University of Strathclyde Glasgow

EngD Medical Devices and Health Technologies

Creating a Real-Time Movement Sonification System for Hemiparetic Upper Limb Rehabilitation for Survivors of Stroke

T.H.Nown

Department of Biomedical Engineering
University of Strathclyde, Glasgow

February 29, 2024

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

Abstract

Upper limb paresis is a common problem for survivors of stroke, impeding their ability to live independently, and rehabilitation interventions to reduce impairment are highly sought after. The use of audio-based interventions, such as movement sonification, may improve rehabilitation outcomes in this application, however, they are relatively unexplored considering the potential that audio feedback has to enhance motor skill learning. Movement sonification is the process of converting movement associated data to the auditory domain and is touted to be a feasible and effective method for stroke survivors to obtain real-time audio feedback of their movements.

To generate real-time audio feedback through movement sonification, a system is required to capture movements, process data, extract the physical domain of interest, convert to the auditory domain, and emit the generated audio. A commercial system that performs this process for gross upper limb movements is currently unavailable, therefore, system creation is required.

To begin this process, a mapping review of movement sonification systems in the literature was completed. System components in the literature were identified, keyword coded, and grouped, to provide an overview of the components used within these systems. From these results, choices for components of new movement sonification systems were made based on the popularity and applicability, to create two movement sonification systems, one termed 'Soniccup', which uses an Inertial Measurement Unit, and the other termed 'KinectSon' which uses an Azure Kinect camera. Both systems were setup to translate position estimates into audio pitch, as an output of the sonification process.

Both systems were subsequently used in a comparison study with a Vicon Nexus system to establish similarity of positional shape, and therefore establish audio output similarity. The results indicate that the Soniccup produced positional shape representative of the movement performed, for movements of duration under one second, but performance degraded as the movement duration increased. In addition, the Soniccup produced these results with a system latency of approximately 230 ms, which is beyond the limit of real-time perception. The KinectSon system was found to produce similar positional shape to the Vicon Nexus system for all movements, and obtained these results with a system latency of approximately 67 ms, which is within the limit of real-time perception. As such, the KinectSon system has been evaluated as a good candidate for generating real-time audio feedback, however further testing is required to identify suitability of the generated audio feedback.

To evaluate the feedback, as part of usability testing, the KinectSon system was used in an agency study. Volunteers with and without upper-limb impairment performed reaching movements whilst using the KinectSon system, and reported the perceived association of the sound generated with the movements performed. For three of the four sonification conditions, a triangular wave pitch modulation component was added to distort the sound. The participants in this study associated their movements with the unmodulated sonification condition stronger than they did with the modulated sonification conditions, indicating that stroke survivors are able to use the KinectSon system and obtain a sense of agency whilst using the system.

The thesis concludes with a discussion of the findings of the contributing chapters of this thesis, along with the implications, limitations, and identified future work, within the context of creating a suitable real-time movement sonification system for a large scale study involving an upper limb rehabilitation intervention.

Contents

List of Figures

List of Tables

Acronyms

9HPT Nine Hole Peg Test. 28, 74

ADL Activities of Daily Living. 2, 5, 6, 10, 23, 30, 172

AHRS Attitude and Heading Reference Systems. 68, 75, 81

ARAT Action Research Arm Test. 23, 139, 144–150, 165–167

BBT Box and Block Test. 27, 28

CIMT Constraint-Induced Movement Therapy. 3, 4

CV Computer Vision. 61, 104–106, 162, 170

DK Development Kit. 107–109

DPDT Double Pole Double Throw. 123

FIR Finite Impulse Response. 110

FMA Fugl-Meyer Assessment. 23, 25

FMA-UE Fugl-Meyer Assessment for Upper Extremity. 5, 18, 29

FoA Feeling of Agency. 127, 128, 130–132

FOV Field of View. 51, 139

GPU Graphics Processing Unit. 109, 162

GRASP Graded Repetitive Arm Supplementary Program. 3, 5

GUI Graphical User Interface. 11

HCI Human-Computer Interface. 11, 129–131, 171, 172

IMU Inertial Measurement Unit. 12, 26, 27, 30, 49, 63, 64, 69, 70, 123, 157, 159, 169, 176

INS Inertial Navigation System. 61, 68, 71, 80

JoA Judgement of Agency. 127, 128, 130, 132, 133, 145, 147–149, 151, 154, 164, 167, 171

MEMS Microelectromechanical Systems. xi, 62

MIDI Musical Instrument Digital Interface. 84, 92, 93, 95, 112, 124, 136

MoCA Montreal Cognitive Assessment. 135, 139, 154, 167

MSE Mean Squared Error. 117, 119–122

MST Music Supported Therapy. 16–19

NGIMU Next Generation Inertial Measurement Unit. 75, 80–84, 87, 125, 227, 229

NMU Number of Movement Units. 151

PD Parkinson's Disease. 5, 14, 16

PSE Patterned Sensory Enhancement. 16, 18, 19

PVAO Position, Velocity, Acceleration, Orientation. 42, 43, 48, 57, 59

RAC Rhythmic Auditory Cuing. 14–16, 18, 19, 23, 174

RCT Randomised Control Trial. 4, 12, 14, 17, 29, 33, 167, 173, 175

RGB Red, Green, Blue. 52, 59, 75, 107

ROM range of motion. 15, 17, 150

SDK Software Development Kit. x, 12, 106–109, 124, 125, 159, 162

SoA Sense of Agency. 125–127, 130–132, 148, 155, 164, 171, 172

TIMP Therapeutic Instrumental Music Performance. 17–19

ToF Time-of-Flight. 105–107

UK United Kingdom. 1, 3

USB Universal Serial Bus. 80

VR Virtual Reality. 3, 5, 6, 52, 59, 106

WMFT Wolf Motor Function Test. 4, 17, 18

YLD Years of Life Lived with Disability. 1

ZUPT Zero Velocity Update. 61, 72, 74, 80, 82–87, 102, 103, 122–124, 159–161, 163, 169

Acknowledgements

As my Eng.D. journey approaches an end, I would like to take a page out of this thesis to thank those that have contributed to the completion of this work. To Christos Tachtatzis and Madeleine Grealy, for taking a schoolboy in academia, and growing him over four years to become a schoolman of academia, and with your endless patience, empathy, and guidance, this experience has been enjoyable and fulfilling. I hope that the benefits and rewards of the work far outweigh the headaches that I have given you during this period of supervision. To Patricia, Steven, and Elise, for the love, understanding, and support you have given me, during this challenging period of my life. I hope that this document finally gives you an answer to all the questions you have about my work. To Andy Kerr, and Ivan Andonovic, for your contributions to the project conception, and your help and guidance through the years. To Tasos, Lesley, for your support, insightful conversation, and friendship. To Glasgow, University of Strathclyde, those responsible for the CDT in Medical Devices and Health Technologies, and the EPSRC, for hosting and funding this work.Finally, to my friends in Glasgow, and those from home, for keeping me human.

Chapter 1

Introduction

Stroke occurs as a consequence of blood not reaching the brain through blockage (ischaemic) or when a blood vessel ruptures and blood accumulates, compressing the surrounding brain tissue (haemorrhagic). Statistics report that stroke is the second leading cause of death worldwide [3], and a leading killer in the United Kingdom (UK) [4]. It is estimated that at least 100,000 people per year in the UK will have a stroke [4], and the average age for men and women to have a stroke is 72 and 77 [5], respectively. The development of stroke treatments along with improvements to hyperacute and acute care have led to a reported decrease in stroke mortality from 2007 to 2017 [6], and an increase in stroke survival rates past three years of incident [7]. Current estimates of stroke survivor numbers in the UK are 1.3 million [4] and are expected to increase due to improving stroke survival rates and increase in population. The ramifications for many of those who survive stroke is disability. This is shown worldwide with an estimated prevalence of 86 million people with an estimated 18 Years of Life Lived with Disability (YLD) as a consequence of stroke, where one YLD is a time-based measure that represents the loss of the equivalent of one year of healthy life due to disability [8].

The growing number of stroke survivors with disability in the UK is putting a substantial burden on the UK economy, as shown by Patel *et al.* [9] with a reported estimate of £25.6 billion attributed in UK per year, consisting of £5.3 billion assigned to incident stroke costs and £20.6 billion for prevalent stroke costs, which is an average cost of £45,409 per person in the first 12 months and £24,778 in subsequent years. In all cases, informal care costs contribute the most to these figures. The same authors project that the cost of stroke will increase to £43 billion in 2025 and £75 billion in 2035, with key reasons stated as: prediction that the number of 'older' people will increase substantially; that care for those people is highly labour intensive, therefore, increasing the prevalent stroke cost; that it is expected that stroke survival rates will improve. Reducing the demand for all forms of care by creating and using a cost-effective rehabilitation pathway toward independent living would effectively reduce the annual societal cost of stroke. One of the most common persisting impairments of stroke is contralateral upper limb weakness as seen in 40% of stroke survivors [10]. This weakness and paralysis results in impaired movement co-ordination patterns, which at a minimum, increases the difficulty of completing Activities of Daily Living (ADL). ADL are tasks that are essential for living independently, and if an individual cannot complete those activities, then a level of care is required to aid in completing those tasks. Therefore, a cost-effective rehabilitation pathway for upper limb weakness is highly sought after.

1.1 Rehabilitation

For stroke survivors with upper limb impairment, rehabilitation aims to improve ADL functionality. The importance of addressing upper limb impairment is underlined through a priority setting investigation completed through the James Lind Alliance in 2021 [11], looking into long-term care and rehabilitation with one of the priority questions stating "What interventions improve arm function after stroke and when should they be provided?". This is further underlined in an earlier priority setting investigation in 2012 with stroke survivors, caregivers and healthcare professionals, based in Scotland, with a resulting priority question stating "What are the best treatments for arm recovery and function, including visual feedback, virtual reality, bilateral training, repetitive task training, imagery or mental practise, splinting, electromechanical and robot-assisted arm training, and botulinum toxin?", reported through Pollock *et al.* [12].

Current recommendations to rehabilitate paretic upper limb functionality consist of intensive, repetitive, task-oriented use for at least three hours, five days a week, as recommended by the National Clinical Guideline for Stroke [13]. Ideally, this would lead to the impaired arm being used more often, and as such, the muscles associated with the movements would become stronger, making future attempts at actions easier, leading to greater motivation to perform movements with the arm in the future. To augment movement practise, therapists can provide treatments to improve functional outcomes of upper limb movement practice. The type of treatment incorporated is typically dependent on the utility of the impaired arm. Stockley *et al.* [14] surveyed UK therapists in 2018, asking about the types of treatments used for upper limb after stroke, with results reporting 30 types used for those with mild paresis, 25 types used for those with moderate paresis, and 16 types for those with severe paresis. Traditional examples of treatments used include Constraint-Induced Movement Therapy (CIMT), Graded Repetitive Arm Supplementary Program (GRASP) and Mirror Therapy. However, these treatments require trained personnel to operate the intervention effectively, and based on the estimates of health professionals that can provide rehabilitation services per country (and therefore, providing an indicator of the level of rehabilitation provision), evidence suggests that there are not enough therapists to meet the need for rehabilitation [15]. Consequently many people with disability, due to stroke or otherwise, are unable to access the rehabilitation services local to them. Developing treatments, that could be self-administered and therefore, enable individuals to access rehabilitation proactively would reduce the disparity between demand and supply of rehabilitation treatments. An example is through the use of Virtual Reality (VR) which is becoming more prevalent in research [16]. However, traditional treatments for upper limb movements are more commonly used due to the associated studies showing positive effects.

Constraint-Induced Movement Therapy

From the perspective of stroke survivors, added difficulty of performing tasks with the impaired arm can lead to 'learned non-use' [17], a common phenomenon where due to the increased difficulty of performing actions with the impaired arm, and as a result of unsuccessful attempts, creates a downwards spiral with suppressed use of the impaired arm, leading to the arm becoming less active and weaker as a consequence, making future movements more difficult [17]. Eventually this spiral leads to the stroke survivor learning to not use the arm [18]. To counter condition learned non-use behaviour, researchers have developed interventions to coerce stroke survivors to use their impaired upper limb [19]. One of the popularly used interventions is known as CIMT and involves constraining the less-affected upper limb with a mitt for 90% of waking hours, referred to as ‘forced use’, whilst also including a package of activities that involves repetitive use of the more-affected upper limb to perform task-oriented training. Evidence from a clinical trial shows that stroke survivors at chronic phases of recovery who undergo a two week programme of CIMT consisting of six hours per day, five days per week, as opposed to ‘usual and customary care’, obtain clinically significant improvements in their upper limb motor function (as determined by Wolf Motor Function Test (WMFT) scores) with persisting improvements, in comparison to stroke survivor controls [20].

As a result of the above work and further research validating the efficacy, CIMT is considered an effective intervention in physical therapy to improve upper limb paresis [21]. Questions remain on how best to utilise this intervention, and what the underlying mechanisms leading to improvement are. As such researchers have developed and tested variations of the intervention, labelled as ‘modified CIMT’, altering intervention variables such as dosage duration per session, amount of sessions per week, and amount of weeks set, with multiple small Randomised Control Trials (RCTs) conducted for these variations. Kwakkel *et al.* [22] published a meta-analysis of these variations, with results of the analysis highlighting no significant differences between types of CIMT procedure, dosage contrast, or timing of intervention post-stroke, as such concluding that important components of this intervention are procedures involving shaping, repetition and, instructions to adapt behaviour. As established, the use of this intervention results in the improvement of upper limb motor function, as evaluated by improvement in functional assessment score, however, these improvements are argued to come more from the increased contribution of compensatory movements, and less from recovery [23], which may explain the lack of reported improvements in ability to manage ADL [24].

Graded Repetitive Arm Supplementary Program

GRASP is a collective of movement activities that aim to prevent learned non-use and improve functionality in the paretic upper limb [25]. The protocol is split into three exercise levels to accommodate difference in impairment and involves, arm and hand strengthening, range of motion training, gross motor skill training and, fine motor skill training [26]. The most recent study reported that for a study in a 'real-world' setting, after one hour per day worth of GRASP exercises for 10 weeks, significant improvements in Fugl-Meyer Assessment for Upper Extremity (FMA-UE) scores were found when compared to baseline measures [27].

Mirror Therapy

A visual stimulation intervention using a mirror placed perpendicular to the chest of the performer, known as mirror therapy, has been well researched as an effective method of improving motor function, reducing impairment, and improving ADL capability [28]. The crux of this intervention is the visual illusion that the more-affected arm, hidden behind the mirror, is able to move and perform tasks as well as the less-affected arm that actually performs the movements.

Virtual Reality

Virtual Reality (VR) is a technology that allows users to interact with a virtual environment enabling the user to access visual, audial, and haptic feedback while interacting in this environment, as dependent on the associated interfaces integrated into the environment. Although this technology is not considered traditional, the use of VR has been used as a treatment approach in rehabilitation for a variety of conditions include Parkinson's Disease (PD) [29], cerebral palsy [30], and stroke rehabilitation [16]. The commercialised technology generally consists of a headset to provide the virtual environment visually and audially, and handheld controllers to allow for interactivity as well as haptic feedback. These systems are considered to be highly immersive and can provide an environment for stroke survivors to practice functional task-oriented movement. The most recent Cochrane review investigating VR-based interventions [16] reported that the use of VR, when compared to conventional therapy, showed a small statistically significant benefit to ADL outcome, and when used as an adjunct to conventional therapy showed a statistically significant moderate benefit to upper limb function.

Summary

These treatments have been shown to be effective at improving upper limb movement to some capacity. However, given the use of clinical assessments as a metric to evaluate functional improvement in these studies, there is difficulty in identifying whether these functional improvements are a result of restitution, or compensation. The inclusion of a physiotherapist to address incoherent movement synergies, spasticity management, and where appropriate introduce treatments through technology or otherwise, could be done to promote restitution. One of the primary tools that a physiotherapist can use to promote restitution is through providing feedback related to the performed movement.

1.1.1 Feedback

Motor skill learning at a behavioural level can be characterised through different frameworks. A commonly used framework is the three phase Fitts and Posner model [31]:

1. Cognitive stage: Early stages of learning movement, requiring high demand on working memory, leading to rapid progression;
2. Associative stage: Demand on working memory diminishes, associations between actions and stimuli improve, movement variability decreases;
3. Autonomous stage: Movements are performed with high consistency with low or no demand on working memory.

In this model, as motor skill learning progresses through stages 1-3 as described, certain performer and performance characteristics evolve as a person becomes more adept at a new movement. Improved competency in detecting and correcting errors are considered to be highly valued properties of this evolution, and this is instigated by feedback.

There are two main types of feedback, intrinsic and extrinsic. Intrinsic feedback, refers to sensory information obtained naturally during movement, such as information from the visual and proprioceptive domains. Extrinsic feedback, refers to information from external sources and are typically referred to as 'augmented', inferring to the improved usefulness of that feedback. Extrinsic feedback can be further subdivided into two types, knowledge of results and knowledge of performance. Knowledge of results provides data or/and information relating to the outcome of an action. Knowledge of performance provides content relating to the process of an action, and this can be applied concurrently, or terminally. For a movement learner, appropriate feedback is a major contributor to movement learning. A question therefore, is how best to provide extrinsic feedback given a type of movement to be learnt. The accumulated feedback from both types (intrinsic and extrinsic) would lead to improved effectiveness of skill acquisition, in comparison to one type of feedback alone. Naturally, providing feedback through knowledge of results has to be through terminal means, as an outcome is required, whereas feedback through knowledge of performance could be done either terminally or concurrently. In the case of providing verbal feedback based on complex movements, the time taken to generate, communicate, and interpret feedback, would result in a large temporal window between the performed movement and the associated extrinsic feedback. As a result the performer is required to remember the movements performed whilst interpreting extrinsic feedback to perform movement correction.

To reduce the temporal window between action and feedback, motion capture systems can be used with an augmented display system to provide movement feedback concurrently. As motion capture systems operate at frequencies much higher than those of human movement, feedback can be generated concurrently to the movement. Commonly these feedback displays are in the visual domain, for example via video capture, which would provide detailed knowledge of performance type feedback through the visual domain through terminal means. However, providing feedback concurrently may prove to be a more effective means of motor skill learning, especially early on in the movement learning process [32]. Creating a concurrent feedback system through the visual domain would be possible, however, many task-related activity practices demand visual attention (for example object handling), as such displaying information visually is likely to create a conflict in attention. Creating a concurrent display through an alternative domain, such as through audio, could resolve this attentional conflict. A common example of audio feedback is with musical instruments, where alterations in actions results in audio changes to pitch, loudness, tempo, and timbre, all of which can be identified simultaneously by listeners. Using an audio feedback system that is concurrent could be of benefit for movement learning as studies have shown the positive effects audio has on motivation [33] and body perception [34], and the negative effects on sports performance that muted audio has [35]. For motor function, neuroscience literature uses the term 'audio-motor coupling' as a hypothesis to explain the connection between audio feedback and movement.

1.2 Audio-Motor Coupling

Audio-motor coupling refers to the neural coupling between the auditory cortex and the motor cortex, resulting in an association between audio perception and motor action, that has been observed in individuals after brief musical training [36]. The auditory cortex is a network of areas that processes auditory information, and include the primary auditory cortex and the secondary auditory cortex. The motor cortex is a network of areas that plan, control, and execute volitional movement. Through recent studies, audio has been shown to impact motor planning [37], and shown to be a stimulus for error detection [38].

The neural mechanisms behind audio-motor coupling have been proposed in the literature, through a recent review written by Damm *et al.* [39]. Auditory stimulus is interpreted by the primary and secondary auditory cortex, which are the initial sites for auditory-motor transformation, before continuing to the planum temporale. The planum temporale is situated in the parietal lobe, posterior to the auditory cortex, and with regards to audio-motor coupling, is of interest for multiple reasons. Firstly, the reported positive correlation of music training and activation response of the planum temporale, an indication of improved competence in auditory processing [40]. Secondly, the planum temporale connects to the dorsal premotor cortex [41], of which reported activity appears to relate to spatial and temporal parameters of audio [42], implying a functional role of disambiguating complex sounds. Lastly, the planum temporale connects indirectly to the ventral premotor cortex via the parietal cortex, relating to the role of sound categorisation [43], as noted by the activity to action-related sounds [44]. As the premotor cortex connects to the primary motor cortex, which is involved in movement execution, indicates that audio could contribute prior and subsequently to movement.

To generate audio-motor coupling, through providing movement feedback via an auditory display, a technique named 'sonification' can be used. Sonification is the translation of data into sound and has been used in a variety of applications to display data (see Dubus *et al.* [45] for a review). Combining sonification and data related to movement kinematics leads to the term 'movement sonification', providing users with a type of audio feedback that represents movement. Movement sonification has been proposed to generate audio-motor coupling [46] as theoretically, augmented continuous information provided through sonification would lead to a richer and more effective internal representation of movement [47], which would enhance online error correction mechanisms [48]. As movement sonification is primarily translating movement kinematics into an accessible form of audio feedback, the technology could be used to provide an action observation platform between demonstrator and patient, which would activate the mirror neuron system [49] and therefore, contribute to movement planning and error identification during and subsequent to movement. Furthermore, the use of movement sonification would produce congruent audiovisual feedback [50], with reported enhanced action observation effects, as shown through functional magnetic resonance imaging activation of various cortex areas including the auditory and motor cortex [51]. To create movement sonification, a system is required to capture movement, translate the data into an auditory display, and subsequently display that data. Operation of a movement sonification system can be achieved independently by a patient without additional risk to safety, allowing for extended periods of movement practice beyond conventional rehabilitation. However, acquisition of a movement sonification system for movement rehabilitation are not commercially available as of present (December 7, 2023), as such creation of a system is required. The impact of an established audio-feedback system along with an established programme backed through evidence to improve upper limb motor function would contribute to improvements in ADL functionality, and therefore, reduce dependency on care.

1.3 Project Motivation and Content

The primary aim of this thesis is to initiate development of a real-time movement sonification system that could be used by stroke survivors to create meaningful audio feedback based on movements performed with their hemiparetic arm. To achieve this aim the following list of objectives were formed:

- Review the systems used in projects in the literature;
- Identify component candidates for new systems;
- Create a real-time movement sonification system;
- Evaluate the extent that the output of the created system to the equivalent output obtained from a gold-standard motion capture system;
- Validate if users with and without hemiparesis can associate their movements with the system output, and therefore, use the system to receive movement-based feedback;

The thesis reports on the creation of a real-time movement sonification system, to be used by clinicians, therapists, and patients alike, that tracks movement of the upper limb, and display a concurrent sound that represents the movement performed. The system itself could be used individually or/and collaboratively with therapists, to identify differences in performed movement, and provide an enriched environment for task-oriented movement practice. As part of the system, a Graphical User Interface (GUI) accompanies the system to provide sonification options for users of the system, allowing for various movement kinematics to be sonified, and for a variety of auditory dimensions to be used as the auditory display.

Considering the commonality of arm impairment persisting into the chronic phase of stroke recovery, and arm recovery stated as a desired priority area of research in post-stroke life [12], provides a level of justification to develop a technology that would be used as part of a rehabilitation intervention. As this technology is not available off-the-shelf, the system would need to be created via the acquisition of a motion capture technology, and the development of a platform to convert data into an auditory display. The finalised system should be low-cost and accessible for the general public to be used as a treatment/therapeutic tool for upper limb rehabilitation.

The thesis as presented commences work on creating a suitable real-time movement sonification system. For this thesis, Chapter 2 details the different types of existing auditory techniques reported in the literature, before elaborating on movement sonification. Chapter 3 details a mapping review of the existing real-time movement sonification systems, presenting an overview of the motion capture technologies used, the sections of anatomy monitored, and the types of sonification configurations applied. This chapter concludes with brief evaluations on the motion capture technologies when considered as appropriate for upper limb motion capture, and comments on the appropriateness of the sonification configurations. The remarks in this chapter are brought forward to subsequent chapters. Chapter 4 displays the technology development of two individual movement sonification systems, one that uses inertial sensor technology and the other that uses markerless computer vision technology. Positional estimates from these systems were compared to an industrial gold-standard motion capture system as a method to evaluate movement sonification accuracy, of which the system incorporating markerless computer vision was judged to be the more accurate. This system was subsequently applied to investigate the usability of the system as a Human-Computer Interface (HCI), through the process of an agency study, as described in Chapter 5. The background, methodology creation, final methodology, results for both non-neurologically impaired participants and stroke survivor participants are contained in this chapter. The thesis concludes with a discussion of each contributing chapter, and a discussion of the work in context to the creation of a large scale RCT study.

1.3.1 Research Contributions

The work presented in this thesis contributes to the movement sonification literature, with the contributions as follows:

- A review of the created systems in the real-time movement sonification literature, detailing the components used in the systems to track human movement, and the sonification mappings used to create audio feedback;
- A novel position estimation system that uses acceleration from a single Inertial Measurement Unit (IMU) sensor, with a push-to-make momentary switch, to estimate position of the hand whilst performing reaching movements. This system is used as part of a movement sonification system termed 'Soniccup' in this thesis;
- Creation of a real-time movement sonification system using the motion capture capabilities of an Azure Kinect Camera and associated Body Tracking SDK to sonify hand position. This system is termed 'KinectSon' in this thesis;
- A comparison study to evaluate the output from the Soniccup and KinectSon, to the Vicon Nexus system which was assumed to be the ground truth;
- An agency study to evaluate acceptability of the KinectSon generated audio feedback through reaching movements performed by volunteers with and without upper limb impairment.

Journal

Published:

- T.H.Nown, P.Upadhyay, A.Kerr, I.Andonovic, C.Tachtatzis, M.A.Grealy, A Mapping Review of Real-Time Movement Sonification Systems for Movement Rehabilitation, *IEEE Reviews in Biomedical Engineering*, 2023, 16, 672-686, `https://doi.org/10.1109/RBME.2022.3187840` [52]

Conferences

- T.H.Nown, A.Kerr, I.Andonovic, M.A.Grealy, C.Tachtatzis, Developing a Real-Time Movement Sonification System for Upper-Limb Rehabilitation for Stroke Survivors, *44th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Piscataway, NJ, 2022.

- T.H.Nown A.Kerr, I.Andonovic, M.A.Grealy, C.Tachtatzis, Creating a Real-Time Movement Sonification System as an Upper-Limb Intervention for Stroke Survivors, *BioMedEng 22*, 2022.

- T.H.Nown A.Kerr, I.Andonovic, M.A.Grealy, C.Tachtatzis, Verifying a Sense of Agency with a Real-Time Movement Sonification System for Stroke Survivors with Hemiparesis, *Joint 2022 SRR & Scottish Stroke AHP Forum Winter Conference: Rehabilitation to Support Self-Management*, 2022.

Chapter 2

Review of Auditory Based Interventions

This chapter reviews the current literature relating to auditory techniques used for rehabilitation, provide a critical analysis of the research into movement sonification, and states the starting point for the contributing chapters described in this thesis.

2.1 Rhythm and Music-Based Interventions

Audio interventions for motor rehabilitation range from the use of simple rhythmic cues to multi-faceted musical components, and have been applied to various rehabilitation interventions with those who have Parkinson's Disease (PD), cerebral palsy, and stroke [53]. A recent Cochrane review [54] concluded that audio interventions show promising results but require high quality RCTs to evaluate efficacy. This section provides a review of the recognised rhythmic and music-based interventions that have been applied to rehabilitate the paretic upper limb of stroke survivors.

2.1.1 Auditory Cuing

Rhythmic Auditory Cuing (RAC) is an example of a passive audio interventions, that uses a metronomic sound byte or music with a salient beat, to instigate synchronised timing of motor execution during specific movements. This technique has been widely researched in a variety of movement disorders, including those due to stroke. Described through Thaut *et al.* [55], four mechanisms are proposed for applications of rhythm:

1. Rhythmic stimulation and entrainment, where an external rhythm is used to regulate physiological and behavioural functions, and provide 'temporal templates' to aid with movement priming, movement anticipation, motor preparation, and potentially bypass infarct areas through the activation of alternate pathways;

2. Patterned information processing, where rhythm is used to create temporal structures to enhance learning and perception;

3. Differential neurological processing, the parallel activation of multiple areas of the cortex may provide alternative transmission routes for information processing and learning;

4. Affective-aesthetic response, where a stimulus that affects arousal, motivation and emotion, which in turn impacts on motivation and learning.

The existing literature reports beneficial effects of RAC in gait training and balance, specifically with improvements in stride symmetry, length of stance phase on the paretic leg, knee angle control, mediolateral and vertical displacement of centre of mass [56], in addition to a significant difference found in gastrocnemius muscle activity compared to without sound [56]. Studies have been undertaken with stroke survivors performing reciprocating reaching movements with their paretic upper limb, Thaut *et al.* [57] and Kim *et al.* [58] each presented a study with volunteers; Malcolm *et al.* [59] conducted a two week RAC pilot study with five volunteers that consisted of three hour sessions, with outcome measures obtained before and after the intervention. Additionally, Sethi *et al.* [60] presented a study, asking volunteers to perform movements at a comfortable speed, maximum speed, and inline with audio cues set to the preferred speed. The reported effect of this intervention by the above studies includes:

- Increase in elbow range of motion (ROM);

- Decreases in compensatory shoulder and trunk movements;

- Decrease in upper limb variability during reaching, resulting in converging movement synergies towards 'normal' reaching movement;
- Decreased standard deviation for optimal maximum acceleration and decreased number of movement units, highlighting smoother movement;
- Decreased movement time and increased mean reaching velocity, resulting in faster movement.

Other projects have included tempo-alternative auditory dimensions with RAC to stimulate movement. A methodology presented by Kang *et al.* [61] used an amalgamation of RAC and variation in pitch, termed 'melodic auditory cuing', to represent shoulder abduction, holding, and adduction for stroke survivors. This methodology can be further elaborated to include harmonic, and dynamic-acoustical patterns of music, to synthesise a more detailed audio description of a movement in space, time, and force, referred as Patterned Sensory Enhancement (PSE) [55]. Hong [62] reported beneficial effects of PSE for stroke survivors compared to a control group, while Han *et al.* [63] presented a small study that reported improved effects of PSE for stroke survivors compared to the effects of RAC. Other examples can be observed with persons with cerebral palsy [64] and PD [65].

2.1.2 Music Supported Therapy (MST)

Music Supported Therapy (MST) is an active audio intervention programme of exercises to train upper limb movement using musical instruments. The original conception of this technique made use of drum exercises that focused on gross motor movement, and musical keyboard exercises that focused on fine motor movement [66]. Unlike the passive auditory techniques mentioned above, performers receive audio feedback from the musical instruments, which can be used to detect and adjust erroneous movements. Under guidance from a therapist, each programme consists of patients performing mass repetition of varied movements and performing music with progressive complexity, resulting in an activity that engages various areas of the brain including the involvement of emotional and motivational areas [67]. Studies involving chronic stroke survivors using MST as an intervention, have shown evidence of cortical motor map reorganisation and enhanced auditory-motor coupling [68, 69]. Tong *et al.* [70] conducted a controlled pilot study with stroke survivors, that isolated the effect of audio in MST, and showed increased WMFT scores for those that obtained audio feedback through instrument use during the programme in comparison to those that used muted instruments. Ripollés *et al.* [71] reported a pre- and post-intervention MST study with chronic stroke survivors, highlighting improvements in motor and cognitive function, along with improved emotional outcomes. Recent research by Gráu-Sanchez *et al.* [72] have further developed MST to:

- increase intensity and include more instruments and hence vary the ROM trained;
- include peer-group sessions and incorporate an artificial intelligence platform to utilise gamification into the programme, to boost intrinsic motivational factors and promote autonomy;
- adapt the programme from a clinical/laboratory environment to a home-based environments.

The authors labelled this approach as 'enriched MST' as part of a presented protocol with stroke survivors.

2.1.3 Therapeutic Instrumental Music Performance

Therapeutic Instrumental Music Performance (TIMP) is a technique which combines the use of external audio cues and music playing through an instrument or digital application, to simulate non-musical movement patterns [55]. Theoretically this combines feedforward motor priming and movement planning component, with the error identification and correction, feedback mechanism associated with music playing. Street *et al.* [73] presented a feasibility RCT study with this technique highlighting the general acceptance with stroke survivors, and further described two case studies using this technique as an intervention [74]. Haire *et al.* [75] presented a study with stroke survivors investigating interventions composed of either TIMP alone, TIMP and motor imagery, or TIMP and motor imagery with metronome cues, with all three conditions producing improvements as identified through FMA-UE and WMFT scores.

2.1.4 Evaluation of Rhythm and Music-Based Interventions

To summarise, there appears to be a trade-off between simplicity and applicability in the use of these techniques. For passive audio techniques, RAC is desirable in many aspects, including the low cost, ease of implementation, general acceptance and feasibility of interventions that use RAC, as well as beneficial impact on motor function. However, the applicability of this technology is limited to reciprocal movements, which for the upper limb generally involves repetitive reaching motions. The application of this intervention for alternative actions is not reported in the literature; this may be due to the unsuitability of rhythmic based intervention on movements that are arrhythmic. Additional auditory elements as shown through PSE provides a viable option for upper limb rehabilitation, as this is also low cost, has general acceptance, with the addition of flexibility to accommodate a variety of movements. The compromise is the increased complexity of accommodating and integrating multiple types of cues in a representative way that the listener can identify and translate auditory information into movement. For active audio techniques, MST uses an approach through musical instrument playing, tied to a variety of functional movements, which requires the acquisition and setup of equipment to operate effectively. Researchers are looking to expand on this approach with the inclusion of additional instruments, increasing variation in movements trained, at the compromise of increased implementation cost. The combinational use of auditory cuing and musical playing shown in TIMP would theoretically produce the beneficial outcomes of the audio cuing and audio feedback aspects, with the compromise of added equipment, and the requisite design and training implementations.

All the described auditory techniques above show promising signs of improving motor rehabilitation, however, conclusive high-quality evidence on the efficacy of these interventions is reported to have a high risk of bias for studies before 2017 [54], and most trials are reported to use small sample sizes. As such, further research is necessary to validate the efficacy of these techniques in large scale studies. However, the limitations and compromises for each of the mentioned techniques, when applied as an intervention, may be exacerbated as the implementation is scaled up for a larger sample size. A viable audio-based alternative may be through the use of sonification, a technique that retains the ease of implementation of RAC, whilst obtaining the versatility of PSE, MST, and TIMP.

2.2 Sonification

As introduced in Section 1.2, sonification is defined as 'the transformation of data relations into perceived relations in an acoustic signal for the purposes of facilitating communication or interpretation' [76] and is a multi-disciplinary topic that involves data science, audio engineering, and psychology, to name a few. Sonification has multiple descriptive labels to organise and describe configurations, for example De Campo [77] proposed keywords *event-based*, *model-based* and, *continuous*. Event-based sonification can be further categorised into *auditory icons* and *earcons*, generally these types of sonification strategies are short snippets of audio that are triggered based on predetermined criteria. Model-based sonification is defined as 'the general term for all concrete sonification techniques that make use of dynamic models which mathematically describe the evolution of a system in time, parameterise and configure them during initialisation with the available data and offer interaction/excitation modes to the user as the interface to actively query sonic responses which depend systematically upon the temporal evolution model.' [78]. The majority of studies in the sonification literature, however, makes use of continuous sonification methods, namely *parameter mapping* which involves the association of information with auditory parameters for the purpose of creating an auditory display. Unlike the setup in model-based sonification which uses a dynamic model to create an audio output, parameter mapping techniques make use of a predetermined static configuration. Configuration designers that use parameter mapping have many options available to create an audio output that relates to the input data, including the data dimension to sonify, the audio properties to alter, choice of positive or negative polarity, scaling, and whether to sonify multiple streams of data in parallel through the use of one-to-one, one-to-many, or/and many-to-one, data-to-audio mappings.

Parameter mapping sonification at a high-level is described as *direct*, associated with communicating movement-associated data, or *error-based*, which is associated with data in relation to a target. Mid-level descriptions of sonification describes the characteristics of audio that are altered, including tonality, spatialisation, and loudness. Whereas low-level descriptions provide detail to the sonification, for example increasing the tempo of generated audio notes for a positive increase in magnitude. In addition, the number of components in a configuration should be considered when designing an auditory display, as increasing dimensionality provides additional information but increases the likelihood of sensory overload for the listener [79].

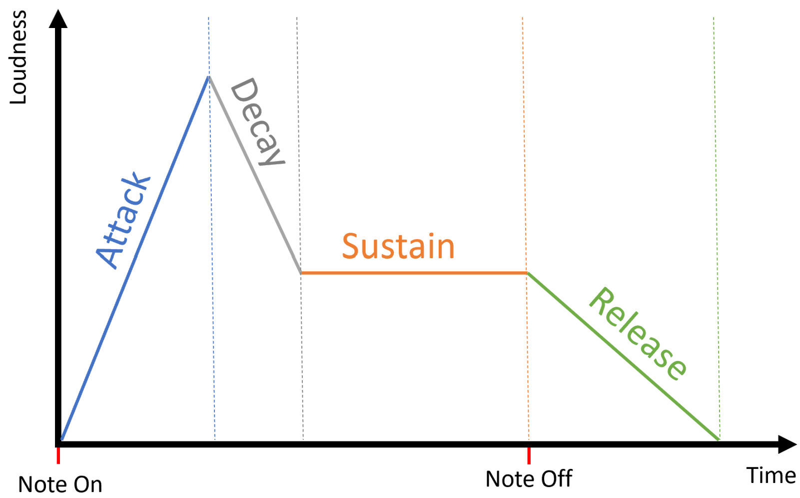

Figure 2.1: Diagram showing segments of a signal envelope related to audio loudness.

An example of the complexities of sonification design was reported by Walker [80]. A simulated factory was monitored by volunteers, of which data dimensions consisting of temperature, pressure, size and, rate, were each parameter mapped to one of either loudness, pitch, tempo, or onset/attack time (see Figure 2.1 for note envelope theory). The investigators assigned each data dimension to a different auditory dimension, and assigned an intuitiveness rank based on how well the mapping corresponded as theorised by the investigators, with the hypothesis that the most intuitive setting would correspond with the best outcomes. Table 2.1 shows their ranking for each configuration.

Table 2.1: Table 1. from Walker's thesis [80], showing the intuitiveness ranking of the sonification configuration as determined by the investigators.

Display Dimension	**Data Dimension**			
	Temperature	Pressure	Rate	Size
"Intuitive" ensemble	Pitch	Onset	Tempo	Loudness
"Okay" ensemble	Loudness	Pitch	Onset	Tempo
"Bad" ensemble	Onset	Tempo	Loudness	Pitch
"Random" ensemble	Tempo	Loudness	Pitch	Onset

Volunteers were asked to listen to a sonification output corresponding to a data dimension and through the audio, maintain data values in a certain range by performing corrective actions (pressing buttons) associated with increasing/decreasing the value. The investigators recorded the response accuracy and response time with outcomes indicating that the sonification configuration design that produced the best results - correct response and lowest reaction time - was ranked low on the intuitiveness as determined by the investigators [81].

Evidently, intuitively selecting a mapping, polarity, and scale, leads to unjustifiable assumptions that the holistic output will be beneficial to novice users. As such, assuming that positive outcomes for various systems in different applications are transferable for upper limb rehabilitation would also not be justifiable. However, creating a testing paradigm for every sonification option is also not feasible [78]. This is justified by the results of a systematic review of sonification mapping strategies presented through Dubus *et al.* [45], where a systematic search identified 60 projects prior to (and including) January 2013 and extracted keywords to identify 33 physical dimensions and 30 auditory dimensions used for sonification, which when considering scaling and polarity choices, leads to a mountainous amount of options to choose from to design a sonification configuration. A perpetual question, therefore, is how best to effectively communicate with a listener when it comes to describing movement.

To summarise, the use of parameter mapping sonification provides potential for creating audio that can be aesthetically pleasing and informative, whilst also flexible to accommodate different strategies, and therefore, could be used as an effective auditory displays to provide information for movement rehabilitation. Consequently, the flexibility of the methodology also leads to increased heterogeneity between different audio outputs, and as sonification intuitiveness is not shared, leads to a need for extensive testing to establish which options are useful for an application. However, there is a consensus on certain aspects of audio feedback for rehabilitation applications, which can be applied to limit options. Firstly, that concurrently presented audio feedback presenting higher quality feedback, increasing the saliency of errors, and enhancing performance monitoring as a result [82], and secondly, that feedback should be perceivably continuous as opposed to discrete, providing a high rate of feedback that can display higher frequency oscillations (associated with jerky paretic movement for example), and additionally follows gestalt principles of perception [77, 78, 83].

2.3 Movement Sonification

The inclusion of audio stimulation through audio-based interventions, could be a catalyst for upper limb motor skill learning. However, a methodology limitation is the compromise between simplicity and applicability. A rehabilitation intervention based on movement sonification may provide the desirable traits of music-based interventions, without many of the associated limitations. This section describes movement sonification systems and known examples of use in rehabilitation literature.

Figure 2.2 shows the rudimentary steps of a movement sonification system, with each step acting sequentially, initiating from the movement of the person. The movement performed is monitored by a chosen motion capture technology setup to track an anatomical segment/landmark during movement. The raw data obtained through this technology is sent to a smart device, typically a host personal computer, for data processing to obtain the desired physical dimension to sonify. Proceeding this is a conversion step where the desired physical dimension is translated into the auditory domain, before emitted through an audio output technology. In the context of this project, a motion capture technology will be used to track the paretic arm of stroke survivors, and as they perform movements, the data obtained is processed to obtain

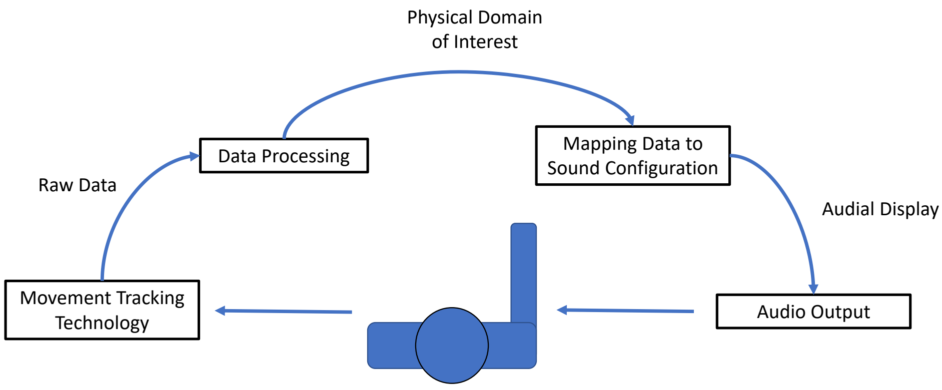

Figure 2.2: Figure depicting a movement sonification system tracking human movement.

movement kinematic data ready for translation to the audio domain. Commonly in rehabilitation literature the translation process makes use of parameter mapping type sonification, however other sonification types have been used as well. Commercial movement sonification systems for the purpose of rehabilitation towards movements associated with ADL completion do not exist at present. Therefore, researchers have created a variety of movement sonification systems as dependent on the intended application(s).

2.3.1 Application

At time of writing, few existing reviews relate directly to the topic of movement sonification. One includes a meta-analysis [79] analysing the effect of real-time movement sonification (along with RAC) on post-stroke upper limb recovery, finding improvements for Fugl-Meyer Assessment (FMA) [84] and Action Research Arm Test (ARAT) [85] scores, albeit with the analysis containing only five studies that use movement sonification (labelled as real-time auditory feedback). Another constitutes a scoping review [86], which summarises the benefits of sonification on physical therapy. Lastly a narrative review has been published [46] describing the progress that has been achieved in the industries of sport and rehabilitation. Snowballing through these reviews and external literature searches, has led to the identification of several projects relating to upper limb rehabilitation and real-time movement sonification.

Maulucci *et al.* [87] published a study into paretic movement performance with stroke survivors using error-feedback sonification, in a forward reaching task. The system comprised a starting touch-plate and three ending touch-plates, with an electromagnetic tracking system tracking the paretic hand, to monitor and record the pathway taken during the reaching movements. A 'normal' pathway was created via 24 right-handed volunteers without neurological impairment performing reaching movements to the three end touch-plates. The normal pathway was used to create a field of 3-dimensional space from the start point to each end point and was the physical metric used for error-feedback synthesis. Sixteen stroke survivors were split into a feedback and control group. Each participant performed reaching movements on 18 sessions, consisting of 42 trials each, of which 24 trials (situated in the middle) had audio feedback for the feedback group only. Audio feedback was emitted if, and when, the hand of the participant ventured outside the field of 'normal movement' as they reached for the intended target. The audio feedback consisted of a rising audio pitch correlating to the measured distance from the field of normal movement. When compared to the control group, results indicated that the feedback group performed more of their reaching movements within the calculated field of normal movement, and at the end of reach the hand was closer to the target touch plate (medial/lateral axis only). However, results indicated no improvement in range of movement towards normal. To evaluate retention, a session was completed two weeks post-intervention, with the same methodology except all participants were without audio feedback. Data obtained from the retention test and at the end of the 18th training session, were compared, showing that the feedback group showed a retained significant decrease in the elevation of the upper arm, indicating an improvement towards normality. However, given the low sample size in each group, the heterogeneity of the participants in terms of activation and capability, and the potential lack of audio feedback emitted in this system, it is unclear if the reported sonification strategy has led to the effect of improved movement, as opposed to the repetitive movement incurred by the study protocol.

A study by Wallis *et al.* [88] detailed a system tailored towards the rehabilitation of reaching and grasping movements. The system used a six camera marker-based motion capture system and extracted various movement features to translate and produce audiovisual feedback to the user. Inline with the recorded movements, the visual display showed either a virtual arm or an altering image. The generated auditory display linked various parts of the tracked upper body to audio properties, consisting of both concurrent sonification and audio alerts. An audio alert of a triangle strike was used to indicate a successful reach. The hand position during reaching was linked to harmonic progression, providing alterations in pitch, whilst the velocity of the hand was linked to note length, providing a temporal alteration. The elbow joint angle was linked to loudness of an orchestral background component. Shoulder and trunk compensatory movements were linked to cymbals and rain-like sounds, respectively, with increasing loudness as the participant moved outside of a predetermined threshold. Three stroke survivors were recruited to test the system with generated feedback recorded and used to formulate future goals for the system. Duff *et al.* [89] used this system to compare rehabilitation outcomes to conventional therapy. Twenty-one stroke survivors were recruited, and divided into the experimental (11 participants) and control (10 participants) group. The experimental group used the system to perform forward reach-and-grasp movement to cone-shaped objects with their paretic arm, receiving concurrent audiovisual stimuli in the process. The control group received dose-matched therapy consisting of object-based reaching tasks. Although results of this study report that when compared to the control group, the experimental group improved their quality of movement (as determined by the Kinematic Impairment Measure score [90]) for certain reaching tasks, results also showed that a larger change occurred with the control group in the motor function section of the FMA and in the Stroke Impact Scale score. The created system in this project consists of a marker-based motion capture system, and through placing retroreflective markers on the movement performer and objects, a virtual environment can be created to provide concurrent audiovisual feedback. However, as these systems are typically high cost to acquire, require mounted multi-camera hardware to be installed and a trained system operator to run the system, further development would be needed to improve the usability and scalability of the system, so that large scale studies using this technology could be feasible.

Robertson *et al.* [91] presented a study identifying whether side of stroke lesion impacted the influence of real-time audio feedback on movement. Sixteen stroke survivors participated in this study, with eight experiencing left hemisphere damage. Reference data was collected from 10 participants without neurological impairment. Each participant was requested to perform reciprocal reaching movements to each of nine targets positioned in front of the participant. An electromagnetic tracking system with multiple tracking sensors placed on the upper body, was used to obtain movement data which was subsequently used as an input to an audio feedback synthesizer. Two sonification configurations were used, one consisted of mapping increasing loudness to the reduced distance between the hand and the target, and the other used this same configuration with an added audio spatial element which was dependent on the orientation of the tracked hand. Results of the investigation showed differences between the two groups. Participants with right hemisphere damage showed improved movement smoothness whilst maintaining pre-intervention velocity, whilst participants with left hemisphere damage showed decreased movement smoothness along with a decrease in peak velocity. Whilst these findings are promising, considering the low sample size for the two groups (left affected hemisphere group versus right affected hemisphere group) with eight participants in each, in addition to the heterogeneity of the participants, produces uncertainty about how replicable these results will be.

A pilot study presented by Schmitz *et al.* [92] investigated the influence of audio feedback on the upper limb movement using a created system involving IMUs [93]. The study recorded the affected hemisphere and hand dominance of the participants. Participants were randomly allocated to an experimental or control group, where the experimental group received concurrent audio feedback whereas the control group did not. The experiment consisted of participants performing investigator-led pointing tasks in a designated space segregated into a 3x3 grid, before progressing to object transferal to locations in the same grid. Movement data captured by IMUs attached to the paretic upper limb of the participant were used to generate audio-feedback in the following sound configuration: arm velocity was mapped to loudness of the audio; elevation angle of the hand to pitch; radial distance of the hand to audio brightness; azimuth angle of the hand to stereo panning. Each sound configuration was individually adjusted to the preferred movement speed of the participant. Box and Block Test (BBT) assessments showed significant improvement in the experimental group, however the experimental sample size of seven was insufficient to draw any conclusions. As such an experimental protocol was created by Schmitz *et al.* [94] with the intent of expanding from the pilot study and recruiting 32 stroke survivors for the study. The system to be used, as described in Schmitz *et al.* [94], makes use of seven IMU sensors to indirectly track the joint angle for each upper limb. Utilising this system for rehabilitation would lead to extended setup time to attach and sync each sensor for motion capture purposes, especially for independent use.

Similarities can be observed in a separate investigation presented in Scholz *et al.* [95], who also uses a system containing IMUs on the upper arm and wrist to sonify upper limb movements. To instruct participants on where to move their wrist, a created 3-dimensional volume was segregated into a 3x3 grid, and extrapolated in the caudal/cranial axis was used. The extrapolation was segregated into six regions each. Audio feedback was produced as participants moved their wrist in the 3-dimensional volume. The sonification configuration consisted of mapping the wrist position to audio brightness along the medial-lateral axis; wrist position to loudness along the proximal-distal axis; wrist position to pitch in the cranial-caudal axis. Twenty-five right-handed stroke survivors with moderate impairment were recruited for the study, of which motor skill learning and movement smoothness were evaluated. Participants were randomly allocated to a movement sonification group (that received the audio feedback as described) and a control group, and assessed pre- and post-experiment, with significant results observed in the dampening of perceived joint pain, and movement smoothness improvement in the movement sonification group. The results also showed a non-significant improvement in hand function in the movement sonification group. This study methodology was used in an expanded study presented by Nikmaram *et al.* [96] who used an additional site of investigation along with a new movement sonification system that used a LEAP motion sensor to capture movement. Forty stroke survivors were recruited and assigned to a treatment or control group. The study methodology reports that the amount of movement performed per session was controlled, with the treatment group receiving additional audio feedback in comparison to the control group. Results of the study report an insignificant improvement in movement smoothness, along with strong appraisal of enjoyment by the participants. Given that each session lasted approximately 30 minutes and the reported heterogeneity of the number of training sessions completed by participants (7 - 46 sessions), makes it difficult to interpret the reported efficacy of movement sonification as a rehabilitation tool.

Friedman *et al.* [97] presented a system for hand rehabilitation, termed MusicGlove. The system made use of a modified glove to detect contact between the distal phalange of the thumb and five anatomical locations on the hand, to interact with a game intended to encourage different functional grip movements of the user. The game required different hand postures in line with coloured notes observed in a visual display. Dynamically correct postures rewarded the user with an increase in music volume and in-game alerts to signify the correct action. Alternatively, an incorrect posture, or incorrect timing, produced a decrease in music volume. For the pilot study, 12 stroke survivors were randomly allocated an order of hand therapies, interchanging from conventional hand exercises, MusicGlove, and an isometric version of the MusicGlove device, termed IsoTrainer. The IsoTrainer provided a reference for the study to identify differences in game performance as a result of proprioceptive richness when using the MusicGlove, as such the IsoTrainer was also an input to the serious game as described above. Each participant completed three training sessions of six hours each, one session for each hand therapy. Results showed a significant improvement in participant hand function after using the MusicGlove – determined by BBT and Nine Hole Peg Test (9HPT) scores – compared to the conventional hand rehabilitation sessions, and a general non-significant improvement when comparing the MusicGlove therapy against the IsoTrainer therapy. The authors reported a linear correlation between the serious game score and BBT score. Due to the small sample size, and relatively small effect size, the reported future project work is to investigate the MusicGlove in a domicile environment with a more general population of participants, as opposed to specifically stroke survivors in the chronic phase of recovery.

Colombo *et al.* [98] presented a validation and feasibility evaluation of a training protocol, termed SonicHand, involving a system with a Leap Motion Controller as the movement capture system. The criteria of feasibility were determined by adverse effects of stroke survivors performing the training protocol, and the qualitative evaluation of an observing therapist, who evaluated the appropriateness of exercises in terms of difficulty and execution. Thirty participants, including 15 stroke survivors with upper limb impairment, participated in a training protocol that involved movement exercises of the wrist and hand. Those with impairment received additional audio feedback. Two sound configurations were used, one configuration associated movements with an arpeggio progression and loudness, the other configuration associated movements with only loudness. The exercise determined the sound configuration used. The concluding remarks of the publication reported the feasibility of the protocol. The study by Raglio *et al.* [99] proceeds from this feasibility evaluation into an RCT, to investigate the effect of this intervention in comparison to conventional interventions, with regards to level of impairment, pain, and perceived quality of life. Sixty-three stroke survivors in the subacute stage of recovery completed a four-week treatment schedule, were randomly allocated into a group that received conventional interventions, and another group that followed the same protocol as described above. In comparison to the conventional treatment group, reported results showed a significant positive effect for the sonification group on FMA-UE total score, with a large contribution coming from the improvement of the distal segments of the assessments. However, the sonification group did not show significant improvement at the proximal level, unlike the control group. Considering that the audio feedback related only to movements in the distal segments of the upper limb, it may be that the extrinsic feedback distracted participants from the movement associated intrinsic feedback, creating a disparity in recovery between the proximal and distal segments of the upper limb.

Peyre *et al.* [100] presented a study to identify preferences in various sonification configurations and to evaluate the influence of audio feedback on forward reaching movements. Thirty participants, of which 15 were stroke survivors, were asked to perform repetitive reaching movements between two arbitrary points at their own preferred speed, with IMUs attached to both upper limbs. Data from the IMUs were used to calculate arm position, which was subsequently used as the input to the sonification configurations. Five sonification configurations were used for this study. Based on the reported preference from the participants, the most popular sonification configuration divided the distance between the arbitrary reaching points into three equal sections, and assigned each section to environmental sounds, namely the sound of wind, a river, and birds. The second most popular sonification configuration emitted a continuous musical phrase as the reaching movement was performed. The remaining three sonification configurations consisted of a discrete musical phrase, pitch alteration, and a tempo alteration of a drum beat. There was also a reported increase in movement time, along with larger variation in movement time, when audio feedback was played during movement in comparison to the trials completed without added feedback. This was to be expected, given the self-directed speed of movement, and the argument that the novel use of receiving concurrent audio feedback would have encouraged playful engagement with the movement sonification system. The findings of this project are of interest considering that the preferred sonification configuration produced the least informative feedback, given that there was only three soundscapes that were alternated between.

Evaluation of the Literature

Research into the use of movement sonification as a tool for upper limb rehabilitation is present in the literature, with projects looking to identify feasibility, changes in quality of movement, and enjoyment, through the addition of audio feedback. The studies as presented above show a wide range of approaches to improving upper limb function in terms of system used and methodology presented, therefore, collating results and drawing conclusions from these studies is not feasible, especially considering the low reported sample size within most of these studies. As such, a large amount of research is required to evaluate efficacy, long-term effectiveness, and fully understand the mechanisms that audio feedback has on upper limb rehabilitation, especially regarding complex gross movements and how they would translate to ADL competency.

2.4 Chapter Summary

Rhythm and music-based interventions are currently being researched with reported positive effects, however these techniques have applicability limitations with respect to the types of movement that can be trained, and the equipment required to perform that training. To alleviate these limitations, research is investigating an alternative technique, movement sonification, to provide real-time audio feedback to a user performing upper limb movements. Existing projects have detailed a variety of systems to create rehabilitation applications, however, due to the limited sample size in these studies, the heterogeneity of the participants, and nature of the study methodology, there is inconclusive evidence that movement sonification is of benefit, at present. To further investigate this area, a real-time movement sonification system is required to synthesise audio feedback, however, a commercialised system does not exist at present, and therefore, creation of a system is required to begin studies. As system creation guidance is sparse, and existing reviews primarily investigating the effects of movement sonification, provides motivation to create a review on real-time movement sonification system setups.

Chapter 3

Mapping Review

Following from the findings of the literature review, it is evident that a real-time movement sonification system is not available commercially and therefore, a system must be created. The creation of such system constitutes a combination of hardware and software elements, with each combination producing different challenges and limitations. At present, a guidance resource to create such a system does not exist. This provided motivation to produce a review with focus on the created systems in the literature.

For this chapter, the following research questions were proposed:

RQ1 What types of motion capture technologies were chosen to create the real-time movement sonification systems found in the literature?

RQ2 To create auditory displays through real-time movement sonification, which auditory dimensions and physical dimensions combinations were chosen in the systems found in the literature?

RQ3 For the above questions, which components would be good candidates to create a new system intended for upper limb rehabilitation?

The work presented in this chapter has been published in IEEE Review in Biomedical Engineering, with the initial search performed on the 14th of January 2021, and the results reflect the outcome of the methodology at the time of search. A subsequent search has since been performed on the 11th of September 2023 using the same methodology to identify systems in articles published between the two search dates. The outcome of this subsequent search is presented in Section 3.5 of this chapter.

3.1 Existing Reviews

As previously mentioned, reviews looking into movement sonification generally report intervention effects. Most of these same reviews also provide description as to the system composition of the reviewed articles. For Ghai *et al.* [79], their systematic review and meta-analysis of sonification and rhythmic auditory stimulation studies assessing recovering arm functions post-stroke included 23 articles, listed five projects using sonification, showing four different sonification configurations. Additionally, Guerra *et al.* [86] published a scoping review on the use of sonification for physical therapy in human movement that contained 35 articles, including 13 RCT showing beneficial effects in each. The review also lists 13 different types of motion capture technologies used in the articles.

Analysing the existing reviews from Ghai *et al.* [79] and Guerra *et al.* [86] show that movement sonification systems used are not commercially available off-the-shelf systems, instead they comprise motion capture systems integrated with another smart device (PC or otherwise) which contains software components to synthesise audio feedback. The use of these systems for movement sonification is also reviewed by Wang *et al.* [101], who investigated system setups for interactive wearable upper body technologies in a rehabilitation context. However, the review only contained seven articles that used auditory feedback. Due to the set focus and the selection criteria applied in each review paper, results are limited in the number of movement sonification articles identified.

Motion capture technologies used in existing rehabilitation research have been developed and used primarily within a laboratory environment, for specific applications, as such many of the systems are inappropriate for alternative environments and limited within rehabilitation for the following reasons: extensive setup, challenging data for sonification (inertial sensors, EMG), limits or constraints movement (ergometer, tablets), high acquisition cost (marker-based motion capture system, goniometer), high environmental dependence (Microsoft Kinect), and/or be unpurchaseable (custom platforms).

The aforementioned reviews which target movement sonification have identified only four options for sonification, which is a low number of approaches considering the high number of results found through Dubus *et al.* [45], which is a dedicated review on sonification alone. Sonification systems are also used extensively outside of the healthcare domain but their potential for rehabilitation has not been assessed.

3.2 Methodology

This chapter has been conducted starting from a global overview of current movement sonification systems irrespective of intended application. It intends to: identify trends in system setups, establish if there are motion capture technologies that have been overlooked for rehabilitation applications, provide scope on technological requirements for next generation rehabilitation technologies, and create a resource that future researchers in movement rehabilitation can utilise to develop appropriate and effective rehabilitation tools. To achieve this three key components of movement sonification systems in the literature are identified and analysed: 1) The types of physical to auditory parameter mapping; 2) The part(s) of the body that are tracked; 3) The types of tracking technology.

3.2.1 Search

The methodology commenced with a systematic search for published articles to identify systems of interest. Components within each system were identified and keyword coded, to form a database of keywords, that were later synthesised into a graphical display. Following PRISMA guidelines (Page *et al.* [102]), database searches were performed on the 14th of January 2021 on the following electronic literature databases: ACM, IEEE Xplore, PubMed, ScienceDirect, SCOPUS, Web of Science. Full search strategies for each database are shown in Table 3.1. For each search strategy the word 'sonification' was included to focus the search on relevant projects, and where possible was shortened to sonif* to include variations of the word (such as sonify), the remaining keywords have been selected to cover a wide area of movement-related keywords to increase search results, where similarly each keyword was reduced to include a wildcard symbol allowing for variations of the word to be included. In addition to articles yielded from the database searches, relevant articles cited in the reference lists of existing literature reviews were also extracted.

3.2.2 Eligibility Criteria

Duplicate articles were removed, and article abstracts were screened to ensure that the articles met the following inclusion and exclusion criteria. Inclusion criteria applied: 1) Written in English; 2) Describes an implemented system; 3) System monitors human anatomical movement; 4) System produces at least one auditory output; 5) Auditory output described provided 'real-time' feedback, i.e. does not exclusively provide terminal feedback or provide feedback that exceed 100ms from the input [103]. Exclusion criteria applied: 6) System only monitored ocular movements; 7) System where the movement was captured exclusively through a computer mouse, computer keyboard, or touchscreen; 8) System described did not mention a connection between physical movements and an auditory output; 9) System tracks an object, where the object was not attached to a human; 10) System used microphones to record musical instruments as a method to monitor movement; 11) Section of tracked human anatomy was not stated; 12) Movement tracking technology was not stated.

The screening of articles for eligibility was carried independently by two researchers. Initially abstracts were considered for eligibility, before the assessment of the full manuscript. In disagreements between the first two reviewers, a third researcher was sought. Following eligibility checks, relevant information was extracted from each article by the primary researcher, and assigned a coded keyword into the appropriate category in a data table. An example project entry is shown in Figure 3.1.

Table 3.1: Full search strategies for each electronic database.

