 'standard' personal computer in the home. Already the prototype version of the directory has 322 registered users, providing a rich source of feedback for the development team.

To date applications of the system include tourism, education, architecture, and planning, with the area of virtual shopping and e-commerce currently under investigation. Other future additions may include the use of avatars, so that visitors to the virtual city may communicate with each other during their visit.

VRML has proved to be a rapidly evolving programming language, which is easy to implement since it does not require compilation, and is platform independent. Its open architecture and Internet compatibility will encourage and sustain its use, and its versatility will see further implementations such as those under development in the VEL.

REFERENCES

From the video "The Glasgow Experience : Building a computer based model of a city.", April 1987
King, S., "The Glasgow Model : An Appraisal" Computer Aided Architectural Design, Special Study Project, April 1987
Day, A., and Radford, A., "An overview of City Simulation", CAADRIA '98, Proceedings of the Third Conference on Computer Aided Architectural Design Research in Asia, p.183, CAADRIA, 1998
Intervista Whitepaper, "VRML Myths", http://www.intervista.com/vrml/whitepapers/index.html
Information regarding the VEL can be found on-line at http://www.strath.ac.uk/VEL
The Glasgow Directory can be found at http://iris.abacus.strath.ac.uk/glasgow/
Lynch, K., "The image of the city", p48, The MIT Press, 1990
Bourdakis, V., "Virtual Reality: a Communication Tool for Urban Planning", CAAD Towards new design conventions, p52, Technical University of Bialystok, 1997
The ABACUS Viewpoint Creator can be found at http://iris.abacus.strath.ac.uk/glasgow/vp.htm
Newsgroups to which ABACUS submitted feedback requests to include alt.planning.urban alt.architecture comp.lang.vrml and alt.3d.

ABACUS (1,3,5)
Department of Architecture Strathclyde University

BioEngineering Unit (2,4,6)
Wolfson Centre, 106 Rottenrow
Strathclyde University
Glasgow
Scotland

DEVELOPMENT OF A WHEELCHAIR VIRTUAL REALITY PLATFORM FOR USE IN EVALUATING WHEELCHAIR ACCESS

C.S.Harrison1, P.M.Dall2, P.M.Grant3, M.H.Granat4, T.W.Maver5, B.A.Conway6

1 http://www.strath.ac.uk/Departments/Architecture/VEL/intro.htm
2 http://www.strath.ac.uk/Departments/BioEng/

ABSTRACT

In the UK the Disability Discrimination Act 1995 aims to end discrimination against disabled people. Importantly the Act gives the disabled community new employment and access rights. Central to these rights will be an obligation for employers and organisations to provide premises which do not disadvantage disabled people. Many disabled people rely on wheelchairs for mobility. However, many buildings do not provide conditions suited to wheelchair users. This project aims to provide instrumentation allowing wheelchair navigation within virtual buildings. The provision of such instrumentation assists architects in identifying the needs of wheelchair users at the design stage. Central to this project is the need to provide a platform which can accommodate a range of wheelchair types, that will map intended wheelchair motion into a virtual world and that has the capacity to provide feedback to the user reflecting changes in floor surface characteristics and slope. The project represents a collaborative effort between architects, bioengineers and user groups and will be comprised of stages related to platform design, construction, interfacing, testing and user evaluation.

KEYWORDS: Virtual Wheelchair Access, Disability, Discrimination, Motion Simulation, Building, Instrumentation

1.0 INTRODUCTION

With the increasing age profile of developed nations due to improved health care reducing mortality rates (Tuljapurkar et al.,2000) there will be a continual increase in the number of people becoming reliant on wheelchairs for mobility as a consequence of disease or injury. In most societies wheelchair users are often discriminated against in areas such as employment and education due in a large part to inadequate access provision within the built environment. This can lead to a below average standard of living and the need for a lifetime of state support. The development of a virtual reality facility which allows wheelchair users to explore virtual representations of the built environment should, if appropriately utilised by user groups and architects, lead to socio-economic benefits to a large group of disabled people through improved building design and therefore equal opportunities. In addition, the use of virtual reality systems where navigation is integrated with the sensing of intended motion of wheelchairs (or other forms of assistive devices) will be of benefit to those groups who need to examine compliance with the requirements of equal opportunity legislation. This could lead to the establishment and widespread adoption of design standards relating to building design for wheelchair access.

The long term objectives of this project are to provide a virtual reality facility that can be used to generate, via an interaction between architects, designers and wheel chair users, guidelines which address the issue of wheelchair access to, and within, the built environment. The project aims to design and build a wheelchair motion platform through which wheelchair users can explore virtual representations of buildings. It is envisaged that such a facility would form a powerful and cost effective means of evaluating wheelchair access provision early in the design of new buildings and in the redevelopment of existing buildings (Forest and Gombas 1995). Accordingly the following preliminary objectives need to be met:

- The design and construction of a manual wheelchair motion platform that can accurately monitor intended wheelchair motion and can provide physical and optical feedback to the wheelchair user on the presence of virtual obstacles or changes in floor coverings or slope.

- Interface the platform with a Silicon Graphics virtual reality facility to provide an immersive virtual environment within which navigation is linked to the intended wheelchair motion.

- Generate virtual representations of a range of building types in order to test and calibrate the performance of the platform and perform an evaluation of the system by wheelchair users.

1.1 WHEELCHAIR DRIVEN VR SYSTEM

The facility is comprised of seven functionally separate elements: projection system; image generator; graphics software; motion simulator; roller system; control system; user. An overview of the system is shown in Figure 1.

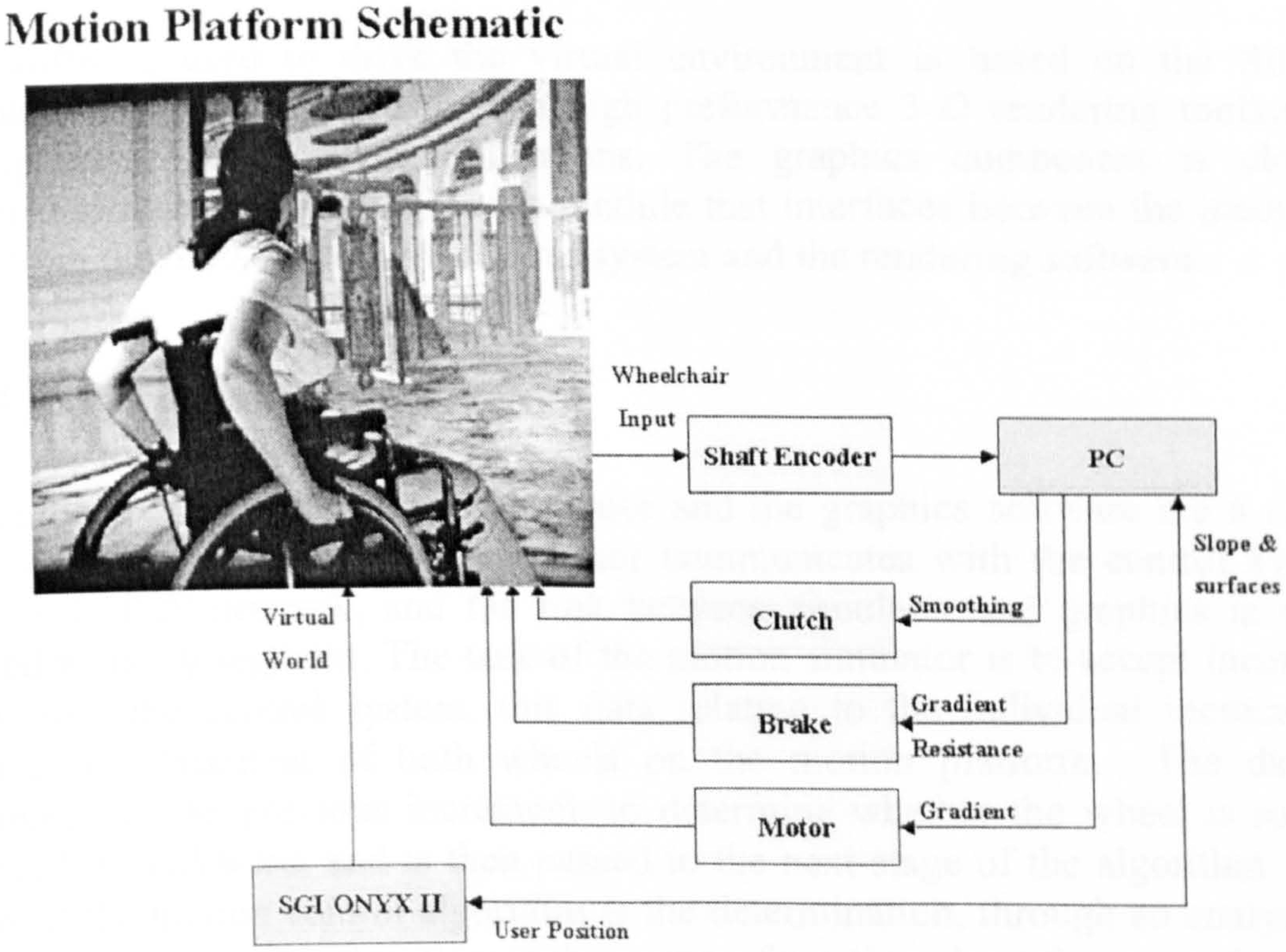


Figure 1. Layout of the wheelchair motion platform and Virtual Environment Laboratory (VEL)

1.1.1 Projection system

The virtual environment is visualised using a three-projector system that provides a 150° by 40°, high-resolution image on a five metre diameter cylindrical screen. Each of the three image channels is edge-blended to provide a seamless display. When viewed from the design eye point the image fills most of the users field of vision providing a highly convincing sense of immersion within the scene.

1.1.2 Image Generator.

Graphics are generated on a twelve-processor Silicon Graphics ONXY II with two graphics pipes. This is capable of processing detailed architectural models at high frame rates in order to provide the desired degree of realism. At each time-step in the simulation the graphics are rendered to three separate output channels, each channel sharing the same eye point but with a different angular offset in azimuth, corresponding to the offsets in the projection system. This circumvents the geometrical distortion inherent in large field of view displays.

1.1.3 Graphics Software

The software used to drive the virtual environment is based on the Silicon Graphics Performer API. This is a high performance 3-D rendering toolkit for multiprocessed interactive applications. The graphics component is closely coupled to a separate asynchronous module that interfaces between the incoming data from the motion platform control system and the rendering software.

1.1.4 Motion Simulator

As outlined above, the motion simulator and the graphics software are a close-coupled system. The motion simulator communicates with the control system over a TCP/IP network, and the link between simulator and graphics is via a shared memory segment. The task of the motion simulator is to accept incoming data from the control system, this data relating to the individual incremental angular displacement of both wheels on the motion platform. The data is compared to the previous increment, to determine whether the wheel is rotated forward or backward, and is then passed to the next stage of the algorithm. The basis of the motion control algorithm is the determination, through an analysis of similar triangles, of the location of the centre of rotation along the rear axle of the virtual wheelchair and the angle through which is turned. From this the transformation of the eye point and rotation of the view vector can be determined.