Electronic Database	**Search Criteria**
ACM	[[Publication Title: sonif*] OR [Abstract: sonif*] OR [Keywords: sonif*]] AND [[All: mov*] OR [All: reach*] OR [All: grasp*] OR [All: point*] OR [All: rotat*] OR [All: acceler*] OR [All: velocit*] OR [All: position] OR [All: danc*] OR [All: kine*]]
IEEE	((((“Publication Title” Sonification) OR “Author Keywords”: Sonification) OR “Abstract”: Sonification) AND (“All Metadata”: mov* OR “All Metadata”: reaching OR “All Metadata”: grasping OR “All Metadata”: rotat* OR “All Metadata”: acceleration OR “All Metadata”: velocity OR “All Metadata”: position “All Metadata”: danc* OR “All Metadata”: kine*))
PubMed	(sonif*[Title/Abstract]) AND (mov* OR reach* OR grasp* OR point* OR rotat* OR acceler* OR velocit* OR position OR danc* OR kine*) Filters: English
ScienceDirect	“Find articles with these terms”: movement OR reaching OR pointing OR rotating OR acceleration OR velocity OR position OR dancing “Title abstract or author-specified keywords”: sonification
SCOPUS	TITLE-ABS-KEY(sonif*) AND (mov* OR grasp* OR reach* OR point* OR rotat* OR acceler* OR velocity* OR position OR danc* OR kine*) AND (LIMIT-TO(LANGUAGE, “English”)) “Filter by subject area”: Exclude: Biochemistry, Genetics and Molecular Biology; Medicine; Mathematics; Physics and Astronomy; Social Sciences; Agricultural and Biological Sciences; Chemistry; Environmental Science; Materials Science; Chemical Engineering; Earth and Planetary Sciences; Health Professions; Pharmacology, Toxicology and Pharmaceutics; Immunology and Microbiology; Decision Sciences; Energy; Nursing; Veterinary; Business, Management and Accounting; Economics, Econometrics and Finance.
Web of science	#1: (ALL = (mov* OR reach* OR grasp* OR point* OR rotat* OR acceler* OR velocit* OR position OR danc* OR kine*)) #2: (ALL = sonif*) #3: (#1 AND #2) AND LANGUAGE: (English) Filter by Research Area: Exclude: Chemistry; Materials Science; Education Educational Research; Environmental Sciences Ecology; Biochemistry Molecular Biology; Marine Freshwater Biology; Medical Informatics; Behavioral Sciences; Fisheries; Oceanography; Cardiovascular System Cardiology; Pharmacology Pharmacy; Veterinary Sciences; Agriculture; Anesthesiology; Energy Fuels; Geochemistry Geophysics; Life Sciences Biomedicine Other Topics; Mathematical Computational Biology; Remote Sensing; Zoology; Anatomy Morphology; Astronomy Astrophysics; Audiology Speech Language Pathology; Automation Control Systems; Biodiversity Conservation; Communication; Cultural Studies; Dermatology; Food Science Technology; History Philosophy of Science; Mathematics; Mechanics; Meteorology Atmospheric Sciences; Mining Mineral Processing; Polymer Science; Psychiatry; Reproductive Biology; Social Issues; Theater.

Project Number	Reference	Year	Published In	Physical Category	Physical Dimension	Auditory Category	Auditory Dimension	Anatomy	Technology	Technology Category	Application
62	H.Brückner et al.	2016	Journal of Multimodal User Interfaces	Kinematics	Position	Loudness-Related	Loudness	Wrist	IMU	Inertial Sensor	Rehabilitation
						Spatial	Stereo Panning				
	H.Brückner et al.	2014	2014 IEEE International Conference on Consumer Electronics (ICCE)			Pitch-Related	Pitch				
					Velocity	Loudness-Related	Loudness				

Figure 3.1: Snippet of database showing completed project data insertion. Each project is allocated a unique identification number, which contains articles identified inside the data columns Reference, Year, Published In, highlighted through light blue. Data from each article was extracted, keyword-coded, and inserted into the remaining data columns, highlighted through dark red.

3.2.3 Keyword Coding

For data extraction purposes, five keyword lists based on the work of Dubus *et al.* [45] have been created (i) Physical Dimension, (ii) Auditory Category, (iii) Anatomy, (iv) Technology and (v) Application. The classification of every article considered in the review after application of inclusions/exclusion criteria is provided in the Appendix.

Physical Dimension

From initial data extraction, nine intermediate-level physical dimension keywords in three high-level categories were selected. The Kinematics category constitutes of Position, Orientation, Joint Angle, Velocity, Acceleration, and Jerkiness. The Kinetics category comprises of Force/Pressure, and Energy. The Other category is set as a catch all category, and keyword, for alternative physical dimensions to the listed above.

Auditory Category

For the auditory domain, six high-level category keywords are selected Pitch-Related, Loudness-Related, Temporal, Spatial, Timbral, and Event-Driven. Each category is defined as follows with reference to the sound generated: Event-Driven - Sound sample played upon a movement parameter-based trigger; Loudness-Related - increase or decrease in perceived audio intensity; Pitch-Related - increase or decrease in perceived audio frequency; Spatial - change in perceived location of sound source; Temporal -

audio alteration in the time dimension; Timbral - audio alteration in the frequency dimension that excludes changes in pitch or loudness.

Anatomy

Sixteen human anatomy keywords taken from Martini *et al.* [104] were selected to accommodate large and small sections of anatomy required to assign appropriate keywords in this section. The contents of the list are: Head (includes movement of the face and neck) Shoulder, Upper Limb, Upper Arm, Elbow, Forearm, Wrist, Hand (includes movement of fingers) Trunk (includes movement of the chest, abdomen, pelvis and back) Hip, Lower Limb, Thigh, Knee, Lower Leg, Ankle, Foot (includes movement of toes). Additionally, to represent projects that use a physical dimension associated with a tracked centre of mass of a person, the keyword 'Centre of Mass' was included. No distinction is made between anterior and posterior sections of each anatomical segment, nor the amount of each segment.

Technology

Thirty-six technology keywords have been assigned to three high level categories labelled as Inertial Sensor, Camera and Other. The Inertial Sensor category contains: Accelerometer, Gyroscope, IMU, Mobile Phone, and Gaming Controller (IS). The Camera category contains: Marker-Based Motion Capture, Virtual Reality Controller, LEAP Motion Controller, Kinect, Infra-Red, Optical Image, LED-Based Optical Capture, Gaming Controller (Ca). The Other category contains: Graphics Tablet, Microphone, Rotary Encoder, Haptic Device, EMG, MMG, Ergometer, Goniometer, Tendon-Based Parallel Robot, Ultrasonic Sensor, Variable-Resistance Elastic, Bend Sensor, Cadence Sensor, Electromagnetic Tracker, Gaming Balance Board, Tension-Activated Switch, Electrical Contacts, Textile Stretch Sensor, Piezoelectric Transducer Pickup, Infra-Red Proximity Sensor, Footswitch Sensor, Customised Speed Sensor, and Force/Pressure Sensor.

Application

Each project included in the review has been assigned a keyword, from a list of 11, to provide context on the type of project that the movement sonification system is used. This list constitutes of: Gait, Sport, Performing Arts, Immersive Environment, Rehabilitation, Body Perception, Balance/Posture, Visual Impairment, Task Performance, Alternative Locomotion, Other.

3.3 Results

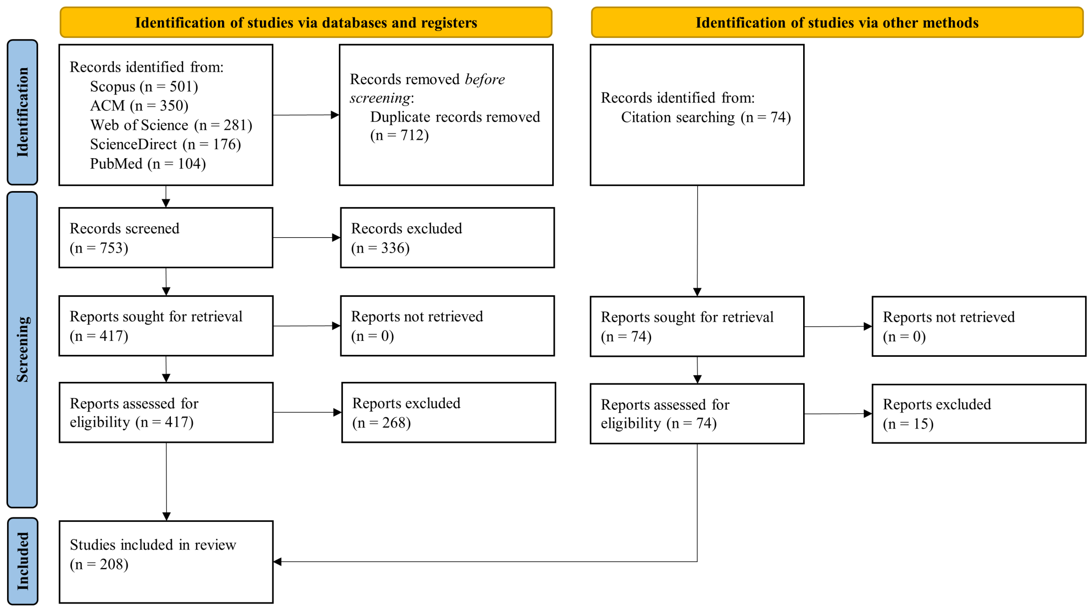

Figure 3.2: PRISMA flow diagram, which include searches of databases and other sources. Resulted in 208 articles brought forward for review.

As shown by Figure 3.2 a total of 1465 articles were identified from the search results, with 712 duplicates, resulting in 753 article abstracts screened for eligibility. The full text of 417 articles were assessed for eligibility, resulting in a total of 149 articles for data extraction. From studies identified outside of the database search 74 were identified, with 59 assessed as eligible for inclusion, leading to final total of 208 articles included in this review. For the following results sections, percentages are used as part of the statistical description for the results, however due to the methodology of the review and the complexity of movement sonification systems in the literature, the projects often recorded multiple elements for each category, and consequently for the following data analysis, the sum of the percentages shown in each statement, may exceed 100%. Based on the analysis of the complete data table, graphical visualisations were created to address the following sections.

Types of Physical to Auditory Parameter Mapping

Keywords entered in the Physical Dimension and Auditory Category data columns have been analysed separately and in combination for each project. Figure 3.3 presents a bubble plot of the chosen movement sonification options with Physical Dimension keywords listed on the vertical axis, and the Auditory Category keywords listed on the horizontal axis. From the 145 projects recorded in the database, 48 distinct types of combinations are recorded, out of a possible 54 - as limited by the keyword categorisation - amounting to a total of 397 recorded combinations within the search. No recording was obtained for the combination of Jerkiness to Loudness-Related, Energy to Spatial, Force/Pressure to Spatial, Jerkiness to Spatial, Joint Angle to Spatial, and Other to Spatial.

The highest number of recordings for the Physical Dimension is Position with 133, amounting to 33.50% of the recorded Physical Dimension keywords, whereas for the Auditory Category, the highest number is Pitch-Related with 105, amounting to 26.45% of the recorded Auditory Category keywords. The combination of Position and Pitch-Related keywords recorded the most with 42 recordings in these results, amounting to 10.58% of all chosen combinations.

Other popular keywords in the Physical Dimensions list are Velocity with 55 recordings, Acceleration with 47 recordings and Orientation with 68 recordings in projects. Likewise, other popular keywords in the Auditory Category list are Timbral with 75 recordings and, Loudness-Related with 73 recordings in the reviewed projects. In contrast, the recordings of Energy and Jerkiness in the Physical Dimensions list, have been recorded on less than 10 occasions in these results, whilst for the

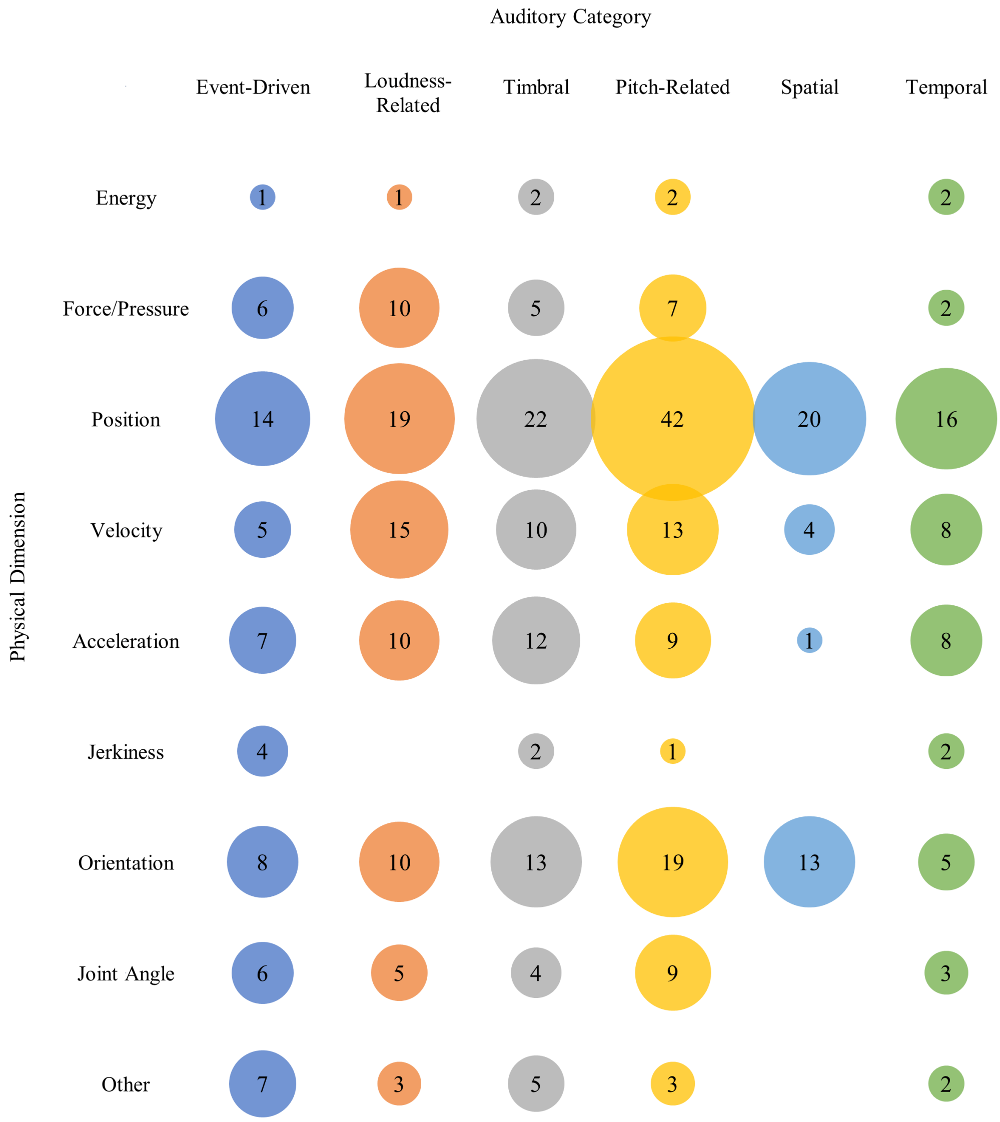

Figure 3.3: Bubble plot visualisation showing the mapping relationship between Physical Dimension keywords and Auditory Category keywords in all projects. Number displayed shows the number of different projects containing that mapping, with bubble plot size proportional to number shown in the centre of each bubble.

Auditory Category list, Spatial shown to be the least recorded with 38. Cumulatively Position, Velocity, Acceleration, Orientation (PVAO) amount to 76.32% of the Physical Dimension keywords recorded in the database.

Part(s) of the Body that are Tracked

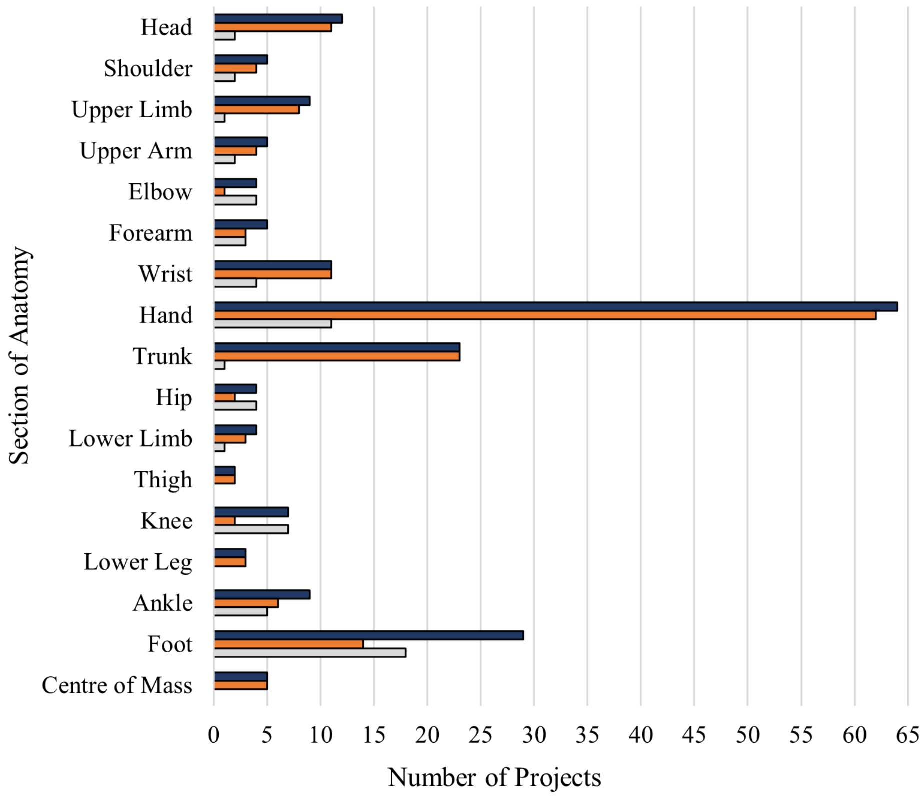

Figure 3.4: Bar chart containing allocated 'Anatomy' keywords for each project: i) data visualised in blue corresponds to all recorded anatomy keywords, ii) data visualised in orange shows all anatomy entries where at least one of Position, Velocity, Acceleration or Orientation was obtained from that anatomy, iii) data visualised in grey shows all anatomy entries where a Physical Dimension alternative to Position, Velocity, Acceleration or Orientation was obtained from that anatomy.

Keywords entered in the Anatomy list were analysed independently and in combination with the popularly used Physical Dimension identified. Figure 3.4 presents three groups of data for this analysis, i) all recorded anatomy keywords, ii) all recorded anatomy keywords with at least one PVAO Physical Dimension, iii) all recorded anatomy keywords with at least one Physical dimension outside of PVAO. Data i) shows all recorded Anatomy keywords consisting of 201 entries from the list of projects. The keyword Hand was recorded the most with 64 entries, which calculates to 44.13% of all projects. Other frequently recorded keywords in this dataset are: Foot with 29 entries calculating to 20.00% of all projects, Trunk with 23 entries calculating to 15.86% of all projects, Head with 12 entries calculating to 8.28% of all projects, and Wrist with 11 entries with 7.59% of all projects. The remaining keywords in this section were each recorded in less than 10 projects. Data ii) shows Anatomy keywords where at least one PVAO Physical Dimension was obtained, amounting to 164 entries from the list of projects. The keyword Hand was recorded the most with 62 entries, which calculates to 42.76% of all projects. Other frequently recorded keywords in this dataset are: Trunk with 23 entries calculating to 15.86% of all projects, Foot with 14 entries calculating to 9.66% of all projects, Head with 11 entries calculating to 7.59% of all projects, and Wrist with 11 entries calculating to 7.59% of all projects. The remaining keywords in this dataset were each recorded in less than 10 projects. Data iii) show Anatomy keywords where at least one Physical Dimension outside of PVAO was obtained from it, amounting to 65 entries from the list of projects. The keyword Foot was recorded the most with 18 entries, which calculates to 12.41% of all projects. Other frequently recorded keywords in this dataset are: Hand with 11 entries calculating to 7.59% of all projects, Knee with seven entries calculating to 4.82% of all projects, Ankle with five entries calculating to 3.45% of all projects. The remaining keywords in this dataset were each recorded in less than five projects.

Types of Tracking Technology

As described in Section 3.2.3 each technology type was classified to three Technology Categories and the analysed results are presented in Figure 3.5, showing 173 entries overall. The figures in this section have been colour coded to represent the technology category assigned. The Inertial Sensor category shown in red contains 59 entries with Accelerometer the most frequently recorded keyword in this category with 19 entries. The Camera category shown in orange contains 58 entries with Marker-Based Motion

Table 3.2: Table detailing the remaining contents of the tracking technology that are not presented in Figure 3.5. Each of the presented technology in this table contain less than five recorded entries and are assorted depending on their Technology Category.

Remaining Inertial Sensors		**Remaining Other**					
Gyroscope	2	Graphic Tablet	4	MMG	1	Tension-Activated Switch	1
Remaining Camera		Microphone	4	Ultrasonic Sensor	1	Electrical Contacts	1
Virtual Reality Controller	1	Rotary Encoder	2	Variable-Resistance Elastic	1	Textile Stretch Sensor	1
		Haptic Device	2	Bend Sensor	3	Piezoelectric Transducer Pickup	1
Infra-Red	1	Ergometer	3	Cadence Sensor	1	Infra-Red Proximity Sensor	1
Gaming Controller (Ca)	2	Goniometer	4	Electromagnetic Tracker	3	Footswitch	1
		Tendon-Based Parallel Robot	1	Gaming Balance Board	1	Speed Sensor	1

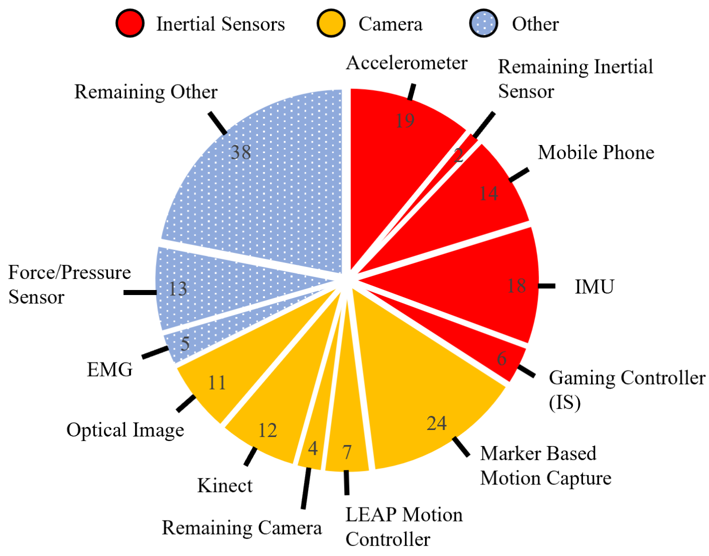

Figure 3.5: Pie chart visualisation showing all tracking technology keyword recorded in the dataset. Data visualised in red corresponds to Technology keywords categorised in Inertial Sensors. Data visualised in yellow corresponds to Technology keywords categorised in Camera. Data visualised in blue corresponds to Technology keywords categorised in Other. For Figure 3.5 all categories with 'Remaining' are detailed in Table 3.2.

Capture the most frequently recorded keyword in this category with 23 entries. The Other category shown in blue contains 56 entries with Force/Pressure Sensor keyword the most frequently recorded keyword in this category with 13 entries. All technology entries that are recorded in less than five projects have been grouped depending on their assigned category and represented by a 'Remaining Inertial Sensor', 'Remaining Camera', or 'Remaining Other' segment, each keyword grouped in this way is detailed in Table 3.2.

Technology categories have also been analysed in combination with the Anatomy keywords that contained more than 10 entries: Hand, Head, Trunk, Wrist and, Foot (Figure 3.6(a), (b), (c), (d), (e), respectively). These figures have each been colour coded in an identical manner as 3.5, with the same key. Figure 3.6(a) also contains 'Remaining Inertial Sensor', 'Remaining Camera' and 'Remaining Other' segments to group together technologies that have been recorded once, these segments have been expanded in Table 3.3.

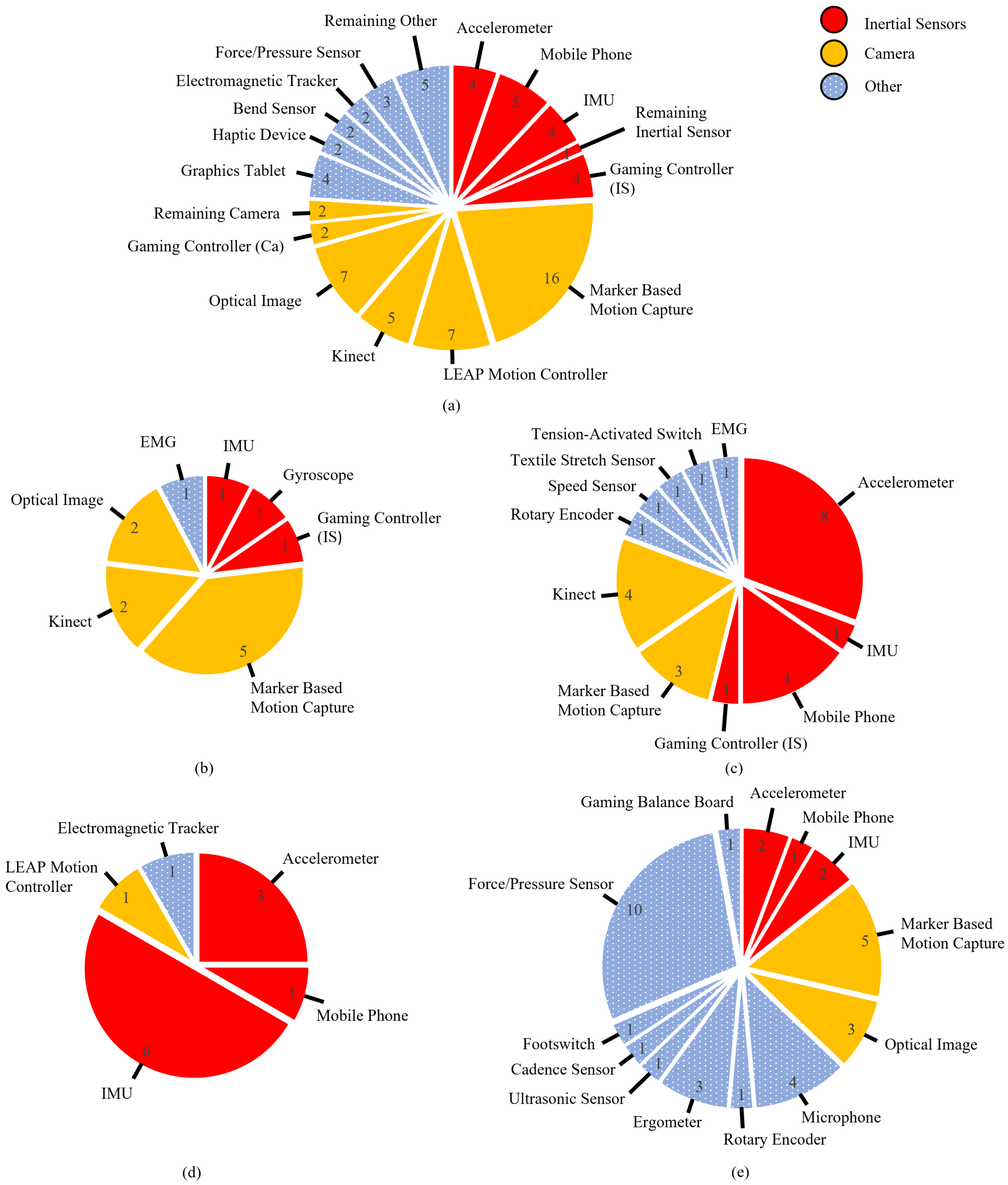

Figure 3.6: Pie chart visualisation showing the tracking technology keywords recorded in the database, filtered to show popularly used tracked anatomy as follows: (a) Hand, (b) Head, (c) Trunk, (d) Wrist, (e) Foot. Data visualised in red corresponds to Technology keywords categorised in Inertial Sensors. Data visualised in yellow corresponds to Technology keywords categorised in Camera. Data visualised in blue corresponds to Technology keywords categorised in Other. For (a), all categories with 'Remaining' are detailed in Table 3.3.

Table 3.3: Table detailing the remaining contents of the tracking technology that are not presented in Figure 3.6(a). Each of the presented technology in this table contain one recorded entry and are assorted depending on their Technology Category

Remaining Inertial Sensors	
Gyroscope	1
Remaining Camera	
Infra-Red	1
Virtual Reality Controller	1
Remaining Other	
Tension-Based Parallel Robot	1
Electrical Contacts	1
Piezoelectric Transducer Pickup	1
Infra-Red Proximity Sensor	1
Rotary Encoder	1

Figure 3.6(a) shows 75 recorded entries, from 23 different technology types split into 17 segments. The Camera category is the most recorded Technology Category with 42 entries calculating to 56.00% of all entries involving the Hand, and the Marker-Based Motion Capture keyword is the most recorded Technology keyword with 15 entries calculating to 20.00% of all entries involving the Hand. Technology keywords associated with Hand also have the highest number of different keywords for each Technology Category recorded. Figure 3.6(b), shows 13 recorded entries from seven different technology types used to monitor the Head. The Camera category is the most recorded Technology Category with nine entries calculating to 69.23% of all entries involving the Head, and the Marker-Based Motion Capture keyword is the most recorded Technology keyword with five entries calculating to 38.46% of all entries involving the Head. Figure 3.6(c), shows 26 recorded entries from 11 different types of technology used to monitor the Trunk. The Inertial Sensor category is the most recorded Technology Category with 14 entries calculating to 53.85% of all entries involving the Trunk, and the Accelerometer keyword is the most recorded Technology keyword with eight entries calculating to 30.77% of all entries involving the Trunk. Figure 3.6(d), shows 12 recorded entries from five different technology types that monitor the Wrist. The Inertial Sensor category is the most recorded Technology Category with 10 entries calculating to 83.33% of all entries involving the Wrist, and the IMU keyword is the

most recorded Technology keyword with six entries calculating to 50.00% of all entries involving the Wrist. Figure 3.6(e), shows a total of 51 entries from 13 different types of technology that are used to monitor the Foot. The Other category is the most recorded Technology Category with 22 entries calculating to 43.14% of all entries involving the Foot, and the Force/Pressure Sensor keyword is the most recorded Technology keyword with 10 entries calculating to 19.60% of all entries involving the Foot.

3.4 Discussion

The results highlight the diversity of components chosen for real-time movement sonification systems in the literature. Each project was keyword-coded based on the system(s) detailed in the composing article(s), and the resulting keyword database was synthesised to produce visual displays in support of addressing the research topics stated in this chapter. Disaggregation of each movement sonification system to the three principal components (motion tracking technology, anatomy and sonification) allows the identification of the components that are most popular in the literature. The present discussion looks to view the identified system components in terms of motion tracking and sonification configuration and provide perspective on appropriateness to movement rehabilitation.

3.4.1 Motion Capture Technology

Existing rehabilitation projects identified in this review have made use of a variety of technologies to monitor PVAO of a tracked segment of anatomy (RQ1). For each technology type, a perspective on the acquisition costs is included where possible, with approximate price of <\$100 labelled as 'low cost', \$100 - \$500 labelled as 'moderate cost', and >\$500 labelled as 'high cost'. Technologies that have an undisclosed acquisition cost, have no label assigned. Due to the categorisation of technologies in this manuscript, there are ranges of costs for most of these technology categories, as dependent on the requirements and capabilities of the products in the category, as such these labels are intended as guides when considering costs of each technology. Similarly, accuracy is an important characteristic to consider when selecting a motion capture system as an input to a real-time movement sonification system, and varies depending on multiple aspects including: technology type, number of units, intended application, choice of kinematics and, capture frequency. Due to this variability, generalising the accuracy of each technology would be inappropriate and is absent from this discussion. Further research into accuracy requirements and competency is strongly recommended before selecting a technology, with information available in existing reviews, such as in Van Der Kruk *et al.* [105]. Finally, although the use of a computer mouse, computer keyboard and, touchscreen technologies have commonly been used as input interfaces for commercialised devices and therefore, are accessible low-cost motion tracking technologies, the use of these devices come with limited tracking volume, and are not considered applicable for functional training, as such these technologies have been excluded from this review.

Inertial sensors embedded in mobile phones used in Spina *et al.* [106] were attached to the wrist and ankle of users to monitor clinical routines, and in Stahl *et al.* [107] to track wrist flexion-extension and radial-ulnar movements. Multiple IMUs were utilised by Bruckner *et al.* [108, 109], Brock *et al.* [93], Schmitz *et al.* [92, 94] and Scholz *et al.* [95] to monitor an upper limb whilst performing task-oriented movements, whereas a single IMU was used in Bevilacqua *et al.* [110] to monitor reaching movements. The use of inertial sensor technologies, such as the IMU and the sensors integrated in a mobile device, allows for a technology that is versatile in tracking gross movement for a low cost, however, for the application of a movement sonification system the data can require extensive filtering and manipulation to obtain the desired physical dimension.

Repurposed gaming controllers are used widely motion capture devices for entertainment-alternative applications, with the acquisition of these technologies available at a low cost being a key reason. Examples include the Wiimote as applied in a rehabilitative context by Alankus *et al.* [111] to detect shoulder abduction/adduction and compensatory trunk movements; used outside of a rehabilitative context by Dotov *et al.* [112], and Seko *et al.* [113]. Other examples can be seen through the Wii Balance Board as demonstrated by Feltham *et al.* [114]; the PlayStation Move motion controllers as demonstrated by Tanaka *et al.* [115]; the Microsoft Kinect as demonstrated by Hebling *et al.* [116]. However, with exception to the PlayStation Move Controllers, all the gaming controllers listed above have been discontinued. The Microsoft Kinect for Xbox on the other hand has a successor named 'Microsoft Azure Kinect SDK' which is available for purchase, and could be used in future motion tracking applications, however as noted in Tölgyessy *et al.* [117], there are limitations with using this technology, including object reflectivity issues, and degraded performance in outdoor environments. In terms of a motion tracking solution, the minimum requirement to use the Azure Kinect body tracking on a Windows PC are as follows: Seventh Gen Intel Core™ i5 Processor (Quad Core 2.4 GHz or faster), 4 GB Memory, NVIDIA GEFORCE GTX 1050 or equivalent, Dedicated USB3 port. A host device with these requirements along with the device itself, leads to a high cost for this technology as a rehabilitation commodity to be sold to the public.

Alternative projects that make use of multiple motion capture technologies include Ghisio *et al.* [118] who used an accelerometer for a synchronous task and a Microsoft Kinect to track sitting posture; Cibrian *et al.* [119], Singh *et al.* [120, 121], Newbold *et al.* [122] developed a framework referred as 'Go-with-the-flow' which used embedded inertial sensors in a mobile phone, or the Microsoft Kinect, to monitor the trunk and upper limb.

Marker-based motion capture systems, is the term selected in this manuscript to represent optical motion capture systems that track retroreflective markers attached to target locations. Although this type of technology is considered the gold-standard of motion capture [123] the system comes at a high cost, requires dedicated space, calibration time and trained personnel to maximise the capabilities of this technology. Examples of use in a rehabilitative context can be seen through: Wallis *et al.* [88] and Chen *et al.* [124], to capture reaching and grasping movements; Vogt *et al.* [125] to capture the upper-body during rehabilitation exercises; Dailly *et al.* [126] to capture hand movements during a figure tracing task.

A LEAP motion controller was utilised by Nikmaram *et al.* [96], and included in a system termed 'SonicHand' by Colombo *et al.* [98], to track hand and wrist movements.

This low cost technology is designed to track the hand of a user, within the Field of View (FOV) of the camera. However as noted by Gamboa *et al.* [127], the limited FOV, dependency on environmental conditions, and performance with objects in FOV, are limitations with using this technology for home-based rehabilitative applications.

Motion capture systems that make use of an electromagnetic field and attachable sensors (that act as markers) have been labelled as an electromagnetic tracker in this manuscript, as shown in Maulucci *et al.* [87] and Robertson *et al.* [91] to monitor reaching movements, such systems are able to monitor the position and orientation of each sensor. However, the resolution of the system is distance dependent from the field source, which restricts the appropriate range of operation for motion capture.

Haptic devices, such as the SensAble PHANToM Desktop haptic device, a computer periphery device that operates by the user moving a stylus attached to a robotic arm, have also been used as a motion capture device. Usually applications with haptic devices will only focus on their haptic feedback capabilities, however Frid *et al.* [128], and Rodriguez *et al.* [129], incorporated additional audio feedback using the motion capture capabilities of the device, to create a multimodal system for their projects.

A graphics tablet, although conventionally used for drawing applications, was used as a motion capture system as part of a writing rehabilitation task by Véron-Delor *et al.* [130], as such capturing the movements of the hand on a 2D-plane, albeit in a limited range of space. The technology is available at low cost, but has a large range of cost as dependent on size of working area, resolution, and quality of product.

Other motion capture technologies have been created as wearable systems for rehabilitative applications, including a garment integrated with stretch sensors was created by Ten Bhömer *et al.* [131] to monitor the upper body during rehabilitative exercises; a bespoke glove with integrated electrical contacts was created in Friedman *et al.* [97, 132] to detect connection between the thumb and specific hand locations in a rhythmical serious game. Both systems show the potential and limitations of wearable systems, with the garment allowing motion capture of the entire upper body with a single item but creating difficulty for a hemiparetic user (who would struggle to clothe) in using the item. In contrast the bespoke glove would be easier to clothe and use, however the motion capture would be restricted to hand movements and postures.

Outside of rehabilitation, existing projects have made use of alternative off-the-shelf technologies to capture human movement for their systems. Virtual Reality (VR) systems and the handheld controllers associated with them are one example. Johnson *et al.* [133] shows an example of a virtual reality sonification system, tested with the Samsung HMD Odyssey Windows Mixed Reality Headset. The VR market is an emerging competitive market, as companies look to provide entertainment experiences through these systems, as such off-the-shelf systems vary in price from moderate to high cost, depending on the desired capabilities and specification of the system. VR either with associated controllers or in combination with a LEAP motion controller could provide an effective environment for real-time audiovisual feedback, and as technology in this area is advancing in quality, with a healthy competitive market, leads to a promising motion capture system for upper limb rehabilitation.

The use of Red, Green, Blue (RGB) camera-based devices, labelled in this manuscript as Optical Image, is an established means of capturing images, however the use of these images as a means of motion tracking is of interest in this review. As observed from the projects identified in search list, there are two methods of using this technology, one is using a mobile camera to track an anatomy (typically the hand holding the camera) in relation to an observable fixed reference (example shown in Ahmetovic *et al.* [134]) , and the other method is with a fixed camera tracking a mobile section of anatomy (example shown in Ramsay *et al.* [135]). The use of this system is observed in many applications including visual impairment [134, 136], sport [135, 137], performing arts [138, 139], gait rehabilitation [140], immersive environments [141], task performance [142], or for other purposes [143, 144]. Likewise smart phones typically contain an RGB camera as standard, providing an accessible means of capturing movement, available at low cost. However, the performance of this technology is dependent on the environment. As the technology market is a competitive market with a range of specifications for desired capabilities and costs, the development and use of these systems show promise for rehabilitative applications.

An ergometer, such as an exercise bike or an indoor rower, although limits the actions of a user to specific activity-dependent movement, are popularly used as exercise equipment. Although other projects make use of an ergometer in their project, only Sigrist *et al.* [145], Schaffert *et al.* [146], O'Brien *et al.* [147] have used the technology for motion capture system in their real-time movement sonification system. The technology is widely available for purchase, however, the systems are unportable and range from moderate to high cost for acquisition.

Rotary encoders, used to determine angular position of a rotational shaft, were applied to a cycling task [148] and a rowing task [50]. Within the projects identified in this review, these encoders are low to moderate cost attachments to existing ergometers, however other projects outside of the remit of this review have made use of encoders as part of robot-assisted lower extremity exoskeleton [149], as such there is vindication of using this technology as a means of capturing movement, but this requires additional integrated components for a usable system.

Goniometers are instruments that when applied to human biomechanics context, are used to measure joint angles. In their primitive analogue form, goniometers are low cost and accessible instruments, but are inadequate for real-time monitoring. Hermann *et al.* [150] demonstrated a setup utilising potentiometers as goniometers to create a real-time system, this type of technology is otherwise known as an electrogoniometer, which are commercially available, however this option comes at a high cost. Examples of use can be seen through Hale *et al.* [151] and Fujii *et al.* [152].

The use of a microphone, as a method of obtaining sound from the foot-ground interaction, has been used in many projects as a means of an input stream for a movement sonification system, with most recent examples including Gomez-Andres *et al.* [34], Tajadura-Jiménez *et al.* [153], Turchet [154], Maculewicz *et al.* [155], all for walking purposes, and Pugliese *et al.* [156] as part of a trampoline sonification system. The use of this technology for motion capture, although innovative in providing motion capture of the foot-ground interaction, would only be applicable for highly specific applications. Although this technology is available at low cost, the use of microphones attributing to moderate to high costs have generally been used.

Force/Pressure sensors are commonly used as motion capture devices in the literature, although none have been recorded for use in an upper body rehabilitative context, examples of use can be found with performing arts [157] [158], to affect body perception [159–161], monitor cycling [146], [147], monitor skiing [162], gait rehabilitative purposes [163–166], sports application [50, 167, 168], or with use as an interface [169]. As the sensor requires compression to change electrical resistance, human motion capture is limited to interaction with a surface, however, due to the low cost and high environmental versatility of the sensor, this remains a popular sensor type for motion capture. As shown in Section 3.3, the use of force or pressure sensors are especially popular in combination with motion tracking of the foot, or feet, of a person.

Ultrasonic sensors utilise ultrasonic waves as a method of measuring distance an example of use can be seen with Akiyama *et al.* [170] to detect foot elevation from the floor whilst walking. Similarly with force/pressure sensors, this technology is considered low cost and versatile, however the application limitations differ as ultrasonic sensors require distance from a perpendicular surface to be utilised effectively.

Bend sensors, otherwise known as flex sensors, are variable resistors with flex-dependent resistance. Projects that use such sensors have applied them to detect postures of the hand [171, 172], and to detect joint angle around the elbow [173]. Although the sensors are low cost, and versatile, multiple flex sensors are required per joint to capture movement in multiple axes.

Electromyography and Magnetomyography (EMG and MMG respectively) are instruments used to detect muscle activation by monitoring the neural signals sent to that muscle. Researchers that use such technologies for motion capture generally use surface electromyography (sEMG) allowing for safer monitoring of muscle activation, these generally have high cost. As many sensors are required to monitor many synergistic or antagonistic muscle groups, and extensive signal processing is required for each sensor, this limits the appropriateness of using such a technology type for complex movements. However, examples of use can be seen with Nakayama *et al.* [174] in a facial expression sonification project and with Donnarumma *et al.* [175].

Other technologies have been applied to the projects identified, these are generally considered to be very specific to the application of the movement sonification system. A tendon-based parallel robot was developed and used in a rowing task [176]. A cadence sensor was used in a cycling motivational investigation [177]. Multiple infra-red proximity sensors were used to capture hand movements in a specific 3D volume [178]. A piezoelectric transducer pickup was used in a 'sonic interactive surface' [179]. There are also recorded projects that made use of switches [180, 181], variable-resistance elastic [182], or a speed sensor [183, 184].

From the existing motion capture systems identified in the literature, several potential technologies could be utilised for motion capture purposes in a home-based stroke rehabilitative context. Inertial sensors are widely accessible with low cost and if raw acceleration or angular velocity are appropriately utilised, these devices provide an excellent candidate technology. However, metrics such as gravity removed linear acceleration and orientation require additional data fusion between the measurands. If these are to be further processed to obtain velocity or position, integration and drift errors accumulate requiring additional calibration, anchoring or use of additional devices that increase the cost and difficulties with setup. The use of camera technologies such as Azure Kinect or LEAP motion has potential, especially with the capability of measuring position leading to greater flexibility in desired physical dimension, however the cons of high cost and environmental dependence could demotivate users. Whereas other technologies have various pros and cons that generally make them a good option depending on the intended application, but not for others. There also remains a possibility to combine the capabilities of multiple motion capture devices to obtain a synergistically superior system. One such example could be through combining portable sensors to an ergometer, allowing for multimodal bilateral training, that is not only available for home use, but could be taken to a gym, or physiotherapy session. Overall, the diversity in motion capture technology chosen in the literature is justified as an ideal motion capture system is still absent.

3.4.2 Sonification Mapping

As mentioned in Section 3.1, existing reviews from Ghai *et al.* [185] and Guerra *et al.* [86] have overall come to positive conclusions with regards to utilising auditory techniques for rehabilitative purposes, however there is minimal spotlight on the sonification configurations utilised in the reviewed articles. Although some studies have been conducted evaluating sonification mapping choices [186, 187] evidence-based guidelines for real-time movement sonification mapping are currently absent. As such creating an effective movement sonification system is likely to require a more trial-and-error iterative approach, as opposed to an efficient systematic approach. The results from this review provides information on the available (or lack of) choices in the existing literature and motivate future system creators to select, test, and compare their system with those identified in the literature, and therefore, future work can provide evidence on the efficacy of these sound configurations for upper limb stroke rehabilitation.

From the existing rehabilitation projects, a wide range of sonification options have been utilised (RQ2). The following projects contain position, the most chosen physical dimension, in simple sonification designs. Maulucci *et al.* [87] and Ten Bhömer *et al.* [131] linked position to pitch. Robertson *et al.* [91] linked position to loudness and orientation to stereo panning. Alankus *et al.* [111] linked position to an audio sample, and orientation to loudness and to trigger an audio sample. Spina *et al.* [106] linked position and velocity to trigger an audio sample. Dailly *et al.* [126] linked position to the addition of noise. Friedman *et al.* [97, 132] linked position to melody, and loudness. Ghisio *et al.* [188] linked position to timbre and loudness, and acceleration to polyphonic content. Bruckner *et al.* [108, 109] linked position to loudness, pitch, and stereo panning, and linked velocity to loudness. Schmitz *et al.* [92, 94], and Brock *et al.* [93] linked position to pitch, stereo panning, and brightness, and linked velocity to loudness. Bevilacqua *et al.* [110] linked position to pitch, tempo, and melody. Nikmaram *et al.* [96], Scholz *et al.* [95] [189] linked position to pitch, brightness, instrumentation, and loudness.

Sonification designs that do not use position appear in this area as well, Rodriguez *et al.* [129] linked velocity to pitch, Véron-Delor *et al.* [130] linked velocity to melody, and force generated is linked to loudness. Stahl *et al.* [107] linked velocity to loudness and tempo, and linked orientation to pitch and timbre.

Other rehabilitation projects include a range of sonification options that are implemented in their system, Cibrian *et al.* [119], Singh *et al.* [120,121], and Newbold *et al.* [122] contains 14 different mappings, Colombo *et al.* [98] contains six different mappings, Vogt *et al.* [125] contains five different mappings, Wallis *et al.* [88] and Chen *et al.* [124] contains seven different mappings. Details for each project are listed in the database attached in the appendix. The use of position as a physical dimension mapping option, and a pitch-related auditory mapping option are predominantly chosen in the literature, either in combination or with other types of mappings.

Velocity as part of a sonification mapping is also favoured, especially in combination with the following auditory categories: loudness as demonstrated through Ghai *et al.* [185,190], Frid *et al.* [191] and Hermann *et al.* [150]; timbral as demonstrated through O'Brien *et al.* [192, 193], Boyer *et al.* [194], Dyer *et al.* [82, 195]; pitch as demonstrated by Jakus *et al.* [196], Gref *et al.* [197], Wang *et al.* [198]. Likewise, orientation is a preferred physical dimension, and is used most in combination with the following auditory categories: timbral as demonstrated by Ikeda *et al.* [199], Tanaka *et al.* [115], Schlegel *et al.* [200]; pitch as demonstrated by Ley-Flores *et al.* [159], Dotov *et al.* [112], Volta *et al.* [201], [202]; spatial as demonstrated by Russell *et al.* [203], Avissar *et al.* [204], Franco *et al.* [205]. Finally, acceleration is generally chosen in combination with the following auditory categories: loudness as demonstrated by Salter *et al.* [206], Baalman *et al.* [207]; timbral as demonstrated by Brazauskayte [208], Lorenzoni *et al.* [209, 210], Burloiu *et al.* [211, 212], Giomi *et al.* [213]; pitch as demonstrated by Schaffert *et al.* [214], Chen *et al.* [215], Wood *et al.* [216].

Outside of the PVAO physical dimensions, data obtained as force or pressure, as dictated by the use of a force or pressure sensor, has been combined with sampled sounds [217], and changes in loudness [158], timbral [114], pitch [145] and, temporal-related [147] auditory feedback. There are also projects that make use of multiple mappings that include force or pressure as a physical dimension, examples include

Bisig *et al.* [157], Gorgas *et al.* [163], Horsak *et al.* [164], Fischer *et al.* [165], Cesarini *et al.* [167, 168].

The angle difference calculated around a joint has been used as an input physical dimension for movement sonification purposes, examples can be seen in combination with sampled sounds [218], and changes in loudness [151], timbral [198], pitch [219] and, temporal-related [220] auditory changes.

Alternative physical dimensions used as the input dimension to sonification mappings have relatively low numbers in comparison to the aforementioned sections. Jerkiness, is calculated and used within six projects [221–231]. Energy is calculated and used within three projects [226–228], [232–234]. Categorised in the 'Other' keyword category are alternative physical dimensions that are highly specific to the application that the system is developed for. These include contact with a surface [34, 153, 235], electromyography signals [174], [236], magnetomyography signals [175], and facial expressions [139].

The choice of physical dimension and auditory dimension to sonify seems to impact the effectiveness of audio feedback, however as of writing, insufficient evidence is available on which combinations provide the most effective results for motor learning, or movement rehabilitation. Based on the results of this review, established combinations of physical dimensions to auditory dimensions can be identified and brought forward for direct comparison studies.

3.5 Updated Search

A subsequent search following the same procedure as described in Section 3.2 was performed on the 11th of September 2023. From the 340 articles obtained in this search, 23 articles [99, 237–258], were identified as suitable through the set inclusion/exclusion criteria, corresponding to the addition of 19 projects and an extension of one project [98], to the results obtained through the previous search. The added projects were processed independently to the data presented in the 2021 search to identify motion capture technologies were used, identify present trends of anatomy segments tracked, and identify sonification configuration mappings used.

50 sonification mapping types were identified in these additional projects. Position (14) and orientation (13) were the most popularly chosen physical dimensions. The auditory categories used were generally split between Event-Driven (11), Pitch-Related (11), Loudness-Related (10), and Temporal (10) audio properties used as part of the sonification configurations in these projects. For sections of anatomy tracked, the most common were Hand (7) and Foot (5), with all occurrences for hand tracking extracting either PVAO, whereas for foot tracking half of the projects did so. For motion tracking technologies chosen, the most common were through RGB camera images, labelled as 'Optical Image' (7) in this work. In addition, emerging technologies from the application of VR, are becoming more prevalent as movement tracking technologies in movement sonification applications, with the inclusion of VR headsets and trackers as part of the created systems.

3.6 Concluding Remarks

This chapter presents a range of available options for the use of movement sonification. Of interest to this thesis are the motion capture technologies identified. Results of the analysis shown in this chapter show 38 technology types that have been used as part of existing systems found in the literature. Filtering of the technology types identified, based on suitability for upper limb motion capture, cost, potential for concurrent sonification, and versatility for application, has led to a number of potential technologies to consider as part of a new system. These include inertial sensor technologies, optical camera technologies, LEAP Motion controller, Kinect camera, goniometer, VR systems, and bend sensors (RQ3). Considering the desire for self-administering therapy for rehabilitation, and the potential for fast development to an initial proof-of-concept system, the choices of inertial sensors and the Kinect camera were brought forward for development.

Chapter 4

Movement Sonification System Development

The aim is to create a prototype proof-of-concept real-time movement sonification system that can later be developed into an accessible, low-cost system, suitable for home environments, with flexibility in training methodologies, minimal setup, and with feasibility for self-initiation and self-operation of the system. Explicitly, the movement sonification system to be developed would have the following specification in mind:

- Capability to sonify data pertaining to physical position, velocity, acceleration, and orientation;
- Capability to obtain multi-axis movement kinematics;
- Capability to monitor movements of either upper limb;
- Acquisition cost of less than £200;
- Environmental versatility in terms of illumination and location;
- Straightforward setup and activation to enable self-directed practice or guided practice with trained personnel;
- Options to edit sonification aesthetics;

- Capability to create auditory displays within real-time audiomotor perception of the user;
- Capability to create auditory displays that users can associate with their movements.

A movement sonification system consists of two principal components; a motion capture system, and a sonification module, of which the motion capture output is directed to a sonification module. For the latter component, multiple design choices are available to a system designer, of which the previous chapter provides commonly used options for consideration, such as the use of physical position, and audio pitch. For the former component, this chapter proposes two individual solutions (selected as the result of the review presented in Chapter 3): the first system termed ‘Soniccup’ relies on Inertial Navigation System (INS) technology, the inclusion of a push-to-make switch, and a data processing methodology based on the Zero Velocity Update (ZUPT) method, whilst the second system referred to as ’KinectSon’, utilises markerless Computer Vision (CV) technology. The position estimates from the two motion capture systems are used to produce audio in the sonification system. Finally, the results of the two systems are compared with a Vicon Nexus, a gold standard marker-based motion capture system, which has been excluded as a candidate outside of laboratory use. The remainder of the chapter provides background on the processing of INS, the methodology adopted and processing used to obtain position.

To begin, movement sonification system development as described in this chapter focuses on the following specification:

- Capability to monitor movements of either upper limb;
- Capability to sonify data pertaining to physical position, along one axis of movement;
- Capability to create auditory display within real-time audiomotor perception of the user, set to below 100 ms [103];

- Capability to create auditory displays that users can associate with their reaching forward movements.

This list is not exclusive to the technology used to acquire movement kinematics, and therefore, this list will be considered throughout the work described in this chapter.

4.1 Motion Capture through an Inertial Navigation System

Inertial sensors refer to two types of sensor, 3-dimensional accelerometers to measure linear acceleration, and 3-dimensional gyroscopes to measure angular velocity. Both are ubiquitous in present-day technologies in the form of Microelectromechanical Systems (MEMS) type of devices, made accessible through bulk manufacturing resulting in low cost sensors. As such, these sensors are highly accessible and potentially suitable for movement rehabilitation purposes, given that they have been developed to be small and lightweight. However, the associated errors with these devices require additional consideration. The miniature size of the MEMS devices makes them susceptible to environmental conditions [259], which can drastically affect output. Noise sources that contribute to the performance of the sensor can be categorised based on their energy type, i.e. mechanical, electrical, coupling; and to their noise source type, i.e. extrinsic (packaging and mounting) and intrinsic (physical interaction of the sensors). Table 4.1 shows the major types of MEMS noises and how they can be categorised [2].

Table 4.1: Noises in MEMS sensors as described by Bhardwaj *et al.* [2]

Energy	**Noise Source**	**Noise Type**
Mechanical	Intrinsic	Brownian motion
	Extrinsic	Sinusoidal noise
		High frequency noise
Electrical	Intrinsic	Shot noise
		Johnson noise
		Flicker noise
Coupling	Intrinsic	Thermal-mechanical noise

Inertial sensors have additional noise components associated with the accelerometer and gyroscope, as shown in Table 4.2. Even with mitigation techniques such as calibration, careful fabrication and construction to minimise intrinsic errors, errors can not be eliminated completely. Improving measurement accuracy can be achieved through statistical model correction or/and through sensor fusion to minimise the effect of errors [260].

Table 4.2: Inertial Sensor errors as described by Bhardwaj *et al.* [2]

Device	**Description**
Gyroscope	g-dependent bias coefficient
	Anisoelastic coefficient
	Anisoinertia coefficient
	Cross coupling coefficient
	In-run random bias
Accelerometer	Scale factor error
	Cross coupling coefficient
	Measurement bias
	Vibro-pendulous coefficient
	In-run random bias

4.1.1 Sensor Fusion

Typically, inertial sensors are packaged together with additional sensors such as a 3-dimensional magnetometer, and temperature sensor. Furthermore, they include digital conversion and data transfer capabilities in a single-package device known as an Inertial Measurement Unit (IMU). The accelerometer measures the linear acceleration experienced by an object, including the gravitational force of the Earth (SI unit of meters per second squared). The gyroscope measures angular velocity (SI unit of radians per second). The magnetometer measures the magnetic field strength and direction. Appropriately combining the measurements of the accelerometer, gyroscope and magnetometer permits orientation measurements with higher accuracy. Orientation refers to an arrangement of points after an applied transformation, and is generally represented through Euler angles, rotation matrices, or quaternions. Obtaining the orientation of the sensors permits the removal of the gravitational contributions to the measured linear acceleration. Through the sensors in an IMU, there are two methodologies to obtain orientation. One method is through integrating the readings of the gyroscope, obtaining relative orientation to a reference. However, due to gyroscope noise and biases, the integration leads to a long-term error in orientation estimation, and requires correction. An alternative method is through the combination of accelerometer and magnetometer data. As the accelerometer obtains the gravitational force of the Earth, and the magnitude is constant and known, this can be used to obtain a reference direction pointing down. The magnetometer is used to identify the direction of the north magnetic pole, assuming the absence of an alternative dominant magnetic field, as such creating a reference direction pointing north. These two reference directions in the Earth coordinate frame (with the third reference direction orthogonal to these two directions, with respect to a right-hand coordinate system), permits the computation of the orientation defined by:

$$\phi = \tan(a_y/a_z)^{-1} \tag{4.1}$$

$$\theta = \tan(-a_x/\sqrt{{a_y}^2 + a_z^2})^{-1} \tag{4.2}$$

$$\psi = \tan(-m_x/m_y)^{-1} + D \tag{4.3}$$

where ϕ represents the roll angle, θ represents the pitch angle, ψ represents the heading (see Figure 4.1 for reference), a_x, a_y, a_z represents the three orthogonal measurements from the accelerometer in the sensor coordinate frame, m_x, m_y represents orthogonal measurements from the magnetometer in the sensor coordinate frame (note that m_z is not required to calculate the heading angle), and D represents the declination angle, which incorporates the tilt of the Earth into heading calculation [261]. However, short-term perturbations in the magnetic field as typical in industrial environments could also lead to erroneous readings, and therefore, orientation calculated through this method typically have worse short-term accuracy than orientation calculated through the integration of gyroscope measurements. The solution therefore, is to combine the short-term performance of the gyroscope methodology, with the long-term performance of the combined accelerometer and magnetometer performance.

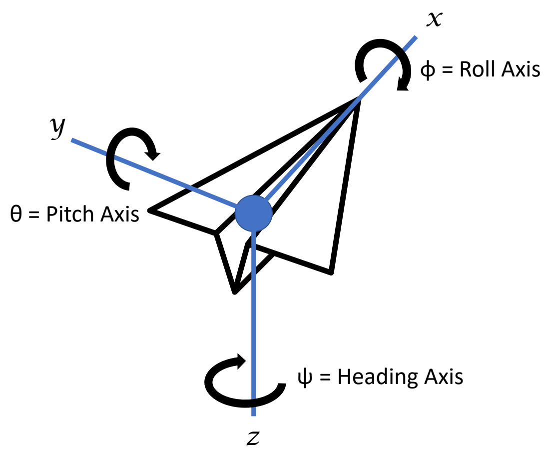

Figure 4.1: Figure illustrating six degrees of freedom on a rigid body in 3-dimensional space. The axis labelled x represents the forward/backward direction, the axis labelled y represents the left/right axis, and the axis labelled z represents the up/down axis.

An equation to obtain orientation through inertial sensor measurements is to combine the independent orientation measures from the gyroscope, and the magnetometer and accelerometer, through a simple weight mean [262] as shown in Equation 4.4,

$$q_{fused}(k) = \gamma \cdot q_g(k) + (1 - \gamma) \cdot q_{a/m}(k) \tag{4.4}$$

where γ is the weight parameter, q_g is the orientation obtained from the gyroscope data in quaternion form, at time k, and $q_{a/m}$ is the orientation obtained from the combination of acceleration and magnetometer data, in quaternion form.

Further improvements to the orientation estimation can be obtained using Kalman filters [263], complementary filters [264], or variations of each [265, 266]. Kalman filters are recursive estimators that work through iterative prediction and correction phases. During the prediction phase, the filter will use initial state conditions, previous prediction estimates, and the most recent sensor measurements, to predict the best current estimates of state. Based on this calculation, the filter formulates an uncertainty estimate. Through the correction phase, the filter receives measurements from an external source and determines how noisy the measurements are, assigning a weight to them. If measurements are deemed to be noisy then the weight value decreases, corresponding to the algorithm relying more heavily on previous predictions to form new predictions. In contrast, if the calculated weight value increases, then future predictions will be influenced greater by the measurement. Subsequently, the received measurements are compared to previous predictions and are brought forward to the next prediction phase, along with the calculated weight. This calculated weight is then used to update the uncertainty of the new prediction. A flow chart diagram of the Kalman filter is illustrated in Figure 4.2 with notation described in Table 4.3.

Table 4.3: Notation used for Figure 4.2.

Mathematical symbol	**Description**
x	State estimate
P	Covariance matrix
F	State transition matrix
H	Measurement function
Q	Process uncertainty
R	Measurement uncertainty
G	Noise distribution matrix
n	Measurement epoch
z	Measurement from external source
K	Kalman gain
I	Identity matrix
(-)	*a priori*
(+)	*a posteriori*

The application of Kalman filters assumes that both the system model and measurement model of the process are linear. In addition, there are other assumptions that Kalman filtering relies on [267]:

- The system noise and measurement noise are uncorrelated zero-mean white noise processes with known autocovariance functions;
- The initial system is a random vector that does not correlate with the process noise or measurement noise;

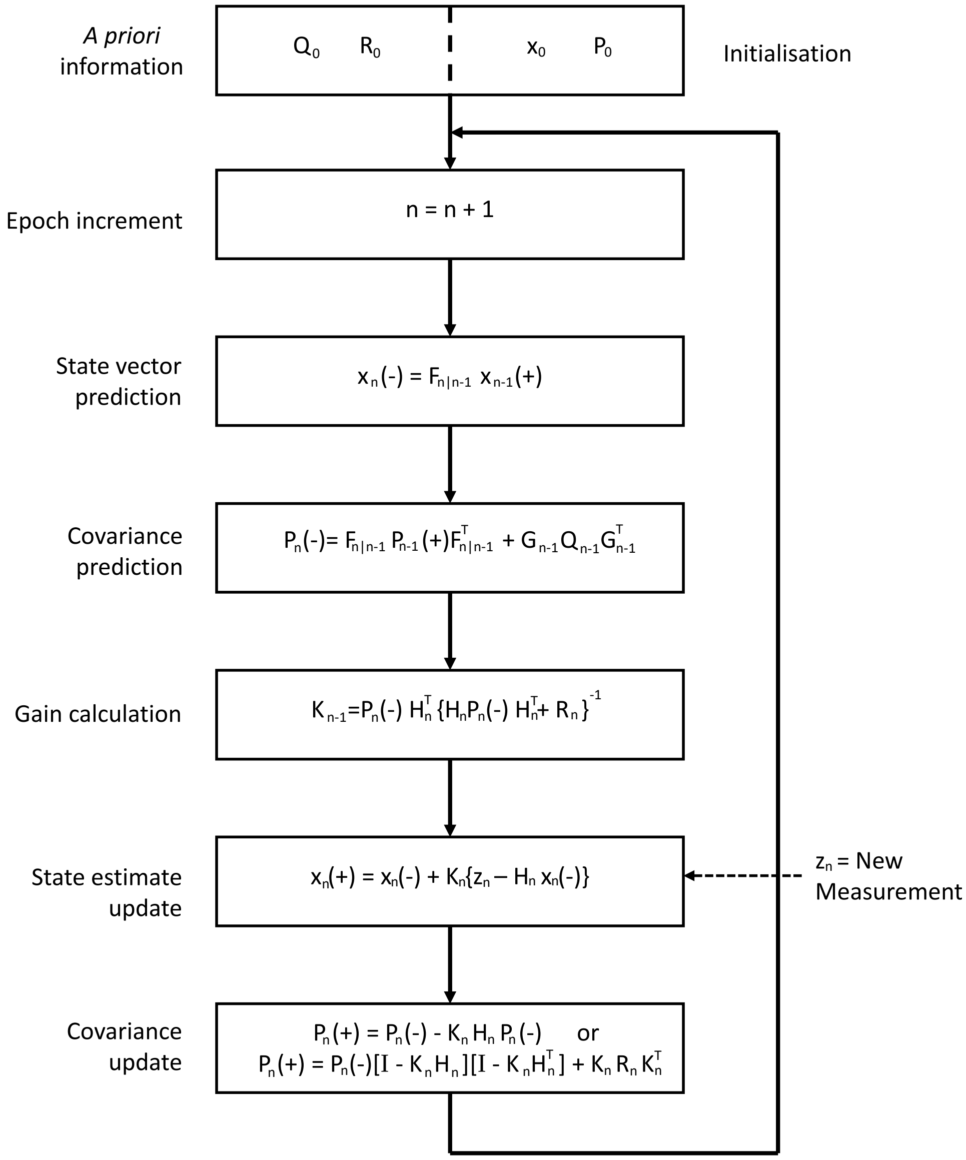

Figure 4.2: Flow diagram showing the steps of a Kalman filter. Notation is described in Table 4.3

.

- The mean value of the initial state and the associated covariance matrix are known.

With these assumptions met, Kalman filtering is a useful methodology to remove errors due to noise, and can be an integral part of sensor fusion within INS.

Another approach to improve accuracy of orientation estimates is through complementary filters [264]. Complementary filters are low computationally-complex sensor fusion algorithms that use standard filters with different characteristics that complement each other and applied in tandem. In the context of inertial sensor fusion, an example of a complementary filter would use a low-pass filter for the accelerometer and magnetometer to remove high frequency components from the data, and a high-pass filter for the gyroscope data to remove drift, before combining the respective filtered outputs and attaining orientation estimates. Methodologies to obtain orientation based on complementary filters can be seen in the literature, with popular examples including the Mahony algorithm [268] and the Madgwick algorithm [269]. Furthermore, Madgwick [270] proposed a separate approach based on the Mahony algorithm [268], which is distinct to his previous work proposed in [269]. This revised Attitude and Heading Reference Systems (AHRS) algorithm [270] includes additional algorithmic improvements such as a faster gain initialisation procedure, sensor-based rejection and compensation mechanisms through filtering and bias adjustments. This AHRS algorithm [270] has since been commercialised in products by x-io Technologies Limited providing a readily available solution for further kinematic research for human motion capture applications, such as sit-to-stand [271], gait tracking [272], and upper limb prosthetic control [273], all of which use an orientation modelling methodology (described in Section 4.1.2). Device availability and adoption of the AHRS methodology by the research community motivate the use of these devices, for the experimentation and analysis presented in this chapter. In summary, the high-level block diagram of the steps to the AHRS algorithm [270] as shown in Figure 4.3, illustrates how raw gyroscope, magnetometer and acceleration measurements are processed to obtain orientation, and linear acceleration in the Earth reference frame (Global acceleration) and the sensor reference frame (Zero-g acceleration).

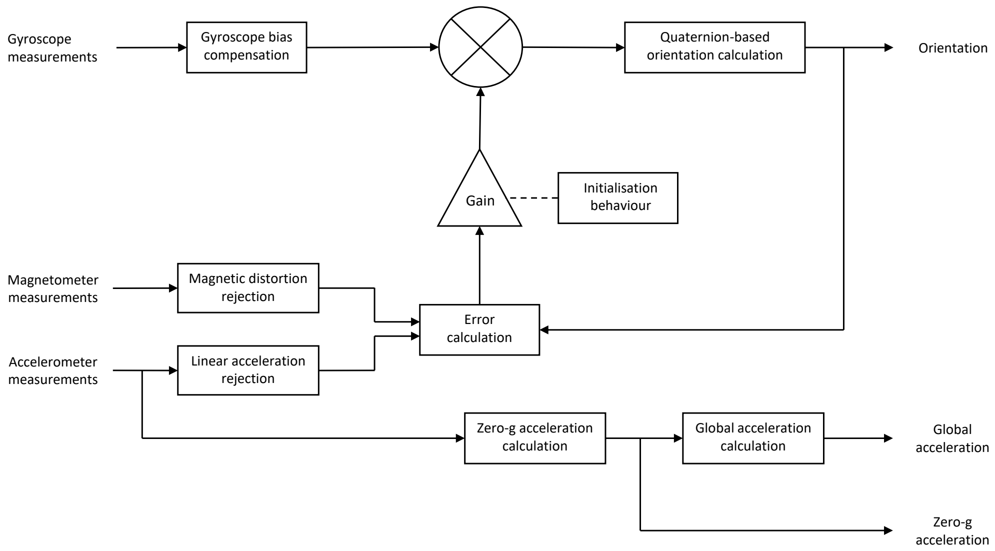

Figure 4.3: Revised attitude and heading reference system algorithm as described in Madgwick [270].

4.1.2 Orientation Modelling

Previous movement sonification work using inertial sensors for rehabilitation applications typically makes use of the orientation output from multiple IMUs in tandem [94,108]. A typical methodology would use Euler angles (or quaternions) of each sensor calculated in the global coordinate system; from there the angular displacement of each proximal sensor is added to each distal sensor to obtain position of the distal sensor. For example, referring to Figure 4.4, to calculate the position of the elbow P_e where the shoulder is the origin, one would calculate the following:

$$P_e = R_{es} L_1 \tag{4.5}$$

where R_{es} is the rotation matrix of the upper arm, and $L_1 = [L_e, 0, 0]^T$ corresponding to the distance between shoulder joint and elbow joint. In turn, the position of the wrist P_w would be calculated through:

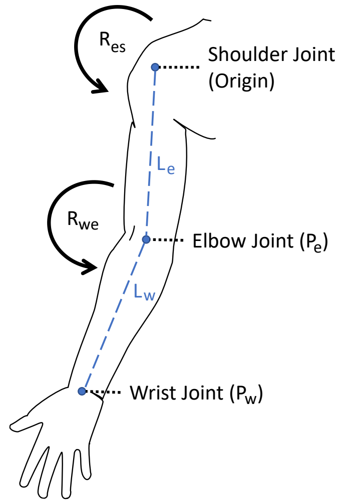

Figure 4.4: Diagram illustrating the variables required to calculate the position of the wrist, with respect to the shoulder, through orientation modelling.

$$P_w = R_{we}L_w + P_e \tag{4.6}$$

where R_{we} is the rotation matrix of the forearm where the elbow is the origin, and $L_2 = [L_w, 0, 0]^T$ corresponding to the distance between the elbow joint and the wrist joint. A detailed description of the process is given in Zhou *et al.* [274]. Although this method has been shown to work effectively, the requisite amount of sensors required to track the distal sectors of anatomy can lead to excessive setup time and a higher acquisition cost. Therefore, a methodology to obtain position (and subsequent time-derivative kinematics) using a single IMU would minimise setup time and acquisition cost, and therefore, maximise the accessibility of this technology.

Movement sonification system development as described in this chapter will explore the creation of position estimates through a single IMU as the motion capture sensor, and if possible, integrate the technology into a movement sonification system.

4.1.3 Physical Dimensions for Movement Sonification

Obtaining the orientation from a single sensor is established in the literature and is readily available in commercial INS in the quaternion or Euler angle form. Velocity and position are commonly used in the literature for sonification purposes [45] and are intuitive metrics for movement. Hence, this section provides the formulation to obtain velocity and position signals from recorded acceleration measurements in the Earth reference frame.

Considering a 1-dimensional movement, velocity and position can be calculated through Equations 4.7 and 4.8:

$$V_n + c = \sum_{n=1}^{\infty} A_n \Delta n, \tag{4.7}$$

$$P_n + d = \sum_{n=1}^{\infty} (V_n + c) \Delta n, \tag{4.8}$$

where P_n is the position, V_n is the velocity, A_n is the acceleration, n is the sample number, and c, d represent the integration error. The effect of c, d can be observed through Figures 4.5 and 4.6.

Figure 4.5 demonstrates a simulated extending motion consisting of a single movement unit with linear increasing acceleration from the beginning of the unit to the first quarter, followed by decreasing acceleration to a negative peak of equal amplitude reached at the third quarter of the unit and finally, increasing acceleration to zero at the unit completion. This results in a velocity profile with a positive peak in the middle of the unit and ideal smoothness. For a simulated retracting motion, the process is reversed. An example of four simulated reaching motions (extending-retracting) is illustrated in Figure 4.5. In the middle plot, the velocity profile is obtained from the integration of the acceleration trace once, while the position plot is the result of double integration of the acceleration trace.

Performing the same methodology on real data would result in the traces shown in Figure 4.6. The data plots shows 15 seconds of recorded acceleration, sampled at 100 Hz, post-filtered through a second order low-pass Butterworth filter with a cutoff

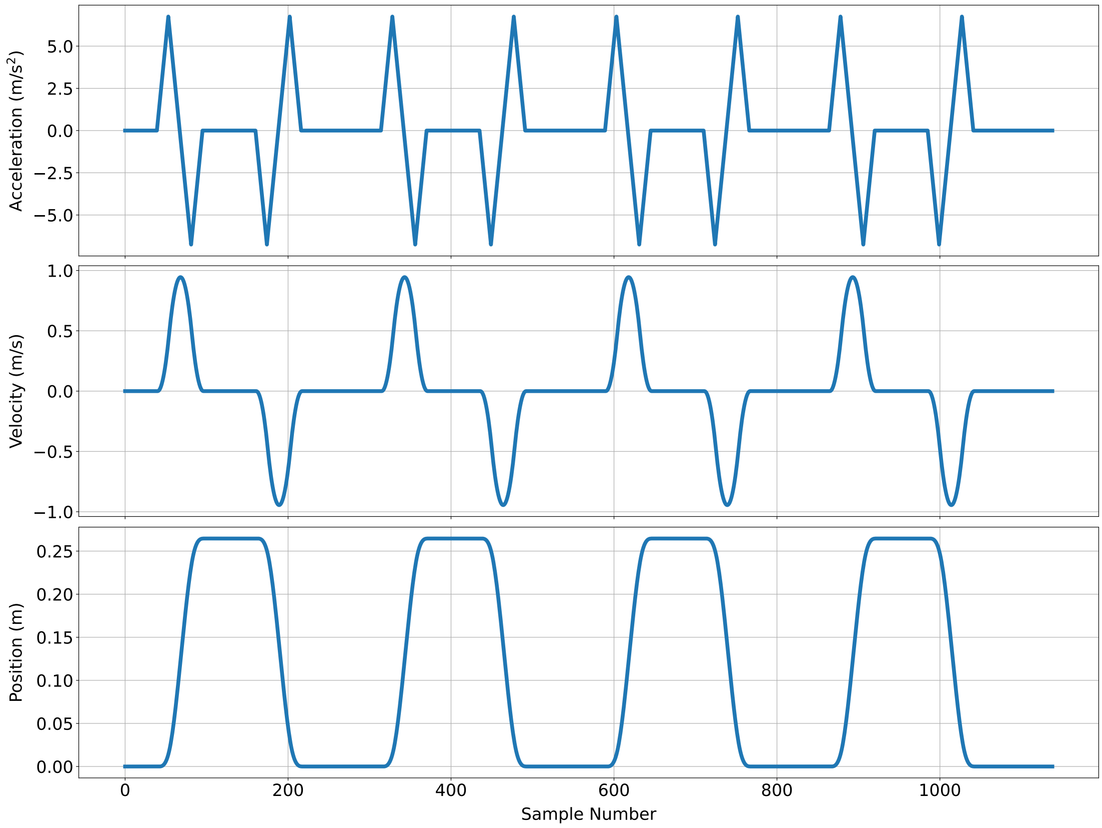

Figure 4.5: Three plots representing simulated reaching movements from the kinematic coordinates of the hand. Top plot represents acceleration, middle plot represent velocity, and bottom plot represents position.

frequency of 6 Hz, in both directions [275]. Integrating the acceleration data to obtain velocity and again to obtain position, results in accumulation of drift error after 15 seconds resulting in a resting position lower than -2.5 m, instead of zero. This example illustrates that double integration of acceleration to position alone does not produce a recognisable signal shape; i.e. similar to the one obtained in the simulated case. One approach to bound drift error is to regularly reset accumulation through additional signal processing such as ZUPT and Kalman filters, which have been used in this work.

4.1.4 Zero Velocity Updates

The Zero Velocity Update (ZUPT) method is a measurement correction method that identifies known stationary periods of data which can then be used to remove error. In

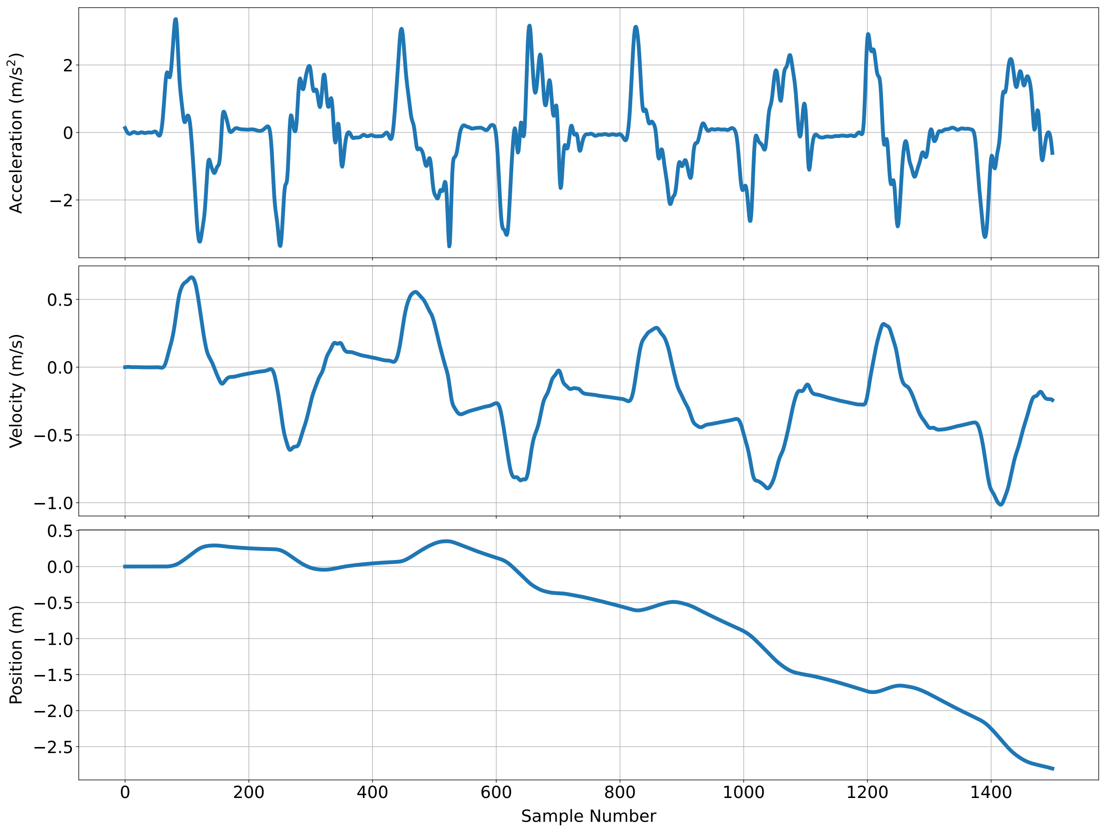

Figure 4.6: Three plots representing a reaching movement recorded from the kinematic coordinates of the hand. Top plot shows a recorded post-filtered acceleration plot, middle plot shows velocity estimates from integrated acceleration, and bottom plot represents position estimates from double integrated acceleration.

context of obtaining position from acceleration, the concept of this methodology can be explained as follows:

- Acceleration is integrated to obtain velocity estimates;
- Velocity estimates contains integration error, obtained by integrating acceleration;
- Known stationary periods of movement in the time domain are recorded and mapped to the velocity;
- Within the stationary periods, velocity is explicitly set to zero, removing integration error;

- Modified velocity estimates are subsequently integrated to obtain position estimates.

The outcome of this methodology results in reduced error in position estimates when converting from velocity. A common application of ZUPT is gait analysis [276, 277], whereby identifying and referencing known stance phases of a gait cycle, position estimates of the foot during normal walking motion were improved. As normal walking patterns are cyclic, this methodology provides a repetitive recalibration procedure during intended use. Application of ZUPT in upper limb tracking also exist in the literature. Comotti *et al.* [278] created an extended arm swing application in the caudal/cranial axis, by applying a 2 Hz low-pass filter to the measured acceleration magnitude and using a value threshold to identify movement and stationary periods. This approach requires post-filtering (after the complete motion) and is therefore, unsuitable for real-time operation. Bai *et al.* [279] created an application for use with the 9HPT, through the use of different value thresholds determined by short-time energy calculations, and a separate threshold to identify values (and therefore, temporal periods) crossing zero. This application requires four samples worth of data sampled at 120 Hz (resulting in latency of $0.03\dot{3}$ s) and could be feasible for real-time operation, however, the energy thresholds require manual selection to operate effectively. Furthermore, given the heterogeneity of hemiparetic stroke survivors and the general characteristics identified for paretic movement the threshold selection introduces challenges for adoption in a rehabilitation tool. Both approaches proposed demonstrate that estimating position from acceleration with ZUPT is feasible. The proposed implementation of the ZUPT approach adopted in this work is detailed in Section 4.3.

4.2 Data Acquisition

To benchmark the motion capture effectiveness of the Soniccup (described in Section 4.3) and KinectSon (described in Section 4.5), an experiment was devised to acquire data from a Vicon Nexus optoelectronic marker-based motion capture system while simultaneously capturing data from the Soniccup and the KinectSon systems. The Vicon Nexus was selected as it is widely accepted in the literature as the gold-standard for motion analysis, with typical associated measurement error in the sub-millimetre range [280, 281]. The Vicon system used is set up at the Wolfson Centre, University of Strathclyde, and uses T-Series cameras that obtain the position of attachable retroreflective markers. For this experiment, a single marker was attached to the Soniccup (specifically on top of the Next Generation Inertial Measurement Unit (NGIMU) sensor) to permit the Vicon system to capture the position of the sensor.

To summarise the data acquisition parameters of these systems, Table 4.4 describes the Soniccup, Table 4.5 describes the KinectSon, and Table 4.6 describes the Vicon system.

Table 4.4: Data acquisition parameters used for the Soniccup.

NGIMU	**Accelerometer**	Range	$\pm 16\ g$
		Sample Rate	100 Hz
		Resolution	16-bit
	AHRS	Gain	0.8
		Gyroscope Bias Correction	True
		Magnetic Field Rejection	20-70 μT
	Analogue Input	Range	0 - 3.1 V
		Sample Rate	100 Hz
		Resolution	10-bit
Switch	**Momentary Push Button Switch**	Operating Force	1.2 N
		Total Travel Distance	2 mm
		Electrical Travel Distance	0.8 mm

Table 4.5: Data acquisition parameters of the KinectSon.

Azure Kinect Camera	**Depth Camera**	Resolution	640x576 px
		Field of Interest	75°x65°
		Frames Per Second	30
		Measuring Distance	$\approx$ 2 m
	RGB Camera	Resolution	1280x720 px
		Nominal Field of View	90°x59°

A volunteer, without neurological or movement deficit, was instructed to perform three sets of fifteen reaching movements with their dominant hand, with each, corresponding to the lifting of the device, extension of the arm, placement of the

Table 4.6: Data acquisition parameters of the Vicon system.

<table>
<tr><td rowspan="6">Vicon T-Series Cameras</td><td rowspan="3">T40</td><td>Resolution</td><td>2352x1728 px</td></tr>
<tr><td>Number of Cameras</td><td>6</td></tr>
<tr><td>Frames Per Second</td><td>100</td></tr>
<tr><td rowspan="3">T160</td><td>Resolution</td><td>4704x3456 px</td></tr>
<tr><td>Number of Cameras</td><td>6</td></tr>
<tr><td>Frames Per Second</td><td>100</td></tr>
</table>

Soniccup at maximum reach, and then lifting of the device, retraction of the arm, placement of the Soniccup at approximately the starting position. The volunteer was instructed to modify the speed of reaching movement so that the first set was performed at a normal speed, the second set was performed at a slow speed, and the third set was performed at a fast speed. The volunteer was instructed to perform the slow movements inline with their own deep breath, inhaling during the extension phase, and exhaling during the retraction phase, and were instructed to perform the fast movements as quickly as able whilst in control of their movements. Further details of the procedure used is explained in Section 5.3.1.

Figure 4.7 shows a velocity comparison of the first four movements in the three sets of movements captured. From the collected movement data of the dominant hand, the first four reaching movements of the set corresponding to normal speed movements have been used to visually display the Soniccup methodology and KinectSon methodology described in Sections 4.3.4 and 4.5, respectively, and the full amount of movements in the normal speed and slow speed sets of movement are used in Section 4.3.6. The set of fast speed movements are not used in the following sections, but have been displayed in Figure 4.7 to provide a velocity visual comparison of the three sets of movement data.

Data obtained from the Vicon system has been assumed to contain no errors and therefore, have been considered accurate. Visual inspection of the traces of position, velocity, and acceleration from all systems confirmed that all systems capture the same pattern for reaching motion.

Figure 4.8 illustrates the data captured via the Soniccup and Vicon system, where the dark traces indicate the primary measurement from each sensing system (position for the Vicon system and acceleration for the Soniccup), while the lighter shaded lines

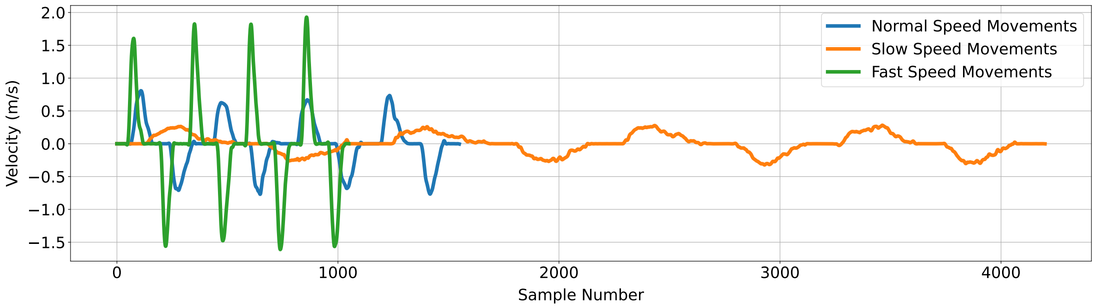

Figure 4.7: Figure showing the calculated velocity of three sets of reaching movements as captured by the Vicon system.

show the derivative domains. From these traces, it is clear that the two systems capture the pattern associated with four separate reaching movements. However, due to the lack of synchronised start (and stop) of the data acquisition, the data from the two systems require manual alignment.

Data alignment was achieved through a two-step process. First, the acceleration data obtained through the Soniccup and the Vicon system were normalised to a magnitude of positive and negative one. Second, cross-correlation calculations were used to identify the temporal shift value that results in the highest similarity. Equations 4.9 and 4.10 mathematically displays this process.

$$x_n = a_t/max(a) \tag{4.9}$$

$$z_k = \sum_{n=0}^{||x||-1} x_n y_{n-k+(N-1)} \tag{4.10}$$

For these equations, a is the acceleration signal, a_t is a sample of the acceleration signal, x and y are signals after the normalisation process, z_k is the cross correlation output at sample difference k, n equals the sample number, x and y are discretely sampled signals, $||x||$ is the length of x, and N is the highest number of samples in x or y. The index corresponding to the maximum cross correlation value was used to identify the sample difference between the two signals. This was visually verified through isolating and viewing the plot of a small subset of the data. Figure 4.9 shows

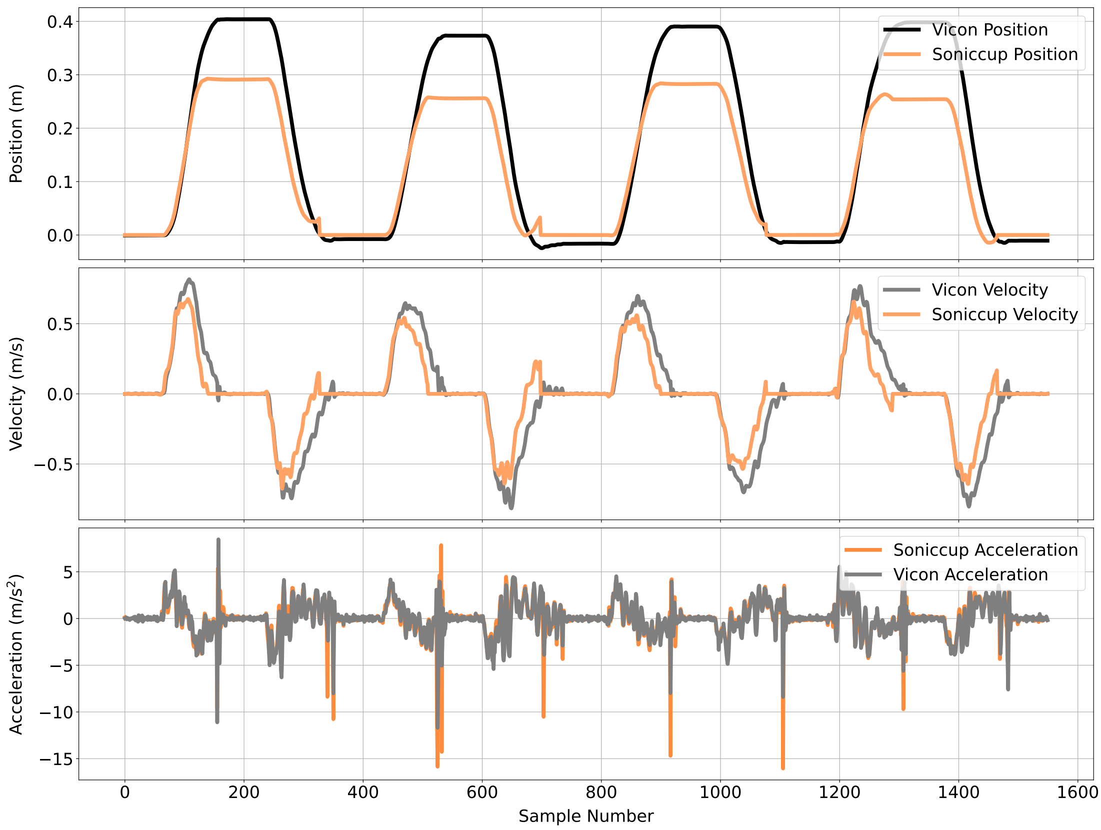

Figure 4.8: Figure showing plots corresponding to the position (top), velocity (middle), and acceleration (bottom) obtained from the Soniccup (orange traces) and Vicon (black traces) systems. Lighter shades of each colour indicate calculated/estimated values for the kinematic.

three plots, the top plot shows the two acceleration data sets, highlighting the sample difference between the two.

Through the cross-correlation process as stated in Equation 4.10, the middle plot shows the overall result, and zooming in to the graph at the maximum cross-correlation, as illustrated through the bottom plot, shows the result of a seven sample lag of the Soniccup data relative to the Vicon system. As such, the acceleration values of the Soniccup must shift seven samples towards the Vicon data set to achieve alignment. It should be noted that the sampling frequency of both the Soniccup and Vicon systems was 100 Hz, and from visual observation, the two data sets do not expand or shrink sufficiently enough to warrant the segmentation of data for alignment purposes.

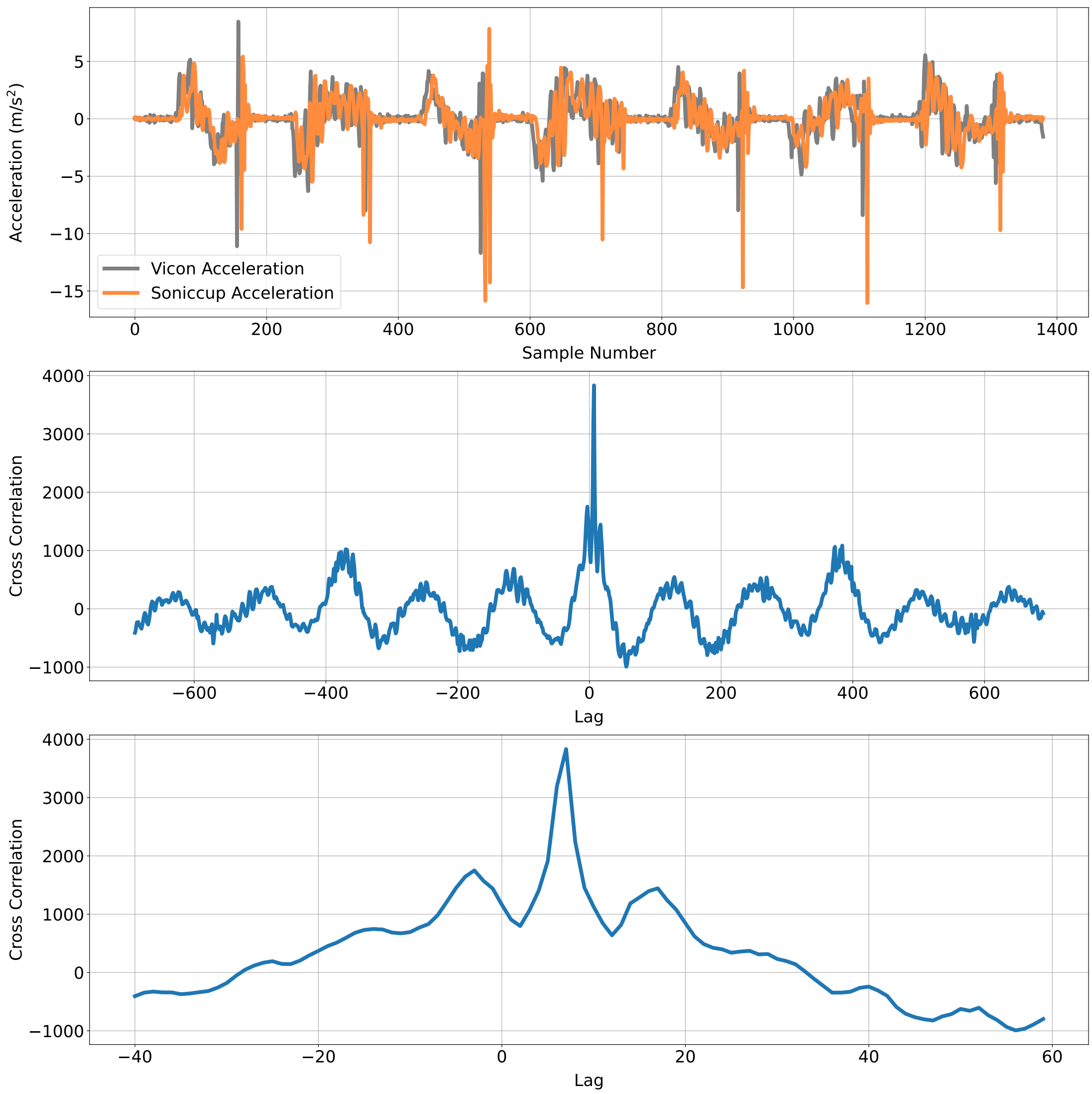

Figure 4.9: Figure showing three plots corresponding to the alignment process between the data obtained from the Soniccup and Vicon systems. Top plot shows the recorded acceleration from the Soniccup (blue), and the calculated acceleration from the Vicon system (grey). Middle plot shows the result of a cross correlation process between the acceleration data from the two systems. The bottom plot zooms in at the peak of this result to show the sample offset of seven, for maximum correlation.

4.3 Soniccup

This section details the design and creation of the Soniccup prototype proposed in the thesis. At the commencement of this work, the following research questions were proposed:

RQ4 Using Kalman filters and the ZUPT method, to what extent can position estimates be extracted from gravity-free linear acceleration?

RQ5 Can the ZUPT method be applied to upper limb movements without limiting to one type of movement?

RQ6 To create audio feedback, what steps must be taken to convert position into an auditory display?

The final Soniccup, comprises an INS, a push-to-make switch, and a 3-dimensional printed object. The device communicates over Wi-Fi with a PC that hosts the position estimation algorithm that utilise acceleration readings in the Earth reference frame, and analogue voltage values corresponding to the closing and opening of a switch and output estimated velocity and position data. The estimated outputs can then be converted to an audio output through the developed sonification platform. The Soniccup operates on movements that contain anchoring and in this thesis the focus is on reaching forward movements delineated by the placement of the Soniccup on surface.

4.3.1 Next Generation IMU

To create a proof-of-concept system, a product was acquired from x-io Technologies Limited, named the Next Generation Inertial Measurement Unit (NGIMU), which contains a 3-dimensional accelerometer and gyroscope that each have a maximum sample rate of 400 Hz, plus a magnetometer at 20 Hz maximum sample rate, along with sensors to monitor pressure, humidity, and temperature, and is powered with a rechargeable 1000 mAh battery, enclosed by plastic housing. The device provides wired connectivity through serial and Universal Serial Bus (USB) connections and wireless connectivity through Wi-Fi, 802.11n, 5 GHz. The Wi-Fi module can operate in

Figure 4.10: Next Generation IMU with housing, image taken from x-io Technologies Limited website [282]

Access Point or Client mode. The implementation described in this thesis operates the device in Wi-Fi Client mode. The NGIMU acquires raw data from sensors and provides processing capabilities to compute derivative and fused metrics such as those from the AHRS; the quaternions, rotation matrix, Euler angles, gravity-free acceleration in the sensor coordinate frame, and gravity-free acceleration in the Earth coordinate frame. An associated software package to the NGIMU permits the configuration of device parameters such as the raw sensor data acquisition frequency, send rates, AHRS fusing parameters [270], enabling and disabling reporting of the derivative and fused metrics. The gravity free acceleration in the Earth coordinate frame, theoretically provides acceleration that is independent to the rotation of the device. This metric was chosen, fundamentally because human biomechanical movement is a combination of anatomical rotations, and considering that the target user audience (stroke survivors) of the system will have incomplete motor control and are therefore, likely to perform compensatory rotations, segregating the effects of the additional rotations from the acceleration plots would increase audio feedback versatility, and could be included in future work involving orientation measurements. The NGIMU communicates to the host computer through Open Sound Control protocol, where packets are received and decoded, allowing data extraction. In terms of peripheral ports, the device also contains an SD card socket, two serial interfaces, and an analogue input connector, which is used in the system described in this chapter. The pinouts for this connector comprise a 3.3 V output, ground, and eight analogue channels to receive inputs. To conclude on this section, as a means of creating a proof-of-concept system, the NGIMU provides an inexpensive and versatile option to investigate different methodologies and strategies, and hence was selected.

4.3.2 Zero Velocity Update Trigger

As a fundamental to the system design, the choice was made to include the ZUPT as a means of improving positional estimation. Therefore, it is necessary to devise an apparatus to detect zero velocity and in the context of upper limb movements, a potential solution is to include a sensor to detect contact with a surface. Two types of sensors were considered, a pressure sensor and a push-to-make switch. Either of these sensors would achieve the goals of obtaining zero velocities, however, the pressure sensor introduces additional complexity in terms of signal conditioning and interfacing and was not adopted in favour of a push-to-make switch. A non-latching double pole single throw switch with 1.2 N of operating force (described in Section A.1.1) was selected that allows the switch to be activated by the weight of the comprised object alone, which is necessary to improve usability for people with hemiparesis. The switch was connected to the analogue inputs available in the NGIMU as shown by Figure 4.11.

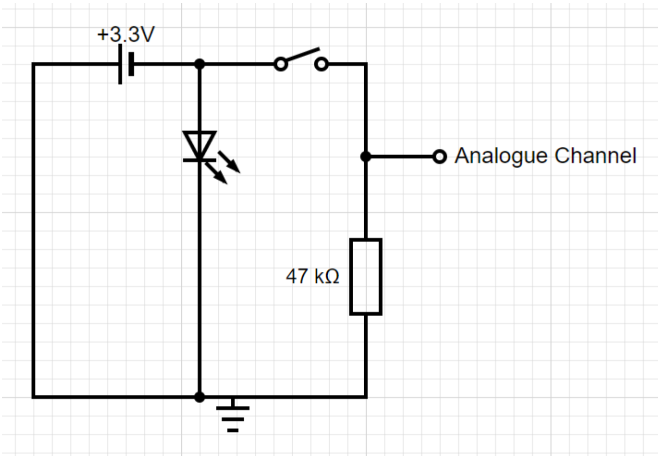

Figure 4.11: Circuit diagram showing the additional electronic components used to create a switch enabled voltage divider for the Soniccup system.

The NGIMU supplies 3.3 V to a voltage divider, and is able to monitor the output, labelled ‘Analogue Channel’, as a consequence of the switch state. As such, a high-voltage recording corresponds to the compression of the switch, and a low value represents the release of the switch.

4.3.3 Soniccup Housing

In regular upper limb practice, movements often involve reaching for objects to be picked up and placed. These objects are generally lightweight, durable, inexpensive, and in contact with a surface. Therefore, creating an object that detects the state of being picked up or placed, in combination with the ZUPT, would produce a suitable methodology to obtain velocity and position estimates. A 3-dimensional printed designed object permits integration of the push-to-make switch and NGIMU in a single package.

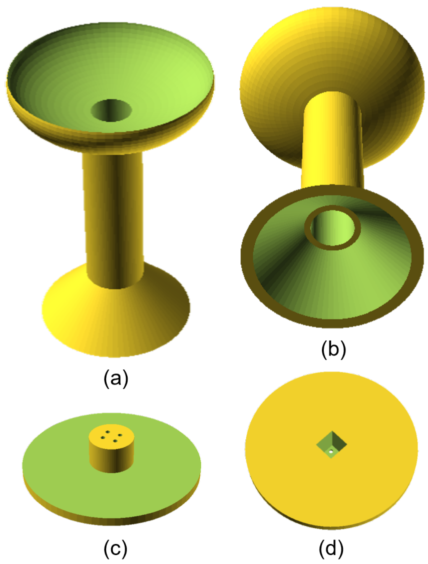

Figure 4.12: Design of two 3-dimensional printed components to be combined to form the housing of Soniccup, (a) and (b) show different perspectives of the body of a chalice shape object, (c) and (d) show different perspective of an attachment base to the chalice shape object, designed to include a specific double pole double throw switch in the prototype model.

Figure 4.12 shows rendering of the 3-dimensional printed object that consists of the primary component in (a) and (b) in two different perspectives. The top side of this component is designed to accommodate the NGIMU and additional circuitry, while the bottom side is designed to attach to the secondary component. The stem of the object is a hollow cylinder to allow connection of the device with the ZUPT trigger (switch), and to provide a robust handle whilst also reducing the mass of the object. Figure 4.12 (c) and (d) shows the secondary component to securely hold in place the push-to-make switch. Special attention was given to the switch placement to ensure the bottom surface was flush when the switch was fully compressed. The exact dimensions of both components are detailed in Section A.1.2. This housing along with the NGIMU constitutes the complete physical object labelled as Soniccup.

4.3.4 Position Estimation

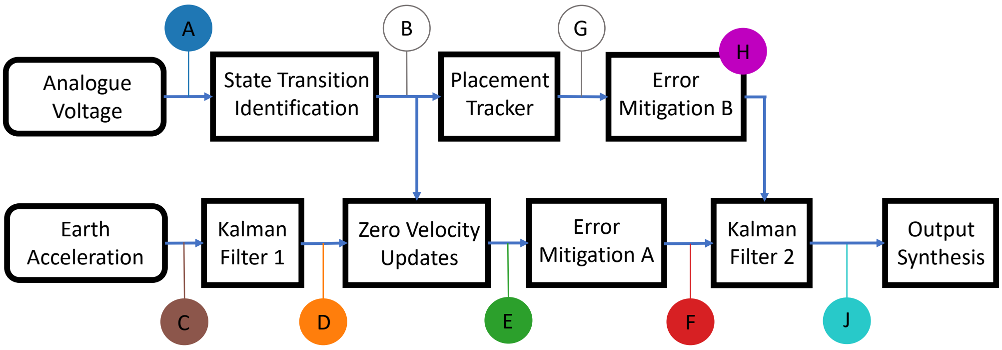

Figure 4.13: Block diagram showing the signal conditioning steps to the sonification stage, starting from analogue input and Earth acceleration.

Soniccup was set to wirelessly send two data streams to a host PC, one contains an analogue voltage measurement corresponding to the switch state, and another is the acceleration measurement in the Earth reference frame, with both data streams measuring at 100 Hz. Figure 4.13 shows an overview of the end-to-end processing to converting these two data streams into estimated position. The estimated position data is subsequently converted into Musical Instrument Digital Interface (MIDI) notes to produce auditory feedback. To further explain the methodology used to estimate position, data obtained through Section 4.2 are displayed in a step-by-step process, firstly describing the generation of the Zero Velocity Update Trigger Signal, and secondly describing the conversion process of acceleration to position. Throughout this description the main focus is the primary direction of travel along the sagittal axis (labelled as the X axis), unless otherwise stated. A positive change in the X axis represents an arm extension movement, whereas a negative change in the X axis represents an arm retraction movement.

State Transition Identification

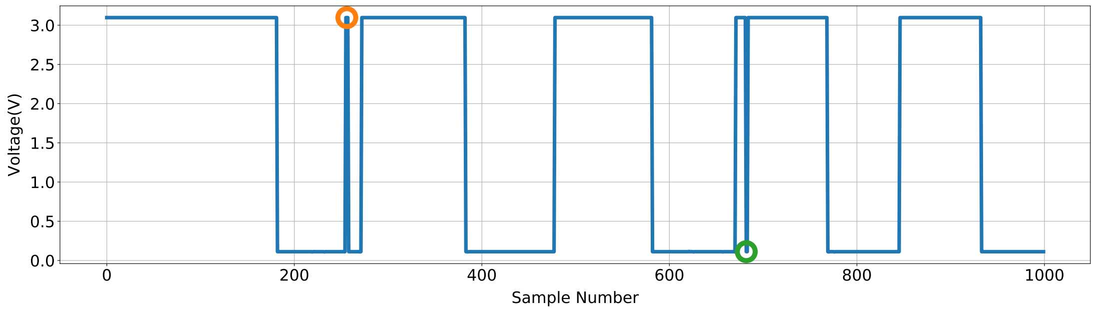

Figure 4.14: Figure depicting the recording of mechanical bouncing. Two events are shown by orange and green circles, corresponding to a momentary placement and momentary lift of the Soniccup, respectively.

Focusing on the analogue voltage data stream, at point 'A' as shown on Figure 4.13, the data is processed to identify states of when the Soniccup is placed or picked up from a surface. Figure 4.14 shows that data stream trace which ranges between 0 and 3.3 V. It can be observed that momentary fluctuations of the analogue voltage occur nears the pick up and placing down events. These events are shown on orange and green markers which correspond to the placement and pick up respectively. The orange marker shows a low-to-high transition, indicating that the switch had been momentarily compressed while the green marker shows a high-to-low transition, indicating the momentary release of the switch. These events typically happen due to imperfect placement or picking up of the Soniccup, resulting in a type of 'biomechanical bouncing' that are separate events from a purposeful placement/pickup. As accurate identification of placements and pickups are necessary for ZUPT, the detection and removal of these events are of interest. To filter out these momentary changes when running the system online a state change is considered only if the analogue voltage signal persists for n consecutive samples. A small value for n runs the risk of momentary changes remaining unfiltered, while a high n will lead to lag between the beginning of the movement and the audio generation which is undesirable. For this system the value of n was selected to be eight after experimentation and results in a delay of 80 ms, which is below the delay threshold for real-time applications [103]. The subsequent stage in the process, at point 'B' on Figure 4.13, receives the filtered switch state signal ready to apply the ZUPT method to the velocity estimates.

Acceleration to Position

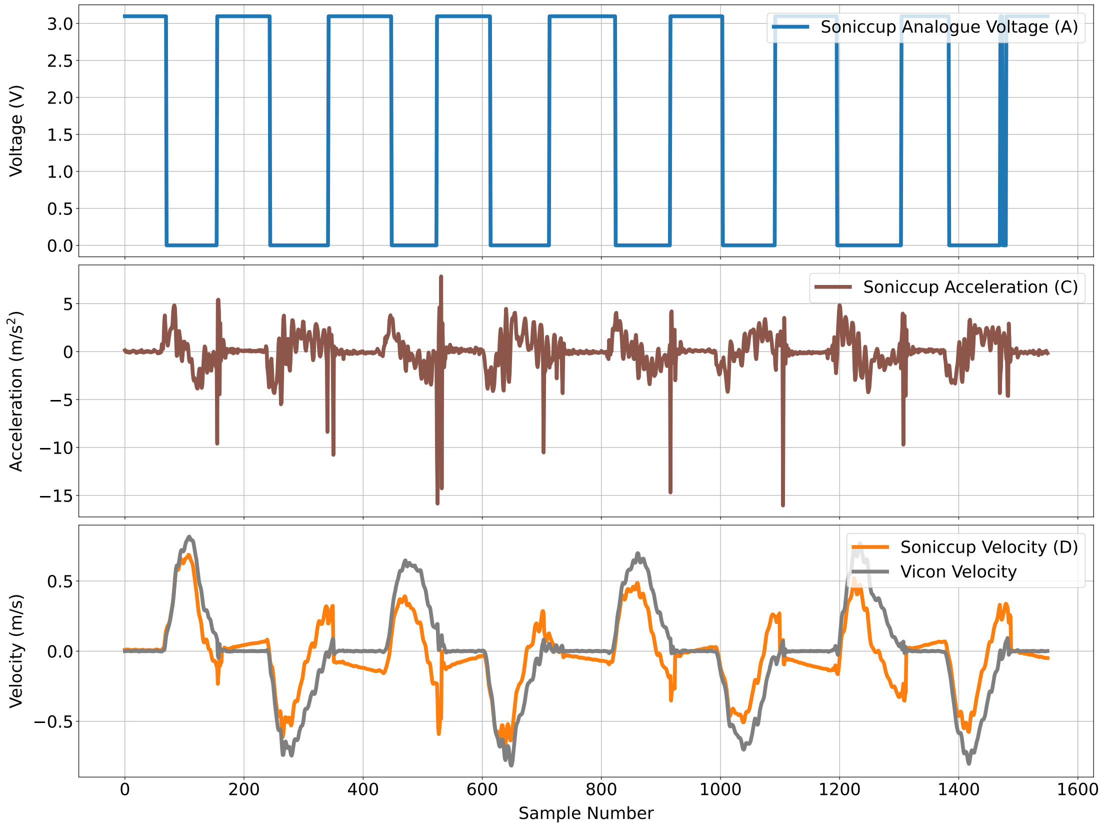

Figure 4.15: Figure showing associated data prior to and with the first designed Kalman filter. Top plot presents analogue voltage obtained through the NGIMU, middle plot presents raw data values corresponding to acceleration in the Earth reference frame obtained through the NGIMU sensor, and bottom plot presents the estimated velocity from the first Kalman filter.

Focusing on the Earth acceleration data stream, referring to point 'C' on Figure 4.13, acceleration data from the sensor is fed through a linear Kalman filter to output estimated velocity. This Kalman filter is setup to integrate acceleration data values that have been filtered through a white noise filter. All the parameters for the Kalman filter are provided in the Appendix A.2.1. Figure 4.15 shows plots associated with input data obtained through the NGIMU sensor, prior to and after the application of the Kalman filter. The top plot shows the analogue switch voltage measurements obtained from the NGIMU, while the middle plot shows the raw Earth acceleration labelled as 'Soniccup Acceleration (C)'. The bottom plot, shows the output of the Kalman filter; i.e. the estimated velocity, labelled as 'Soniccup Velocity (D)', and the velocity obtained via the Vicon system after differentiation of position. Directly comparing the velocity plots between the Soniccup and the Vicon, shows observable differences. These are attributed to the measurement errors associated with the Soniccup data as described in Section 4.1. To remove these errors, and obtain a better match to the Vicon velocity, 'Soniccup Velocity (D)' is processed using the ZUPT methodology. In particular, the use of the analogue voltage signal from the non-latching switch along with the 'State Transition Identification' process as described in Section 4.3.4 were used to identify periods of zero velocity (point 'B' of Figure 4.13). As the stationary phases have been identified, the estimated velocity values obtained during these phases are explicitly set to zero, hence removing integration errors that have accumulated up to this point. Additionally, the Kalman filter is reinitialised at the start of the stationary phase to reset the filter state.

The results of these changes are shown in Figure 4.16 (top) as 'Soniccup Velocity (E)', with the applied ZUPT method. The effect of this processing is that during stationary periods the velocity is now zero. As the point of reset occurs when the device has been placed, and the point of placement and pickup is not instantaneous, this would lead to an error component that distorts the data. To remove this accumulated error component, velocity values are continuously estimated through the stationary periods, and treated as error values so that on transition to a movement phase, the error values are subtracted from the corresponding estimated velocity data, to improve velocity estimates. The outcome of this error subtraction mechanism can be seen through Figure 4.16 (middle) labelled as 'Soniccup Velocity (F)'. This resultant data shows

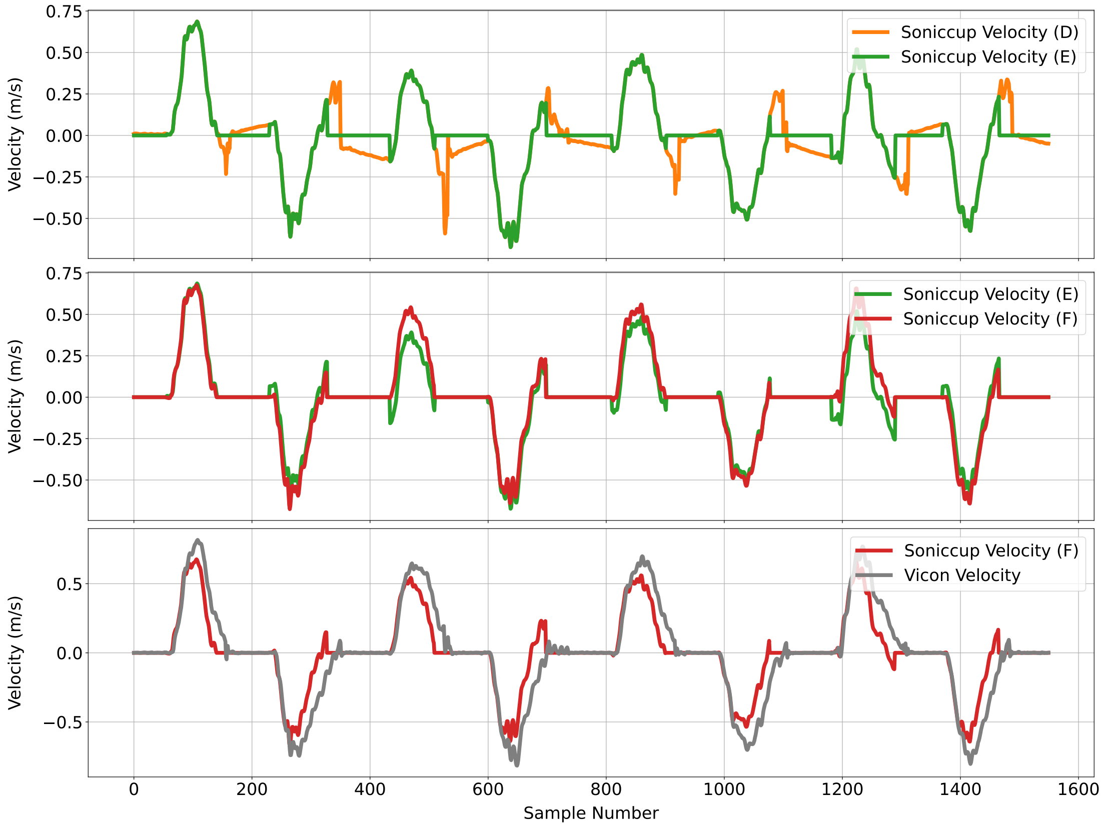

Figure 4.16: Figure showing data associated with signal conditioning velocity estimates. Top plot presents velocity estimates before (orange) and after (green) the application of zeroing by zero velocity updates. Middle plot shows before (green) and after (red) further signal conditioning to remove accumulated stationary errors. Bottom plot compares the conditioned velocity estimates (red) with the calculated velocity obtained through Vicon data (grey).

a shift towards congruence with the calculated velocity data associated with the Vicon system as highlighted through Figure 4.16 (bottom).

Subsequent steps refer to data associated with a second Kalman filter, with the properties as described by Equations A.7 - A.12 in Section A.2.2. Figure 4.17 shows the performance of the second Kalman filter conversion from conditioned velocity data into estimated position data. The top plot shows 'Soniccup Velocity (F)' which is used as an input to the Kalman filter. The bottom plot shows a comparison between the estimated position output from the Kalman filter along the primary axis of movement (X) labelled as 'Soniccup Position (H)', and the recorded position data from the Vicon

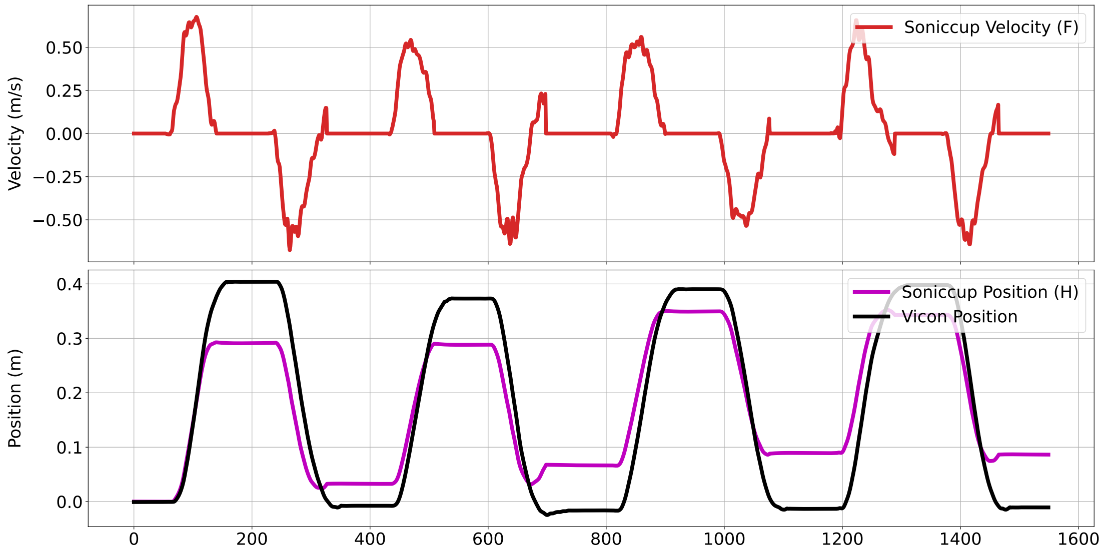

Figure 4.17: Figure showing associated data with the second Kalman filter. Top plot presents conditioned velocity obtained at point (F). Bottom plot presents the corresponding position estimates from the second Kalman filter (purple), compared to the position measurements from the Vicon system (black).

system. The 'Soniccup Position (H)' plot in Figure 4.17 shows a resemblance to the position plot obtained through the Vicon system, however, there are many differences between the two plots, including a difference in position range, movement start and end points, and the appearance of 'lumps' on 'Soniccup Position (H)'. The difference in position range is a consequence of the methodology used to obtain position estimates, and the lumps coincide with the nonperfect shape of the conditioned velocity where the estimated velocity goes beyond zero when returning from maxima/minima velocity. Further conditioning of the velocity plots would result in the removal of these lumps. Observable from the bottom plot the average values of the estimated position per movement fluctuates due to the magnitude of the values measured through one direction not equalling the magnitude of the values measured in the opposite direction. The mismatch of values over time accumulates and becomes salient when observing the integration of the velocity plots. To remove this error the system tracks the amount of placements completed, with the assumption that the first placement corresponding to the end of the extension phase and the second placement corresponding to end of the retraction phase of movement. Upon placement of the device at the end of the

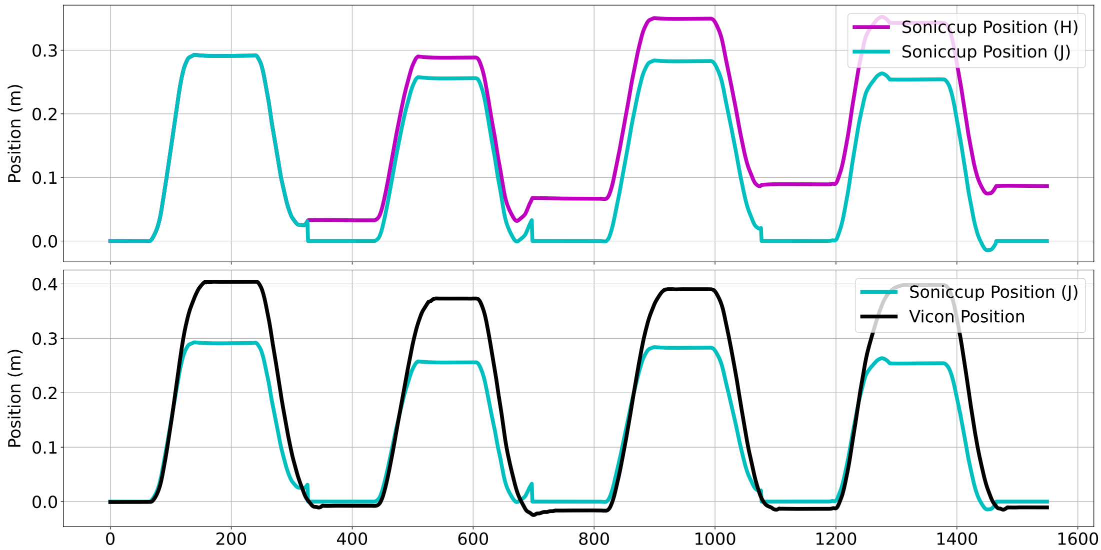

Figure 4.18: Figure showing position estimates from the second Kalman filter and the recorded data from the Vicon system. Top plot showing estimated position from second Kalman filter without (purple) and with (blue) reinitialising the second Kalman filter. Bottom plot shows a comparison of position values from the Soniccup (blue) and the Vicon system (black).

retraction phase, the second Kalman filter is reinitialised. Figure 4.18 shows the effect of reinitialising the second Kalman filter on the estimated position. The top plot shows a direct comparison of the estimated position without ('Soniccup Position (H)') and with ('Soniccup Position (J)') the implemented mechanism. The bottom plot compares this new estimated position with the measured position of the Vicon system. This estimated position is the output of the data processing stage for the movement sonification system and is the input for the sonification module.

Z-Axis Position Estimation

Figure 4.19 shows the result of the same position estimation process applied to the data corresponding to the cranial/caudal (Z) axis. In contrast to the outcome of the methodology for data related to anterior/posterior (X) axis, the estimated position in the Z axis show large differences compared to the Vicon system (bottom plot). It can be observed that there is an offset during the first placement at the end of the extension phase which is subsequently affecting the position estimates during the

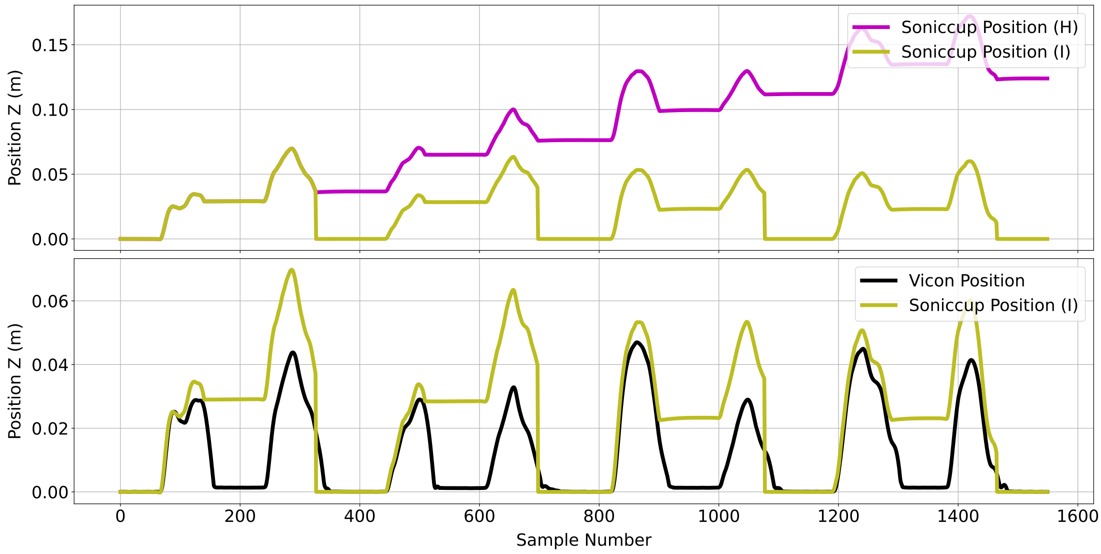

Figure 4.19: Figure showing the predicted position along the Z axis from the second Kalman filter and the recorded data from the Vicon system. Top plot shows estimated position from the second Kalman filter with (green) and without (purple) resetting the Kalman filter at every second placement. The bottom plot shows a comparison of the Soniccup position (green) and the measured Vicon position (black), at this stage.

retraction phase. At the end of the retraction phase, explicit zeroing realigns the two traces and misalignment starts again at the end of the next extension phase. To mitigate these differences, an additional processing step is introduced to eliminate the offset; the Z-axis processing pipeline is similar to that presented in Figure 4.13 but modified to introduce the ‘Error Mitigation C’ step as shown in Figure 4.20. The ‘Error Mitigation C’ step effectively resets the Z axis position after every placement. Figure 4.21 shows the position estimate result for the Z axis with the modified processing pipeline and its comparison to the Vicon system position.