The graphics application requires the Cartesian co-ordinates of the eye point, plus the yaw, pitch and roll angles of the direction of view. Given the yaw angle the remaining two parameters can be calculated based on the wheelchair's attitude on the floor plane. In the database traversal three rays corresponding to the contact patch of each of the rear wheels and the midpoint of the front axle, are intersected with the floor. The normal vector of the ground plane at these points can then be used to calculate the roll and pitch of the chair and the corresponding view. The same intersection procedure can also be used to identify the surface under each wheel, this information can then be used to index material properties, such as rolling resistance, which are passed back to the control system.

1.1.5 Roller System

The roller system is housed within a framework that supports the wheelchair and occupant, and converts wheel motion into an instrumented rotation of the main shaft. The system is duplicated for each wheel of the wheelchair. Mounted on each shaft are the brake, clutch, encoder, inertial mass and the take off for the motor drive. A detailed discussion of the design rationale and function is given in section 3.

1.1.6 Control System

The control system is based on a standard PC with purpose written software that interfaces with the image generator via a network link using TCP/IP and with the instrumentation via a General Purpose Interface Board (GPIB). The control system feeds the motion engine of the image generator with incremental readings from the rotary encoders on the motion platform whilst controlling the feedback stimuli to the wheelchair on the basis of data received from the simulation in order to effect changes in floor conditions or potential collisions.

1.1.7 Wheelchair User

Each of the above elements forms a linked system that is controlled by the bidirectional flow of information from the wheelchair to the virtual environment, and from the virtual environment to the wheelchair. The feedback loop is by the users visual perception of progress through the virtual environment and by the perceived proprioceptive changes associated with alterations in the rolling resistance of the wheelchair. By closing the feedback loop with a human rather than a further pair of sensor connections, it is expected that any minimal latency or hysteresis in the rest of the communications path will be compensated for by the user.

1.2 DESIGN OF THE MOTION PLATFORM

1.2.1 Manual Wheelchair Interface

The design of the interface between the manual wheelchair and the virtual environment has been specified so that the wheelchair remains fixed in place, with movable driving wheels. In this way the user is not limited within the virtual environment by the constraints of the physical environment. This interface was required to fulfil two main functions. Firstly, it had to be able to transfer the rotation of the driving wheels to provide realistic navigation around the virtual world (Hofstad and Patterson 1994 for modelling characteristics). Secondly, it was to provide additional non-visual (proprioceptive) feedback to the user of the environment represented by the virtual world, in order to match optic flow and visual perception with voluntary motor effort, and thereby enhance the users' experience of navigating the virtual world. Finally, in order that a user could retain their own wheelchair whilst navigating in the virtual environment, the interface has been designed to accommodate a wide range of manual wheelchairs.

Figure 2 shows the detailed layout of the principle components of the wheelchair interface.

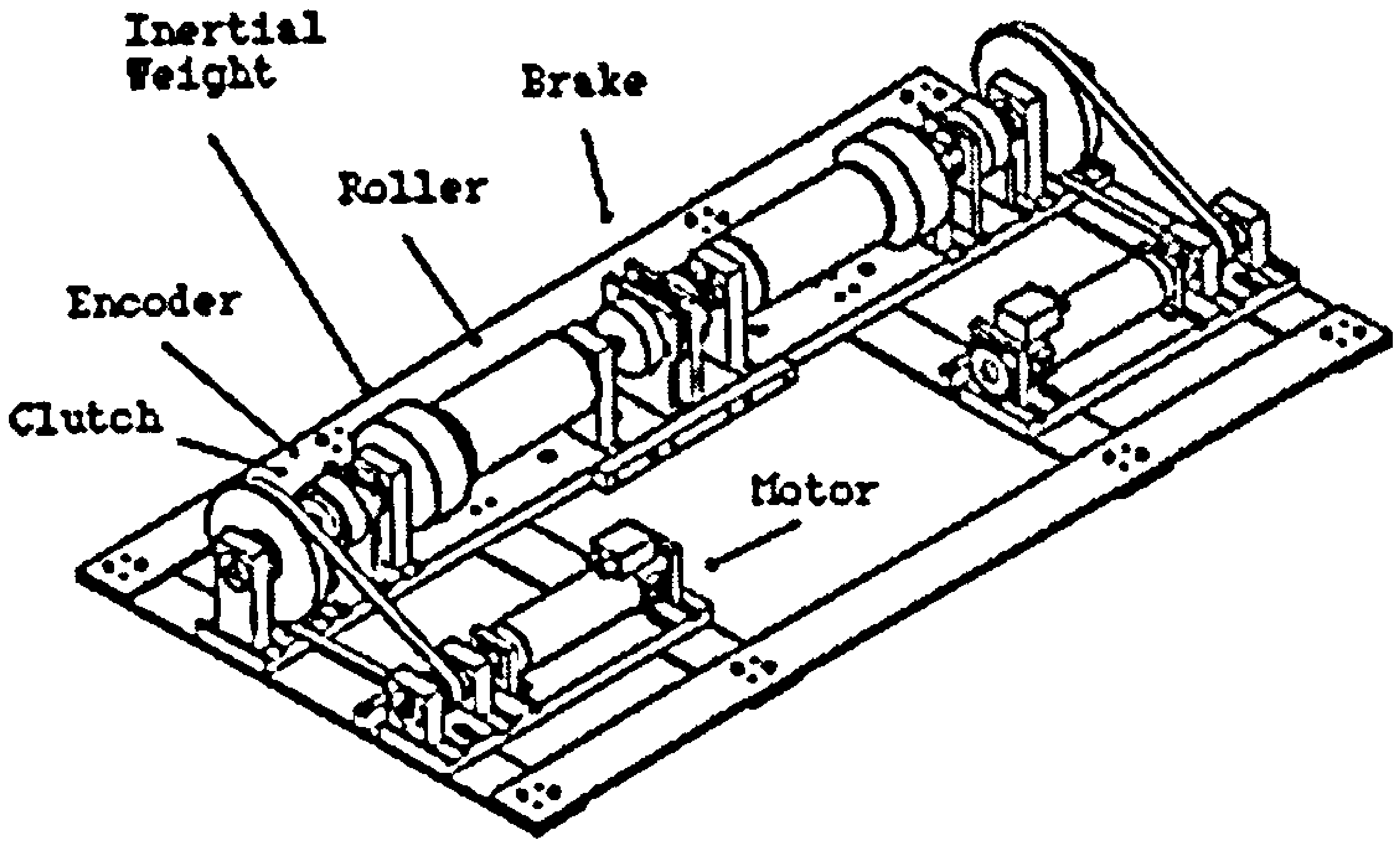


Figure 2. Physical Components of the Wheelchair Interface

The physical structure of the wheelchair interface is based around a pair of rollers mounted so that one roller is in tangential contact with each of the driving wheels. Frictional contact between the tyre and the roller is sufficient to ensure that the roller rotates simultaneously with each wheel and so could then be used to navigate within the virtual environment. The use of two rollers was required so that differential motion of the driving wheels of the wheelchair could be distinguished to detect turning.

A large number of environmental features were identified for which accessory physical feedback could enhance the visual feedback of the virtual environment. These included slopes and cambers, kerbs, uneven surfaces and different ground surfaces. For the initial design of the wheelchair interface it was decided to concentrate on providing feedback for different ground surfaces and different grades of slope, as these are the features most commonly encountered by the wheelchair user. Different floor surfaces are simulated by altering the resistance to motion of the rollers, when pushing the wheels. A variable torque hysteresis brake is used to provide resistance to motion at each roller. The brake is used to provide increased resistance and so simulate the effect of gravity on the wheelchair when moving up a slope. Simulation of the wheelchair moving down a slope requires active input into the system, providing a torque against which the user can control their movement down the gradient. A variable torque motor is used to provide this input for different grades of slope. Switching between use of motor and brake is accomplished by monitoring the position of the rollers.

An early design proposal suggested that the wheelchair platform should tilt in response to changes in gradient within the computer model. This may have been advantageous because users often vary their seating position to alter their centre of gravity when traversing slopes. However due to the increased complexity of such an implementation, and the level of visual feedback already available, this feature was discounted.