Final Position Output

Figure 4.22 reiterates the final position estimation plots for all three axes and the radial distance, for the Soniccup and the Vicon System. An efficient linear reaching motion (potentially performed by a motorised system) would result in the highest amplitude on the X axis, zero amplitude on the Y axis, and a small amplitude on the Z axis. Considering that the motion in this example was performed by a person, some

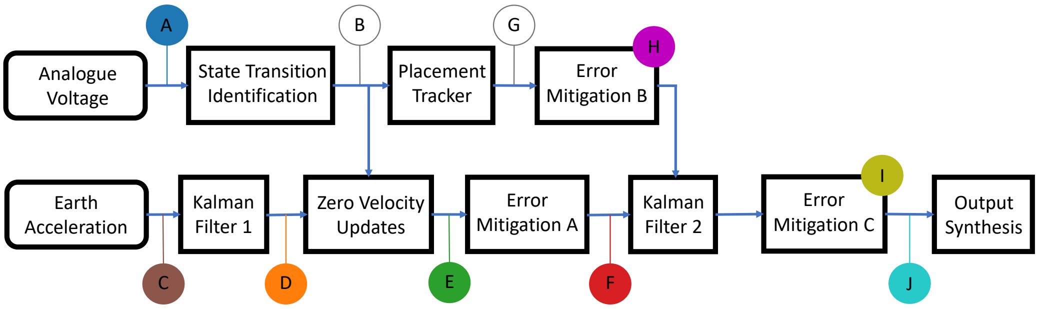

Figure 4.20: Block diagram showing the signal conditioning steps for data associated with the cranial/caudal (Z) axis to the sonification stage, starting from analogue input and Earth acceleration.

amplitude is observed on the Y axis, however, this is not significant for a reaching motion, while the X and Z axis are largely in agreement with the Vicon system. For the purposes of this study, the X axis is used for audio feedback, without loss of generality (for example, the Y axis may be chosen as an input to the sonification system to create audio feedback on lateral motion).

4.3.5 Audio Output

The final stage of the movement sonification process is to translate the motion signals into a form of audio feedback. In this particular study, the position estimates are chosen to generate the audio feedback but other motion related variables such as velocity or acceleration could be chosen. The position estimates are then sonified through parameter mapping to convert data to audio output. The mapping strategy implemented for the Soniccup system maps linearly the estimated position to audio pitch. As the movement performer extends their arm, the audio pitch rises, and when the movement performer retracts their arm, the opposite occurs. Audio output is presented through the use of 'MIDI Notes' [283]. Musical Instrument Digital Interface (MIDI) is a communication standard for digital musical instruments and related devices for playing, editing, and recording music. MIDI note values range from 0 to 128, which is wider than the grand piano, of which the lowest note (A0) corresponds to MIDI note 21, and the highest note (C8) corresponds to MIDI note 108. Notes at the high end of the spectrum generate high treble which may create discomfort to the user, while

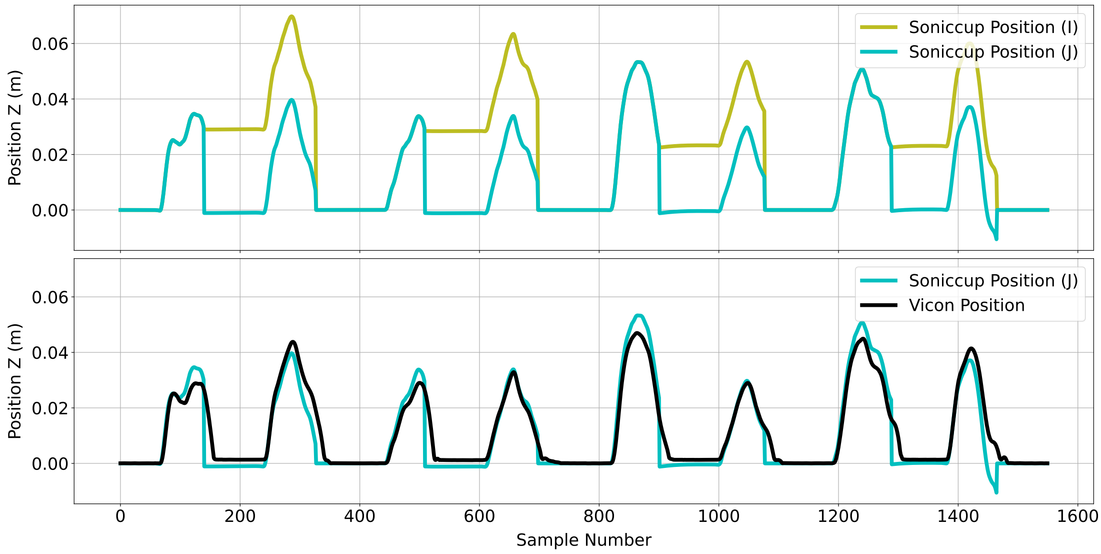

Figure 4.21: Figure illustrating the additional correction mechanism that is implemented for data associated with the cranial/caudal (Z) axis. Top plot shows the effects before (green) and after (blue) the correction mechanism. Bottom plot show a comparison between the final estimated position via the Soniccup and the measured position from the Vicon system.

notes at the low end of the spectrum are harder to perceive. For that reason, a design decision was made to sonify movements using three octaves, with the lowest at MIDI note 48 (C3), and the highest at MIDI note 84 (C6). This range of notes is used for the remainder of this study.

The MIDI note mapping is performed using Equations 4.11, 4.12 and 4.13 to convert positional data to audio output. Initially, the note range $note_{min}$ to $note_{max}$ is scaled over the position range p_{min} to p_{max}, to obtain the scaling factor r,

$$r = \frac{(note_{max} - note_{min})}{(p_{max} - p_{min})} \tag{4.11}$$

where $note_{min} = 48$ and $note_{max} = 84$, which correspond to the upper and lower boundary of MIDI notes available as output, $p_{min} = 0$ and $p_{max} =$ maximum position, which correspond to the start and end values of the reaching motion. Specifically, p_{max} is a value estimated based on the range of movements captured previously through the Soniccup, therefore, trial movements are required prior to online sonification. The

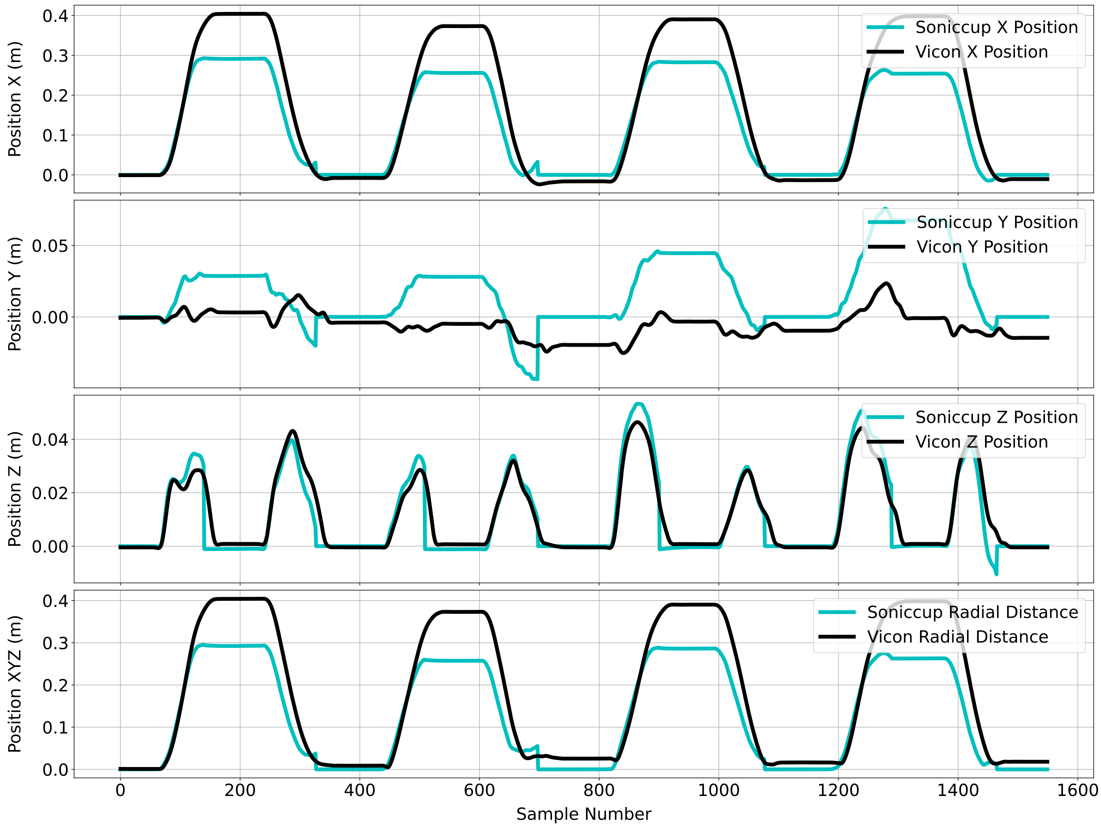

Figure 4.22: Figure containing four plots, corresponding to the estimated position from the Soniccup (blue) and the measured position from the Vicon system (black). Top plot corresponds to the frontal/parietal (X) axis, second plot corresponds to the medial/lateral (Y) axis, third plot corresponds to the cranial/caudal (Z) axis, and the bottom plot corresponds to the radial distance.

scaling factor r is calculated at the system initialisation, and hence requires position estimates for the start and end of movement.

With the aid of the r scaling factor, the audible note n_i at position p_i can be obtained using the equation:

$$n_i = \lfloor r \cdot p_i \rfloor + note_{min} \tag{4.12}$$

Note that the product $r \cdot p_i$ is floored to quantise the note output, as decimal notes are not meaningful. This will result in all 36 notes in the range $note_{min}$–$note_{max}$ to be audible and small deviations in p_i will result in fluctuating audio output. To avoid noisy audio output, note quantisation can be made coarser by increasing the number of steps (n_{steps}) between audible note changes. To achieve this, Equation 4.12 was modified to:

$$n_i = n_{steps} \cdot \lfloor \frac{r \cdot p_i}{n_{steps}} \rfloor + note_{min} \tag{4.13}$$

The n_{steps} variable in essence controls the audible note resolution and in these experiments has been selected equal to three resulting in 12 audible MIDI notes.

Through experimentation, it was observed that through the majority of the reaching motion the audio output changes smoothly, but at the near maximum extremity of the reach, the audible output created trills (rapid alternation between two notes). The trills was a result of the hand hovering around a threshold between MIDI note changes when the arm extension reached 100% and resulted in the movement performer overextending, to mitigate the trill. To mitigate the trill through the sonification configuration, the position estimate was updated to saturate at 95% of the maximum position.

$$p_i = min(p_i, 0.95 \cdot p_{max}) \tag{4.14}$$

The saturation on the high end, helps to discourage over-extension of the movement performer and results in the removal of the last note of the MIDI range (in the case of experiments, from the number of notes are reduced from 12 notes to 11). Furthermore, saturating the position estimate leads to narrowing the reaching extension which in turn, promotes a healthy reaching goal for users of the system. On the lower extremity of reach, when the arm retracts beyond the starting position the position estimate becomes negative and sonification results in emitted notes below the selected MIDI range. To maintain the $note_{min}$ as the lowest audible note, lower side saturation is performed using the following equation:

$$p_i = max(p_i, p_{min}) \tag{4.15}$$

Through Equations 4.14 and 4.15, a sonification range has been created where movement captured outside of designated area occur saturation to sustain a constant audio note range. To benchmark the audio output of the Soniccup, the same audio conversion strategy was used for data acquired via the Vicon system. Naturally, the two systems, Soniccup and Vicon, result in different p_{min} and p_{max} parameters. The aim of the movement sonification system is to capture the relative motion patterns and generate similar audio feedback for similar pattern, therefore, the position estimates for both system are normalised using max normalisation and results in values in the range [0,1] using the equation:

$$p_{norm} = p_i / p_{max} \tag{4.16}$$

where p_{norm} is the normalised position value.

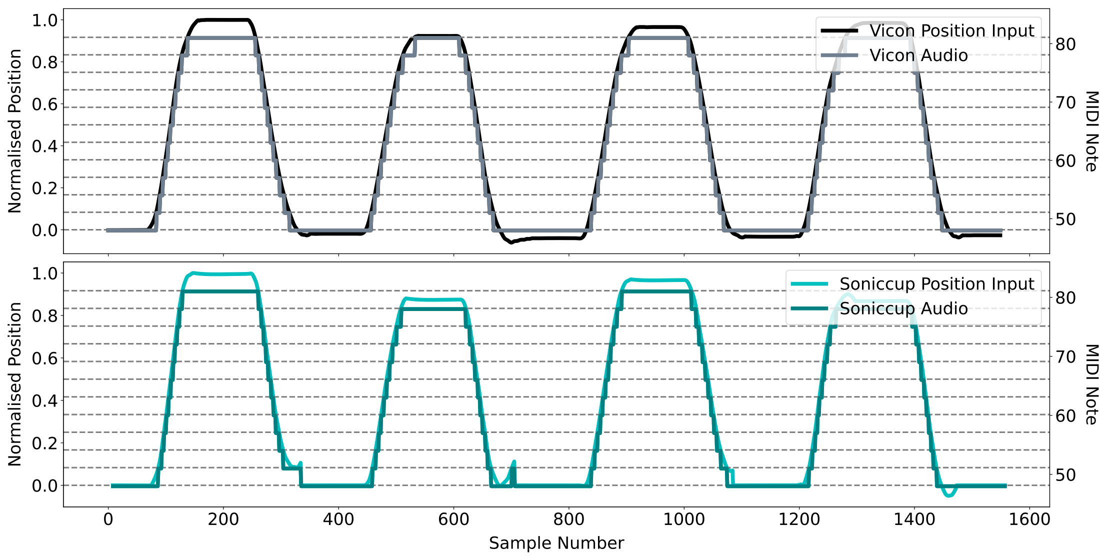

Figure 4.23: Figure showing two plots corresponding to the conversion process from normalised position data to MIDI notes. Top plot shows the input and output for the Vicon system, the bottom plot shows the input and output for the Soniccup.

Equations 4.14 and 4.15 are applied sequentially to obtain the final position estimate, and this is combined with Equation 4.13 to obtain the note output. The effect of saturating the position with Equations 4.14 and 4.15 results in the last note audible note either at 81 or 48 for the top and low ends respectively. Figure 4.23 displays the calculated audible notes at every point in time for four extraction/retraction motions for both the Vicon and the Soniccup, alongside the corresponding relative position, as calculated through Equation 4.16. For Figure 4.23 the plot shows the effect of converting normalised position to audio notes for the Vicon system (top plot), and the Soniccup system (bottom plot), where dotted lines are drawn to show the data cut-off threshold for each audio note. Figure 4.24 shows a plot comparison of the two systems with regards to the normalised position (top plot), and the audio notes (bottom plot).

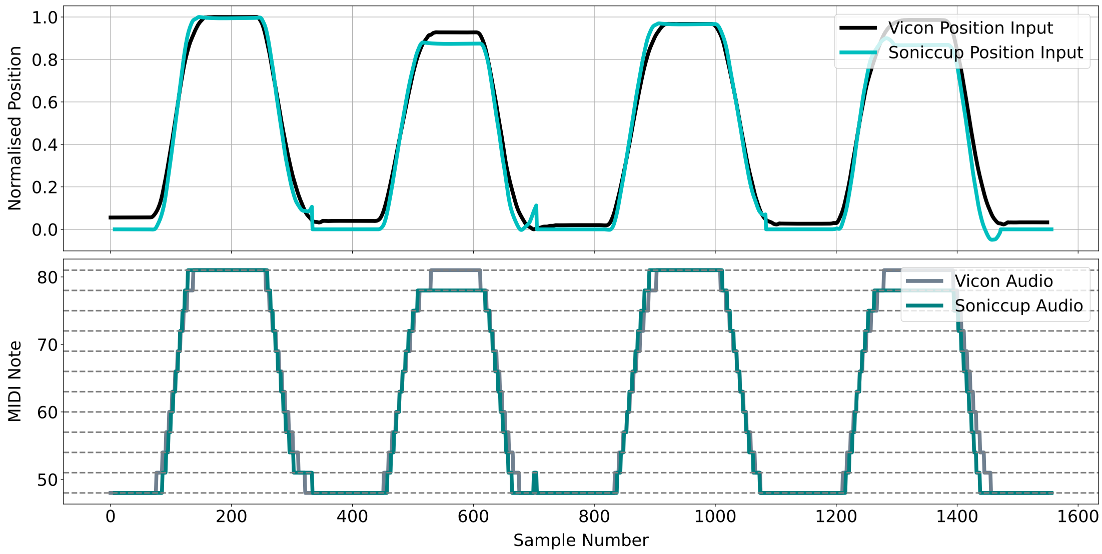

Figure 4.24: Figure showing two plots that provide a comparison between the Vicon and Soniccup systems for the normal speed set. Top plot shows normalised position plots that are used as inputs for the audio conversion process. Bottom plot shows the MIDI notes generated from the normalised position data.

Figure 4.25 displays multiple plots highlighting the impact of increasing the n_{steps} variable through the sonification configuration. As the resolution parameter increases from one to nine (corresponding to (a) and (f) in the figure) the amount of audio notes displayed decreases from 33 to 4, resulting in an increase in pitch interval for each note transition and an increase in the amount of distance required per change of note.

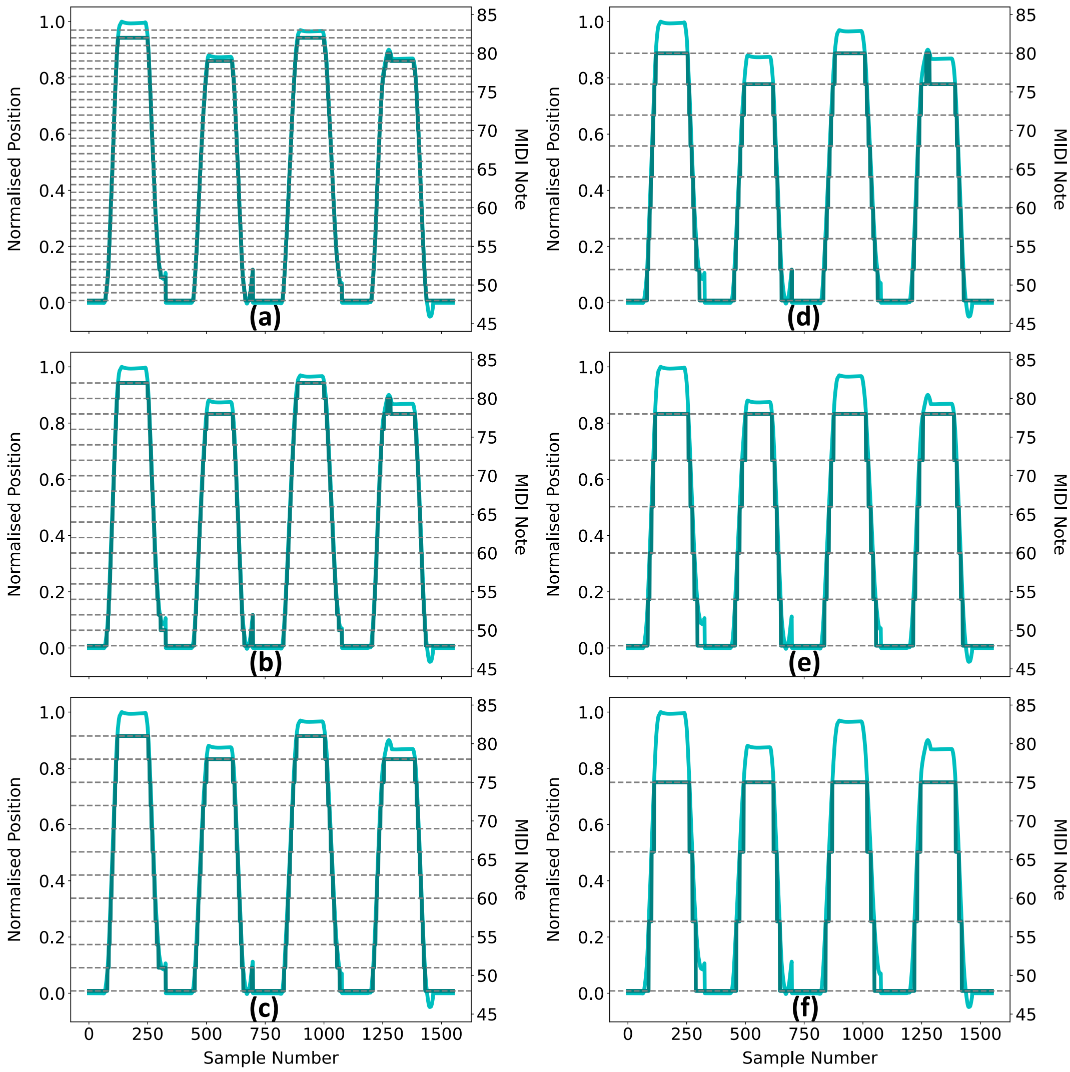

Figure 4.25: Figure showing nine plots corresponding to the effect of altering the audio resolution parameter as stated in Equation 4.12 on the normal speed set. Labels (a), (b), (c), (d), (e) and (f) correspond to the numeric values 1, 2, 3, 4, 6 and 9 used for the n_{steps} parameter.

4.3.6 Achieving Maximum Congruence

To estimate position with maximum congruence to the position data captured through the Vicon system, the sensor streams corresponding to the acceleration and switch state must be temporally aligned at the start of movement. Failure to achieve this produces position estimates through the Soniccup methodology that do not represent the movements performed.

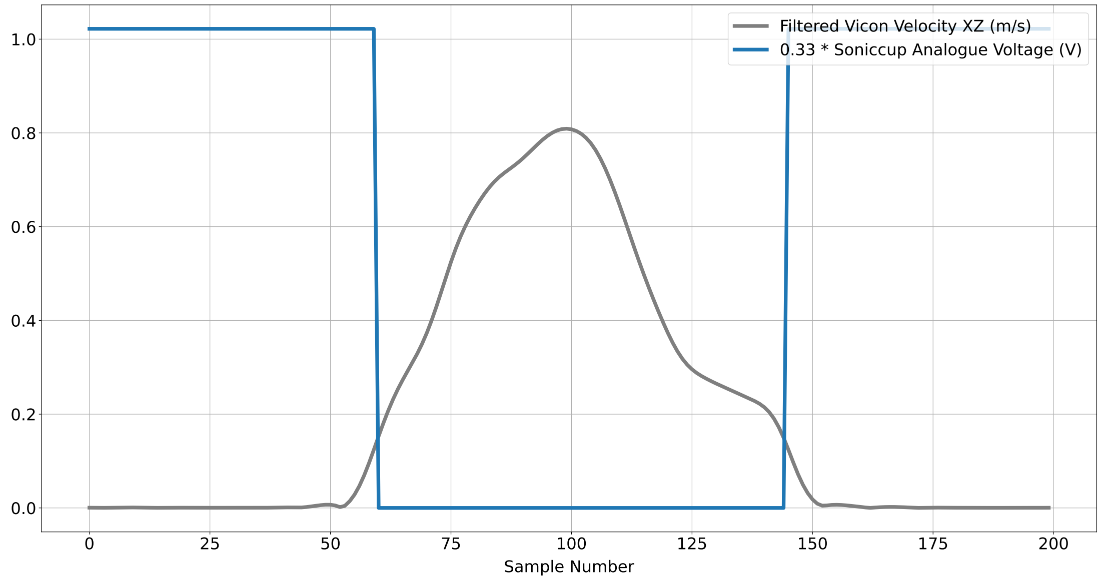

Figure 4.26: Figure displaying the calculated velocity obtained from the Vicon system, and the analogue voltage recordings, corresponding to the switch state, obtained through the Soniccup system, with alignment and magnitude reduction of this trace by a third.

Figure 4.26 shows an example of an extension movement that has been captured through the Soniccup and the Vicon system in parallel. Data from the two systems have been temporally aligned using the same process as described in Section 4.2. The grey trace in this figure shows the magnitude of the hand velocity (combined X and Z axis) obtained via the Vicon system, low-pass filtered at 6 Hz in both directions [275] to create a zero phase output. The blue trace in this figure shows the switch state prior to the filtering process described in Section 4.3.4 and has been amplitude scaled by a third, for illustrative purposes. As can be observed in the figure, at the start of movement the switch transitions from high to low values later than the initial rise of velocity values. This observation is consistent with all movements captured through this process. Addressing this start of movement temporal misalignment is a necessity for the Soniccup methodology.

To remove the effects of temporal misalignment requires the delaying of a signal so that the start of movement at both data streams happens at the same time. To achieve this, the acceleration data stream has to be delayed by X_i samples to match the switch state signal. However, the value of X_i varies for different reaching movements. Understanding the magnitude and range of values of X_i, in-part permits evaluation of the Soniccup.

To obtain an estimate of X_i, all reaching movement data obtained via Section 4.2 for the normal speed set and the slow speed set were analysed to identify start of movement from the Vicon system (velocity magnitude of X and Z axis) and Soniccup system (switch state). The slow speed set represents reaching movements performed at a slow pace, and were included for this estimation process as it was expected that the slower movement resulted in larger temporal misalignment between the two signals, compared to the normal speed set. For data associated with the Vicon system, the first value above a velocity threshold of 0.02 m/s [275] was used to determine start of movement, and for data associated with the Soniccup the first value below a threshold of 0.1 V was used. The extracted sample numbers were then compared. This process resulted in a total of 30 data points for X_i in the normal speed set and 30 data points in the slow speed set. Of the data points obtained in each set of data, statistical outliers were identified using Equations 4.17 and 4.18,

$$LO = Q1 - (1.5 \cdot IQR) \tag{4.17}$$

$$HO = Q3 + (1.5 \cdot IQR) \tag{4.18}$$

which identified one outlier in the slow speed set that was excluded from the analysis. No outliers were identified for the normal speed set. The resulting data for X_i are displayed through Figure 4.27. Figure 4.27 shows the measured variable X_i for each movement within the normal speed set (top) and the slow speed set (bottom), along with a density estimate curve for each. The means and standard deviations for each density estimate curves are shown in Table 4.7.

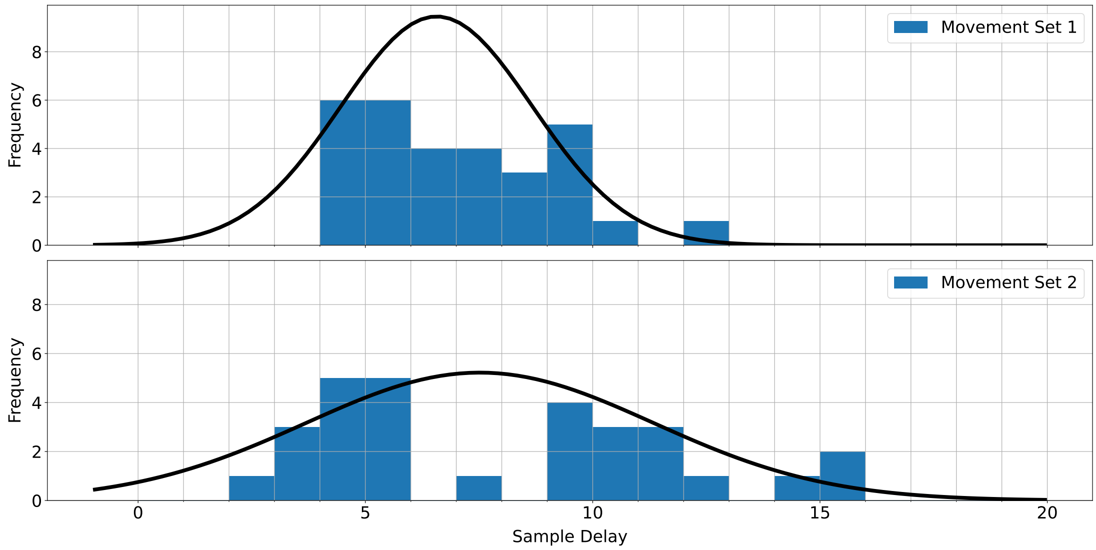

Figure 4.27: Figure illustrating the variation of sample offset between theoretical start of movement, and the change of switch state due to the lifting of the Soniccup, for the normal speed set (top) and the slow speed set (bottom). Density plot is shown in both plots as a black trace. Positive offset indicates that the movement begun before the switch changed state.

Table 4.7: Table showing means and standard deviations for the density estimate curves shown in Figure 4.27.

Movement Extremity	**Movement Set**	**Mean**	**Standard Deviation**
Start of Movement	Normal Speed	6.567	2.108
	Slow Speed	7.517	3.820

In the current implementation of the Soniccup, a design choice was to set a constant system latency. To set this for the Soniccup, X_i must also be modelled as a constant (X_{const}). Setting X_{const} to be a value smaller than the actual delay would result in the zeroing of movement data at the early stages of movement and needs to be avoided for the position estimates to effectively represent the performed movements. Setting X_{const} to be a value larger than the actual delay, results in the inclusion of data prior to the start of movement in the synthesis of position estimation. Given that the ZUPT methodology explicitly zeros these values, the inclusion of these extra values does not affect the position estimates, and therefore, overestimating X_{const} is preferable to underestimating. The compromise however, is that a larger X_{const} value leads to a larger system latency. Assuming that these values form a normal distribution, a value of X_{const} corresponding to 95% of the data values can be calculated using the mean and standard deviation of an attained density estimation curve. For this analysis, the measurements from the slow speed set were used to obtain a conservative estimate of X_{const} via Equation 4.19,

$$X_{const} = \mu + Z\sigma \tag{4.19}$$

where μ is the mean value, σ is the standard deviation unit value, Z is the Z-score corresponding to a level of confidence of 95%. From this equation $X_{const} = 15.004$.

X_{const} has been calculated from data obtained comparing sample numbers corresponding to the start of movement from the switch state data obtained through the Soniccup system, and the *velocity* data obtained via the Vicon system. To convert the calculated measures to compare the switch state and acceleration data streams associated with the Soniccup (X), the calculated mean value, and therefore, the calculated sample discrepancy value, are reduced by one for the start of movement. Equation 4.20 shows the transition from X_{const} to X,

$$X = \lfloor X_{const} - 1 \rfloor \tag{4.20}$$

where the intermediary value has been rounded down to the nearest integer. Therefore, X as calculated through a combination of Equations 4.19 and 4.20 is 14 samples.

4.3.7 System Latency

For a movement sonification system to be considered a real-time system, the latency must be considered. For the Soniccup, a major source of latency comes with the position estimation methodology as described in Section 4.3.4. The described latency is the time between the motion capture measurement and the synthesis of the output position. The cumulative delay as created by factors external to the position estimation methodology are considered to be less than 1 ms, and therefore, have not been included. Figure 4.28 shows a model of the accumulated delay through the stages of the Soniccup position estimation methodology, where one sample corresponds to a delay of 10 ms. Sources of

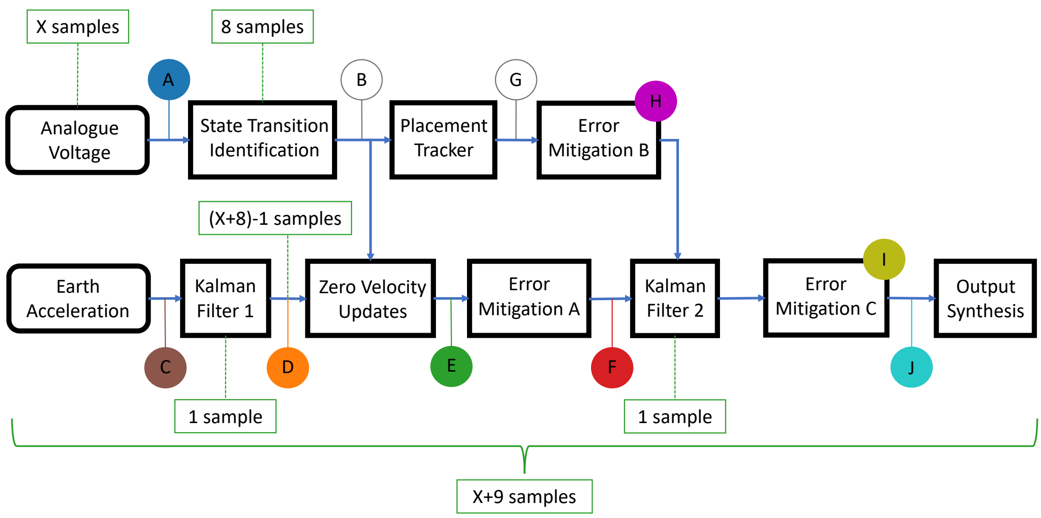

Figure 4.28: Flowchart presenting sources of delay associated with the start of movement, that accumulate during the Soniccup methodology. This figure shows Figure 4.20 with added annotations corresponding to sources of delay. Total delay is shown through 'X+9 samples', and one sample worth of delay equals 10 ms.

delay include the filtering mechanism to remove momentary changes of state, resulting in a delay of eight samples, and the integration mechanism performed through the two Kalman filters, which induce a delay of one sample each. To achieve maximum congruence with the measured position from the Vicon system, an additional delay of 14 samples is included into the model as described in Section 4.3.6, to align the input data streams at the start of movement. Subsequently, to temporally align the signals for the ZUPT methodology, a delay mechanism for the acceleration data stream of '(X+8)-1 samples' is included in this model.

The total latency of this system is an accumulation of three sources as shown through Figure 4.28:

- 'X samples' shown above Analogue Voltage

- ‘8 samples’ shown above State Transition Identification
- ‘1 sample’ shown below Kalman Filter 2

resulting in a total of ‘X+9 samples’ as shown in the bottom annotation of the figure. As such for the Soniccup methodology as presented, the current system latency is calculated to be 23 samples, which corresponds to a delay time of 230 ms from the algorithm receiving sensor inputs, to obtaining position estimates.

4.4 Motion Capture Through a Camera

In parallel to the development of the Soniccup, a second movement sonification system was developed using a Microsoft Azure Kinect camera as the motion capture component. Predecessors of this type of camera-based technology, as identified in Chapter 3, have been used in many movement sonification projects, and as such induces confidence that the motion capture technology can be integrated into a movement sonification system to produce reliable outputs. Creating a movement sonification system alternative to the Soniccup, using a different motion capture technology, improves the chances that at least one system would be a suitable candidate as a rehabilitation tool. This section starts with an introduction into how motion capture can occur through camera-based technology, before proceeding to the development of a movement sonification system named ‘KinectSon’.

4.4.1 Computer Vision

Computer Vision (CV) is the field of study that enables computers to process, analyse, and interpret digital images, with the goal of automating tasks that could be performed by the human visual system. In context of human motion capture, the goal of CV is to identify human anatomical segments of interest and temporally monitor the difference in location of these anatomical segments between subsequent captured images, with example applications including sports performance [105], biomechanical evaluations [284], and human-computer interactions [221]. Human motion capture through CV requires a camera-based technology to capture and extract human anatomical landmarks in three dimensions during operation. These landmarks are used to formulate a representation of the human tracked. This is typically achieved through a model-based approach, of which one of either volumetric, contour-based, or skeleton-based models, are chosen.

To extract the position of these landmarks, a choice can be made between marker-based and markerless camera-based systems. Commercial marker-based systems typically use multiple cameras, infra-red light, and markers physically attached to the object of interest, to obtain 3-dimensional position. Although systems that are reported to be the 'gold-standard' of motion capture use this type of methodology [280,281], there are limitations with the use of such systems for human motion capture, including the violation of the rigid body assumption that is used in analysis (as markers are attached to the skin), long preparation time, and the physical restraints imposed by the markers, leading to compromised movement execution [285].

Commercial markerless CV systems make use of human pose detection libraries, such as OpenPose [286], as part of their image processing strategy to identify human anatomical segments for motion tracking; however, this strategy only provides 2-dimensional data, in absence of depth measurements. Data corresponding to depth, can be acquired from two types of methods, triangulation-based methods [287] and the Time-of-Flight (ToF) principle [288, 289].

Triangulation based methods can be split into passive, and active techniques. Passive triangulation techniques, such as stereovision, uses multiple cameras to observe a scene at different viewpoints. This setup obtains different images containing corresponding points, of which depth data can be extracted. Examples of commercialised cameras that utilise stereovision include the Stereolabs Zed series [290] and Intel RealSense Stereo Depth Camera series [291]. Active triangulation techniques, such as structured light sensing, utilises a light source to project a known illusionary texture onto the scene. The projected texture is altered due to the distance and shape of the objects contained in the scene, of which a camera captures the altered texture, and depth data is extracted from the calculated differences in texture [292]. An example that uses structured light is with the TrueDepth camera that is a part of Apple's 'Face ID' technology [293].

The ToF principle is a time-based method that uses a light source and a receiver to measure depth within a scene. Light emitted onto a scene is reflected back in different segments based on the objects in the scene, and the relative time difference recorded between the different segments is used as a form of measurement that corresponds to depth [294]. Examples of commercialised cameras that use ToF sensing include the LIPSedge DL series [295] and VZense series [296]. Combining depth data with 2-dimensional coordinates permits 3-dimensional motion capture through markerless CV.

The use of commercial markerless systems for human motion analysis are of interest for rehabilitation, as the absence of physical markers and reduction of cameras required, shortens the preparation time and reduces the acquisition cost for implementation. Taken from Chapter 3, two commercially available markerless CV technologies found in the literature that have been used in existing movement sonification systems to track the upper limb, the LEAP Motion Controller, and the Kinect systems.

The LEAP Motion Controller is a camera-based sensor developed to track the hands of a person. The camera uses infra-red emitters and cameras, along with algorithms for hand pose estimation, named Ultraleap Gemini in the most recent implementation. The technology can be used as a standalone, or in combination with Virtual Reality (VR) to create a visually augmented environment. As mentioned in Chapter 3, a key limitation of the technology is the limited field of view of the sensor, restricting the range of movement available for interactive feedback on movements, and therefore, was not chosen for this application.

The Kinect systems associated with Microsoft are popularly used motion capture devices. Originally released in 2010, Microsoft Xbox Kinect for Xbox 360 was a periphery device, providing an alternative interactive controller for entertainment purposes. This commercialised technology achieved 3-dimensional motion capture through the use of structured light sensing. The novelty of this device, along with the creation of associated SDK, attracted developers and researchers to apply this device to alternative human motion capture applications. Inline with the release of a new entertainment gaming console from Microsoft in 2013 a newer version of the Kinect, titled Microsoft Xbox Kinect for Xbox One ,was released with improved functionality to its predecessor, but with the technology of the ToF principle to achieve it. These technology systems have been used to capture human movement as part of movement sonification systems, referred simply as 'Kinect' or 'Kinect-based' system. Singh *et al.* used it as part of their created 'Go-with-the-Flow' system for chronic pain management [121], Yang *et al.* used the technology to monitor repetitive arm movements [236], and Wang *et al.* applied the technology as part of a system to enhance awareness of movement [198]. However, both of these Microsoft Xbox Kinect products have been discontinued since 2017, and as such would require second-hand acquisition. The most recent successor is the Microsoft Azure Kinect Development Kit

Figure 4.29: Image of the Azure Kinect SDK camera.

(DK). An image of the associated motion capture camera is shown in Figure 4.29. The technology was released in 2019 aimed at applications outside of entertainment, and is currently priced at £355.00. The device contains a 1-megapixel depth sensor, 12-megapixel RGB camera, an infra-red emitter, inertial sensors, a 7-microphone array, and external synchronisation pins to synchronise multiple devices. Among the commercial tools that have been created to work with the system, a Body Tracking SDK has been developed that uses a skeleton-based model consisting of 32 joints, emanating from the estimated 'pelvis' joint. Figure 4.30 illustrates the joint hierarchy of this generative model [1]. The technology has been used in multiple human motion capture applications as reported in the literature, including gait analysis [297], hand posture tracking [298], and an upper limb rehabilitation programme [299].

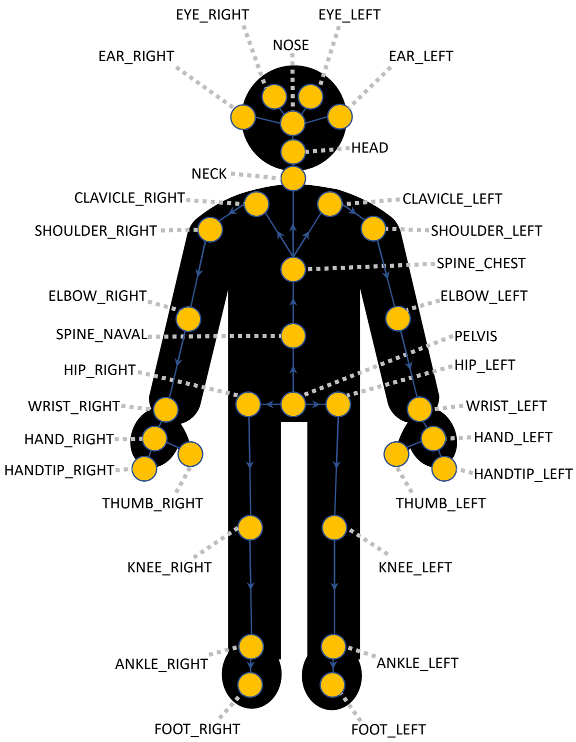

Figure 4.30: Skeleton-based model used by Azure Kinect Body Tracking SDK [1].

Given the capabilities of the technology as described from the literature, the Microsoft Azure Kinect DK along with the Body Tracking SDK was chosen to create a movement sonification system to provide online audio feedback on upper limb movement. Upon commencement of this work, the following research questions were formed:

RQ7 What steps are required to create real-time audio feedback via the motion capture capabilities of the Kinect?

RQ8 What are the limitations of using the Kinect and the associated Body Tracking SDK as a movement sonification system?

4.5 KinectSon

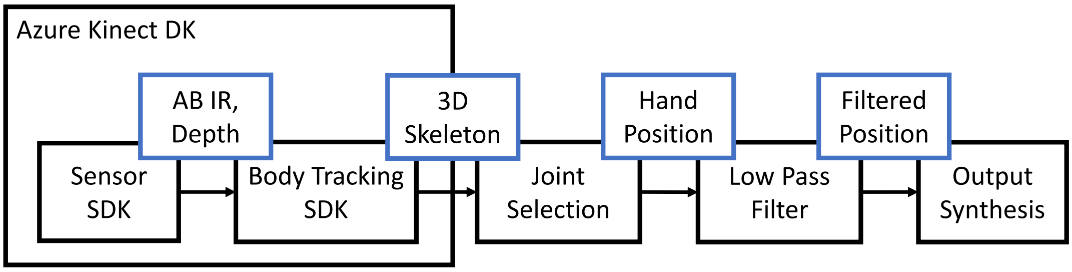

Figure 4.31: Flowchart showing the steps used to create a prototype real-time movement sonification system using the Azure Kinect DK.

This section describes the composition of the movement sonification system termed 'KinectSon'. The Microsoft Azure Kinect camera was connected to a PC with an attached NVIDIA Titan RTX Graphics Processing Unit (GPU), allowing the accompanied Body Tracking SDK to operate at 30 Hz. Camera settings used for data capture are described in Section A.3. Figure 4.31 shows the stages of the KinectSon system. Data capture is achieved through the associated Sensor SDK, and processed through the Body Tracking SDK, to create a skeleton-based model. Subsequently, positional coordinates of the anatomical joint/landmark of interest were extracted and low-pass filtered, before sent through to the sonification module for output synthesis.

The addition of the low-pass filter to the KinectSon system has been implemented to address the errors associated with the depth sensor enclosed in the device. These errors can be segregated into two types, systematic error and random error [300]. Systematic error is the measurable difference between the recorded values and ground truth. Random error is caused by 'shot noise', and is dependent on the number of photons hitting the sensor, which can be calculated through Equation 4.21, where E_r is the random error, d_t is the depth data captured at time t, N is the number of depth measurements, and $\overline{d}$ is the computed mean depth value [300].

$$E_r = \sqrt{\frac{\Sigma_{t=1}^{N}(d_t - \overline{d})^2}{N}} \tag{4.21}$$

The effect of these error types (most notably shot noise) on the data can lead to the sonification of additional frequency components beyond that of human movement, especially when the tracked hand is stationary, and hence impact the output of the system. Filtering would lead to a greater signal to noise ratio as an input for sonification, however, to retain a linear phase response of the filter, a Finite Impulse Response (FIR) filter is required. A limitation with this type of filter is the induced delay time, which can be calculated through Equation 4.22,

$$\tau = \frac{N-1}{(2 * F_s)} \tag{4.22}$$

where τ is the time delay, N is the number of taps, and F_s is the sampling frequency. To maintain a system latency of <100 ms, and therefore, retain real-time performance of the movement sonification system, a 5th order low-pass FIR filter at 1 Hz cutoff was implemented.

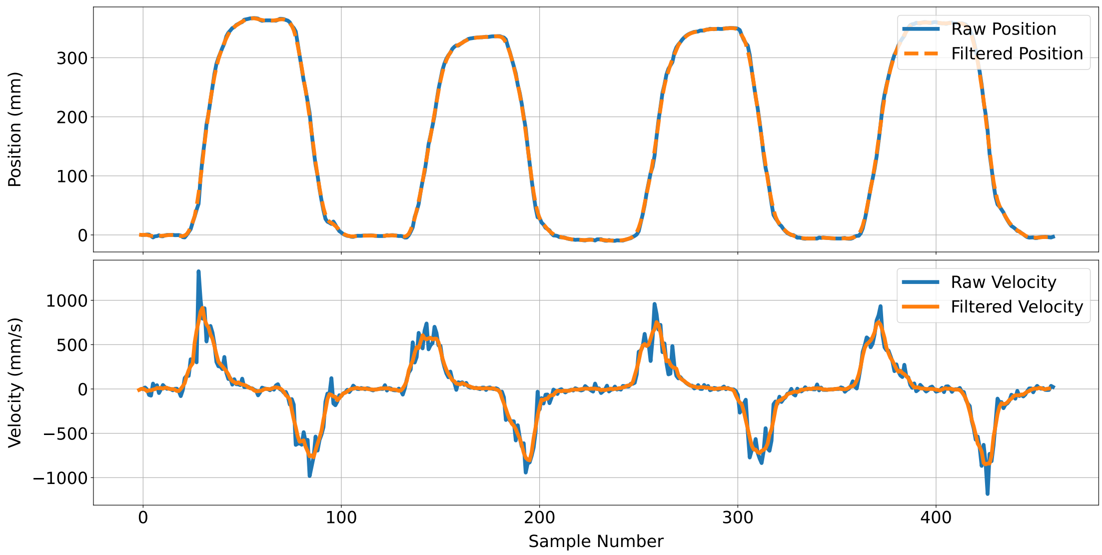

Figure 4.32: Figure illustrating four forward reaching movements captured through the KinectSon system, where in each plot the blue trace shows data before low-pass filtering, and orange shows the resulting data after low-pass filtering. Top plot shows position, bottom plot shows derived velocity.

The effects of the filtering can be seen through Figure 4.32, which presents the same four normal speed reaching movements as captured in Section 4.2, that has been captured by the Kinect camera. This figure presents two plots, the top plot shows position data before and after applying the low-pass filter, the bottom plot shows the derived velocity using the same position data. The figure shows the minimal impact of the filter on position data, but saliently filters out high frequency components to output velocity data with less variability. The inclusion of this filter potentially reduces the trill effect that would occur due to the effect of noise oscillating across an audio note boundary. Figure 4.33 displays the calculated signal delay before and after filtering, using Equations 4.9 and 4.10 as described in Section 4.2.

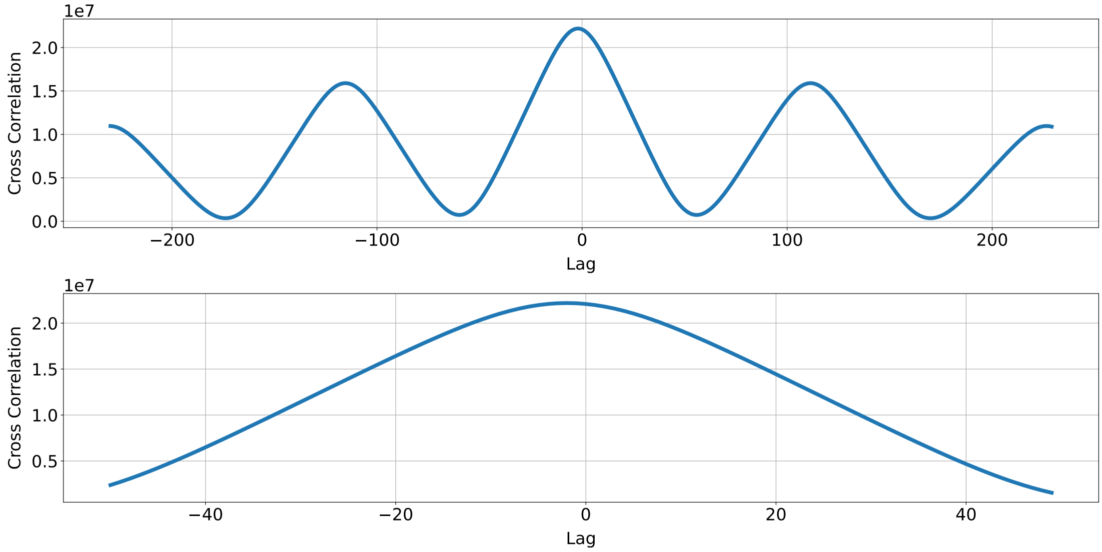

Figure 4.33: Figure showing two plots showing the resulting delay caused by the low-pass filter. Top plot shows the shows the result of a cross-correlation process between the position data before and after the application of the two data streams. The bottom plot zooms in at the peak of this result to show a sample offset of two samples for maximum correlation.

The value of signal lead/lag obtained through the low-pass filter is shown in the bottom plot of Figure 4.33 to be two samples, amounting to a temporal delay of $66.\dot{6}$ ms. The filtered data then proceeds to the output synthesis step.

To create audio output, position data obtained from the KinectSon system was used as input into the same configuration strategy described in Section 4.3.5. To achieve this for the KinectSon system, prior to movement sonification operation a calibration process was completed involving the movement performer placing and holding the tracked hand at the starting position, and then maximum reach, for 100 samples each. The mean average positional value from these two positions is used for p_{min} and p_{max} in Equation 4.11, respectively.

4.6 GUI and Sonification Module

For both the Soniccup and KinectSon systems, a GUI created through PyQt [301] was developed to provide a visual display of the data, and provide options for sonification. Figures 4.34 and 4.35 shows the GUI used with the Soniccup and KinectSon systems, respectively.

In addition, a sonification module was developed through Python, using Pyo from AJAX Sound Studio [302], a module that permits the creation of sounds through signal processing and audio synthesis. The current sonification implementation of the Soniccup and KinectSon systems uses an audio synthesizer that alters note pitch as dependent on the value of the input data. Data obtained from the associated motion capture system were normalised to a minimum-maximum range of [0,1] and linearly converted into MIDI notes as described in Section 4.3.5, before conversion to a frequency value through Equation 4.23,

$$f = 440 \cdot (2^{\frac{n-69}{12}}) \tag{4.23}$$

where n is the MIDI note value, and f is the corresponding frequency. The frequency value is subsequently used to create an audio byte through an audio synthesizer within Pyo. Implementation details for the sonification module are shown in Section A.4.

4.7 Comparison Study

Sections 4.3 and 4.5 describe two separate movement sonification systems that are potential candidates as a rehabilitation tool. Both systems provide position data tracking the hand, that can be used as an input into a sonification conversion module. To evaluate if the systems are fit-for-purpose, a comparison study was completed using

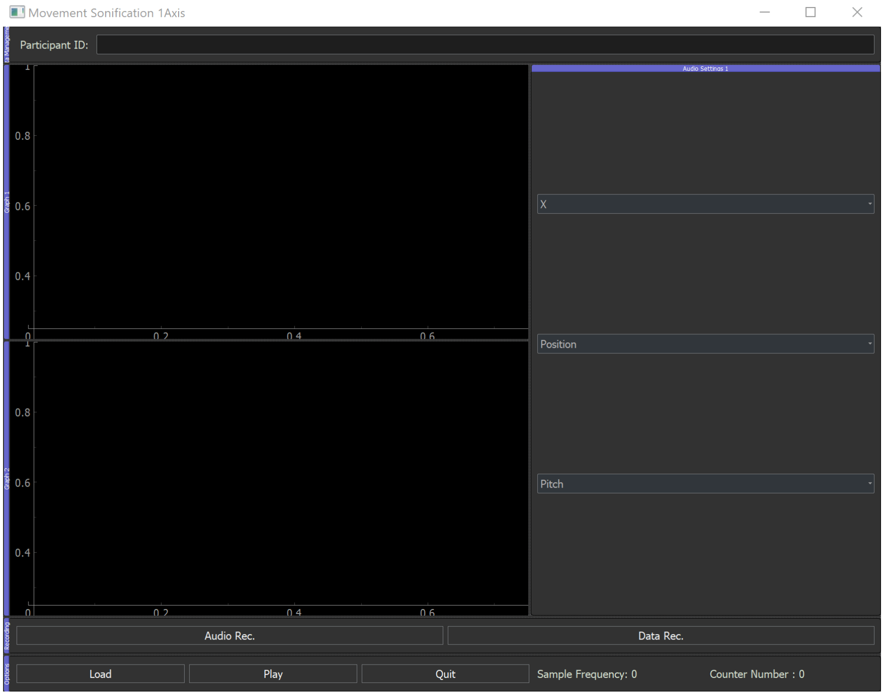

Figure 4.34: Figure displaying GUI used with the Soniccup system. Options include axis data, movement kinematic, and sonification configuration to use.

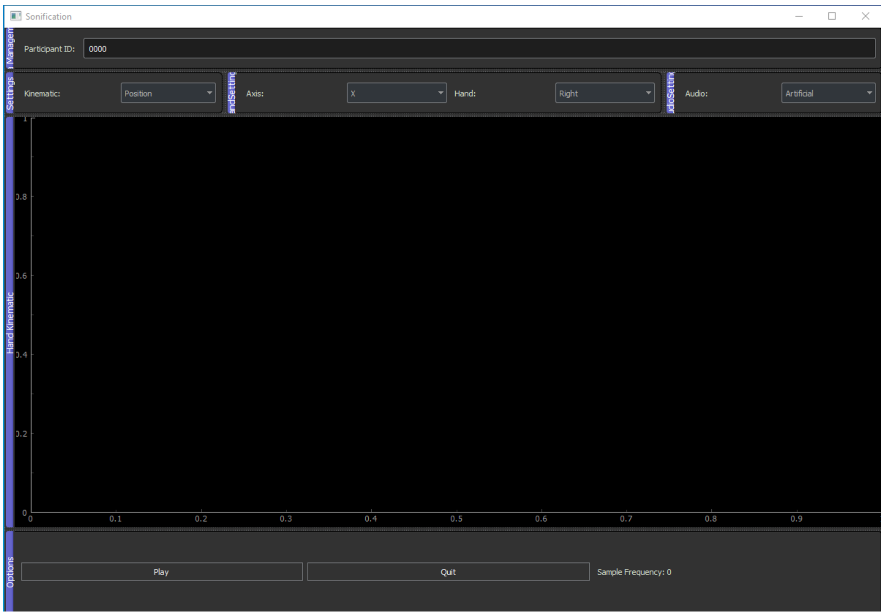

Figure 4.35: Figure displaying GUI used with the KinectSon system. Options include axis data, movement kinematic, hand to track, and sonification configuration to use.

the Soniccup and the KinectSon system to measure the similarity of the position values compared to the associated measurements obtained from a Vicon Nexus system. The Vicon system, as previously described, is a gold-standard system for motion capture and was assumed to be the ground truth in this study for comparison purposes. As sonification is a process of converting data into an audial display, the minima and maxima of the measured signal can be set to an arbitrary minimum and maximum of the auditory dimension, therefore, similarity of the position values is compared, as opposed to the actual position values. This study is separate to the experiment described in Section 4.2, but used the same procedure to capture movement, as such this section extends from Section 4.2 by describing the methodology for data capture, the proceeding data analysis procedure to obtain similarity measurements and the results of the study. The following research question moulded the work completed in this section:

RQ9 For an arbitrary sonification configuration using position estimates, how similar will the auditory displays obtained from the Soniccup and KinectSon, be to the auditory display from a gold-standard motion capture system?

4.7.1 Procedure

A table and chair were set up in the middle of the Vicon system tracking space at the Wolfson Centre, University of Strathclyde, with the Kinect system setup adjacent to the target tracking space. The Kinect camera faced towards the table, so that it captured a side view of the participant in the tracking area. Calibration procedures for each system were completed prior study commencement. A volunteer sat at the table with the Soniccup positioned in front of them on the table at the closest table edge. With all systems online, the volunteer performed three sets of 15 reaching movements with the Soniccup in hand, using their dominant hand. Each reaching movement consisted of simultaneously raising the Soniccup from the table whilst extending their arm to an approximate full reach, before placing the Soniccup on to the table. The Soniccup was then simultaneously raised whilst the arm retracted back to the starting position, and then placed on the table. For the first set of movements the volunteer was instructed

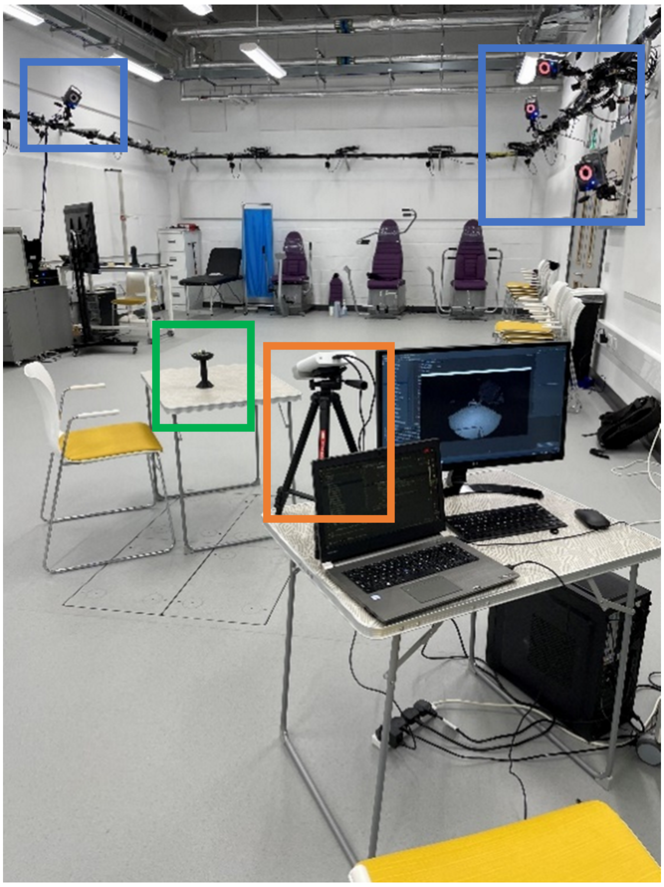

Figure 4.36: Initial investigation setup showing apparatus. Green box shows position of Soniccup placed on a table in front of the participation area. Orange box shows position of the Kinect camera, positioned and oriented to capture upper limb movements of the participant. Blue boxes show some of the cameras in the Vicon Nexus system, which captured the position of a retroreflective marker placed on top of the Soniccup.

to perform movements at a normal speed, for the second set of movements at a slow speed, and for the third set of movements at a fast speed.

4.7.2 Data Processing and Analysis

Parameter mapping sonification, chosen for this project, used an arbitrary scaling when converting from data to sound, as such the relative values of the data, corresponding to the shape of the data, are of importance. For this comparison study, the similarity of position data obtained through the Soniccup and KinectSon systems are compared to the data obtained through the Vicon system, to evaluate the motion capture capabilities of the two systems as part of a movement sonification system, with the Vicon system assumed to be the ground truth.

Interpolation and Sample Alignment

The Kinect system had a sampling frequency of 30 Hz, whereas the Soniccup and Vicon systems had a sampling frequency of 100 Hz. As such, data obtained from the Kinect system were up-sampled to 100 Hz using a linear interpolation method. Data from the three systems were temporally aligned through the use of cross-correlation calculations and verified through observation of drawn plots.

Data Processing

Using the pressure values from the Soniccup as a reference, periods of movement and non-movement were labelled onto data associated with each motion capture system. Identified periods of non-movement were removed from the captured data to produce temporally aligned movement data. As axis alignment of the systems for this study was completed manually, data obtained from the motion capture systems were converted to radial distance through a process of Equations 4.24 and 4.25. Equation 4.24 was used to align the origin of data from each system,

$$\begin{aligned} X_c &= X_j - X_0 \\ Y_c &= Y_j - Y_0 \\ Z_c &= Z_j - Z_0 \end{aligned} \tag{4.24}$$

where X_j, Y_j, Z_j are data associated with the X, Y, and Z axis at sample j, X_0, Y_0, Z_0 is the first value in the corresponding axis, and X_c, Y_c, Z_c is the corrected data. The radial distance r is then computed through equation 4.25.

$$r = \sqrt{X_c^2 + Y_c^2 + Z_c^2} \tag{4.25}$$

Following this conversion, the data is normalised to a maximum of one, through Equation 4.26.

$$r_{norm} = r/r_{max} \tag{4.26}$$

Both r and r_{norm} are subsequently processed through Equation 4.27 to calculate similarity through a Mean Squared Error (MSE) approach, where n is the total number of samples, d corresponds to the data used i.e. r or r_{norm}, and i is the ith sample of data.

$$MSE = \frac{1}{n} \sum_{i=0}^{n-1} (d_i - \hat{d}_i) \tag{4.27}$$

4.7.3 Results

Table 4.8 presents statistics obtained from the study. Included in this table are average mean, standard deviation, and cumulative, MSE metrics for three sets of movement data performed at normal (Movement 1), slow (Movement 2), and fast (Movement 3) speed. To evaluate the similarity of the data obtained from the motion capture systems, statistical metrics were performed on absolute value data, and data through a normalisation process. Considering the application that these motion capture systems are used for, results will focus on metrics obtained for data that has been normalised. Figure 4.37 presents a visual display of the MSE of normalised data versus the peak movement velocity of that movement for data associated with the Soniccup and KinectSon systems. Figure 4.38 presents a visual display of the same data for the KinectSon only. Boxes are drawn and labelled on both figures to show the associated movement set that the data points belong to.

Within this study, the KinectSon system has lower MSE values relative to the Soniccup system for the same movement. The MSE values of normalised data for the data associated with the KinectSon range from 0.0001 - 0.0030 with an overall average mean of 0.0012 (rounded to two s.f.), and the same measurement for the data associated with the Soniccup range from 0.0006 - 0.2247 with an overall average mean of 0.0037 (rounded to two s.f.). These ranges translate to lower MSE for the

Table 4.8: Table of key results obtained from the study.

Movement Set	**Normal Speed (1)**		**Slow Speed (2)**		**Fast Speed (3)**	
Mean (SD) Movement Duration (s)	0.91(0.08)		2.94(0.37)		0.53(0.04)	
Mean (SD) Peak Speed (mm/s)	796.17(68.01)		283.69(23.40)		1705.35(139.26)	
Technology	**Soniccup**	**KinectSon**	**Soniccup**	**KinectSon**	**Soniccup**	**KinectSon**
Average Mean of MSE of Normalised Data	0.0034	0.0008	0.1030	0.0012	0.0057	0.0015
Standard Deviation of MSE of Normalised Data	0.0019	0.0004	0.0640	0.0007	0.0058	0.0008
Accumulation of MSE of Normalised Data	0.1005	0.0229	3.0893	0.0348	0.1697	0.0437
Average Mean of MSE (mm^2)	7413.21	707.28	32994.65	789.24	10314.38	1174.40
Standard Deviation of MSE (mm^2)	2838.81	199.80	13644.69	186.35	5494.36	394.57
Accumulation of MSE (mm^2)	222396.22	21218.51	989839.36	23677.09	309431.51	35231.98

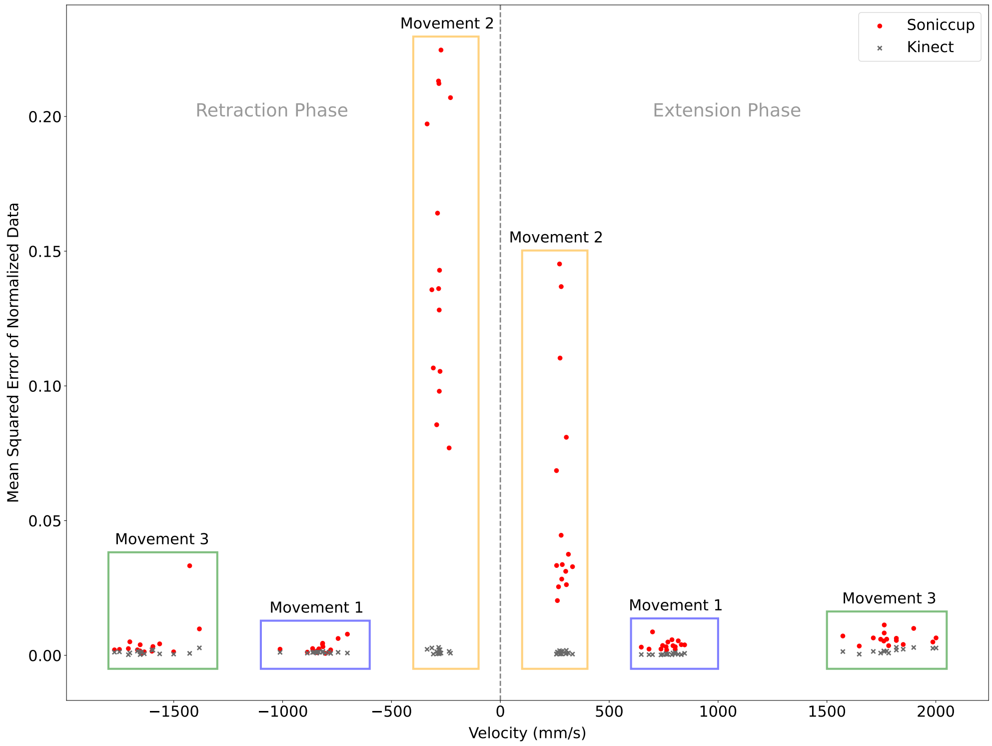

Figure 4.37: Scatter plot presenting the calculated MSE of the Soniccup and KinectSon for each movement. Data points with a positive velocity are from the extension phase of the reaching movement, whilst data points with negative velocity correspond to the retraction phase of the reaching movement. Red dots correspond to movement captured through the Soniccup, grey crosses correspond to movement captured through the KinectSon. Boxes enclose segments of the plot, and labelled with association with a movement set, 'Movement 1' corresponds to normal speed movement, 'Movement 2' corresponds to slow speed movement, 'Movement 3' corresponds to fast speed movement.

KinectSon in comparison to the Soniccup. With focus on the results per movement set, calculated statistical metrics for the Soniccup and Kinect have similar values for Movement Set 1 and Movement Set 3, for the majority of movements. This observation can be seen through Figure 4.37. For Movement Set 2, referring to data associated with the Soniccup only, the associated statistical MSE metrics indicate that the data is dissimilar, represented by the average mean of MSE equalling 0.1030 which is two orders of magnitude greater than the next greatest average mean of MSE (0.0057). The KinectSon system, however, retains similar statistical metrics to the other movement sets. To validate these findings, Figure 4.39 shows two plots of the first four movements

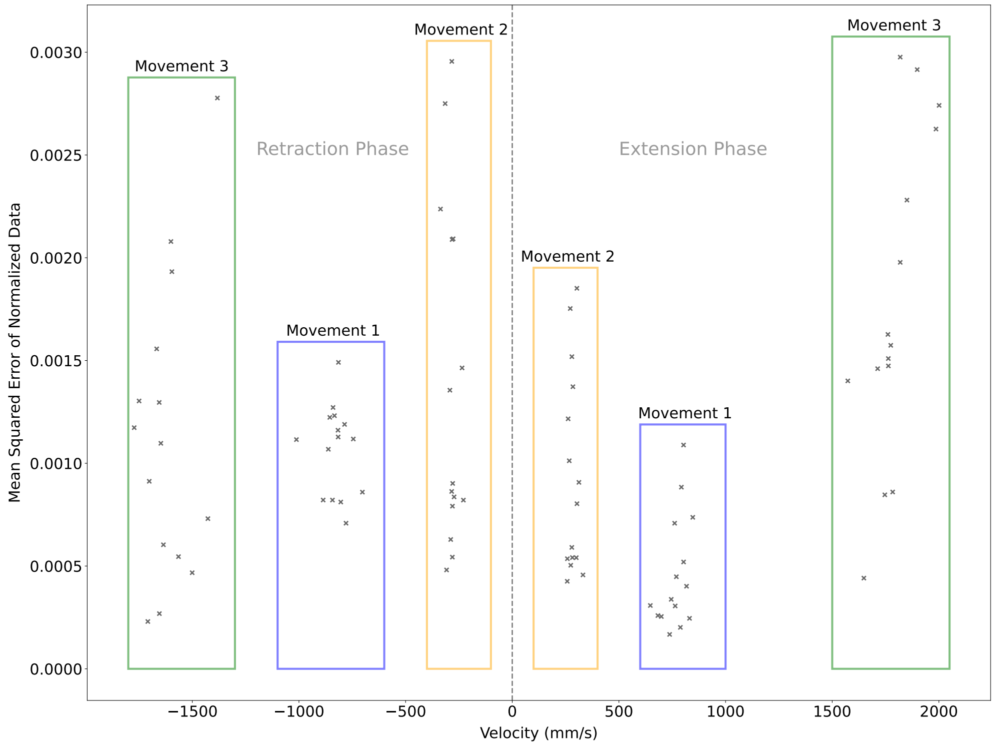

Figure 4.38: Scatter plot presenting the calculated MSE of the KinectSon for each movement. Data points with a positive velocity are from the extension phase of the reaching movement, whilst data points with negative velocity correspond to the retraction phase of the reaching movement. Boxes enclose segments of the plot, and labelled with association with a movement set, 'Movement 1' corresponds to normal speed movement, 'Movement 2' corresponds to slow speed movement, 'Movement 3' corresponds to fast speed movement.

captured by each system, and provides a visual representation of the performance of the Soniccup (top) and KinectSon (bottom) systems regarding this movement set. The effect of the normalisation process (see Equation 4.26) on the Soniccup data can be seen saliently in this figure, as the trace associated with the Soniccup does not reach a value of 0.7. This is due to two reasons, firstly the value of r_{max} exists on a movement after the first four movements, and secondly the relative values obtained through the Soniccup system are highly variable between each movement. The traces corresponding to the Vicon and Kinect systems do not show the same issue in this figure, as the data obtained through these systems have much less variance between movements. These observations help to explain the comparatively large MSE for normalised data values measurement

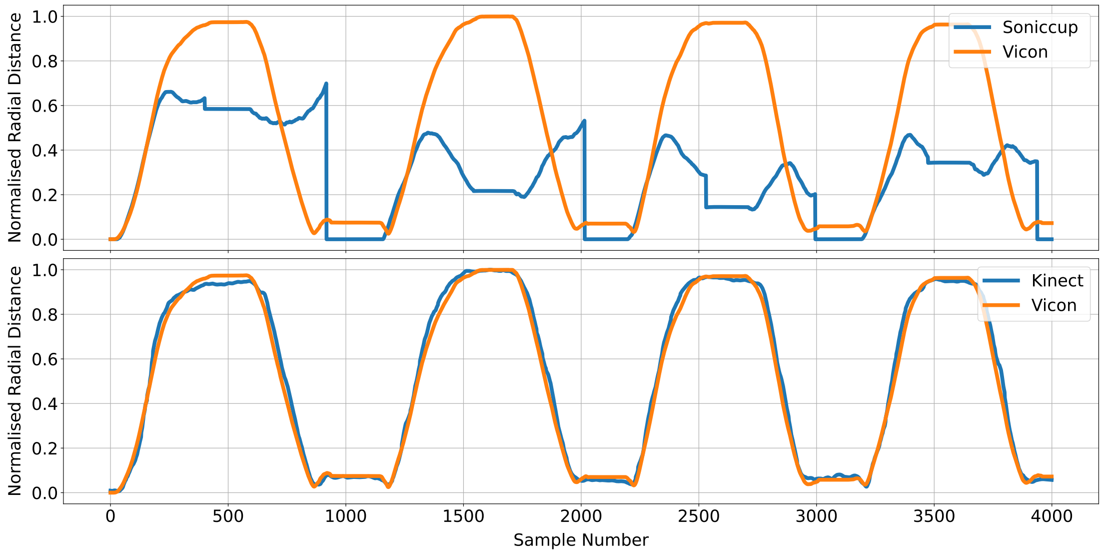

Figure 4.39: Figure presenting the radial distance obtained through the motion capture systems of the Soniccup, KinectSon, and Vicon systems, for four movements within Movement Set 2. Top plot shows traces associated with the Soniccup (blue) and Vicon (orange), bottom plot shows traces associated with the KinectSon (blue) and Vicon (orange). Data associated with each trace has been normalised so that the maximum data point in the 15 captured reaching movements is equal to one, resulting in the trace associated with the Soniccup showing all data points in the first four reaching movements to be less than 0.7.

for the Soniccup compared to the KinectSon system. The MSE for normalised data values for each movement shown in Figure 4.39 captured by the Soniccup are 0.0337, 0.1361, 0.1452, and 0.2122, whereas the same metric for the KinectSon are 0.0014, 0.0030, 0.0018, and 0.0021.

4.7.4 Discussion

To identify suitability of the Soniccup and KinectSon systems to sonify movements performed by people with a range of functional competency, and in turn to identify potential suitability for audio feedback based rehabilitation, a study was completed analysing the position data obtained through the two prototype systems, and comparing them to position data obtained through a Vicon Nexus system. For this study a volunteer completed slow, normal, and fast speed, forward reaching movements. As the values obtained within each movement sonification system would be converted into an auditory display, similarity metrics were used.

As part of the study methodology the movement performer was instructed to perform 15 forward reaching movements at slow, normal, and fast speed, resulting in an average mean movement speed of 283.69 mm/s, 796.17 mm/s, and 1705.35 mm/s, respectively. As visually presented in Figure 4.37, the speed of the performed movements was reasonably consistent for each instruction. As such, current conclusions are based on a dataset that has three distinct ranges of speed. From observing the average mean values, the KinectSon system produced a positional output of greater similarity to the Vicon system in comparison to the Soniccup system, for every speed of movement. The performance of the KinectSon was also more consistent than the Soniccup, as shown by the lower calculated standard deviation. For the Soniccup, although the MSE statistical metrics are larger than the KinectSon, results indicate relatively strong similarity with the Vicon system for forward reaching movements performed at normal to fast speed. This can be observed by the overlap of scatter plot points in Figure 4.37 within boxes labelled ‘Movement 1’ and ‘Movement 3’. However, the performance of the system decreased with regards to slow movements, as represented by the greater average mean of MSE and greater standard deviation of MSE showing that the system in its current iteration is not suitable to capture forward reaching movements for slow speed performers. A primary reason for this decreased performance in the slower movement is through the extended time periods between each placement of the Soniccup object, and hence extended periods of time between activation of the ZUPT method. These results corroborates with the findings of Bai *et al.* [303] that highlight the effectiveness of the technique provided that regular zero velocity periods can be identified.

To conclude, the KinectSon has shown a greater similarity to the Vicon system for obtaining position data, than the Soniccup system. Although the Soniccup system has shown results to indicate similar output to the Vicon system for normal to fast movements, the results have shown that for slow movements, the position estimates during slow movement are strongly dissimilar, and as such converting this data to the audio domain would result in a sonification output that would appear to be unrelated to the movement performed. Further work is required to improve the slow movement performance of this system before it can be considered as a rehabilitation tool.

4.8 Summary

The chapter has described the creation of two proof-of-concept movement sonification systems, the Soniccup and the KinectSon systems. The Soniccup uses a combination of a single IMU, a Double Pole Double Throw (DPDT) momentary switch, and a 3-dimensional printed object to house the system, that allows users of the system to perform reaching movements and obtain online audio feedback based on the relative position of the forward reach. The system makes use of the ZUPT method to correct position estimates at the beginning and end of each extension and retraction phase, when the Soniccup object has been placed. As verified through the comparison study with the Vicon system the position estimates, and therefore, audio output, obtained through the Soniccup is representative of the movements performed, provided that the movement performed is above a certain speed. As such a limitation of the system is the performance at slower speeds, corresponding to movements with a greater duration of time, resulting in position estimates that do not represent the movements performed, which limits the applicability of the system as a rehabilitation tool for those with slower movement (for example, people with upper limb paresis [275]). An additional limitation is with the perceivable delay time from movement to sound. The theoretical delay to generate position output, as calculated by the time required to identify non-momentary changes of switch-state and the start of movement sample discrepancy between the Earth acceleration and switch state, in this system design is in excess of 230 ms which is too large to be considered real-time.

To answer research questions RQ4 - RQ6 stated in Section 4.3, the Soniccup was designed to include a push-to-make switch implemented at the base of a handheld object to detect placement and pickup events during movements. The sensor was used to detect periods of placement, and therefore, zero velocity, so that the ZUPT method could be implemented to condition velocity estimates. As a rigid horizontal surface is required to compress the switch on placement, the created system extends implementations of the ZUPT for upper limb tracking found in the literature (RQ5), which are specific to the types of movement performed. The conditioned velocity estimates are translated into position estimates prior to audio synthesis. From visual inspection of the position estimates, the shape of the signal is similar, but contain differences, to the equivalent measurements obtained through the Vicon system (RQ4). Subsequent to the position estimation process, a sonification module receives position estimates and creates MIDI notes through a parameter mapping strategy. The strategy used for the Soniccup linearly converts a range of position values into 11 MIDI notes where the lowest note occurs at the starting position of a reach, and the highest note occurs at 95% of the maximum reaching position. The design choices of this strategy creates a concurrent auditory display, however, further testing is required to verify if the audio feedback is suitable (RQ6).

To answer RQ7 stated in Section 4.5 the KinectSon system uses the motion capture capabilities of the Microsoft Azure Kinect camera, along with an associated purpose-built Body Tracking SDK, to obtain position coordinates of human anatomical landmarks captured through the camera. Hand position data are low-pass filtered to reduce the effects of noise, before proceeding to the sonification module to create real-time audio feedback through linearly converting position data to audio pitch. For reaching movement of all speeds tested, the system obtained relative position data that is representative of the movements performed. Given the results from the comparison study, there is a higher chance that the audio feedback generated through the KinectSon, as opposed to the Soniccup, would be relatable to the performed movement as perceived by the performer (RQ9). However, this is not guaranteed based on this study. For this reason, the KinectSon has been selected as a movement sonification system for a subsequent study to identify acceptability of feedback through an agency study.

Chapter 5

Agency Study

At present two movement sonification systems have been created, one through the motion capture capabilities of the NGIMU, termed Soniccup, and the other through the motion capture capabilities of the Microsoft Azure Kinect SDK, termed KinectSon. Using a Vicon Nexus motion capture system as a reference, a study was completed to identify which of the Soniccup or KinectSon systems produced a positional output that was most representative of the reaching movements being performed. The KinectSon system was shown to obtain an output that was consistently more similar to the reference, than the Soniccup system, and was able to achieve this with less system latency between movement and audio. Hence the KinectSon system was brought forward for further testing.

The next stages of development are to evaluate the suitability of the KinectSon system as a tool for rehabilitation, through determining feasibility and acceptability of this system. To begin, a study methodology was created to investigate whether novice volunteers obtained a Sense of Agency (SoA) whilst using the KinectSon system, through the identification of their own movements from the audio feedback generated.

The KinectSon system was setup to capture forward reaching movements performed by non-neurologically impaired volunteers and stroke survivor hemiparetic volunteers and produce audio feedback with various sonification configurations as dictated by the study methodology. Volunteers were asked to report how strongly they associated the perceived sound to the movement(s) they performed. Results of this study are described in this chapter, along with a discussion of the implications of the findings.

For the work in this section, the following research question was proposed:

RQ10 Do users with movement impairment as a result of stroke, and users without movement impairment, obtain a SoA whilst using the KinectSon?

5.1 Developing a Complex Intervention

Interventions can be described as complex as determined by the intervention characteristics. Examples of characteristics to consider include the number and variability of outcomes, number of targeted groups or organisational levels, and the degree of flexibility or tailoring permitted [304]. Effectiveness of a complex intervention could be dependent on these characteristics, and as such developing an effective complex intervention is a challenge. Movement sonification, if applied as a rehabilitation tool, would be viewed as a complex intervention as effectiveness may depend on the movement capabilities of the patient, the musical competency of the patient, the type of motion capture technology used, the sonification configuration used etc.

Guidance for developing a complex intervention from the Medical Research Council [304], along with an update from Skivington *et al.* [305], highlight the need to establish intervention feasibility along with evaluating efficacy. Intervention feasibility encompasses multiple factors, such as cost effectiveness, capacity to deliver the intervention, and acceptability. Acceptability, in the context of complex interventions, is the term used to describe how likely the intervention is to receive a positive response from the target population. High acceptability increases the chance that the implemented intervention would be adhered to when applied in a real-world settings [306], and as such desirable from a research perspective.

In the context of developing of a movement sonification system, underlining whether the movement performer believes that the audio feedback obtained through the movement sonification system represents their movements, creates a foundation to work from for further system development. At present, the KinectSon system has shown to produce positional outputs with high similarity to the positional outputs obtained through a Vicon Nexus system, leading to the assumption that the KinectSon would generate audio feedback that a movement performer can anticipate and understand. To test this assumption, a study is required to establish whether the KinectSon sonification output produces a SoA with movement performers.

5.1.1 Agency

Agency, or the Sense of Agency (SoA), is the feeling of control when performing a volitional action. Literature on this topic describes two theoretical models to explain the origination of agency; predictive and postdictive [307]. The predictive model involves a generated internal forecast of an action along with the consequential sensory outcome and comparing this to the actual outcome of the action. The SoA would emerge if the comparison results in the prediction matching the outcome. The postdictive model is based on reflective assumptions during and after movement. This model relies on three principles [308]:

- Priority - conscious intention to perform the act immediately prior to performing
- Consistency - sensory outcome fits the predicted outcome
- Exclusivity - only apparent cause of the outcome is through the thoughts of the performer

Agency emerges based on the sensory evidence that supports these principles. Both of these models are considered valid in the literature of agency [309], with other theoretical models of agency suggesting that the SoA is an amalgamation of both [307].

Agency has been segregated into two concepts, labelled as the Feeling of Agency (FoA) and the Judgement of Agency (JoA) [310]. The FoA is considered to be a low level sensorimotor process based on the comparison between the predicted and actual sensory event from an action. Those who are unable to obtain this process, such as people with passivity symptoms due to schizophrenia, struggle to identify their own movements [311]. Therefore, the FoA is imperative to the feeling of control when performing actions. The JoA is considered a higher level cognitive process [312], which uses more causal processes to assign agency to an action. The judgement is considered a reflective process that is dependent on the low level sensorimotor process, contextual knowledge, and belief reasoning [310]. These concepts have been used to explain the predictive (associated with FoA) and postdictive (associated with JoA) models.

A method of measuring the FoA is through identifying the 'intentional binding' effect [313], a phenomenon where participants perceive a shortened time period between their own voluntary actions (action binding) and the sensory consequences (outcome binding). These are interpreted as indicators of agency emergence. The opposite effect to the phenomenon occurs for involuntary actions whereby a longer delay between action and consequence is perceived, which indicates a lack of emerging agency. Methods to obtain JoA measurements are through asking participants to rate how closely they associate an outcome to their action [179, 314, 315]. Methods to obtain measurements for FoA and JoA are markedly different, and as such, the observed cognitive processes of the FoA and the JoA are markedly different as well. This difference has been represented through activated brain regions which has been described to form a rostrocaudal gradient [316], which for the FoA is largely situated posterior to the parietal cortex, and for the JoA situated in the prefrontal areas. In the context of neurological damage as caused by stroke, these activated brain regions may be compromised for stroke survivors, as dependent on the location and extent of the lesion. Although both of these concepts are stated to be important to determine self-agency, it is evident that these concepts relate to different roles and use different neurophysiological processes. However, it is clear that agency is closely linked to movement.

Agency and Movement

To perform a movement, an accepted theory proposed by von Holst [317] describes that efferent signals are required to travel to the associated muscles, and afferent signals are required to travel from the associated muscles. As movement is performed, an 'image' of the efferent signals is retained, that forms a template for future movement to be guided from. This image copy is referred to as the 'efference copy' [317]. It is hypothesised that during the intention phase of movement through the 'forward model' [318] that the cerebro-cerebellar pathways use an efference copy to predict the outcome of a movement. Subsequent to the feedforward control of the performed movement through the efferent signal pathways, the outcome of the movement, presented as feedback stimuli through the afferent signal pathways, is presented to the cerebellum. With the prediction and outcome information, the cerebellum is hypothesised to act as a comparator for the movement performed, with the ability to update the efference copy as dependent on the disparity of prediction and outcome information. This process allows for the refinement of a movement template, of which future movements of that type are constructed from.

In the context of conscious control of movement, the cerebellum is considered a predictor and comparator. Evidence points to the role of the cerebellum in motor planning [319], temporal control [320], and receiving sensory information, such as vision, audio, balance, and proprioception [321]. The cerebellum is therefore, frequently associated with mismatch detection, which is a strong contributor in the emergence of agency [318]. As the sensory information that the cerebellum can process comes in a variety of forms, it follows that agency can emerge through a variety of stimuli types. One example is shown through Nahab *et al.* [322], with a study that modulated the visual feedback of a finger tapping action, and observed increased activation in various brain regions, including the cerebellum, for modulated feedback. For this study, the visual feedback provided consisted of a virtual hand, with a Cyberglove (Cyberglove Systems) used as an input controller, as such in addition to the imaging results obtained in this study, the study also showed results that indicated that agency, as reported by participants, could be elicited through a novel HCI as dependent on the quality of the feedback.

Agency and Human-Computer Interface

Interaction between a user physically moving a computer mouse, and a virtual pointer on a computer screen is an example of a HCI. It is expected that a working computer mouse would have real-time performance and high accuracy related to the control of the user, and therefore, users would perceive that the movement of the virtual pointer on screen, is a direct result of the movement of the hand, despite no physical link existing between the two. In other words, the user should elicit a SoA based on the feedback of the virtual pointer via the control of the computer mouse.

New HCI technologies are rapidly becoming available to consumers, including voice assistants [323], and virtual reality technologies [324], with each new interface type producing challenges for technology developers to enable users with a SoA. In general, research links HCI acceptability with levels of SoA, whereby if users of a technology interface perceive they have a low level of control, and therefore, a low level of agency, they are less likely to use that technology [325]. The background of agency with HCI research details various factors which affect agency elicitation [326], these factors include input modalities [327], feedback quality [328], reliability [313], and latency [313].

Coyle *et al.* [327] completed a study comparing agency emergence on two input modalities, a keypad button press, and a skin-based input device. The skin-based input device consisted of a piezo-electric microphone fixed on the forearm of the user, that detected vibrations from user generated finger tapping. The study procedure used a Libet clock [329] and an auditory stimulus that was emitted 250 ms after the user action. Participants reported the time on the Libet clock corresponding to their action, or the auditory stimulus. Results from this study indicate that the skin-based input device, compared to the keypad button press, elicited a stronger FoA as determined by a greater intentional binding effect.

Sato *et al.* [328] presented a study highlighting the effect of mismatch detection on agency. Participants were instructed to press buttons, that corresponded to feedback through auditory tones. Under the training phase, the participants associated one button with a 600 Hz tone, and another button with a 1000 Hz tone. The study investigated participant JoA, through the reporting of agreement to the phrase "I was the one who produced the tone", whilst switching the associated auditory tones between the buttons. Results indicate a significant decreasing effect when the tones were switched, in comparison to the learnt tones associated in the training phase.

Haggard *et al.* [313] conducted a study relating to intentional binding, with varying temporal delays of 250, 450, and 650 ms as the independent variable. The study was split in part into two sections, the first section composed of consistent temporal delays in blocks of trials, and the second section randomised the temporal delays. Results of this study show two aspects, firstly that the intentional binding effect was reduced for the randomised temporal delays in comparison to the consistent temporal delays, and secondly that a consistent reduction in intentional binding for increased temporal delay was found. As such the maximum intentional binding effect found, corresponding to the strongest FoA emergence, occurred when the delay between the action and stimulus, was at the shortest delay period, and when the delay period was consistent. These results have also been obtained through work presented by Sato *et al.* [328].

To summarise, to create HCI technologies that elicit strong SoA, the action-elicited response is required to be expected, consistent, and to occur within a short time-frame from the performed action. However, as movement sonification systems are also a type of HCI with the intended functionality to generate audio as a means to provide sufficient feedback to the user, there remains some unknowns on agency generation with this type of HCI, namely with regards to the Gulf of Execution and Gulf of Evaluation [326,330]. The Gulf of Execution describes the process from the intentions of a user to the input mechanism of the interface, and is of relevance given the work completed through Coyle *et al.* [327] that shows that different input interfaces elicit different levels of agency, and therefore, it would be ill-advised to assume that results from studies that use button presses can be translated into mid-air complex movements such as a forward reach. The Gulf of Evaluation describes the process from the output of the system, to the user evaluation of that output with respect to the original intention. This process and the resulting agency was highlighted through the work of Sato *et al.* [328], however, the feedback presented consisted of an auditory tone, whereas movement sonification systems are able to generate continuous concurrent audio as movement feedback, and therefore, it would be ill-advised to assume that the results of this study would translate well between the two.

The relationship of HCI with agency is untested in context to movement sonification systems, and given that different system based factors affects agency elicitation, it would be difficult to theorise whether the created KinectSon system would induce a SoA with movement performers. As such, a study is required to establish whether users of the

KinectSon system are able to associate their movements with the audio output, elicit a SoA through this process, and therefore, evaluate if movement performers could use the system to receive feedback on their movements.

5.2 Existing Agency Methodologies

Existing methodologies to determine agency evocation are categorised into implicit and explicit methods. Implicit methods are generally associated with measuring FoA, whereas explicit methods are generally associated with measuring JoA. For implicit methods, studies revolve around time-perception, popularly achieved by identification of intentional binding. However, these types of studies generally involve the temporal-manipulation of movement-induced stimuli, and recording the perceived delay from the participant. As such an implicit method to test stimuli quality has not been reported. Explicit methods of studying agency consist of asking the user a question similar to `"Did your action cause this outcome to happen?"`, and recording the response. Multiple studies have made use of explicit methods to obtain JoA measurements, although many have used these measurements to compare with implicit method measurements, two notable studies that use explicit methods are through Farrer *et al.* [314] and Tajadura-Jiménez *et al.* [179].

Farrer *et al.* [314] investigated the effect that spatial and temporal information had on self-recognition, by implementing a methodology that displayed a virtual hand moving a joystick, that related to the participant physically holding a joystick that was obscured from their vision. Each trial consisted of the participant performing movements with the joystick and observing the display which presented either, a direct correlation to their movements, an angular alteration to their movements, or a temporal alteration to their movements. Participants were instructed to report either that the display showed, themselves moving without alterations (labelled as 'self'), themselves moving with alterations (labelled as 'bias'), or another person moving (labelled as 'other'). For temporal alterations, participants responded 'self' significantly more for added latency of 0-50 ms, and 'bias' significantly for added latency of 150-1100 ms. For angular alterations, participants responded 'self' significantly more for angular alterations of 0-15°, 'bias' significantly more for angular alterations of 20-30°, and 'other' significantly more for angular alterations of 50-60°.

Tajadura-Jiménez *et al.* [179] investigated the impact of altered sound spatialisation and synchronicity in an audio-tactile study. Participants were seated at a table, and performed tapping actions on the table at set locations, at a set rhythm. Audio feedback corresponding to the tapping actions were played through headphones, either in real-time (synchronous) or with an added delay of 300-800 ms (asynchronous). After every set of actions, a questionnaire was filled out with responses taken on a 7-point Likert scale, which included the statement `"During the audio-tactile stimulation it seemed like the sound I heard was caused by me."`, where results from the obtained responses showed a significant difference between synchronous and asynchronous feedback for their experiment, providing further evidence that adding a perceivable artificial delay to audio feedback reduces agency.

At time of searching (February 29, 2024), an existing agency study coupled with movement sonification has not been reported, and creation of a new study is required to evaluate acceptability of the system. Both above reported studies researching agency have primarily presented altered/unaltered stimuli based on the movement of the participant, and used surveys as a means to obtain JoA scores from the movement performer. These elements have been taken and used to form a new study related to an acceptability study for the KinectSon.

5.3 Agency Study

A new study design to evaluate agency based on the quality of audio feedback was created and is detailed in this section. Prior to evaluating the KinectSon system with stroke survivors, an initial study with volunteers without movement disability was completed to evaluate the study design. Subsequent to this study, stroke survivors with upper limb paresis were recruited.

5.3.1 Materials and Methods

Participants

For this study, two sets of volunteers were recruited. The first set of volunteers were recruited with the following inclusion criteria:

- Ability to understand and communicate in English;
- Normal upper limb movement in both upper limbs;

and exclusion criteria:

- History of neurological condition;
- Existing musculoskeletal condition affecting the upper limbs;
- Severe balance issues;
- Visual impairment that is not corrected by glasses/lenses;
- Auditory impairment, that dramatically impacts perception of sound;
- Impairment of haptic sensitivity;
- Pregnant;
- Unable to provide consent on your own;
- Recent injurious fall without medical assessment;

of which 10 volunteers from the University of Strathclyde were recruited. The second set of volunteers were recruited with the following inclusion criteria:

- Diagnosis of stroke and upper limb deficit resulting from stroke;
- Living in the community (Greater Glasgow);
- Ability to understand and communicate in English;

and exclusion criteria:

- History of or existing neurological condition other that stroke;
- Not discharged from NHS rehabilitation services;
- Electrocardiogram changed suggesting recent myocardial infarction;
- Diagnosed cardiovascular disorder;
- Extreme obesity(>159kg);
- Suspected of known dissecting aneurysm;
- Acute infections;
- Recent injurious fall without medical assessment;
- Severe cognitive impairment (assessed through the Montreal Cognitive Assessment (MoCA), with a score <22);
- Condition that affects the muscle or joints of the paretic upper limb;
- Severe balance issues;
- Pregnant;
- Unable to provide consent on your own;
- Time period of less than three months post-stroke;
- Visual impairment that is not corrected by glasses/lenses;
- Auditory impairment, that dramatically impacts the perception of sound;
- Impairment of haptic sensitivity;

of which 14 volunteers with self-reported stroke-related upper limb deficits, were recruited for an agency study. Volunteers were recruited from a variety of sources, including the Sir Jules Thorn Co-Creation Centre for Rehabilitation Technology, Chest Heart & Stroke Scotland, advertisements placed on public forums. All participants signed a consent form approved by the University of Strathclyde University Ethics Committee.

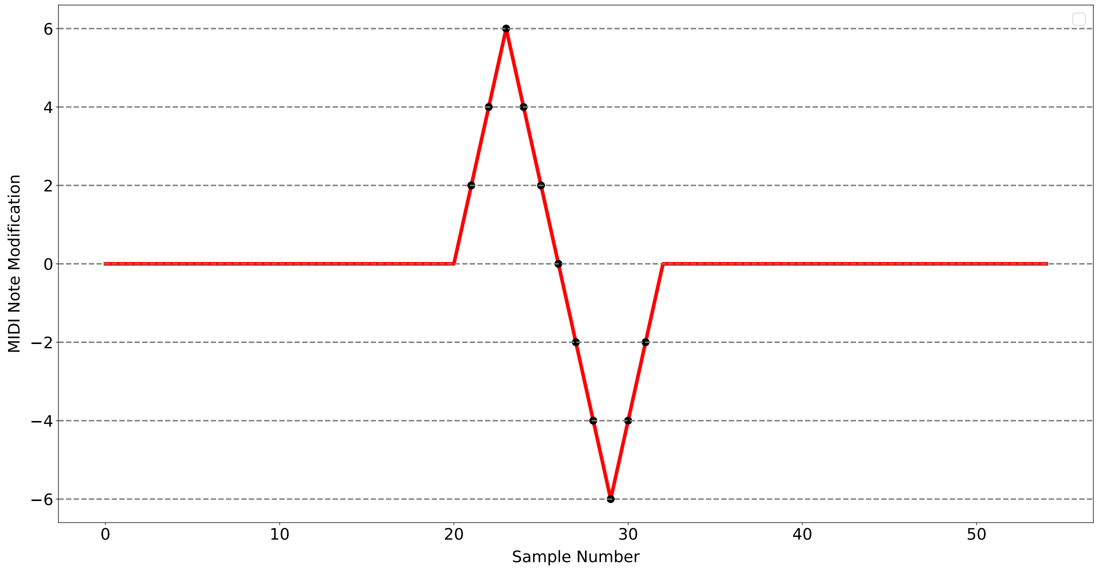

Figure 5.1: Visual representation of the pitch-modulating triangular wave component. Black dots represent sample values that are used to modify the MIDI note values generated in the No Modulation condition.

Sonification Configuration

The primary sonification configuration used for audio feedback in this study mapped the hand position of the volunteer to audio pitch, and only emitted sound on a change of audio note, resulting in an output that emitted a rising pitch as a volunteer performed the extension phase of a forward reach, and emitted a falling pitch as a volunteer performed the retraction phase of a forward reach. This sonification configuration follows the same synthesis process as described in Section 4.5, as such the interval of each change in audio pitch was three MIDI notes, from a low of 48 - corresponding to C3, to a high of 81 - corresponding to A4. For this study, this sonification configuration and three variants of this configuration were used as independent variables. Furthermore, for the remainder of this chapter, these four configurations will be referred to as ‘sonification conditions’.

For three of the sonification conditions, a triangular wave pitch modulating component was added to distort the linear relationship of position to pitch. Figure 5.1 presents a trace representing the triangular wave component.

Black dots on Figure 5.1 represents the numeric value used to distort the sonification output. Equation 5.1 shows the mathematical process of creating distorted audio,

$$n_e = n_i + \lfloor t_i \rfloor \tag{5.1}$$

where n_e is the distorted note, n_i is the original note, and t_i is the value obtained from the triangular wave component. The summation of this component begins when the movement performer reaches 40%-60% of the maximum reach, as such the audio output will initially represent the movement performed until the hand reaches this area, of which the pitch-modulation component begins distorting the audio output. For the example in Figure 5.1, where i is the first black dot, t_i will initially increase n_i by two and then t_{i+1} will increase n_{i+1} by four, etc. This distortion continues for the duration of the triangular wave component, assuming that the generated note does not equal the previous note (i.e. the movement performer continues their movement). Figure 5.2 provides visual illustrations of the effects of the triangular wave pitch modulation component on the audio output.

Figure 5.2(a),(c),(e) and (g) shows visual representations of the triangular wave pitch modulation component at different amplitudes. Figure 5.2(a) shows the component with an amplitude of zero, and hence this sonification condition is the baseline variable of the study, labelled as No Modulation. The corresponding audio output for a full forward reaching movement (i.e. extension and retraction phase) with this condition is shown in Figure 5.2(b) as an example. Subsequent sonification conditions in this figure use the same full reaching movement for the purpose of comparison. Figure 5.2(c),(e) and (g), displays the pitch-modulation component with an amplitude greater than zero, and affects the sonification output by summing a distortion value to the note of the baseline sonification condition. Figure 5.2(d) represents a forward reaching action with the addition of pitch modulating component (c) and is referred to as Subtle Modulation. Figure 5.2(f) represents a forward reaching action with the addition of pitch modulating component (e) and is referred to as Moderate Modulation. Figure 5.2(h) represents a forward reaching action with the addition of pitch modulating component (g) and is referred to as Severe Modulation.

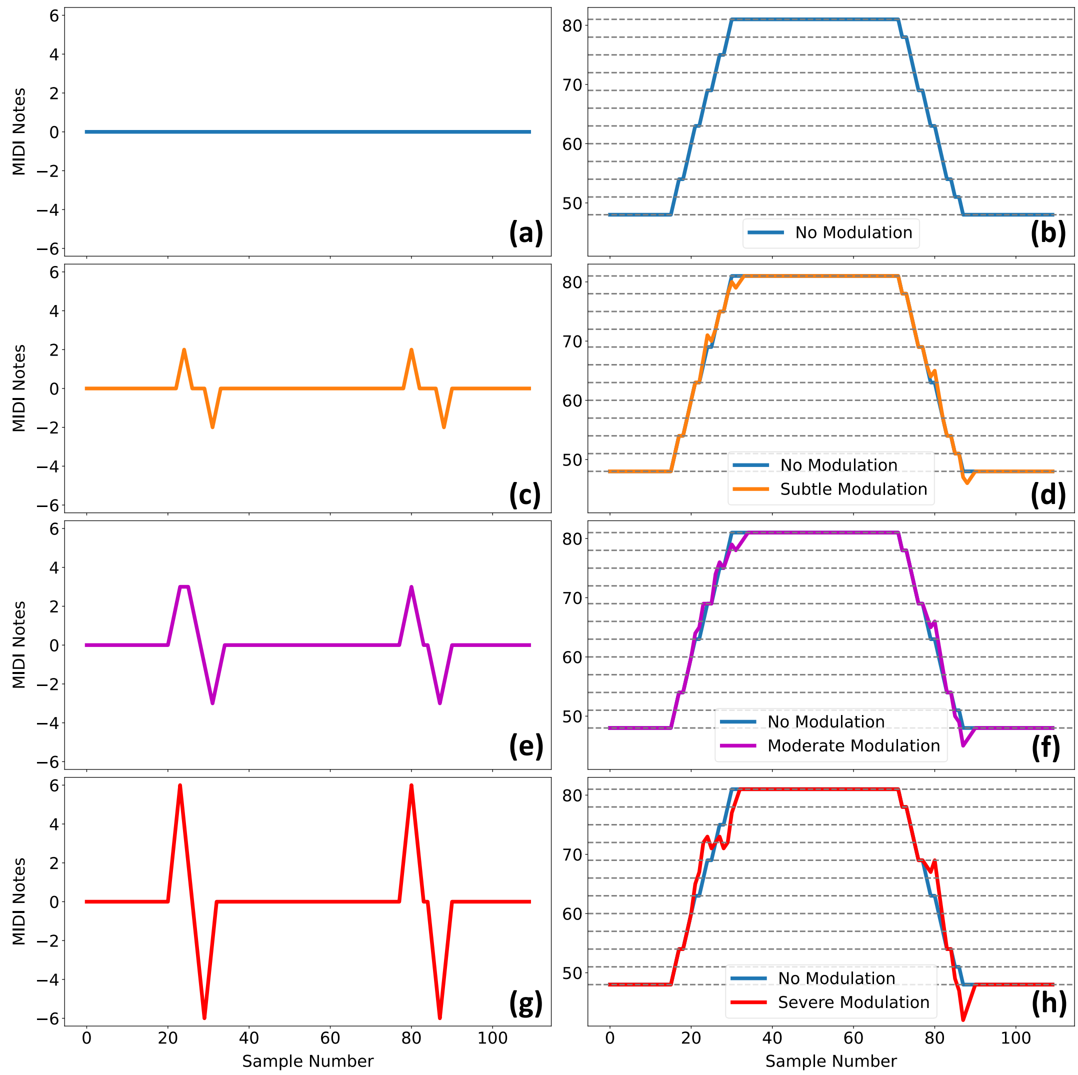

Figure 5.2: Visual representation of the resulting addition of the pitch-modulating triangular wave component onto the audio output. Plots (a),(b) represents the No Modulation condition, plots (c),(d) represents the Subtle Modulation condition, plots (e),(f) represents the Moderate Modulation condition, plots (g),(h) represents the Severe Modulation condition.

Procedure

Each participant attended a single one hour session and reported that they met the respective inclusion/exclusion criteria as stated above. For the second set of volunteers recruited, subsequent to a signed consent form, scores attaining to the ARAT and MoCA were obtained in a screening stage of the procedure. Each participant was instructed to sit in front of a desk. For the volunteers without upper limb impairment they were asked to reach with their right hand, and for the the volunteers with upper limb impairment they were asked to reach with their impaired arm. The KinectSon system was positioned approximately 1.5 m away from the participation area, facing the participation area so that a forward reach movement travelled horizontally across the FOV of the camera. The camera settings were as described in Section A.3, and set to track the position of the chosen hand. The system was calibrated to record the range of reaching movement for each participant, to scale the percentage of reach to the sound condition chosen. This was achieved by taking the average mean starting position of the chosen hand over 100 samples, and the average mean position at maximum forward reach over 100 samples, before the trials started.

Each participant was briefed about the KinectSon system with an explanation on the movements to perform to use the system, and the audio output to expect from the system. Verbal descriptions of the audio for the sonification condition labelled as No Modulation were provided. Participants were not informed about the number of other sonification conditions used in this study nor how those sonification conditions were composed. The final preparation stage was a brief training period that allowed the participant to try-out the system with two different sound conditions, one was the No Modulation condition, and the other was the Severe Modulation condition. Upon instruction from the investigator, participants were asked to perform reaching forwards and backwards movements in the frontal/parietal direction at a comfortable speed with option to pause in-between each reaching movement. Participants were not limited in the amount of reaching movements they could perform in each trial. The study contained 40 trials comprising 10 trials of each of the four sonification conditions (as described in Section 5.3.1) presented in a random order. After each trial participants were instructed to report on a 7-point Likert scale their perceived association with the audio output generated through their reaching movements. All trials were completed through one session.

Data Analysis

Obtained scores from the feedback sheets for every participant were collated and converted to a numerical score whereby, Strongly Disagree = 1 and Strongly Agree = 7. Results were categorised based on the associated sonification condition, and a mean average score for each sonification condition was obtained for each participant, corresponding to four mean scores for each participant. Mean scores from the four conditions were analysed using the Friedman test to assess if a significant difference in scores was observed for any of the groups. If significant differences were observed, a Nemenyi post-hoc test were completed to identify which groups were different. Statistical analyses was completed using the programming language R (version 4.3.1) [331] in conjunction with RStudio (version 2023.6.0.421) [332].

5.3.2 Results

Participants without Upper Limb Deficit

Table 5.1 shows the mean scores from each participant in the first set of volunteers for each sonification condition, along with the average mean score from these 10 participants. Standard deviation of each mean score are also shown. Figure 5.3 visualises the mean scores obtained from the 10 non-impaired participants, along with their average mean score for each sonification condition, shown as a black trace. Error bars shown in this figure correspond to the standard error of each mean score.

Results from the Friedman test on the feedback scores revealed a statistically significant difference ($\chi^2 = 25.2$, $p < 0.001$) in the four groups. Post-hoc analysis with Nemenyi test revealed significant differences between Severe Modulation - No Modulation ($p < 0.001$), Moderate Modulation - No Modulation ($p < 0.002$), and Severe Modulation - Subtle Modulation ($p < 0.016$).

Table 5.1: Mean(SD) agency scores obtained from participants without upper limb deficit.

Participant	**Severe Modulation**	**Moderate Modulation**	**Subtle Modulation**	**No Modulation**
1	2.1(0.831)	2.8(1.249)	4.5(1.025)	7.0(0.000)
2	2.4(0.916)	3.0(1.414)	4.6(1.625)	6.3(1.269)
3	1.1(0.300)	1.8(0.872)	2.7(1.005)	6.0(1.000)
4	2.3(1.100)	2.4(1.200)	3.8(1.833)	6.5(0.671)
5	2.1(1.221)	3.8(1.939)	5.1(1.221)	6.8(0.400)
6	2.6(1.428)	2.8(0.748)	4.9(1.300)	4.7(1.418)
7	2.0(0.632)	2.5(1.565)	4.1(1.446)	4.5(1.746)
8	2.6(1.356)	3.8(1.939)	4.5(1.628)	5.6(1.020)
9	2.0(0.632)	2.7(1.418)	4.1(1.044)	6.8(0.600)
10	3.9(1.921)	3.7(1.792)	3.7(2.052)	5.0(1.844)
Mean	**2.3(0.667)**	**2.9(0.628)**	**4.2(0.657)**	**5.9(0.875)**

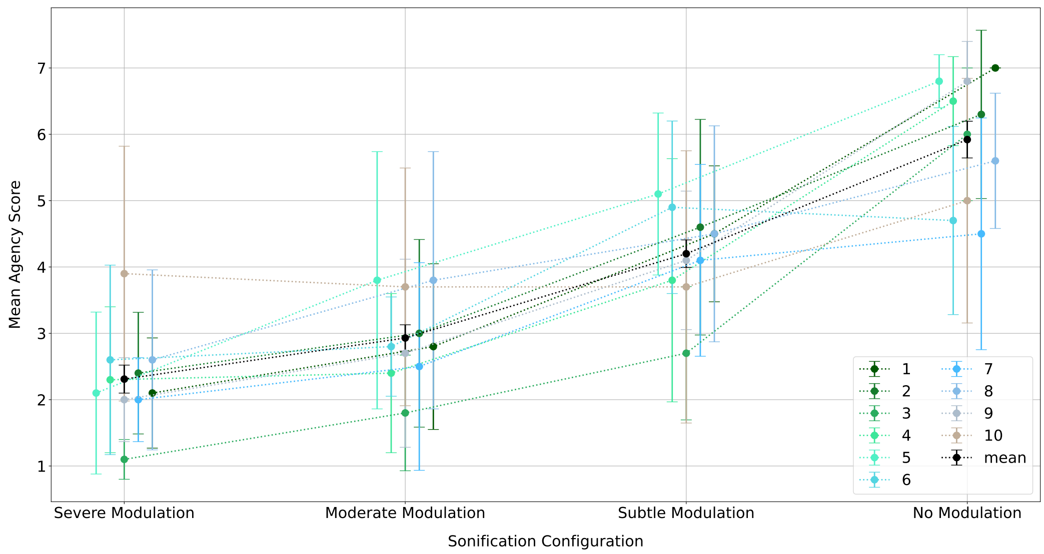

Figure 5.3: Mean response scores of the 10 volunteers for each sonification condition. Error bars represent standard error.

These results indicate that users without movement disability are able to identify their movements using the audio output strategy set on the KinectSon. Figure 5.3 shows a general trend where the agency scores increase with the reduction of the triangular pitch modulation mechanism. For nine of the participants, the highest average mean score was obtained for the No Modulation condition, with the other highest score obtained in the Subtle Modulation condition. For eight of the ten participants the scores of Severe Modulation < Moderate Modulation < Subtle Modulation < No Modulation. Assorting these categories based on the mean scores of the four categories for all participants, as shown through Figure 5.3, presents the trending relationship between the agency score and the sonification condition. For two of the participants, results recorded Severe Modulation score < No Modulation score, but the results for these participants did not follow the general trend as reported through the other participants.

The purpose of this study was to evaluate the study design. A major unknown was whether the triangular wave component generated and emitted with the movement sonification output was salient enough to be identified by the volunteers, without being obvious to the extent that movement associated agency was not examined. Statistical analysis revealed significant differences in the mean scores of three pairs of categories, Severe Modulation - No Modulation, Severe Modulation - Subtle Modulation, and Moderate Modulation - No Modulation, with the rest of the condition pairings not obtaining significant differences. As none of the statistically significant pairings are 'neighbouring', implies that the independent variable used in this study was noticeable, but not obvious.

Participants with Upper Limb Deficit

Table 5.2 shows the mean scores from each participant in the second set of volunteers for each sonification condition, along with the average mean score from these 14 participants. Standard deviation of each mean score are also shown. Figure 5.4 visualises the mean scores obtained from the 14 stroke survivors, along with the average mean score for each sonification condition, shown as a black trace. Error bars shown in this figure correspond to the standard error of each mean score.

Table 5.2: Mean(SD) agency scores obtained from participants with upper limb deficit.

Participant	**Paretic Hand**	**ARAT**	**MoCA**	**Severe Modulation**	**Moderate Modulation**	**Subtle Modulation**	**No Modulation**
1	Left	54	28	1.2(0.400)	1.1(0.300)	1.5(0.671)	5.5(1.204)
2	Left	48	26	3.6(1.960)	3.2(1.833)	4.0(2.000)	5.2(1.600)
3	Left	48	29	1.0(0.000)	3.1(2.700)	4.7(2.610)	7.0(0.000)
4	Left	57	27	1.1(0.300)	1.0(0.000)	1.3(0.458)	6.8(0.400)
5	Right	57	26	1.6(0.663)	1.1(0.300)	2.8(1.536)	6.1(0.831)
6	Left	45	27	6.2(1.661)	5.8(1.720)	6.5(1.500)	6.9(0.300)
7	Right	20	27	1.2(0.400)	1.6(0.490)	1.7(0.640)	6.9(0.300)
8	Right	24	25	3.1(1.136)	4.4(0.663)	4.4(0.663)	5.8(0.600)
9	Right	55	28	1.3(0.458)	1.4(0.490)	2.9(1.700)	5.9(0.700)
10	Left	16	28	3.6(1.744)	3.9(2.587)	4.6(1.800)	6.0(1.844)
11	Right	4	27	3.1(1.700)	4.2(2.441)	5.7(2.369)	6.3(1.792)
12	Right	51	23	6.7(0.458)	5.6(1.960)	6.0(1.414)	6.1(1.136)
13	Right	21	28	2.6(1.200)	5.2(1.720)	4.6(1.428)	6.7(0.640)
14	Left	42	29	2.7(1.676)	1.5(0.671)	3.6(1.685)	5.8(0.872)
Mean	-	-	-	**2.8(1.757)**	**3.1(1.724)**	**3.9(1.602)**	**6.2(0.548)**

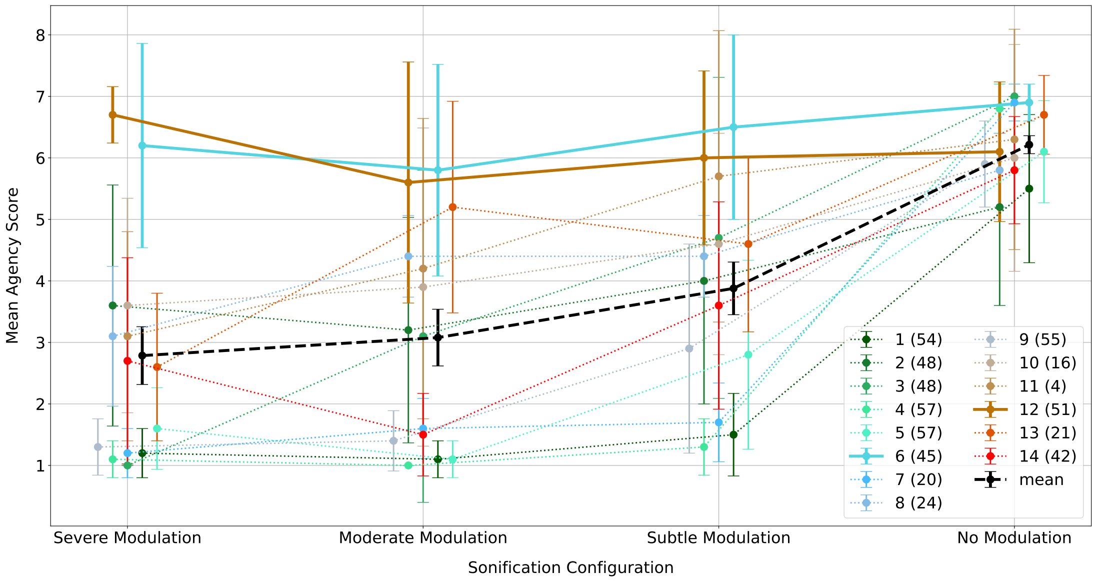

Figure 5.4: Mean response scores of the 14 volunteers for each sonification condition, plus mean response score for each sonification condition. Number in brackets correspond to participant ARAT score. Error bars represent standard error.

Results from the Friedman test on the feedback scores revealed a statistically significant difference ($\chi^2 = 31.1$, $p < 0.001$) in the four conditions. Post-hoc analysis with Nemenyi test revealed significant differences between Severe Modulation - No Modulation ($p < 0.001$), Moderate Modulation - No Modulation ($p < 0.001$), and Subtle Modulation - No Modulation ($p < 0.024$).

This agency study looked to assess whether stroke survivors with upper limb deficit were able to identify their movements through the control of the KinectSon system using their paretic upper limb, and therefore, determine whether the audio feedback generated was deemed acceptable. The inclusion of stroke survivors with upper limb hemiparesis was chosen because the intended application of the movement sonification system was for upper limb rehabilitation, and also to stress test the system by allowing users with different functional capacities to perform reaching movements, and as such provide varied inputs to the system. The participants recruited were diverse in terms of movement performance ability as judged by ARAT scores. The ARAT score range for this set of participants, representing the range of ability to perform functional upper limb tasks, was from 4 - 57. The ARAT is frequently used in stroke rehabilitation as a method of assigning a numerical score from 0 (low functionality) - 57 (high functionality) based on the level of impairment of the person, and was chosen for this study for the same purpose. Two participants who obtained a score of 57 on the ARAT were retained for this study for the reasons that the participant(s) reported upper limb impairment, and the known factor that the ARAT has a ceiling effect [333]. All participants were able to perform a forwarding reaching extension and retraction movement of their own volition, without assistance from their other limb, or from an external agent.

These results indicate that users with hemiparesis are able to associate their movements using the audio output strategy set on the KinectSon. The average mean for all reported scores in this study shows a reduction in agency scores for an increase in amplitude for the triangular wave pitch modulation component. The highest mean score for 13 of the 14 participants occurred in the No Modulation sonification condition. The overall mean obtained for this study showed scores of Severe Modulation < Moderate Modulation < Subtle Modulation < No Modulation, of which 5 of the 14 participants reported scores that match the same trend. However, 12 of the participants reported scores of Moderate Modulation < Subtle Modulation < No Modulation. Statistical results comparing scores of Severe Modulation - Moderate Modulation result in a p-value of $p = 0.934$, highlighting the similarity of the scores given by the majority of participants for these two sonification conditions.

Observable through Table 5.2 and Figure 5.4, the reported JoA scores per sonification condition are markedly different between the participants. To discuss the results regarding each participant, participants are grouped into four groups as dependent on their reported mean score.

Figure 5.5 shows participant grouping through colour coding, and is referred to in the following descriptions. The first group (colour coded blue) comprised participants 1, 4, 5, 7, and 9. Four of the five participants in this group have obtained maximum or close to maximum score (> 53) on the ARAT, indicating mild paresis in their upper limb, with the exception obtaining a score of 20, indicating a more severe paresis. This

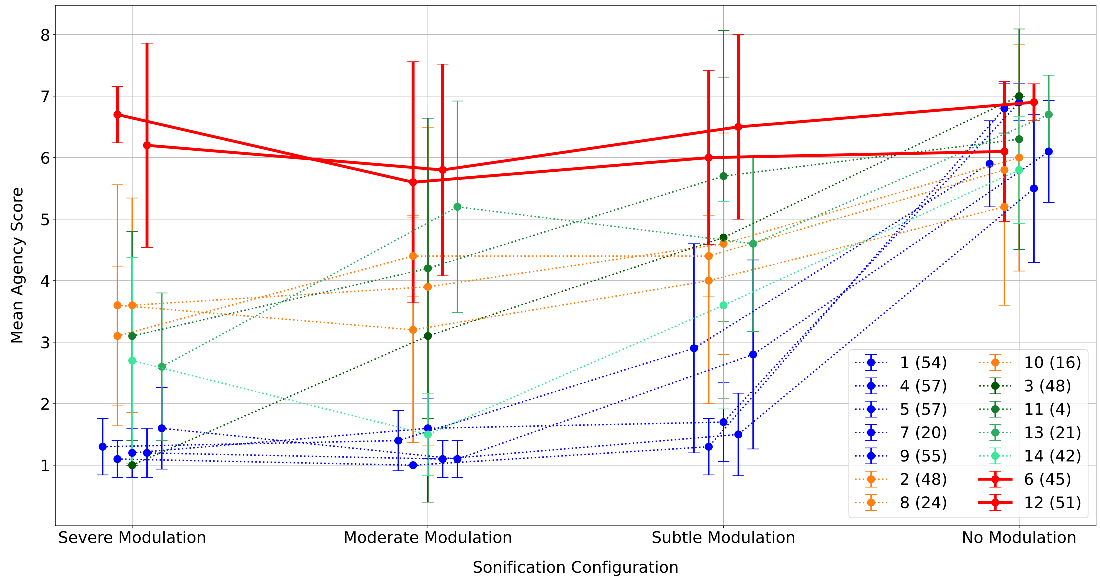

Figure 5.5: Mean response scores of the 14 volunteers for each sonification condition as shown in Figure 5.4, colour coded based on assigned group for explanation purposes. Number in brackets correspond to participant ARAT score. Error bars represent standard error.

group has a higher average ARAT score than the other groups to be described. Each participant in this group reported low scores (< 3) through the sonification conditions that contained the pitch-modulation component, and reported a high score (> 5) with the sonification condition without modulation, implying that they were able to strongly dissociate their movements with the addition of the pitch-modulation component.

The second group (colour coded orange) comprised participants 2, 8, and 10. This group contains one participant with minor-moderate severity paresis (ARAT score of 48) and two participants with more severe paresis (24 and 16). Each participant in this group reported mean scores between 3 and 5 for the sonification conditions that contained pitch-modulation, and reported a high mean score for the sonification condition without modulation, implying that there is a dissociation between movement and sound with the addition of the pitch-modulation component, but the dissociation is not as strong as shown in the first group.

The third group comprising participants (colour coded shades of green) 3, 11, 13, and 14, reported mean scores that do not align with the two prior groups of participants, and requires an individual analysis of each participant. Participants 3 and 14 scored 48 and 42 on the ARAT, respectively, indicating mild-moderate impairment. Participants 11 and 13 scored 4 and 21 on the ARAT indicating moderate-severe impairment. Participant 3 (darkest green) reported mean scores that implies varying severity of dissociation as dependent by the amplitude of the pitch-modulation component, however, the standard error measurement for this participant is the highest for any participant, in the Moderate Modulation (0.722) and Subtle Modulation (0.697) categories, whereas they were the lowest for the Severe Modulation (0.000) and No Modulation (0.000) categories. These results state that this participant reported either Strongly Disagree or Strongly Agree for the majority of trials, irrespective of JoA certainty for the Moderate Modulation and Subtle Modulation but was able to associate/dissociate their movement with the sound created with the extreme variables of this study. Participant 11 (dark green) reported scores in the Severe Modulation and Moderate Modulation condition that are comparable to the scores obtained with the second group of participants highlighted. The distinction however, is with the mean scores obtained with the Subtle Modulation and No Modulation categories, which show a difference of 0.6. This small difference does not suggest a difference in JoA between these two categories for this participant, which implies that for this participant, their JoA process interpreted that the additional low amplitude pitch-modulation component was caused by their movement. Participants 13 (green) and 14 (turquoise), reported mean scores for Severe Modulation, Subtle Modulation, and No Modulation, which show mean scores in each condition that are comparable to the second group of participants, however, the distinction is with the mean score of the Moderate Modulation condition, which for participant 13 is higher than the associated score for the Severe Modulation and Subtle Modulation categories, and for participant 14 is lower than the associated score for the Severe Modulation and Subtle Modulation categories. Reasons for these anomalous results would be of benefit to understand how JoA fluctuates for these participants with respect to the pitch-modulation component, however, this would require a separate study that extends the results found here.