1.3 PHYSICAL STRUCTURE OF THE WHEELCHAIR INTERFACE

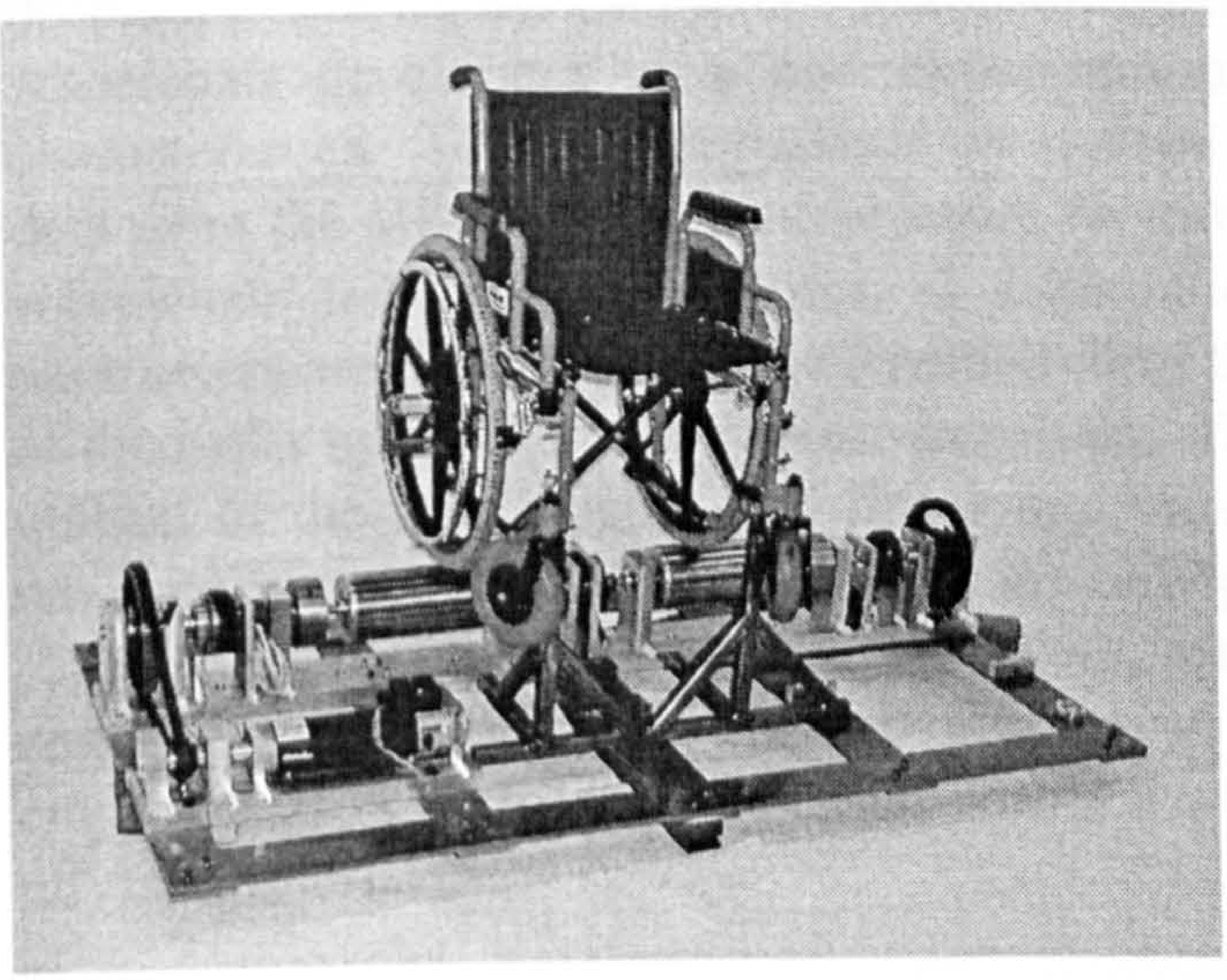


Figure 3. Photograph of Wheelchair Motion Platform

The physical structure of the wheelchair platform is based around a pair of rollers. These are mounted on separate shafts so that one roller is under each driving wheel of the wheelchair. The rollers are 300 mm long so that a range of wheelchair widths can be supported. Each roller is constructed from seamless steel and aluminium discs were inserted into the ends of each tube to close the roller, and to provide a bush to which the axle could be secured.

The roller shaft is supported by a pair of single row radial ball bearings mounted in support pillars, fixed to a solid base plate, as illustrated in figure 3. The roller and space for an inertial mass is between the two bearings. The maximum size of the mass that could be accommodated was a cylinder 65mm long with a diameter of 240mm. Outside the lateral ball bearing, the shaft was machined to accommodate a hollow shaft encoder. The body of the encoder is held with respect to the base plate, while the hollow shaft has been clamped to the roller shaft. Each brake is rigidly mounted coaxial to the roller shaft. The motor is geared to the roller shaft using a toothed belt and is coupled by an electromagnetic clutch.

The entire structure is enclosed by a wooden cover so that the user is protected from the moving parts. Two rectangular holes in the cover allow the rollers to stand slightly proud of the surrounding surface allowing wheel contact. Adjustable straps and bars ensure that the wheelchair is held in place on the rollers, and a ramp allows the user to gain access to the facility. The system is electrically isolated.

1.3.1 Design Issues for the Wheelchair Interface

Normal translation in a manual wheelchair is a discontinuous motion. Between propulsive pushes the user needs to reposition the arms and during this period the wheelchair decelerates at a rate dependant on environmental and wheelchair characteristics. To simulate realistic navigation in the virtual world using encoder data from the rollers as the control input, requires that roller motion simulates real wheelchair motion with respect to kinematic parameters. These kinematic parameters are dependant on the interaction of the inertia and resistance to motion of the roller system. In the real environment, wheelchair motion consists of rotary motion of the four wheels, and linear motion of the wheelchair and user. When navigating in the virtual environment, the only physical motion is provided by the rotary motion of the rear wheels, the rollers, inertial masses, the brakes and, when engaged, the motors. To provide basic navigation through the virtual environment, a high inertia coupled with a low rolling resistance is needed in the physical interface.

The inertia of the roller system is predominantly determined by the fixed inertia of the contributing components. Provision has been made for the inclusion of an inertial mass for each roller. This allows the inertia of the system to be increased, should the basic inertia of the roller system prove too low, and provides a measure of adjustment to the system.

The biggest design issue has been the reduction of resistance of motion of the system to an acceptable level. The acceptable level was defined as that required for the system to decelerate with similar characteristics to the wheelchair and user when on a smooth level surface such as linoleum. A belt connection considerably increased the overall resistance to motion of the system and was avoided where possible, and therefore the brake was mounted directly coaxial to the roller shaft. The motor needed to be geared with respect to the roller and thus a belt connection was unavoidable. However counteracts the increase in system resistance by increasing the torque provided. An early prototype used needle bearings to mount the roller shaft. The characteristics of the needle bearings resulted in a failure of the roller to be able to match the required deceleration rates, and ball bearings were used in subsequent designs.

Additionally, a single roller per wheel was used rather than two. While a pair of parallel rollers per wheel would offer simple control over wheel positioning, a single roller system gives less rolling resistance. Therefore the platform has had to incorporate a set of adjustable bars and straps which fix the wheelchair in the correct position over the roller systems.

1.3.2 Matching the real world to the Virtual Environment

Defining the kinematics of wheelchair motion with respect to user input, surface and slope conditions, and different wheelchairs can be an involved task with many variables (Kauzlarich and Thacker 1985; Frank and Abel 1989; Hofstad and Patterson 1994). The creation of algorithms to encompass such variability was outwith the scope of this project, and an empirical approach was adopted towards the determination of control settings for various surface considerations. This involved the measurement of the motion of a single wheelchair when there was no user input, for real motion over various surface conditions and when on the wheelchair interface.

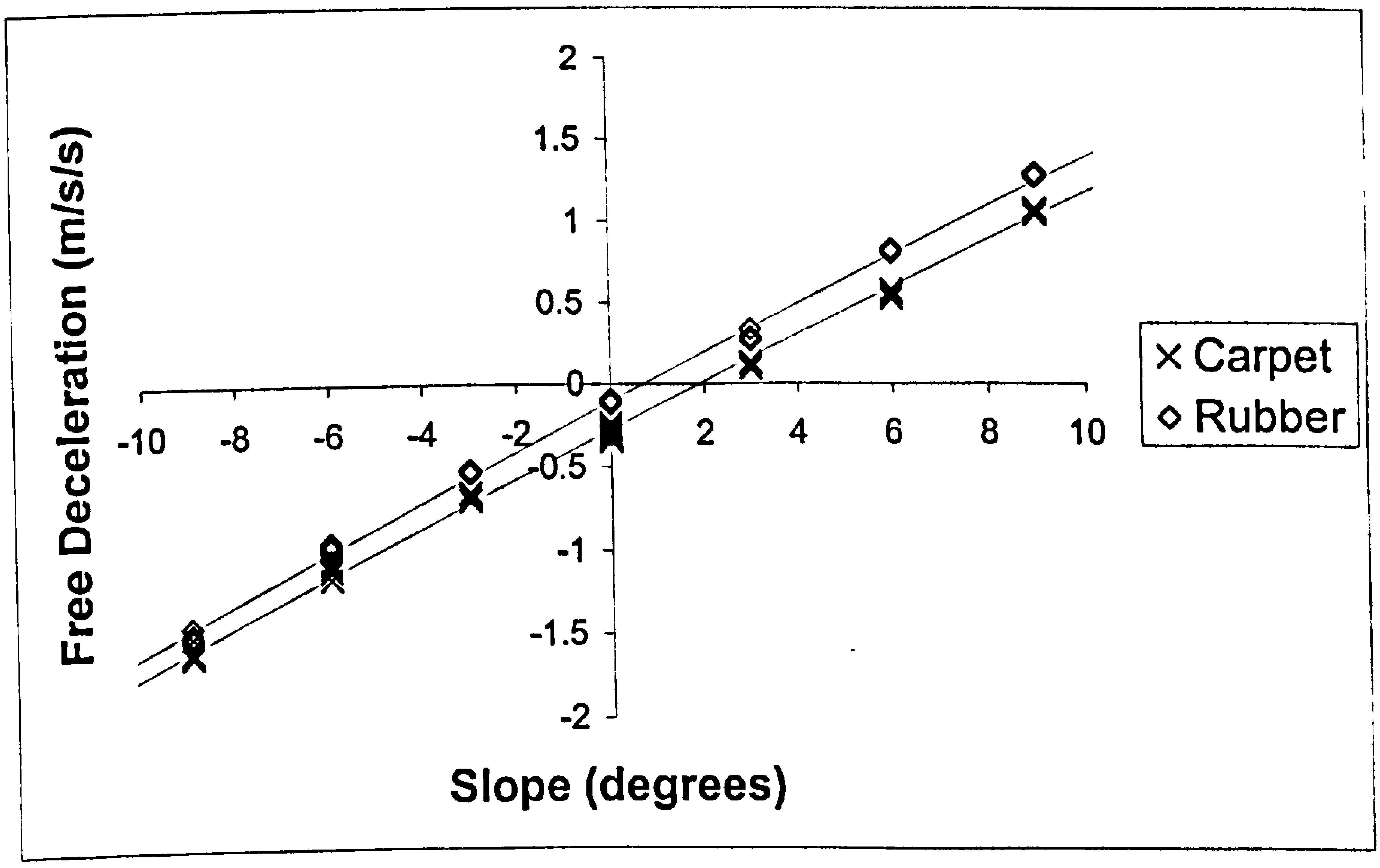


Figure 4. Measured Free Acceleration for Carpet and Ribbed Rubber on Various Grades of Slope

The condition of no user involvement was used so that it could be reliably reproduced in all experimental circumstances. A rotary encoder was attached to the axle and spokes of one wheel of a basic wheelchair to investigate this relationship. The wheelchair was rolled with no external input over a range of surface coverings, for example ribbed rubber and carpet, and several different grades of slope. The acceleration of the period when there was no external input was calculated. Data has been collected for slope and surface combinations and a best-fit line for free acceleration against slope for each surface was calculated (figure 4). The graph shows a linear relationship between the free deceleration of the wheelchair and the grade of slope. This basic relationship is subject to an offset due to negative acceleration equal to the free deceleration of the wheelchair over each floor surface. A database of the free deceleration of a wheelchair in the real environment has been obtained from which the required free deceleration of the roller system for any particular surface and slope combination can be extrapolated. By knowing the free deceleration of the wheelchair and the roller system, the virtual environment can then be simulated by a range of voltage inputs for the brake and motor.

The kinematics of deceleration also vary depending on wheelchair characteristics, and these must be taken into account when setting up a session with the wheelchair interface. A calibration procedure is required when a user is introduced to the interface for the first time. The diameter of the wheel also needs to be input into the control so that the ratio between roller and wheel for different wheelchairs is respected.

1.4 SYSTEM INTEGRATION

1.4.1 Integrating the Mechanical motion Platform and Virtual Environment Laboratory

It was decided at a relatively early stage to use a personal computer as the host for the motion platform and electromechanical devices, as this is a computer that is relatively straightforward to implement using commercially available software and peripherals. The PC could also provide a multi-role capability.
The team had earlier experience using sockets in a Unix to PC environment and this was one of the reasons for using this proven route again. For this reason a PC based solution was chosen, although many other integration routes are possible (Stredney et al 1995), the modular approach means that in the future interfacing to other VEL computers should be straightforward.