All participants in these three groups had scores that in some capacity indicate that they are able to associate their movements with the audio feedback presented, and obtain a SoA whilst using the system. The final group (colour coded red) comprising participants 6 and 12 reported consistently high agency scores irrespective of the modulation, and are highlighted as solid traces on Figures 5.4 and 5.5. The respective ARAT scores obtained by these participants are 45 and 51, indicating mild-moderate impairment, as such, the study protocol was deemed to be reasonable for these participants in terms of requested movement. The lowest mean score obtained for these participants were 5.6 and 5.8, and these were both observed in the Moderate Modulation condition. These results are contrasting to the JoA scores obtained through the other participants, where at least one mean score in the Severe Modulation or the Moderate Modulation condition was below four. These results imply that every auditory stimulus in this study, elicited a SoA as determined by JoA, from their perspective. Identifying the mechanisms behind this observation could aid in determining who is likely to benefit from the use of movement sonification generated audio feedback as a rehabilitation intervention, and therefore, aid in developing recommendations of use. Theories to explain the results of participants 6 and 12 can be split into two broad proposals, one is that the movements performed by these participants affected the sound emitted through the KinectSon system and therefore, the reported scores reflected the compromised audio output. Alternatively, the sound generated through the KinectSon system worked as intended, i.e. the audio output was saliently affected by the pitch-modulation component, but these participants were unable to report JoA scores that indicated a dissociation with agency.

To investigate each proposal, the following theories were considered, where Items (i) and (ii) relate to the first proposal, and Items (iii) and (iv) relate to the second proposal:

(i) The audio notes generated were indistinguishable from one another, resulting in a lack of clarity in the audio feedback. This is most likely due to a short duration of the movements performed;

(ii) Quality of movements performed were dependent on the sonification condition. This may be due to unconscious cues given by the investigator, which affected the quality of the performed movement, and therefore, affected the sound generated;

(iii) The cognitive process of perceiving audio and determining a JoA, was unable to distinguish between the different audio stimuli generated;

(iv) The participant did not understand the protocol instructions.

Kinematic Analysis

To identify the plausibility of Items (i) and (ii), motion capture data obtained (from the Azure Kinect camera) for all participants were analysed. Movements obtained from each participant were identified based on sections of speed data continuously above a magnitude of 0.02 m/s [275]. This process identified each extension and retraction phase of the reach, and as such each complete reaching movement would produce two sections of data. From these sections of data corresponding to every movement, the following metrics were calculated:

1. Distance of movement performed along X axis;
2. Movement duration;
3. Maximum speed;
4. Smoothness of movement - as determined by number of peaks observed in acceleration magnitude data converted from position.

Item 3 was obtained by calculating the maximum absolute velocity (speed) of each movement, and item 4 was obtained through calculating the acceleration magnitude of the data, and using the 'find_peaks' function [334] inside the scipy.signal module under the Python coding language, with a minimum height set to 0.001 m/s^2 to obtain movement-associated peaks.

Participants 6 and 12 obtained ARAT scores of 45 and 51 respectively, indicating a mild-moderate level of upper limb paresis, therefore, it is expected that observations of some of the movement performance metrics above for these participants would be found to lie in between the extremes of each measurement. Other measures that are not strictly associated with the ARAT may also be relevant, such as ROM, that could affect the audio feedback produced throughout the movement.

Item (i), assumes that the movements performed have a short movement duration, that resulted in the generation of audio notes at a high enough rate that made the auditory display indiscernible, and therefore, the addition of the pitch-modulation component was also indiscernible from the other components of the condition. To test this theory, data associated with the movement duration, was used to create Figure 5.6.

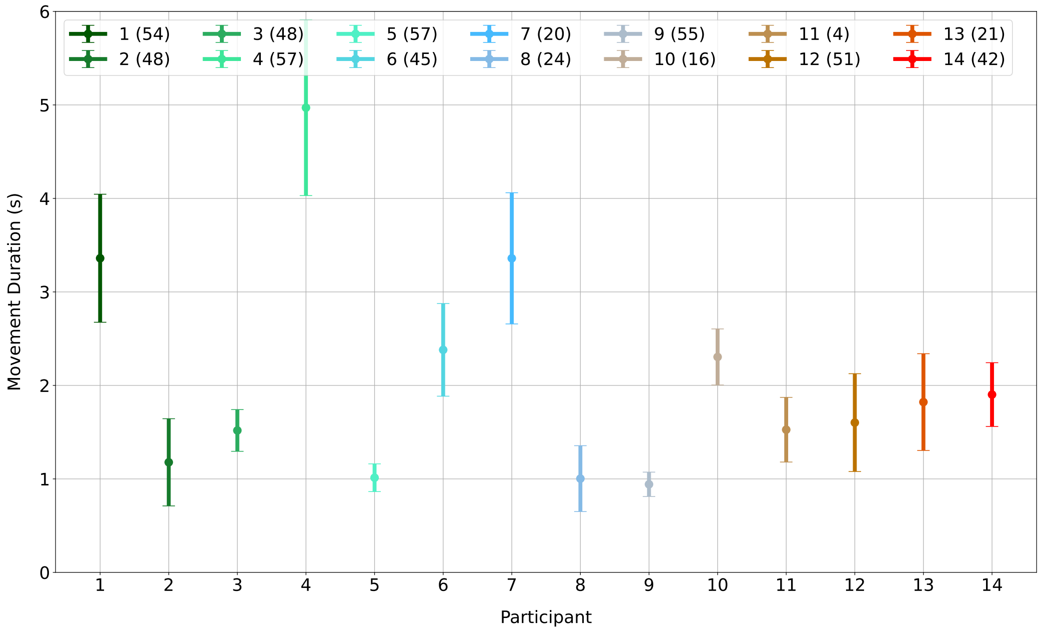

Figure 5.6: Figure showing the average mean data corresponding to the duration of movement of the 14 participants. Standard error of each mean score is shown as error bars.

From this figure, the shortest recorded movement durations were performed by participant 9 (μ = 0.942, s.d. = 0.130), participant 8 (μ = 1.003, s.d. = 0.352), participant 5 (μ = 1.013, s.d. = 0.148) and participant 2 (μ = 1.178, s.d. = 0.466). All of these participants reported a change in agency in line with the pitch-modulation component. Participant 12 (μ = 1.602, s.d. = 0.523) and participant 6 (μ = 2.380, s.d. = 0.496) performed movements with a higher duration than these four participants.

This result infers that participants 6 and 12 performed movements that corresponded to a sufficiently low rate of audio notes that were discernible from each other.

The second theory, that the participants altered their movements based on the sonification condition set, and as such associated a different movement with a different sound, is impossible in theory, given that participants were not informed about which sonification condition was set in each trial, nor informed about how exactly the conditions differ. However, as the trial setup was conducted by the investigator, it is possible that an unconscious cue was obtained from the trial setup during the study. To test this theory, data obtained for each participant was mean averaged for each sonification condition, and has been presented in Figures 5.7 - 5.10. The error bars in these figures represent the standard deviation of each mean calculation. In all figures, participants 6 and 12 are shown as solid traces.

Considering that participants were not informed of the sonification condition for each trial, the expectancy of each figure is to show horizontal traces, i.e. movement kinematic measurements are independent of the sonification condition. For the majority of traces displayed on these figures, the results match the expectancy. The largest change in movement kinematics was observed with participant 4, which showed differences in movement duration (Figure 5.7) and Number of Movement Units (NMU) (Figure 5.10) between the sonification conditions. For participants 6 and 12, the results indicate that they performed movements independently of the sonification condition set, although minor differences can be observed.

To summarise, there is a lack of evidence to suggest that the movements performed by participants 6 and 12 are the core reason for the consistently high JoA score. Both participants performed movements with a duration in excess of one second, with several other participants able to dissociate movements from the modified sounds, in a shorter time frame. The movement metrics calculated per sonification condition also show that the movements performed were consistent throughout the trials for these participants in particular. The two remaining theories therefore, are that Item (iii) the participant's cognitive process of perceiving audio and determining JoA was unable to distinguish between the different audio stimuli generated; Item (iv) the participants did

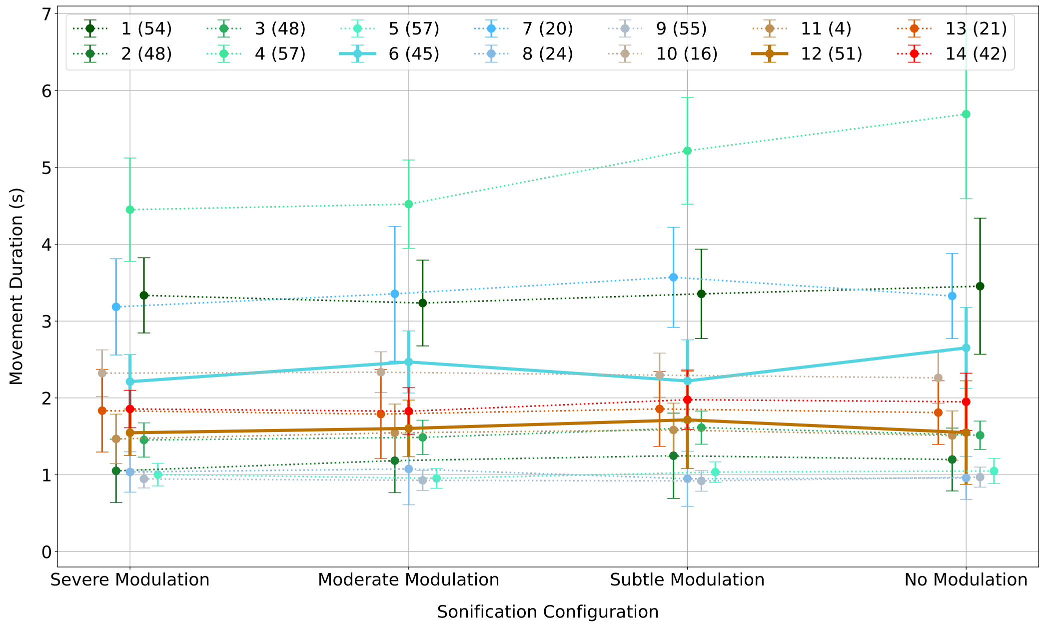

Figure 5.7: Average mean data corresponding to the duration of movement of the 14 volunteers for each sonification condition. Standard deviation of each mean score is shown as error bars.

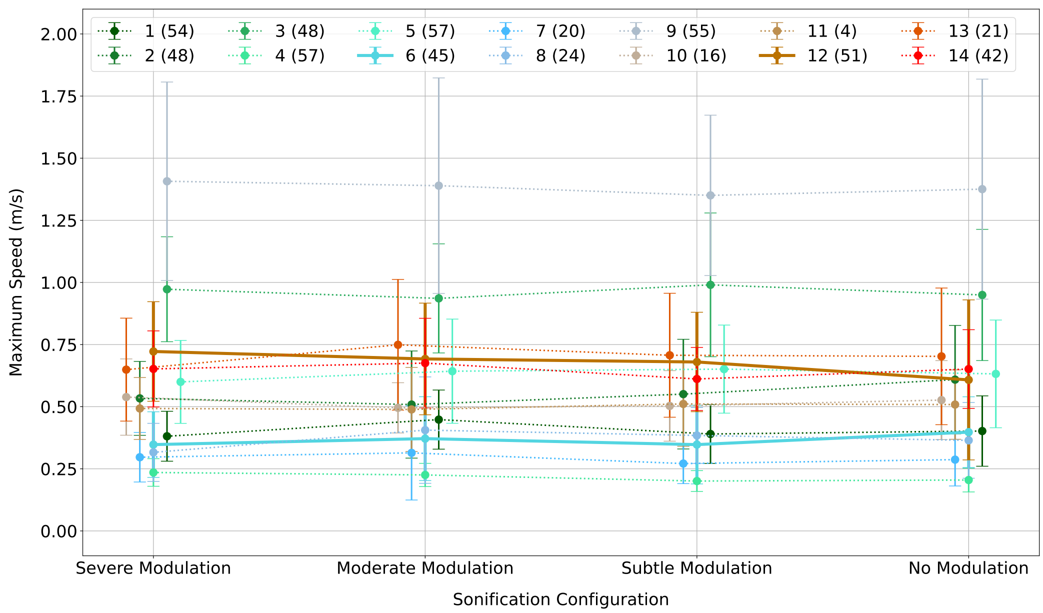

Figure 5.8: Average mean data corresponding to the maximum speed of movement of the 14 volunteers for each sonification condition. Standard deviation of each mean score is shown as error bars.

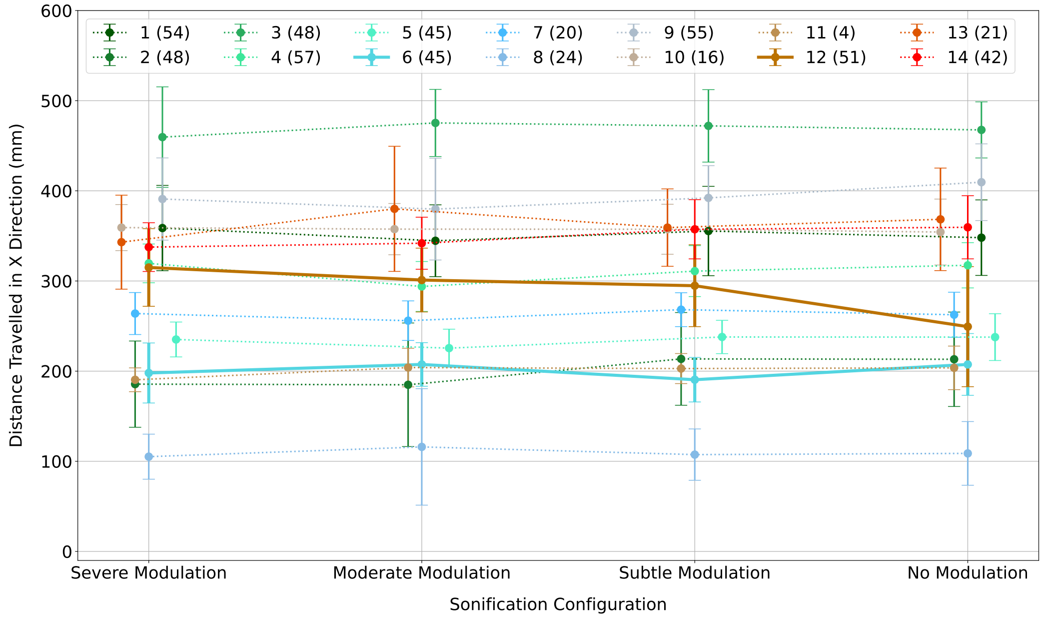

Figure 5.9: Average mean data from the 14 volunteers corresponding to the movement distance travelled in the primary direction of travel, for each sonification condition. Standard deviation of each mean score is shown as error bars.

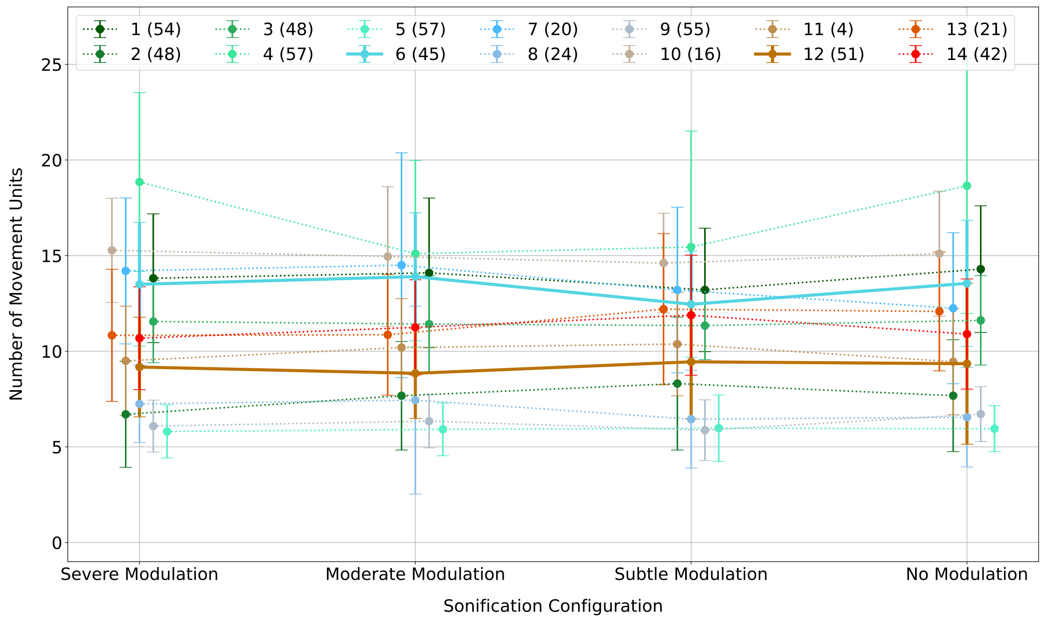

Figure 5.10: Average mean data of the 14 volunteers corresponding to the number of movement units, for each sonification condition. Standard deviation of each mean score is shown as error bars.

not understand the protocol instructions. Both items point to the cognitive ability of the participant, of which a MoCA score was obtained prior to the trials. As the MoCA score for participant 12 was lower that the associated scores of the other participants, a theory is that cognitive deficits played a role in the anomalous JoA scores. This is discussed in greater detail in Chapter 6.

5.4 Conclusion

The findings within Sections 5.3 show that the majority of users of the prototype movement sonification system, the KinectSon, were able to associate their reaching movements with a known sonification condition, and were able to dissociate their reaching movements with a modified version of that sonification condition, which implies that the unmodified movement feedback they received was judged to be representative of their movements (RQ10). However, not every participant reported a dissociation between their movements and the modified movement feedback, and identifying the reasons behind this lack of dissociation would allow for future recommendations on who is most likely to benefit from the introduction of concurrent movement-based audio feedback through real-time movement sonification, as a rehabilitation intervention.

Chapter 6

Discussion

The work presented in this thesis aimed to create a real-time movement sonification system that could be used to rehabilitate the paretic upper limb of stroke survivors. The previous chapters have described work that has begun this process by reviewing existing system components in the field (Chapter 3), creating two new movement sonification systems (Chapter 4), and creating a new methodology to evaluate SoA (Chapter 5). The aims and findings of these chapters are initially summarised, before limitations and suggestions for future work are discussed.

6.1 Mapping Review

The primary objectives of Chapter 3 were to review the systems used in the projects found in the literature, and identify component candidates that could be used to create new movement sonification systems for upper limb rehabilitation. To achieve these objectives, a mapping review was created that collated systems associated with real-time movement sonification that monitored human movement, and categorised these systems based on technology type, human anatomy tracked, physical dimension used for sonification, and auditory dimension used for audio output. These were used to present visual displays that represent the systems used in the literature. The review also points movement sonification system designers to literature relevant to their intended application, and provides an overview of the components that designers can use to create their movement sonification system. Although these variables do not provide a complete picture on the systems that were created, it extends the current body of knowledge of movement sonification systems from reviews that primarily focused on the effects of movement sonification.

The results presented in this study highlight the heterogeneity of the systems created to sonify human movement in real-time. There were 48 different types of physical dimension to auditory category combinations recorded, 17 anatomical segments tracked, and 35 motion capture technologies used in these systems. Given the variation of the identified systems, it would be difficult to extrapolate results from these studies. Therefore, for the application of movement rehabilitation, comparison studies and efficacy studies that minimise independent variables are highly sought after to create new movement sonification systems and provide an evidence-based foundation to develop rehabilitation interventions from. Until then, the results from this review aims to guide new system designers on component choices for their constructed system.

As real-time movement sonification systems for gross movement are not available commercially, system creation is required. To choose components for the new movement sonification systems described in Chapter 4, popularly chosen components as determined by the results of Chapter 3 were a key factor for decision making. Other factors including the feasibility of implementing a large-scale study with the system were also considered. The reason for selecting the component based on this criterion is through the assumption that these components are commonly selected in existing systems because of their suitability as judged by researchers in the field, and therefore, they would be suitable for the creation of a new system for the application of paretic upper limb motion capture.

The review first presented results relating to sonification configurations used, that show a clear preference to using position as the movement kinematic of choice and the use of audio pitch as the altering audio property. The choice of position could be explained through the competency that humans generally have to identify position of their anatomy through proprioception and vision, and that observations of time derivative kinematics such as speed, require additional cognitive processing. The choice of audio pitch as a means to provide information, allows for a greater perceivable range of information available given the audio interpretation capabilities of the auditory cortex.

The review subsequently presented results relating to the anatomical segment used as a target for motion capture, during movement sonification. As the project is focused on the use of the upper limb, results of interest include the shoulder, upper limb, upper arm, elbow, forearm, wrist, and hand. From these options, the majority of projects tracked the hand as the anatomical section of interest. As such, a design choice was made to also target the hand when creating the movement sonification systems in Chapter 4.

Lastly, the review presents results associated with the motion capture technologies used. Given that the hand is the anatomical segment of interest, the results as shown through Figure 3.6(a) highlight marker-based motion capture systems as the most commonly selected technology. Although this type of system is within a competitive market with multiple companies looking to improve performance and reduce costs of their systems, this type of system would not be considered feasible for large-scale movement sonification studies at present given the high acquisition cost for these systems. Commonly used alternatives (Leap Motion controller, Kinect, optical image camera, inertial sensors) however, do not have the same magnitude of acquisition cost, and as such were considered as part of movement sonification system creation (RQ3). For the creation of two movement sonification systems in the project, the Kinect and IMU motion capture systems were chosen.

The primary purpose of this review is to provide a resource for new movement sonification system designers to help choose suitable components for their system. This has been achieved by collating movement sonification systems in the literature, categorising them, and presenting the results. However, this only provides a starting level of justification for these component choices, and so system designers are encouraged to research the components of interest before including them into their design. In addition, this review does not contribute to the efficacy or effectiveness aspect of the literature.

Whilst key components of each movement sonification system have been identified and categorised, not every stage of the movement sonification process has been covered in this review. The review categorises sonification configurations into combinations of physical dimensions and auditory dimensions, and these auditory dimensions have been subsequently categorised into one or many auditory categories. However, polarity and scaling for each individual sonification mapping in each configuration was not recorded and analysis of combinations of sonification mappings were not analysed. Therefore, the results presented highlight the variety of sonification mappings used, but do not provide a complete review of the sonification configurations used. The review also categorises anatomical segments tracked in the projects reviewed, however the review does not distinguish between different areas of each section of human anatomy which may be of relevance for certain applications (for example anterior/posterior sections of the trunk). The review also does not describe the programme or method used to convert data into sound, of which options such as Pyo [302], Max/MSP/Jitter [335], or CSound [336], to name a few, are available.

To provide recommendations of components with a deeper level of justification, evidence-based studies that compare the effects of different movement sonification systems and different sonification configurations are needed. To start, isolating and identifying the 'best' type of physical dimension for the intended application should be an aim. From this information, comparison studies of the auditory dimensions could be performed, and in parallel a convergence in motion capture system choices would allow for studies with similar systems to be implemented.

Given the low sample sizes in studies involving movement sonification, future researchers in this area are likely in need of agreeing to a standardised system so that multiple institutions can contribute their results to a meta-analysis, and therefore, evaluate whether movement sonification can used as an effective rehabilitation tool. The standardised system would need to be affordable, accessible, and versatile, to permit multi-institution collaborative studies.

6.2 Movement Sonification System Development

Chapter 4 described two created movement sonification systems. One system termed Soniccup, used a single IMU sensor, a switch, a 3D-printed object, and a position estimation algorithm, that was able to obtain position estimates from acceleration readings. The other system, termed KinectSon, made use of an Azure Kinect camera and an associated Body Tracking SDK to obtain position estimates of anatomical landmarks. Both systems were developed with the objective to generate interpretable real-time audio feedback of hand position estimates. To evaluate the suitability of the motion capture systems as a part of a movement sonification system, the shape of the position output from the two systems were compared to the shape of the measured position obtained through a Vicon Nexus system, which was assumed to be the ground truth for this comparison.

Soniccup

For the Soniccup system, the shape of the positional output was shown to have high similarity with the Vicon Nexus measurements for movements performed at a normal to fast pace. However, the shape of the positional output for slow movements was shown to have low similarity with the Vicon Nexus measurements. The low similarity observed for the slow speed of movements can be explained through the increased time periods of movement. As this duration increases, errors as a result of double integration accumulate further, resulting in signal shapes that do not represent the performed movements. Bai *et al.* [303] recently compared various drift correction techniques to improve position estimation and concluded that the use of ZUPT provided the most effective method to improve position estimates, provided that regular period of non-movement can be identified. Although other techniques have not been researched for this application, the results of Section 4.7 support the findings of Bai *et al.* [303], and show that appropriate use of ZUPT could lead to improved position estimates for reaching movements with a movement duration of less than one second. Improving the performance of reaching movements with durations greater than one second remains a challenge to implement effectively, especially for a real-time system, and therefore, remains a barrier for users to practice reaching movements at a slow speed whilst obtaining acceptable audio feedback.

The results from the Soniccup system were obtained with a system delay of approximately 230 ms from the movement performed, representing a system that does not provide real-time audio feedback. As the position estimation algorithm in the Soniccup system used the Zero Velocity Update (ZUPT) method to condition intermediary velocity estimates, a push-to-make switch was used to detect periods of zero velocity. The inclusion of the push-to-make switch in this system was the primary reason for the large system delay for two reasons. The first, is that to validate that an accidental change of state did not occur in the system, a filter was required and implemented to remove momentary changes in switch state, which created a delay of 8 samples (80 ms). The second, is from the start of movement misalignment between the acceleration and switch, and therefore, required the acceleration data to be delayed prior to processing.

The switch chosen to create the prototype system contained many desirable properties, including low cost, low actuation force, and the momentary activation, however the non-negligible difference of the switch electrical distance with respect to the total distance, was also a contributing factor to the system. The Soniccup was setup so that the object sits flush on a surface as the switch was compressed to the total travel distance. The total travel distance of this switch is 2 mm, whereas the electrical distance, i.e. the distance required to change switch state, is 0.8 mm, resulting in a 1.2 mm gap between the Soniccup lifting off the table and the switch changing state. To resolve this issue with the same components, an adjustment to the depth of the cavity on the base of the Soniccup, so that the switch only reaches the electrical distance when the Soniccup sits flush on the surface, would suffice. This change would also result in the reduction of the sample discrepancy at the start of movement, and therefore, decrease the Soniccup system latency for position estimation. Alternatively replacing the chosen switch with another switch with similar properties but with a smaller difference between electrical distance and total distance (ideally <0.01 mm), would produce the same effect.

Another large component of this system that creates delay is the need to check for purposeful changes of switch state. For this application a design decision was made to use the present sample and eight previous samples to identify a purposeful change of switch state, resulting in a delay of 80 ms. Substituting the use of a push-to-make switch with an alternative technology as a ZUPT sensor, such as a depth sensor, could be an effective solution that requires less samples to validate purposeful changes of state.

Establishing a setup and algorithm that can obtain position estimates with latency less than 100 ms, would be the next goal for future developments. Beyond this, creating an integrated system that in addition to the motion capture component, includes signal processing, audio conversion, and audio emitting components, would allow for a singular real-time movement sonification system to be created, where audio would emit at the location of the hand position, without the need for an additional PC. This type of system would have multiple benefits compared to the existing version of the movement sonification system described in this thesis. A key benefit would be through the improved sensory congruence of the hand for the visual, proprioceptive, and augmented auditory sensory feedback, allowing for easier spatial and temporal integration of the sensory information [337], and therefore, improve the quality of movement feedback during movement [47]. Another notable benefit would be through the reduction in different system components to commence rehabilitation, and therefore, improve the feasibility of using this type of system as a rehabilitation tool.

KinectSon

For the KinectSon system, all movements captured showed to have high similarity with the Vicon Nexus system, and therefore, was chosen as the movement sonification system to progress with for Chapter 5. The results from the KinectSon system were obtained with a delay of approximately 67 ms from the movement performed, representing a system that generates audio feedback in real-time. However, given the low capture frame rate of the camera (30 Hz), time derivative metrics of the obtained position resulted in extended system latency between the movement and corresponding audio. For velocity this would be approximately 100 ms, and for acceleration would be approximately 133 ms. The extension to the system latency could result in perceivable audio feedback delay in users, which might affect the acceptability and efficacy of the system as a rehabilitation tool. With the assumption that position is the physical dimension of interest for sonification, the results presented in Section 4.7 provide a good indication that the position obtained can be converted into the auditory domain in real-time, and be interpretable as movement feedback by the movement performer, due to the low magnitude and consistency of the error when compared to the Vicon Nexus system. However, a separate study to evaluate audio feedback acceptability through the KinectSon is still required, given the novelty of the system.

The current version of the KinectSon system uses the Azure Kinect Body Tracking SDK to extract hand position of movement performers. Currently this technology attempts to identify anatomical landmarks of the whole body, through a skeleton-based model, resulting in unstable measurements of the hand in scenarios where sections of the body are occluded (for example when a person is seated at a desk (RQ8)). In the work presented, the impact of the occlusion was mitigated through either maximising the view of the legs through careful positioning of the camera, or through using the desk to completely occlude the lower body from the camera. In scenarios when the lower body was in view, movements in the lower body affected the mapped body pose, which may have been a component in the estimation of the hand position. As such, future work to develop the KinectSon system for the application of upper limb rehabilitation, would be to use a different CV model to estimate the anatomical landmarks of upper-body exclusively, in real-time. An example of such a model has been created by Tsai *et al.* [338] with depth images, as such the use of this type of model along with an Azure Kinect camera could be an improved method of capturing upper limb movements.

An additional limitation of using this system is the required hardware needed to run the technology at 30 Hz, which includes the need of a PC hosting a powerful GPU. The total cost of acquiring these components, therefore, would be a barrier to most of the general public (RQ8). To ensure that such a system is affordable and accessible, future work would be to develop a system that retains the competency of the current system in an accessible technology such as a smart phone.

Comparison Study

A Vicon Nexus system was used to compare the performance of the Soniccup and KinectSon systems with respect to their position estimates. Motion capture of reaching movements at various speeds were obtained and the shape of the position estimates from these systems were compared. As parameter mapping sonification is a linear process, the similarity of position estimates shape is an indicator that the corresponding audio output would be the same, as such comparison of the audio output was not necessary to evaluate the performance of the systems.

The variance in speed of the movements performed were a necessity to represent the variety in speed of the upper limb movements typically performed by paretic stroke survivors. From the performed movements of a volunteer without movement disability, the data shows three groups of data as segmented by their peak speed, where the main discriminant between the Soniccup and KinectSon systems is shown in the movements performed at slow speed. From the result obtained there is a clear difference in the Soniccup performance between the normal and slow speed movements, however a gap in the data emerges around the 500 mm/s movement speed region, which limits the identification of a boundary between good and poor performance for the Soniccup system. Future versions of this type of study could use a visual guide to dictate the desired movement speed, and therefore, avoid gaps in the data. As mentioned, the performance of the Soniccup system is likely to be affected by movement duration and therefore, length of time between ZUPT activations, this metric was assumed to be inline with the performed movement speed, given that the range of movement for the performer was also assumed to be constant, and as such visual displays of the similarity metrics with respect to movement duration were not required.

6.3 Agency Study

To validate if users could obtain interpretable movement-based feedback from the KinectSon system, Chapter 5 reports an agency study to evaluate whether users with and without movement deficits associated their reaching movements to the sounds generated through the KinectSon system. Existing agency studies typically use time-based methods to implicitly measure SoA. These methods are generally preferred to methods that explicitly measure SoA, as agency measurements seem to depend on the methodology used [339], and explicit measures corresponding to JoA measurements, are at increased risk of influence by factors outside of the methodology, for example participants adjusting their reported scores as dependent on an ulterior motive. However, as the intention of this study was to evaluate the suitability of the audio output of the KinectSon system, and the output of the KinectSon system is continuous auditory stimuli, using a time-based method was not feasible. An alternative implicit method could use sensory attenuation, a phenomenon where voluntarily actions incurs a reduced intensity in stimulation feedback. For sensory attenuation of the auditory domain, a voluntary action would result in reduced loudness, hence the reduced loudness of the auditory stimuli would signpost to SoA. However, existing studies that combined sensory attenuation with auditory stimuli used auditory tones triggered through button presses [340–342], and an existing study that examines sensory attenuation through the use of continuous auditory feedback is currently absent. In theory, creating such a study would be feasible with a real-time movement sonification system, provided that the pitch-loudness inequality is addressed for the full range of audio notes [343]. Given the complexity of producing audio in such a way, a choice was taken to use an explicit method of obtaining SoA measurements, through asking the participant to report scores with their perceived association of the movement performed and sound heard.

The methodology presented is the first agency study to use pitch-modulated audio feedback as an independent variable. Results of this study indicated that users of the system were able to dissociate their movements with modulated sonification configurations that added a pitch-modulation component to the sonification output generated, were able to associate their movements with an unmodulated sonification configuration. These findings imply that the audio feedback generated through the system was interpretable by users, as determined by the difference in reported agency scores between the unmodulated and modulated sonification configurations. The findings also imply that the KinectSon system can be used by people with upper limb impairment to perform reaching movements and receive real-time audio feedback, in a laboratory environment (RQ10).

The variation in movement sonification mapping in the literature has presented difficulties in ascertaining types of suitable sonification configurations for movement feedback. The findings of this study indicate that mapping hand position to audio pitch, and emitting sound only during movement, is a suitable configuration for those with and without upper limb movement deficits. However, identifying the most effective audio property to communicate movement feedback, and enhance motor skill learning, would be of great benefit to movement research. Future studies could use a similar methodology as described in Section 5.3 with the use of different audio properties (for example loudness, spatial, temporal) and variations of these audio properties, to identify the property that shows the clearest movement association/dissociation for the participants for the same severity of variation.

For this study, the ARAT was chosen to measure the upper limb functional capacity of the stroke survivors with self-reported upper limb impairment. For the recruited participants, the ARAT scores ranged from 16, indicating moderate-severe impairment in the upper limb, to 57 indicating a lack of impairment in the upper limb. All of the participants were able to perform extension and retraction movements independently throughout the study, however, this may not be the case for other people with moderate-severe impairment as measured through the ARAT. As part of the ARAT assessment, many of the tasks involved picking up and placing objects and as such requiring the use of their fine motor skills (i.e. finger movements, grasping, etc.) in addition to their gross motor skills to perform the actions. As such, some participants in this study had higher quality reaching extension and retraction movements than their ARAT score would suggest. One example can be seen with a participant in this study with an ARAT score of 20, indicating moderate-severe impairment, who was able to strongly dissociate their movements with the modulated sonification configurations, in part due to their reaching movement kinematics being comparable to another participant with an ARAT score of 54. Other hemiparetic stroke survivors with similar ARAT scores could have lower quality reaching movements which may translate to increased difficulty in using the KinectSon system. Future studies could therefore, measure and display the gross movement subset of the ARAT in combination with subsets from alternative clinical assessments such as the upper extremity section of the Fugl-Meyer Assessment [84], to present a more comprehensive measure of upper limb gross motor skill capacity and present a better representation of the gross motor skill capability of each participant.

The results of this study was obtained with minimal training periods for the participants. Each participant was given two opportunities to practice reaching movements, one with the unmodulated sonification configuration, and the other with the most severe modulation of the sonification configurations. In each case, participants were encouraged to perform multiple reaching actions and observe the sound created, with each training period lasting less than one minute per configuration. The results therefore, must be considered in context of the small training period in this methodology, and it may be the case that stronger association/dissociation would be observed with participants after an extensive training period.

Within the reported scores from the different participants, a general trend can be identified with the strongest association of agency obtained with the known sonification configuration, with a decline in scores as the amplitude of the pitch-modulation component increased. Most saliently for the volunteers with deficits in the upper limb, certain groups of participants were able to associate/dissociate their movements to a stronger effect relative to the other participants. As reaching movements were used as an input to the system, one theory for this observation is that high consistency of the performed reaching movements presented consistent audio feedback, and therefore, alterations in the sonification configuration were perceived more saliently to these participants, then they did to other participants that performed movements with higher variability. This theory would extend to suggest that extended training periods to understand how the movement variability affected the audio output, would create a stronger association/dissociation of the audio feedback for the other participants. An alternative theory is that these participants have enhanced auditory processing capabilities, likely as a result of their music education [344], and so were able to identify differences in audio strategy much easier than other participants, resulting in a stronger change in association/dissociation with the audio feedback. As musical competence is likely to be a factor for studies involving movement sonification, future studies should incorporate a musical competence assessment, such as The Musical Ear Test, as described through Wallentin *et al.* [345], in addition to the MoCA and the ARAT (or equivalent assessment for upper limb functionality).

As reported in the study results, some participants did not report a strong difference in JoA scores between the different categories, therefore, testing these theories would aid in identifying stroke survivors that are most likely and least likely to receive benefit from the inclusion of real-time movement sonification as a rehabilitation tool. A limitation of this study therefore, is that the mechanisms behind the inter-participant differences are not identifiable through this study. To address this limitation, changes to the study protocol through a refined eligibility criteria list, and through additions/improvements to the baseline assessments used, would be recommended for future movement sonification studies. Along with the addition of the musical competency assessment and amendment to the assessment used for upper limb motor competency, as stated above, other recommended changes would be for eligibility clarification, for example 'Severe balance issues' could be replaced with 'Inability to sit upright at rest' to clarify the need for the participant maintain balance with a seated position, as opposed to balance throughout walking.

6.4 General Discussion

The aspirations of this project, of which this work contributes towards, is to conduct and complete a large-scale RCT study that uses movement sonification as an audio based upper limb rehabilitation tool, to evaluate effectiveness. In preparation for this type of study, identifying study feasibility, system acceptability, and intervention efficacy, are all important components to predict intervention effectiveness. A current limitation for a large-scale study is the feasibility of obtaining many systems to give to many participants. A commercial system to provide real-time movement sonification for movements involving gross motor skills is currently absent and the challenge for researchers therefore, is in creating a suitable system for their intended application. The difficulty of the challenge increases when considering the need for decisions based on evidence of efficacy/effectiveness, and increases further when considering the need for such a system to be environmentally versatile (for example laboratory and home-based environments).

As reviewed in Chapter 3 many systems have been created for a variety of applications such as upper limb rehabilitation, gait rehabilitation, and sport. On one hand, the observation that many researchers have used movement sonification as a method of obtaining real-time audio feedback, reporting positive effects, indicates that there is potential with using this technology for upper limb rehabilitation. On the other hand, the broad range of components used to create a system, and a lack of consensus on chosen components, indicates that a one-size-fits-all solution does not exist, and that research has not come to any conclusions on the best system(s) to create.

Ideally, creating a low cost system using sensors that are widely available would allow for mass production of a movement sonification system design, and as a result increases the feasibility of implementing large scale studies to investigate the effect of movement sonification as a rehabilitation tool. In principle, this could be achieved through using inertial sensors without many limitations, given their availability in devices such as smart phones. However, a question remains as to the type of movement data that is most useful for movement performers when converting to concurrent audio feedback. Although an efficacy-based answer was not obtained, as part of the results presented in Chapter 3, the majority of projects made use of position, velocity, acceleration, and/or orientation. As inertial sensors include a gyroscope, orientation can be obtained readily, and signal processing from raw angular velocity, to gravity-free orientation is described extensively in the literature [262]. The challenge therefore, is how to extract accurate velocity and position data, from acceleration data, considering the attached integration error.

The findings of Chapter 4, Sections 4.3 and 4.7 indicate that it is possible to condition acceleration data, using Kalman filters and the ZUPT methodology, to create position estimates that represent the performed movement through translation to the audio domain. It may also be possible to achieve this in real-time, given a suitable method of identifying periods of zero velocity. However, to effectively obtain position estimates through the ZUPT method, regular zero velocity markers are required, which would be viable for repetitive pickup and placement exercises. If the duration of movement extends beyond a second, then performance of the system will decline resulting in output substantially affected by error. In context of movement rehabilitation, not every movement will consist of zero velocity periods (for example rotational movement), as such this methodology has limitations in representing certain types of movement. In context of movement metrology, the reported results show a discrepancy in distance between the estimated position and the measured position from the Vicon system, as such this system in its current version would not be suitable for applications that require high accuracy of the reaching movements performed.

To summarise, obtaining accurate position estimates through acceleration in real-time remains aspirational. With the aim of obtaining position estimates of a forward reaching action, position estimates obtained from the Soniccup system only represent the movements performed to some capacity if the movement duration for each extension/retraction movement is below a second. However, the creation of the Soniccup, along with the reported performance of movements of short durations, is a progressive step towards creating an appropriate system for large scale studies involving movement sonification. The system is able to output position estimates online using accelerometer measurements from a single IMU sensor, with approx. 230 ms of delay from movement to audio, explainable in part by the choice of the ZUPT sensor, and as such the methodology to achieve these results is a contribution to the literature. Further development to this type of system, including an alteration in ZUPT sensor used, decreasing system latency, further signal conditioning to improve position estimates, are warranted prior to testing for acceptability, feasibility and efficacy.

The use of markerless CV technologies is likely to be a faster way to obtain accurate movement kinematics for movement sonification. Commercial packages that include hardware and CV algorithms for markerless motion capture have been available for some time. An example is with the Microsoft Kinect camera for Xbox, a commonly chosen technology with researchers, that was acquired and repurposed from its original conception in the entertainment industry. The production of the technology has since ended, and replaced with a new version, the Azure Kinect. Comparisons of this technology to its predecessors have shown mixed results, on the one hand Tölgyessy *et al.* [117] reported an improvement in repeatability measures, and Albert *et al.* [297] reported increased accuracy of motion capture of the feet in a gait analysis study, however, the latter publication also highlighted decreased accuracy in the motion capture of the upper extremities. Even with these reported results, the choice of selecting the Azure Kinect package as the motion capture system for a second movement sonification system, the KinectSon, was justifiable given the known use of the technology, and the tracking capability. The findings of Chapter 4, Sections 4.5 and 4.7 indicate that position measurements from this technology are able to represent the reaching movements performed, regardless of the speed or duration of the movement. As important as these results are for movement sonification development, the cost to produce multiple versions of this type of system is infeasible considering the cost of the technology and the necessary hardware to run the technology at 30 Hz. Therefore, a challenge remains to reproduce the same technology capabilities of the Azure Kinect in a system without the high acquisition cost, and preferably with a CV model tailored to the application of upper limb rehabilitation. As such the current system would not be suitable for large-scale studies, but may be useful to determine system acceptability, which would form a basis for future development.

Establishing acceptability of the system would be a key marker for creating a movement sonification based rehabilitation intervention. Although the motion capture and data processing elements of the movement sonification system may alter as dependent on the choices of the next system design, the audio conversion and display are largely independent from this alteration and can be evaluated with the current KinectSon system. Acceptability studies can involve qualitative methods to engage with the intended audience and extract data pertaining to the research questions of interest. With upper limb rehabilitation, a study involving a user-centred design approach with clinicians, therapists, and hemiparetic users, is recommended. One study could involve using the KinectSon to enable augmented audio feedback during movement practice, whilst feedback could be obtained through direct questioning and through methods such as the think-aloud protocol [346], to obtain perspectives on system usability and to mould design requirements for future iterations of the system.

Prior to this type of engagement with users, a study was completed to evaluate the KinectSon system as a HCI, through a quantitative methodology to evaluate the SoA experienced by users of the system, which is inline with the literature of SoA studies. The methodology presented in Chapter 5, Section 5.3 is the first type of agency study to use audio quality as an independent variable. For this study methodology, the sonification configurations were varied by the addition of a pitch modulation component, with an amplitude as dictated by the variable group assigned, which altered the audio note played to the movement performer. The choice of amplitude in the most severe category was made in aim so that individuals with relatively low auditory processing skills could identify the difference between the most extreme versions of the independent variables. Two other categories were created with amplitude values of the pitch modulation component between the extreme categories to obtain additional data on the sensitivity to change in the JoA scores compared to the severity of the change in the modified notes.

The results of this study, shown first through participants without arm impairment, indicate that the alterations in the audio notes by the pitch modulation component were perceived and interpreted as an effect outside of their control, resulting in a decline in the reported score JoA that correlated with the increase in the amplitude of the pitch modulation. The findings from this group of participants indicate that the selection of pitch-modulation amplitude and triangular-wave shape, provided enough of an alteration for the majority of participants to dissociate their movements with the sound, but not enough alteration that the results showed a binary-type plot for any of the participants.

The findings of participants with arm impairment showed that some participants were able to identify all of the pitch-modulated trials with ease, whilst others were unable to. Evidently, factors related to each participant have impacted the results obtained in this study, and more studies are required to identify the most impactful mechanisms underlying the variance of these results. A proposed factor is musical competence, Bianchi *et al.* [347] described a pitch discrimination study and showed that participants with years of formal musical training were able to identify changes in pitch more precisely than those without musical training. With the assumption that the inclusion of formal musical training leads to increased musical competence, a question remains on if there is a link between SoA of an audiomotor HCI technology, and musical competence.

Aside from testing system acceptability, creation of a suitable intervention using a movement sonification system is necessary prior to studies that investigate effect. For the application of upper limb rehabilitation, notable aspects of creating an effective rehabilitation intervention would include varied movement practice to improve ADL function, participant motivation, and avoiding the guidance effect.

For the work described in this thesis, the movements performed were forward reaching actions. As the movement action can be represented through travel of one axis, this was judged as a suitable movement to begin work on sonification. However, movement practice for rehabilitation encompasses a variety of different movements to elicit functional improvement, as such, audio translation for these different movements would need to be considered prior to large scale studies. For the sonification configuration used in this thesis, the expected output corresponds to a linear increase in audio pitch as the movement performer extended their arm, whilst the opposite occurred as the movement performer retracted their arm. The results from the described agency study provides evidence that this audio strategy is suitable for forward reaching actions. However, the effect of using this sonification strategy on other movements can not be extrapolated from the results of this study alone.

Design of an audio strategy for different movements would require research in the literature to identify suitable strategies, and testing of strategies if necessary. Ideally, the strategies for different movements can be integrated together to form an augmented auditory environment where movements are communicated through audio feedback that is intuitive through play. An aspiration for movement sonification designers therefore, is to create a standardised evidence-based sonification strategy guideline. A pathway to creating such a guideline could be achieved through studies similar to Vinken *et al.* [186], who investigated whether naive participants could identify upper limb movements based on the audio feedback produced through different amounts of combinations of sonification components, therefore, investigating the effect of sonification dimensionality. In this study the following sonification components were used: (i) cranial/caudal direction was mapped to audio pitch, (ii) medial/lateral direction was mapped to stereo panning, (iii) anterior/posterior direction was mapped to spectral composition, and (iv) absolute velocity of movement was set to loudness. The findings of this study indicated that identification rates were not related to the dimensionality of the sonification configuration, and the use of (iv) with either (i), (ii), or (iii) was sufficient enough to display different types of movement, i.e. slow rotational movements from fast reciprocal movements. Although the findings of this study indicate that participants were able to discriminate different actions based on the generated sound, significant effects of parameter mapping combinations were not found within this study. As such further studies could use a similar methodology with different audio strategies, and compare with these results, to identify if there are strategies that are particularly suitable/unsuitable as audio feedback.

The creation of an acceptable audio strategy, along with the creation of a feasible movement sonification system, will affect participant motivation to complete a rehabilitation programme using movement sonification. Recent studies shows the difficulty in retaining participants for long term studies. Raglio *et al.* [99] described a study that originally recruited 65 participants for a RCT involving movement sonification, and included a retention assessment one month post intervention, of which 29 participants completed, showing that 55.4% of participant dropped out of the study. Hankinson *et al.* [348] presented a relatively smaller study, primarily involving Rhythmic Auditory Cuing (RAC) and detailed the limited adherence to an intervention involving an inertial sensor-based movement sonification system. In this study, one participant declined to engage with the intervention post consent whilst two others had issues with shortages of trained clinicians to implement the intervention. Participants in this study were cited to enjoy the music elicited through the movement sonification system, but were bored of the repetitive movements that they had to perform. Given that the participants were asked to perform 20 minutes a day, three days a week in this study and current guidelines for motor rehabilitation recommend therapy for at least three hours per day, five days per week [13], providing motivation to perform movements for the recommended amount of time will be a challenge for researchers to address, even with augmented audio feedback.

The use of parameter-mapping in real-time movement sonification as a method of providing concurrent audio feedback on movement has been highlighted for its potential in human motor skill learning in this thesis. The rehabilitation pitfall of using this type of technology however, comes with the guidance hypothesis [349]. Sigrist *et al.* [350] describes that task learning involves integration of optimal sources of information, and as the created audio feedback is designed to be an extrinsic optimal source, this can override the use of suboptimal intrinsic sources. This can result in improved performance with the guidance of the augmented audio feedback, and then decreased performance in the absence of the guided feedback. As the intended application is to create a rehabilitation tool, as opposed to a sensory neuroprosthesis, creators of an intervention using concurrent audio feedback need to consider the regularity of feedback within the intervention. Furthermore retention studies post-intervention are a necessity for studies of rehabilitation effect, considering the possible effect of guidance.

6.5 Conclusions

Audio-based techniques show potential for upper limb sensorimotor rehabilitation, with real-time movement sonification identified as a feasible method of generating augmented audio feedback on the movements performed. As real-time movement sonification systems for upper limb movements are not commercially available, the thesis presents work to begin the process of creating an effective system and therefore, meets the primary aim of the work as detailed in Section 1.3. Beyond the review of the most relevant literature, a review of the systems used was completed, highlighting the components used in each system, along with a description of the components used in these systems. The work progresses from the review, to show the creation and testing of two systems, the Soniccup and KinectSon systems. The Soniccup, using inertial sensor technology, was shown to obtain position estimates with good correlation to the measured position from a Vicon Nexus system, for movements with a duration up to one second, but was also shown to have poor performance for movements of longer duration, and operated with a system latency beyond that of real-time perception. The KinectSon, using markerless computer vision technology, obtained position measurements with strong correlation to the corresponding measurements obtained from the Vicon Nexus system, and operated within real-time perception. Subsequent to the comparison study, an agency study was completed with volunteers with and without arm impairment, as part of an evaluation of the feedback generated through the system. The findings from this study are that participants were able to dissociate their movements with pitch-modulated variants of a sonification configuration that involved hand position to audio pitch, but were able to associate their movements with an unmodulated version of the sonification configuration. Future work would look to extend from the findings in this thesis, to develop a feasible and acceptable real-time movement sonification system, create an effective intervention based on the created system, and conduct a large scale RCT study to evaluate effectiveness of real-time audio feedback for upper limb rehabilitation for stroke survivors.

6.5.1 Recommendations

Further to the work described in this thesis, the following is a list of recommendations for future work:

- Continued development towards a real-time movement sonification system that can be operated through a ubiquitous camera-based technology, such as a smart phone;
  - This may include the development and inclusion of an upper body exclusive body tracking program;
- Continued development towards a real-time movement sonification system that obtains position estimates from linear acceleration obtained through a single IMU, with the aim of improving performance for slow and long duration movements;
  - This may include a new approach to detecting zero velocity updates through an alternative sensor;
  - This may include improvements to the signal conditioning approach described with the Soniccup;
- Proceed to studies that further apply movement sonification systems to assess acceptability, feasibility, efficacy, and effectiveness;
  - This may include studies that compare different sonification configurations to identify strong candidates for audio feedback;
  - This may include an investigation to identify users that are most likely to benefit from augmented audio feedback incorporated into movement practice;
- Consultation with end users to obtain preferences of use, mechanisms of interest, and explore opportunities to integrate movement sonification into their practice.

Bibliography

[1] "Azure Kinect body tracking joints - Microsoft Learn," 2022. [Online]. Available: https://learn.microsoft.com/en-us/azure/kinect-dk/body-joints

[2] R. Bhardwaj, N. Kumar, and V. Kumar, "Errors in micro-electro-mechanical systems inertial measurement and a review on present practices of error modelling," *Transactions of the Institute of Measurement and Control*, vol. 40, no. 9, pp. 2843–2854, 6 2018. doi: 10.1177/0142331217708237

[3] V. L. Feigin, M. Brainin, B. Norrving, S. Martins, R. L. Sacco, W. Hacke, M. Fisher, J. Pandian, and P. Lindsay, "World Stroke Organization (WSO): Global Stroke Fact Sheet 2022," *International Journal of Stroke*, vol. 17, no. 1, pp. 18–29, 1 2022. doi: 10.1177/17474930211065917

[4] "Stroke statistics — Stroke Association." [Online]. Available: https://www.stroke.org.uk/what-is-stroke/stroke-statistics#UK%20summary

[5] R. K. Akyea, Y. Vinogradova, N. Qureshi, R. S. Patel, E. Kontopantelis, G. Ntaios, F. W. Asselbergs, J. Kai, and S. F. Weng, "Sex, Age, and Socioeconomic Differences in Nonfatal Stroke Incidence and Subsequent Major Adverse Outcomes," *Stroke*, pp. 396–405, 2021. doi: 10.1161/STROKEAHA.120.031659

[6] OECD Health Statistics, "Mortality following ischaemic stroke," 2019.

[7] J. Aked, H. Delavaran, and A. G. Lindgren, "Survival, causes of death and recurrence up to 3 years after stroke: A population-based study,"

European Journal of Neurology, vol. 28, no. 12, pp. 4060–4068, 12 2021. doi: 10.1111/ene.15041

[8] A. Cieza, K. Causey, K. Kamenov, S. W. Hanson, S. Chatterji, and T. Vos, "Global estimates of the need for rehabilitation based on the Global Burden of Disease study 2019: a systematic analysis for the Global Burden of Disease Study 2019," *The Lancet*, vol. 396, no. 10267, pp. 2006–2017, 12 2020. doi: 10.1016/S0140-6736(20)32340-0/ATTACHMENT/61C6E611-76A5-4194-96FE-9AA2848066DE/MMC1.PDF

[9] A. Patel, V. Berdunov, D. King, Z. Quayyum, R. Wittenberg, and M. Knapp, "Current, future and avoidable costs of stroke in the UK. Executive Summary Part 2: Societal costs of stroke in the next 20 years and potential returns from increased spending on research," Stroke Association, London, Tech. Rep., 2017.

[10] H. Nakayama, H. Stig Jørgensen, H. Otto Raaschou, and T. Skyhøj Olsen, "Recovery of upper extremity function in stroke patients: The Copenhagen stroke study," *Archives of Physical Medicine and Rehabilitation*, vol. 75, no. 4, pp. 394–398, 4 1994. doi: 10.1016/0003-9993(94)90161-9

[11] James Lind Alliance, "Stroke Rehabilitation and Long-term Care Top 10 Priorities — James Lind Alliance," Tech. Rep., 2021. [Online]. Available: https://www.jla.nihr.ac.uk/priority-setting-partnerships/stroke/stroke-rehabilitation-and-long-term-care-top-10-priorities.htm

[12] A. Pollock, B. St George, M. Fenton, and L. Firkins, "Top ten research priorities relating to life after stroke," *The Lancet Neurology*, vol. 11, no. 3, p. 209, 3 2012. doi: 10.1016/S1474-4422(12)70029-7

[13] London: Intercollegiate Stroke Working Party, "National Clinical Guideline for Stroke for the United Kingdom and Ireland," Tech. Rep., 2023. [Online]. Available: https://www.strokeguideline.org/chapter/rehabilitation-and-recovery-principles-of-rehabilitation/#229

[14] R. Stockley, R. Peel, K. Jarvis, and L. Connell, “Current therapy for the upper limb after stroke: a cross-sectional survey of UK therapists,” *BMJ Open*, vol. 9, no. 9, p. e030262, 9 2019. doi: 10.1136/BMJOPEN-2019-030262

[15] N. Gupta, C. Castillo-Laborde, and M. D. Landry, “Health-related rehabilitation services: assessing the global supply of and need for human resources.” *BMC health services research*, vol. 11, no. 1, p. 276, 10 2011. doi: 10.1186/1472-6963-11-276

[16] K. E. Laver, B. Lange, S. George, J. E. Deutsch, G. Saposnik, and M. Crotty, “Virtual reality for stroke rehabilitation.” *The Cochrane database of systematic reviews*, vol. 11, no. 11, p. CD008349, 11 2017. doi: 10.1002/14651858.CD008349.pub4

[17] E. Taub, G. Urswatte, V. Mark, and D. Morris, “The learned nonuse phenomenon: implications for rehabilitation,” *Europa Medicophysica*, vol. 42, no. 3, pp. 241–255, 9 2006.

[18] R. J. Nudo, “Recovery after brain injury: mechanisms and principles,” *Frontiers in Human Neuroscience*, vol. 7, no. DEC, p. 887, 12 2013. doi: 10.3389/fnhum.2013.00887

[19] J. Liepert, I. Uhde, S. Gräf, O. Leidner, and C. Weiller, “Motor cortex plasticity during forced-use therapy in stroke patients: a preliminary study,” *J Neurol*, vol. 248, pp. 315–321, 2001.

[20] S. L. Wolf, C. J. Winstein, J. P. Miller, E. Taub, G. Uswatte, D. Morris, C. Giuliani, K. E. Light, and D. Nichols-Larsen, “Effect of Constraint-Induced Movement Therapy on Upper Extremity Function 3 to 9 Months After Stroke: The EXCITE Randomized Clinical Trial,” *JAMA*, vol. 296, no. 17, pp. 2095–2104, 11 2006. doi: 10.1001/JAMA.296.17.2095

[21] P. Langhorne, J. Bernhardt, and G. Kwakkel, “Stroke rehabilitation,” *The Lancet*, vol. 377, no. 9778, pp. 1693–1702, 2011. doi: 10.1016/S0140-6736(11)60325-5

[22] G. Kwakkel, J. M. Veerbeek, E. E. H. van Wegen, and S. L. Wolf, "Constraint-induced movement therapy after stroke," *The Lancet Neurology*, vol. 14, no. 2, pp. 224–234, 2 2015. doi: 10.1016/S1474-4422(14)70160-7

[23] T. Kitago, J. Liang, V. S. Huang, S. Hayes, P. Simon, L. Tenteromano, R. M. Lazar, R. S. Marshall, P. Mazzoni, L. Lennihan, and J. W. Krakauer, "Improvement after constraint-induced movement therapy: Recovery of normal motor control or task-specific compensation?" *Neurorehabilitation and Neural Repair*, vol. 27, no. 2, pp. 99–109, 2 2013. doi: 10.1177/1545968312452631

[24] D. Corbetta, V. Sirtori, G. Castellini, L. Moja, and R. Gatti, "Constraint-induced movement therapy for upper extremities in people with stroke," *Cochrane Database of Systematic Reviews*, vol. 2017, no. 9, p. CD004433, 10 2015. doi: 10.1002/14651858.CD004433.pub3

[25] M. Y. Pang, J. E. Harris, and J. J. Eng, "A Community-Based Upper-Extremity Group Exercise Program Improves Motor Function and Performance of Functional Activities in Chronic Stroke: A Randomized Controlled Trial," *Archives of Physical Medicine and Rehabilitation*, vol. 87, no. 1, pp. 1–9, 1 2006. doi: 10.1016/j.apmr.2005.08.113

[26] J. E. Harris, J. J. Eng, W. C. Miller, and A. S. Dawson, "A Self-Administered Graded Repetitive Arm Supplementary Program (GRASP) Improves Arm Function During Inpatient Stroke Rehabilitation," *Stroke*, vol. 40, no. 6, pp. 2123–2128, 6 2009. doi: 10.1161/STROKEAHA.108.544585

[27] C.-L. Yang, M.-L. Bird, and J. J. Eng, "Implementation and Evaluation of the Graded Repetitive Arm Supplementary Program (GRASP) for People With Stroke in a Real World Community Setting: Case Report," *PTJ: Physical Therapy & Rehabilitation Journal — Physical Therapy*, vol. 101, pp. 1–8, 2021. doi: 10.1093/ptj/pzab008

[28] H. Thieme, N. Morkisch, J. Mehrholz, M. Pohl, J. Behrens, B. Borgetto, and C. Dohle, "Mirror therapy for improving motor function after stroke,"

Cochrane Database of Systematic Reviews, vol. 2018, no. 7, 7 2018. doi: 10.1002/14651858.CD008449.pub3

[29] K. Dockx, E. M. Bekkers, V. Van den Bergh, P. Ginis, L. Rochester, J. M. Hausdorff, A. Mirelman, and A. Nieuwboer, "Virtual reality for rehabilitation in Parkinson's disease." *The Cochrane database of systematic reviews*, vol. 12, no. 12, p. CD010760, 12 2016. doi: 10.1002/14651858.CD010760.pub2

[30] B. M. Ökmen, M. D. Aslan, G. F. Nakipoğlu Yüzer, and N. Özgirgin, "Effect of virtual reality therapy on functional development in children with cerebral palsy: A single-blind, prospective, randomized-controlled study," *Turkish Journal of Physical Medicine and Rehabilitation*, vol. 65, no. 4, p. 371, 2019. doi: 10.5606/TFTRD.2019.2388

[31] P. Fitts and M. Posner, *Human Performance.* Brooks/Cole Publishing Co., 1967.

[32] D. G. Liebermann, L. Katz, M. D. Hughes, R. M. Bartlett, J. McClements, and I. M. Franks, "Advances in the application of information technology to sport performance," *Journal of Sports Sciences*, vol. 20, no. 10, pp. 755–769, 1 2002. doi: 10.1080/026404102320675611

[33] T. Särkämö, "Cognitive, emotional, and neural benefits of musical leisure activities in aging and neurological rehabilitation: A critical review," *Annals of Physical and Rehabilitation Medicine*, vol. 61, no. 6, pp. 414–418, 11 2018. doi: 10.1016/J.REHAB.2017.03.006

[34] A. Gomez-Andres, J. Grau-Sánchez, E. Duarte, A. Rodriguez-Fornells, and A. Tajadura-Jiménez, "Enriching footsteps sounds in gait rehabilitation in chronic stroke patients: a pilot study," *Annals of the New York Academy of Sciences*, vol. 1467, no. 1, pp. 48–59, 5 2020. doi: 10.1111/nyas.14276

[35] T. Takeuchi, "Auditory information in playing tennis." *Perceptual and motor skills*, vol. 76, no. 3 Pt 2, pp. 1323–1328, 10 1993. doi: 10.2466/pms.1993.76.3c.1323

[36] M. Bangert, T. Peschel, G. Schlaug, M. Rotte, D. Drescher, H. Hinrichs, H.-J. Heinze, and E. Altenmüller, "Shared networks for auditory and motor processing in professional pianists: Evidence from fMRI conjunction," *NeuroImage*, vol. 30, no. 3, pp. 917–926, 4 2006. doi: 10.1016/j.neuroimage.2005.10.044

[37] C. M. Peters and C. M. Glazebrook, "Rhythmic auditory stimuli heard before and during a reaching movement elicit performance improvements in both temporal and spatial movement parameters," *Acta Psychologica*, vol. 207, p. 103086, 6 2020. doi: 10.1016/J.ACTPSY.2020.103086

[38] A. M. Proverbio, M. Cozzi, A. Orlandi, and M. Carminati, "Error-related negativity in the skilled brain of pianists reveals motor simulation," *Neuroscience*, vol. 346, pp. 309–319, 3 2017. doi: 10.1016/J.NEUROSCIENCE.2017.01.030

[39] L. Damm, D. Varoqui, V. C. De Cock, S. Dalla Bella, and B. Bardy, "Why do we move to the beat? A multi-scale approach, from physical principles to brain dynamics," *Neuroscience & Biobehavioral Reviews*, vol. 112, pp. 553–584, 5 2020. doi: 10.1016/j.neubiorev.2019.12.024

[40] M. Meyer, S. Elmer, and L. Jäncke, "Musical expertise induces neuroplasticity of the planum temporale," *Annals of the New York Academy of Sciences*, vol. 1252, no. 1, pp. 116–123, 4 2012. doi: 10.1111/j.1749-6632.2012.06450.x

[41] J. D. Schmahmann, D. N. Pandya, R. Wang, G. Dai, H. E. D'Arceuil, A. J. De Crespigny, and V. J. Wedeen, "Association fibre pathways of the brain: parallel observations from diffusion spectrum imaging and autoradiography," *Brain*, vol. 130, no. 3, pp. 630–653, 3 2007. doi: 10.1093/BRAIN/AWL359

[42] L. Vaquero, N. Ramos-Escobar, C. François, V. Penhune, and A. Rodríguez-Fornells, "White-matter structural connectivity predicts short-term melody and rhythm learning in non-musicians," *NeuroImage*, vol. 181, pp. 252–262, 11 2018. doi: 10.1016/j.neuroimage.2018.06.054

[43] G. Hickok and D. Poeppel, "Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language," *Cognition*, vol. 92, no. 1-2, pp. 67–99, 5 2004. doi: 10.1016/J.COGNITION.2003.10.011

[44] J. L. Chen, V. B. Penhune, and R. J. Zatorre, "The Role of Auditory and Premotor Cortex in Sensorimotor Transformations," *Annals of the New York Academy of Sciences*, vol. 1169, no. 1, pp. 15–34, 7 2009. doi: 10.1111/J.1749-6632.2009.04556.X

[45] G. Dubus and R. Bresin, "A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities," *Plos One*, vol. 8, no. 12, 2013. doi: 10.1371/journal.pone.0082491

[46] N. Schaffert, T. B. Janzen, K. Mattes, and M. H. Thaut, "A Review on the Relationship Between Sound and Movement in Sports and Rehabilitation," *Frontiers in Psychology*, vol. 10, no. FEB, pp. 1–20, 2 2019. doi: 10.3389/fpsyg.2019.00244

[47] D. M. Wolpert, J. Diedrichsen, and J. R. Flanagan, "Principles of sensorimotor learning," *Nature Reviews Neuroscience*, vol. 12, no. 12, pp. 739–751, 12 2011. doi: 10.1038/nrn3112

[48] J. F. Dyer, P. Stapleton, and M. W. M. Rodger, "Sonification as Concurrent Augmented Feedback for Motor Skill Learning and the Importance of Mapping Design," *The Open Psychology Journal*, vol. 8, no. 1, pp. 192–202, 12 2015. doi: 10.2174/1874350101508010192

[49] E. Kohler, C. Keysers, M. A. Umiltà, L. Fogassi, V. Gallese, and G. Rizzolatti, "Hearing sounds, understanding actions: action representation in mirror neurons." *Science (New York, N.Y.)*, vol. 297, no. 5582, pp. 846–8, 8 2002. doi: 10.1126/science.1070311

[50] A. O. Effenberg, U. Fehse, G. Schmitz, B. Krueger, and H. Mechling, "Movement Sonification: Effects on Motor Learning beyond Rhythmic Adjustments,"

Frontiers in Neuroscience, vol. 10, no. MAY, pp. 1–18, 5 2016. doi: 10.3389/fnins.2016.00219

[51] G. Schmitz, B. Mohammadi, A. Hammer, M. Heldmann, A. Samii, T. F. Münte, and A. O. Effenberg, "Observation of sonified movements engages a basal ganglia frontocortical network," *BMC Neuroscience*, vol. 14, no. 1, pp. 1–11, 3 2013. doi: 10.1186/1471-2202-14-32/FIGURES/5

[52] T. H. Nown, P. Upadhyay, A. Kerr, I. Andonovic, C. Tachtatzis, and M. A. Grealy, "A Mapping Review of Real-Time Movement Sonification Systems for Movement Rehabilitation," *IEEE Reviews in Biomedical Engineering*, vol. 16, pp. 672–686, 2023. doi: 10.1109/RBME.2022.3187840

[53] T. Braun Janzen, Y. Koshimori, N. M. Richard, and M. H. Thaut, "Rhythm and Music-Based Interventions in Motor Rehabilitation: Current Evidence and Future Perspectives," *Frontiers in Human Neuroscience*, vol. 15, p. 843, 1 2022. doi: 10.3389/FNHUM.2021.789467/BIBTEX

[54] W. L. Magee, I. Clark, J. Tamplin, and J. Bradt, "Music interventions for acquired brain injury." *The Cochrane database of systematic reviews*, vol. 1, no. 1, p. CD006787, 1 2017. doi: 10.1002/14651858.CD006787.pub3

[55] M. H. Thaut, *Rhythm, music and the brain: scientific foundations and clinical applications*, 1st ed. New York: Routledge, 2008.

[56] M. H. Thaut, G. C. McIntosh, S. G. Prassas, and R. R. Rice, "Effect of Rhythmic Auditory Cuing on Temporal Stride Parameters and EMG. Patterns in Hemiparetic Gait of Stroke Patients," *Neurorehabilitation and Neural Repair*, vol. 7, no. 1, pp. 9–16, 1 1993. doi: 10.1177/136140969300700103

[57] M. H. Thaut, G. P. Kenyon, C. P. Hurt, G. C. McIntosh, and V. Hoemberg, "Kinematic optimization of spatiotemporal patterns in paretic arm training with stroke patients," *Neuropsychologia*, vol. 40, no. 7, pp. 1073–1081, 1 2002. doi: 10.1016/S0028-3932(01)00141-5

[58] J.-r. Kim, M.-y. Jung, E.-y. Yoo, J.-H. Park, S.-H. Kim, and J. Lee, "Effects of Rhythmic Auditory Stimulation during Hemiplegic Arm Reaching in Individuals with Stroke: An Exploratory Study," *Hong Kong Journal of Occupational Therapy*, vol. 24, no. 2, pp. 64–71, 12 2014. doi: 10.1016/j.hkjot.2014.11.002

[59] M. P. Malcolm, C. Massie, and M. Thaut, "Rhythmic Auditory-Motor Entrainment Improves Hemiparetic Arm Kinematics During Reaching Movements: A Pilot Study," *Topics in Stroke Rehabilitation*, vol. 16, no. 1, pp. 69–79, 1 2009. doi: 10.1310/tsr1601-69

[60] A. Sethi, N. Stergiou, T. S. Patterson, C. Patten, and L. G. Richards, "Speed and Rhythm Affect Temporal Structure of Variability in Reaching Poststroke: A Pilot Study," *https://doi.org/10.1080/00222895.2016.1219304*, vol. 49, no. 1, pp. 35–45, 1 2016. doi: 10.1080/00222895.2016.1219304

[61] S. Kang, J. H. Shin, I. Y. Kim, J. Lee, J. Y. Lee, and E. Jeong, "Patterns of enhancement in paretic shoulder kinematics after stroke with musical cueing," *Scientific Reports 2020 10:1*, vol. 10, no. 1, pp. 1–13, 10 2020. doi: 10.1038/s41598-020-75143-0

[62] M. S. Hong, "The Development and Effects of an Upper Extremity Exercise Program Based on Patterned Sensory Enhancement for Home-bound Stroke Patients," *Journal of Korean Academy of Community Health Nursing*, vol. 22, no. 2, p. 192, 4 2011. doi: 10.12799/jkachn.2011.22.2.192

[63] S. J. Han, A. J. Kwon, and H. Y. Park, "Immediate Effect of Patterned Sensory Enhancement (PSE) on Upper Limb Function after Stroke," *Journal of Music and Human Behavior*, vol. 11, no. 1, pp. 1–19, 2014. doi: https://doi.org/10.21187/jmhb.2014.11.1.001

[64] 김보연, 김수지, and 여명선, "경직형 뇌성마비 아동의 상지기능 변화를 위한 패턴화된 감각증진(PSE) 음악 적용 연구," *Korean Journal of Physical, Multiple, & Health Disabilities*, vol. 62, no. 4, pp. 257–274, 10 2019. doi: 10.20971/kcpmd.2019.62.4.257

[65] A. A. Bukowska, P. Krężałek, E. Mirek, P. Bujas, and A. Marchewka, "Neurologic Music Therapy Training for Mobility and Stability Rehabilitation with Parkinson's Disease – A Pilot Study," *Frontiers in Human Neuroscience*, vol. 9, no. JAN2016, 1 2016. doi: 10.3389/fnhum.2015.00710

[66] S. Schneider, P. W. Schönle, E. Altenmüller, and T. F. Münte, "Using musical instruments to improve motor skill recovery following a stroke," *Journal of Neurology*, vol. 254, no. 10, pp. 1339–1346, 10 2007. doi: 10.1007/s00415-006-0523-2

[67] J. Grau-Sánchez, T. F. Münte, E. Altenmüller, E. Duarte, and A. Rodríguez-Fornells, "Potential benefits of music playing in stroke upper limb motor rehabilitation," *Neuroscience & Biobehavioral Reviews*, vol. 112, pp. 585–599, 5 2020. doi: 10.1016/j.neubiorev.2020.02.027

[68] J. L. Amengual, N. Rojo, M. Veciana de las Heras, J. Marco-Pallarés, J. Grau-Sánchez, S. Schneider, L. Vaquero, M. Juncadella, J. Montero, B. Mohammadi, F. Rubio, N. Rueda, E. Duarte, C. Grau, E. Altenmüller, T. F. Münte, and A. Rodríguez-Fornells, "Sensorimotor Plasticity after Music-Supported Therapy in Chronic Stroke Patients Revealed by Transcranial Magnetic Stimulation," *PLOS ONE*, vol. 8, no. 4, p. e61883, 4 2013. doi: 10.1371/JOURNAL.PONE.0061883

[69] S. Ghai, F. D. Maso, T. Ogourtsova, A. X. Porxas, M. Villeneuve, V. Penhune, M. H. Boudrias, S. Baillet, and A. Lamontagne, "Neurophysiological Changes Induced by Music-Supported Therapy for Recovering Upper Extremity Function after Stroke: A Case Series," *Brain sciences*, vol. 11, no. 5, 5 2021. doi: 10.3390/BRAINSCI11050666

[70] Y. Tong, B. Forreider, X. Sun, X. Geng, W. Zhang, H. Du, T. Zhang, and Y. Ding, "Music-supported therapy (MST) in improving post-stroke patients' upper-limb motor function: a randomised controlled pilot study," *Neurological Research*, vol. 37, no. 5, pp. 434–440, 5 2015. doi: 10.1179/1743132815Y.0000000034

[71] P. Ripollés, N. Rojo, J. Grau-Sánchez, J. L. Amengual, E. Càmara, J. Marco-Pallarés, M. Juncadella, L. Vaquero, F. Rubio, E. Duarte, C. Garrido, E. Altenmüller, T. F. Münte, and A. Rodríguez-Fornells, "Music supported therapy promotes motor plasticity in individuals with chronic stroke," *Brain Imaging and Behavior*, vol. 10, no. 4, pp. 1289–1307, 12 2016. doi: 10.1007/s11682-015-9498-x

[72] J. Grau-Sánchez, E. Segura, D. Sanchez-Pinsach, P. Raghavan, T. F. Münte, A. M. Palumbo, A. Turry, E. Duarte, T. Särkämö, J. Cerquides, J. L. Arcos, and A. Rodríguez-Fornells, "Enriched Music-supported Therapy for chronic stroke patients: a study protocol of a randomised controlled trial." *BMC neurology*, vol. 21, no. 1, p. 19, 1 2021. doi: 10.1186/s12883-020-02019-1

[73] A. J. Street, W. L. Magee, A. Bateman, M. Parker, H. Odell-Miller, and J. Fachner, "Home-based neurologic music therapy for arm hemiparesis following stroke: results from a pilot, feasibility randomized controlled trial," *Clinical Rehabilitation*, vol. 32, no. 1, pp. 18–28, 1 2018. doi: 10.1177/0269215517717060

[74] A. J. Street, J. Fachner, and W. L. Magee, "Upper limb rehabilitation in chronic stroke using neurologic music therapy: Two contrasting case studies to inform on treatment delivery and patient suitability," *Nordic Journal of Music Therapy*, vol. 28, no. 5, pp. 382–404, 10 2019. doi: 10.1080/08098131.2019.1606848

[75] C. M. Haire, V. Vuong, L. Tremblay, K. K. Patterson, J. L. Chen, and M. H. Thaut, "Effects of therapeutic instrumental music performance and motor imagery on chronic post-stroke cognition and affect: A randomized controlled trial," *NeuroRehabilitation*, vol. 48, no. 2, pp. 195–208, 1 2021. doi: 10.3233/NRE-208014

[76] T. L. Bonebright, P. R. Cook, J. H. Flowers, N. E. Miner, J. G. Neuhoff, R. Bargar, S. Barrass, J. Berger, G. E. Evreinov, W. T. Fitch, M. Grohn, S. Handel, H. G. Kaper, H. Levkowitz, S. K. Lodha, B. G. Shinn-Cunningham, M. Simoni, and S. Tipei, "Sonification Report: Status of the Field and Research Agenda Prepared

for the National Science Foundation by members of the International Community for Auditory Display," 1999.

[77] A. De Campo, "Towards a data sonification design space map." in *Proceedings of the 13th International Conference on Auditory Display*, Montreal, Canada, 2007.

[78] T. Hermann, A. Hunt, and J. Neuhoff, *The Sonification Handbook*, T. Hermann, A. Hunt, and J. G. Neuhoff, Eds. Berlin: Logos Verlag, 2011.

[79] S. Ghai, "Effects of Real-Time (Sonification) and Rhythmic Auditory Stimuli on Recovering Arm Function Post Stroke: A Systematic Review and Meta-Analysis," *Frontiers in Neurology*, vol. 9, no. JUL, p. 488, 7 2018. doi: 10.3389/fneur.2018.00488

[80] B. Walker, "Magnitude Estimation of Conceptual Data Dimensions for Use in Sonification," Ph.D. dissertation, Houston, Texas, 9 2000. [Online]. Available: http://sonify.psych.gatech.edu/~walkerb/research/phd/index.html

[81] B. N. Walker and G. Kramer, "Sonification Design and Metaphors: Comments on Walker and Kramer, ICAD 1996," *ACM Transactions on Applied Perception*, vol. 2, no. 4, pp. 413–417, 2005. doi: 10.1145/1101530.1101535

[82] J. F. Dyer, P. Stapleton, and M. W. M. Rodger, "Advantages of melodic over rhythmic movement sonification in bimanual motor skill learning," *Experimental Brain Research*, vol. 235, no. 10, pp. 3129–3140, 10 2017. doi: 10.1007/s00221-017-5047-8

[83] L. Harrar and T. Stockman, "Designing auditory graph overviews: An examination of discrete vs. continuous sound and the influence of presentation speed," in *Proceedings of the 13th International Conference on Auditory Display*, Montreal, Canada, 2007.

[84] A. R. Fugl-Meyer, L. Jääskö, I. Leyman, S. Olsson, and S. Steglind, "The post-stroke hemipegic patient. I: A method for evaluation of physical performance." *Scand J Rehab Med*, no. 7, pp. 13–31, 1975.

[85] R. C. Lyle, "A performance test for assessment of upper limb function in physical rehabilitation treatment and research," *International Journal of Rehabilitation Research*, vol. 4, no. 4, pp. 483–492, 12 1981. doi: 10.1097/00004356-198112000-00001

[86] J. Guerra, L. Smith, D. Vicinanza, B. Stubbs, N. Veronese, and G. Williams, "The use of sonification for physiotherapy in human movement tasks: A scoping review," *Science & Sports*, vol. 35, no. 3, pp. 119–129, 2020. doi: 10.1016/j.scispo.2019.12.004

[87] R. A. Maulucci and R. H. Eckhouse, "Retraining reaching in chronic stroke with real-time auditory feedback," *NeuroRehabilitation*, vol. 16, no. 3, pp. 171–182, 10 2001. doi: 10.3233/NRE-2001-16306

[88] I. Wallis, T. Ingalls, T. Rikakis, L. Olsen, Y. Chen, W. Xu, and H. Sundaram, "Real-Time Sonification of Movement for an Immersive Stroke Rehabilitation Environment," in *Proceedings of the 13th International Conference on Auditory Display, (ICAD-2007)*, 2007.

[89] M. Duff, Y. Chen, L. Cheng, S.-M. Liu, P. Blake, S. L. Wolf, and T. Rikakis, "Adaptive Mixed Reality Rehabilitation Improves Quality of Reaching Movements More Than Traditional Reaching Therapy Following Stroke," *Neurorehabilitation and Neural Repair*, vol. 27, no. 4, pp. 306–315, 5 2013. doi: 10.1177/1545968312465195

[90] Y. Chen, M. Duff, N. Lehrer, H. Sundaram, J. He, S. L. Wolf, and T. Rikakis, "A Computational Framework for Quantitative Evaluation of Movement during Rehabilitation," *AIP Conference Proceedings*, vol. 1371, no. 1, p. 317, 6 2011. doi: 10.1063/1.3596656

[91] J. V. Robertson, T. Hoellinger, P. Lindberg, D. Bensmail, S. Hanneton, and A. Roby-Brami, "Effect of auditory feedback differs according to side of hemiparesis: a comparative pilot study," *Journal of NeuroEngineering and Rehabilitation*, vol. 6, no. 1, p. 45, 12 2009. doi: 10.1186/1743-0003-6-45

[92] G. Schmitz, D. Kroeger, and A. O. Effenberg, “A mobile sonification system for stroke rehabilitation,” in *The 20th International Conference on Auditory Display (ICAS-2014)*, 2014.

[93] H. Brock, G. Schmitz, J. Baumann, and A. O. Effenberg, “If motion sounds: Movement sonification based on inertial sensor data,” *Procedia Engineering*, vol. 34, pp. 556–561, 2012. doi: 10.1016/j.proeng.2012.04.095

[94] G. Schmitz, J. Bergmann, A. O. Effenberg, C. Krewer, T.-H. Hwang, and F. Müller, “Movement Sonification in Stroke Rehabilitation,” *Frontiers in Neurology — www.frontiersin.org*, vol. 1, p. 389, 2018. doi: 10.3389/fneur.2018.00389

[95] D. S. Scholz, S. Rohde, N. Nikmaram, H.-P. Brückner, M. Großbach, J. D. Rollnik, and E. O. Altenmüller, “Sonification of Arm Movements in Stroke Rehabilitation – A Novel Approach in Neurologic Music Therapy,” *Frontiers in Neurology* —, vol. 7, p. 106, 2016. doi: 10.3389/fneur.2016.00106

[96] N. Nikmaram, D. S. Scholz, M. Großbach, S. B. Schmidt, J. Spogis, P. Belardinelli, F. Müller-Dahlhaus, J. Remy, U. Ziemann, J. D. Rollnik, and E. Altenmüller, “Musical Sonification of Arm Movements in Stroke Rehabilitation Yields Limited Benefits,” *Frontiers in Neuroscience*, vol. 13, p. 1378, 12 2019. doi: 10.3389/fnins.2019.01378

[97] N. Friedman, V. Chan, A. N. Reinkensmeyer, A. Beroukhim, G. J. Zambrano, M. Bachman, and D. J. Reinkensmeyer, “Retraining and assessing hand movement after stroke using the MusicGlove: comparison with conventional hand therapy and isometric grip training,” *Journal of NeuroEngineering and Rehabilitation*, vol. 11, no. 1, p. 76, 4 2014. doi: 10.1186/1743-0003-11-76

[98] R. Colombo, A. Raglio, M. Panigazzi, A. Mazzone, G. Bazzini, C. Imarisio, D. Molteni, C. Caltagirone, and M. Imbriani, “The SonicHand Protocol for Rehabilitation of Hand Motor Function: A Validation and Feasibility Study,”

IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 27, no. 4, pp. 664–672, 4 2019. doi: 10.1109/TNSRE.2019.2905076

[99] A. Raglio, M. Panigazzi, R. Colombo, M. Tramontano, M. Iosa, S. Mastrogiacomo, P. Baiardi, D. Molteni, E. Baldissarro, C. Imbriani, C. Imarisio, L. Eretti, M. Hamedani, C. Pistarini, M. Imbriani, G. L. Mancardi, and C. Caltagirone, "Hand rehabilitation with sonification techniques in the subacute stage of stroke," *Scientific Reports*, vol. 11, no. 1, p. 7237, 12 2021. doi: 10.1038/S41598-021-86627-Y

[100] I. Peyre, A. Roby-Brami, M. Segalen, A. Giron, B. Caramiaux, V. Marchand-Pauvert, P. Pradat-Diehl, and F. Bevilacqua, "Effect of sonification types in upper-limb movement: a quantitative and qualitative study in hemiparetic and healthy participants," *Journal of NeuroEngineering and Rehabilitation*, vol. 20, no. 1, p. 136, 10 2023. doi: 10.1186/s12984-023-01248-y

[101] Q. Wang, P. Markopoulos, B. Yu, W. Chen, and A. Timmermans, "Interactive wearable systems for upper body rehabilitation: a systematic review," *Journal of NeuroEngineering and Rehabilitation*, vol. 14, no. 1, p. 20, 12 2017. doi: 10.1186/s12984-017-0229-y

[102] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *BMJ*, vol. 372, p. n71, 3 2021. doi: 10.1136/bmj.n71

[103] F. T. van Vugt and B. Tillmann, "Thresholds of Auditory-Motor Coupling Measured with a Simple Task in Musicians and Non-Musicians: Was the Sound Simultaneous to the Key Press?" *PLoS ONE*, vol. 9, no. 2, p. e87176, 2 2014. doi: 10.1371/journal.pone.0087176

[104] F. H. Martini, J. L. Nath, and E. F. Bartholomew, *Fundamentals of Anatomy & Physiology, Global Edition*, 11th ed. Pearson Education Limited, 2018.

[105] E. Van Der Kruk and M. M. Reijne, "Accuracy of human motion capture systems for sport applications; state-of-the-art review," *European Journal of Sport Science*, vol. 18, no. 6, pp. 806–819, 2018. doi: 10.1080/17461391.2018.1463397

[106] G. Spina, G. Huang, A. Vaes, M. Spruit, and O. Amft, "COPDTrainer," in *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing.* New York, NY, USA: ACM, 9 2013. doi: 10.1145/2493432.2493454 pp. 597–606.

[107] B. Stahl and I. Zmölnig, "Musical Sonification in Electronic Therapy AIDS for Motor-functional Treatment - a Smartphone Approach," in *SMC 2016 - 13th Sound and Music Computing Conference, Proceedings*, 2016. doi: 10.5281/zenodo.851307 pp. 448–454.

[108] H.-P. Brückner, S. Lesse, W. Theimer, and H. Blume, "Design space exploration of hardware platforms for interactive low latency movement sonification," *Journal on Multimodal User Interfaces*, vol. 10, no. 1, pp. 1–11, 3 2016. doi: 10.1007/s12193-015-0199-y

[109] H.-P. Bruckner, W. Theimer, and H. Blume, "Real-time low latency movement sonification in stroke rehabilitation based on a mobile platform," in *2014 IEEE International Conference on Consumer Electronics (ICCE).* IEEE, 1 2014. doi: 10.1109/ICCE.2014.6775997 pp. 264–265.

[110] F. Bevilacqua, I. Peyre, M. Segalen, P. Pradat-Diehl, V. Marchand-Pauvert, and A. Roby-Brami, "Exploring different movement sonification strategies for rehabilitation in clinical settings," in *MOCO '18: Proceedings of the 5th International Conference on Movement and Computing*, 2018. doi: 10.1145/3212721.3212881 pp. 1–6.

[111] G. Alankus and C. Kelleher, "Reducing compensatory motions in video games for stroke rehabilitation," in *Proceedings of the SIGCHI Conference on Human*

Factors in Computing Systems. New York, NY, USA: ACM, 5 2012. doi: 10.1145/2207676.2208354 pp. 2049–2058.

[112] D. Dotov and T. Froese, "Entraining chaotic dynamics: A novel movement sonification paradigm could promote generalization," *Human Movement Science*, vol. 61, pp. 27–41, 10 2018. doi: 10.1016/j.humov.2018.06.016

[113] K. Seko and K. Fukuchi, "A guidance technique for motion tracking with a handheld camera using auditory feedback," in *Adjunct proceedings of the 25th annual ACM symposium on User interface software and technology - UIST Adjunct Proceedings '12*. New York, New York, USA: ACM Press, 2012. doi: 10.1145/2380296.2380339 p. 95.

[114] F. Feltham, J. Curtis, and L. Loke, "The Prefix/Suffix Model," in *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction*. New York, NY, USA: ACM, 2 2020. doi: 10.1145/3374920.3374985 pp. 545–550.

[115] Y. Tanaka, H. Kon, and H. Koike, "A real-time golf-swing training system using sonification and sound image localization," in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. New York, NY, USA: ACM, 11 2018. doi: 10.1145/3281505.3281604 pp. 1–2.

[116] E. D. Hebling, E. Partesotti, C. P. Santana, A. Figueiredo, C. G. Dezotti, T. Botechia, C. A. P. da Silva, M. A. da Silva, D. Rossetti, V. A. W. de Oliveira, S. Cielavin, A. S. Moroni, and J. Manzolli, "MovieScape," in *Proceedings of the 9th International Conference on Digital and Interactive Arts*, ser. ARTECH 2019. New York, NY, USA: ACM, 10 2019. doi: 10.1145/3359852.3359883 pp. 1–7.

[117] M. Tölgyessy, M. Dekan, Chovanec, and P. Hubinský, "Evaluation of the azure kinect and its comparison to kinect v1 and kinect v2," *Sensors (Switzerland)*, vol. 21, no. 2, pp. 1–25, 2021. doi: 10.3390/s21020413

[118] S. Ghisio, E. Volta, P. Alborno, M. Gori, and G. Volpe, "An open platform for full-body multisensory serious-games to teach geometry in primary school," in *Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education*. New York, NY, USA: ACM, 11 2017. doi: 10.1145/3139513.3139523 pp. 49–52.

[119] F. L. Cibrian, J. Ley-Flores, J. W. Newbold, A. Singh, N. Bianchi-Berthouze, and M. Tentori, "Interactive sonification to assist children with autism during motor therapeutic interventions," *Personal and Ubiquitous Computing*, vol. 25, no. 2, pp. 391–410, 4 2021. doi: 10.1007/s00779-020-01479-z

[120] A. Singh, N. Bianchi-Berthouze, and A. C. Williams, "Supporting Everyday Function in Chronic Pain Using Wearable Technology," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, vol. 2017-May. New York, NY, USA: ACM, 5 2017. doi: 10.1145/3025453.3025947 pp. 3903–3915.

[121] A. Singh, S. Piana, D. Pollarolo, G. Volpe, G. Varni, A. Tajadura-Jiménez, A. C. Williams, A. Camurri, and N. Bianchi-Berthouze, "Go-with-the-Flow : Tracking, Analysis and Sonification of Movement and Breathing to Build Confidence in Activity Despite Chronic Pain," *Human–Computer Interaction*, vol. 31, no. 3-4, pp. 335–383, 7 2016. doi: 10.1080/07370024.2015.1085310

[122] J. W. Newbold, N. Bianchi-Berthouze, N. E. Gold, A. Tajadura-Jiménez, and A. C. Williams, "Musically Informed Sonification for Chronic Pain Rehabilitation," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, 5 2016. doi: 10.1145/2858036.2858302 pp. 5698–5703.

[123] G. Ozkaya, H. R. Jung, I. S. Jeong, M. R. Choi, M. Y. Shin, X. Lin, W. S. Heo, M. S. Kim, E. Kim, and K.-K. Lee, "Three-dimensional motion capture data during repetitive overarm throwing practice," *Scientific Data*, vol. 5, no. 1, p. 180272, 12 2018. doi: 10.1038/sdata.2018.272

[124] Y. Chen, H. Huang, W. Xu, R. I. Wallis, H. Sundaram, T. Rikakis, T. Ingalls, L. Olson, and J. He, "The design of a real-time, multimodal biofeedback system for stroke patient rehabilitation," in *Proceedings of the 14th annual ACM international conference on Multimedia - MULTIMEDIA '06.* New York, New York, USA: ACM Press, 2006. doi: 10.1145/1180639.1180804 p. 763.

[125] K. Vogt, D. Pirrò, I. Kobenz, R. Höldrich, and G. Eckel, "PhysioSonic - Evaluated Movement Sonification as Auditory Feedback in Physiotherapy," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2010, vol. 5954 LNCS, pp. 103–120. [Online]. Available: http://link.springer.com/10.1007/978-3-642-12439-6_6

[126] A. I. Dailly, R. Sigrist, Y. Kim, P. Wolf, H. Erckens, J. Cerny, A. Luft, R. Gassert, and J. Sulzer, "Can simple error sonification in combination with music help improve accuracy in upper limb movements?" in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob).* IEEE, 6 2012. doi: 10.1109/BioRob.2012.6290908 pp. 1423–1427.

[127] E. Gamboa, A. Serrato, D. Toro, and M. Trujillo, "Advantages and limitations of leap motion for developing physical rehabilitation exergames (PREGs)," in *Proceedings of the 5th Workshop on ICTs for improving Patients Rehabilitation Research Techniques.* New York, NY, USA: ACM, 9 2019. doi: 10.1145/3364138.3364149 pp. 43–46.

[128] E. Frid, J. Moll, R. Bresin, and E.-L. Sallnäs Pysander, "Haptic feedback combined with movement sonification using a friction sound improves task performance in a virtual throwing task," *Journal on Multimodal User Interfaces*, vol. 13, no. 4, pp. 279–290, 12 2019. doi: 10.1007/s12193-018-0264-4

[129] M. C. Rodriguez, M. Rossini, G. Caruso, G. Samali, C. Giovanzana, F. Molteni, and M. Bordegoni, "Sound Feedback Assessment for Upper Limb Rehabilitation Using a Multimodal Guidance System," in *ICCHP 2016: Computers Helping*

People with Special Needs, 2016, vol. 9759, pp. 529–536. [Online]. Available: http://link.springer.com/10.1007/978-3-319-41267-2_74

[130] L. Véron-Delor, S. Pinto, A. Eusebio, J. Azulay, T. Witjas, J. Velay, and J. Danna, "Musical sonification improves motor control in Parkinson's disease: a proof of concept with handwriting," *Annals of the New York Academy of Sciences*, vol. 1465, no. 1, pp. 132–145, 4 2020. doi: 10.1111/nyas.14252

[131] M. Ten Bhömer, O. Tomico, and C. Hummels, "Vigour: Smart textile services to support rehabilitation," in *Nordic Research Conference 2013*, 2013, pp. 505–506.

[132] N. Friedman, D. Reinkensmeyer, M. Bachman, and H. Samueli, "A Real-Time Interactive MIDI Glove for Domicile Stroke Rehabilitation," in *International Conference on Human-Computer Interaction*, vol. 6764, 2011, pp. 151–158.

[133] D. Johnson, D. Damian, and G. Tzanetakis, "OSC-XR: A Toolkit for Extended Reality Immersive Music Interfaces," in *Proceedings of the Sound and Music Computing Conferences*, 2019, pp. 202–209.

[134] D. Ahmetovic, D. Sato, U. Oh, T. Ishihara, K. Kitani, and C. Asakawa, "ReCog: Supporting Blind People in Recognizing Personal Objects," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, 4 2020. doi: 10.1145/3313831.3376143 pp. 1–12.

[135] J. Ramsay and H. J. Chang, "Body Pose Sonification for a View-Independent Auditory Aid to Blind Rock Climbers," in *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 3 2020. doi: 10.1109/WACV45572.2020.9093462 pp. 3403–3410.

[136] K. Lee, J. Hong, S. Pimento, E. Jarjue, and H. Kacorri, "Revisiting Blind Photography in the Context of Teachable Object Recognizers," in *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*, ser. ASSETS '19. New York, NY, USA: ACM, 10 2019. doi: 10.1145/3308561.3353799 pp. 83–95.

[137] R. Kirby, "Development of a real-time performance measurement and feedback system for alpine skiers," *Sports Technology*, vol. 2, no. 1-2, pp. 43–52, 4 2009. doi: 10.1002/jst.85

[138] A. Aska and M. Ritter, "Approaches to real time ambisonic spatialization and sound diffusion using motion capture," in *ICMC 2016 - 42nd International Computer Music Conference, Proceedings*, 2016, pp. 327–332.

[139] R. Valenti, A. Jaimes, and N. Sebe, "Sonify your face," in *Proceedings of the international conference on Multimedia - MM '10*. New York, New York, USA: ACM Press, 2010. doi: 10.1145/1873951.1874219 p. 1363.

[140] P. Manikashani and J. E. Boyd, "A phase-entrained particle filter for audio-locomotion synchronization," in *Proceedings - 2016 13th Conference on Computer and Robot Vision, CRV 2016*, 2016. doi: 10.1109/CRV.2016.24 pp. 242–249.

[141] E. Kabisch, F. Kuester, and S. Penny, "Sonic panoramas," in *Proceedings of the 2005 international conference on Augmented tele-existence - ICAT '05*, vol. 157. New York, New York, USA: ACM Press, 2005. doi: 10.1145/1152399.1152428 p. 156.

[142] N. Konttinen, K. Mononen, J. Viitasalo, and T. Mets, "The Effects of Augmented Auditory Feedback on Psychomotor Skill Learning in Precision Shooting," *Journal of Sport and Exercise Psychology*, vol. 26, no. 2, pp. 306–316, 6 2004. doi: 10.1123/jsep.26.2.306

[143] T. Hermann, T. Henning, and H. Ritter, "Gesture desk - An integrated multi-modal gestural workplace for sonification," in *Gesture-Based Communication in Human-Computer Interaction*, ser. Lecture Notes in Artificial Intelligence, A. Camurri and G. Volpe, Eds., 2003, vol. 2915, pp. 369–379. ISBN 3-540-21072-5. [Online]. Available: https://link.springer.com/content/pdf/10.1007%2Fb95740.pdf

[144] T. Hermann, C. Nölker, and H. Ritter, "Hand Postures for Sonification Control," in *Gesture and Sign Language in Human-Computer Interaction*. Springer Verlag, 2002, vol. 2298, pp. 307–316. ISBN 9783540436782. [Online]. Available: http://link.springer.com/10.1007/3-540-47873-6_32

[145] R. Sigrist, S. Fox, R. Riener, and P. Wolf, "Benefits of Crank Moment Sonification in Cycling," *Procedia Engineering*, vol. 147, pp. 513–518, 2016. doi: 10.1016/j.proeng.2016.06.230

[146] N. Schaffert, A. Godbout, S. Schlueter, and K. Mattes, "Towards an application of interactive sonification for the forces applied on the pedals during cycling on the Wattbike ergometer," *Displays*, vol. 50, pp. 41–48, 12 2017. doi: 10.1016/j.displa.2017.09.004

[147] B. O'Brien, R. Hardouin, G. Rao, D. Bertin, and C. Bourdin, "Online sonification improves cycling performance through kinematic and muscular reorganisations," *Scientific Reports*, vol. 10, no. 1, p. 20929, 12 2020. doi: 10.1038/s41598-020-76498-0

[148] P.-J. Maes, V. Lorenzoni, and J. Six, "The SoundBike: musical sonification strategies to enhance cyclists' spontaneous synchronization to external music," *Journal on Multimodal User Interfaces*, vol. 13, no. 3, pp. 155–166, 9 2019. doi: 10.1007/s12193-018-0279-x

[149] I. Sanz-Pena, J. Blanco, and J. H. Kim, "Computer Interface for Real-Time Gait Biofeedback Using a Wearable Integrated Sensor System for Data Acquisition," *IEEE Transactions on Human-Machine Systems*, vol. 51, no. 5, pp. 484–493, 10 2021. doi: 10.1109/THMS.2021.3090738

[150] T. Hermann and S. Zehe, "Sonified Aerobics-Interactive Sonification of Coordinated Body Movements," in *Proceedings of the 17th Annual Conference on Auditory Display*, 2011, pp. 1–6.

[151] R. F. Hale, S. Dorgo, R. V. Gonzalez, and J. Hausselle, "The Efficacy of Simultaneously Training 2 Motion Targets During a Squat Using Auditory

Feedback," *Journal of Applied Biomechanics*, vol. 37, no. 1, pp. 6–12, 2 2021. doi: 10.1123/jab.2019-0276

[152] S. Fujii, T. Lulic, and J. L. Chen, "More Feedback Is Better than Less: Learning a Novel Upper Limb Joint Coordination Pattern with Augmented Auditory Feedback," *Frontiers in Neuroscience*, vol. 10, no. JUN, 6 2016. doi: 10.3389/fnins.2016.00251

[153] A. Tajadura-Jiménez, J. Newbold, L. Zhang, P. Rick, and N. Bianchi-Berthouze, "As Light as You Aspire to Be," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. New York, NY, USA: ACM, 5 2019. doi: 10.1145/3290605.3300888 pp. 1–14.

[154] L. Turchet, "Interactive sonification and the IoT," in *Proceedings of the 14th International Audio Mostly Conference: A Journey in Sound*. New York, NY, USA: ACM, 9 2019. doi: 10.1145/3356590.3356631 pp. 252–255.

[155] J. Maculewicz, C. Erkut, and S. Serafin, "An investigation on the impact of auditory and haptic feedback on rhythmic walking interactions," *International Journal of Human-Computer Studies*, vol. 85, pp. 40–46, 1 2016. doi: 10.1016/j.ijhcs.2015.07.003

[156] R. Pugliese and T. Takala, "Sonic Trampoline: How Audio Feedback Impacts the User's Experience of Jumping," *IEEE MultiMedia*, vol. 22, no. 1, pp. 74–79, 1 2015. doi: 10.1109/MMUL.2015.13

[157] D. Bisig and P. Palacio, "Sounding feet," in *Proceedings of the 15th International Conference on Audio Mostly*. New York, NY, USA: ACM, 9 2020. doi: 10.1145/3411109.3411112 pp. 222–228.

[158] T. Grosshauser, B. Bläsing, C. Spieth, and T. Hermann, "Wearable sensor-based real-time sonification of motion and foot pressure in dance teaching and training," *AES: Journal of the Audio Engineering Society*, vol. 60, no. 7-8, pp. 580–589, 2012.

[159] J. Ley-Flores, F. Bevilacqua, N. Bianchi-Berthouze, and A. Taiadura-Jimenez, "Altering body perception and emotion in physically inactive people through movement sonification," in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 9 2019. doi: 10.1109/ACII.2019.8925432 pp. 1–7.

[160] I. Bergström and M. Jonsson, "Sarka," in *Proceedings of the 3rd International Symposium on Movement and Computing*, vol. 05-06-July. New York, NY, USA: ACM, 7 2016. doi: 10.1145/2948910.2948922 pp. 1–8.

[161] L. Turchet, "Custom made wireless systems for interactive footstep sounds synthesis," *Applied Acoustics*, vol. 83, pp. 22–31, 9 2014. doi: 10.1016/j.apacoust.2014.03.005

[162] S. Hasegawa, S. Ishijima, F. Kato, H. Mitake, and M. Sato, "Realtime sonification of the center of gravity for skiing," in *Proceedings of the 3rd Augmented Human International Conference on - AH '12*, ser. AH '12. New York, New York, USA: ACM Press, 2012. doi: 10.1145/2160125.2160136 pp. 1–4.

[163] A.-M. Gorgas, L. Schön, R. Dlapka, J. Doppler, M. Iber, C. Gradl, A. Kiselka, T. Siragy, and B. Horsak, "Short-Term Effects of Real-Time Auditory Display (Sonification) on Gait Parameters in People with Parkinsons' Disease—A Pilot Study," in *Biosystems and Biorobotics*, 2017, vol. 15, pp. 855–859. [Online]. Available: http://link.springer.com/10.1007/978-3-319-46669-9_139

[164] B. Horsak, R. Dlapka, M. Iber, A.-M. Gorgas, A. Kiselka, C. Gradl, T. Siragy, and J. Doppler, "SONIGait: a wireless instrumented insole device for real-time sonification of gait," *Journal on Multimodal User Interfaces*, vol. 10, no. 3, pp. 195–206, 9 2016. doi: 10.1007/s12193-016-0216-9

[165] T. Fischer, A. Kiselka, R. Dlapka, J. Doppler, M. Iber, C. Gradl, A.-M. Gorgas, T. Siragy, and B. Horsak, "An Auditory Feedback System in Use with People Aged +50 Years: Compliance and Modifications in Gait Pattern," in *Biosystems*

and Biorobotics. Springer International Publishing, 2017, vol. 15, pp. 881–885. [Online]. Available: http://link.springer.com/10.1007/978-3-319-46669-9_143

[166] M. Batavia, J. G. Gianutsos, A. Vaccaro, and J. T. Gold, "A do-it-yourself membrane-activated auditory feedback device for weight bearing and gait training: A case report," *Archives of Physical Medicine and Rehabilitation*, vol. 82, no. 4, pp. 541–545, 4 2001. doi: 10.1053/apmr.2001.21931

[167] D. Cesarini, D. Calvaresi, C. Farnesi, D. Taddei, S. Frediani, B. E. Ungerechts, and T. Hermann, "MEDIATION : An eMbEddeD System for Auditory Feedback of Hand-water InterAcTION while Swimming," *Procedia Engineering*, vol. 147, pp. 324–329, 2016. doi: 10.1016/j.proeng.2016.06.301

[168] D. Cesarini, B. E. Ungerechts, and T. Hermann, "Swimmers in the loop: Sensing moving water masses for an auditory biofeedback system," in *2015 IEEE Sensors Applications Symposium (SAS)*. IEEE, 4 2015. doi: 10.1109/SAS.2015.7133578 pp. 1–6.

[169] T. Hermann, J. Krause, and H. Ritter, "Real-Time Control of Sonification Models with an Audio-Haptic Interface," in *Proceedings of the 2002 International Conference on Auditory Display*, Kyoto, Japan, 2002, pp. 1–5.

[170] S. Akiyama, K. Sato, Y. Makino, and T. Maeno, "EffectON — Enriching impressions of motions and physically changing motions via synchronous sound effects," in *The 6th International Conference on Soft Computing and Intelligent Systems, and The 13th International Symposium on Advanced Intelligence Systems*, ser. Joint International Conference on Soft Computing and Intelligent Systems SCIS and International Symposium on Advanced Intelligent Systems ISIS. IEEE, 11 2012. doi: 10.1109/SCIS-ISIS.2012.6505218 pp. 856–860.

[171] A. G. E. Mulder, S. S. Fels, and K. Mase, "Design of virtual 3D instruments for musical interaction," in *Graphics Interface '99, Proceedings*, ser. Proceedings - Graphics Interface, I. S. MacKenzie and J. Stewart, Eds., 1999, pp. 76–83. ISBN 1-55860-632-7

[172] D. Brown, C. Nash, and T. J. Mitchell, "Was that me?" in *Proceedings of the 15th International Audio Mostly Conference.* New York, NY, USA: ACM, 9 2020. doi: 10.1145/3411109.3411137 pp. 168–174.

[173] N. Diniz, P. Coussement, A. Deweppe, M. Demey, and M. Leman, "An embodied music cognition approach to multilevel interactive sonification," *Journal on Multimodal User Interfaces*, vol. 5, no. 3-4, pp. 211–219, 5 2012. doi: 10.1007/s12193-011-0084-2

[174] Y. Nakayama, Y. Takano, M. Matsubara, K. Suzuki, and H. Terasawa, "The sound of smile: Auditory biofeedback of facial EMG activity," *Displays*, vol. 47, pp. 32–39, 4 2017. doi: 10.1016/j.displa.2016.09.002

[175] M. Donnarumma, B. Caramiaux, and A. Tanaka, "Muscular Interactions: Combining EMG and MMG sensing for musical practice," in *NIME '13: Proceedings of the International Conference on New Interfaces for Musical Expression*, 2013.

[176] R. Sigrist, G. Rauter, L. Marchal-Crespo, R. Riener, and P. Wolf, "Sonification and haptic feedback in addition to visual feedback enhances complex motor task learning," *Experimental Brain Research*, vol. 233, no. 3, pp. 909–925, 3 2015. doi: 10.1007/s00221-014-4167-7

[177] B. van der Vlist, C. Bartneck, and S. Mäueler, "moBeat: Using Interactive Music to Guide and Motivate Users During Aerobic Exercising," *Applied Psychophysiology and Biofeedback*, vol. 36, no. 2, pp. 135–145, 6 2011. doi: 10.1007/s10484-011-9149-y

[178] C. Miyama, "Peacock: A non-haptic 3D performance interface," in *International Computer Music Conference, ICMC 2010*, 2010, pp. 443–445.

[179] A. Tajadura-Jiménez, M. Tsakiris, T. Marquardt, and N. Bianchi-Berthouze, "Action sounds update the mental representation of arm dimension: contributions of kinaesthesia and agency," *Frontiers in Psychology*, vol. 6, 5 2015. doi: 10.3389/fpsyg.2015.00689

[180] R. J. Bood, M. Nijssen, J. van der Kamp, and M. Roerdink, “The Power of Auditory-Motor Synchronization in Sports: Enhancing Running Performance by Coupling Cadence with the Right Beats,” *PLoS ONE*, vol. 8, no. 8, p. e70758, 8 2013. doi: 10.1371/journal.pone.0070758

[181] L. Baudry, D. Leroy, R. Thouvarecq, and D. Chollet, “Auditory concurrent feedback benefits on the circle performed in gymnastics,” *Journal of Sports Sciences*, vol. 24, no. 2, pp. 149–156, 2 2006. doi: 10.1080/02640410500130979

[182] A. Godbout and J. E. Boyd, “Rhythmic sonic feedback for speed skating by real-time movement synchronization,” *International Journal of Computer Science in Sport*, vol. 11, no. 3, pp. 37–51, 2012.

[183] D. Chollet, J. P. Micallef, and P. Rabischong, “Effects of two types of biomechanical bio-feedback on crawl performance,” in *Biomechanics and Medicine in Swimming: Swimming Science VI*, D. Maclaren, T. Reilly, and A. Lees, Eds. Spon, 1992, pp. 57–62.

[184] D. Chollet, J. P. Micallef, and P. Rabischong, “Biomechanical signals for external biofeedback to improve swimming techniques,” *Biomechanics and Medicine in Swimming: Swimming Science V*, pp. 389–396, 1988.

[185] S. Ghai, G. Schmitz, T.-H. Hwang, and A. O. Effenberg, “Auditory Proprioceptive Integration: Effects of Real-Time Kinematic Auditory Feedback on Knee Proprioception,” *Frontiers in Neuroscience*, vol. 12, no. MAR, 3 2018. doi: 10.3389/fnins.2018.00142

[186] P. M. Vinken, D. Kröger, U. Fehse, G. Schmitz, H. Brock, and A. O. Effenberg, “Auditory Coding of Human Movement Kinematics,” *Multisensory Research*, vol. 26, no. 6, pp. 533–552, 2013. doi: 10.1163/22134808-00002435

[187] D. S. Scholz, L. Wu, J. Pirzer, J. Schneider, J. D. Rollnik, M. Großbach, and E. O. Altenmüller, “Sonification as a possible stroke rehabilitation strategy,” *Frontiers in Neuroscience*, vol. 8, no. OCT, p. 332, 2014. doi: 10.3389/FNINS.2014.00332/BIBTEX

[188] S. Ghisio, P. Coletta, S. Piana, P. Alborno, G. Volpe, A. Camurri, L. Primavera, C. Ferrari, C. M. Guenza, P. Moretti, V. Bergamaschi, and A. Ravaschio, "An open platform for full body interactive sonification exergames," in *2015 7th International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN)*, 6 2015, pp. 168–175.

[189] D. S. Scholz, S. Rhode, M. Großbach, J. Rollnik, and E. Altenmüller, "Moving with music for stroke rehabilitation: a sonification feasibility study," *Annals of the New York Academy of Sciences*, vol. 1337, no. 1, pp. 69–76, 3 2015. doi: 10.1111/nyas.12691

[190] S. Ghai, G. Schmitz, T.-H. Hwang, and A. O. Effenberg, "Training proprioception with sound: effects of real-time auditory feedback on intermodal learning," *Annals of the New York Academy of Sciences*, vol. 1438, no. 1, pp. 50–61, 2 2019. doi: 10.1111/nyas.13967

[191] E. Frid, R. Bresin, P. Alborno, and L. Elblaus, "Interactive Sonification of Spontaneous Movement of Children—Cross-Modal Mapping and the Perception of Body Movement Qualities through Sound," *Frontiers in Neuroscience*, vol. 10, no. 521, 11 2016. doi: 10.3389/fnins.2016.00521

[192] B. O'Brien, B. Juhas, M. Bieńkiewicz, F. Buloup, L. Bringoux, and C. Bourdin, "Online sonification for golf putting gesture: reduced variability of motor behaviour and perceptual judgement," *Experimental Brain Research*, vol. 238, no. 4, pp. 883–895, 4 2020. doi: 10.1007/s00221-020-05757-3

[193] B. O'Brien, B. Juhas, M. Bieńkiewicz, F. Buloup, L. Bringoux, and C. Bourdin, "Sonification of Golf Putting Gesture Reduces Swing Movement Variability in Novices," *Research Quarterly for Exercise and Sport*, vol. 92, no. 3, pp. 301–310, 7 2021. doi: 10.1080/02701367.2020.1726859

[194] O. Boyer, F. Bevilacqua, P. Susini, and S. Hanneton, "Investigating three types of continuous auditory feedback in visuo-manual tracking," *Experimental Brain Research*, vol. 235, no. 3, pp. 691–701, 3 2017. doi: 10.1007/s00221-016-4827-x

[195] J. Dyer, P. Stapleton, and M. Rodger, "Transposing musical skill: sonification of movement as concurrent augmented feedback enhances learning in a bimanual task," *Psychological Research*, vol. 81, no. 4, pp. 850–862, 7 2017. doi: 10.1007/s00426-016-0775-0

[196] G. Jakus, K. Stojmenova, S. Tomažič, and J. Sodnik, "A system for efficient motor learning using multimodal augmented feedback," *Multimedia Tools and Applications*, vol. 76, no. 20, pp. 20 409–20 421, 10 2017. doi: 10.1007/s11042-016-3774-7

[197] A. A. Gref, L. Elblaus, and K. F. Hansen, "Sonification as Catalyst in Training Manual Wheelchair Operation for Sports and Everyday Life," in *Proceedings of the Sound and Music Computing Conference*, 2016, pp. 9–14.

[198] Q. Wang, P. Turaga, G. Coleman, and T. Ingalls, "SomaTech," in *CHI '14 Extended Abstracts on Human Factors in Computing Systems.* New York, NY, USA: ACM, 4 2014. doi: 10.1145/2559206.2581339 pp. 1765–1770.

[199] A. Ikeda, Y. Tanaka, D.-H. Hwang, H. Kon, and H. Koike, "Golf training system using sonification and virtual shadow," in *ACM SIGGRAPH 2019 Emerging Technologies.* New York, NY, USA: ACM, 7 2019. doi: 10.1145/3305367.3327993 pp. 1–2.

[200] A. Schlegel and C. Honnet, "Digital Oxymorons," in *Proceedings of the 4th International Conference on Movement Computing*, vol. Part F1291. New York, NY, USA: ACM, 6 2017. doi: 10.1145/3077981.3078040 pp. 1–8.

[201] E. Volta, P. Alborno, M. Gori, and G. Volpe, "Designing a Multisensory Social Serious-Game for Primary School Mathematics Learning," in *2018 IEEE Games, Entertainment, Media Conference (GEM).* IEEE, 8 2018. doi: 10.1109/GEM.2018.8516442 pp. 1–9.

[202] E. Volta, P. Alborno, M. Gori, S. Ghisio, S. Piana, and G. Volpe, "Enhancing children understanding of mathematics with multisensory technology," in

Proceedings of the 5th International Conference on Movement and Computing. New York, NY, USA: ACM, 6 2018. doi: 10.1145/3212721.3212889 pp. 1–4.

[203] S. Russell, G. Dublon, and J. A. Paradiso, "HearThere," in *Proceedings of the 7th Augmented Human International Conference 2016*, vol. 25-27-Febr. New York, NY, USA: ACM, 2 2016. doi: 10.1145/2875194.2875247 pp. 1–8.

[204] D. Avissar, C. Leider, C. Bennett, and R. Gailey, "An audio game app using interactice Movement sonification for Targeted posture control," in *Proceedings of the International Conference on Auditory Display*, 2013, pp. 45–48.

[205] C. Franco, A. Fleury, P. Y. Gumery, B. Diot, J. Demongeot, and N. Vuillerme, "iBalance-ABF: A Smartphone-Based Audio-Biofeedback Balance System," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 1, pp. 211–215, 1 2013. doi: 10.1109/TBME.2012.2222640

[206] C. L. Salter, M. A. J. Baalman, and D. Moody-Grigsby, "Between Mapping, Sonification and Composition: Responsive Audio Environments in Live Performance," in *Computer Music Modeling and Retrieval. Sense of Sounds.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, vol. 4969 LNCS, pp. 246–262. [Online]. Available: http://link.springer.com/10.1007/978-3-540-85035-9_17

[207] M. A. J. Baalman, D. Moody-Grigsby, and C. L. Salter, "Schwelle," in *Proceedings of the 7th international conference on New interfaces for musical expression - NIME '07.* New York, New York, USA: ACM Press, 2007. doi: 10.1145/1279740.1279774 p. 178.

[208] Y. Brazauskayte, "Undla," in *Proceedings of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction*, ser. TEI '20. New York, NY, USA: ACM, 2 2020. doi: 10.1145/3374920.3375287 pp. 655–660.

[209] V. Lorenzoni, P. Van den Berghe, P.-J. Maes, T. De Bie, D. De Clercq, and M. Leman, "Design and validation of an auditory biofeedback system for modification of running parameters," *Journal on Multimodal User Interfaces*, vol. 13, no. 3, pp. 167–180, 9 2019. doi: 10.1007/s12193-018-0283-1

[210] V. Lorenzoni, P.-J. Maes, P. Van den Berghe, D. De Clercq, T. de Bie, and M. Leman, "A biofeedback music-sonification system for gait retraining," in *Proceedings of the 5th International Conference on Movement and Computing.* New York, NY, USA: ACM, 6 2018. doi: 10.1145/3212721.3212843 pp. 1–5.

[211] G. Burloiu, "Adapting the SoundThimble movement sonification system for young motion-impaired users," in *Proceedings - 2018 IEEE 14th International Conference on Intelligent Computer Communication and Processing, ICCP 2018*, 2018. doi: 10.1109/ICCP.2018.8516435 pp. 153–157.

[212] G. Burloiu, V. Mihai, and Damian, "Layered motion and gesture sonification in an interactive installation," *AES: Journal of the Audio Engineering Society*, vol. 66, no. 10, pp. 770–778, 2018. doi: 10.17743/jaes.2018.0047

[213] A. Giomi and F. Fratagnoli, "Listening Touch," in *Proceedings of the 5th International Conference on Movement and Computing.* New York, NY, USA: ACM, 6 2018. doi: 10.1145/3212721.3212815 pp. 1–8.

[214] N. Schaffert, A. Engel, S. Schlüter, and K. Mattes, "The sound of the underwater dolphin-kick: developing real-time audio feedback in swimming," *Displays*, vol. 59, pp. 53–62, 9 2019. doi: 10.1016/j.displa.2019.08.001

[215] S. Chen, J. Bowers, and A. Durrant, "'Ambient walk'," in *Proceedings of the 2015 British HCI Conference.* New York, NY, USA: ACM, 7 2015. doi: 10.1145/2783446.2783630 p. 315.

[216] C. M. Wood and K. Kipp, "Use of audio biofeedback to reduce tibial impact accelerations during running," *Journal of Biomechanics*, vol. 47, no. 7, pp. 1739–1741, 5 2014. doi: 10.1016/j.jbiomech.2014.03.008

[217] R. Bresin, S. Papetti, M. Civolani, and F. Fontana, "Expressive Sonification of Footstep Sounds," *Interactive Sonification Workshop*, no. May 2014, pp. 51–54, 2010.

[218] M. Matsubara, T. Oba, H. Kadone, H. Terasawa, K. Suzuki, and M. Iguchi, "Wearable Auditory Biofeedback Device for Blind and Sighted Individuals," *IEEE MultiMedia*, vol. 22, no. 1, pp. 68–73, 1 2015. doi: 10.1109/MMUL.2015.20

[219] J. Pietschmann, F. Geu Flores, and T. Jöllenbeck, "Gait Training in Orthopedic Rehabilitation after Joint Replacement - Back to Normal Gait with Sonification?" *International Journal of Computer Science in Sport*, vol. 18, no. 2, pp. 34–48, 9 2019. doi: 10.2478/ijcss-2019-0012

[220] J. Anlauff, J. R. Cooperstock, and J. Fung, "Augmented feedback for learning single-legged stance on a slackline," in *2013 International Conference on Virtual Rehabilitation (ICVR)*. IEEE, 8 2013. doi: 10.1109/ICVR.2013.6662104 pp. 162–163.

[221] J. Françoise and F. Bevilacqua, "Motion-Sound Mapping through Interaction," *ACM Transactions on Interactive Intelligent Systems*, vol. 8, no. 2, pp. 1–30, 7 2018. doi: 10.1145/3211826

[222] J. Françoise, S. Fdili Alaoui, T. Schiphorst, and F. Bevilacqua, "Vocalizing dance movement for interactive sonification of laban effort factors," in *Proceedings of the 2014 conference on Designing interactive systems*. New York, NY, USA: ACM, 6 2014. doi: 10.1145/2598510.2598582 pp. 1079–1082.

[223] J. Danna and J.-L. Velay, "On the Auditory-Proprioception Substitution Hypothesis: Movement Sonification in Two Deafferented Subjects Learning to Write New Characters," *Frontiers in Neuroscience*, vol. 11, no. MAR, 3 2017. doi: 10.3389/fnins.2017.00137

[224] J. Danna, M. Fontaine, V. Paz-Villagrán, C. Gondre, E. Thoret, M. Aramaki, R. Kronland-Martinet, S. Ystad, and J.-L. Velay, "The effect of real-time auditory feedback on learning new characters," *Human Movement Science*, vol. 43, pp. 216–228, 10 2015. doi: 10.1016/j.humov.2014.12.002

[225] J. Danna, V. Paz-Villagrán, A. Capel, C. Pétroz, C. Gondre, S. Pinto, E. Thoret, M. Aramaki, S. Ystad, R. Kronland-Martinet, and J.-L. Velay, "Movement

Sonification for the Diagnosis and the Rehabilitation of Graphomotor Disorders," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2014, vol. 8905, pp. 246–255. ISBN 9783319129754. [Online]. Available: http://link.springer.com/10.1007/978-3-319-12976-1_16

[226] P. Palacio and D. Bisig, "Piano & Dancer - Interaction between a dancer and an acoustic instrument," in *Proceedings of MOCO'17*, vol. Part F1291, 2017. doi: 10.1145/3077981.3078052

[227] D. Bisig, P. Palacio, and M. Romero, "Piano&Dancer - Interaction Between a Dancer and an AcousticInstrument," in *19th Generative Art Conference GA2016*, 2016.

[228] A. Camurri, C. Canepa, N. Ferrari, M. Mancini, R. Niewiadomski, S. Piana, G. Volpe, J.-M. Matos, P. Palacio, and M. Romero, "A system to support the learning of movement qualities in dance," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. New York, NY, USA: ACM, 9 2016. doi: 10.1145/2968219.2968261 pp. 973–976.

[229] T. Smith, S. J. Bowen, B. Nissen, J. Hook, A. Verhoeven, J. Bowers, P. Wright, and P. Olivier, "Exploring Gesture Sonification to Support Reflective Craft Practice," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, vol. 2015-April. New York, NY, USA: ACM, 4 2015. doi: 10.1145/2702123.2702497 pp. 67–76.

[230] S. Barrass, N. Schaffert, and T. Barrass, "Probing Preferences between Six Designs of Interactive Sonifications for Recreational Sports, Health and Fitness," in *Proceedings of the ISon 2010, 3rd Interactive Sonification Workshop*. Interactive-Sonification, 2010, pp. 23–29.

[231] D. Bradshaw and K. Ng, "Motion capture, analysis and feedback to support learning conducting," in *Proceedings of the 2009 International Computer Music Conference, ICMC 2009*, 2009, pp. 307–310.