1.4.2 Host Personal Computer

Based on a networked Dell 220 Precision, the PC has a Digital to Analogue (D/A) card driving the clutches, brakes and motors through a purpose built power supply unit, and a commercial motor interface unit. The optical encoders provide feedback to the PC on the position of the rollers in space, and hence the position and orientation of the wheelchair, which is then be sent to the Silicon Graphics machine for conversion into 3D world co-ordinates. Standard geometrical treatments are available which describe the wheelchair dynamics (Stredney et al 1995).

1.4.3 Communications

The system uses sockets which provide either UDP (User Datagram Protocol) or TCP (Transmission Control Protocol) in order to establish communication between the PC and the VR Platform, currently a Silicon Graphics machine.

The use of this standard technology means that if the VR host is upgraded, the wheelchair motion platform will still be able to be commissioned with little or no difficulty.

1.5 CONCLUSIONS

The mechanical construction of the wheelchair motion platform is complete, although a certain amount of user calibration is required. The design and manufacture of the platform has proved to be a more complicated task than first estimated due to the requirement to provide low rolling resistance and minimal deflection when loaded by the user. This has meant working to high levels of machining accuracy, and assembly alignment for the eight separate bearings. Since in this application the user must be able to steer, the roller assembly is duplicated for each side resulting in a high quality assembly platform which can accommodate a wide range of wheelchairs, and which has good mechanical performance.

The design has enabled the independent use of brake, clutch, and motors which means that a wide variety of surface and slope conditions can be simulated in as natural a way as possible. These include uphill and downhill slopes combined with smooth and rough surfaces. The design can be evaluated by the user community, which includes architects, building design engineers and healthcare professionals as well as wheelchair users. Uniquely this system gives a sense of feeling of what a proposed design will be like to use in practice, which cannot be duplicated by any other method. Whilst it is possible to evaluate designs using conventional methods, none conveys the actual physical experience of wheelchair use in proposed buildings, which provides solid education and feedback outcomes for the user community.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the financial support of the Extending Quality of Life (EQUAL) EPSRC research program (GR/ M05416). We would also like to acknowledge the help and advice of Spinal Injuries Scotland (SIS) and Disability (Scotland).

REFERENCES

G. Forrest, G. Gombas, (1995), Wheelchair accessible Housing: Its role in cost containment in Spinal Cord Injury, Archives of Physical Medicine and Rehabilitation, 76, 5, pp450-452.

T. Frank, E. Abel, (1989), Measurement of the turning, rolling and obstacle resistance of wheelchair castor wheels, Journal of Biomedical Engineering, 11, 462-466.

M. Hofstad, P. Patterson, (1994), Modelling the propulsion characteristics of a standard wheelchair, Journal of Rehabilitation Research Developments, 31, 2, pp129 – 137.

J. Kauzlarich, J Thacker, (1985), Wheelchair tire rolling resistance and fatigue, Journal of Rehabilitation Research Developments, 22, 3, pp25 – 41.

D. Stredney, W. Carlson, J. Edward Swan II and Beth Blonstein, (1995), The determination of Environmental Accessibility and ADA Compliance Through Virtual Wheelchair Simulation, Presence: Teleoperators and Virtual Environments, 4, 3, pp297-305

S. Tuljapurkar, N. Li and C. Boe, (2000), A universal pattern of mortality decline in the G& countries, Nature, 405, pp789-792.

ABACUS
Department of Architecture and Building Science,
Strathclyde University,
Glasgow

HAPTIC INTERFACES FOR WHEELCHAIR NAVIGATION IN THE BUILT ENVIRONMENT

M.Grant, C.Harrison, B.Conway

ABSTRACT

In the UK the Disability Discrimination Act 1995 aims to end discrimination against disabled people. Importantly the Act gives the disabled community new employment and access rights. Central to these rights will be an obligation for employers and organisations to provide premises, which do not disadvantage disabled people. Many disabled people rely on wheelchairs for mobility. However, many buildings do not provide conditions suited to wheelchair users. This project aims to provide instrumentation allowing wheelchair navigation within virtual buildings. The provision of such instrumentation assists architects in identifying the needs of wheelchair users at the design stage. Central to this project is the need to provide a platform which can accommodate a range of wheelchair types, that will map intended wheelchair motion into a virtual world and that has the capacity to provide feedback to the user reflecting changes in floor surface characteristics and slope. The project represents a collaborative effort between architects, bioengineers and user groups and will be comprised of stages related to platform design, construction, interfacing, testing and user evaluation.

KEYWORDS:
Virtual Wheelchair Access, Disability, Discrimination, Motion Simulation, Building,
Instrumentation.

1.0 INTRODUCTION

Over the last 30 years there has been an increasing awareness of the lack of provision of rights for people with disabilities and this has led to the introduction of a number of important legal devices. For example, in the USA the Americans with Disability Act (ADA) (http://www.usdoj.gov/crt/ada/adahom1.htm/) has introduced legislation aiming to end the discrimination in education, employment, access to public and private accommodation and transport faced by the disabled community. Similarly, in the United Kingdom the Disability Discrimination Act (1995) (http://www.disability.gov.uk/dda/) has begun to phase in legislation that by 2004 will require service providers to make reasonable adjustments to the physical features of their premises in order for disabled people to overcome physical barriers to access.

Improved access to the build environment has therefore become, not only an important issue to the disabled community, but also an issue that must be addressed by professionals involved in the design, construction and management of the built environment. Many disabilities relate to problems of mobility and assistive technology in the form of wheelchairs can dramatically impact on the ability of users to engage more fully in society. In the USA estimates based on data collated in 1994 indicate that across all age groups 1.5 million people use wheelchairs (REF Russell et al., 1997) while in the UK the Prosthetic and Wheelchair Committee (1996) reported that more than 1.5% of the population require to use wheelchairs. With current demographic trends towards an older population in the developed world (REF Tuljapurkar et al., 2000) and the health problems associated with this these figures will have increased making wheelchair users a substantial and important population.

Nevertheless, many features of the external and internal build environment fail to accommodate the access needs of wheelchair users and this often leads to many wheelchair users achieving a below average standard and quality of living. Adoption of inclusive design together with the enforcement of current legislation should offer equal opportunities for disabled people and through their full integration in society lessen social security and welfare cost. Accordingly, there are important social and economic reasons for the development of tools, which could allow architects and others to explore access issues early in the design process of new buildings or in the redevelopment of existing building stock (domestic and commercial).

Virtual reality (VR) simulations of the build environment may provide the basis of such tools through accurate visualisation and user interaction. However, to capitalise on the potential for using VR as a tool to improve building accessibility it is vital that navigation through the VR world should be driven by systems that accurately reflect the mode of transport used by intended users. Accordingly, we report here on the development of an haptic interface which allows wheelchair users to navigate within VR simulations of buildings through the use of their own wheelchair and which provides the user with feedback that is related to the sense of effort needed in propelling the wheelchair over changes in floor surface and slope.

Normally, immersion in VR systems is generated by optic flow driven by a navigation device. The provision of non-visual sensory feedback in a form that alters with the physical demands associated with normal wheelchair motion augments immersion and provides an extremely powerful tool that allows wheelchair users to directly participate at an early stage in the architectural design process. This paper presents an outline of the development of the haptic wheelchair interface, its integration within a virtual reality laboratory and the results of a wheelchair user group evaluation on its use.

1.1 SYSTEM OVERVIEW

The entire system is comprised of three mutually inter-linked components;
the motion platform
the graphics system
the control system

Each of the above elements forms a linked system that is controlled by the bidirectional flow of information from the wheelchair to the virtual environment, and from the virtual environment to the wheelchair. The feedback loop is via the users visual perception of progress through the virtual environment and by the perceived proprioceptive changes associated with alterations in the rolling resistance of the wheelchair.

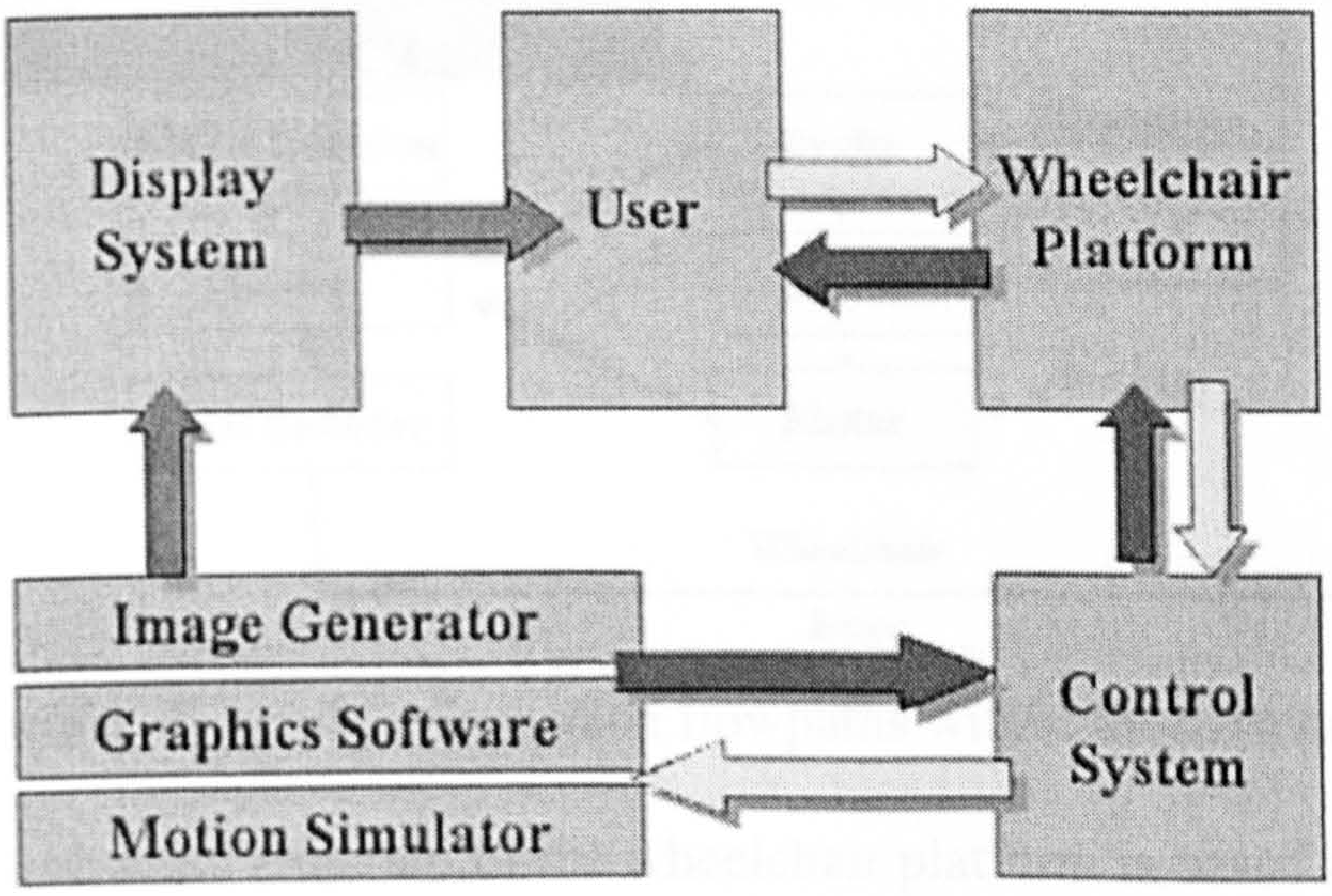


Figure 1. System overview

By closing the feedback loop with a human rather than a further pair of sensor connections, it was found that any minimal latency or hysteresis in the rest of the communications path can be compensated for by the user.

1.2 MOTION PLATFORM

The design of the interface between the manual wheelchair and the virtual environment has been specified so that the wheelchair remains fixed in place, with movable driving wheels. In this way the user is not limited within the virtual environment by the constraints of their physical environment. This interface was required to fulfil two main functions.

Firstly, it had to be able to transfer the rotation of the driving wheels to provide realistic navigation around the virtual world (REF Hofstad and Patterson, 1994, for modelling characteristics).

Secondly, it was to provide additional non-visual (proprioceptive) feedback to the user on the basis of their interaction with the virtual environment. This served to match optic flow and visual perception with voluntary motor effort, and thereby enhance the users experience of navigating the virtual world.
Finally, in order that a user could retain their own wheelchair whilst navigating in the virtual environment, the interface has been designed to accommodate a wide range of manual wheelchairs.

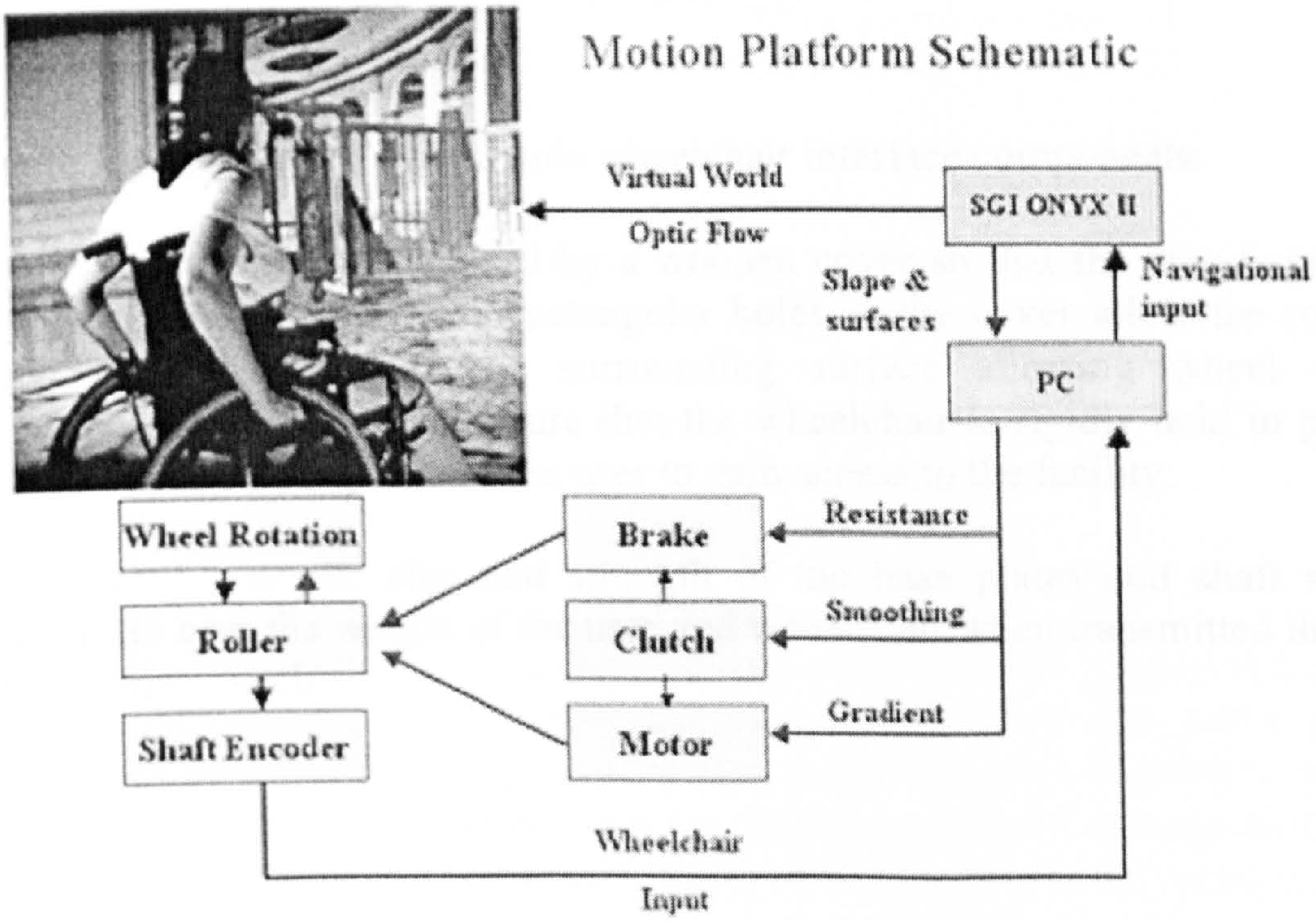


Figure 2. Principle information flowpaths within the system.

The physical structure of the wheelchair platform is based around a pair of rollers. These are mounted on separate shafts so that one roller is under each driving wheel of the wheelchair. The rollers are 300 mm long so that a range of wheelchair widths can be supported. The roller shaft is supported by a pair of single row radial ball bearings mounted in support pillars, fixed to a solid base plate. The roller, and space for an inertial mass, is situated between the two bearings. The maximum size of the mass that could be accommodated was a cylinder 65mm long with a diameter of 240mm. Outside the lateral ball bearing, the shaft was machined to accommodate a hollow shaft encoder. The body of the encoder is held with respect to the base plate, while the hollow shaft has been clamped to the roller shaft. Brakes are rigidly mounted coaxial to the roller shaft. Motors are geared to each roller shaft using a toothed belt and coupled through an electromagnetic clutch.

The disposition of these components is shown in Figure 3.

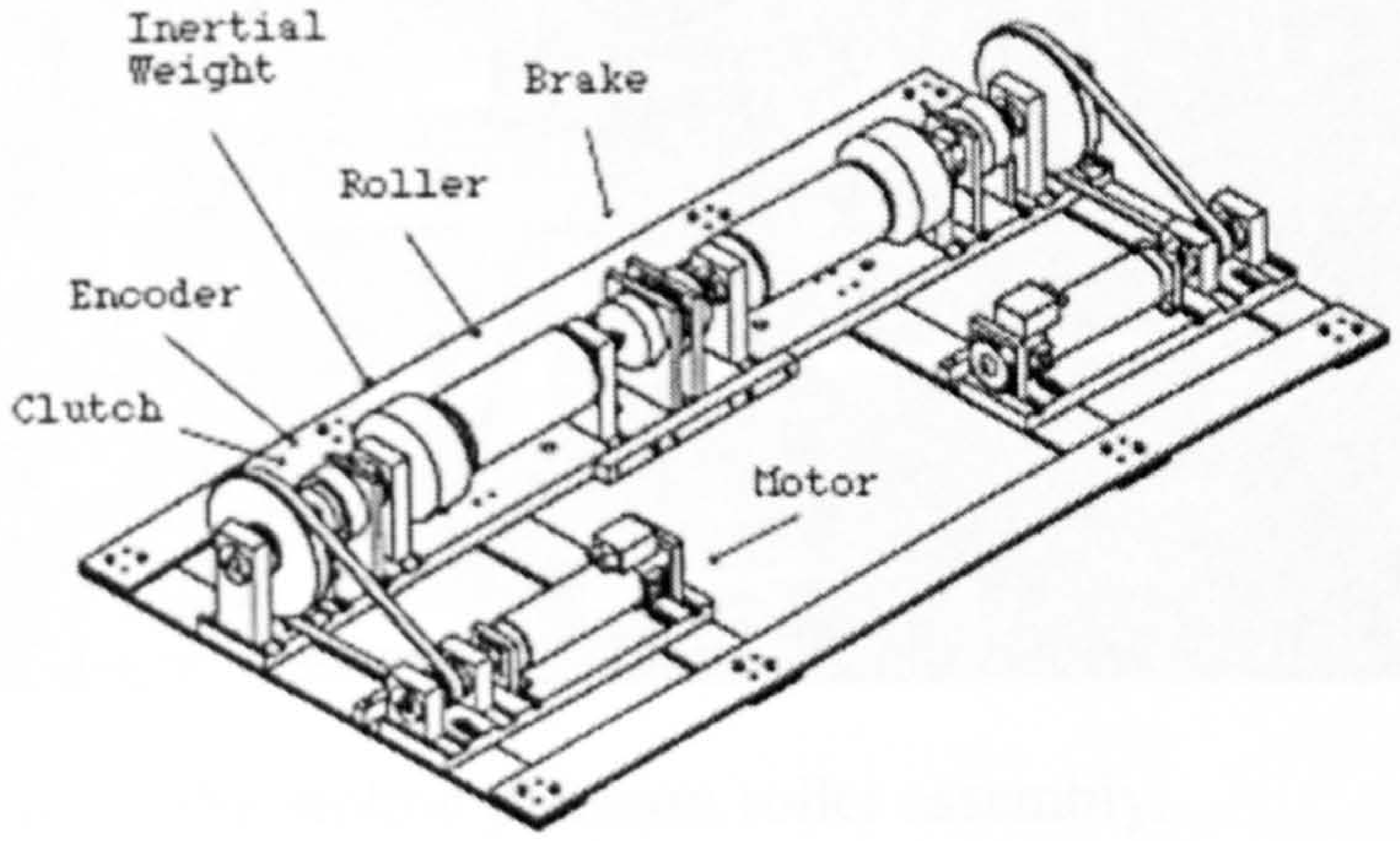


Figure 3. Detail layout of principle wheelchair interface components.

The entire structure is enclosed by a wooden cover so that the user is protected from the moving parts. Two rectangular holes in the cover allow the rollers to stand slightly proud of the surrounding surface allowing wheel contact. Adjustable straps and bars ensure that the wheelchair is rigidly held in place on the rollers, and a ramp allows the user to gain access to the facility.

In Figure 4 note the size and strength of the base plates and shaft supports required to bear the weight of the user and wheelchair when transmitted through a small contact patch.

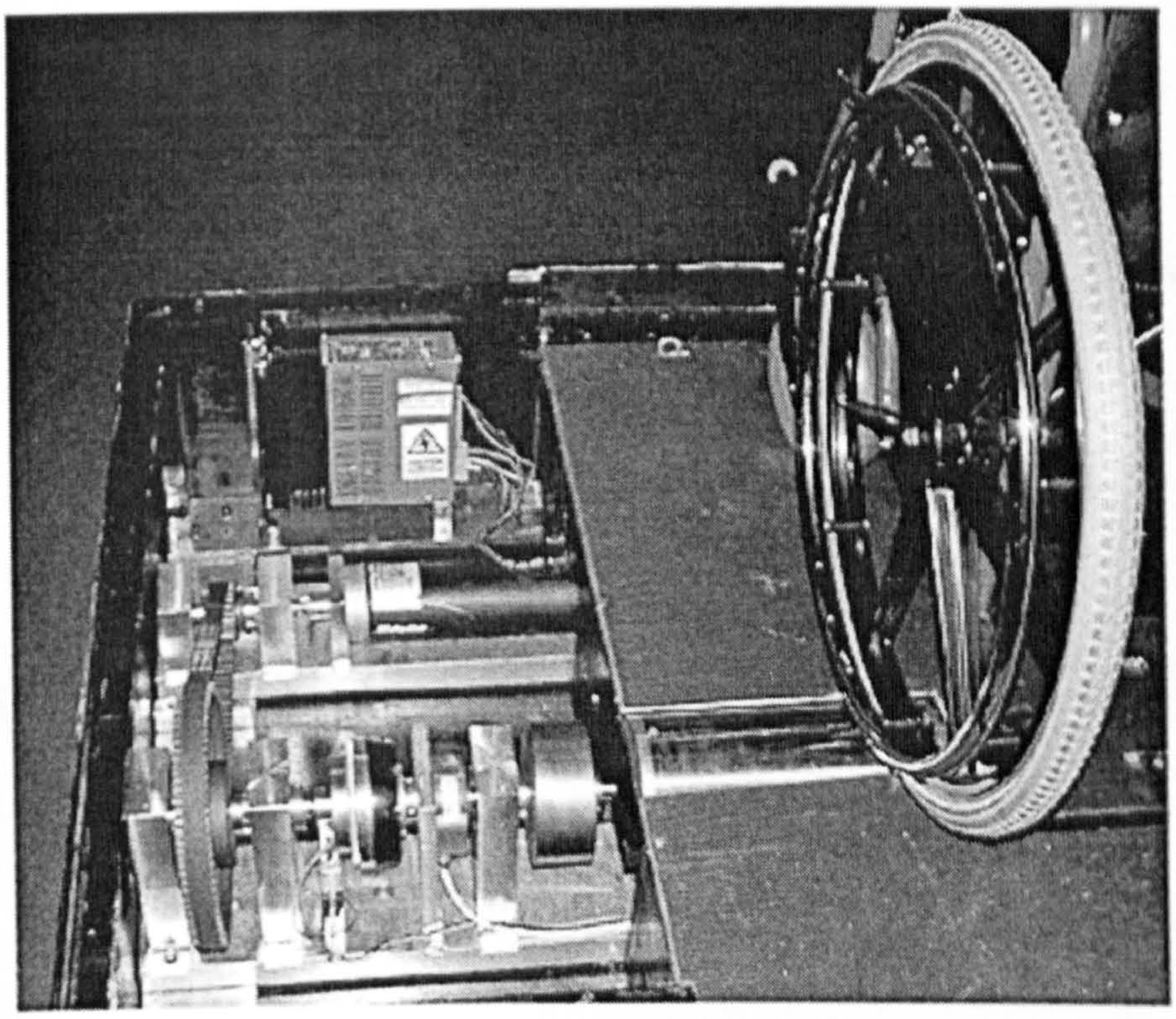


Figure 4. The motion platform roller assembly.

1.2.1 Graphics System

The role of the graphics system is to generate a virtual world that can realistically represent the built environment both visually and physically and so provide feedback along two separate paths.

1.3 VISUAL SIMULATION

The virtual environment is visualised using a three-projector system that provides a 150 degree by 40 degree, high resolution image on a five metre diameter cylindrical screen. Each of the three image channels is edge blended to provide a seamless display. When viewed from the design eye point the image fills most of the users field of vision providing a highly convincing sense of visual immersion within the scene. Graphics are generated on a twelve- processor Silicon Graphics ONXY II with two graphics pipes. This is capable of processing detailed architectural models at high frame rates in order to provide the desired degree of realism. At each time-step in the simulation the graphics are rendered to three separate output channels, each channel sharing the same eye point but with a different angular offset in azimuth, corresponding to the offsets in the projection system. This circumvents the geometrical distortion inherent in large field of view displays.

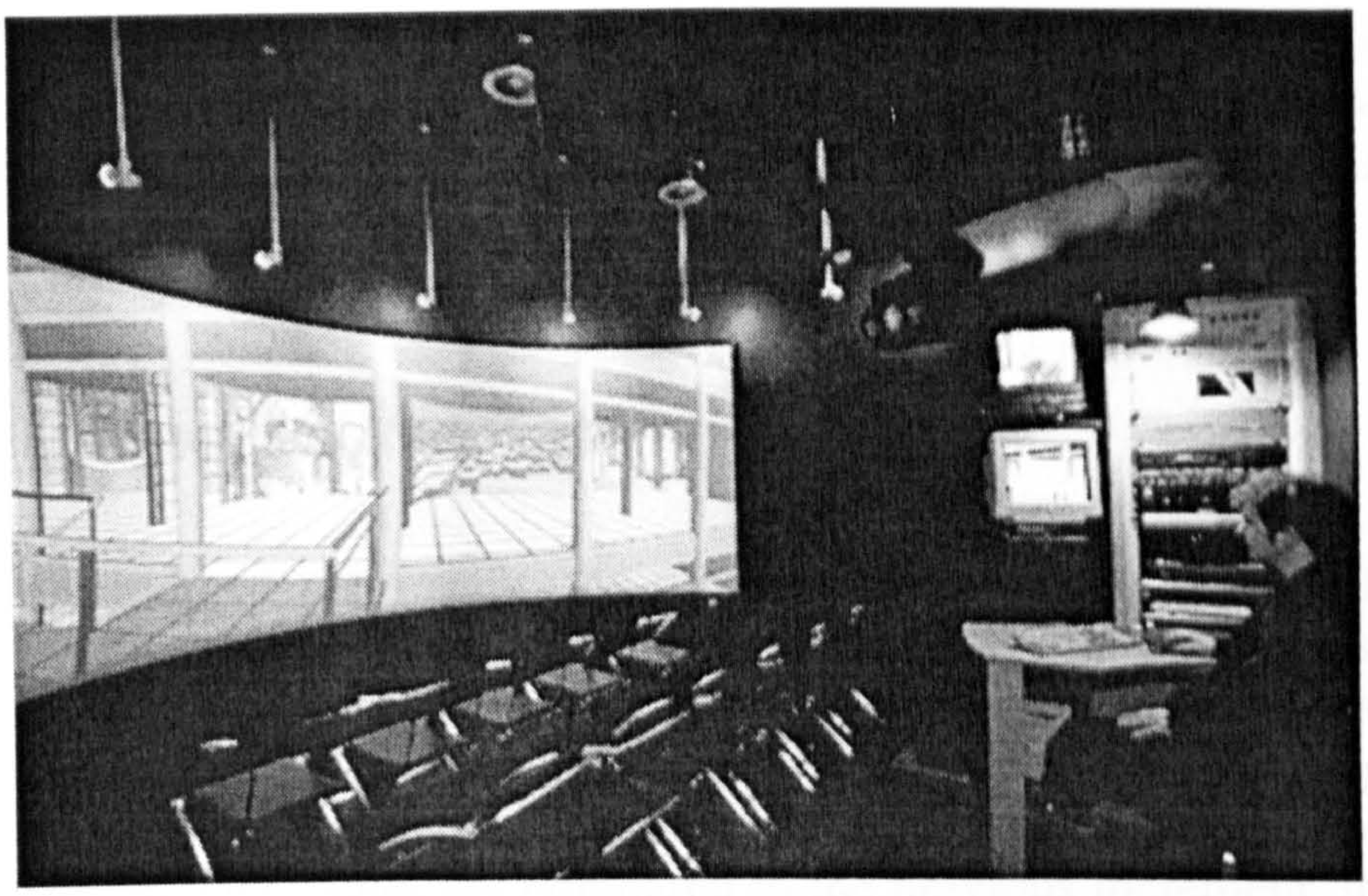


Figure 5. Photograph of Virtual Reality theatre at Strathclyde University.

The software used to drive the virtual environment is based on the Silicon Graphics Performer API. This is a high performance 3-D rendering toolkit for multiprocessed interactive applications. The graphics component is closely coupled to a separate asynchronous module that interfaces between the incoming data from the motion platform control system and the rendering software.

1.4 PHYSICAL SIMULATION

The graphics application requires the Cartesian co-ordinates of the eye point, plus the yaw, pitch and roll angles of the direction of view. Given the yaw angle the remaining two parameters can be calculated based on the wheelchairs attitude on the floor plane. In the database traversal three rays corresponding to the contact patch of each of the rear wheels and the midpoint of the front axle, are intersected with the floor. The normal vector of the ground plane at these points can then be used to calculate the roll, pitch and altitude of the chair and hence the corresponding view. The same intersection procedure can also be used to identify the surface under each wheel, this information then being used to index material properties, such as rolling resistance and surface texture which can be passed back to the control system. Because the system does not support object to object collision detection a further class of intersection appraisal was implemented. This provided a matrix of randomly "jittered" rays aligned with the current direction of motion which proved to be capable of detecting collisions between the wheelchair and even relatively narrow vertical or horizontal obstructions. The length of these rays was varied according to the distance traversed by the wheelchair between frame updates. On detecting a collision condition a flag is also passed back to the control system.
Control System

A large number of environmental features were identified for which accessory physical feedback could enhance the visual feedback of the virtual environment. These included object collisions, slopes and cambers, kerbs, uneven surfaces and different ground surfaces.

1.5 MOTION SIMULATION

As outlined previously, the motion simulator and the graphics software are a close-coupled system. The motion simulator communicates with the graphics system over a TCP/IP network. The task of the motion simulator is to accept incoming data from wheelchair platform, this data relating to the individual incremental angular displacement of both wheels on the motion platform. The current data values are compared to the previous increment, to determine whether either wheel is rotated forward or backward, this information then being passed to the next stage of the algorithm. The basis of the motion control algorithm is the determination, through an analysis of similar triangles, of any translation and also, using the location of the centre of rotation along the rear axle of the virtual wheelchair, the angle through which is turned. These values are passed to the graphics system where the transformation of the eye point and rotation of the view vector can be determined. Feedback from the graphics system determines whether the brakes, clutch or motors should be actuated to provide a physical level of feedback to the user.

1.5.1 Platform Control

The platform control system is based on a standard Personal Computer, running purpose written software, interfacing with the virtual world via a network link using TCP/IP and also with the platform instrumentation via a General Purpose Interface Board (GPIB). The control system monitors the user input by taking incremental readings from the rotary encoders on the motion platform whilst simultaneously controlling the feedback stimuli to the wheelchair on the basis of feedback data received from the graphics system.

The motors are independently controlled for each wheel simply by setting the suitable voltages on the motor control unit, this can be achieved by using the GPIB card and controlled via the software drivers. Thus control of the entire wheelchair platform is accomplished at the physical level by simply writing values on the GPIB card which are generated by software logic controlled by events in the virtual world. Since the user controls the position of the chair in the world realistic force feedback is provided.

1.5.2 Ascending and Descending a steep gradient.

Varying floor surfaces are simulated by altering the resistance to motion of the rollers, using the brakes. Simulation of the wheelchair moving down a slope requires active input into the system, providing a torque against which the user can control their movement down the gradient. Similarly a resistive torque is supplied when ascending a gradient. A variable torque motor is used to provide this input for different grades of slope. When the graphics system detects a collision event a flag is passed back to the control system which then resets the users position to the last "non collided" location and cycles the brakes on the rollers. This temporarily locks the wheels on the wheelchair and coincidentally provides convincing audio feedback.

User Evaluation

Fifteen volunteer manual wheelchair users were recruited to participate in an evaluation of the interface. The evaluation was conducted with the approval of the University of Strathclyde Ethics Advisory Committee and written and informed consent was obtained from each participant. Summary details of the volunteers and the type of wheelchair used by individuals is given in Table 1.

Age Range (years)	Duration of wheelchair usage (years)	Wheelchair models used by volunteers
19 – 59 (mean 41)	0.75 - 54 (mean 18)	Quickie (n=8) RGK (n=3) Elite (n=2) Suntec (n=2)

Table 1: Summary of volunteer profiles.

In evaluating the system volunteers were asked to navigate around the VR model described previously and on completion were asked to rate how realistic various features of the simulated environment compared with 'real life' experiences/expectations. In total thirteen features (Table 2) relating to the performance of the haptic interface and the simulation of wheelchair mobility were rated by each user. This was achieved by use of a questionnaire in which each user was required to rate how realistic each feature was. A simple rating scale was used to achieve this (Table 2).

INTERFACE FEATURE	RATING SCALE RESPONSE
A. Negotiation of kerbs	1. Unable to judge
B. Wheelchair coasting/freewheeling	2. Very unrealistic
C. Negotiation of doorways	3. Unrealistic
D. Turning	4. Moderately realistic
E. Negotiation of corridors	5. Realistic
F. Effort for propulsion	6. Very realistic
G. Changes in floor surface	
H. Ascending slopes	
I. Descending slopes	
J. Perception of motion	
K. Quality of VR image	
L. Collisions with objects	

Table 2: List of features rated by the users and the rating scale used.

Data from the completed questionnaires were then collated and the results are summarised in the histogram shown in Figure 5. Out of the 15 subjects, a single user reported sensations of nausea due to motion sickness and this subject was unable to complete the evaluation. The data from this subject has been included in Figure 5 but the majority of responses were in the 'unable to judge' category. Importantly, when asked about susceptibility to motion sickness the subject indicated that he often became ill during car, train and boat journeys. However, the remaining 14 subjects reported no incidences of motion sickness and their feedback from the questionnaire provides an important pointer to the performance of the system.

The features rated are listed in Fig 5 on the x-axis in ascending (from left to right) order of 'realism'. Thus on average the negotiation of kerbs was the least realistic feature of the interface with collisions rated as the most realistic feature. The method of depicting a wheelchair traversing a kerb in our system relies exclusively on a transient visual stimulation that correlates with the change in wheelchair tilt that would accompany rolling off or over a kerb. There is no additional haptic feedback provided as the platform itself is stationary and does not have a tilting mechanism incorporated in its design. It is because of this lack of non-visual feedback that features such as kerbs are poorly rated. However, note that the average response for this feature is 'moderately realistic'. Features that utilise haptics such as changes in floor surface, changes in slope and collisions are all considered as realistic and it is the combination of the sense of effort experienced by the user together with an accurate visual representation of expected motion that provides the perception of reality and is a vital component in making the overall VR simulation truly immersive.

Refer to Table 2. for key to axis anotation.

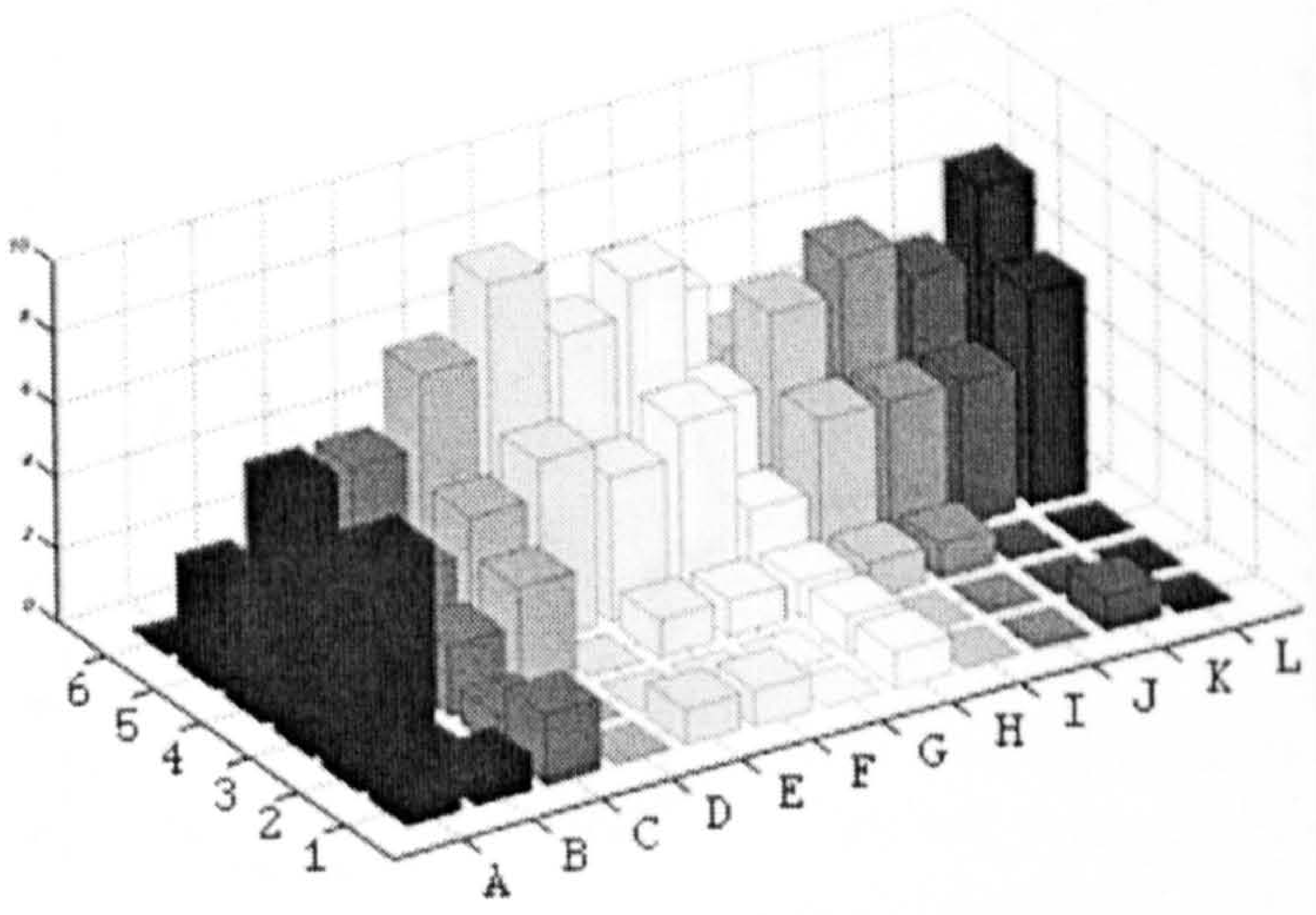


Figure 5. Histogram charting perceived realism of features by user response.

Based on the responses to the questionnaire we believe that the developed interface can provide an important tool for examining access issues in the built environment. The system also has the ability to provide quantitative information on the navigation a wheelchair user employs as they manoeuvre within the simulated built environment. During a test session data relating to the position and heading of the wheelchair are continuously logged by the system creating a file that can be used to recreate the navigation pathway taken by a user. This file also logs the collision points that occur between wheelchair and the features of the VR world (e.g. walls, desks, doorframes etc.). Figure 6 illustrates a map of the VR model used in our evaluation with the navigation paths of the 15 subjects participating shown.

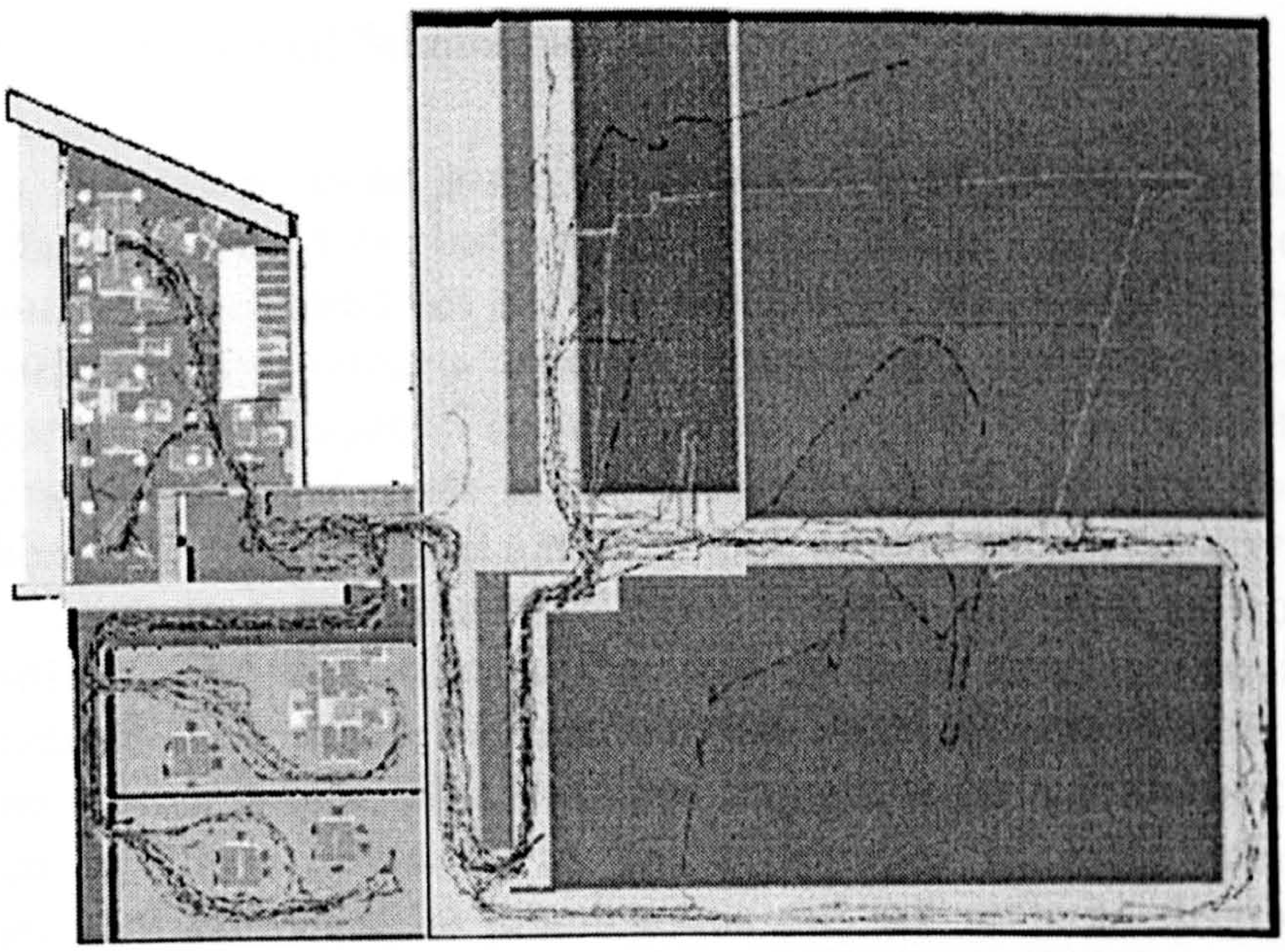


Figure 6. Navigation paths within the Virtual Environment.

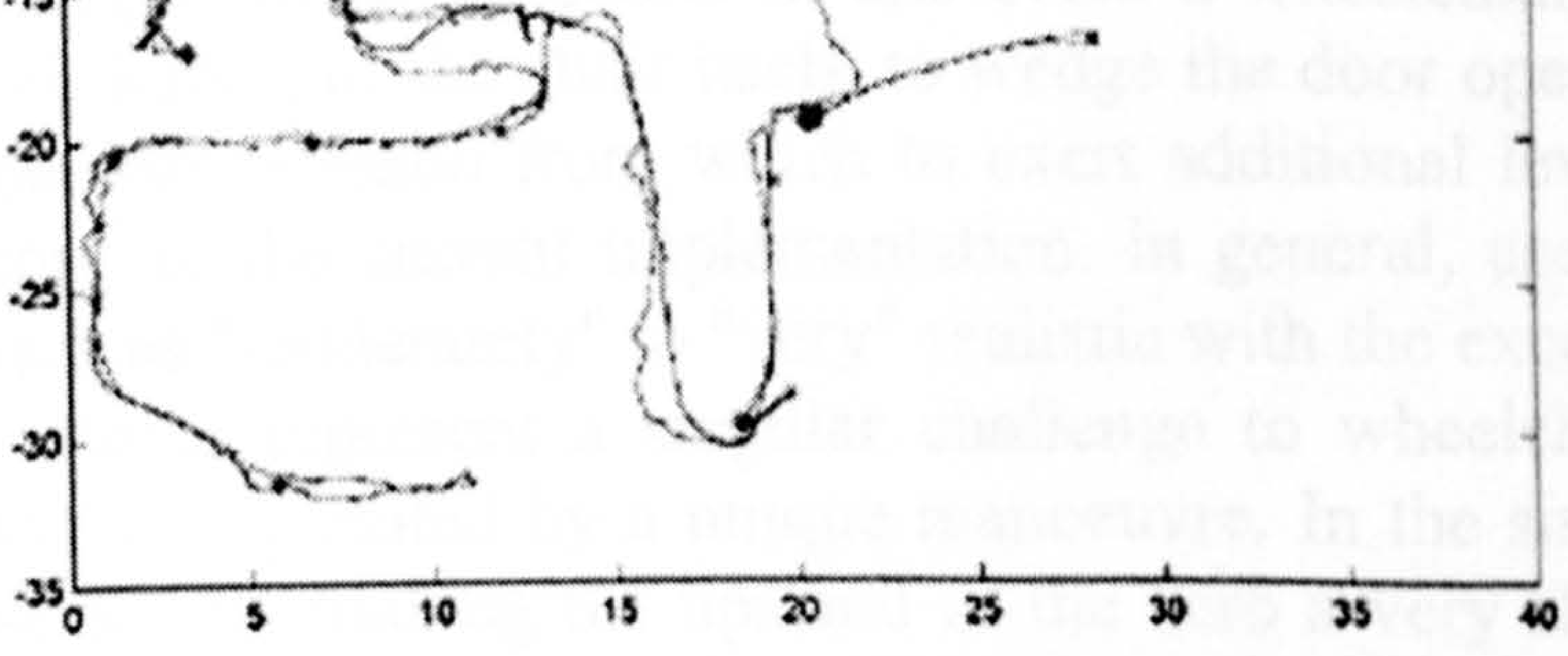


Figure 7. Log of Wheelchair motion within the Virtual Environment.

Importantly, the logged data can be used to analyse the kinematics of the users wheelchair motion and because collisions with objects, walls, doors etc are also logged (see Figure 7, red points indicate collisions) the potential exists to explore in a rigorous and cost effective way how wheelchair users cope with different building layouts

1.6 CONCLUSIONS

1.6.1 The Visual Simulation

The realism of even a simple model proved to be adequate in generating a sufficiently complex, built environment, which could provide all the necessary visual cues required for surface recognition and spatial navigation. The display system demonstrated a mix of positive and negative aspects relating to the technology employed. The wide angle of view was a benefit as this allowed the users direction of view to be decoupled from the direction of motion of the wheelchair. This enabled a user to look around within the environment rather than be constrained to the narrow view frustum common to conventional graphics displays. The downside is that this form of display provides an "out of the window" view which separates a user from objects, which would otherwise be within arms length. This had been thought to be a drawback as wheelchair users tend to make reference to the extremities of their chair when negotiating obstacles. However, in practice this did not seem to disadvantage users of the system. A head mounted display was considered as a possible solution but in trials failed to bring substantive additional benefits.

1.6.2 The Physical Simulation

The physical simulation is based on the interaction of geometry within the virtual world. As such, refinement or extension of these capabilities would usually be just a matter of writing additional software. However the project did highlight some fundamental limitations with this concept in that the system was primarily related to the interaction of the wheelchair with the environment as opposed to modelling the users personal interactions. This is exemplified by the manner in which a user might negotiate swing doors. In this event a wheelchair occupant might tend to use their knees, or the chair itself, to wedge the door open while manoeuvring for a favourable position from which to exert additional leverage. This was beyond the scope of the current implementation. In general, users regarded the physical feedback as "moderately" to "very" realistic with the exception of the treatment of kerbs. Kerbs represent a singular challenge to wheelchair users and are either avoided or negotiated by a unique manoeuvre. In the simulation this feature was represented by making the upstand of the kerb a very short, but steep, incline as opposed to being truly vertical. This allowed the software to treat kerbs in the same manner as all other inclines, requiring substantial input to climb the obstruction, but not faithfully mimicking real world practice.

1.6.3 The Control System

The control system is responsible for interfacing between the motion platform and the virtual world. One of its tasks is to translate the sensed wheel rotations from the incremental movement of the rotary encoders into translation and rotation of the wheelchair via a motion model. This function accurately modelled the gross behaviour of the wheelchair but initially neglected to account for the subtle influence of the castoring front wheels. This castoring action induces two further complications to the model. Firstly, the castors tend to transfer torque between the driving wheels. This tends to stabilise the heading of the chair, an effect that increases with speed. Secondly, the orientation of the castors is a function of the previous direction of motion; any subsequent movement on a new heading must first re-align the castors with the new direction. This can result in unexpected deviations from the desired course, especially among new wheelchair users. When the simulation failed to take account of the first feature it was difficult to maintain a constant heading but a simple algorithm was introduced to mimic the effect of torque transfer and this succeeded in damping out the oscillations. The second feature is more subtle and perhaps of greater concern to powered wheelchair users. In the event this was not corrected and the trial users did not comment on its omission. This will be addressed in a future phase of the project.

1.6.4 The Motion Platform

The motion platform proved to be a complicated electro-mechanical device. Throughout the project it performed to its design potential but also exhibited some limitations. The rollers, brake, clutch, encoder and motor drive all shared the same axis of rotation which necessitated the use of a composite axle shaft. In practice it proved to be difficult to maintain the perfect alignment of all these components. In order to provide minimal friction so as to allow freewheeling, the clutch and brake had to rotate with minimal drag. This meant that they had to be aligned with no more than 0.1 mm axial run-out and therefore required frequent adjustment. Other than the mechanical issues the platforms contribution to the realism and perception of immersion within, and interaction with, the virtual world proved to be the key issue in allowing users to explore the (virtual) built environment in a manner which would allow them to make qualitative judgements on issues of accessibility.

REFERENCES

JN Russell, GE Hendershot, F LeClere & LJ Howie (1997) Trends and differential use of assistive technology devices: United States, 1994. National Center for Health Statistics, No. 292.

Prosthetic and Wheelchair Committee (1996) National prosthetic & wheelchair services report, 1993-1996. London: College of Occupational Therapists.

G. Forrest, G. Gombas, (1995), Wheelchair Accessible Housing: Its role in cost containment in spinal cord injury, Archives of Physical Medicine and Rehabilitation, 76, 5, pp450-452.

T. Frank, E. Abel, (1989), Measurement of the turning, rolling and obstacle resistance of wheelchair castor wheels, Journal of Biomedical Engineering, 11, 462-466.
M. Hofstad, P. Patterson, (1994), Modelling the propulsion characteristics of a standard wheelchair, Journal of Rehabilitation Research Developments, 31, 2, pp129 - 137.
J. Kauzlarich, J Thacker, (1985), Wheelchair tire rolling resistance and fatigue, Journal of Rehabilitation Research Developments, 22, 3, pp25 - 41.
D. Stredney, W. Carlson, J. Edward Swan II and Beth Blonstein, (1995), The determination of Environmental Accessibility and ADA Compliance Through Virtual Wheelchair Simulation, Presence: Teleoperators and Virtual Environments, 4, 3, pp297 - 305
S. Tuljapurkar, N. Li and C. Boe, (2000), A universal pattern of mortality decline in the G3 countries, Nature, 405, pp789 - 792.

Acknowledgements

The authors would like to acknowledge the financial support of the Extending Quality of Life (Equal) EPSRC research program (GR/ M05416). We would also like to acknowledge the help and advice of Spinal Injuries Scotland (SIS) and Disability (Scotland).