[232] R. Niewiadomski, M. Mancini, A. Cera, S. Piana, C. Canepa, and A. Camurri, "Does embodied training improve the recognition of mid-level expressive movement qualities sonification?" *Journal on Multimodal User Interfaces*, vol. 13, no. 3, pp. 191–203, 2019. doi: 10.1007/s12193-018-0284-0

[233] G. Varni, G. Dubus, S. Oksanen, G. Volpe, M. Fabiani, R. Bresin, J. Kleimola, V. Välimäki, and A. Camurri, "Interactive sonification of synchronisation of motoric behaviour in social active listening to music with mobile devices," *Journal on Multimodal User Interfaces*, vol. 5, no. 3-4, pp. 157–173, 5 2012. doi: 10.1007/s12193-011-0079-z

[234] M. Fabiani, R. Bresin, and G. Dubus, "Interactive sonification of expressive hand gestures on a handheld device," *Journal on Multimodal User Interfaces*, vol. 6, no. 1-2, pp. 49–57, 2012. doi: 10.1007/s12193-011-0076-2

[235] A. Tajadura-Jiménez, M. Basia, O. Deroy, M. Fairhurst, N. Marquardt, and N. Bianchi-Berthouze, "As Light as your Footsteps," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, vol. 2015-April. New York, NY, USA: ACM, 4 2015. doi: 10.1145/2702123.2702374 pp. 2943–2952.

[236] J. Yang and A. Hunt, "Real-time sonification of biceps curl exercise using muscular activity and kinematics," in *Proceedings of the 21st International Conference on Auditory Display*, 2015, pp. 289–293.

[237] K. Bakogiannis, A. Andreopoulou, and A. Georgaki, "The development of a dance-musification model with the use of machine learning techniques under COVID-19 restrictions," in *Audio Mostly 2021*, ser. AM '21. New York, NY, USA: ACM, 9 2021. doi: 10.1145/3478384.3478407 pp. 81–88.

[238] A. Bergsland, "DESIGNING INTERACTIVE SONIFICATIONS FOR THE EXPLORATION OF DANCE PHRASE EDGES," in *Proceedings of the Sound and Music Computing Conferences*, 2022, p. 178 – 185.

[239] J. E. Bruen, H. Kwon, and M. Jeon, "Cro-Create: Weaving Sound Using Crochet Gestures," in *Creativity and Cognition*, ser. C&C '23. New York, NY, USA: ACM, 6 2023. doi: 10.1145/3591196.3596816 pp. 334–337.

[240] Z. Gao, H. Wang, G. Feng, and H. Lv, "Exploring Sonification Mapping Strategies for Spatial Auditory Guidance in Immersive Virtual Environments," *ACM Trans. Appl. Percept.*, vol. 19, no. 3, 9 2022. doi: 10.1145/3528171

[241] K. Groß-Vogt, I. Svoronos-Kanavas, M. Weger, and C. Amon, "The Augmented Floor - Assessing Auditory Augmentation," in *AudioMostly 2022*, ser. AM '22. New York, NY, USA: ACM, 9 2022. doi: 10.1145/3561212.3561219 pp. 7–14.

[242] X. Hu, A. Song, Z. Wei, and H. Zeng, "StereoPilot: A Wearable Target Location System for Blind and Visually Impaired Using Spatial Audio Rendering," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 1621–1630, 9 2022. doi: 10.1109/TNSRE.2022.3182661

[243] M. Iber, B. Dumphart, V.-A. de Jesus Oliveira, S. Ferstl, J. M. Reis, D. Slijepčević, M. Heller, A.-M. Raberger, and B. Horsak, "Mind the Steps: Towards Auditory Feedback in Tele-Rehabilitation Based on Automated Gait Classification," in *Proceedings of the 16th International Audio Mostly Conference*, ser. AM '21. New York, NY, USA: Association for Computing Machinery, 2021. doi: 10.1145/3478384.3478398 pp. 139–146.

[244] Y. Ishii, H. Zheng, K. Tsubaki, C. Kato, and K. Aoki, "Demonstration of the Improvement of Golf Swing Practice Effectiveness Using Sonification Feedback," in *Lecture Notes in Networks and Systems*, 2022, vol. 364 LNNS, pp. 344–352. [Online]. Available: https://link.springer.com/10.1007/978-3-030-92604-5_31

[245] Y. Jiang and D. Tetteroo, "Serenity: Exploring audio-based gaming for arm-hand rehabilitation after stroke," in *ASSETS 2022 - Proceedings of the 24th*

International ACM SIGACCESS Conference on Computers and Accessibility, 2022. doi: 10.1145/3517428.3550388

[246] V. Kalampratsidou, S. Kemper, and E. B. Torres, "Real-Time Proxy-Control of Re-Parameterized Peripheral Signals using a Close-Loop Interface," *Journal of Visualized Experiments*, no. 171, 5 2021. doi: 10.3791/61943

[247] B. Kibushi and J. Okada, "Auditory sEMG biofeedback for reducing muscle co-contraction during pedaling," *Physiological Reports*, vol. 10, no. 10, 5 2022. doi: 10.14814/phy2.15288

[248] C. Landelle, J. Danna, B. Nazarian, M. Amberg, F. Giraud, L. Pruvost, R. Kronland-Martinet, S. Ystad, M. Aramaki, and A. Kavounoudias, "The impact of movement sonification on haptic perception changes with aging," *Scientific Reports*, vol. 11, no. 1, p. 5124, 3 2021. doi: 10.1038/s41598-021-84581-3

[249] J. Ley-Flores, L. Turmo Vidal, N. Berthouze, A. Singh, F. Bevilacqua, and A. Tajadura-Jiménez, "SoniBand: Understanding the Effects of Metaphorical Movement Sonifications on Body Perception and Physical Activity," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI '21. New York, NY, USA: Association for Computing Machinery, 2021. doi: 10.1145/3411764.3445558

[250] K. Masai, T. Kajiyama, T. Muramatsu, M. Sugimoto, and T. Kimura, "Virtual Reality Sonification Training System Can Improve a Novice's Forehand Return of Serve in Tennis," in *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 10 2022. doi: 10.1109/ISMAR-Adjunct57072.2022.00182. ISSN 2771-1110 pp. 845–849.

[251] Y. Miura, E. Wu, M. Kuribayashi, H. Koike, and S. Morishima, "Exploration of Sonification Feedback for People with Visual Impairment to Use Ski Simulator," in *Proceedings of the Augmented Humans International Conference 2023*, ser. AHs '23. New York, NY, USA: Association for Computing Machinery, 2023. doi: 10.1145/3582700.3582702 pp. 147–158.

[252] Y. Miura, M. Kuribayashi, E. Wu, H. Koike, and S. Morishima, "A Study on Sonification Method of Simulator-Based Ski Training for People with Visual Impairment," in *SIGGRAPH Asia 2022 Posters*, ser. SA '22. New York, NY, USA: Association for Computing Machinery, 2022. doi: 10.1145/3550082.3564172

[253] R. Sun, A. V. Wallop, G. Leslie, and E. Y.-L. Do, "SoniSpace: Expressive Movement Interaction to Encourage Taking Up Space with the Body," in *DIS 2023 Companion: Companion Publication of the 2023 ACM Designing Interactive Systems Conference*, 2023. doi: 10.1145/3563703.3596659 p. 279 – 283.

[254] P. Van den Berghe, R. Derie, P. Bauwens, J. Gerlo, V. Segers, M. Leman, and D. De Clercq, "Reducing the peak tibial acceleration of running by music-based biofeedback: A quasi-randomized controlled trial," *Scandinavian Journal of Medicine & Science in Sports*, vol. 32, no. 4, pp. 698–709, 4 2022. doi: 10.1111/sms.14123

[255] P. Van den Berghe, R. Derie, J. Gerlo, P. Bauwens, V. Segers, M. Leman, and D. De Clercq, "Influence of music-based feedback training on peak tibial acceleration during running outside of the biomechanics laboratory," *Footwear Science*, vol. 13, no. sup1, pp. S39–S41, 7 2021. doi: 10.1080/19424280.2021.1917671

[256] L. van Renswouw, J. Neerhof, S. Vos, P. van Wesemael, and C. Lallemand, "Sensation: Sonifying the Urban Running Experience," in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA '21. New York, NY, USA: Association for Computing Machinery, 2021. doi: 10.1145/3411763.3451788

[257] E. Wu, F. Perteneder, H. Koike, and T. Nozawa, "How to VizSki: Visualizing Captured Skier Motion in a VR Ski Training Simulator," in *Proceedings of the 17th International Conference on Virtual-Reality Continuum and its Applications in Industry*. New York, NY, USA: ACM, 11 2019. doi: 10.1145/3359997.3365698 pp. 1–9.

[258] A. Yadaei, M. R. Azadehfar, and B. Alizadehashrafi, "Creation of Melodic and Rhythmic Patterns Based on Cultural Heritage Sonification of Azerbaijani Dance Movements," *Journal of Sound and Music in Games*, vol. 3, no. 2-3, p. 1 – 27, 2022. doi: 10.1525/jsmg.202232-31

[259] N. Yazdi, F. Ayazi, and K. Najafi, "Micromachined Inertial Sensors," 1998.

[260] S. Han, Z. Meng, O. Omisore, T. Akinyemi, and Y. Yan, "Random Error Reduction Algorithms for MEMS Inertial Sensor Accuracy Improvement—A Review," *Micromachines 2020, Vol. 11, Page 1021*, vol. 11, no. 11, p. 1021, 11 2020. doi: 10.3390/MI11111021

[261] P. I. Cooper, "The absorption of radiation in solar stills," *Solar Energy*, vol. 12, no. 3, pp. 333–346, 1 1969. doi: 10.1016/0038-092X(69)90047-4

[262] E. M. Diaz, F. de Ponte Muller, A. R. Jimenez, and F. Zampella, "Evaluation of AHRS algorithms for inertial personal localization in industrial environments," in *2015 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 3 2015. doi: 10.1109/ICIT.2015.7125605 pp. 3412–3417.

[263] J. Gośliński, W. Giernacki, and A. Królikowski, "A Nonlinear Filter for Efficient Attitude Estimation of Unmanned Aerial Vehicle (UAV)," *Journal of Intelligent & Robotic Systems 2018 95:3*, vol. 95, no. 3, pp. 1079–1095, 11 2018. doi: 10.1007/S10846-018-0949-7

[264] P. Narkhede, S. Poddar, R. Walambe, G. Ghinea, and K. Kotecha, "Cascaded Complementary Filter Architecture for Sensor Fusion in Attitude Estimation," *Sensors*, vol. 21, no. 6, 2021. doi: 10.3390/s21061937

[265] Ammar Assad, W. Khalaf, and I. Chouaib, "Novel Adaptive Fuzzy Extended Kalman Filter for Attitude Estimation in Gps-Denied Environment," *Gyroscopy and Navigation*, vol. 10, no. 3, pp. 131–146, 7 2019. doi: 10.1134/S2075108719030027

[266] O. De Silva, G. K. Mann, and R. G. Gosine, "The Right Invariant Nonlinear Complementary Filter for Low Cost Attitude and Heading Estimation of Platforms," *Journal of Dynamic Systems, Measurement and Control, Transactions of the ASME*, vol. 140, no. 1, 1 2018. doi: 10.1115/1.4037331/371546

[267] A. Noureldin, T. B. Karamat, and J. Georgy, "Kalman Filter," *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*, pp. 225–245, 2013. doi: 10.1007/978-3-642-30466-8_7

[268] R. Mahony, T. Hamel, and J. M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Transactions on Automatic Control*, vol. 53, no. 5, pp. 1203–1218, 2008. doi: 10.1109/TAC.2008.923738

[269] S. O. H. Madgwick, A. J. L. Harrison, and R. Vaidyanathan, "Estimation of IMU and MARG orientation using a gradient descent algorithm," in *2011 IEEE International Conference on Rehabilitation Robotics.* IEEE, 6 2011. doi: 10.1109/ICORR.2011.5975346 pp. 1–7.

[270] S. O. H. Madgwick, "AHRS algorithms and calibration solutions to facilitate new applications using low-cost MEMS," Ph.D. dissertation, 2014.

[271] A. Nagasubramony, R. F. Player, C. E. Westling, K. Galvin, and H. J. Witchel, "Using wearable inertial sensors to detect different strategies for the sit-to-stand transition in multiple sclerosis," *ACM International Conference Proceeding Series*, 4 2021. doi: 10.1145/3452853.3452862

[272] J. Marín, T. Blanco, J. J. Marín, A. Moreno, E. Martitegui, and J. C. Aragüés, "Integrating a gait analysis test in hospital rehabilitation: A service design approach," *PLOS ONE*, vol. 14, no. 10, p. e0224409, 10 2019. doi: 10.1371/journal.pone.0224409

[273] M. Legrand, M. Merad, E. Montalivet, A. Roby-Bramiand, and N. Jarrassé, "Movement-based control for upper-limb prosthetics: Is the regression technique the key to a robust and accurate control?" *Frontiers in Neurorobotics*, vol. 12, pp. 1–13, 7 2018. doi: 10.3389/FNBOT.2018.00041/BIBTEX

[274] H. Zhou, T. Stone, H. Hu, and N. Harris, "Use of multiple wearable inertial sensors in upper limb motion tracking," *Medical engineering & physics*, vol. 30, no. 1, pp. 123–133, 1 2008. doi: 10.1016/J.MEDENGPHY.2006.11.010

[275] M. Alt Murphy, S. Murphy, H. C. Persson, U.-B. Bergström, and K. S. Sunnerhagen, "Kinematic Analysis Using 3D Motion Capture of Drinking Task in People With and Without Upper-extremity Impairments," *Journal of Visualized Experiments*, vol. 2018, no. 133, pp. 3–9, 3 2018. doi: 10.3791/57228

[276] "GitHub - xioTechnologies/Gait-Tracking-With-x-IMU." [Online]. Available: https://github.com/xioTechnologies/Gait-Tracking-With-x-IMU

[277] X. Yun, J. Calusdian, E. R. Bachmann, and R. B. McGhee, "Estimation of Human Foot Motion During Normal Walking Using Inertial and Magnetic Sensor Measurements," *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 7, pp. 2059–2072, 7 2012. doi: 10.1109/TIM.2011.2179830

[278] D. Comotti, M. Caldara, M. Galizzi, P. Locatelli, and V. Re, "Inertial based hand position tracking for future applications in rehabilitation environments," *Proceedings - 2015 6th IEEE International Workshop on Advances in Sensors and Interfaces, IWASI 2015*, pp. 222–227, 8 2015. doi: 10.1109/IWASI.2015.7184959

[279] L. Bai, M. G. Pepper, Y. Yan, S. K. Spurgeon, M. Sakel, and M. Phillips, "Quantitative Assessment of Upper Limb Motion in Neurorehabilitation Utilizing Inertial Sensors," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 2, pp. 232–243, 2015. doi: 10.1109/TNSRE.2014.2369740

[280] P. Merriaux, Y. Dupuis, R. Boutteau, P. Vasseur, and X. Savatier, "A Study of Vicon System Positioning Performance," *Sensors 2017, Vol. 17, Page 1591*, vol. 17, no. 7, p. 1591, 7 2017. doi: 10.3390/S17071591

[281] M. Windolf, N. Götzen, and M. Morlock, "Systematic accuracy and precision analysis of video motion capturing systems—exemplified on the Vicon-460 system," *Journal of Biomechanics*, vol. 41, no. 12, pp. 2776–2780, 8 2008. doi: 10.1016/J.JBIOMECH.2008.06.024

[282] "NGIMU – x-io Technologies." [Online]. Available: https://x-io.co.uk/ngimu/

[283] MIDI Association, "MIDI Association: Official MIDI Specifications." [Online]. Available: https://midi.org/specifications

[284] B. Debnath, M. O'Brien, M. Yamaguchi, and A. Behera, "A review of computer vision-based approaches for physical rehabilitation and assessment," *Multimedia Systems*, vol. 28, no. 1, pp. 209–239, 2 2022. doi: 10.1007/s00530-021-00815-4

[285] S. L. Colyer, M. Evans, D. P. Cosker, and A. I. T. Salo, "A Review of the Evolution of Vision-Based Motion Analysis and the Integration of Advanced Computer Vision Methods Towards Developing a Markerless System," *Sports Medicine - Open*, vol. 4, no. 1, p. 24, 12 2018. doi: 10.1186/s40798-018-0139-y

[286] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172–186, 1 2021. doi: 10.1109/TPAMI.2019.2929257

[287] J. Davis, D. Nehab, R. Ramamoorthi, and S. Rusinkiewicz, "Spacetime stereo: A unifying framework for depth from triangulation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 2, pp. 296–302, 2 2005. doi: 10.1109/TPAMI.2005.37

[288] R. Lange, "3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology," Ph.D. dissertation, University of Siegen, 2000.

[289] S. Foix, G. Alenyà, and C. Torras, "Lock-in time-of-flight (ToF) cameras: A survey," *IEEE Sensors Journal*, vol. 11, no. 9, pp. 1917–1926, 2011. doi: 10.1109/JSEN.2010.2101060

[290] "Stereolabs - Capture the World in 3D." [Online]. Available: https://www.stereolabs.com/

[291] "Stereo Depth Solutions from Intel RealSense." [Online]. Available: https://www.intelrealsense.com/stereo-depth/

[292] Y. Zhang, Z. Xiong, Z. Yang, and F. Wu, "Real-time scalable depth sensing with hybrid structured light illumination," *IEEE Transactions on Image Processing*, vol. 23, no. 1, pp. 97–109, 2014. doi: 10.1109/TIP.2013.2286901

[293] M. Kriman, A. Magen, A. K. N. Chakravartula, Z. Ge, G. A. Cohoon, E. H. O'Connor, C. K. Wong, C. F. Mischke, C. Hu, and Y. Z. Alnahhas, "LIGHT PROJECTOR WITH INTEGRATED INTEGRITY SENSOR," pp. 1–17, 5 2020.

[294] H. Sarbolandi, M. Plack, and A. Kolb, "Pulse Based Time-of-Flight Range Sensing." *Sensors (Basel, Switzerland)*, vol. 18, no. 6, 5 2018. doi: 10.3390/s18061679

[295] "3D ToF Camera, Best-in-One Time-of-Flight Camera — LIPSedge DL Camera." [Online]. Available: https://www.lips-hci.com/lipsedge-dl

[296] "Vzense 3D Time-of-Flight Camera." [Online]. Available: https://www.vzense.com/

[297] J. A. Albert, V. Owolabi, A. Gebel, C. M. Brahms, U. Granacher, and B. Arnrich, "Evaluation of the Pose Tracking Performance of the Azure Kinect and Kinect v2 for Gait Analysis in Comparison with a Gold Standard: A Pilot Study," *Sensors 2020, Vol. 20, Page 5104*, vol. 20, no. 18, p. 5104, 9 2020. doi: 10.3390/S20185104

[298] G. Amprimo, C. Ferraris, G. Masi, G. Pettiti, and L. Priano, "GMH-D: Combining Google MediaPipe and RGB-Depth Cameras for Hand Motor Skills Remote Assessment," *Proceedings - 2022 IEEE International Conference on Digital Health, ICDH 2022*, pp. 132–141, 2022. doi: 10.1109/ICDH55609.2022.00029

[299] M. Antico, N. Balletti, A. Ciccotelli, M. Ciccotelli, G. Laudato, A. Lazich, M. Notarantonio, R. Oliveto, S. Ricciardi, S. Scalabrino, and J. Simeone, "2Vita-

B Physical: An Intelligent Home Rehabilitation System Based on Microsoft Azure Kinect," *Frontiers in Human Dynamics*, vol. 0, p. 54, 9 2021. doi: 10.3389/FHUMD.2021.678529

[300] "Azure Kinect DK depth camera — Microsoft Learn," 9 2021. [Online]. Available: https://learn.microsoft.com/en-us/azure/kinect-dk/depth-camera

[301] "Introduction — PyQt Documentation v5.15.4." [Online]. Available: https://www.riverbankcomputing.com/static/Docs/PyQt5/introduction.html

[302] O. Bélanger, "Pyo 1.0.4 documentation." [Online]. Available: http://ajaxsoundstudio.com/pyodoc/#

[303] L. Bai, M. G. Pepper, Z. Wang, M. D. Mulvenna, R. R. Bond, D. Finlay, and H. Zheng, "Upper Limb Position Tracking with a Single Inertial Sensor Using Dead Reckoning Method with Drift Correction Techniques," *Sensors*, vol. 23, no. 1, p. 360, 12 2022. doi: 10.3390/s23010360

[304] P. Craig, P. Dieppe, S. Macintyre, S. Mitchie, I. Nazareth, and M. Petticrew, "Developing and evaluating complex interventions: the new Medical Research Council guidance," *BMJ*, vol. 337, no. 7676, pp. 979–983, 9 2008. doi: 10.1136/BMJ.A1655

[305] K. Skivington, L. Matthews, S. A. Simpson, P. Craig, J. Baird, J. M. Blazeby, K. A. Boyd, N. Craig, D. P. French, E. McIntosh, M. Petticrew, J. Rycroft-Malone, M. White, and L. Moore, "A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance," *BMJ*, vol. 374, p. n2061, 9 2021. doi: 10.1136/bmj.n2061

[306] M. Sekhon, M. Cartwright, and J. J. Francis, "Acceptability of healthcare interventions: An overview of reviews and development of a theoretical framework," *BMC Health Services Research*, vol. 17, no. 1, pp. 1–13, 1 2017. doi: 10.1186/S12913-017-2031-8/TABLES/3

[307] M. Synofzik, G. Vosgerau, and M. Voss, "The experience of agency: an interplay between prediction and postdiction," *Frontiers in Psychology*, vol. 4, 2013. doi: 10.3389/fpsyg.2013.00127

[308] D. M. Wegner, "The mind's best trick: how we experience conscious will," *Trends in Cognitive Sciences*, vol. 7, no. 2, pp. 65–69, 2 2003. doi: 10.1016/S1364-6613(03)00002-0

[309] P. I. Cornelio Martinez, "Examining the sense of agency in human-computer interaction," Ph.D. dissertation, University of Sussex, 2020.

[310] M. Synofzik, G. Vosgerau, and A. Newen, "Beyond the comparator model: A multifactorial two-step account of agency," *Consciousness and Cognition*, vol. 17, no. 1, pp. 219–239, 3 2008. doi: 10.1016/j.concog.2007.03.010

[311] M. Synofzik, P. Thier, D. T. Leube, P. Schlotterbeck, and A. Lindner, "Misattributions of agency in schizophrenia are based on imprecise predictions about the sensory consequences of one's actions," *Brain*, vol. 133, no. 1, pp. 262–271, 1 2010. doi: 10.1093/brain/awp291

[312] J. W. Moore, "What Is the Sense of Agency and Why Does it Matter?" *Frontiers in Psychology*, vol. 7, no. AUG, 8 2016. doi: 10.3389/FPSYG.2016.01272

[313] P. Haggard, S. Clark, and J. Kalogeras, "Voluntary action and conscious awareness," *Nature Neuroscience*, vol. 5, no. 4, pp. 382–385, 4 2002. doi: 10.1038/nn827

[314] C. Farrer, M. Bouchereau, M. Jeannerod, and N. Franck, "Effect of Distorted Visual Feedback on the Sense of Agency," *Behavioural Neurology*, vol. 19, no. 1-2, pp. 53–57, 2008. doi: 10.1155/2008/425267

[315] R. Nataraj, S. Sanford, A. Shah, and M. Liu, "Agency and Performance of Reach-to-Grasp With Modified Control of a Virtual Hand: Implications for Rehabilitation," *Frontiers in Human Neuroscience*, vol. 14, p. 126, 4 2020. doi: 10.3389/fnhum.2020.00126

[316] S. Seghezzi, E. Zirone, E. Paulesu, and L. Zapparoli, “The Brain in (Willed) Action: A Meta-Analytical Comparison of Imaging Studies on Motor Intentionality and Sense of Agency,” *Frontiers in Psychology*, vol. 10, 4 2019. doi: 10.3389/fpsyg.2019.00804

[317] E. von Holst, “Relations Between the Central Nervous System and the Peripheral Organs,” *The British Journal of Animal Behaviour*, vol. 2, pp. 89–94, 4 1954.

[318] Q. Welniarz, Y. Worbe, and C. Gallea, “The Forward Model: A Unifying Theory for the Role of the Cerebellum in Motor Control and Sense of Agency,” *Frontiers in Systems Neuroscience*, vol. 15, 4 2021. doi: 10.3389/fnsys.2021.644059

[319] Z. Gao, C. Davis, A. M. Thomas, M. N. Economo, A. M. Abrego, K. Svoboda, C. I. De Zeeuw, and N. Li, “A cortico-cerebellar loop for motor planning,” *Nature*, vol. 563, no. 7729, pp. 113–116, 11 2018. doi: 10.1038/s41586-018-0633-x

[320] M. Manto, J. M. Bower, A. B. Conforto, J. M. Delgado-García, S. N. F. da Guarda, M. Gerwig, C. Habas, N. Hagura, R. B. Ivry, P. Mariën, M. Molinari, E. Naito, D. A. Nowak, N. Oulad Ben Taib, D. Pelisson, C. D. Tesche, C. Tilikete, and D. Timmann, “Consensus Paper: Roles of the Cerebellum in Motor Control—The Diversity of Ideas on Cerebellar Involvement in Movement,” *The Cerebellum*, vol. 11, no. 2, pp. 457–487, 6 2012. doi: 10.1007/s12311-011-0331-9

[321] O. Baumann, R. J. Borra, J. M. Bower, K. E. Cullen, C. Habas, R. B. Ivry, M. Leggio, J. B. Mattingley, M. Molinari, E. A. Moulton, M. G. Paulin, M. A. Pavlova, J. D. Schmahmann, and A. A. Sokolov, “Consensus Paper: The Role of the Cerebellum in Perceptual Processes,” *The Cerebellum*, vol. 14, no. 2, pp. 197–220, 4 2015. doi: 10.1007/s12311-014-0627-7

[322] F. B. Nahab, P. Kundu, C. Gallea, J. Kakareka, R. Pursley, T. Pohida, N. Miletta, J. Friedman, and M. Hallett, “The Neural Processes Underlying Self-Agency,” *Cerebral Cortex*, vol. 21, no. 1, pp. 48–55, 1 2011. doi: 10.1093/cercor/bhq059

[323] G. Terzopoulos and M. Satratzemi, “Voice Assistants and Smart Speakers in Everyday Life and in Education,” *Informatics in Education*, vol. 19, no. 3, pp. 473–490, 2020. doi: 10.15388/infedu.2020.21

[324] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, “A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda,” *Computers and Education*, vol. 147, p. 103778, 4 2020. doi: 10.1016/J.COMPEDU.2019.103778

[325] B. Berberian, “Man-Machine teaming: a problem of Agency,” *IFAC-PapersOnLine*, vol. 51, no. 34, pp. 118–123, 2019. doi: 10.1016/j.ifacol.2019.01.049

[326] H. Limerick, D. Coyle, and J. W. Moore, “The experience of agency in human-computer interactions: a review,” *Frontiers in Human Neuroscience*, vol. 8, no. AUG, p. 643, 8 2014. doi: 10.3389/fnhum.2014.00643

[327] D. Coyle, J. Moore, O. Kristensson, P. C. Fletcher, and A. F. Blackwell, “I did that! Measuring Users’ Experience of Agency in their own Actions,” *CHI 2012*, 2012.

[328] A. Sato and A. Yasuda, “Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership,” *Cognition*, vol. 94, no. 3, pp. 241–255, 1 2005. doi: 10.1016/J.COGNITION.2004.04.003

[329] B. LIBET, C. A. GLEASON, E. W. WRIGHT, and D. K. PEARL, “TIME OF CONSCIOUS INTENTION TO ACT IN RELATION TO ONSET OF CEREBRAL ACTIVITY (READINESS-POTENTIAL),” *Brain*, vol. 106, no. 3, pp. 623–642, 9 1983. doi: 10.1093/brain/106.3.623

[330] D. A. NORMAN, “Cognitive Engineering,” in *User Centered System Design*. Boca Raton: CRC Press, 1 1986, pp. 31–62.

[331] R Core Team, “R: A Language and Environment for Statistical Computing.” Tech. Rep., 2023. [Online]. Available: https://www.R-project.org/

[332] Posit team, "RStudio: Integrated Development Environment for R," Posit Software, PBC, Boston, MA, Tech. Rep., 2023. [Online]. Available: https://www.posit.co/

[333] I. Carpinella, D. Cattaneo, and M. Ferrarin, "Quantitative assessment of upper limb motor function in Multiple Sclerosis using an instrumented Action Research Arm Test," *Journal of NeuroEngineering and Rehabilitation*, vol. 11, no. 1, p. 67, 4 2014. doi: 10.1186/1743-0003-11-67

[334] "scipy.signal.find_peaks — SciPy v1.11.1 Manual." [Online]. Available: https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.find_peaks.html

[335] "What is Max? — Cycling '74." [Online]. Available: https://cycling74.com/products/max

[336] "Home — Csound Community." [Online]. Available: https://csound.com/

[337] M. O. Ernst and H. H. Bülthoff, "Merging the senses into a robust percept," *Trends in Cognitive Sciences*, vol. 8, no. 4, pp. 162–169, 4 2004. doi: 10.1016/j.tics.2004.02.002

[338] M.-H. Tsai, K.-H. Chen, and I.-C. Lin, "Real-time upper body pose estimation from depth images," in *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 9 2015. doi: 10.1109/ICIP.2015.7351198 pp. 2234–2238.

[339] J. A. Dewey and G. Knoblich, "Do Implicit and Explicit Measures of the Sense of Agency Measure the Same Thing?" *PLoS ONE*, vol. 9, no. 10, p. e110118, 10 2014. doi: 10.1371/journal.pone.0110118

[340] A. Sato, "Action observation modulates auditory perception of the consequence of others' actions," *Consciousness and Cognition*, vol. 17, no. 4, pp. 1219–1227, 12 2008. doi: 10.1016/j.concog.2008.01.003

[341] C. Weiss, A. Herwig, and S. Schütz-Bosbach, "The self in action effects: Selective attenuation of self-generated sounds," *Cognition*, vol. 121, no. 2, pp. 207–218, 11 2011. doi: 10.1016/j.cognition.2011.06.011

[342] M. Fukuoka, F. Nakamura, A. Verhulst, M. Inami, M. Kitazaki, and M. Sugimoto, "Sensory Attenuation with a Virtual Robotic Arm Controlled Using Facial Movements," *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–17, 2023. doi: 10.1109/TVCG.2023.3246092

[343] "ISO 226:2023 - Acoustics — Normal equal-loudness-level contours." [Online]. Available: https://www.iso.org/standard/83117.html

[344] S. Swaminathan and E. G. Schellenberg, "Musical Competence is Predicted by Music Training, Cognitive Abilities, and Personality," *Scientific Reports 2018 8:1*, vol. 8, no. 1, pp. 1–7, 6 2018. doi: 10.1038/s41598-018-27571-2

[345] M. Wallentin, A. H. Nielsen, M. Friis-Olivarius, C. Vuust, and P. Vuust, "The Musical Ear Test, a new reliable test for measuring musical competence," *Learning and Individual Differences*, vol. 20, no. 3, pp. 188–196, 6 2010. doi: 10.1016/J.LINDIF.2010.02.004

[346] C. D. Güss, "What Is Going Through Your Mind? Thinking Aloud as a Method in Cross-Cultural Psychology," *Frontiers in Psychology*, vol. 9, 8 2018. doi: 10.3389/fpsyg.2018.01292

[347] F. Bianchi, S. Santurette, D. Wendt, and T. Dau, "Pitch Discrimination in Musicians and Non-Musicians: Effects of Harmonic Resolvability and Processing Effort," *Journal of the Association for Research in Otolaryngology*, vol. 17, no. 1, pp. 69–79, 2 2016. doi: 10.1007/s10162-015-0548-2

[348] K. Hankinson, A. Shaykevich, A.-M. Vallence, J. Rodger, M. Rosenberg, and C. Etherton-Beer, "A Tailored Music-Motor Therapy and Real-Time Biofeedback Mobile Phone App ('GotRhythm') to Promote Rehabilitation Following Stroke: A Pilot Study," *Neuroscience Insights*, vol. 17, 1 2022. doi: 10.1177/26331055221100587

[349] A. W. Salmoni, R. A. Schmidt, and C. B. Walter, "Knowledge of results and motor learning: A review and critical reappraisal," *Psychological Bulletin*, vol. 95, no. 3, pp. 355–386, 5 1984. doi: 10.1037/0033-2909.95.3.355

[350] R. Sigrist, G. Rauter, R. Riener, and P. Wolf, "Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review," *Psychonomic Bulletin & Review*, vol. 20, no. 1, pp. 21–53, 2 2013. doi: 10.3758/s13423-012-0333-8

Appendix A

Implementation

A.1 Soniccup

A.1.1 DPDT Switch

Make: C & K Momentary Push Button Switch, Through Hole, DPDT, 32V dc

A.1.2 Soniccup 3D Model

Table A.1: Dimensions of Soniccup - Component 1.

Component	**Dimension**	**Magnitude (mm)**
Handle	Inner radius	10.00
	Outer radius	15.00
	Height	100.00
Base	Bottom outer radius	40.00
	Bottom inner radius	35.00
	Top outer radius	15.00
	Top inner radius	10.00
	Height	25.00
Bowl	Top radius	50.00
	Bottom radius	10.00
	Height	25.00

Table A.2: Dimensions of Soniccup - Component 2.

Component	**Dimension**	**Magnitude (mm)**
Base	Radius	40.00
	Height	4.00
Cylinder	Radius	10.00
	Height	17.74
Cavity	Width	10.00
	Length	10.55
	Height	15.84
Switch Leg Insert	Radius	1.00
	Height	4.50
	Position Coordinates [Length,Width]	[3.00,2.50]
		[-3.00,2.50]
		[-3.00,-2.50]
		[3.00,-2.50]

A.2 Kalman Filter

A.2.1 Kalman Filter 1

The following matrices describe the composition of the first Kalman filter used as part of the Soniccup signal conditioning process that receives raw earth acceleration from the NGIMU and outputs an estimated velocity plot, which is then edited.

$$Q = \begin{bmatrix} 2.5e^{-9} & 5.0e^{-7} & 5.0e^{-5} & 0 & 0 & 0 & 0 & 0 & 0 \\ 5.0e^{-7} & 1.0e^{-4} & 1.0e^{-2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 5.0e^{-5} & 1.0e^{-2} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2.5e^{-9} & 5.0e^{-7} & 5.0e^{-5} & 0 & 0 & 0 \\ 0 & 0 & 0 & 5.0e^{-7} & 1.0e^{-4} & 1.0e^{-2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 5.0e^{-5} & 1.0e^{-2} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2.5e^{-9} & 5.0e^{-7} & 5.0e^{-5} \\ 0 & 0 & 0 & 0 & 0 & 0 & 5.0e^{-7} & 1.0e^{-4} & 1.0e^{-2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 5.0e^{-5} & 1.0e^{-2} & 1 \end{bmatrix} \tag{A.1}$$

$$H = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.2}$$

$$F = \begin{bmatrix} 1 & \delta t & 0.5 * (\delta t)^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \delta t & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \delta t & 0.5 * (\delta t)^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \delta t & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \delta t & 0.5 * (\delta t)^2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.3}$$

$$R = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \tag{A.4}$$

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.5}$$

$$X = \begin{bmatrix} x & \dot{x} & \ddot{x} & y & \dot{y} & \ddot{y} & z & \dot{z} & \ddot{z} \end{bmatrix} \tag{A.6}$$

where X = state mean, P = state covariance, Q = process covariance, F = state transition function, H = measurement function.

A.2.2 Kalman Filter 2

The following matrices describe the composition of the first Kalman filter used as part of the Soniccup signal conditioning process that receives raw earth acceleration from the NGIMU and outputs an estimated velocity plot, which is then edited.

$$Q = \begin{bmatrix} 2.5e^{-9} & 5.0e^{-7} & 5.0e^{-5} & 0 & 0 & 0 & 0 & 0 & 0 \\ 5.0e^{-7} & 1.0e^{-4} & 1.0e^{-2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 5.0e^{-5} & 1.0e^{-2} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2.5e^{-9} & 5.0e^{-7} & 5.0e^{-5} & 0 & 0 & 0 \\ 0 & 0 & 0 & 5.0e^{-7} & 1.0e^{-4} & 1.0e^{-2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 5.0e^{-5} & 1.0e^{-2} & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2.5e^{-9} & 5.0e^{-7} & 5.0e^{-5} \\ 0 & 0 & 0 & 0 & 0 & 0 & 5.0e^{-7} & 1.0e^{-4} & 1.0e^{-2} \\ 0 & 0 & 0 & 0 & 0 & 0 & 5.0e^{-5} & 1.0e^{-2} & 1 \end{bmatrix} \tag{A.7}$$

$$H = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{A.8}$$

$$F = \begin{bmatrix} 1 & \delta t & 0.5 * (\delta t)^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & \delta t & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \delta t & 0.5 * (\delta t)^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \delta t & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \delta t & 0.5 * (\delta t)^2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \delta t \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.9}$$

$$R = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \tag{A.10}$$

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{A.11}$$

$$X = \begin{bmatrix} x & \dot{x} & \ddot{x} & y & \dot{y} & \ddot{y} & z & \dot{z} & \ddot{z} \end{bmatrix} \tag{A.12}$$

where X = state mean, P = state covariance, Q = process covariance, F = state transition function, H = measurement function.

A.3 Azure Kinect Camera Settings

The following settings were used with the Azure Kinect camera throughout the work described in this thesis.

Table A.3: Azure Kinect Camera Settings.

Depth Mode	Near Field of View Unbinned
Colour Resolution	720 p
Frame Rate	30 Hz

The following variables were used with the KinectSon associated low pass filter.

Table A.4: Properties of Low Pass Filter.

Order	5
Cutoff Frequency	1 Hz
Width	2 samples
Window	Hamming
Sample Frequency	30 Hz

A.4 Pyo Sonification Settings

The following sonification configuration settings were used throughout this thesis.

Table A.5: Properties of generated audio note, f denotes the output from Equation 4.23.

Envelope	Attack	Time = 0.01 s
	Decay	Time = 0.01 s
	Sustain	Time = 0.1 s
	Release	Time = 0.01 s
	Duration	Time = 0.13 s
Waveform Generator	SawTable	Order = 8
Audio Output	Signal Frequency	f, $1.005 \cdot$ f
	PLAY	Current Note $\neq$ Previous Note
	STOP	Current Note = Previous Note

Appendix B

Ethics

OFFICE USE ONLY
UECREF
Date
Paper

Ethics Application Form

Please answer all questions

1. Title of the investigation
A feasibility study on new movement sonification systems for rehabilitating the upper limb hemiparesis in stroke survivors.
Please state the title on the PIS and Consent Form, if different: A feasibility study on new audio-feedback movement systems for rehabilitating the arm movements of stroke survivors.

2. Chief Investigator (must be at least a Grade 7 member of staff or equivalent)
Name: Madeleine Grealy ☒ Professor ☐ Reader ☐ Senior Lecturer ☐ Lecturer ☐ Senior Teaching Fellow ☐ Teaching Fellow Department: School of Psychological Science and Health Telephone: +44 (0)141 548 4885 E-mail: m.grealy@strath.ac.uk

3. Other Strathclyde investigator(s)
Name: Thomas Nown Status (e.g. lecturer, post-/undergraduate): EngD Student Department: Department of Biomedical Engineering Telephone: 07519745527 E-mail: thomas.nown@strath.ac.uk Name: Dr Christos Tachtatzis Status (e.g. lecturer, post-/undergraduate): Reader Department: Department of Electrical and Electronic Engineering Telephone: 01415482625 E-mail: christos.tachtatzis@strath.ac.uk Name: Dr Andrew Kerr Status (e.g. lecturer, post-/undergraduate): Senior Lecturer Department: Department of Biomedical Engineering

Telephone: 01415482855
E-mail: a.kerr@strath.ac.uk

Name: Prof Ivan Andonovic
Status (e.g. lecturer, post-/undergraduate): Professor
Department: Department of Electrical and Electronic Engineering
Telephone: 01415482537
E-mail: i.andonovic@strath.ac.uk

4. Non-Strathclyde collaborating investigator(s) (where applicable)

Name:
Status (e.g. lecturer, post-/undergraduate):
Department/Institution:
If student(s), name of supervisor:
Telephone:
E-mail:
Please provide details for all investigators involved in the study:

5. Overseas Supervisor(s) (where applicable)

Name(s):
Status:
Department/Institution:
Telephone:
Email:
I can confirm that the local supervisor has obtained a copy of the Code of Practice: Yes ☐ No ☐
Please provide details for all supervisors involved in the study:

6. Location of the investigation

At what place(s) will the investigation be conducted?

WC109, Wolfson Centre, University of Strathclyde, 106 Rottenrow East, Glasgow, G4 0NW

If this is not on University of Strathclyde premises, how have you satisfied yourself that adequate Health and Safety arrangements are in place to prevent injury or harm?

7. Duration of the investigation

Duration(years/months): 0 years, 5 months

Start date (expected): 01/02/2023 Completion date (expected): 01/07/2023

8. Sponsor
Please note that this is not the funder; refer to Section C and Annexes 1 and 3 of the Code of Practice for a definition and the key responsibilities of the sponsor.

Will the sponsor be the University of Strathclyde: Yes ☒ No ☐
If not, please specify who is the sponsor:

9. Funding body or proposed funding body (if applicable)

Name of funding body:
Status of proposal – if seeking funding (please click appropriate box):
☐ In preparation
☐ Submitted
☐ Accepted
Date of submission of proposal: / / Date of start of funding: / /

10. Ethical issues

Describe the main ethical issues and how you propose to address them:

1. Highlighting movement deficits. The device we are creating will convert a person's movement into a sound so that they can hear their movement in real-time. For stroke survivors with movement problems providing this type of feedback may emphasize the nature of their deficit and they may find this distressing. In order to minimise this, we will only ask participants to perform movements they are comfortable with and that they can perform with relative ease.

2. Raising expectations. As the goal of this project is to develop a device that can be used in rehabilitation it is possible that some stroke survivors may hope that participating in this feasibility study may provide some therapeutic benefit to them. In order to minimise this we will ensure that the participant is fully aware of why the trial is taking place, and what to expect, it will be clarified to the volunteer (prior to and, after giving consent) which stage of development this system is at, and what the purpose of the investigation is. This will provide transparency to the participant as to why the study is taking place.

3. Participants may feel obliged to contribute to the study. To avoid participants feeling like they are trapped in the study or are forced to do something they are not comfortable with, it will be reassured that the participation is voluntary, and they can withdraw from the study at any point.

4. Participants may experience fatigue during the session. If this occurs, we will encourage them to rest. If a participant feels unwell prior to a session we will advise them not to do it, or to cancel the laboratory visit. Similarly, if the participant feels unwell during the session, we will advise them to stop and we will assist them in returning home or contact the emergency services if required. This information is on the PIS.

5. Participants may be concerned with COVID-19 infection risk. To minimise the risk of COVID-19 transmission, in addition to the existing guidelines enforced by the university, the participation environment will be disinfected and wiped down before and after each participation session. Participants will be made aware of the steps taken to ensure their safety, and if they are uncomfortable with the steps taken, they are reminded that there is no obligation to contribute to the study. A further point that we have considered is the anonymity of the data collected. We will make it clear to participants that we will keep personal data such as their name and email address in a password protected file for the duration of the investigation so that we can contact them if the need arises, but this will be deleted at the end of the study. We will also inform them that they will be given a participant number, and this will be used on all study files that will allow us to trace their data in case they wish to withdraw from the study. This information is on the PIS. Establishing a secure method of data capture and data storage will stop participant's contact details being given to other companies/institutions without their consent.

11. Objectives of investigation (including the academic rationale and justification for the investigation) Please use plain English.

A stroke is a serious life-threatening medical condition that occurs when the blood supply to part of the brain is cut off. Approximately 70% of stroke survivors are left with movement problems and whilst approximately 60% of them regain walking ability, only approximately 5% regain full arm function. The arm and hand movement deficit that stroke survivors experience interfere with their ability to carry out essential, everyday tasks, reducing independence and quality of life.

Beyond the primary source of care, stroke survivors will require additional physiotherapy sessions to continue their rehabilitation. Currently movement feedback in physiotherapy sessions are limited to real-time verbal feedback from the physiotherapist or recorded video feedback. Both types of feedback have limitations in communicating to a stroke survivor about their movements. Video feedback provides a rich level of information from the angle captured of the movements, however for the user to obtain this feedback, their attention must be shifted away from the action being performed. Verbal feedback from the physiotherapist would be less rich in terms of information, but alternatively the stroke survivor's attention could remain on the action being performed.

To retain a rich level of information for feedback, and to not shift attention away from the user's actions, new feedback systems are being developed that makes use of sonification, to emit auditory feedback to the users. This system converts a person's movement pattern into a sound so that auditory feedback about their movement is generated in real-time. At this stage in academic research, several types of movement sonification systems have been developed in different institutions, however, there isn't a consensus on how best to map different movement parameters (e.g. position, velocity, smoothness) to different musical parameters (e.g. pitch, volume, instrument).

The new systems will use either low-cost sensors and a switch integrated into a 3D printed object shaped like a saucer glass – of which we've termed 'Sonicup', or a camera-based technology called Azure Kinect, to obtain movement data from the hand of the user and transmit data to a smart device, which converts and plays sound during its use. The objectives of this study are therefore:

- To verify that the newly developed motion tracking systems, Sonicup and Azure Kinect, produces similar positional/velocity/acceleration data to a gold-standard movement tracking system from Vicon.
- To investigate the result of different sonification (converting data into sound) algorithms on movements exerted by people with no motor deficit and mild upper limb motor deficit.
- To establish if it is feasible for stroke survivors to use the new system, and that it provides them with meaningful feedback about their movements.

12. Participants

Please detail the nature of the participants:

This investigation will have several stages of progress whilst the technological system is in development. Phase 1 will primarily focus on the development of the audio feedback system and recruit volunteers for the testing and evaluation of the present system. At the earlier stages of development, the research team will look to recruit with no history of a diagnosed neurological condition or motor deficit.

Phase 2 will focus on the feasibility of the created system when used by a stroke survivor with mild upper limb motor deficit. Therefore, for the testing of Phase 2, the research team will look to recruit volunteers who are stroke survivors in the chronic stage of recovery and have mild upper limb motor deficit.

Summarise the number and age (range) of each group of participants:
Number: See section 16. Age (range): 18-80.
Please detail any inclusion/exclusion criteria and any further screening procedures to be used:

Phase 1 (individuals with no history of neurological conditions):
Inclusion criteria:

- Understand English, and ability to follow instructions
- 'Normal' range of upper limb movement in both of their upper limbs
- Have accepted permission for access to campus, and Wolfson Building

Exclusion criteria:

- History of a neurological condition
- Existing musculoskeletal condition that affects both upper limbs
- Severe balance issues
- Visual impairment that is not corrected by glasses/lenses
- Auditory impairment, that prevents them from hearing the sounds generated by the system
- Impairment of haptic sensitivity
- Pregnant
- Participants unable to provide consent on their own
- Recent injurious fall without medical assessment

Phase 2 (stroke survivors)

Inclusion criteria:

- A diagnosis of stroke and upper limb deficit resulting from stroke
- Ability to follow and understand instructions in English

Exclusion criteria:

- Neurological condition other than stroke
- Not discharged from NHS rehabilitation services
- Electrocardiogram changes suggesting recent myocardial infarction
- Diagnosed cardiovascular disorder
- Extreme obesity (>159kg)
- Suspected or known dissecting aneurysm
- Acute infections
- Recent injurious fall without medical assessment
- Severe cognitive impairment (assessed using the Montreal Cognitive Assessment (MoCA)).
- Musculoskeletal condition that effects hemiparetic upper limb (e.g. arthritis, fracture)
- Severe balance issues
- Pregnant
- Participants unable to provide consent on their own
- Time period of less than three months post-stroke
- Visual impairment that is not corrected by glasses/lenses
- Auditory impairment, that prevents them from hearing the sounds created by the system
- Impairment of haptic sensitivity
- Severity of upper-limb deficit prevents them from moving the arm without assistance

13. Nature of the participants
Please note that investigations governed by the Code of Practice that involve any of the types of participants listed in B1(b) must be submitted to the University Ethics Committee (UEC) rather than DEC/SEC for approval.

Do any of the participants fall into a category listed in Section B1(b) (participant considerations) applicable in this investigation?: Yes ☒ No ☐
If yes, please detail which category (and submit this application to the UEC):
v. have a physical disability or a chronic physical condition relevant to the subject of the investigation and for whom participation in the investigation may pose a risk to their wellbeing

14. Method of recruitment

Describe the method of recruitment (see section B4 of the Code of Practice), providing information on any payments, expenses or other incentives.

Participants will be recruited through a variety of channels including, from a pool of volunteers available through the University of Strathclyde 'Sir Jules Thorn Centre for Co-Creation of Rehabilitation Technology', meetups with peer support groups and exercise classes associated with Chest, Heart and Stroke Scotland, emails, social media and advertisements. Social media posts will be made to various appropriately selected Facebook groups, including the Strathclyde Biomedical Engineering Society. Advertisements will be made to the local newspapers, appropriate community boards and the Chest, Heart and Stroke Scotland Website. Volunteers registered with the University of Strathclyde will also be sent information sheets.

Potential participants will be contacted by their preferred method of communication, such as email or by phone and provided with the details of the current study. For the purposes of sending through the PIS, should communication be received from a phone call, the receiver will ask for the email address of the volunteer, and on the occasion that an email address cannot be given, the receiver will ask for the address of the volunteer. If a volunteer meets the inclusion and exclusion criteria, they will be provided with a consent form and asked to sign it. Participants from Phase 1 will not receive financial payment in exchange for their participation in the study. Participants from Phase 2 are entitled to a reimbursement of up to £20 for a taxi journey (each way).

15. Participant consent

Please state the groups from whom consent/assent will be sought (please refer to the Guidance Document). The PIS and Consent Form(s) to be used should be attached to this application form.

Informed consent will be sought from all volunteers wishing to participate. Participants will be made aware that they are under no obligation to complete any of the exercises they may feel uncomfortable with. The research team foresees no special issues surrounding potential participant's ability to provide informed consent. Any potential participants unable to provide consent for themselves will not be allowed to participate in the study. There are no issues surrounding children, legal guardians or the use of deception. All participants will have the opportunity to withdraw their data, from the date of data collection, up to one year from that date, if they so wish.

16. Methodology
Investigations governed by the Code of Practice which involve any of the types of projects listed in B1(a) must be submitted to the University Ethics Committee rather than DEC/SEC for approval.

Are any of the categories mentioned in the Code of Practice Section B1(a) (project considerations) applicable in this investigation? ☐ Yes ☒ No
If 'yes' please detail:

Describe the research methodology and procedure, providing a timeline of activities where possible. Please use plain English.

The investigation will be an iterative process and may take any number of trials to complete, depending on the decisions that are made during the development progress of the technology.

Potential participants who express interest in taking part in the study will be provided with a PIS. Those who then contact the research team wishing to participate will then be invited to the laboratory and the investigator will ask the participant to confirm that they meet the inclusion/exclusion criteria that has been detailed in the PIS (with exception to the MoCA assessment if applicable). Upon verbal confirmation, the investigator will then provide information to the participant with regards to the project, and the developing system's status, as well as demonstrate the types of exercises that will be asked in the experiment and, the equipment used to record those movements.

Volunteers will be given the option of providing consent on the day of visiting the laboratory or taking up to one week to make their decision. If consent is not given, then the volunteer would not be allowed to participate, and they would be thanked for their time. If consent is given later than the day of visiting the laboratory, then a time and date will be arranged with them to participate.

Once the potential participant has provided consent and before the data collection stage begins, if the potential participant is looking to meet the criteria of Phase 2, a MoCA assessment will also be conducted to assess the cognitive ability of the participant. If successful, an Action Research Arm Test score will be obtained, to assess severity of upper limb deficit, and the experiment will progress to the next stage. If unsuccessful the participant will not be allowed to progress to the next stage of the experiment and will be thanked for their time.

The participant will be asked to perform tasks primarily with their preferred upper limb for Phase 1, or primarily with their impaired upper limb for Phase 2. Calibration of the systems will be done prior to the investigation. The session will consist of asking the participant to complete repetitive upper-limb reaching movements. Up to three motion tracking systems (Sonicup, Azure Kinect, and the Vicon motion capture system) may operate independently in parallel, to obtain the position of the user's movements.

When the participant moves their arm, the sound generated through one of the systems will be played through speakers that are positioned in front of the participant, at an appropriate pitch and volume, the participant will be made aware which system is generating the sound. The data recorded through each system will be obtained and stored in a portable storage device.

The final stage of the study will involve the investigator asking the participant with about their experience of this system and their perception of the feedback generated, a form will be filled out by either the participant or by the investigator and verified with the participant.

The duration of the study will not exceed 1 hour.

What specific techniques will be employed and what exactly is asked of the participants? Please identify any non-validated scale or measure and include any scale and measures charts as an Appendix to this application. Please include questionnaires, interview schedules or any other non-standardised method of data collection as appendices to this application.

The session will consist of asking the participant to complete a number of movements such as:

- Moving the object further and closer
- Moving the object to the right and left
- Raising and lowering the object between two surfaces

After completion of the repetitive exercises a feedback form will be completed based on the aesthetics of the sounds generated and the experience of the system.

Where an independent reviewer is not used, then the UEC, DEC or SEC reserves the right to scrutinise the methodology. Has this methodology been subject to independent scrutiny? Yes ☐ No ☒ If yes, please provide the name and contact details of the independent reviewer:

17. Previous experience of the investigator(s) with the procedures involved. Experience should demonstrate an ability to carry out the proposed research in accordance with the written methodology.
Prof Madeleine Grealy has over twenty years of experience working with stroke survivors and she is an HCPC registered Sport and Exercise Psychologist. Dr Andrew Kerr is an experienced physiotherapist that currently supervises multiple projects involving the use of technology to support and optimise neurorehabilitation. Madeleine and Andrew will supervise and advise Thomas Nown on the most appropriate ways to interact with participants throughout the study. Dr Christos Tachtatzis and Prof Ivan Andonovic have an extensive portfolio of research involving sensor technologies and will advise Thomas Nown in the development of the audio feedback movement system.

18. Data collection, storage and security
How and where are data handled? Please specify whether it will be fully anonymous (i.e. the identity unknown even to the researchers) or pseudo-anonymised (i.e. the raw data is anonymised and given a code name, with the key for code names being stored in a separate location from the raw data) - if neither please justify. Participants will be given a study code and this will be used to store their raw data. The key for code names will be kept in a locked cupboard separate from the data and will only be accessible to the investigators. All data recorded throughout the duration of this study will be kept strictly confidential and anonymity of all participants will be maintained. The pseudo-anonymised data will be securely stored on password protected electronic devices. After a period of 1 year from data collection, the pseudo-anonymised data will become fully anonymous.
Explain how and where it will be stored, who has access to it, how long it will be stored and whether it will be securely destroyed after use: Data obtained from the three movement capturing systems will be securely stored on password protected electronic devices and stored in a locked cupboard in the Technology Innovation Centre. Data obtained from the feedback forms will be tabulated and stored onto the same electronic device. Only members of the research team will have access to the research data. The data will be stored as pseudo-anonymised for a period of 1 year from the data being collected, after which the data will become anonymised, and will be stored indefinitely.
Will anyone other than the named investigators have access to the data? Yes ☐ No ☒ If 'yes' please explain:

19. Potential risks or hazards
Briefly describe the potential Occupational Health and Safety (OHS) hazards and risks associated with the investigation:

For every participant, an investigator with access to the laboratory will meet each participant at the entrance to the building.

The testing laboratory in WC109 contains a variety of equipment and furniture, there is therefore risk that participants may hurt themselves by colliding with this equipment. All investigators will ensure that the laboratory is free from clutter and safe for participants to freely move around.

As one of the primary objectives is to provide informative auditory feedback to the participant, there is a risk that the sound emitted from the speakers will be overwhelming in terms of amplitude and pitch, therefore prior to the investigation, the maximum volume and pitch will be set through the speakers to a comfortable audio level, even at the most extreme of scenarios.

Due to the existing risk of COVID-19, mitigations are in place to reduce the potential spread of the virus by ensuring that all persons are well before entering the laboratory; participation environment is cleaned with disinfectant and wiped down before and after every participant session; masks are worn by investigators and participants upon reaching the environment; hands are sanitised by all before entering the laboratory.

Please attach a completed eRisk Assessment for the research. Further Guidance on Risk Assessment and Form can be obtained on Occupational Health, Safety and Wellbeing's webpages

20. What method will you use to communicate the outcomes and any additional relevant details of the study to the participants?

At the end of their participation, each participant will be asked if they would like their contact emails and addresses to remain on a mailing list through which they will receive updates on the outcomes of the study or studies of which they were a part of.

21. How will the outcomes of the study be disseminated (e.g. will you seek to publish the results and, if relevant, how will you protect the identities of your participants in said dissemination)?

Upon conclusion of Phase 2, the final system will be the primary topic for any publishing. Dissemination of the data obtained through the motion capture system and feedback form, will only occur if the need arises for validating design decisions. If such reasons for dissemination do occur, the data that will be provided will be anonymised.

Checklist	**Enclosed**	**N/A**
Participant Information Sheet(s)	☒	☐
Consent Form(s)	☒	☐
Sample questionnaire(s)	☒	☐
Sample interview format(s)	☐	☒
Sample advertisement(s)	☒	☐
OHS Risk Assessment (S20)	☒	
Any other documents (please specify below)	☐	☐

	☐ ☐ ☐ ☐	☐ ☐ ☐ ☐

22. Chief Investigator and Head of Department Declaration
Please note that unsigned applications will not be accepted and both signatures are required

I have read the University's Code of Practice on Investigations involving Human Beings and have completed this application accordingly. By signing below, I acknowledge that I am aware of and accept my responsibilities as Chief Investigator under Clauses 3.11 – 3.13 of the Research Governance Framework and that this investigation cannot proceed before all approvals required have been obtained.

Signature of Chief Investigator

Please also type name here: Madeleine Grealy

I confirm I have read this application, I am happy that the study is consistent with departmental strategy, that the staff and/or students involved have the appropriate expertise to undertake the study and that adequate arrangements are in place to supervise any students that might be acting as investigators, that the study has access to the resources needed to conduct the proposed research successfully, and that there are no other departmental-specific issues relating to the study of which I am aware.

Signature of Head of Department

Please also type name here Dr Allan Hewitt

Date: 19 / 11 / 2019

23. Only for University sponsored projects under the remit of the DEC/SEC, with no external funding and no NHS involvement

Head of Department statement on Sponsorship
This application requires the University to sponsor the investigation. This is done by the Head of Department for all DEC applications with exception of those that are externally funded and those which are connected to the NHS (those exceptions should be submitted to R&KES). I am aware of the implications of University sponsorship of the investigation and have assessed this investigation with respect to sponsorship and management risk. As this particular investigation is within the remit of the DEC and has no external funding and no NHS involvement, I agree on behalf of the University that the University is the appropriate sponsor of the investigation and there are no management risks posed by the investigation.

If not applicable, tick here ☐

Signature of Head of Department

Please also type name here	Dr Allan Hewitt
Date:	19 / 11 / 2019

For applications to the University Ethics Committee, the completed form should be sent to ethics@strath.ac.uk with the relevant electronic signatures.

24. Insurance
The questionnaire below must be completed and included in your submission to the UEC/DEC/SEC:

Is the proposed research an investigation or series of investigations conducted on any person for a Medicinal Purpose? Medicinal Purpose means: ▪ treating or preventing disease or diagnosing disease or ▪ ascertaining the existence degree of or extent of a physiological condition or ▪ assisting with or altering in any way the process of conception or ▪ investigating or participating in methods of contraception or ▪ inducing anaesthesia or ▪ otherwise preventing or interfering with the normal operation of a physiological function or ▪ altering the administration of prescribed medication.	No

If "**Yes**" please go to **Section A (Clinical Trials)** – all questions must be completed
If "**No**" please go to **Section B (Public Liability)** – all questions must be completed

Section A (Clinical Trials)

Does the proposed research involve subjects who are either: i. under the age of 5 years at the time of the trial; ii. known to be pregnant at the time of the trial	No

*If "**Yes**" the UEC should refer to Finance*

Is the proposed research limited to: iii. Questionnaires, interviews, psychological activity including CBT; iv. Venepuncture (withdrawal of blood); v. Muscle biopsy; vi. Measurements or monitoring of physiological processes including scanning; vii. Collections of body secretions by non-invasive methods; viii. Intake of foods or nutrients or variation of diet (excluding administration of drugs).	Yes

*If "**No**" the UEC should refer to Finance*

Will the proposed research take place within the UK?	Yes

*If "**No**" the UEC should refer to Finance*

Title of Research	
Chief Investigator	Professor Madeleine Grealy
Sponsoring Organisation	University of Strathclyde

Does the proposed research involve:	
a) investigating or participating in methods of contraception?	No
b) assisting with or altering the process of conception?	No
c) the use of drugs?	No
d) the use of surgery (other than biopsy)?	No
e) genetic engineering?	No
f) participants under 5 years of age(other than activities i-vi above)?	No
g) participants known to be pregnant (other than activities i-vi above)?	No
h) pharmaceutical product/appliance designed or manufactured by the institution?	No
i) work outside the United Kingdom?	No

If **"YES"** to **any** of the questions a-i please also complete the **Employee Activity Form** (attached).
If **"YES"** to **any** of the questions a-i, and this is a follow-on phase, please provide details of SUSARs on a separate sheet.
*If "**Yes**" to any of the questions a-i then the UEC/DEC/SEC should refer to Finance (insurance-services@strath.ac.uk).*

Section B (Public Liability)	
Does the proposed research involve :	
a) aircraft or any aerial device	No
b) hovercraft or any water borne craft	No
c) ionising radiation	No
d) asbestos	No
e) participants under 5 years of age	No
f) participants known to be pregnant	No
g) pharmaceutical product/appliance designed or manufactured by the institution?	No
h) work outside the United Kingdom?	No

*If **"YES"** to any of the questions the UEC/DEC/SEC should refer to Finance (insurance-services@strath.ac.uk).*

For NHS applications only - Employee Activity Form

Has NHS Indemnity been provided?	Yes / No
Are Medical Practitioners involved in the project?	Yes / No
If YES, will Medical Practitioners be covered by the MDU or other body?	Yes / No

This section aims to identify the staff involved, their employment contract and the extent of their involvement in the research (in some cases it may be more appropriate to refer to a group of persons rather than individuals).

Chief Investigator		
Name	**Employer**	**NHS Honorary Contract?**
		Yes / No
Others		
Name	**Employer**	**NHS Honorary Contract?**
		Yes / No
		Yes / No
		Yes / No
		Yes / No

Please provide any further relevant information here: