Segmentation and quantification of oropharynx and larynx tumours from MRI data

---

Trushali Doshi

A thesis submitted for the degree of

Doctor of Philosophy

Centre for Signal and Image Processing

Department of Electronic and Electrical Engineering

University of Strathclyde

Glasgow

2016

Declaration

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.

Trushali Doshi 18 May 2016

Acknowledgements

This PhD thesis would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this PhD.

First and foremost, I offer my sincerest gratitude to my first supervisor, Professor John Soraghan, for not only providing such an interesting topic to tackle but for the continuous stream of encouragement and advice offered throughout the project. His enthusiasm, knowledge and commitment to highest standards inspired and motivated me throughout this PhD.

My thanks also go to my second supervisor Dr Lykourgos Petropoulakis and Dr Gaetano Di Caterina for their technical help and improving my written work.

I sincerely thank by sponsors Dr Derek Grose from the Beatson West of Scotland Cancer Centre, Glasgow and Professor Kenneth MacKenzie from the Royal Infirmary Glasgow for always being helpful and for their ideas and contributions in this research.

I would like to thank Dr Christina Wilson for the collection of MRI data, for her help in interpretation of the data and her substantial input for the paper submitted to the Clinical Oncology Journal. In addition, I would also like to express my gratitude to Dr Claire Paterson, Dr Carolynn Lamb and Dr Allan James for their time and manual segmentation of MRI data.

I am also grateful to Sean Campbell for his admirable work in his final year project that made Chapter 8 of this thesis possible.

I would like to thank my family, for their love, support and encouragement, despite the distance. Most importantly, I would like to thanks my husband, Prasad Doshi, for his understanding, patience, and encouragement.

Lastly, I would like to acknowledge the Beatson Charity Fund and the University of Strathclyde for their financial support with this study.

Abstract

Radiation therapy (RT) is often offered as the primary treatment for the head and neck cancer. Quantitative analysis and volumetric measurements in RT require segmentation of a tumour (gross tumour volume (GTV)) and other anatomical structures (organs at risk). The current tumour segmentation technique, manual segmentation, using medical imaging is subject to high observer variability. Thus, this thesis describes new image processing methods for oropharynx and larynx tumours (segmentation and quantification) analysis from magnetic resonance imaging (MRI). An integrated approach has been developed, including data size and MRI artefacts reduction, throat region detection, extraction and segmentation of tumour (GTV) region with 3D reconstruction and quantification.

Initially, a novel 2D automatic pharyngeal and laryngeal cancer segmentation framework (PLCSF) is proposed for oropharynx and larynx tumours segmentation from contrast enhanced T1-weighted axial MRI slices. In PLCSF, prior to segmentation, local entropy minimisation technique is employed to reduce intensity inhomogeneity and new fuzzy rules based method is used for the throat region detection. Moreover a novel modified fuzzy c-means (FCM) clustering algorithm is presented that is shown to robustly extract a tumour region compared to standard FCM clustering. Then a tumour segmentation boundary (outline) is obtained using region-based level set method after noise removal using non-linear filtering. The advantage of the proposed PLCSF lies in its ability to obtain a comparable tumour outlines even in presence of artefacts, heterogeneous tumour profile and fuzzy boundaries.

Further an approach for three dimensional (3D) reconstruction and quantification of a tumour is presented. In this approach tumour outlines obtained from contiguous 2D slices are reconstructed in 3D using interpolation and rectangular mesh generation technique and volume is calculated using slice profile. Experimental results of PLCSF with volumetric measurements for oropharynx and larynx tumours on synthetic and real MRI data demonstrate that this tool may help reduce observer variations and can assist clinical oncologists with time-consuming, complex

radiotherapy treatment planning. Finally, a novel automatic 3D throat region segmentation algorithm is presented. This algorithm efficiently combines Fourier interpolation and 3D level set segmentation technique to improve the accuracy of the segmentation results.

Contents

List of figures

List of tables

List of acronyms

2D	Two Dimensional
3D	Three Dimensional
AAM	Active Appearance Model
ACM	Active Contour Model
ANN	Artificial Neural Network
ASM	Active Shape Model
BBPHE	Brightness Background Preserving Histogram Equalisation
BoT	Base of Tongue
CO	Clinical Oncologist
CoV	Coefficient of Variation
CT	Computed Tomography
DFT	Discrete Fourier Transform
DICOM	Digital Imaging and Communications in Medicine
DSC	Dice Similarity Coefficient
CTV	Clinical Target Volume
FCM	Fuzzy c-means
FI	Fourier Interpolation
FID	Free Induction Decay
FFT	Fast Fourier Transform
GTV	Gross Tumour Volume
GUI	Graphical User Interface
HE	Histogram Equalisation
HNC	Head and Neck Cancers
IIH	Intensity Inhomogeneity
IMRT	Intensity-Modulated Radiation Therapy
LSM	Level Set Method
MFCM	Modified Fuzzy c-means
MHD	Modified Hausdorff Distance
MRF	Markov Random Field

MR/MRI	Magnetic Resonance/Magnetic Resonance Imaging
MS	Mean Shift
Ncut	Normalised Cut
OAR	Organs at risk
PCC	Pearson Correlation Coefficient
PET	Positron Emission Tomography
PLCSF	Pharyngeal and Laryngeal Cancer Segmentation Framework
PTV	Planning Target Volume
PVE	Partial Volume Effect
ROI	Region of Interest
RT	Radiotherapy/Radiation Therapy/Radiation Oncology
RTP	Radiotherapy Treatment Planning
SUSAN	Smallest Univalue Segment Assimilating Nucleus
SVM	Support Vector Machine
T1+Gd	Gadolinium Enhanced T1-Weighted
TE	Echo Time
ThDA	Throat Detection Algorithm
TNM	Tumour-Neck-Metastases
TR	Repetition Time

List of symbols

Section 2.5

B_o	External magnetic field
B_1	Radiofrequency field
${}^1H$	Hydrogen nuclei
M_o	Net magnetisation vector
M_{xy}	Transverse magnetisation vector
M_z	Longitudinal magnetisation vector
ω_o	Characteristic/resonance frequency
$\wp$	Gyromagnetic ratio
α	Flip angle

Section 2.6

I	Image
o	Original signal
S_{bg}	Proportion of background (air) in a voxel/pixel
S_{fg}	Proportion of foreground (tissue) in a voxel/pixel
w_{bg}, w_{fg}	Weighting parameter of foreground and background proportion in a pixel
x, y, z	Spatial coordinates
β	Bias field/Intensity inhomogeneity
η	Noise

Section 3.3

log	Logarithmic scale
τ_1, τ_2	Knots in a spline
χ_{Nor}	Normalisation constant

Section 3.4

L	Interpolation factor
N	Total number samples in a signal (length of signal)
O	Fourier transform of a signal o
O_0	Zero padded Fourier transform of a signal o
o_0	Inverse Fourier transform of O_0

Section 3.5

c	Number of clusters
$d(.,.)$	Euclidean distance
E_{snake}	Energy functional for active contour model (snakes)
E_{image}	External energy
E_{int}	Internal energy
J_T	Objective function of fuzzy c-means clustering
K	Binary mask
m	Fuzziness factor
N	Total number of pixels in an image/slice
N_x	Size of x-dimension or Number of columns
N_y	Size of y-dimension or Number of rows
R	Region
t	Time
Th	Threshold value
v	Cluster centre
V	Speed function in evolution equation of level set function
ϑ	Closed curve
ϑ'	First derivative of v
ϑ''	Second derivative of v
ϖ_R	Mean intensity of a region
ξ	Measure of dissimilarity
∂	Partial derivative
μ	Membership matrix
∇	Gradient operator
ϕ	Level set function

Section 3.8

D_1, D_2, D_3	Diameters of a ellipsoid
T_{vol}	Volume of a tumour

Section 4.2

ϖ_{Hl}	Mean intensity of a healthy tissue region
ϖ_T	Mean intensity of a tumourous tissue region
σ_{bg}	Standard deviation of intensity of a background region

Section 4.5

h	Hausdorff distance
K_{auto}	Automatic binary segmentation mask
K_{ref}	Reference binary segmentation mask
N_{auto}	Total number of boundary points in K_{auto}
N_{ref}	Total number of boundary points in K_{ref}
P_{auto}	Set of boundary points of K_{auto}
P_{ref}	Set of boundary points of K_{ref}
ρ	Statistical significance (Type I error)
γ	Strength of Pearson Correlation Coefficient

Section 5.3

A_a	Active area
A_{in}	Area of a local interior region
A_{out}	Area of a local exterior region
BB	Bounding box
BB_d	Height of a bounding box
K_{mask}	Mask of a ball
w	Weight (smoothing) parameter
δ	Dirac function
ϖ_{inx}	Mean intensity of a local interior region
ϖ_{outx}	Mean intensity of a local exterior region

Section 6.2

N	Number of knots
R	Region under consideration
R''''	Fourth derivative value of region under consideration
ε	Error value
$\sigma_{corrected}$	Standard deviation of intensity of a tumour region of IIH corrected slice
$\varpi_{corrected}$	Mean intensity of a tumour region of a IIH corrected slice
χ	Constant

Section 6.3

F_H	Fuzzy horizontal line function
F_{VL}	Fuzzy vertical line function
N_{xc}	Centre of 1: number of columns N_x

Th_u, Th_l	Threshold values for μ_{VL}
Th_{u1}, Th_{l1}	Threshold values for μ_H
w_1	Weighting parameter of μ_{VL}
μ_H	Membership values of F_H
μ_{VL}	Membership values of F_{VL}

Section 6.4

J_{ST}	Modified objective function of fuzzy c-means clustering
R_x	Mean x-coordinate of throat region
R_y	Mean y-coordinate of throat region
p, q	Control parameters of spatial fuzzy c-means clustering
λ	Lagrange multiplier
Λ	Spatial function

Section 7.3

$C - auto$	Tumour outlines obtained by the framework (PLCSF)
$C - gold$	Consensus manual tumour outlines
$C1 - A$	First set of tumour outlines by Clinical oncologist 1 (C1)
$C1 - B$	Second set of tumour outlines by Clinical oncologist 1 (C1)
$C2 - A$	First set of tumour outlines by Clinical oncologist 2 (C2)
$C2 - B$	Second set of tumour outlines by Clinical oncologist 2 (C2)

Section 7.4

C_{vol}	Volume concordance
$T_{volauto}$	Automatic tumour volume
T_{volref}	Reference (manual) tumour volume

Section 8.2

G	Fourier transform of multidimensional array g
G_0	Zero padded Fourier transform obtained from G
g_0	Inverse Fourier transform of G_0
$N_1 ... N_n$	Size of N-dimensional array
X	Vector of x, y, z
$\Re$	Set of real numbers
ϖ_R	Mean intensity of region
σ_R	Standard deviation of a region

Chapter 1
Introduction

1.1 Preface

According to the World Health Organisation, cancer is a leading cause of death worldwide, with approximately 14 million new cases and 8.2 million cancer related deaths in 2012 [1]. Advancement in diagnostic imaging technologies, such as computed tomography (CT), magnetic resonance imaging (MRI), have enabled the visualisation, characterisation and quantification of cancer regions at cellular levels. However, manually identifying and delineating (segmenting) cancer regions from this large medical imaging data which is a main step in treating the cancer, using treatment options such as Radiotherapy (RT) and Surgery, is time consuming and usually depend on clinical oncologists (CO) expertise and available facilities. Thus, these difficulties in manual interpretation of the medical images has led to the need for the development of automated cancer analysis tools [2-3]. With automated tools, the goal is to precisely identify and delineate cancer regions so that the treatment will remove the cancer region, while preserving the functions of the neighbouring healthy anatomical structures. However, no general automated (computer-aided) framework is available to characterise and quantify all cancer regions from the human anatomy because cancer areas vary considerably across patients in terms of size, shape, extension and localisation.

Applying a range of image processing techniques to assist the interpretation of magnetic resonance (MR) images in presence of head and neck cancers (HNC) is considered in this research. In particular, this research aims to develop the system for automatic segmentation and quantification for the oropharynx and larynx tumours from MR images. This system could be used to improve the characterisation and quantification of cancer regions and therefore the treatment of cancer with ultimate goal to improve patient care and quality of life. It can be noted that this system is not intended to replace an expert (CO), but would provide a reliable assistance to CO to reduce the work load and improve the RT treatment plan.

1.2 Research motivation

Over the past decade, HNC RT treatment has quickly shifted from two-dimensional (2D) radiotherapy to Intensity-Modulated Radiotherapy (IMRT), resulting in better organs at risk (OAR) sparing and organ preservation [4]. IMRT planning basically produces steep dose gradients, particularly at the border between target volumes [5] and OAR. Thus, accurate target volume delineation (segmentation) from medical imaging such as MRI is of the utmost importance to ensure the improved therapeutic ratio of this technique [4]. However, in a current clinical routine, as well as in clinical studies, a common approach to delineate the target volumes, predominantly gross tumour volume (GTV), involves CO manually drawing the boundary of the tumour on every MRI slice for a patient. This segmentation process is time-consuming, [6], [7] mostly in the complex head and neck region such as oropharynx and larynx. Further, delineation is also prone to large inter- and intra- observer variations [8], [9]. Thus, there is a need to replace the current delineation process with highly automatic, accurate and reproducible delineation algorithms.

Image processing techniques that can automatically segment and quantify head and neck tumours from MRI would be of enormous potential value for improved diagnosis, treatment planning, and follow-up of individual patients because to the best of our knowledge no known (commercially available) system exist to automatically delineate (segment) the GTV of HNC from MRI.

1.3 Research hypothesis and objectives

The hypothesis of this research is that, an automatic (computer-aided) MR-based segmentation and quantification tool, developed using image processing techniques, for the oropharynx and larynx tumours can obtain similar results when compared to the current manual approach defined by an expert CO but in time-efficient and reproducible manner with application to improve RT treatment plan.

To assess the above stated hypothesis, the objectives of this research were to:

1. Develop novel image processing algorithms to reduce MRI artefacts such as intensity inhomogeneity (IIH) [10], to extract geometrically variable tumour regions and to obtain appropriate outlines of the extracted tumour regions from 2D MRI slices.

2. Integrate newly-developed image processing algorithms into a novel framework for automatic segmentation and quantification of the oropharynx and larynx tumours from real MRI data.
3. Assess the clinical feasibility (accuracy, time-saving and reproducibility) of the automatic framework by comparing automatic tumour segmentation and quantification results to the manual approach defined by different expert CO and consensus manual outline agreed by different CO on real (clinical) MRI dataset.

1.4 Summary of original contributions

The main research contributions of the thesis are described below:

1.1. A novel pharyngeal and laryngeal cancer segmentation framework (PLCSF) for automatic segmentation and quantitative analysis for oropharyngeal and laryngeal tumours from 2D MRI slices is developed (Chapter 5). The PLCSF encompasses pre-processing techniques to reduce MRI artefacts such as noise and IIH, detection of anatomical structure (throat region) close to the expected tumour location, extraction of tumour region using modified fuzzy c-mean (MFCM) clustering and refinement of detected tumour boundary using non-linear filtering and localised region based LSM. Further, 3D reconstruction and volume calculation of a tumour region is presented to gain additional information of a tumour region. To the best of our knowledge this is the first automatic tool developed and validated against manual approach for the segmentation of oropharyngeal and laryngeal GTV using gadolinium-enhanced T1-weighted (T1+Gd) MRI slices.

1.2. A novel method to determine knot spacing/spline distance for IIH estimation using bicubic spline from a MRI slice is proposed (Chapter 6). The spacing between consecutive knots (spline distance) is determined using derivative and filtering method for each MRI slice as opposed to a single value of knot spacing for all MRI slices [10]. The results of the algorithm on real MRI dataset affected with IIH confirm that the determination of spline distance for each MRI slice reduces the coefficient of variation (CoV) of a region of interest (ROI) compared to a single value for all MRI slices.

1.3. A novel throat region detection algorithm (ThDA) which combines thresholding

with fuzzy rules based system is presented (Chapter 6). The positional information of the throat region is embedded in terms of two fuzzy rules to robustly detect the desired region. The detected region is used as supplementary information for the extraction of the tumour region from a MRI slice. Comparison of the proposed ThDA algorithm with other existing techniques (hessian analysis [11], fuzzy connectedness (FC) [12]) on real MRI dataset demonstrates superior results.

1.4. A novel modified fuzzy c-means (MFCM) clustering technique is proposed to separate different tissue types in different clusters and subsequently to extract expected tumour regions (Chapter 6) from MRI slices. In MFCM, standard fuzzy c-means (FCM) clustering [13] algorithm is modified to include the spatial prior knowledge of the expected tumour regions. Robustness of MFCM to extract a tumour region is demonstrated by comparing it with the standard FCM [13].

2. A validation that the automatic framework (PLCSF) can produce segmentations results comparable to manual contouring in the context of RT planning on real MRI data is provided (Chapter 7). For validation, automatic segmentation results are compared with manual segmentation from expert CO using different quantitative metrics on synthetic and real MRI datasets.

3. A novel algorithm (I+3DLSM) for three-dimensional (3D) throat region segmentation from a MRI volume is proposed (Chapter 8). A Fourier interpolation (FI) that scales to any dimensionality is described. The proposed algorithm is a generalised approach that performs FI followed by 3D LSM segmentation. The significance of an interpolation prior to 3D LSM segmentation is illustrated. Experimental results authenticate the accuracy of the algorithm in presence of pathology near the desired region and change of topology.

1.5 Author's publications

Journal papers

1. T. Doshi, C. Wilson, C. Paterson, et al. "Validation of MRI-based auto-contouring software tool for gross tumour delineation in head and neck cancer radiotherapy planning", *Clinical Oncology* – **under second review.**

2. T. Doshi, J. Soraghan, L. Petropoulakis, et al., "Automatic Pharyngeal and Laryngeal Cancer Segmentation Framework (PLCSF) on Contrast Enhanced MR Images", *Biomedical Signal Processing and Control* – **under review.**
3. T. Doshi, J. Soraghan, L. Petropoulakis et al., "3-Dimensional Throat region Segmentation from MRI data based on Fourier Interpolation and 3-Dimensional Level Set Methods" - **in preparation.**

IET Electronics Letters

1. T. Doshi, G. Di-Caterina, J. Soraghan, et al. "Combining interpolation and 3D level set method (I+3DLSM) for medical image segmentation", *IET Electronics Letters*, vol. 52, no. 8, pp. 592-294, April 2016.

Technical conference papers

1. S. Campbell, T. Doshi, J. Soraghan, et al., "3-Dimensional Throat region Segmentation from MRI data based on Fourier Interpolation and 3-Dimensional Level Set Methods", *in Proc. of IEEE EMBC*, Milan, Italy, pp. 2419-2422, 2015.
2. T. Doshi, J. Soraghan, D. Grose, et al. "Automatic detection of larynx cancer from contrast-enhanced magnetic resonance images", *in Proc. of SPIE Medical Imaging*, Orlando, Florida, pp. 94142N-1-9412N-9, Feb. 2015.
3. T. Doshi, J. Soraghan, L. Petropoulakis, D. Grose and K MacKenzie, "Modified Fuzzy C-means Clustering for Extraction of Tongue Base Tumour from MRI data", *in Proc. of IEEE EUSIPCO,* Lisbon, Portugal, pp. 2460-2464, Sept. 2014.
4. T. Doshi, J. Soraghan, L. Petropoulakis, D. Grose and K. MacKenzie "Semi-Automatic Segmentation of Tongue Tumor from Magnetic Resonance Imaging", *in Proc. of IEEE IWSSIP,* Bucharest, Romania, pp. 143-146, July 2013.

Medical conference abstracts

1. T. Doshi, D. Grose, K. MacKenzie, et al., "MRI based 3D reconstruction of pharyngeal cancer to aid clinical oncologists in radiotherapy treatment planning", *3rd ESTRO Forum*, Barcelona, Spain, April 2015.
2. T. Doshi, J. Soraghan, D. Grose, K. MacKenzie, and L. Petropoulakis, "Automatic Delineation of the base of Tongue Tumour from Gadolinium-enhanced T1-weighted MRI slices", *ECHNO*, Liverpool, UK, April 2014.

3. T. Doshi, J. Soraghan, D. Grose, K. MacKenzie, and L. Petropoulakis, "Semi-Automatic Delineation of Tongue Tumour from MRI Slices", *BAHNO*, London, UK, April, 2013.

1.6 Thesis organisation

The reminder of this thesis is organised as follows:

Chapter 2 provides fundamental information about the anatomy, physiology and common pathologic conditions (tumour) of the human head and neck region. It is then followed by the introduction of RT and its treatment planning procedure. Further, information on MRI such as principle behind image reconstruction and some MRI artefacts is presented.

Chapter 3 reviews image processing techniques in various stages of the automatic segmentation and quantification of a tumour (GTV) region from a human anatomy. Artefacts reduction is required at the initial stage of tumour segmentation and the related techniques are therefore reviewed. An overview of segmentation techniques for the throat region, useful information in tumours extraction, is presented. MRI tumour segmentation techniques are discussed, which inspired us to propose a novel framework of oropharynx and larynx tumours segmentation (Chapter 5 and Chapter 6). Finally, the chapter finishes with a brief review of 3D reconstruction and volume calculation that motivates our work on 3D visualisation and quantification of a tumour region.

Chapter 4 introduces the real (clinical) MRI data collection from the Beatson West of Scotland Cancer Centre, Gartnavel Hospital, Glasgow. As the evaluation of developed algorithms in controlled circumstances is useful, generation of the Synthetic MRI data is also presented in this chapter. The chapter concludes with a description of the evaluation parameters used to assess the performance of the proposed algorithms.

Chapter 5 introduces the PLCSF for oropharynx and larynx tumours segmentation from T1+Gd MRI slices. The three steps of the PLCSF which includes pre-processing, separation of expected tumourous tissues from other tissue types and obtaining outline of expected tumour region are introduced. 3D reconstruction and volume calculation algorithm for quantification of a tumour is presented.

Chapter 6 presents the novel methodology for the extraction of oropharynx and larynx tumours from T1+Gd MRI. Initially, the estimation of IIH using bicubic spline with adaptive knot spacing and reduction using entropy minimisation technique is presented. Further, fuzzy rules based throat region detection (ThDA) and extraction of cancer region using MFCM clustering is discussed. The efficiency and robustness of proposed algorithms is validated by comparing the algorithms with the existing techniques.

Chapter 7 presents an experimental evaluation of the proposed framework on synthetic and real MRI data. This includes qualitative and quantitative comparison of the PLCSF results with manual outlines and state-of-the art algorithms. Chapter 8 describes a novel approach (I+3DLSM) for the segmentation of 3D throat region from a MRI volume. It focuses on the description of N-dimensional Fourier interpolation, reconstruction of a MRI volume, and segmentation using 3D LSM. The chapter finishes with the illustration of experimental results on 12 patients. Chapter 9 presents discussion of the thesis reflecting on research objectives, key contributions and findings, generalisation and limitations of the PLCSF. Overall conclusions for this thesis are reported in Chapter 10, along with suggestions for possible future work.

Chapter 2

Head and neck cancers (HNC) and magnetic resonance imaging (MRI)

2.1 Introduction

The head and neck region is one of the most complicated anatomical regions and functionally important structures of the human body. It contains vital organs and major pathways of the nervous, respiratory, digestive, vascular, lymphatic, and endocrine systems. Cancer also referred as a tumour/neoplasm/pathology/lesion in head and neck part has a significant impact on respiration, speech, swallowing and quality of life [14]. The importance of studying cancer regions is to differentiate between normal and cancer cells with aim to eradicate cancer cells.

Different treatment options to eradicate cancer cells are employed with one of the major treatment being RT [15], which kill cancer cells via ionising radiation. RT (currently referred as Radiation Oncology) treatment was first introduced by Dr Emil Grubbe [16] for a cancer region in the expectation of favourable therapeutic effects. Radiotherapy treatment planning (RTP) required for RT treatment involves working a lot of steps, with medical imaging to its core, from a detection of cancer regions to the delivery of a radiation dose.

MRI is a non-invasive medical imaging technique that creates detailed images of organs with specific tissue characterisation and thus provides a good assessment of cancer tissues. The earliest study on the use of MR imaging to produce images of body was conducted by Dr Paul Lauterbug [17]. This was followed by mathematical reconstruction algorithms by Sir Peter Mansfield [18].

In the remainder of this chapter, anatomy and physiology of head and neck regions are introduced in Section 2.2. Section 2.3 describes HNC, medical imaging and treatment options available for HNC. RTP is presented in Section 2.4. Section 2.5 explains basic principle of MRI with its acquisition process. MR image artefacts and artefacts sources are described in Section 2.6. Section 2.7 describes MR imaging in oropharynx and larynx tumours. Section 2.8 concludes the chapter.

2.2 Anatomy and physiology of the head and neck region

Head and neck region for appropriate cancer (TNM (Tumour-Node-Metastasis)) staging [14] is anatomically divided into three major sites [19]; 1. Lip and Oral cavity, 2. Pharynx and 3. Larynx (Fig. 2.1). The pharynx is the region that connects the nasal and oral cavities with the larynx and oesophagus. The pharynx is anatomically adapted to various functions: swallowing, maintenance of the pharyngeal airway, and participation in speech and other respiratory performances. The pharynx region is divided into three regions: nasopharynx, oropharynx and hypopharynx. Each of these regions is further divided into specific sub-sites (Fig. 2.1). Oropharynx is a middle part of the pharynx with complex anatomy that includes the base of tongue (BoT), the tonsils, the soft palate, and the walls of the pharynx. The larynx is located directly above trachea and in front of the pharynx. The principal function of the larynx is to protect the lower respiratory tract. Larynx also acts as a passageway for air and respiratory tract; produces sound; protects trachea from foreign objects.

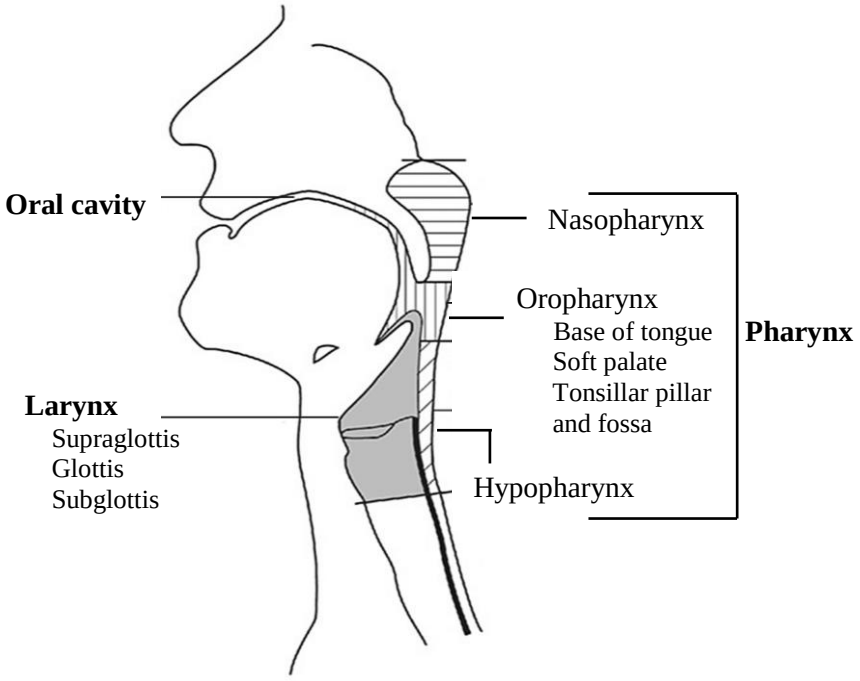


Fig. 2.1: Anatomic sites and sub-sites of the head and neck region. Head and neck region for cancer staging is anatomically divided into three major sites; 1. Oral cavity, 2. Pharynx and 3. Larynx. Cancers of pharynx and larynx regions are of interest in this thesis. Figure adapted from [20].

The throat region contains pharynx and larynx. The term 'throat region' in this thesis is reserved for the respiratory passageway that provides communication between oral cavity and oesophagus and term 'pharynx' and 'larynx' are reserved for a cancer region part.

2.3 Head and neck cancers (HNC)

Abnormal growth of cells in the human body is known as cancer. Cancer of head and neck are mostly (about 90%) squamous cell carcinomas arising from mucosa, or lining, of the upper aero digestive tract. Approximately 13000 new cases for HNC are reported each year in the United Kingdom with over 3,300 deaths per year [21]. The most common HNC in the United Kingdom is of the oral cavity followed by the larynx and then the oropharynx [21] (Fig. 2.2).

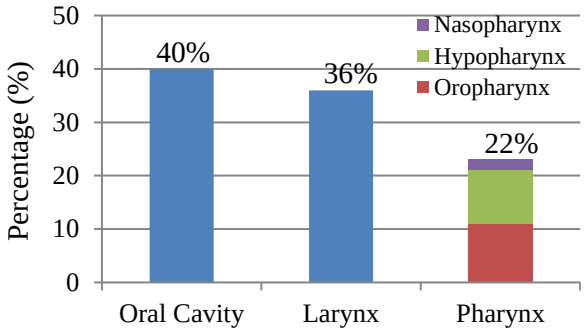


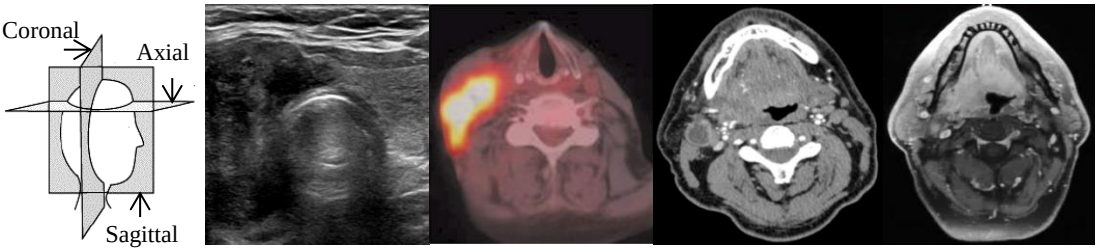
Fig. 2.2: Statistics of head and neck cancers in the United Kingdom (13000 new cases/year) (the approximate distribution is, oral cavity - 40%; larynx - 36%; and pharynx - 22%; (oropharynx - 10%; hypopharynx - 9%; and nasopharynx - 3%)) (Statistics taken from [21]).

The main aetiology for these types of cancer includes cigarette smoking and excessive alcohol consumption, genetic susceptibility, poor diet and recently human papillomavirus. Most HNC present with symptoms from the primary site, for example, hoarseness, difficulty in swallowing, or pain in the ear. In other cases, with primary asymptomatic, lymph node/s cause swelling which is noted by the patient. Initial diagnosis of HNC will normally include clinical examination by an experienced clinician using fibre optic endoscopy, which may be followed by fine needle aspiration/core biopsy of any neck masses. Patients with confirmed cancer further undergo radiological staging by CT or MRI [22]. HNC are staged by the Union for International Cancer Control: TNM classification of malignant tumours, which

describes the anatomical extent of a disease based on an assessment of the extent of the primary tumour, the absence or presence and extent of regional lymph nodes metastasis and the absence or presence of distant metastasis [22]. These individual TNM classifications are assembled into four stage groups (stages I–IV).

2.3.1 Medical imaging in diagnosis and staging of HNC

Medical imaging is a technique of creating detail images of internal organs at tissue (cellular) level in different planes (Fig. 2.3(a)). The main advantages of medical imaging are their ability to show in detail the normal anatomy and the exact extent of a low-lying tumour with deep sub-mucosal structures and spaces. Detection (delineation) of cancer regions and relevant structures (lymph nodes and OAR) from medical imaging is a fundamental task for TNM staging, for quantification of cancerous regions, to decide appropriate cancer treatment and post treatment follow-up [14]. Various imaging techniques are used to asses HNC including ultrasound, positron emission tomography (PET), PET-CT, CT and MRI [14] (Fig. 2.3(b)-(e)).



(a) (b) Ultrasound (c) PET-CT (d) CT (e) MRI

Fig. 2.3: (a) anatomical planes (sagittal, coronal and axial) used to obtain medical images; (b)-(e) different imaging modalities (Ultrasound, PET-CT, CT and MRI) in axial plane used in diagnosis, staging and treatment of head and neck cancers. ((b), (c), (d) and (e): Courtesy of The Beatson West of Scotland Cancer Centre, Glasgow)

Ultrasound (Fig. 2.3(b)) is usually combined with fine needle aspiration cytology for staging of (small) lymph nodes than assessment of primary tumours (which is the main aim of this thesis) as it provides detail information of the internal architecture and vascular pattern of lymph nodes for accurate staging [14]. PET is a functional imaging technique that depicts tissue metabolic activity and uses short-lived radioisotopes that contain protons that decay emitting positrons. PET is normally utilised to detect small secondary deposit of cancer after curative treatment, distant

metastasis or unknown primary tumour. The main limitation of PET is the tumour localisation due to poor anatomical information (resolution). PET scanning can be combined with CT (Fig. 2.3(c)) to provide more accurate anatomical localisation. Compared to PET-CT, MRI gives greatly enhanced soft-tissue contrast and resolution that is particularly valuable in delineating the primary tumours [14].

CT (Fig. 2.3(d)) and MRI (Fig. 2.3(e)) are anatomical imaging modalities which primarily rely on differences in either electron density (CT) or water hydrogen density (MRI) to obtain images. In terms of sensitivity and specificity, CT and MRI are equal to detect tumour regions [23]. CT images demonstrate good bone-tissue contrast and are commonly used for staging HNC as images are obtained rapidly decreasing movement and respiratory artefacts and chest and liver can be imaged at the same time as the neck for distant metastases. Compared to MRI, the main disadvantages of CT includes the use of ionising radiation, artefacts from dental amalgam and swallowing and low contrast between soft tissues. Further, for staging of primary oropharyngeal tumours, MRI (77%) is more accurate than CT (67%) [23] particularly for BoT tumours [24]. For laryngeal tumours, MRI is used in assessing laryngeal cartilage invasion [14] as it decides the outcome of RT treatment. Furthermore, the combination of clinical information with MRI is demonstrated to be more accurate (87%) than with CT (80%) for staging of a laryngeal tumours [25].

2.3.2 Treatment options for HNC

Selection of a treatment modality for HNC is usually based on the size and location of the primary tumour, the status of the regional lymph nodes (TNM stage) and the general condition of the patient. The primary goal of treatment of HNC is to eliminate the cancer while preserving the function of the nearby nerves, organs, and tissues. The main treatment options for HNC are surgery, RT and chemotherapy [14]. For early disease, HNC can be treated with single treatment - either surgery or RT. However, advanced disease may require combination of these treatments. Surgery is an operation with an aim to remove cancerous tumour and some surrounding healthy tissues. Chemotherapy use drugs to destroy cancer cells, usually by stopping the cancer cells' ability to grow and divide. Although HNC are sensitive to chemotherapy (they shrink in size), chemotherapy is not a curative treatment. RT uses high-energy x-rays or other

particles to destroy cancer cells [15]. It can be the main treatment for HNC, or it can be used after surgery to destroy small areas of cancer that cannot be removed surgically.

Currently the preferred treatment modality is RT compared to surgery, in an effort to preserve an organ [26]. Approximately 60% of HNC patient receive some form of radiotherapy as a component of their treatment [26]. For example, in data for head and neck oncology and radiotherapy dataset [27], of the 6443 head and neck tumours, 3976 tumours or 61.7% had RT treatment with highest for oropharynx tumours (Fig. 2.4). The RTP process required for RT treatment is described in next section (Section 2.4).

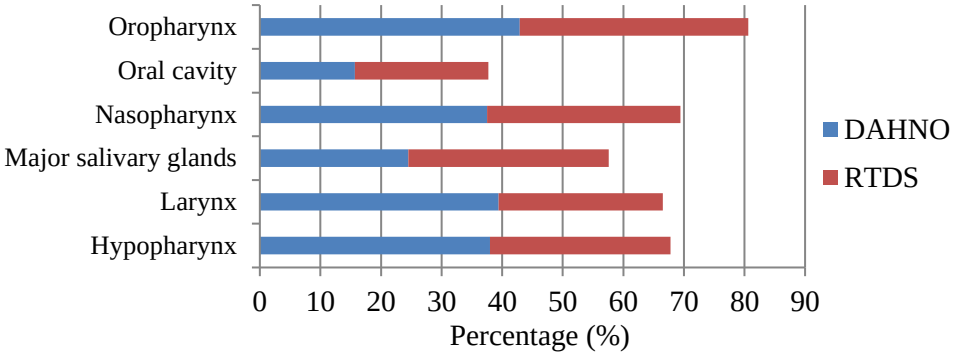


Fig.2.4: Proportion of head and neck tumours, recorded by data for head and neck oncology (DAHNO) and radiotherapy dataset (RTDS) report for England, having radiotherapy treatment, by tumour site group (Statistics taken from [27]).

2.4 Radiotherapy treatment planning (RTP)

The aim of RT is to deliver a precisely measured dose of irradiation to a defined tumour volume with as minimal damage as possible to the surrounding healthy tissues, resulting in eradication of a tumour [28]. New radiotherapy techniques such as IMRT and image guided RT greatly enhances the ability to precisely treat the defined target volumes and limit the radiation to critical normal structures. However, this requires precise treatment plan for accurate treatment delivery. RTP (Fig. 2.5(a)) process involves determination of patient position, construction of immobilisation devices, calculation of optimal treatment parameters such as target volumes (Fig. 2.5(b)), dose planning (dose distribution and dose fractionation using a computerised RTP system (Fig. 2.6)), plan verification before treatment delivery [5].

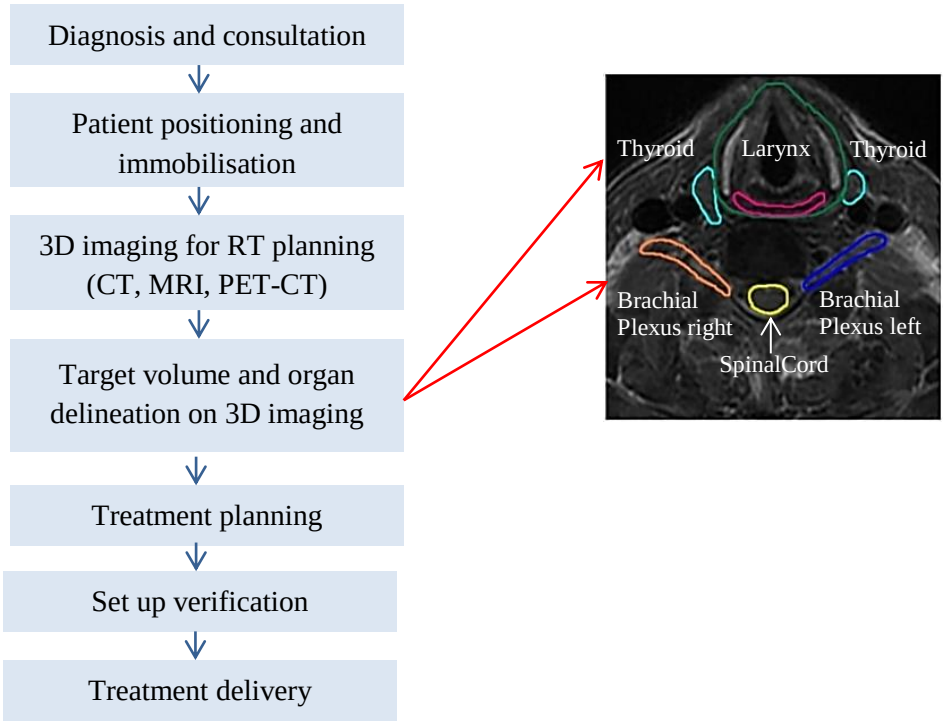


Fig. 2.5: (a) Steps of radiotherapy treatment planning (RTP). Precise target volumes and organs at risk (OAR) delineation is essential in RTP to deliver a high dose to the tumour while limiting the dose to healthy tissues. (b) Axial MRI slice demonstrating target volumes and OAR delineation (different colour outlines).

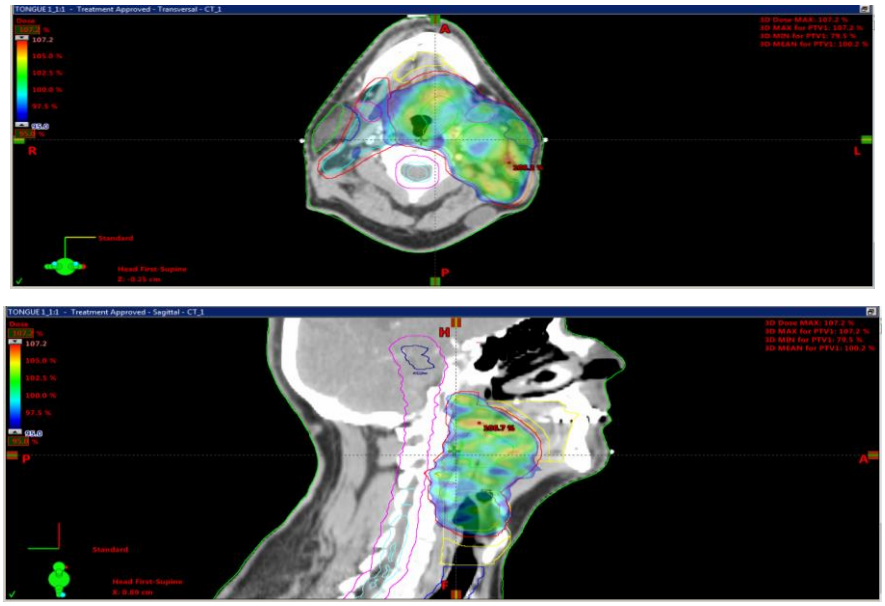


Fig.2.6 : Screen shot of radiation treatment plan on axial and sagittal CT for radiotherapy treatment of base of tongue (BoT) tumour on Eclipse treatment planning software (Courtesy of The Beatson West of Scotland Cancer Centre, Glasgow)

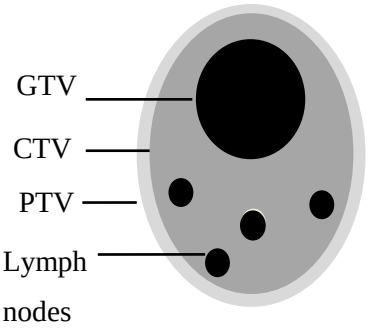
Use of medical imaging data in RTP allows improved tumour characterisation, delineation and treatment verification. Of all these parameters, the determination of target volumes is the most important and highly uncertain parameter [28]. Information of geometry of a tumour and its relation to surrounding tissues and organs from clinical examination and a range of imaging modalities (CT, MRI, PET-CT) are used to define target volumes and critical structures (OAR) (Fig. 2.5(b)) on treatment planning imaging modality. Definition of target volumes and current procedure used in RT for target volume delineation is described in next Section (Section 2.4.1).

CT is presently a treatment planning imaging modality used from initial delineation of the tumour volumes to the creation of digitally reconstructed radiographs for patient treatment setup and treatment verification. CT allows tumour visualisation and reconstruction of patient model in three-dimension (3D) with anatomical and geometrical data and has intrinsic information on the relative electron density of the various tissues used for the dose calculation algorithms. However, in recent years, dedicated MRI scanners [29], that can completely eliminate the use of CT, are being developed to smoothly integrate MRI in radiotherapy workflow. This is because MRI reduces the risk of healthy tissue involvement as it yields superb soft-tissue visualisation and provides several imaging modalities such as T1, T2, FLAIR for the identification of movements, function and physiology.

2.4.1 Target volumes and organs at risk (OAR) delineation in RTP

The target volumes consists of a volumes that includes the tumour demonstrated through imaging or other means and its spread to the surrounding tissues or lymphatic's. The International Committee of Radiation Units and Measurements (ICRU) have developed nomenclature and procedures for defining tumour volumes and normal tissue structures in RT [28]. There are three main volumes in RTP; gross tumour volume (GTV) that denotes demonstrable tumour (which can be imaged), clinical tumour volume (CTV) denotes GTV with volume of tissue with suspected tumour and affected lymph nodes and planning tumour volume (PTV) includes CTV plus margins for organ motion and geometric uncertainties (Fig. 2.7). Besides defining

target volumes, OAR have to be defined. OAR are any organs which might be significantly impacted by the radiation dose delivered. In the head and neck, the spinal cord, brainstem and salivary glands such as the parotid glands are OAR.

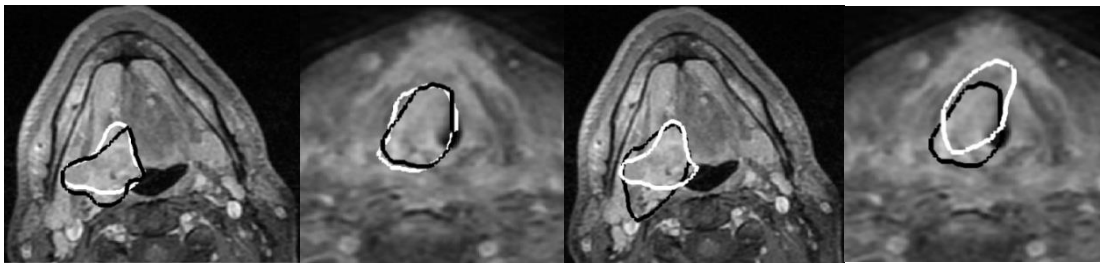


Term	Definition
Gross tumour volume (GTV)	Macroscopic tumour volume as detected by clinical examination and medical imaging (ultrasound, CT, MRI, PET)
Clinical target volume (CTV)	GTV with margin for subclinical microscopic spread and affected lymph nodes
Planning target volume (PTV)	CTV with margin added for organ motion and set up inaccuracy (usually CTV+3mm)

Fig.2.7: Definitions of target volumes (GTV, CTV and PTV) used in RTP (radiation treatment planning). In this thesis GTV is the volume of interest.

2.4.1.1 Current delineation methods

Although there have been significant improvements in RT delivery techniques over the past years, the target delineation method has changed little. GTV delineation is central for RTP. In the present day clinic this volume is manually contoured on medical images by a CO (Fig. 2.8). This manual contouring process is subjective that usually depends on expertise and experience of CO. Also, manual delineation from large medical image dataset is time-consuming process. The target volumes and OAR manual delineation can take from 45 minutes to 3 hours per patient depending on complexity of cancer regions and information availability [6], [7], [30].



(a) (b)

Fig.2.8: (a) intra- and (b) inter-variability in manual delineation (white and black outline) in GTV of base of tongue (BoT) and larynx tumours from MRI (Courtesy of The Beatson West of Scotland Cancer Centre, Glasgow)

Further, the presence of a tumour often distorts local anatomy, making definition of the tumour boundary more challenging. This distortion, combined with factors such as necrosis (death tissues), inflammation, and uncertainty in mucosal spread, lead to high intra- and inter-observer variability in manual target volume delineation. As an example, Fig. 2.8 shows qualitative results of intra-observer variability in two contour volumes from same CO (Fig. 2.8(a)) and inter-variability among two COs (Fig. 2.8(b)) in defining BoT and larynx tumour on MR images obtained from the Beatson West of Scotland Cancer Centre, Glasgow. As for quantitative results, [31] demonstrated mean intra-observer error of 13% for a pharyngeal cancer, [8] demonstrated coefficient of variation of 12.9% for an oropharyngeal tumours and 13.8% for a laryngeal tumours respectively in manual delineation of GTV from MRI. Thus, to reduce this manual bias (human subjectivity), the method to determine the gold standard for target volumes and OAR followed in the current clinical practice in the Beatson West of Scotland Cancer Centre, Glasgow is:

1. CO, subspecialising in head and neck malignancy, delineate target volumes (GTV, CTV and PTV) and OAR manually in slice-by-slice fashion on medical imaging using the guidelines described in [32], [33].
2. Target volumes and OAR outlined by CO are discussed in head-and-neck weekly team meeting (which consist of COs, radiologists, surgeons and radiation oncologists) and consensus manual outline is obtained. This consensus manual outline is considered as a gold standard for RT treatment in a current clinical practice and is also considered as a gold standard to validate the automatic delineation results obtained in this thesis. Thus, a major goal of current RT procedures and systems is to reduce the uncertainty and delineation time in target volumes and OAR by developing automatic systems for delineation [9].

Software tools (Table 2.1) are developed and validated for automatic delineation of OAR and/or lymph nodes of HNC from CT images and are used in a current clinical practice. This software tools have allowed time saving in range of 30% to 40 % [6], [7], [30] with agreement (dice similarity coefficient) ranging from 0.68±0.25 to 0.97±0.03 (1 indicates perfect agreement) when compared with manual delineation. Most segmentation results obtained from software tools, however, require

careful reviewing and editing by an expert. Further, this tools utilise atlas-based techniques [6], [7], [30] for delineation. In atlas-based methods (described in more detail in Chapter 3), a reference or atlas is created from manual segmentation of the region of interest. The image-based deformable registration is used to transfer this segmentation to the target image. These atlas-based methods are suitable for contouring anatomical structures on CT that do not exhibit extensive geometric and texture variations, such as OAR. For GTV with extensive geometric, anatomic and texture variations, these atlas based methods demonstrate significant variation when compared to reference segmentation results [34].

Table 2.1: Few commercially available systems for automatic atlas based segmentation of OAR and/or lymph nodes of HNC from CT images (For full list please refer [9], Table 2)

Systems	References	Description
Atlas-based automatic segmentation (ABAS) by Elekta	[30]	Atlas-based technique: hierarchical approach for non-rigid registration to transform atlas contours to new subject. Fusion of multiple atlas to increase accuracy.
SPICE (Smart Probabilistic Image Contouring Engine) by Philips Healthcare	[6]	Atlas-based technique: deformable registration to transfer contours from atlas to the target image, model based segmentation and probabilistic refinement results.
Varian smart segmentation knowledge based contouring by Varian Medical Systems	[7]	Atlas-based technique: deformable registration combined with multi-resolution approach and modified to include active forces.

Few PET, PET-CT or MRI based techniques [35]–[37] for semi-automatic or automatic delineation of OAR or lymph nodes of HNC are described in this section. Researchers in [35] validated ten PET segmentation tools developed using different thresholding techniques for lymph node metastases and concluded that PET can be useful for the detection, but has no added value for segmentation of lymph node metastases compared to CT. Texture classification was performed in [37] on co-registered PET-CT to discriminate normal and abnormal tissues and established that

this generates better discrimination results than using textural features from a single modality. However, this requires further extension for the segmentation of a target structures for use in HNC treatment. Semi-automatic probabilistic approach developed in [36] for the segmentation of RT related anatomical structures utilised seven MRI sequences with overall accuracy of 80% when compared with manually segmented structures. Obtaining multi-sequence (T1, T2, FLAIR) MRI, however, is not always feasible. These algorithms [35]–[37] for GTV segmentation demonstrate initial results and are not utilised in a current clinical practice.

Table 2.2: Few semi-automatic or automatic delineation (segmentation) algorithms for GTV of HNC from CT, PET or PET-CT

Imaging modality	References	Limitations
CT	[38]	Require user-defined region of interest or information from different imaging modality to segment GTV as low contrast between GTV and surrounding tissues
PET	[39]	Most threshold based algorithms, thus GTV sensitive to the selected threshold value, low anatomical information and may underestimate GTV compared to CT
PET-CT	[40]	In some cases PET/CT-based GTV can be significantly smaller than CT

Automatic or semi-automatic delineation (segmentation) tools are also developed for GTV of HNC from CT, PET and PET-CT (Table 2.2). Algorithms developed for GTV segmentation from CT normally rely on user defined region of interest [38] or information from different imaging modality due to limiting contrast between different tissue types on CT. For PET, algorithms [39] mostly utilise threshold based techniques as metabolic activity of a tumour region is high and is clearly demonstrated on PET as high intensity region. PET normally underestimates the GTV compared to CT and segmentation results are often sensitive to the selected threshold value (for example 5% change in threshold contour level can lead to 200% increase in volume [39]). This limitation is reduced by registering PET with CT and segmenting

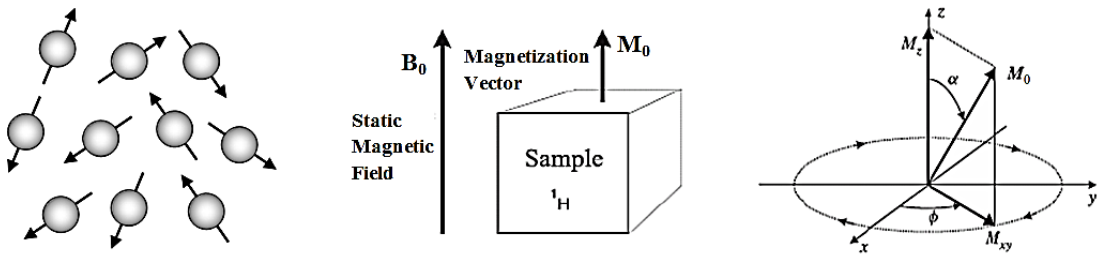
GTV on registered image [40]. However, in some cases PET/CT-based GTV can be significantly smaller than CT. Algorithms developed for semi-automatic or automatic GTV segmentation from axial MR images are described in Chapter 3.

2.5 Magnetic resonance imaging (MRI)

MRI is a non-invasive imaging technique used in radiology to investigate, in detail, the anatomy and physiology of the body in both health and disease. Clinical MRI uses the magnetic properties of hydrogen atoms present in the human body and its interaction with both a large external magnetic field and a radio frequency (RF) field to produce highly detailed images of region of interest.

2.5.1 Basic principle of MR imaging

The basic principle of MR imaging rely on the spinning motion of a specific nuclei in biological tissues, known as MR active nuclei. The hydrogen nucleus (with one proton and no neutrons) (^{1}H) is the MR active nucleus used in clinical MRI because it is abundant in the human body (in fat and water; 70% of the body is made up of water). Due to the presence of only one proton (positive charge), the hydrogen atom has angular momentum or spin. Any electrically charged particles which moves, creates a magnetic field called magnetic moment. In general, the random orientation of millions of these spins (Fig. 2.9(a)) results in no net magnetic field. However, when a human body is placed in a large external magnetic field B_o, many of the free hydrogen nuclei (^{1}H) align themselves with the direction of the magnetic field (z-axis) (Fig. 2.9(b)).



(a) (b) (c)

Fig. 2.9: Behaviour of hydrogen nuclei (a) with no external magnetic field B_0, spins has random orientations (b) external B_0 produce net magnetisation M_0 (c) The magnetisation vector M_0 precesses about the z-axis M_0 [41]. Copyright © 2006 Pearson Education, Inc.

Further, the angular momentum of hydrogen nuclei causes the spins to precess at a certain frequency. This characteristic frequency (ω_0) is given by the Larmor equation [42]:

$$\omega_0 = \wp B_0 \qquad 2.1$$

where ω_0 is the resonance frequency, B_0 is the magnetic field strength in Tesla, $\wp$ is the gyromagnetic ratio and has a unit of radians per second per tesla (T). The gyromagnetic ratio of hydrogen is 2.68×10^8 radians/s/T (or 42.57MHz/T) [41] .

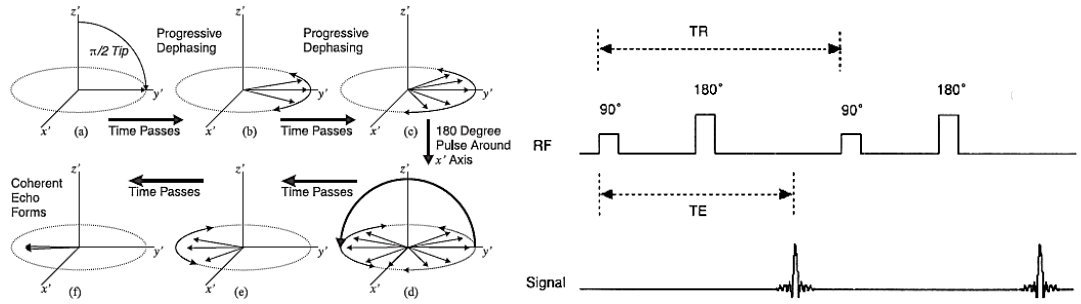
In normal room temperature there are few more spins in the direction to the magnetic field B_0 than in opposite direction. This slight majority produces a small net magnetic field, the net magnetisation M_0 (Fig. 2.9(b)), in the direction of B_0. To observe this net effect (M_0), the second field called B_1 is introduced in the direction perpendicular to B_0. This field B_1 is applied at the resonant frequency of the spins, which for the proton is in the radio-frequency (RF) range. Thus, is also known as RF field. As a result of this B_1, the spins move out of alignment into transverse (x-y) plane, so that they can be measured. The angle α (Fig. 2.9(c)) at which spins moves out of alignment from z-axis and into x-y plane is often referred as 'flip angle'. The M_o now has two components the transverse magnetisation component M_{xy} and the longitudinal magnetisation component M_z (Fig. 2.9(c)).

A receiver coil is situated in the transverse plane. The spins rotate around the transverse plane, as a result of resonance magnetisation component M_{xy} passes across the receiver coil, inducing a voltage in it according to Faraday's law. This voltage is the MR signal. Once the RF pulse is removed, the energy of the spins given by the RF pulse starts to decrease because the spins try to realign with B_0. The amplitude of the MR signal consequently decreases. This is called free induction decay (FID).

The process of decreasing the spin energy is called relaxation. This relaxation causes the loss of phase coherence in the transverse (x-y) plane which is called transverse relaxation time T2 and the exponential recovery of longitudinal magnetisation which is called longitudinal relaxation time T1. The received signal actually decays more rapidly than T2 and that time is known as $T2^*$. The time allowed

for the recovery of longitudinal magnetisation is called the repetition time (TR) and is measured in milliseconds (ms).

The FID decays faster, with time constant $T2^*$ due to the fixed perturbations in the magnetic field. These perturbations cause the precession of some spins to speed up and others to slow down. A short duration 180^o (RF) pulse is applied that turns the spins 180° so that the phases are reversed. Now, faster spins lag and slow spins lead forming a spin-echo (Fig. 2.10(a)). The time interval from initial 90^o pulse to the formation of spin-echo is called echo time (TE) and given in milliseconds (ms). This use of a 90° excitation pulse followed by a 180° RF pulse is known as a spin-echo sequence (SE) (Fig. 2.10(b)). This sequence is repeated to produce an image. For instance, to produce an image with a resolution of 128x128 pixels, the entire procedure for excitation and signal acquisition must be repeated at least 128 times.



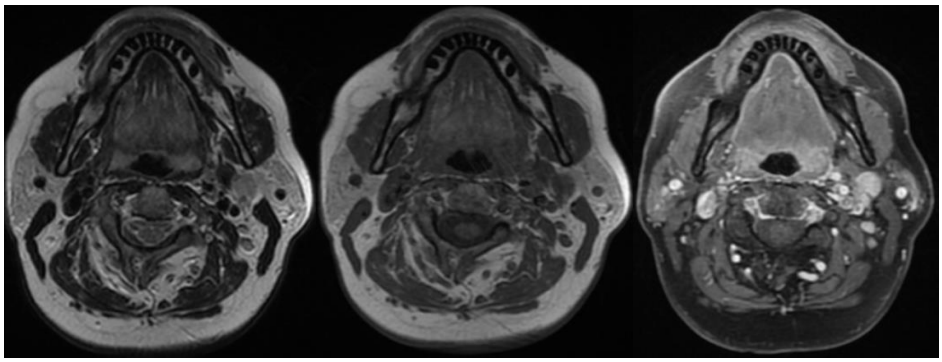
(a) (b)

Fig. 2.10: MR sequence (a) formation of spin-echo (b) spin-echo MR pulse sequence [41] Copyright © 2006 Pearson Education, Inc.

2.5.2 Contrast mechanism

One of the main advantages of MRI is its excellent soft-tissue contrast, which can be widely manipulated. The signal measured will depend on the number of spins present (the proton density, PD), as well as their T1 and T2 relaxation times. By altering the sequence timings (TE and TR) and sometimes the flip angles too, the image contrast can be weighted so that one of these effects predominates, although all maintain a contribution [43]. For example, to acquire T1-weighting in the image, a short TE will reduce the effects of T2-weighting, whereas a long TR will enhance the differences in T1 relaxation times.

The different tissue signals within an image can be further manipulated by using exogenous contrast agents. The most common are the gadolinium-based agents. Gadolinium (Gd) is a paramagnetic ion, with a high relaxivity; that is, it disturbs the local magnetic field of nearby protons and results in a shortening of T1 and T2 relaxation times. In a T1-weighted image with Gd, referred as T1+Gd, produces an increase in signal intensity. They are often used in cancer, as the increased blood supply to rapidly growing tumours preferentially takes up the agent and thus tumours can be easily noticed. Fig. 2.11 shows T2, T1 and T1+Gd weighted images with contrast in different tissue types demonstrated by different grey levels. These different weighted images (T1, T2, T1+Gd) can be considered together for the assessment of cancer regions and are sometimes referred as multi-spectral, multi-channel or multi-sequence MRI.



(a) (b) (c)

Fig. 2.11: MR images with different contrast mechanisms (a) T2 (b) T1 and (c) T1+Gd. T1+Gd are often used to identify tumour as tumourous tissues on T1+Gd MRI slice appear bright (higher grey level) compared to healthy tissues and thus can be easily noticed.

2.5.3 MRI data acquisition

MRI use both Larmor frequency and the phase of the transverse magnetisation to encode spatial position [43]. In this section, frequency and phase encoding of spatial position and fundamental of MR image formation is discussed. Although MRI can obtain images at arbitrary location and orientation, MR images are normally obtained in axial, coronal and sagittal planes (Fig. 2.3(a)).

2.5.3.1 Slice selection

The MR signal recorded thus far has been from the entire sensitive volume of the (RF) receiver coil. In order to form an image, the signal must be spatially localised. This is achieved using a method known as slice selection [44], whereby only a finite section of spins (the image slice) is excited into the transverse plane. First, a gradient, or a linear change in the B_0 field, is created along the slice direction causing the frequency of the spins to change in this direction. A specially shaped RF pulse, called a sinc pulse, containing a bandwidth of frequencies will then excite only the spins at corresponding resonance. The width (slice thickness) and location of the slice can be altered by changing the gradient and the bandwidth. Due to practical constraints the sinc pulse has to be truncated so that the profile is no longer a perfect square, leading to cross-excitation of adjacent spins. To improve imaging coverage without adding slices or increasing thickness, the slice gap (spacing in between slices) is increased.

2.5.3.2 Frequency and phase encoding

Two additional gradients are required to spatially encode the image in the two dimensions (x and y) [43]. The first gradient is applied at the time of signal measurement, called the readout or frequency-encoding gradient. This changes the frequency along the direction of the gradient so that the frequency differences in the final signal correspond linearly with spatial location. For the remaining direction, the signal is encoded by changes of phase. This is done by applying the gradient at the start of the sequence and repeating it many times, each with a different increment of the gradient amplitude. The number of pixels (image matrix) in this direction determines the number of phase-encoding steps required.

2.5.3.3 2D and 3D imaging

Multiple slice imaging is achieved by utilising the period between the echo collection and before the next excitation pulse (TR–TE) to excite adjacent slices. To account for imperfect excitation profiles, slices are usually acquired with an inter-slice gap. In 3D imaging, phase encoding is also utilised in the through-plane direction. This improves the spatial resolution in this dimension, with slice-selection (typically 2-5mm), to that of the in-plane (x-y) dimensions (<1mm). As a result the imaging voxels (pixels) are more isotropic and the three-dimensional visualisation and reconstruction of the data are much better. The scan time increases by a factor equal to the number of phase

encodings in the third dimension but the increased number of signal measurements means that signal-to-noise is better. However, to reduce scan time, MRI slices are normally obtained with large spacing in between slices thus, producing anisotropic voxels with slice spacing of 4-6 mm and in-plane (x-y) resolution of 0.43x0.43-0.94x0.94 mm^2.

Fig. 2.12 demonstrate a MRI slice and its associated terms such a pixel, voxel, slice thickness (in mm), in-plane resolution (Fig. 2.12(a)) and sequence of MRI slices (MRI volume) with spacing in between slices (in mm) (Fig. 2.12(b)).

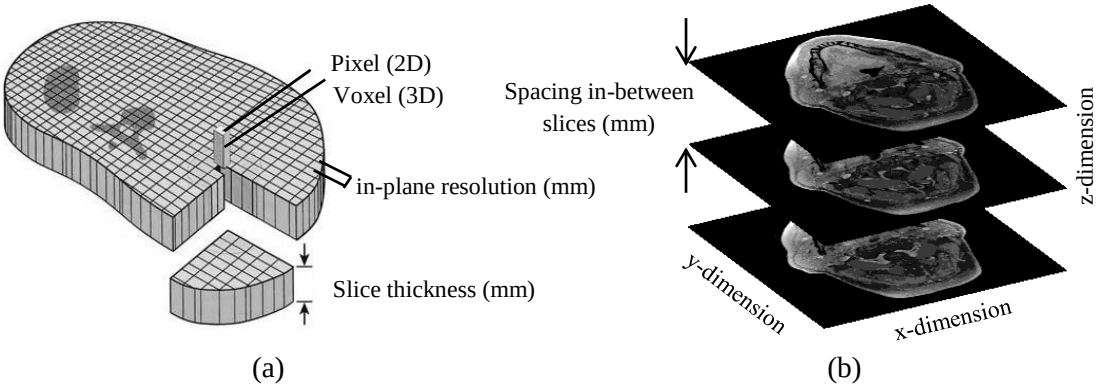


Fig. 2.12: (a) MRI slice and its associated terms such as resolution and slice thickness in mm (millimetre), pixel and voxel (b) MRI volume (sequential MRI slices from a single patient) demonstrating large spacing in between slices in mm.

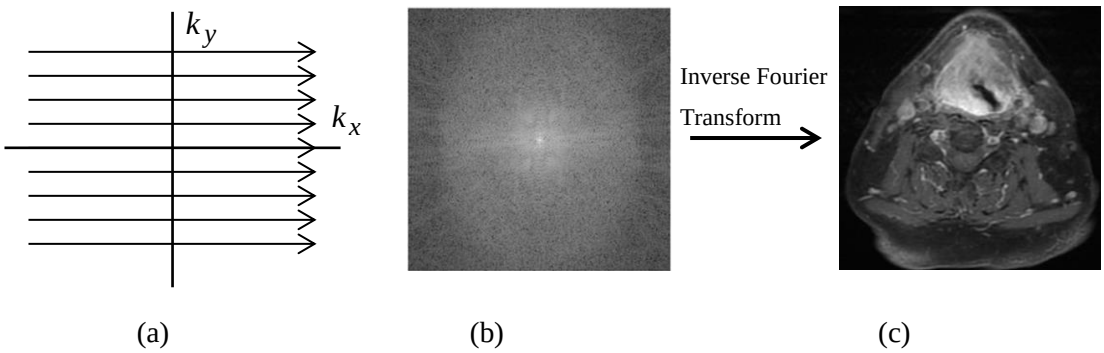
2.5.4 Image reconstruction

As the data acquired from 2-D MR imaging pulse sequence can be interpreted as a scans of Fourier space, the image reconstruction algorithm in MRI is the inverse Fourier transform (FT) [44]. FT is used in MRI to decode the detected signal and spatially assign each pixel, a row and column in the final image.

k-space

k-space is an array of numbers representing spatial frequencies in the MR image. It is often represented as a diagram (Fig. 2.13) with the x- and y-axis to illustrate the acquisition of MR data, although elements between the image-space and k-space do not correspond one-to-one. Each data point in k-space contains information about every pixel in the image. A row of k-space corresponds to the echo data collected with each application of the phase-encoding gradient. To produce an image, the full area of

k-space must be traversed. Data near the centre of k-space correspond to small amplitudes of the phase encoding gradient, and the bulk of the image signal. Towards the edges of k-space, data are acquired with high phase encoding amplitudes, corresponding to low signal but high spatial detail. The inverse Fourier transform is directly applied on the collected k-space data for image reconstruction.



(a) (b) (c)

Fig.2.13: Acquisition of MR image (a) and (b) k-space, the data from each measurement fills a different horizontal line (c) spatial domain representation of an acquired MR image.

2.6 MR image artefacts

Artefacts can be defined as an unintended, unexpected features [45] seen on MRI slices which can lead to inaccurate diagnosis, affect quality of image or may be confused with pathology. These artefacts are caused due to various reasons such as the underlying physics of the energy-tissue interaction, data acquisition errors (such as patient motion), or image reconstruction algorithm's. Examples of different artefacts in MR image include noise, aliasing, motion artefacts, chemical shift, metal artefact, partial volume, Gibbs phenomenon (ringing artefact), and intensity inhomogeneity. This section introduces the artefacts observed in MR images used in this thesis (Fig 2.14) with detail description in [45]. For automatic MR image analysis, ignoring these artefacts can introduce significant errors in quantitative measurements, such as estimations of the tumour volumes.

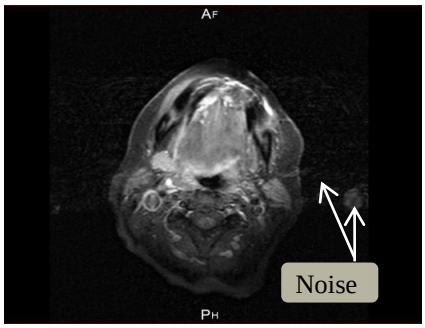
2.6.1 Noise

In MRI, noise can refer to the physiological and respiratory distortions occurred due to patient's breathing or eddy currents produced due to changing magnetic fields or many other distortions. The dominant noise in MRI, however, is the thermal (Johnson) noise [41]. The principal source of this noise in most MRI scans is the subject (patient)

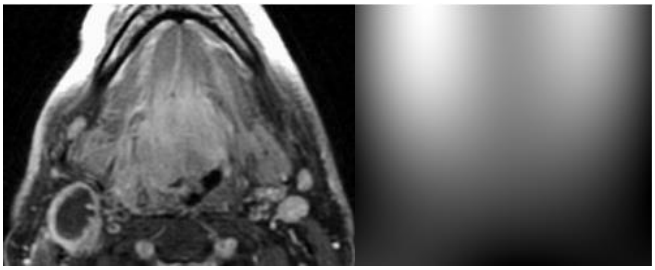
to be imaged, followed by receiver coil and electronics attached to it. Usually all the sources of thermal noise are considered in one additive component $\eta(x,y)$ given as:

$$I(x,y) = o(x,y) + \eta(x,y) \qquad 2.2$$

where, $I(x,y)$ is the image rendered by MRI system, $o(x,y)$ is the original signal and $\eta(x,y)$ is the noisy component of the reconstructed signal. These noise components generate unusual observations, or statistical outliers, on MR image that differ significantly from most MR images that do not contain those noisy components. For instance, neck MRI image (Fig 2.14(a)) normally comprises some pixels without signal (normally referred as background); hence the output signal acquired in such regions must be uniformly zero. However, due to noise the output signal in such regions having varying (intensity) values (Fig. 2.14(a)).

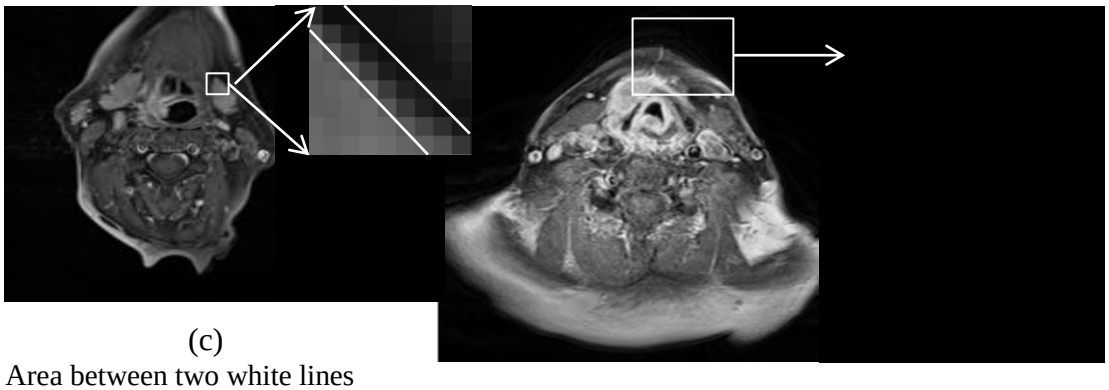


(a)



MRI slice Bias field (IIH)

(b)



(c)

Area between two white lines demonstrate partial volume voxels

Gibbs ringing

(d)

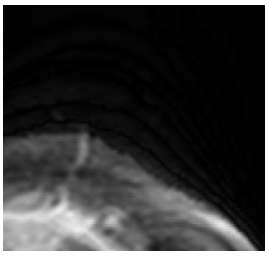


Fig.2.14: Main artefacts observed on MRI slices used in this thesis (a) noise (b) intensity inhomogeneity (IIH) (c) partial volume (d) Gibbs ringing.

2.6.2 Intensity inhomogeneity (non-uniformity)

Intensity inhomogeneity (IIH) [46] is the low frequency smooth intensity variation often seen in MRI slices (Fig. 2.14(b)). IIH is also often known as bias field or shading artefact. IIH is caused by factors such as RF excitation field inhomogeneity, use of phased array coils, non-uniform reception coil sensitivity, eddy currents driven by field gradients, as well as electro-dynamic interactions with the object often described as RF penetration and standing wave effect [47]. In modern MRI scanners these variations are often low enough that they are difficult to detect visually, however, they affect the automated image analysis techniques, particularly segmentation techniques that assume homogeneity of intensity within each tissue class. A widely used way to represent IIH [46], for image analysis, is to consider all sources in one multiplicative factor $\beta(x,y)$. Image obtained from MRI scanner can thus be modelled as:

$$I(x,y) = o(x,y)\beta(x,y) \qquad 2.3$$

2.6.3 Partial volume

The partial volume effect (PVE) [48] is the loss of contrast between two adjacent tissues in an image caused by insufficient resolution so that more than one tissue type occupies the same voxel (or pixel). That may induce a partial volume artefact, dependent on the size of the image voxel. For example, if air and tissue signal occupy the same voxel (Fig. 2.14(c)), the intensity of that voxel depends not only on the imaging sequence and tissue properties, but also on the proportions of air (S_{bg}) and

tissue (S_{fg}) signal present in the voxel [48]. Thus, MR signal (intensity) from the entire voxel $o(x,y)$ will then reflect the weighted average of signals S_{bg} and S_{fg} from the two components and can be given as:

$$o(x,y) = w_{bg} S_{bg}(x,y) + w_{fg} S_{fg}(x,y) \quad 2.4$$

The partial volume effect is minimal with thin slice thickness and sufficiently high resolution, so that fat and water or other different structures are unlikely to occupy the same voxel.

2.6.4 Gibbs phenomenon (ringing artefact)

The MR image is reconstructed in spatial domain from k-space, which is a finite sampling of the continuous signal, using inverse Fourier transform. At high-contrast boundaries (jump discontinuity in mathematical terms) the Fourier transform corresponds to an infinite number of frequencies, and since sampling is finite, the discrepancy appears in the image in the form of a series of lines. These spurious oscillations (lines) are called the Gibbs ringing artefact in MRI [49] and are prevalent particularly at the tissue boundaries. Fig. 2.14(d) shows prominent light and dark line along the sides that fade as they approach the sides of the image. This problem can only be resolved by smoothing filters or with a higher acquisition matrix, to smooth the object.

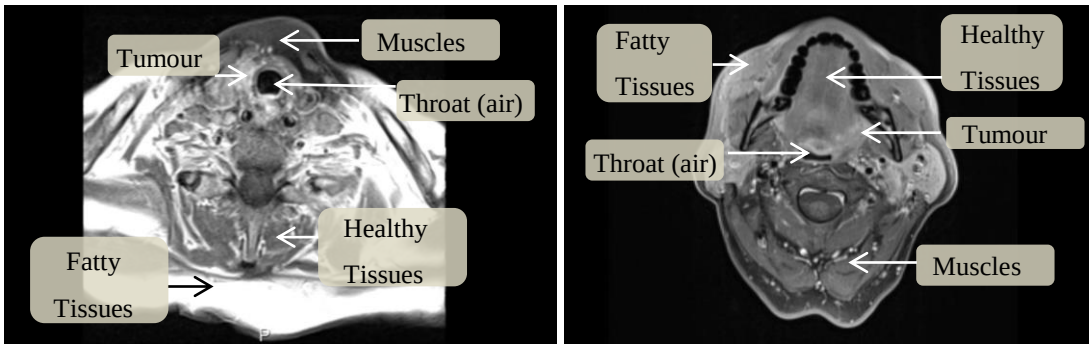
Some image processing solutions to reduce the effect of these artefacts for quantitative MR image analysis are discussed in Chapter 3.

2.7 MR imaging in HNC

MRI slices for the assessment of HNC are normally acquired with the patient in horizontal (supine) position, and during quiet respiration using dedicated head and neck receiver coil [50]. To increase signal to noise ratio and to speed up MR image acquisition process, small phased-array surface coils are used inside the head or neck coil. However these phase array coils can induce IIH. MR images are obtained in three orthogonal planes (axial, sagittal, and coronal) (Fig. 2.3(a)). Common MR sequences used for HNC assessment and staging include T2, T1 and T1+Gd. Depending on the investigated region, axial slices are normally obtained in sequential fashion from

superior (head) to inferior (feet) direction with optimal slice thickness of 3–4 mm, spacing in between slices of 0%–50% [50]. The imaging matrix is at least 256**x**256.

For assessment of oropharynx and larynx tumours, according to the head and neck MR imaging protocol defined by [51], require obtaining axial T1+Gd sequence for both tumours. Advantages of T1+Gd include improvement of soft tissue contrast and cancer margin definition due to administration of contrast agent (gadolinium), as blood vessels and part of cancer region appear hyper-intense (bright/high intensity) area as compared to the normal regions and is often helpful in delineating primary tumours (GTV). Axial MRI slices of oropharynx and larynx region in presence of tumour (GTV) in Fig. 2.15 demonstrate different tissue types observed. There are five main tissue types indicated by five different grey levels: 1. bright intensity level include Fatty tissues and blood vessels, 2. high intensity region include tumourous tissues, 3. mid intensity region include healthy oropharynx or larynx region, 4. low intensity level include muscles and 5. no signal level include background or air region. Further, from Fig. 2.15 it can be noted that oropharynx particularly BoT and larynx tumour are close (adjacent/near) to the throat region and in the upper middle part of the MRI slice. This information is useful in automatic delineation (segmentation) of oropharynx and larynx tumours.



(a) (b)

Fig.2.15: Five main tissue types (labelled) observed on axial MRI slice of (a) larynx and (b) oropharynx (BoT) in presence of a primary tumour (GTV).

2.8 Conclusion

The anatomy and physiology of the head and neck region along with the common HNC, with particular emphasis on oropharynx and larynx cancer, were presented in

this chapter. Medical imaging (ultrasound, CT, PET, MRI) and treatment options (surgery, chemotherapy and RT) available for HNC were introduced. This chapter, further, discussed RT, which is currently a preferred treatment modality in an effort to preserve organ, and planning (RTP) required for RT treatment. The importance of automated (computer-aided) techniques for the delineation (segmentation) of target volumes in RTP was also illustrated. Finally, this chapter presented the MR imaging technology including the basic principle of magnetic resonance, data acquisition and image reconstruction. In next chapter (Chapter 3) various image processing techniques applied to MRI data for segmentation and quantification of a tumour are presented.

Chapter 3
MRI cancer segmentation and quantification techniques

3.1 Introduction

The previous chapter (Chapter 2) introduced HNC and available treatment options for HNC particularly the RT treatment. The main step in RT treatment is the delineation (segmentation) of tumour regions (GTV) from medical imaging. Some computer-aided algorithms developed for the segmentation of tumours of HNC from medical imaging such as CT, PET-CT were also discussed in the previous chapter. The purpose of this chapter is to give the review of image processing techniques developed for the tumour (GTV) segmentation, 3D reconstruction and volume calculation (quantification) from MRI.

The remainder of this chapter is organised as follows: Section 3.2 introduces general MRI cancer segmentation and quantification framework. Section 3.3 gives the overview of pre-processing techniques from MRI slices. Spatial interpolation techniques are presented in Section 3.4. Section 3.5 gives overview of medical image segmentation techniques. Section 3.6 reviews throat region segmentation methods with review of cancer segmentation methods in Section 3.7. 3D reconstruction and volume calculation techniques are presented in Section 3.8 with conclusion of this chapter in Section 3.9.

3.2 MRI cancer segmentation framework

As stated in Section 2.5.3, for a single patient, MRI scan near the cancer regions is obtained as a sequential MRI slices with a slice spacing of few millimetres (mm). From image processing point of view, two approaches [52] (Fig. 3.1) are considered for cancer regions segmentation and quantification from these MRI slices:

1. Each slice is considered as a separate 2D image and cancer region (GTV) is segmented separately from each slice (Fig. 3.1 Method 1) (2D segmentation) to obtain contours (outlines) of the region. Further interpolation (to add contours in

between real contours) and 3D reconstruction of the obtained contours is performed for the quantification and 3D visualisation (Fig. 3.1 Method 1).

2. All sequential slices are considered together as one volume. Initially, interpolation is performed to add slices in between real MRI slices and volume is reconstructed from real and interpolated slices. Cancer region is segmented from the reconstructed MRI volume (Fig. 3.1 Method 2) (3D segmentation) for further quantification and 3D visualisation.

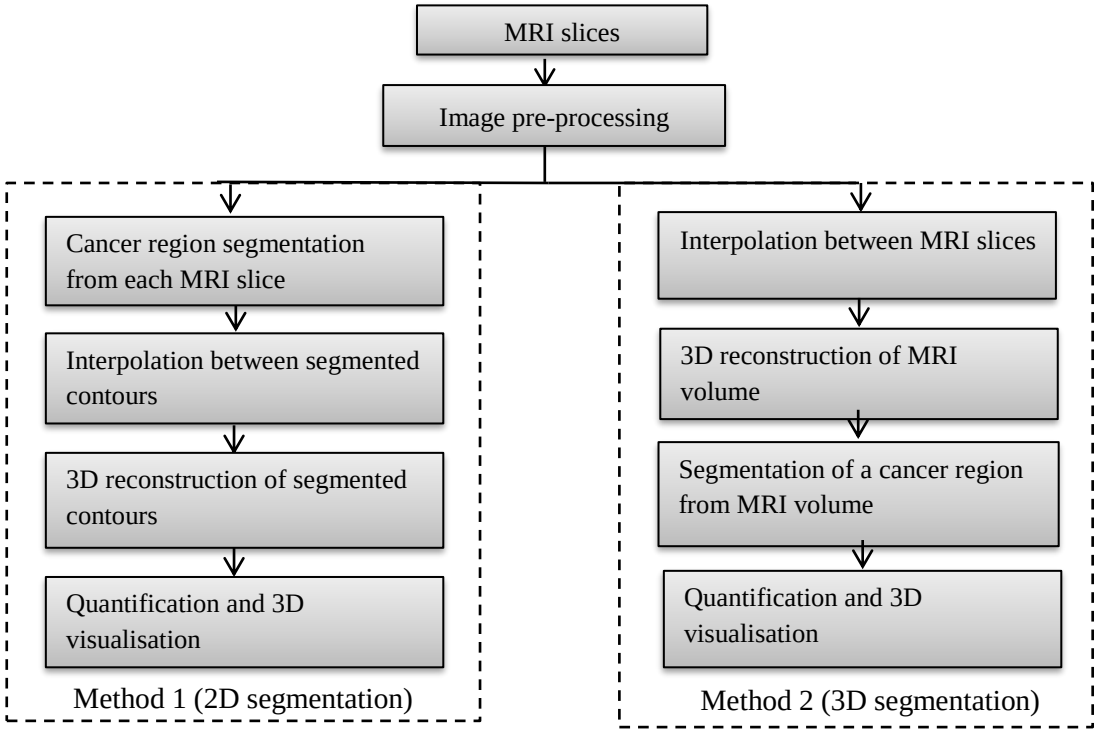


Fig. 3.1: General framework for MRI cancer segmentation and quantification

3.3 Image pre-processing techniques

Pre-processing of MRI slices is an important task before applying segmentation algorithms because of the challenges of automatically segmenting a cancer region. Cancer regions on MRI slices are not well-defined structures with weak edges and poor contrast and are degraded due to presence of artefacts. Thus, the principle objectives of image pre-processing techniques [53] are to enhance the quality of required features such as edges and contrast between different tissue types and to reduce the effect of unwanted features (artefacts) such as noise and intensity variation from an image to make image suitable for the segmentation algorithms. This section reviews such image enhancement and artefacts reduction techniques.

3.3.1 Image enhancement techniques

Histogram equalisation (HE) [54] is widely used technique to improve tissue contrast in MRI slice [55], [56]. Histogram of grayscale (intensity) image is a graph showing number of pixels at each intensity value found in that image and HE is a technique which stretches the histogram of the image, to make use of the entire intensity range of the image. The traditional HE [54], however, over-enhances the image background. Thus, non-linear HE technique was proposed in [57], which equalises background and foreground histogram separately, thus preserving background brightness while improving the contrast in the foreground regions.

Other techniques for image enhancement and noise suppression include the extended neighbourhood filter [58] or the edge preserving smoothing filter [59] to improve the contrast of brain MR images. In [58], enhanced image intensity at a given pixel was calculated as a weighted summation of extended neighbourhood intensities. The weights of the neighbourhood were expressed as a function of distance and intensity gradient. It was shown that this filter works better for unconnected regions with small spatial extent (for example: small lesions). In [59], edge map was obtained through fuzzy edge estimate and intensity information. This map, which separates noise and prominent edges, was thresholded to reconstruct image with optimal contrast and smoothness. The threshold was determined based on objective function that maximises contrast between small lesions and other brain tissues. Thus these techniques [58], [59] are suitable for brain MR images particularly in presence of small lesions. Other techniques, which consider frequency or wavelet domain for image enhancement, can be found in [53].

3.3.2 Artefacts reduction techniques

Table 3.1 gives an overview of the methods that have been proposed to reduce the artefacts from MRI slices.

Noise The main methods applied to reduce noise from a MRI slice are based on non-linear filtering and morphological filtering [60]. Non-linear filtering such as SUSAN (Smallest Univalue Segment Assimilating Nucleus) technique [61] acts like Gaussian smoothing but reduces the strength of diffusion near edge pixels, and thus acts as an

edge-preserving smoothing filter. Yan et al. [62] combined non-local means filter and SUSAN technique to reduce noise while preserving fine anatomical structures with enhanced edges. Morphological filtering which is based on mathematical morphology uses a structuring element to remove small regions while preserving topology of large regions. In [63], a 3D morphological closing operation (with disk shape structuring element) was applied to brain MRI volume to remove vessels and nerves to simplify cerebrospinal fluid volume segmentation. Residual (background) noise was removed in [64] by adding the original image to the difference between morphological open and close (morphological filter) transformed image. A two stage iterative morphological filter using a variable structuring elements was used in [64].

Table 3.1: Summary of MR artefacts reduction techniques

Artefact	References	Techniques
Noise	[62]	Non-linear filtering – non-local means and SUSAN technique
	[63][64]	Morphological filtering – closing and opening operations using a structuring element
Intensity inhomogeneity (IIH)/Bias field	[65], [66]	Low pass filtering – in spatial domain [65] or in log domain [66]
	[10], [67], [68]	Polynomial and/or spline model estimation and entropy minimisation technique
Gibbs ringing	[69]	Matlab toolbox - Filtering and reconstruction techniques
	[70]	Addition of Gaussian shaped slopes

Intensity inhomogeneity (IIH)

Most IIH (bias field) (Fig. 2.14(b)) reduction techniques [10], [65]–[68] rely on the assumption that IIH is a low frequency smooth spatially varying function. Thus, low-pass filtering methods were employed in [65], [66] to reduce IIH. In [65], regions of small intensity variation were extracted using averaging filter. These extracted regions were then smoothed using Gaussian filter. This smooth data was further combined with original data to estimate the bias field. In homomorphic filtering [66], log transformed input image ($\boldsymbol{log(}\, I\,\boldsymbol{(}\, x\boldsymbol{,}\, y\,\boldsymbol{))}$) was subtracted from its low-pass filtered (LPF) version (

$LPF(\mathbf{log}(I(x,y)))$). Normalisation constant χ_{Nor} was added to preserve the mean or maximum intensity of the corrected image $o(x,y)$ as:

$$log(o(x,y)) = log(I(x,y)) - LPF(log(I(x,y))) + \chi_{Nor} \quad 3.1$$

Finally the corrected image $o(x,y)$ was obtained by exponentiation as:

$$o(x,y) = exp(log(o(x,y))) \quad 3.2$$

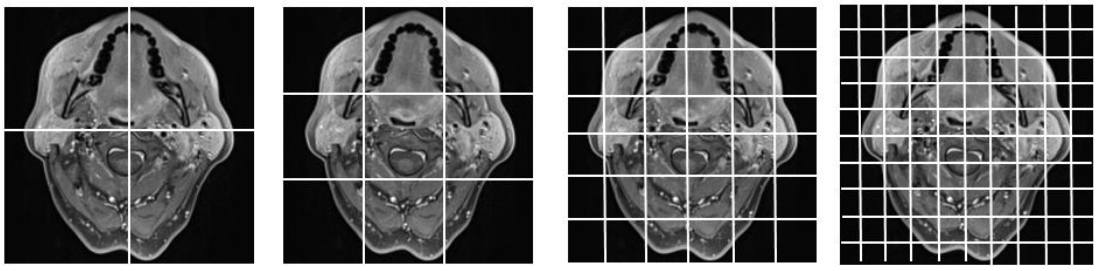
However, low frequency assumption is valid only if the anatomical structures in MRI slices are small and thus contain no low frequencies that may be filtered by a low-pass filter. For most MRI slices with large structures (such as neck muscles) this does not hold, which result in overlap frequencies of anatomical structures and bias field. Furthermore, edge artefacts can result from low-pass filtering.

Other techniques [10], [67], [68] estimated the smooth variation bias field using polynomials or spline techniques. Further, it was also observed that bias field increase the entropy (information) content of the IIH free image [10], [67], [68]. Thus, in [67] the estimation of bias field was modelled as energy minimisation approach. The bias field was estimated using polynomials. The coefficients of polynomials were determined by optimisation based on a special valley function, which was shaped by the a priori given mean intensities and standard deviations of the main tissue classes. However, the polynomial functions are insufficient to express the complex bias field observed, when MR images are obtained using phased array coils. In [68], spline model was used to estimate the bias field. The optimisation was done using stochastic annealing method such that the information of the corrected image was minimised while the global intensity statistic of the image was preserved. However, in [68] optimisation was performed on entire dataset, regardless of SNR (signal to noise ratio) variation. Thus, the estimated bias field was corrupted by low SNR regions. To avoid this corruption, local entropy minimisation technique was proposed in [10]. The review of IIH reduction techniques can be found in [46].

In [10], the initial bias field $\beta(x,y)$ (Eq. 2.3) was calculated by fitting a fourth order polynomial function to tissue pixels excluding background (near zero intensity) pixels. The polynomial function was fitted to tissue pixels using least square approximation. This initial field was then refined using bicubic spline defined by grid

of knots evenly spread out across the image by taking into account the value of the bias field at knot locations. The bicubic spline estimate of the bias field was optimised so as to minimise the entropy of the image [10]. The technique used a piecewise, optimise and merge algorithm. The optimisation started by identifying the knot τ_1 in the area of the highest SNR (highest $\beta(x,y)$ value) and minimising the entropy of the corrected image $o(x,y)$. Then knot τ_2 with the next highest SNR was recognised and the entropy of the area under τ_1 and τ_2 was minimised. This was repeated until all knots are optimised. Additional passes were performed starting again from the area of the highest SNR. Optimisation process stopped when there was no significant different between either the amplitude of knots and image entropy.

In this method [10], parameter knot spacing (spline distance) (Fig. 3.2) is important for the correct estimation of the bias field. In [10], this distance was set to a constant value of 21 mm for all MRI dataset. Boyes et al. [71] evaluated the performance of bias field correction algorithm for 3.0T brain MRI data with variation in the spline distance, i.e., the distance between the knots defining the B-spline mesh. Authors in [71] concluded that greater field strength (3.0 Tesla) generate high bias field. Thus, decreasing the spline distance from 200 mm to the range 50–100 mm (Fig. 3.2) led to improved results. However, authors in [72] showed that much smaller spline distances can cause algorithm to fail.



(a) (b) (c) (d)

Fig. 3.2: Illustration of knot spacing (distance between consecutive knots) to estimate bias field using spline technique from a MRI slice. Knot spacing of (a) 150 mm, (b) 100 mm, (c) 50 mm, and (d) 30 mm

Gibbs ringing

A Matlab toolbox developed by Sarra [69] includes a range of techniques from filtering to reconstruction to reduce Gibbs ringing from MRI data. One of the main

reasons for Gibbs ringing is that expansion (Fourier) coefficients decay slowly due to their global nature. Filtering techniques alter these coefficients so that they decay faster and reduce the effect of discontinuity. However, filtering does not completely remove the Gibbs phenomenon and gives blurring artefact. Reconstruction methods, which recover the image inside its different regions of smoothness, can completely remove Gibbs ringing. However they require the knowledge of the exact location of all discontinuities. Gaussian-shaped slopes at the beginning and end of the MRI volume were added to avoid unwanted Gibbs ringing in [70], while performing sinc-based interpolation to add slices in-between real slices in MRI volume.

3.4 Spatial interpolation techniques

Interpolation [73] is a technique to add new data points within the range of a discrete set of known data points. Interpolation in medical imaging is normally employed for resampling the data [73]. Soleimanifard et al. [74] used sinc interpolation [75] to resample MRI volume so that voxel size in all three spatial dimensions (x, y, z) is similar. However, sinc interpolation involves convolution with sinc kernel which is computationally expensive. Different standard interpolation options to resample MRI data were provided in [76]. Standard techniques [73] such as nearest neighbour provide jagged artefacts, linear and spline based techniques involves inherent blurring and distorted edges [70].

A different sinc interpolation technique (Fourier interpolation) which includes transitioning to the frequency domain from spatial domain and its inverse was utilised in [70], [77], [78]. In this interpolation technique [79], the original data is transformed to Fourier domain using the Fast Fourier transform (FFT). Zeros are appended to centre of the data, with the Nyquist criterion being satisfied in the case of an even sized dimension, and the inverse FFT is performed. This interpolation technique for one dimension can be represented mathematically as:

$$O_0(k) = \begin{cases} O(k), & for \quad k = 0,\ldots,\frac{N}{2} - 1 \\ O(\frac{N}{2})/2, & for \quad k = \frac{N}{2}, NL - \frac{N}{2} \\ 0, & for \quad k = \frac{N}{2} + 1,\ldots,NL - \frac{N}{2} - 1 \\ O(k + N - NL), & for \quad k = NL - \frac{N}{2} + 1,\ldots,NL - 1 \end{cases} \tag{3.3}$$

where $O(k)$ is the FFT of the input signal $o(n)$ of length N and $O_0(k)$ is the zero-padded FFT of $O(k)$ of length NL, with L denoting the interpolation factor. The interpolated signal $o_0(n)$ is obtained by taking the NL-point inverse FFT of $O_0(k)$. The $1/L$ amplitude loss made by interpolation is compensated by multiplying signal $o_0(n)$ by L [79]. Defining $O_0(k)$ as in (Eq. 3.3) maintains the complex-conjugate circular symmetry which characterises the FFT of real-valued signals. Fourier interpolation [79] has number of advantages including its accuracy compared to cubic spline and linear interpolation and the relative ease with which it can be extended to encompass multiple dimensions [80].

3.5 Overview of image segmentation techniques

Image segmentation methods [52] based on the underlying methodology can be divided into four broad categories: (1) Intensity-based, (2) Clustering and Classification, (3) Model based, and (4) Atlas-based techniques. For medical image segmentation, a slice from a medical volume, for example an MRI slice, can be seen as a 2D discrete image $I(x,y)$ with $x \in [0, N_x - 1]$ and $y \in [0, N_y - 1]$. N_x and N_y are the number of pixels in x and y dimension respectively and $I(x,y)$ present feature value (such as intensity, texture) at (x,y) spatial location. In this thesis, these segmentation techniques are explained using 2D image $I(x,y)$, however, these techniques can be easily extended into 3- or more dimensions.

3.5.1 Intensity based techniques

In intensity based methods, intensity (grey-levels) is the primary feature used to separate the region of interest.

Thresholding

Thresholding technique [54] convert a given grey scale image $I(x,y)$ into binary mask $K(x,y)$ by assigning value either 0 or 1 to the pixel depending on pixel value, whether value is smaller or greater than certain threshold Th, and can be given as:

$$K(x,y) = \begin{cases} 1, & if \quad I(x,y) > Th \\ 0, & otherwise \end{cases} \qquad 3.4$$

Threshold value Th can be determined interactively or using automated techniques.

Region growing

Region growing [54] is another intensity based method in which, given an initial region or seed, at each iteration the neighbouring pixels (4-connected component), are compared to the seed or region with respect to their similarity. If there is a high similarity, region R grows to include similar neighbouring pixels for next iteration or else the process stops. Usually, the difference between a pixel intensity value $I(x,y)$ and the region mean ϖ_R is used as a measure of similarity $\xi(x,y)$ (Eq. 3.5). The pixels with similarity value $\xi(x,y)$ less than threshold Th are assimilated in the region R (Eq. 3.6).

$$\xi(x,y) = \left| I(x,y) - \varpi_R \right| \tag{3.5}$$

$$I(x,y) = \begin{cases} \in R, & if \quad \xi(x,y) < Th \\ \notin R, & otherwise \end{cases} \tag{3.6}$$

Fuzzy analogies (fuzzy connectedness) to region growing, have also been developed [52]. In fuzzy connectedness (FC) [81], along with intensity difference, spatial distance between region and pixel (fuzzy adjacency) is also considered in terms of membership value (fuzzy affinity) for image segmentation.

3.5.2 Clustering and Classification techniques

In clustering and classification techniques, images are partitioned into regions or classes, composed of pixels that have close features values either using data directly provided from an image (clustering) or using known data labels (classification).

3.5.2.1 Clustering techniques

In clustering methods, images are clustered without the use of training data. Thus, they are termed as unsupervised methods. Common clustering algorithms are k-means and FCM. Mean-shift (MS) and spectral clustering are also employed for image segmentation.

k-means clustering

The k-means clustering algorithm [52] clusters data by iteratively computing a mean intensity for each class and segmenting the image by classifying each pixel in the class with the closest mean. In k-means, pixels in an image can belong to only one cluster.

Fuzzy c-means (FCM) clustering

The FCM algorithm [13] generalises the k-means algorithm, allowing for soft segmentations (pixels in an image to belong to more than one cluster) based on fuzzy set theory. The FCM [13] method is an iterative process that produces optimal C clusters of the greyscale image $\{I_k\}_{k=1}^{N}$ where N is number of pixels, by minimising the following objective function:

$$J_T = \sum_{i=1}^{C} \sum_{k=1}^{N} \mu_{ik}^{m} d^2(v_i, I_k) \tag{3.7}$$

where $\{v_i\}_{i=1}^{C}$ is the cluster centre of cluster i. Here C is the number of tissue types in the image. $\mu_{ik} \in [0,1]$ is the membership degree of k^{th} pixel to i^{th} cluster. m (> 1) controls the fuzziness of membership function μ_{ik}. $d(v_i, I_k)$ is the Euclidean distance between cluster centre v_i and pixel intensity I_k. The membership functions are subject to the following constraints:

$$\sum_{k=1}^{N} \mu_{ik} > 0 \quad \forall i \in \{1,...,C\}; \quad \sum_{i=1}^{C} \mu_{ik} = 1 \quad \forall k \in \{1,...,N\} \tag{3.8}$$

The membership functions μ_{ik} and cluster centres v_i are updated iteratively as follows:

$$\mu_{ik} = \frac{(d^2(v_i, I_k))^{-1/(m-1)}}{\sum_{j=1}^{C} (d^2(v_j, I_k))^{-1/(m-1)}} \quad \text{and} \quad v_i = \frac{\sum_{k=1}^{N} \mu_{ik}^{m} I_k}{\sum_{k=1}^{N} \mu_{ik}^{m}} \tag{3.9}$$

The FCM is optimised when high membership values are assigned to the pixels close to their centroid and low membership values to the pixels away from the centroid. The objective function of FCM (Eq. 3.7) is modified to include additional information to make it robust against noise and intensity variations.

Mean-shift (MS) clustering

MS clustering [82] is a non-parametric algorithm that does not require prior knowledge on the number of clusters or cluster positions. This is an iterative method which start with an initial estimate and determines the weight of nearby points for re-estimation of the mean. The method basically uses a gradient ascent technique to detect local

maxima of data density in feature space. The data points associated with same local maxima are considered as a member of the same cluster.

Spectral clustering

Spectral clustering clusters points based on eigenvectors of similarity matrices. These similarity matrices are derived from input data. Normalised cut (Ncut) [83], is a popular spectral clustering method. Ncut view an image as a graph, where image pixels represent nodes and edges between nodes are weighted according to the similarity between pixels. The aim of Ncut is to find a partition of a graph that minimises the normalised cut between two parts. The normalised cut avoids bias for partition of small set of points normally observed in graph-cut approach [83]. The partition of the image is obtained by thresholding the resultant eigenvectors.

3.5.2.2 Classification

Classification (supervised) techniques differ from clustering methods through the use of labelled training data. These techniques involves both a training phase that uses labelled data to learn a model that maps from features (such as intensity) to labels, and a testing phase that is used to assign labels to target image based on the measured features. Few classification techniques include support vector machines (SVM) [84] and markov random fields (MRF) [85].

Support vector machines (SVM)

SVM [84] is a binary classifier in which training examples are labelled to one of the two categories. Based on training algorithm a model is built that assigns new examples into one category or the other. In SVM, separate categories are divided by a clear gap (hyperplane) that is as wide as possible.

Markov random fields (MRF)

In MRF [85], image is modelled as an undirected graph where all pixels are represented as vertices and the neighbourhood interactions are modelled using edges between these vertices. The spatial interactions between neighbouring or nearby pixels are obtained in terms of joint probability distribution. These local correlations make MRF robust to noise and intensity variation. The segmentation is then obtained by maximising a posteriori probability of the segmentation given the image data using iterative methods such as iterated conditional modes or simulated annealing.

3.5.3 Model based techniques

In model-based segmentation, a model is built for a specific anatomic structure by incorporating a priori knowledge of the object such as shape, location, and orientation. Common model based techniques are deformable and statistical models.

3.5.3.1 Deformable models

Deformable models [86] delineate the region boundaries using closed parametric curves or 3D surfaces placed on image that deform under the influence of internal and external energy [86]. Internal energy is computed from the curve or surface geometry to keep it smooth throughout the deformation. External energy is usually derived from the image information such as gradient or region statistics to drive the curve or surface towards the desired feature of interest. Active contour model (ACM) and LSM are widely employed deformable models.

Active contour models (ACM) (snakes)

ACM also known as snakes introduced by Kass et al [87] represents a closed curve v in a parametric form: $\vartheta(s) = (x(s), y(s)), s \in [0,1]$. This curve is then deformed through the spatial domain of an image to minimise the following energy functional:

$$E_{snake} = \int_0^1 E_{snake}(\vartheta(s))ds = \int_0^1 (E_{int}(\vartheta(s)) + E_{image}(\vartheta(s)))ds \qquad 3.10$$

where $E_{int}(\vartheta(s))$ is the internal energy of the contour with respect to elastic deformation and bending of a contour and is given as:

$$E_{int} = \frac{1}{2}(w_1(s)|\vartheta'(s)|^2 + w_2(s)|\vartheta''(s)|^2) \qquad 3.11$$

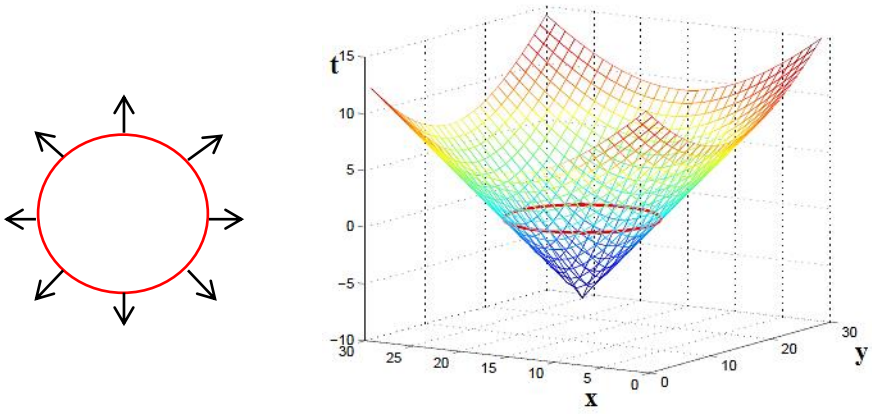
where $\vartheta'(s)$, the first derivative, makes model act like elastic string and $\vartheta''(s)$, the second derivative makes model act like a rigid rod. $w_1(s)$ and $w_2(s)$ are weighting parameters and are often chosen to be constant. The $E_{image}(\vartheta(s))$ term is the external energy that pulls ACM towards the desired features in the image such as image intensity gradient [87].

$$E_{image} = -|\nabla I(x, y)|^2 \qquad 3.12$$

During the optimisation process, the active contour is deformed with respect to the features to be located.

Level set methods (LSM)

LSM [88] evolves a curve using geometric measures, resulting in an evolution that is independent of the parameterisation. In LSM, the closed curve v, propagating with a speed F in its normal direction, is represented implicitly by a zero level set $\vartheta(t)=\{(x,y)|\phi(t,x,y)=0\}$ of higher dimensional function, known as level set function ϕ (Fig. 3.3). This implicit representation allows change of topology and is useful for shape recovery of complex anatomical structures.



(a) (b)

Fig. 3.3 : Example of embedding curve as a level set of higher dimensional function (a) a closed curve (b) the level set function (signed distance function) of a curve in (a)

Instead of tracking a curve through time, the LSM evolves a curve by updating the level set function at fixed coordinates through time t. Normally signed distance function, based on Euclidean distance, is utilised as the level set function ϕ.

$$\phi(t,x,y)=\pm d \qquad 3.13$$

where d is the signed distance (Fig. 3.3) between position (x,y) and zero level set at time t. An initial value partial differential equation can be obtained for the evolution of ϕ as [88]:

$$\frac{\partial\phi}{\partial t}=V|\nabla\phi| \qquad 3.14$$

where V is a speed function and like snake energy contains internal term that act as smoothness regulator and image information (data term) that drive ϕ to desired object boundaries.

In the implementation of LSM for segmentation the smoothness regulator does not change much and is often curvature based. The contribution comes from data driven term. Based on data driven term used in V two main categories exist for LSM: edge-based and region-based. Edge-based LSM use image gradients in order to identify object boundaries. Due to this highly localised information, edge based LSM has been found to be very sensitive to image noise and highly dependent on initial curve placement. Region based LSM model use the foreground and background regions statistics to segment the foreground. Advantages of region-based approaches compared to edge-based methods include robustness against initial curve placement and insensitivity to image noise.

3.5.3.2 Statistical models

Statistical models incorporate prior statistical information drawn from a population of training datasets. The statistical parameterisation provides global constraints and allows the model to deform only in ways implied by the training sets.

Active shape models (ASM) and active appearance models (AAM)

ASM [89] is a deformable model which uses 2D or 3D shape variation from training dataset and iteratively deforms to fit the new image using statistical information from training dataset. The constraint is based on point distribution model which is obtained by principal component analysis that provides number of constraint parameters that controls the variation found in training dataset.

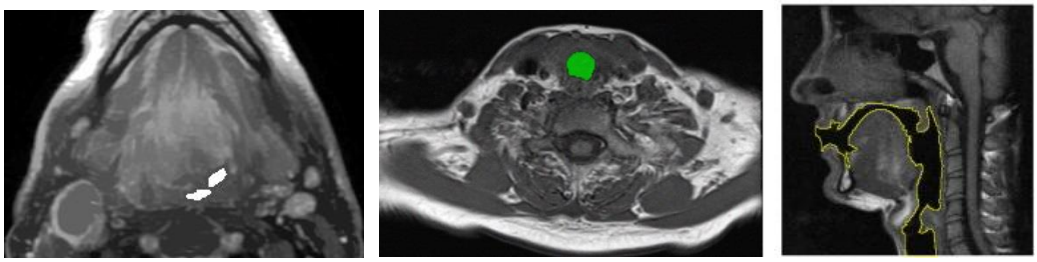
In AAM [90], instead of only shape modelling as in ASM, the image intensity values of the structures and surroundings are taken into account to build the complete model. AAM has better convergence than ASM but are much slower.

3.5.4 Atlas-based techniques

In atlas-based segmentation methods, image segmentation is regarded as a registration problem. The methods rely on the existence of a reference image (or the atlas) in which the structure of interest is labelled manually, segmenting an actual image. It is then used to find the transformation that maps the pre-segmented atlas to the study image. Therefore the key to the atlas based segmentation is to design registration method capable of computing the transformation between atlas and the image to be segmented (target image).

3.6 Throat region segmentation techniques

Segmentation of a throat region (Fig. 3.4(a)), also known as pharynx or upper airway region, is important in assessing speech production analysis [91]–[93], sleep related disorders [94]–[97], or as the guidance for computer-aided segmentation of HNC in RTP [11]. The trachea (Fig. 3.4(b)) and the vocal tract (Fig. 3.4(c)) have same characteristics, such as low intensity region, as the throat region on the MRI slices.



(a) (b) (c)

Fig. 3.4: MR image showing segmented (a) throat region (white region) [11] © 2014 IEEE (b) trachea region (green region) [95] © 2011 IEEE (c) vocal tract outline (yellow outline) [93] © 2009 IEEE.

Algorithms developed for the throat or the trachea or the vocal tract segmentation from MR images are indicated in Table 3.2. Some techniques utilised the knowledge that the throat region is the region with low intensity values present in the centre part of the image for the segmentation from 2D MRI slices. For instance, in [11] only the centre part of the image was considered. In this part, the hole filling operation was performed to make region homogenous and then the Hessian matrix of the low intensity pixels was analysed for the pixels of the throat region. The region with largest repeated eigenvalues was categorised as the throat region. This technique [11] can efficiently detect multiple throat regions on an MRI slice. However, a drawback of this technique is the detection of small throat regions (area<50mm^2) due to rapid intensity change in or near the region and thus the region, failing to satisfy the homogeneity criteria.

Semi-automatic pipeline developed in [96] utilised operator marked locations and k-means clustering algorithm for initial segmentation of a slice. Further, a rule based system was applied to the connected components of the darkest (lowest intensity) class for detection of a pharynx region. Similar approach was applied by the

same authors in [95] for trachea segmentation with further graph-cut technique to obtain the final boundary. However, k-means algorithm does not consider the partial volumes observed particularly in the boundary pixels in MRI slices and thus, underestimate the throat region. Further, these techniques [95]-[96] makes few assumptions about roundness (circularity) and area of trachea which cannot be assumed for the throat regions of various shapes and sizes.

Table 3.2: Summary of the techniques for the throat or trachea or vocal tract segmentation from MR images

Techniques	References	Description
Hessian based analysis	[11]	Analysis of low intensity pixels. Region with largest repeated eigenvalues is categorised as the throat region.
k-means clustering	[95], [96]	Image is divided into 3 classes. A rule based system is used for the analysis of connected component of the darkest class.
Region growing	[12], [94], [97]	Volume of interest or seed points specified by the operator. Region growing according to the specified similarity features.
Deformable models	[91]–[93],	Segmentation is modelled as a tracking of deformable objects. Live wire or ASM or AAM methods to tract the object.
	[77],[98]	LSM to segment the region.
Part of complex organ system	[99]	Knowledge of neighbouring structure for initialisation of snake in 2D and ASM model to extract surface in 3D

System for upper airway segmentation using fuzzy connectedness was described in [12]. This technique utilised both T1- and T2- weighted MR images to generate fuzzy affinity map for airway. An initial 3D region growing operation was performed in [97] to highlight the approximate location of the trachea. The cost function image created from this approximate information was transformed to cylindrical coordinates and graph based method was used to find the least cost surface. Another 3D region growing approach was utilised in [94] for the segmentation of

airway for assessment of obstruction during sleep from dynamic MRI. These region growing techniques can appropriately segment geometrically variable upper airway [12], [100] and trachea [97] regions with few tuning parameters, however they often generates irregular boundaries and small holes due it's sensitivity to noise.

Deformable models, due to their robustness in presence of noise and handling of topological changes, were also utilised for the upper airway or trachea segmentation. In [98] region based LSM was used for a trachea segmentation in guinea pig to evaluate the effect of drug. The region statistics required for the evolution of LSM was estimated from watershed algorithm. LSM was utilised in [77], with manual initialisation, for throat region segmentation from MRI volume. For speech analysis, upper airway or vocal tract segmentation [91]–[93] was modelled as a tracking of deformable objects in real time sagittal MRI sequences. In [93], variables were defined to describe the shape of the vocal tract and an unsupervised regional segmentation technique was adopted to track the contours of the variables. However this method [93] requires considerable supervision to initialise the variables of the vocal tract. Z. Raeesy et al. [91] used statistically modelling of shapes (ASM) for the segmentation of vocal tract. Initially, in training phase, shape based models of vocal tracts were constructed using shape statistics from labelled landmarks. In the testing phase, live wire contour tracking approach was employed for the recognition and segmentation of vocal tract on target image based on defined shape models. Similarly, AAM model was employed in [92] for vocal tract segmentation. Compared to ASM, AAM consider all the available data instead of just the landmark neighbourhoods and thus is more robust than ASM. However, ASM and AAM techniques are suitable for segmentation of a upper airway with known shape or appearance, which is not the case in presence of adjacent cancer regions near the airway.

In other automatic system [99], the trachea was segmented as a part of a complex organ system segmentation from sagittal MR image sequence. In, [99] the initial landmarks to identify the trachea in 2D were obtained from the neighbouring structures. The obtained landmarks were used as an initialisation for active contours to identify the borders between the trachea, the spinal column and the spinal cord. This information was then propagated to the 3D segmentation module implemented with ASMs to efficiently extract the 3D surface contours of the desired regions. This

approach efficiently combines segmentation algorithms, object recognition and anatomical knowledge to segment and reconstruct structures of cervical spine (trachea, the spinal column and the spinal cord). However, the assumptions made about the anatomy in the neighbourhood make this approach infeasible to segment single structure such as trachea.

3.7 Cancer segmentation techniques

A wide variety of tumour segmentation techniques from MRI has been proposed with review in [2], [100]. As described in Section 3.4, cancer segmentation methods can also generally be categorised into intensity based, clustering and classification, model based and atlas based techniques. Although each technique is described separately, multiple techniques are often employed to improve the segmentation results.

3.7.1 Intensity based techniques for cancer segmentation

The most common feature used for the cancer region segmentation is the image intensities. This is based on the assumption that tumour tissues have different grey-levels compared to normal tissues. For instance, intensity based thresholding was applied to a manually selected region of interest in [101] to obtain initial approximate of tumour pixels. This initial estimate was further refined using morphological operations and region growing technique to segment brain tumour from T1W post-contrast MR images. Other intensity based thresholding technique for breast mass lesion segmentation from dynamic MRI was proposed in [102]. In this technique, threshold was estimated by averaging the image intensity at the boundary candidate points which are obtained using ray casting technique. Further, connected component analysis and morphological opening operation were performed in [102] to avoid leakage through weak boundaries. Generally, intensity based thresholding methods do not consider any relationships between the pixels and thus, are unable to exploit all the information provided by an image.

Region growing variants, seed growing and region deformation techniques, were used in [103] for tongue cancer segmentation from T2W MRI. In [103], the region was grown or deformed from manual initialisation based on the similarity of the neighbouring pixels. These techniques correctly segment spatially separated cancer regions having the same properties. The manual initialisation in [103] was removed in

[104] by automatic seed selection using the characteristics of the cancer region and edge image. This seed was used as an initialisation for improved fuzzy connectedness to segment brain cancer. The fuzzy connectedness was improved by adding gradient information and using substituted pixel intensity in fuzzy affinity to avoid leakage through weak boundaries. This technique [104] is independent of cancer pixel intensities but has no spatial constraints. Further, thresholding and region growing techniques are useful for segmenting clearly defined cancer structures, but these algorithms commonly fail on cancer regions with fuzzy boundaries and overlapping intensities, as partial voluming and pixels of similar intensity belonging to different regions are not considered in algorithms' mathematics.

3.7.2 Clustering and classification techniques for cancer segmentation

Tumour can be considered as an additional cluster or class whose features such as intensity, textures are different from healthy tissue clusters or classes. Thus, clustering and classification segmentation techniques aim to separate this tumour class and process it further to obtain the final segmentation results. Table 3.3 shows the summary of the clustering and classification techniques for cancer segmentation described in this section.

3.7.2.1 Clustering techniques for cancer segmentation

k-means is simple clustering algorithm with relatively low complexity and the application is more practical even with large dataset. Thus, Moftah et al. [105] proposed the use of k-means algorithm to segment MR image with tumour. After separate cluster regions were found using k-means, the tumour was recognised as the brightest and circular object from the cluster regions. Compared to k-means, FCM clustering techniques are prevalent for the cancer segmentation task, due to their robustness and efficiency. FCM clustering also considers PVE in terms of membership function in a segmentation task and the objective function of FCM clustering can be modified [107]-[108] to include any suitable information or prior information can be considered [109]-[110] in FCM to increase the robustness of the technique for cancer segmentation with fuzzy boundaries in presence of noise and IIH. Further, FCM can be easily combined with other techniques such as knowledge based techniques [106] or ACM [111] to obtain the final segmentation boundary.

Table 3.3: Summary of clustering and classification techniques for tumour segmentation

	References	Description
Clustering techniques		
1) k-means	[105]	Initial k-means clustering. Tumour recognised as brightest and circular object from the cluster regions
2) fuzzy c-means (FCM)	[106]	Multi-spectral FCM. Output from FCM to knowledge based system. Histogram analysis for cancer separation.
	[107],[108]	Additional neighbourhood information in FCM to reduce its sensitivity to noise and IIH
	[109]	Available prior information from cluster centre in FCM to consider variation in data acquisition
	[110]	Genetic-induced fuzzy clustering for robust initialisation of FCM clustering
	[111]	Combine FCM and possibilistic c-means to reduce effect of noise and to avoid detection of overlapping clusters
3) mean-shift	[112]	Initial multi-spectral mean-shift clustering. Detection of cluster with suspicious lesion using graph-cut
4) spectral (graph-cut)	[113]	Interactive graph-cut. Minimisation of energy formed from region statistics using graph-cut.
	[114]	Tumour-cut – combining tumour segmentation with cellular automata
Classification techniques	[115]	Two class (tumour and non-tumour) SVM trained on intensity feature.
Support vector machines (SVM)	[116]	One class (tumour) SVM trained on intensity feature.
Markov random fields (MRF)	[117]	Multi-channel MRF reduced to single channel MRF and trained using kinematic features.

Clark et al. [106] used multispectral (T1W, T2W, Proton Density) FCM clustering to obtain initial clusters from MR image with glioblastoma-multiforme tumours. The clustered image along with class centre was further given to a rule-based

system. Multispectral histogram analysis was exploited to separate suspected cancer region from the rest of the region. Additional power and flexibility for the system was achieved with use of knowledge of nearby anatomical structures and few (17 slices) training examples. This system can, however, only process axial slices starting from an initial slice 7 to 8 cm from the top of the brain and upward.

FCM clustering was used in [107] to segment brain cancer image. In [107], FCM technique was modified by adding neighbourhood spatial information in the objective function of FCM to simultaneously correct IIH and segment a brain cancer image. This neighbourhood information influence the membership function of pixel under consideration, thus obtaining a piecewise homogenous segmentation in presence of noise and IIH. FCM technique was also modified in [108] by considering neighboured information for the segmentation of brain tumour. The initial clusters for FCM were obtained using histogram based technique. The obtained clusters were further processed to correctly classify ambiguous pixels and to obtain shaper boundaries between different tissue classes. These techniques [107], [108], however, works only on MR images corrupted with salt and pepper noise and cannot compensate for severe IIH.

Prior information on cluster centres was considered in FCM in [109] to segment the tumour in the presence of noise and variation in data acquisition. This technique [109] is good only if cluster centres information is known prior. Zhou et al. [110] proposed algorithm that use FCM for the segmentation of tongue carcinoma and then refined results using artificial neural network (ANN). Genetic algorithm based automated classification was performed in [110] initially to obtain inputs for FCM clustering. Features selected for ANN such as symmetric measure, difference of masses, and distance from centre make this technique suitable only for that particular type of tumours.

The segmentation pipeline (Fig. 3.5) which combines clustering and deformable model for a tumour segmentation was proposed in [111]. The initial tumour is detected using a clustering technique which combines FCM and possibilistic c-means algorithms. Thus, this technique is robust to noise and does not detect the overlapping clusters. Parametric ACM constrained by edge and spatial information of the initial tumour was deformed on the detect cluster to obtain the final result.

However, FCM clustering requires initial knowledge on the number of clusters and cluster centres as it detect local minima.

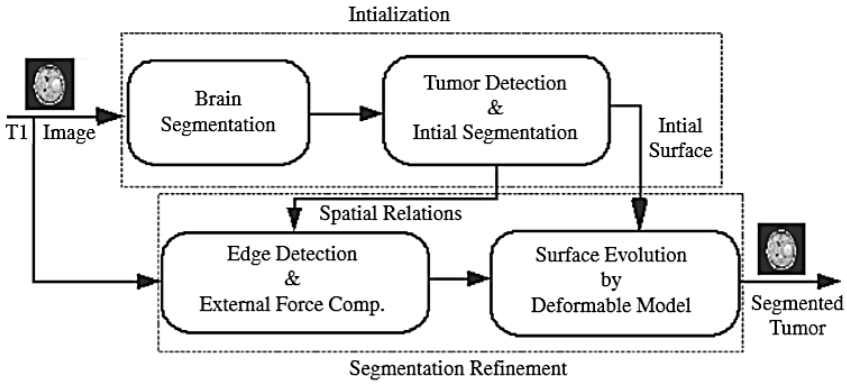


Fig. 3.5: Segmentation pipeline for brain tumour proposed in [111]. Segmentation pipeline combines clustering and deformable techniques to obtain accurate tumour outline. Copyright © 2009 Elsevier (License number: 3831990049366)

MS clustering is an alternative clustering technique for cancer segmentation that produces spatially contiguous regions and does not require the number of clusters be set a priori. This clustering was employed by McClymont et al. [112] as an initial step for clustering similar voxels from multi-spectral MRI data (T1W, T2W and dynamic). Further, graph-cuts segmentation was applied to the region adjacency graph defined on these clusters to identify sets of clusters that are likely to be suspicious tissue for breast lesion segmentation. However, MS clustering usually generates a greater number of clusters than tissue types and with different number of clusters generated for each MRI slice, selecting appropriate cluster with cancer region is difficult.

Spectral clustering (graph-cut), that require no assumptions on the form of clusters, was presented in [113] for tumour segmentation. The region statistics was calculated from the user provided initial mask. The energy function created from this region statistics was minimised using graph-cut. A variation of graph-cut, tumour-cut, was introduced in [114], which combines the tumour segmentation with cellular automata. The tumour probability map in [114] was obtained by assigning each voxel a tumour label if there was a shorter path from that voxel to a seed (supplied by user) than to any background seed, where paths were weighted by image content. The spatial smoothness was ensured by evolving level set on the tumour probability map.

Although spectral clustering (graph-cut tumour-cut), are robust to detect any arbitrarily shape clusters, these methods are computationally expensive for large dataset, can be quite unstable under different choices for parameters for neighbourhood graphs and does not always guarantee convergence.

3.7.2.2 Classification techniques for cancer segmentation

Supervised classification approach such as SVM, MRF are often used for tumour segmentation when training data is available. SVM is robust for cancer segmentation as it does not depend on specific assumptions about the distribution of the data and can perform well when applied to data outside training dataset. Zhou et al. [115] trained two-class (tumour and non-tumour) SVM on features (intensity) extracted from user drawn rectangular boxes and used this training data to segment nasopharyngeal tumour from MR images. Further, Zhang et al. [116] from the same group demonstrated that one-class (tumour) SVM performs better than two-class SVM for tumour segmentation as it removes the assumption of equally balanced data-classes in case of severely under sampled data. However, SVM techniques do not consider labels of adjacent pixels and assume that each pixel has the same probability distribution as the others and is mutually independent. These assumptions are not appropriate for tumour segmentation where neighbouring pixels usually belong to same class.

MRF technique, that can capture features that are particularly important for MR images such as non-parametric distribution of tissue intensities and spatial interaction between neighbouring pixels, thus, can overcome spatial inhomogeneity and noise, is used in [117] for breast tumour segmentation using kinematic features. In this [117], it is demonstrated that any multi-channel MRF can be reduced to a single channel MRF provided the features in different channels are conditionally independent. The output of MRF model was a probability estimate that the super-pixel (group of pixels) belongs to a tumour region. However, a limitation in MRF is the selection of the parameters that control the strength of spatial interactions, which can result in very soft segmentation and a loss of structural details. Further, MRF needs to be combined with other techniques such as Expectation maximisation for the segmentation task.

3.7.3 Model based techniques for cancer segmentation

Table 3.4 shows the summary of model (deformable and statistical) based segmentation methods. Model based techniques can integrate prior information from MRI data in terms of energy, shape or appearance for the segmentation of tumour region and thus are robust than intensity based methods.

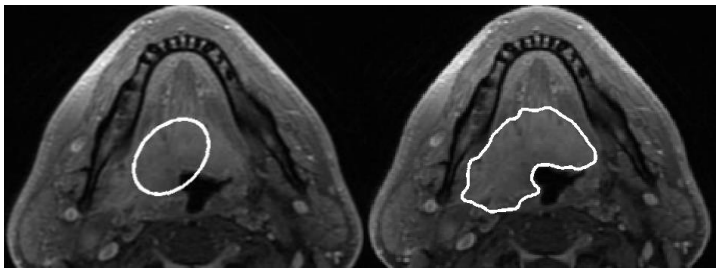
Table 3.4: Summary of model based tumour segmentation methods

Model based techniques	References	Description
Deformable models		
1) active contour algorithms (snake)	[118]	Initialisation of snake: thresholding and morphological operations. Energy from gradient of an image.
	[119]	Energy is generated dynamically to drive the contour evolution.
	[120]	Energy from both edge and region information
2) level set methods (LSM)	[121]	Energy from boundary and region information
	[122]	Local region information around each point in the contour for level set evolution
	[123]	Level set evolution as region competition. Tumour probability map as initialisation for LSM
	[124]	Signed pressure function formulated using local statistics in LSM.
	[125]	Generation of template and atlases using appearance of anatomical structures. These templates and atlas in variational framework.
Statistical models		
Active shape model (ASM)	[126]	Shape boundaries parameterised as radial time series. Gaussian process to create model of the shape which is used as regularisation in dynamic active contours.

Snakes (ACM), a parametric deformable model, due its ability to independently and adaptively search for the object boundary using underlying image information is suitable for tumour segmentation with weak and unclear edges. Zhu, et al. [118] used an ACM (snakes) to detect a tumour boundary in the brain MRI and applied a Hopfield

neural network to speed up computation efficiency. For initialisation of ACM in [118], rough estimate of tumour was obtained using thresholding and morphological operations on a manually selected slice. The image energy in ACM was dependent on the gradient of an image. The tumour boundary obtained on selected slice was subsequently interpolated to next slice and used as initialisation to obtain tumour boundary on that slice. The main disadvantage of standard snake (ACM) [90] is that they cannot deform easily into boundary concavities. Thus, ACM was improved in [119] for concave shape brain tumours segmentation. In [119], external (image) energy was generated dynamically to drive the contour evolution. ACM was also applied on generated parametric eigenvector image for breast lesion segmentation from dynamic MRI [120]. Spectral embedding, a nonlinear dimensionality reduction scheme, was applied in [120] to the dynamic MRI time series in a voxel wise fashion to reduce several time point images to a single parametric image where every voxel is characterised by the three dominant eigenvectors. The energy function for ACM in [120] was derived from both edge (gradient) and region information (hybrid ACM). Constraints of ACM to handle multiple objects and complex geometry limits its applicability for the segmentation of tumour regions with varied topology.

Thus, to segment multiple tumour regions or tumour regions with oedema from MRI slices, LSM technique was used in [121]-[125] for cancer segmentation. As the curve is embedded in higher dimensional function for evolution in LSM, LSM technique can handle splitting or merging of a cancer region. Further, integration of prior information such as gradient and/or region statistics in terms of energy function in LSM allow segmentation of inhomogeneous cancer regions with irregular boundaries. Xie et al. [121] employed hybrid energy function in level set based technique for segmentation of brain tumours. In [121], region information, which was based on homogeneity criteria, served as a propagation force, and boundary information served as stopping functional. Localised region based LSM [127] which used local region information around each point in the contour was used for tongue cancer segmentation from MRI in [122]. In both approaches [121], [122], the user selected ROI was used as an initial zero level set for level set evolution (Fig. 3.6).



(a) (b)

Fig. 3.6: Tumour region segmentation using LSM (a) manual initialisation for LSM (ellipse with white outline) (b) final tumour boundary obtained after level set evolution of ellipse (white outline)

Ho et al. [123] associated level set evolution with region competition. The algorithm [123] used two images (T1W with and without contrast agents) and calculated a tumour probability map using classification, histogram analysis and the difference between the two images, and then this map was used as the zero level of the level set evolution. Thapaliya at el. [124] introduced a new signed pressure function in level set evolution to stop evolution at weak and blurred edges of the tumour. This pressure function was formulated using local statistics information from the object. This technique [124] works on MR images with good contrast between tumour and background. In [125], prior knowledge of the appearance of anatomical structures in the normal brain, in the form of templates and atlases was used. These templates and atlases were then fully integrated into a level set variational segmentation for the tumour. However, the major challenge with the LSM is to find an initial tumour region or boundary which is not too far from the desired tumour features, otherwise LSM may deform to an incorrect local region.

Statistical shape models, which are robust for segmentation of tumour region with known shape as the mean shape and variations in shapes are learned from available training data, was proposed in [126] for brain tumour segmentation. In [126], shape boundaries were parameterised as radial time series. This time series was combined with Gaussian process to develop regularisation term which was used in dynamic contour models for lesion segmentation. The above work [126] used shape information based on the assumption that lesions have similar shapes in different images. However, for most of the cancer region, shape assumption does not hold. Further, shape modelling requires large training dataset to cover all variation in

geometry and robust shape initialisation in a new MRI slice can be challenging even with user interaction.

3.7.4 Atlas-based techniques for cancer segmentation

In atlas based tumour segmentation techniques, atlas is generally created from healthy subjects and tumour segmentation is modelled as a lesion growth [128], [129] or a tumour in a target image is either considered as an outlier [130], [131] compared to healthy atlas. In [128], [129], the registered image (obtain by registration between atlas and target image) was seeded manually by selecting a voxel of tumour regions. Then a non-rigid deformation method at seeded voxel was performed with two forces: the demons force outside the tumour and a prior model of tumour growth inside it. The radial growth model was considered in [128], while, in [129], hierarchical displacement model was considered for tumour growth. However, this growth models are appropriate for certain types of tumours only (for example circular tumours).

In [130], an initial registration was used to align the atlas of healthy brain image to the target image containing brain tumour. Then robust estimator was employed to estimate the tissue probability density functions using healthy samples from atlas. Outliers to this estimation were considered as tumour tissues. A similar approach was utilised in [131] as shown in Fig. 3.7. However, to increase robustness, multi-atlas approach was used instead of single atlas. In multi-atlas label propagation [131], first manually annotated atlases were individually aligned with the un-segmented target image. The propagated segmentations were then used to calculate per-voxel density functions. Next, these per-voxel density functions and the target information were used to estimate the observed and expected density functions. These two density functions were then used to construct a voxel-wise estimate of the out-of-atlas likelihood. Lastly, the background and edge effects were reduced through a post-processing smoothing step. In these approaches [130], [131] large variations of brain structures can lead to incorrect outlier detections. Furthermore, the deformable registration of the atlas to target images with tumour is still an extremely challenging task because of the intensity variations around the tumour caused by dead tissue and the deformations of healthy tissue morphology caused by the tumour mass effect.

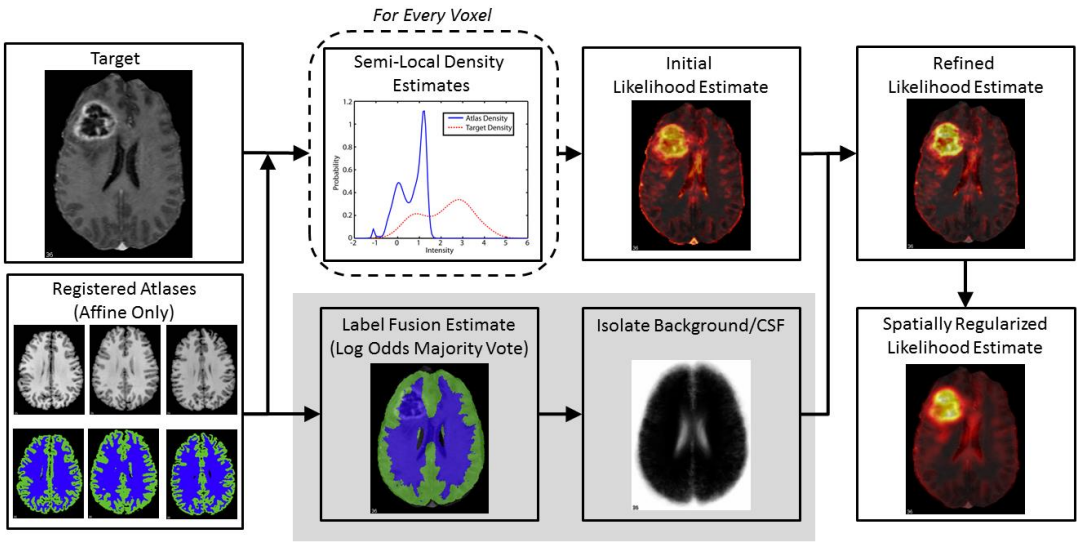
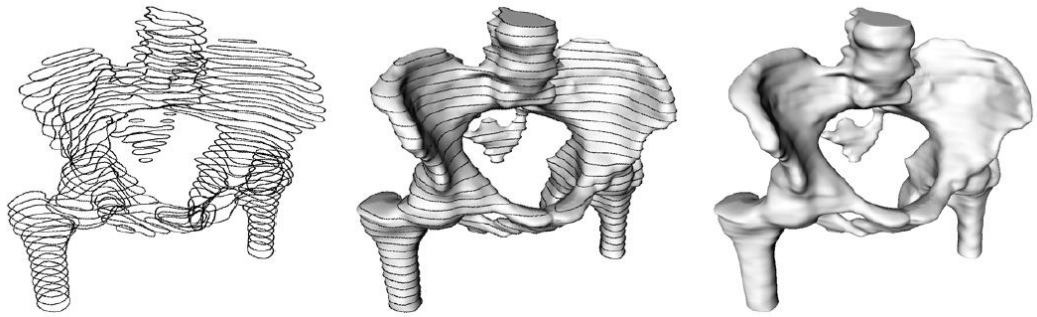


Fig. 3.7: Atlas based segmentation procedure for brain tumour from T1W MR image [131] © 2012 IEEE.

3.8 Three-dimensional (3D) reconstruction and volume calculation

3D reconstructed cancer regions is useful for visualisation of pathological characteristics that are not easily visible from 2D slices, to study morphological changes in the region and to appreciate spatial relationship of tumour with respect to other anatomical structures for RTP and for surgical planning [132].



(a) (b) (c)

Fig. 3.8: 3D reconstruction from 2D contours of pelvic region [133] (a) 2D contours (b) reconstructed surface with contours overlaid (c) 3D reconstructed surface. Copyright © by The Eurographics Association 2006.

3D reconstruction is a process of creating a 3D surface from a set of contours that have been manually or automatically extracted from 2D images (Fig. 3.8). Common technique for contour-based 3D reconstruction is to connect (stitch) the vertices of neighbouring contours into a mesh [134], [135], [136]. In [134], contour

reconstruction was performed by stitching successive contours with a triangular mesh. A mesh contains straight lines and flat surfaces as it is defined by a set of vertices and faces. The resolution of the mesh is limited by the number and density of input points. The continuity in generated mesh in [134] was accomplished by smoothing the mesh. Marching cubes were employed in [135] for the mesh generation between the contours for 3D reconstruction of the brain tumour. The marching cubes algorithm operates on logical cube created from eight pixels; four pixels from two adjacent slices. It processes one cube at a time and determines how the surface intersects each cube using the isovalue of the surface and cube-isosurface intersection patterns. This marching cube algorithm results in a large number of triangles and thus requires mesh simplification algorithm [135]. In [136], meshes were reconstructed using Delaunay triangulation which generate much less triangles than marching cubes. Further, Laplacian smoothing was applied to smooth noisy surface. However, this approach fails to deal effectively with some contour meshing problems such as contour pairs that are either too different from each other, or that overlap [133].

The volumetric approach for surface reconstruction was proposed in [137]. In this implicit method, a 2D distance field for each contour was created using a signed-distance function. These 2D fields were stacked and interpolated in the z-direction with cubic B-splines. The reconstructed surface was extracted from the resulting volume as the zero isosurface. Other implicit surface reconstruction pipeline was proposed in [133]. In [133], initially input binary contours were smooth with multi-level partition of unity (MPU) method. In MPU method, a set of overlapping patches is fitted to the points and then these patches are combined together in order to produce a smooth implicit representation of the contour. A 2D distance field was calculated around the smooth contour and then was filtered to remove the medial axis discontinuities. These 2D fields were then interpolated with splines to produce a 3D field. This spline interpolation allows smooth blending between widely spaced input contours. The reconstructed surface was the zero level set of the 3D field. The advantage of these implicit methods is that they can easily deal with reconstruction of contours even in presence of multiple contours on the same slice or in contours that overlap (Fig. 3.8(a)).

Along with 3D reconstructed cancer region, quantitative value such as tumour volume of the reconstructed region is also useful in RTP and surgery planning [138]. Diameter-based measurement [139] for tumour volume is an established method in current clinical practice as it is easy and fast to perform in a busy clinical practice as does not require outline of the tumour region [139]. In this method, three largest orthogonal diameters of the tumour are obtained and the volume T_{vol} is calculated using ellipsoid volume formula as:

$$T_{vol} = \frac{\pi}{6} D_1 D_2 D_3 \tag{3.15}$$

where D_1, D_2 and D_3 are three diameters. The disadvantage of this method is that the measurement may be less accurate because tumour shape is of irregular configurations that deviate from idealised ellipsoid volume.

Sorensen et al. [140] proposed a perimeter based method in which a volume is calculated by counting the number of voxels in tumour region and multiplying the count by the volume of a single voxel and compared this method with diameter-based method [139] on MRI slices. In [140] tumour region was outlined manually. The correlation analysis between these methods in [140] suggests that diameter method underestimate the tumour volume compared to perimeter based method and perimeter method reduce inter- and intra- reader volume variation. Diameter method was also compared with ROI-based method in [141] for the measurement of primary tumour volume for cervical cancer from MRI. In ROI based measurement, volume was calculated by summing of all tumour areas in each slice and multiplying area by the slice thickness plus spacing in between slices. There preliminary results suggested that, for small (< 40 cm^3) and large (≥ 100 cm^3) tumour size, initial tumour volume can be estimated by diameter-based measurement. However for (40-99 cm^3) intermediate size range, ROI-based method with MRI during RT is needed to quantitatively analyse tumour regression rate for the prediction of treatment outcome.

3.9 Conclusion

This chapter reviewed the state-of-art MRI based image pre-processing techniques to reduce artefacts such as IIH, throat region detection techniques, cancer segmentation, 3D reconstruction and volume calculation techniques. Based on review in Section 3.3

for IIH reduction, technique that estimate IIH using bicubic spline and reduce using local entropy minimisation have shown to effectively reduce the complex bias field. However the important parameter, knot spacing, in this technique is set to a constant value for all MRI slices, which reduce its effectiveness in the presence of geometrically variable cancer regions on MRI slices.

The review for throat detection techniques in Section 3.6 have revealed that region growing and LSM can successfully segment the throat in case of topological changes (splitting and merging), but these algorithms require some form of user interaction from an expert for its initialisation. Further, statistical models such as ASM and AAM can obtain promising results across a wide range of throat shapes. However, most statistical methods bias the solution towards a model of shapes and appearances learned a priori which is not reliable in presence of adjacent cancer regions. It may therefore be advantageous to develop a new throat detection algorithm that does not depend on shape, appearance or user interaction.

The review of cancer segmentation techniques in Section 3.7 have shown that unsupervised methods such as FCM, MS and spectral clustering are preferred compared to supervised methods (SVM and MRF) for the dataset with substantial variations in the cancer shape and intensity, as appropriate models of the cancer region are hard to learn from a finite training set. Further in clustering techniques, intensity or texture based objective measure can handle more complicated cases of fuzzy boundaries and overlapping intensities of a cancer region. However, these techniques do not obtain accurate outline of a cancer region as they are pixel based as opposed to boundary based techniques and do not utilise known spatial extent of the expected cancer region for a robust segmentation. Deformable models such as active contours (snakes) and LSM have shown to obtain accurate outline of a cancer region by incorporating any image information in terms of energy function and geometry constraints on the solution. The advantages of LSM over snakes are implicit handling of topological changes and easy extension to higher dimension. The major challenge with LSM is to find an initial model for correct convergence to the desired boundaries.

Further from the review in Section 3.8 it is clear that mesh generation technique for 3D reconstruction is appropriate for cancer regions with no overlapping contours and ROI-based volume calculation method is suitable to quantitatively analyse cancer

region during RT treatment. Finally, integrating these techniques (pre-processing, throat detection and cancer segmentation and quantification) into a reliable framework to segment and quantify ill-defined oropharynx and larynx cancer regions from MRI slices, that are obtained using different imaging protocols from different MRI scanners (Chapter 4), with no user interaction, is a non-trivial (difficult) task. However, this framework would be a valuable tool for clinicians treating cancer patients to speed up and improve RTP.

Chapter 4

MRI dataset and evaluation parameters

4.1 Introduction

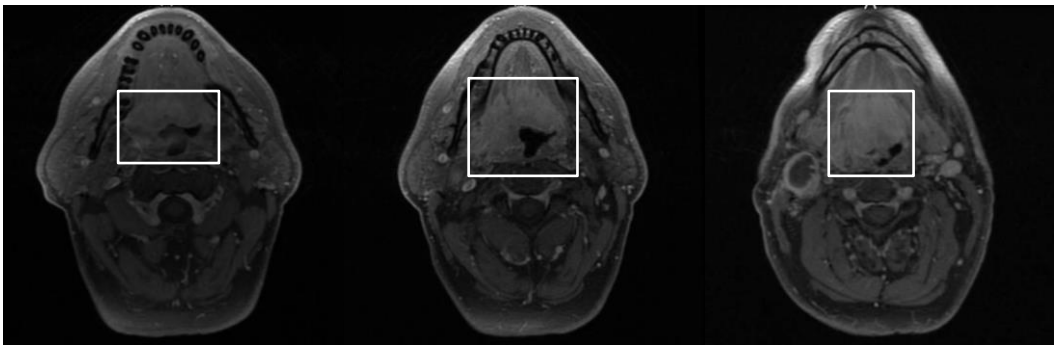
The previous chapter reviewed the current computer-aided MRI tumour segmentation and quantification techniques. The main objective of our research is to develop a system to automatically segment and quantify oropharynx and larynx tumours (GTV) from MRI dataset which will assist a CO and it is easy to use in clinic. At present there exists no publicly available collection of the MRI data that might serve as a suitable database for this research. A key factor in the success of this research is the development and testing of the algorithms on real MRI data from patients with oropharynx or larynx tumour present. This real MRI data were collected at the Beatson West of Scotland Cancer Centre, Gartnavel General Hospital, Glasgow. The collection and detail analysis of this real data is presented in the first section (Section 4.2) of this chapter. The developed algorithms should be compared with the current practice used in the clinic (manual segmentation) to validate the success of the algorithms [9]. Section 4.3 explains the procedure used in this research to obtain the manual segmentation results.

Synthetic MRI data constitute a valuable tool to find out about the algorithms behaviour under the controlled circumstances [142]. The synthetic MRI dataset used to analyse the behaviour of the algorithms developed in this research is presented in Section 4.4.

Further, in any algorithm development it is important to assess the capability of the proposed algorithm with some standard evaluation testing. Section 4.5 introduces and justifies quantitative (evaluation) parameters used to assess the performance of the proposed algorithms.

4.2 Real dataset and ethics statement

T1+Gd MRI data was obtained from The Beatson West of Scotland Cancer Centre. The data used in this research is the standard contiguous axial (short axis) MRI slices (Fig. 4.1).



Patient 1 Patient 1 Patient 1

Fig. 4.1: Typical real axial (short-axis) T1+Gd MRI slices from a single patient with oropharynx tumour (white outline). Dissimilar intensity (grey-level) values for same tissue type (for example: tumourous tissue) can be observed due to non-uniform coil sensitivity.

Ten patients previously treated, from 2010 to 2013, with radical chemo-radiotherapy for Stage II/III head-and-neck squamous cell carcinoma (6 oropharynx and 4 larynx) were randomly selected and included. Typical information, such as age, imaging protocols, number of axial MRI slices with tumour and quality of slices, for each patient is shown in Table 4.1. The mean age of patients was 56 years (range 52-72 years). The MRI scans for these patients were obtained from three different 1.5Tesla (1.5T) MRI scanners namely Magnetom Avanto from Siemens, Intera Neuro coils from Philips Medical Systems, and Signa HDxt from GE Medical Systems. The MRI scans were performed using a dedicated head and neck coils with the patient lying in the supine position. The MRI scans were undertaken pre- and post- intravenous gadolinium (Gd) contrast and were acquired after 15-20 minutes of intravenous injection of 0.1ml/kg. The range of other imaging parameters were 3-5mm slice thickness, 3.3-6 mm spacing in between slices, 9.06-20ms echo time, 542-1066ms repetition time, 90^{o}-150^{o} flip angle, 0.43x0.43-0.94x0.94mm^2 in-plane resolution, 256x256-512x512 acquisition matrix, 97.65-221 Hz/pixel bandwidth. We note that this imaging parameters were not tuned for algorithm development and testing and are typical of those used for clinical work. All images were stored in the standard digital imaging and communications in medicine (DICOM) format. Only DICOM images with visible tumour regions (GTV) were processed by the proposed algorithms. Of the 10 patients assessed, there was a total of 102 axial slices containing tumours with range of 7-17 slices in each patient (Table 4.1).

The quality of MRI slices with tumour for each patient was assessed using signal to noise ratio (SNR) and contrast to noise ratio (CNR) (Table 4.1). To calculate SNR and CNR, for each patient, a region of interest (ROI) was manually selected defining pixels of relatively pure healthy tissues (oropharynx or larynx), tumourous tissues (oropharynx or larynx) and background (air). In choosing the background ROI, the regions that might contain motion artifacts (ghosting and ringing) were avoided. SNR and CNR were calculated as follows [143]:

$$SNR = \frac{\varpi_{Hl}}{\sigma_{bg}} \quad \text{and} \quad CNR = \frac{\varpi_{T} - \varpi_{Hl}}{\sigma_{bg}} \tag{4.1}$$

where ϖ_{Hl} is the mean intensity value of the healthy tissues, ϖ_{T} is the mean intensity value of tumorous tissues and σ_{bg} is the standard deviation value of the background. The CNR values for this dataset are wide-ranging (Table 4.1) with lowest value of 9.60 for Patient 6 and highest value of 33.64 dB for Patient 5. The SNR values for this dataset are standard with minimum value of 27.32 dB and maximum value of 37.36 dB (Table 4.1) even though different imaging parameters are used to obtain MRI slices.

Typical T1+Gd MRI axial slices from different patients with tumour are shown in Fig 4.2. T1+Gd MRI compared to unenhanced (normal) T1, proton-density and T2 MRI is superior to define tumour spread for BoT and larynx tumours [144], [145], as it significantly improves soft tissue contrast and tumour margin definition. However, even with T1+Gd MRI (Fig. 4.2), oropharynx and larynx tumours segmentation is a challenging task due to variability in its geometry (Patient 6: small tumour region and Patient 2: large tumour region), presence of necrotic tissues in the region, overlap of feature values of the tumour and non-tumour pixels (Fig. 4.2(a)), no distinct boundaries between tumourous and healthy tissues (Fig. 4.2(b)) and the presence of MRI artefacts. Further, significant inter- (Fig. 4.2) and intra- (Fig. 4.1) intensity variations across the patients with no predictable pattern and low frequency intensity variation (IIH) across the slice (Fig. 4.2) complicates the segmentation of a tumour region. Furthermore, inconsistency in the throat regions (low intensity region in the centre of the slice), with regard to shapes and sizes, between different patients require additional processing for detection of the throat regions.

Table 4.1: Typical information of the 10 patients used in this research. Patient No. was randomly assigned to the patient for this research project. Different 1.5 Tesla (T) MRI scanners with different imaging protocols used to obtain MRI slices and low values of CNR, particularly for Patient 4 and 6, can be noted from the Table.

Patient No.	**Gender**	**Age** (years)	**Type of Cancer**	**1.5T MRI scanner**	**No. of slices with tumour**	**Slice thickness** (mm)	**Spacing between slices** (mm)	**TE/TR** (seconds)	**Flip Angle** (degree)	**Pixel spacing** (mm)	**Matrix size** (pixels x pixels)	**Band-width** (Hz/pixel)	**CNR** (dB)	**SNR** (dB)
1	M	52	BoT	SMA	8	4	4.4	18/995	150	0.94x0.94	256x256	220	26.89	37.36
2	M	48	BoT	Intera	12	5	6	14/640	90	0.43x0.43	512x512	195	23.28	35.43
3	M	72	BoT	GE	7	3	4	9.33/1021	90	0.47x0.47	512x512	97.65	17.24	27.42
4	M	53	BoT	GE	12	3	4	9.33/766	90	0.47x0.47	512x512	97.65	13.24	35.94
5	F	55	BoT	GE	7	3	4	9.33/766	90	0.47x0.47	512x512	97.65	33.64	37.27
6	M	49	BoT	GE	9	3.5	4.5	9.17/640	90	0.47x0.47	512x512	97.65	09.60	37.74
7	F	66	Larynx	Intera	12	3	4	14/640	90	0.43x0.43	512x512	195	19.88	29.75
8	F	66	Larynx	SMA	7	3	3.3	20/750	150	0.63x0.63	384x384	221	29.23	31.33
9	F	64	Larynx	SMA	17	3	3.3	17/542	150	0.78x0.78	256x256	201	21.07	27.64
10	M	54	Larynx	GE	11	3	4	9.06/1066	90	0.47x0.47	512x512	97.65	32.36	30.87

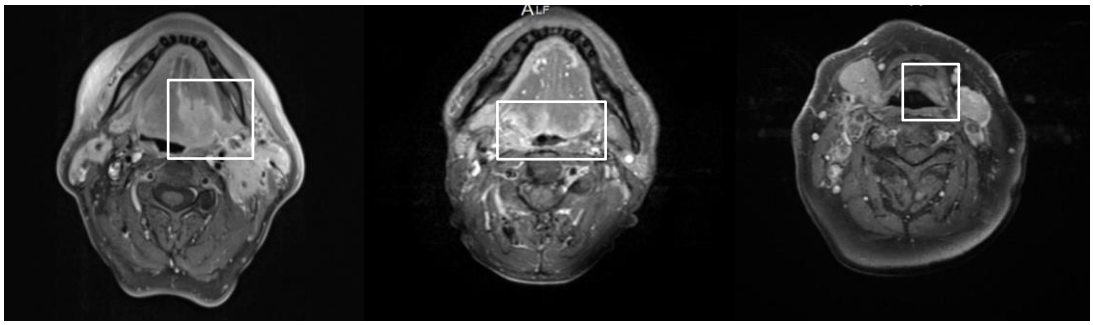
M - Male BoT – Base of Tongue SMA - Siemens Magnetom Avanto (Erlangen Germany) TE – Echo Time CNR – Contrast to noise ratio

F - Female

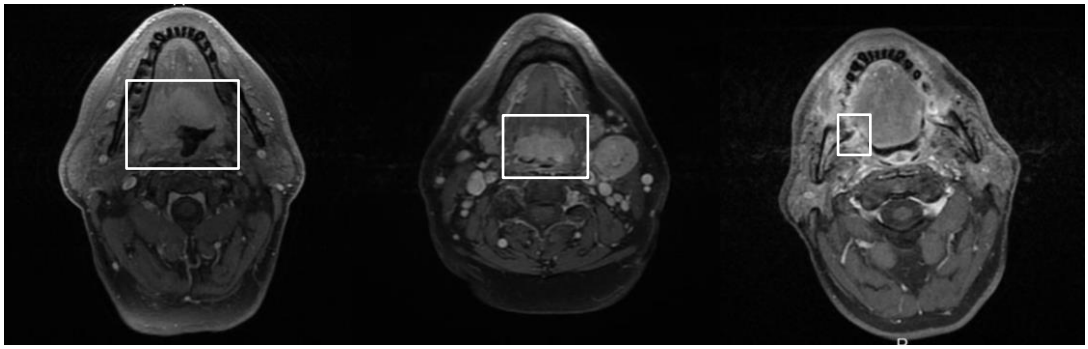
Intera Neuro coils – Philips Medical Systems

GE - Signa HDxt 1.5-Tesla scanner - GE Medical Systems

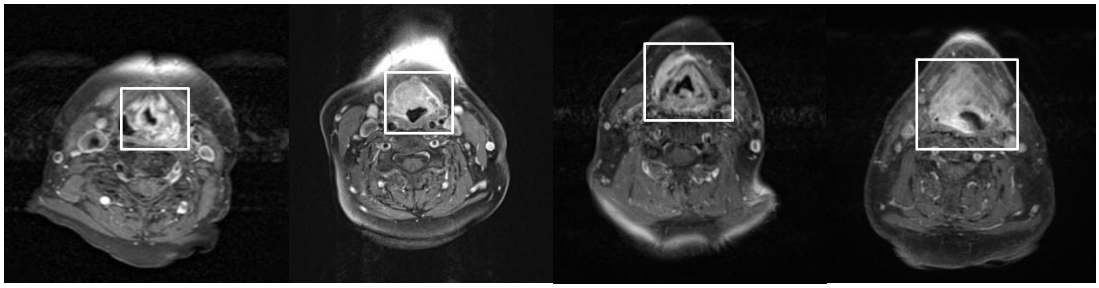
TR – Repetition Time SNR – Signal to noise ratio



Patient 1 (a) (Oropharynx) Patient 2 (Oropharynx) Patient 3 (Oropharynx)



Patient 4 (b) (Oropharynx) Patient 5 (Oropharynx) Patient 6 (Oropharynx)



Patient 7 (Larynx) Patient 8 (Larynx) Patient 9 (Larynx) Patient 10 (Larynx)

Fig. 4.2: Typical real T1+Gd MRI slices, obtained from different MRI scanners with different imaging protocols, with oropharynx and larynx tumours present (white outline), from each of 10 patients used in this research. Different contrast uptake, shape and size, intensity overlap with neighbouring structures of the tumour regions (GTV) can be observed (a) MRI slice with tumour that is not fully enhanced. (b) MRI slice with tumour diffused in normal tissues with no distinct edge.

Ethics statement

This research did not require formal ethical approval as all the MR images utilised in the research were fully anonymised, archival images obtained from the research databases in the Clinical Oncology department of The Beatson West of Scotland Cancer Centre, Glasgow.

4.3 Procedure to obtain reference segmentation on real data

Several studies [2], [30], [114] have shown that manual technique can be used to assess the performance of any proposed algorithm. Thus, manual segmentation results obtained from medical experts was used as reference segmentation in this research to validate the automatic segmentation results. The MRI scans of all 10 patients were peer reviewed in the oncology centre at the head-and-neck radiotherapy weekly team meeting. From that meeting, expert general consensus on the tumour outline on all 2D (axial) slices (102 axial slices), which is considered as a gold standard in current clinical practice, was obtained. The tumour outline obtained using the developed automatic algorithms were qualitatively and quantitatively compared to this consensus manual outline (gold standard) on slice-by-slice basis.

Further, two independent clinical oncologists' (C1 and C2) manually traced the outline of tumour on all 102 axial MRI slices using a built-in Matlab GUI (Appendix A) according to the published guidelines [32], [33]. In order to analyse inter- and intra-variability in manual segmentation results, C1 and C2 repeated this procedure on the same datasets approximately one month later. C1 and C2 were provided with no prior information on patients either clinical or radiological. C1 and C2 have expertise in upper gastrointestinal radiotherapy. The algorithm GTV results were also compared to GTV outlines obtained from C1 and C2.

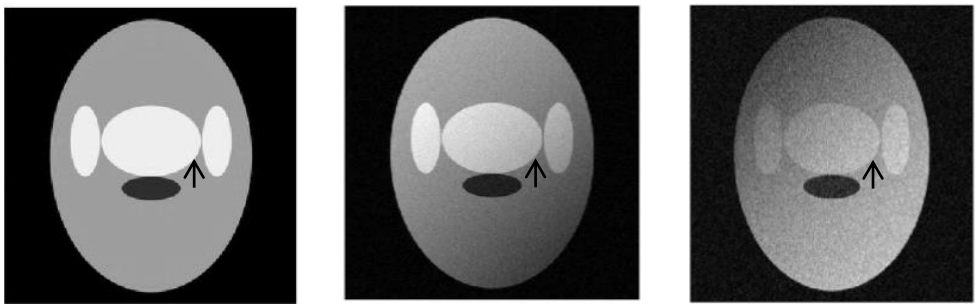
For the evaluation of automatic throat segmentation results, outline of the throat regions obtained from one experienced CO was used as a reference standard.

4.4 Synthetic dataset

Synthetic MRI data is generated analytically from k-space coordinates using the Shepp and Logan head phantom [146] function as described in [147]. From this raw data, MRI slice is reconstructed using FFT. This synthetic MRI slice (Fig. 4.3(a)) is superposition of five ellipses. Large ellipse represents head and small ellipses inside this large ellipse represent features likes tumour region, salivary glands and throat region. The resolution of the slice is 256x256 pixels.

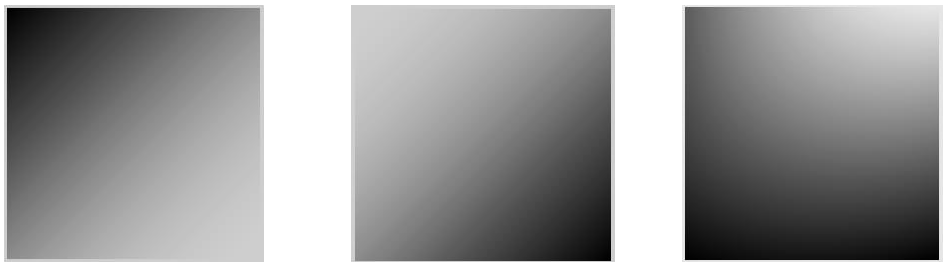
Further, to validate the robustness of the algorithm in presence of MRI artefacts, the generated MRI slice is corrupted with an increasing level of Rician noise (with standard deviation of 5-10-15), decreasing level of tumour-to-background

contrast ratio (0.5-0.3-0.2) and increasing level of IIH (sin-cos-sincos). An example of a bias field generated using sin, cos and sincos function is shown in Fig 4.4. A total of 27 synthetic MRI slices were created to test the system (Fig. 4.3(b) and (c)). The reference slice for this evaluation is the reconstructed MRI slice without the addition of any artefacts (Fig 4.3(a)).



(a) (b) (c)

Fig. 4.3: Examples of synthetic MRI slice (a) ground truth (b) MRI slice with noise of standard deviation: 5, contrast ratio: 0.5, and IIH: cos (c) MRI slice with noise of standard deviation: 15, contrast ratio: 0.2, and IIH: sin (black arrow indicate cancer region ellipse)



(a) (b) (c)

Fig. 4.4: Example of IIH used in synthetic MRI data set (a) sin (b) cos (c) sincos function

Synthetic volume

To validate the 3D throat region segmentation algorithm, synthetic MRI volume was also generated similar to synthetic MRI slice data that include simulated throat region. The throat region was constructed as low intensity ellipses of varying sizes, approximating the shape of a conical sphere (Fig. 4.5). In total 32 slices were reconstructed with varying size of the throat region with in-plane resolution of 256x256 pixels, thus creating volume of 256x256x32 voxels. The generated MRI volume was corrupted with similar artefact as synthetic MRI slice. Thus, a total of 27 synthetic MRI volumes were created to test 3D throat region segmentation algorithm.

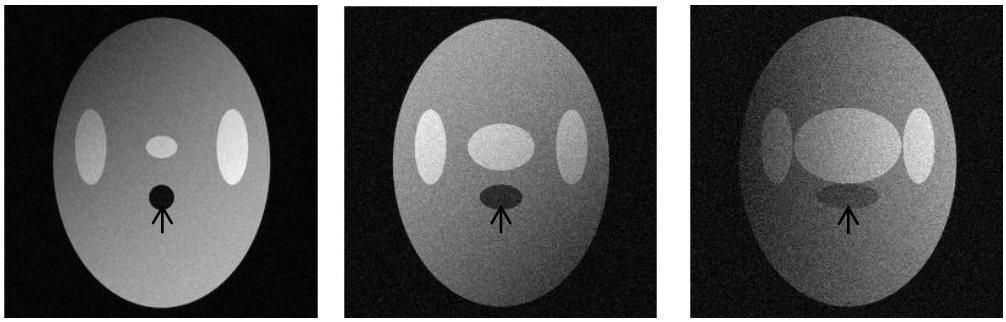


Fig. 4.5: Example of MRI slices from synthetic MRI volumes with addition of MRI artefacts (black arrow indicate the throat region as low intensity ellipses of varying sizes)

4.5 Evaluation parameters

The performance of the developed algorithms in this work is evaluated, by comparing, for a given MRI slice; the segmentation results obtained from developed algorithms and manual (reference) segmentation results, using a number of region and distance-based quantitative (evaluation) measures. The important purpose of an automatic framework in this context is to separate the desired region from the rest of anatomical structures and obtain the accurate outline of the separated region. Therefore, we study the dice similarity coefficient (DSC), F-measure, modified Hausdorff distance (MHD), and correlation analysis which are common metrics to assess the segmented area and outline of the segmented region [2], [77], [117].

Dice Similarity Coefficient (DSC)

DSC [148] is chosen as one of the evaluation parameter in this work as it is one of the best metric [3] to measure the spatial overlap (alignment) of the two segmented areas. DSC normalises the number of true positives to the average size of the segmented areas. For a given automatic binary segmentation mask K_{auto} and a reference mask K_{ref}, DSC is given as:

$$DSC(A,B) = \frac{2(K_{auto} \cap K_{ref})}{K_{auto} + K_{ref}} \qquad 4.2$$

where $\cap$ is intersection and $+$ is addition operator. A DSC value of 1 indicates perfect spatial agreement between segmentation and reference segmentation mask, while 0 indicates no agreement. A DSC value of ≥ 0.7 was indicated as an acceptable

agreement in [2], [30], [114] for a cancer region. A drawback a DSC metric is that, it does not capture shape variations in the outline of the segmentation results.

Modified Hausdorff Distance (MHD)

In RTP, CO is interested in the accuracy of the segmentation outline of the tumour. Any over-inclusion will result in healthy tissue being attacked with radiation; any under-inclusion will lead to that part of the tumour not being treated as it should. Thus, a surface distance measure like MHD metric, which capture shape variation in the segmentation outline, is used in this work to evaluate the accuracy of the segmentation outline.

The Hausdorff distance [149] measure the maximum of the minimum distance between two sets of points of segmentation outlines. MHD [149] reduces impact of outliers and noise compared to original Hausdorff distance. MHD between two point sets, $P_{auto} = \{k_1, k_2, ...k_{Nauto}\}$ and $P_{ref} = \{k_1, k_2, ...k_{Nref}\}$ is given as:

$$d(P_{auto}, P_{ref}) = \frac{1}{N_{auto}} \sum_{k \in P_{auto}} d(k, P_{ref}) \text{ and } d(P_{ref}, P_{auto}) = \frac{1}{N_{ref}} \sum_{k \in P_{ref}} d(k, P_{auto})$$

$$h(d(P_{auto}, P_{ref}), d(P_{ref}, P_{auto})) = max(d(P_{auto}, P_{ref}), d(P_{ref}, P_{auto})) \quad 4.3$$

where $d(k, P_{ref})$ is minimum Euclidean distance between point k in set P_{auto} to all points in P_{ref}. Low variation in segmentation outlines yield low MHD values (< 1mm) [130].

F-measure

Precision (positive predictive value) and recall (true positive rate) (sensitivity) are useful measures of segmentation quality in terms of over and under segmentation [150]. Precision measures the percentage of pixels in the automatic segmentation that correspond to pixels in the reference segmentation and is sensitive to over-segmentation. A low precision indicate large false positive pixels (over segmentation). Recall measures the percentage of pixels in the reference segmentation that were detected in the automatic segmentation and is sensitive to under-segmentation. F-measure [77] is the harmonic mean of precision and recall.

Precision, recall and F-measure, using true positive (TP), true negative (TN), false positive (FP) and false negative (FN) pixels (Fig. 4.6), are calculated as:

$$precision = \frac{TP}{TP + FP}, \quad recall = \frac{TP}{TP + FN}$$

$$F - measure = 2.\frac{precision.recall}{precision + recall} \qquad 4.4$$

Similar to DSC, F-measure value of 1 indicate perfect agreement, while 0 indicates no agreement.

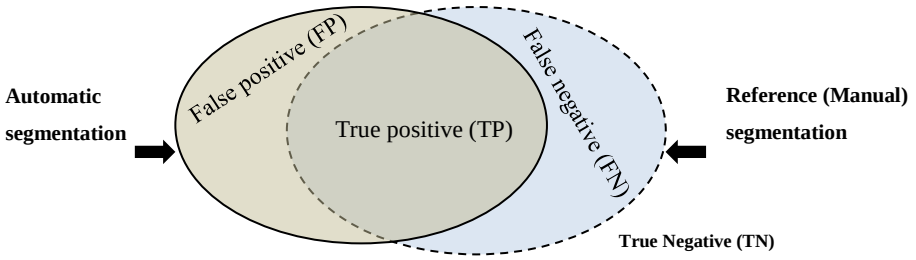
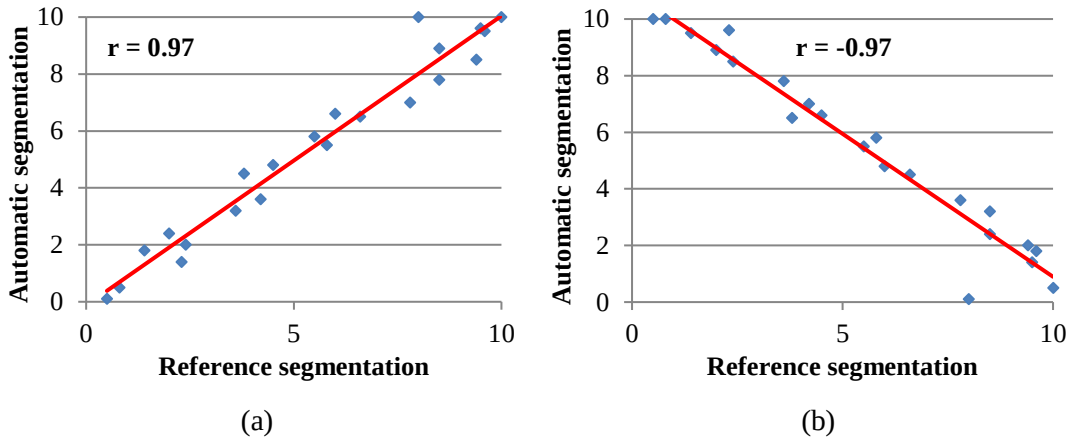


Fig. 4.6: An illustration of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) pixels used to calculate precision and recall.

Correlation analysis

The Pearson correlation coefficient (PCC) r measure the strength of a linear association between two variables [117] (for instance, association between segmented cancer areas by automatic and manual method). The r value ranges from -1 to 1 with sign indicating the direction of linear relationship between two variables. The r value of -1 indicates perfect negative correlation, 0 indicates no correlation and 1 indicates perfect positive correlation. A value of >0.6 indicates a strong positive correlation and >0.8 very strong [117].



(a) (b)

Fig. 4.7: Linear correlation analysis (a) positive (b) negative (Correlation coefficient (r) indicates the strength of correlation, red line indicate direction of the data)

Box plot

Box plot [151] is a graphical tool used to show the general distribution of a dataset (Fig. 4.8). The bottom dash line in the graph is data minimum: the bottom of the box is the estimated 25% point; the middle line (red line) in the box is data median; the top of the box is the estimated 75% point and the top dash line is the data maximum. The red plus represent the outliers of the sample data. Boxplot can provide information about the shape of a data set. This can be interpreted by looking at the position of the median in the box. If the median line within the box is not equidistant from the hinges then the data is skewed.

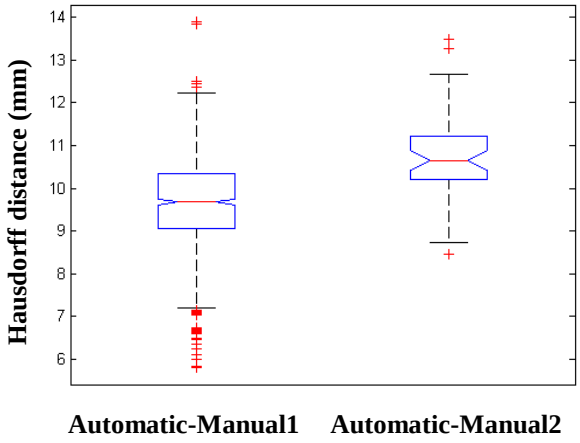


Fig. 4.8: Example of Box plot showing range of Hausdorff distance (mm) values with general indication of their distribution. The centre red line in the box shows median value, the ends of the box shows one and three quarters of the way through dataset, and end of whiskers show minimum and maximum values of data, red plus are the outliers of the data.

Statistical analysis

Statistical analysis of the segmentation results are performed using Minitab Statistical Software (version 17.0). Each data set is first given an Anderson-Darling (AD) test in order to ascertain if the data were normally distributed. If the data were not normally distributed, the normality was assumed as t-test and linear regression are robust to the assumption of normality for large sample size (N>=80) [152]. After normality test, the comparison dataset used a two tail (paired) independent t-test to test the statistical significance. The alpha level was set at 0.05. The null hypothesis was H0: no difference between the two segmentation results, and this hypothesis was rejected only when ρ -value <0.05.

For PCC, the correlation was assumed to statistically significant when ρ-value <0.05. Further, to verify if correlation coefficient is statistically significant for given sample size, sample size was calculated based on level of significance ρ (Type I error), a power value (Type II error) and effect size. Effect size in this case is measured as r and represents the strength of the relationship. These effect sizes for the PCC are 0.10 for a small effect size, 0.30 for a medium effect size, and 0.50 for a large effect size [153]. Thus, for a ρ of 0.05, a power of 0.80, and a medium effect size, sample size of 85 is required [153].

4.6 Conclusion

In this chapter the real T1+Gd MRI data collection of oropharynx and larynx cancer patients was presented. The synthetic MRI data generated to validate the robustness of the developed algorithms in presence of MRI artefacts was also described. Some evaluation metrics were introduced to assess the performance of the proposed algorithms. In the following chapters (Chapter 5 and Chapter 6) the novel techniques developed for the segmentation and quantification oropharynx and larynx tumours (GTV) will be discussed and implemented on MRI dataset discussed in this Chapter.

Chapter 5

Novel oropharynx and larynx tumours segmentation and quantification framework

5.1 Introduction

In Chapter 4, the collection and analysis of real (clinical) and generation of synthetic MRI data used in this thesis was described. The evaluation parameters used to investigate the performance of the proposed algorithms were also introduced in Chapter 4. The objective of this work is to develop an automatic, reliable and quantitative assessment system for oropharynx and larynx tumours from MRI datasets described in Chapter 4. Few systems discussed in Section 3.7 developed for the segmentation of tongue (oropharynx) [103], [110] or larynx tumours [154] from MRI are either semi-automatic [103] which require user input from an expert or require multi-modality MRI [110]. Further, these systems [103], [110], [154] lack through validation with manual approach used in a current clinical practice. Thus, this chapter presents the framework (PLCSF) which includes the automatic extraction and segmentation of oropharyngeal and laryngeal tumour regions in 2D with 3D reconstruction and volume calculation from T1+Gd MRI slices.

The remainder of this chapter is organised as follows. Section 5.2 describes the reason to develop 2D segmentation framework. Section 5.3 presents the proposed framework (PLCSF). Section 5.4 describes the 3D reconstruction and volume calculation algorithm for segmented tumour outlines with conclusion of this chapter in Section 5.5.

5.2 Why 2D segmentation framework?

In the proposed framework, the tumour regions from each axial MRI slice are segmented separately (2D segmentation Section 3.2) and then reconstructed in 3D for further visualisation and quantification as oppose to direct 3D segmentation (Section 3.2). A major difficultly of the dataset used in this work is that the sequential MRI slices obtained using same imaging protocols from a single MRI scanner for one

patient show dissimilar intensity values for the same tissue types (tumourous tissue) (Fig. 4.2) with no predictable pattern. Thus, to consider all MRI slices in one volume for 3D cancer segmentation, intensity standardisation that reduces intra-intensity variation across MRI slices need to be employed. However, intensity standardisation in presence of large pathologies (tumours) remains a very challenging problem [34]. Further, MRI data in this study is highly anisotropic with large spacing in between consecutive slices (maximum spacing 6 mm), which may provide inaccurate results if using initial 3D segmentation. For instance, 3D level set segmentation with large slice spacing may cause collapsing of level set function and missing the desired region in the slice. Moreover, the 3D segmentation approaches use global parameter values (such as a single threshold value for all MRI slices) which may not be suitable for slices with non-standardised intensity range. Thus, using a 2D approach in this work, we aim is to obtain segmentation results in close agreement with the manual outlines in the limited time frame.

5.3 Pharyngeal and laryngeal cancer segmentation framework (PLCSF)

The flowchart in Fig. 5.1 illustrates the steps of the proposed framework (PLCSF) for oropharynx and larynx tumours (GTV) segmentation from axial T1+Gd MRI slices. The proposed PLCSF takes a single axial MRI slice as input and produce the outline of the expected cancer region as output. For this, initially the selected region of interest (ROI) from MRI slice is processed to increase the contrast (intensity difference) between different tissues and to reduce background noise and IIH. Anatomical (Throat) regions close to the expected tumour location are detected from the ROI. MFCM clustering [11] which includes a squared Euclidean distance measure created from the average point of the throat region is used to separate different tissue types in different clusters. Further, SUSAN technique [61] and morphological filtering [60] are applied to the selected cluster to reduce noise and wherever appropriate concavity checks are utilised for labelling of the expected tumour regions. Finally, the labelled region is used as an initialisation for the localised region based level set evolution [122] which converges to obtain the final segmentation boundary.

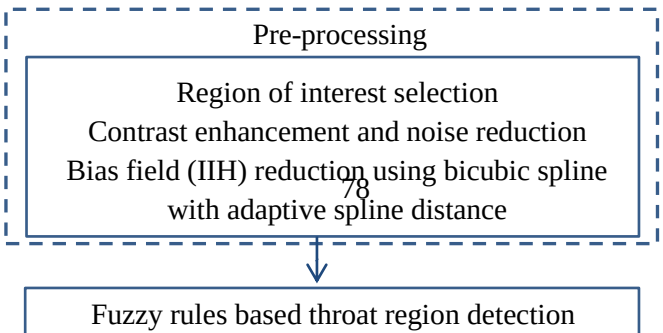


Fig. 5.1: PLCSF flowchart for oropharynx and larynx tumours segmentation from T1+Gd MRI slices.

5.3.1 Pre-processing

The ROI $R(x,y)$[11] is obtained automatically from a MRI slice $I(x,y)$ (Fig. 5.2(a)) to reduce the computational complexity and processing time of the algorithm using an Otsu threshold value Th [155] and knowledge that tumour is present in the upper part of a slice. To obtain ROI, first the binary mask $K(x,y)$ is calculated as:

$$K(x,y)=\begin{cases}1, & if \quad I(x,y)\geq Th\\ 0, & otherwise\end{cases} \tag{5.1}$$

Then, the height $BB_d(k)$ of the bounding box [156] of each connected component in $K(x,y)$ is obtained as:

$$BB_d(k)=height\{BB_k\} \qquad k\in 1...N \tag{5.2}$$

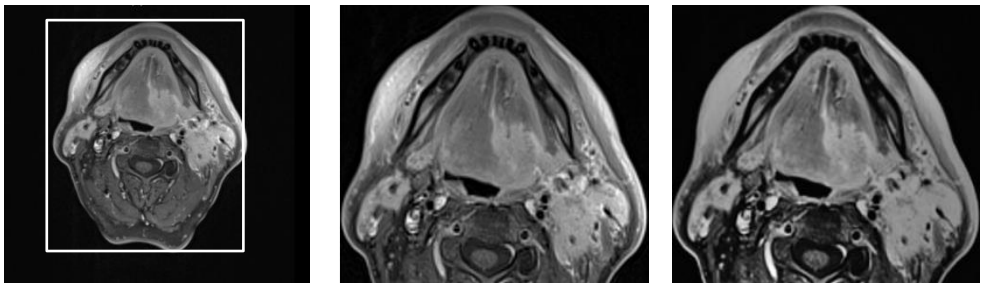
where BB_k represents a bounding box of the k^{th} connected component and M represents the number of connected components in $K(x,y)$. The active area $A(x,y)$ (Fig. 5.2(a) white box) is selected as the region from $I(x,y)$ corresponding to the connected component in $K(x,y)$ with the bounding box of the largest height.

$$A(x,y) = \underset{BB_d(k)}{arg\,max}\, I(x,y) \tag{5.3}$$

The ROI $R(x,y)$ (Fig. 5.2(b)) is obtained as 1 to 70 percent (%) of height and entire width of the active area $A(x,y)$.

$$R(x,y) = A(x,y) \qquad x \in 1..N_x \text{ and } y \in 1..0.7 * N_y \tag{5.4}$$

where N_x the number of columns (width) and N_y represents the number of rows (height) of the $A(x,y)$.



(a) (b) (c)

Fig. 5.2: (a) Original MRI slice (white box: active area) (b) original ROI (c) pre-processed ROI which is further employed for segmentation of a tumour (ROI: region of interest).

The obtained ROI (Fig. 5.2(b)) have poor contrast (intensity difference) between different tissue types. Thus the contrast among different tissue types present in a ROI is increased using background brightness preserving histogram equalisation (BBPHE) [57]. In BBPHE, ROI is divided into background and foreground regions. HE is applied separately to these regions and then they are combined to get the final output ROI. Unlike traditional HE [54], BBPHE avoid enhancement of background noise while increasing the contrast among different tissue types in the foreground. For BBPHE, in this work the ROI is divided into foreground and background using the first local minima of the ROI histogram. The background noise is further reduced from ROI $R(x,y)$ using the log-exponential transform [54] as:

$$R(x,y) = exp(log(R(x,y)) \tag{5.5}$$

In this transform, ROI is log transformed and then corresponding exponential transformed of the log transformed ROI is obtained. The purpose of this transform is to compress low intensity pixels and enhance high intensity pixels thus, to increase the contrast and reduce the background noise.

Finally, IIH observed in a contrast-enhanced noise removed ROI is estimated using bicubic spline where the placement of knots is determined using derivative and filtering technique (Section 6.2). The estimated IIH is reduced using local entropy minimisation technique. A detail description of IIH estimation and reduction technique is presented in Section 6.2 of this thesis. A typical pre-processed ROI is illustrated in Fig. 5.2(c).

5.3.2 Expected tumour region extraction from pre-processed ROI

From a pre-processed ROI, spatial and intensity information of the expected tumour region is utilised to extract the region. Initially, the throat region (anatomical structure) close to the expected tumour location is detected from the ROI using thresholding and fuzzy rules based system (Section 6.3). The distance measure created from the average throat region pixel is embedded in FCM clustering to assign high membership values to the pixels close to the throat region. Using this modified FCM (MFCM) (Section 6.4) ROI is divided into five clusters to separate different tissue types in different clusters. Then, the expected tumour pixels (Fig. 5.3(a)) are labelled from the selected cluster using the membership values of the pixels. Detection of the throat region and MFCM clustering with extraction of a tumour region has been described in detail in Section 6.3 and Section 6.4 respectively of this thesis.

5.3.3 Labelled tumour region boundary refinement

Noise and isolated speckles from the labelled tumour pixels are reduced using SUSAN technique and morphological filtering. The SUSAN algorithm uses a non-linear filtering to reduce noise while preserving edges. A mask of 3x3 and brightness threshold of 10 is used to discriminate noise and underlying image features. Then the morphological filtering is applied to remove small regions and to preserve topology of the large regions. Morphological filtering here is a closing operation followed by an opening process with disk shape structuring element of radius 5.

For some MRI slices, the labelled tumour region is around the throat region. To decide the throat region inclusion in a detected region, a concavity measure [54] of the detected region is computed. Concavity measure is obtained by dividing the area of region with its convex area. If the concavity of the detected region is less than certain

value (0.7), the throat region is included in the labelled tumour region, otherwise it is excluded.

(a) (b) (c)

Fig. 5.3: (a) labelled pixels for the candidate as the tumour pixels (b) initialisation (zero level set) for LSM obtained after filtering (a) (c) final boundary of the labelled tumour region obtained using PLCSF (black outline) and manual outline (white)

Finally, the localised region based LSM where zero (initial) level set (Fig. 5.3(b)) represents the boundary of the noise removed labelled tumour pixels is evolved to obtain the final smooth boundary. In this work, the signed distance function is utilised as a level set function and the region statistics such as mean and area of the local interior and exterior region around each point (Fig. 5.4) on the zero level set is considered as a force V (Eq. 3.14) for the level set evolution. This localised region based model successfully segments the heterogeneous region with weak edges, where both global region-based and local edge based methods (Section 3.5.3) fail.

Thus the evolution equation for local region based LSM, similar to Eq. 3.14, is given as [122]:

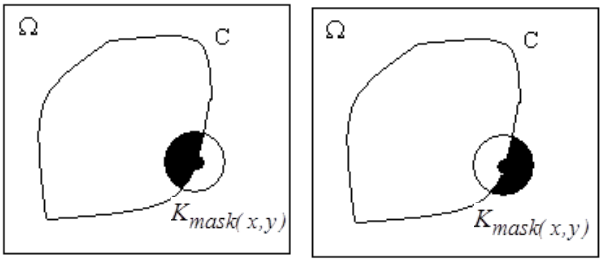
$$\frac{\partial \phi}{\partial t}(x) = -\delta(\phi(x))\nabla V + w\delta(\phi(x))div(\frac{\nabla \phi(x)}{|\nabla \phi(x)|}) \qquad 5.6$$

where $\delta(\phi(x))$ is Dirac delta function which specify the area just around the curve. The first term on the right hand side of Eq. 6.6 is the external force term ∇V which drives level set to the desired boundary and second term $div(\frac{\nabla \phi(x)}{|\nabla \phi(x)|})$ is the curvature of the level set function ϕ which keeps ϕ smooth and is penalised by parameter w.

The external term is given as:

$$\nabla V = \int_{\Omega_y} \delta(\phi(y))K_{mask}(x,y)\left(\frac{(I(y)-\varpi_{inx})^2}{A_{\mu_{in}}}\right)dy - \int_{\Omega_y} \delta(\phi(y))K_{mask}(x,y)\left(\frac{(I(y)-\varpi_{outx})^2}{A_{\mu_{out}}}\right)dy \quad 5.7$$

where ϖ_{inx} is the mean and $A_{\mu_{in}}$ is the area of interior and ϖ_{outx} and $A_{\mu_{out}}$ of the exterior local region $M(x,y)$.



(a) (b)

Fig. 5.4: Area used in level set evolution (a) Local interior is shaded part of the mask and in (b) local exterior is shaded part of the mask along the contour C [122] © 2013 IEEE.

In this work, the radius of local region $K_{mask}(x,y)$ (Fig. 5.4) is 1% of the number of elements (foreground pixels) in the initialisation of LSM (Fig. 5.3(b)). The mean separation energy where optimisation of energy occurs when foreground and background have maximally separate mean intensities is used to stop the evolution [157]. When the surface evolution converges after 500 iterations, the final tumour segmentation boundary is obtained. Obtaining the smooth boundary of the tumour region is important in radiation oncology as clinical expert outlines the cancer borders using contouring for radiotherapy planning as opposed to pixel by pixel labelling as done in clustering algorithms. This smoothing process is expected to obtain a tumour boundary close to the expert's manual outline. The qualitative result of level set evolution is shown in Fig. 5.3(c).

5.4 3D reconstruction, smoothing and volume calculation of segmented tumour outlines

The process of reconstructing the tumour in 3D, for visualisation, from 2D contours (outlines) is illustrated in Fig. 5.5. The individual tumour contours (Fig. 5.5(a) blue outlines) obtained from a sequential MRI slices are stacked to produce a 3D dataset

(Fig. 5.5(b)). The slice location and spacing between slices are determined from the DICOM image header to stack the contours (Fig. 5.5(b)).

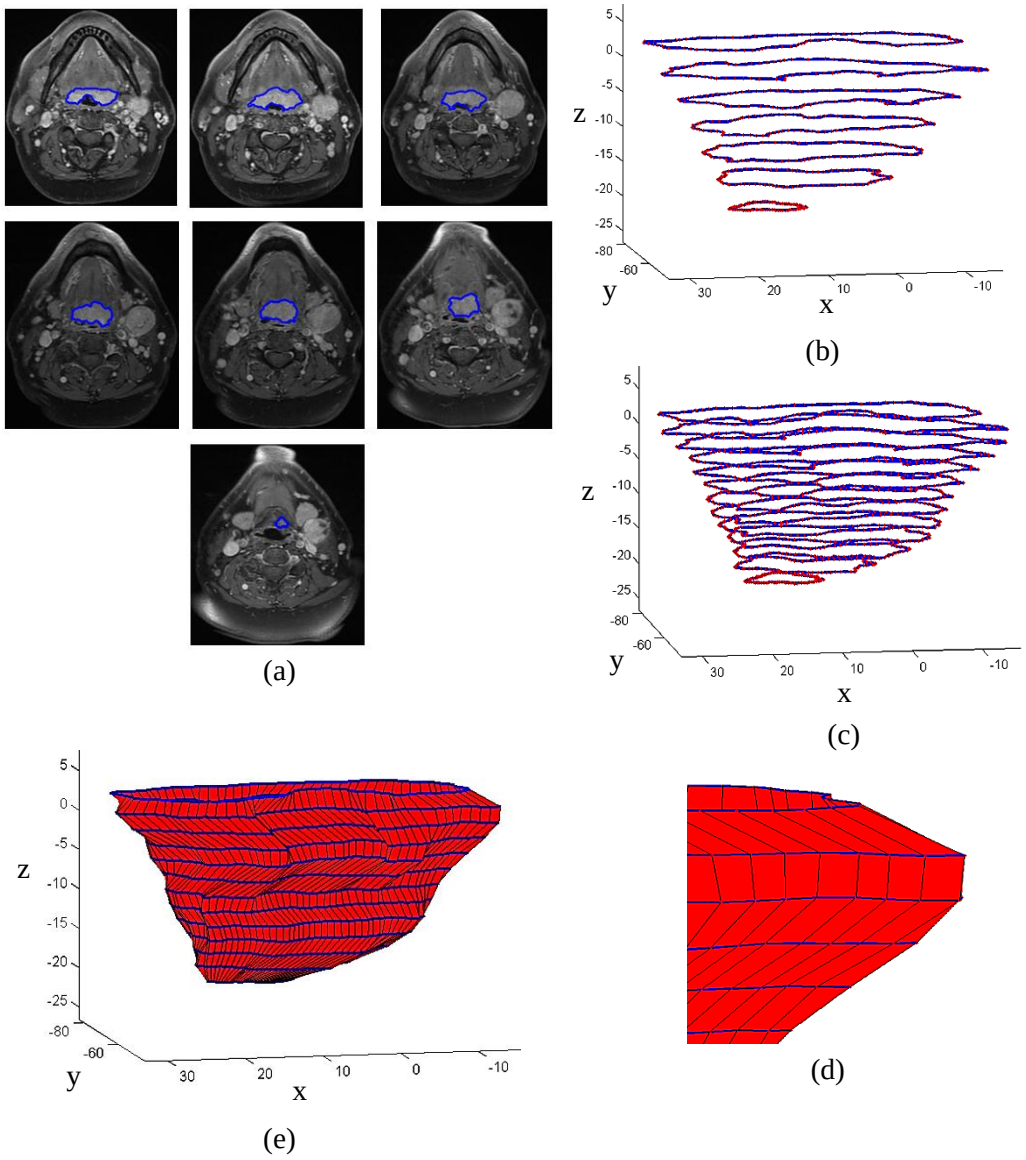


Fig. 5.5: 3D reconstruction of a tumour from 2D contours (a) tumour contours (blue outlines) obtained using PLCSF on 2D axial MRI slices from a patient (b) tumour contours in 3D dataset (blue outlines) with indices indicated by red dots (c) interpolation between contours (d) rectangular mesh generation (e) 3D reconstructed tumour region.

The indices (red dot on blue outlines in (Fig. 5.5(b)) of the contour pixels along with slice location allow converting the contour data into a set of points in 3-dimensions. As there is a large spacing (3.5-6 mm) in between the contours, in our implementation the information between the indices of pair of contours (outlines) is interpolated (Fig.

5.5(c)) to obtain a smooth 3D structure. The interpolation is performed using cubic spline technique [158]. Cubic spline interpolation is a piecewise interpolation process that relies on constructing a polynomial of third degree between each pair of indices. This interpolation obtains a smoother curve compared to first and second degree polynomial and avoids overshoot observed in higher degree polynomial. Further, a mesh (Fig. 5.5(d) red patch) is produced between a pair of contours using indices of the contours. The mesh process joins two contours by creating a strip of rectangles using the correspondences between indices. Mesh generation is repeated for each pair of contours to reconstruct a region. The reconstructed tumour region is illustrated in Fig. 5.5(e).

As tumour region on each axial slice is segmented separately without considering spatial information from neighbouring slices, the irregularities in terms of slight over and under segmentation of tumour regions are observed in sequential 2D contours which are reflected in a 3D reconstructed tumour (Fig. 5.6(a) black arrows). In this work, these irregularities in 2D contours are reduced by automatic smoothing of the 3D reconstructed volume (Fig. 5.6(b)). The smoothing is performed using a two-step Gaussian algorithm described in [159]. In this two-step approach, alternating smoothing and un-smoothing process is implemented to preserve the shape more realistically than the conventional Gaussian smoothing filter. Thus, 3D smoothing removes high frequency variations (such as sharp spikes and tail regions) that may be present on 2D tumour outline and produce smooth 3D structure of a tumour region. However, the limitation of this 3D smoothing is that it sometimes over smooth the region which results in the inclusion of other anatomical structure in the tumour outline (Fig. 5.6(d)). This is because, the 3D smoothing does not utilise any underlying grey-level information from the slices as it is only based on positional information.

Finally, the volume of the reconstructed region is calculated in cubic centimetres (cm^3) by the summation of all 2D tumour areas in each slice and multiplication by the slice profile (slice thickness plus spacing in between slices (mm)). The 2D tumour area is obtained by summing the number of pixels inside tumour contour multiplied by the pixel spacing (mm).

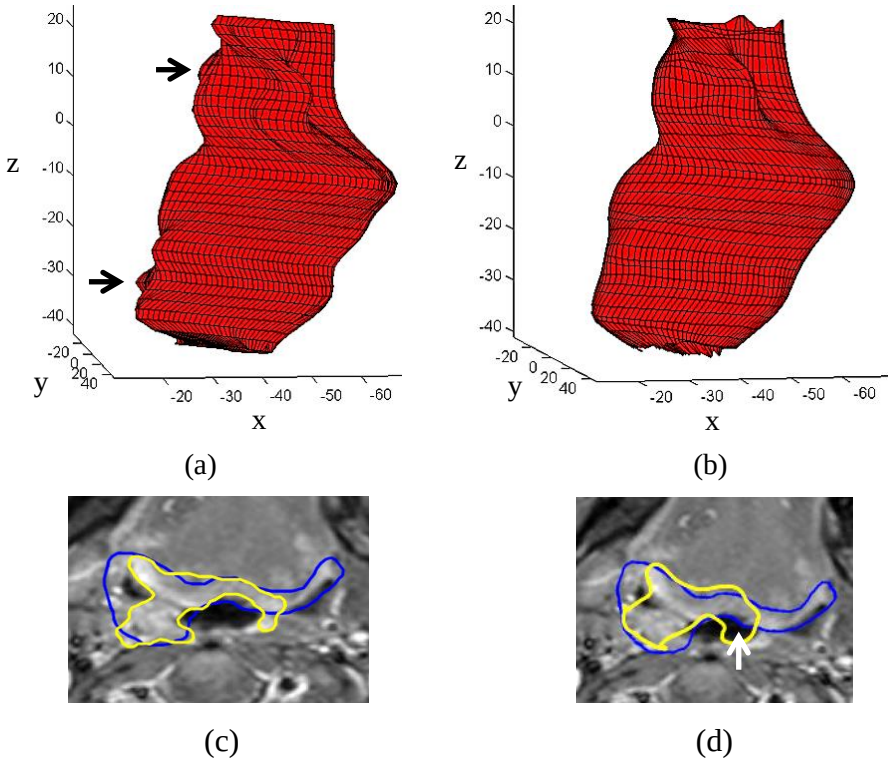


Fig. 5.6: Improvement and limitation of 3D smoothing

Improvement - 3D reconstructed tumour region (a) before smoothing (irregularities indicated by black arrows) (b) after smoothing (irregularities are reduced).

Limitation - Tumour outline (c) before smoothing (throat region is not included in the tumour (yellow) outline (d) after smoothing (throat region (white arrow) is included in the tumour (yellow) outline).

5.5 Discussion

Segmentation of oropharynx and larynx tumours is particularly difficult due to the presence of MRI artefacts, noise, enhancements of other non-tumours regions (blood vessels, salivary glands), geometric variability and weak edges of the tumour regions across the patients. An unsupervised segmentation framework (PLCSF) was presented in this Chapter for this task that does not require any manual intervention or training data. This framework makes no assumption about the shape or size of the tumour regions, thus can successfully segment the tumour regions with geometric variability. The cases used in this study are representative of everyday clinical challenges. PLCSF can also perform robustly against the variations caused by different MRI protocols with different manufacturer and scanner models. To our knowledge, this is the first

study that validates the use of an MRI-based automatic segmentation tool for tumour (GTV) segmentation (delineation) and 3D visualisation in oropharyngeal and laryngeal cases.

In this framework (PLCSF), a novel adaptive determination of parameter spline distance (knot spacing) allowed the correction of complex bias field (IIH) present in MRI slices used in this work. Detection of the throat region using fuzzy rules based technique allowed the knowledge of the approximate tumourous position to be embedded in the system, particularly in MFCM, thus reducing further processing steps to eliminate healthy tissues from tumour detected clusters that are away from the throat region. The continuity and spatial smoothness of the tumour boundary was ensured by evolving the level set surface on the detected region.

This proposed system enables 3D reconstruction of a tumour (GTV) which can provide clinicians with additional information not easily available from 2D MRI slices. Automatic tumour volume measurement could contribute to further tumour staging and prognostication. It has been shown in previous studies [38], [103] that tumour volume is a significant factor in determining the outcome following primary radiotherapy in HNC.

Further, using a single modality (T1+Gd) in RTP can reduce scanning and processing time of MRI slices and increase the computational efficiency. Thus this tool that integrates novel image processing technique can segment the cancer region with substantial variation in intensity, size and shape from 2D MRI slices.

5.6 Conclusion

This chapter presented a new automatic, reliable system (framework) for MRI oropharynx and larynx tumours. The system can effectively reduce artefacts, extract the expected tumour region using prior knowledge and obtain smooth outline of the tumour boundary by integrating image processing techniques such as FCM clustering and LSM from 2D axial T1+Gd MRI slices. Further, system reconstruct tumour in 3D to obtain additional information not available from 2D slices. The novel methodologies incorporated in this framework for IIH reduction, throat region detection and extraction of tumour region are explained in next Chapter (Chapter 6).

Chapter 6

Novel oropharynx and larynx tumours extraction algorithm

6.1 Introduction

The previous chapter introduced the framework for the assessment of oropharyngeal and laryngeal tumour regions (GTV) from axial T1+Gd MRI slices. This chapter presents the novel algorithms included in the framework to correct IIH (Section 5.3.1) and for the tumour region (GTV) extraction (Section 5.3.2) from T1+Gd MRI slices using FCM clustering with the additional distance measure.

As mentioned in Section 5.3.1, the IIH (bias field) (Section 2.6) from MRI slices is estimated using bicubic spline with adaptive spline distance and reduced using entropy minimisation technique. In Section 3.3.2 it was mentioned that for IIH reduction using spline techniques [10], [71], [72], parameter spline distance is important for the correct estimation of a bias field. However, in literature [10], [71], [72] this parameter is set to a constant value for all MRI slices which limits it applicability in intra-intensity variations between slices and presence of tumour regions. Thus, a novel technique to determine spline distance for IIH estimation from each MRI slice is proposed in this chapter.

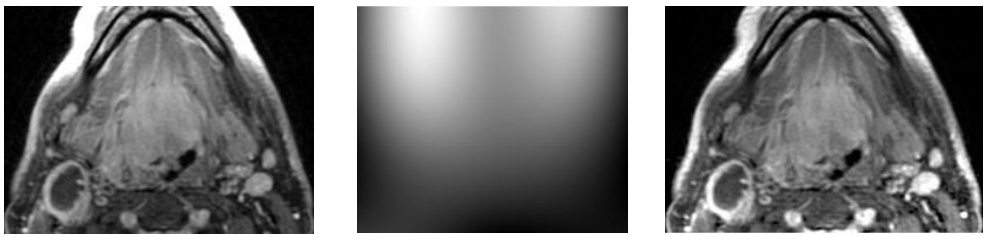
Further, to embed spatial information in the extraction of tumour regions, the throat region (anatomical structure) close to the expected tumour location is detected. The throat segmentation methods described in Section 3.6 either assume shape information [91]–[93], [95] which is violated in presence of pathology near the region or require manual interaction [12], [94], [97]. Thus, a new method for the throat region detection (ThDA) based on fuzzy rules is presented in this chapter. Finally, a MRI slice is divided into different tissue types (clusters) using novel MFCM clustering and tumour region is extracted automatically from the selected cluster. In MFCM clustering, spatial information of the expected tumour region is integrated in FCM clustering [13] to extract the tumour regions. Tumour segmentation techniques using FCM [106]–[111] described in Section 3.7 does not

utilise any spatial information of the expected tumour region.

The remainder of this chapter is organised as follows. Section 6.2 describes the IIH reduction technique with adaptive spline distance and its comparison with constant spline distance. Section 6.3 presents novel throat detection algorithm (ThDA) and its experimental results. Novel MFCM clustering technique employed for tumour (GTV) region extraction and validation of its robustness compared to standard FCM [13] is described in Section 6.4. Section 6.5 concludes the chapter.

6.2 Intensity inhomogeneity (IIH) reduction from a MRI slice

Bias field (IIH) artefact was observed in the MRI data discussed in Section 4.2. To eliminate the effect of this artefact in subsequent analysis of a tumour region, in this work the method described in [10] is used as a basis to reduce the IIH from a MRI slice (Fig. 6.1(a)). This technique [10] was chosen as this technique is developed particularly for the neck region. It considers the challenges faced for IIH (Fig. 6.1(b)) correction in a neck region such as pixels (throat region) that are near the tumour region are void of signal and it can correct complex bias field which is observed in MRI slices in this work.

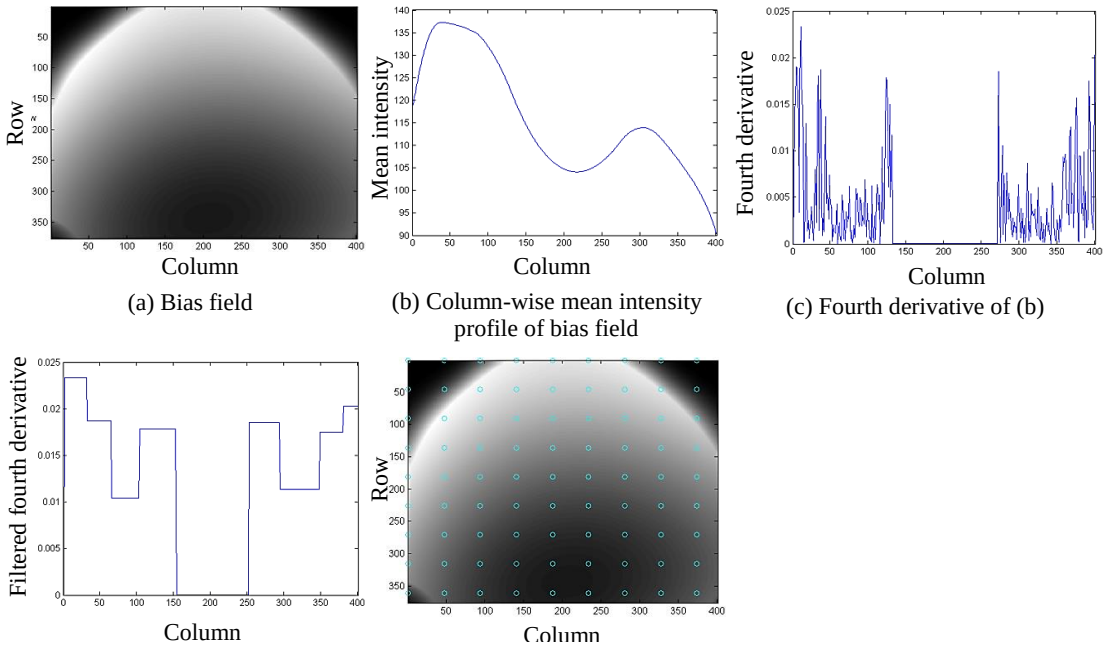


(a) (b) (c)

Fig. 6.1: (a) MRI slice (b) estimated bias field for Fig 5.1(a) (c) IIH reduced slice (roughly obtained by dividing Fig. 5.1(a) by Fig. 5.1(b)) [11] © 2014 IEEE.

Thus, the initial bias field (IIH) from a MRI slice is calculated by fitting a third order polynomial function to all tissue pixels excluding background (near zero intensity) pixels. The polynomial function is fitted to tissue pixels using the least square approximation. This initial estimate of the bias field (IIH) is then refined using bicubic spline defined by grid of knots evenly spread out across the slice by taking into account the value of the bias field at knot locations. In [10], the spacing between the consecutive knots (spline distance) is set to a constant value (21 mm) for all MRI slices. In this work, this spacing between knots is determined automatically

(adaptively) for each MRI slice based on underlying data [160]. The spline distance is determined separately for rows and columns (Fig. 6.2(a)).



(a) Bias field

(b) Column-wise mean intensity profile of bias field

(c) Fourth derivative of (b)

(d) Filtered data from (c)

(e) knot location marked with 'o'

Fig. 6.2: The procedure to determine spline distance parameter (distance between consecutive knots) for bias field estimation from a MRI slice.

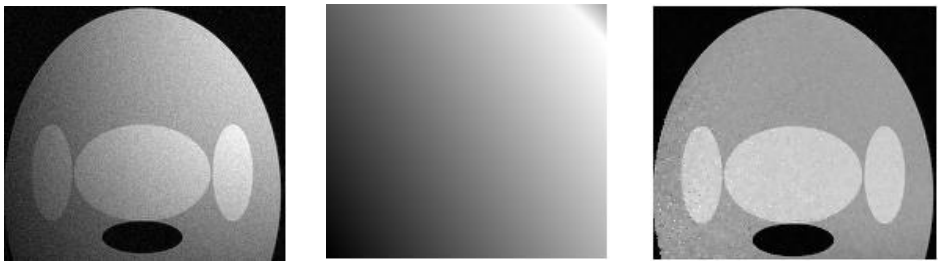
To calculate this distance, initially the fourth derivative of the data (Fig. 6.2(c)) is calculated using central difference method. This data is divided into distinct regions after passing it through a maximum filter [60] of window size 20 and merging small regions into nearby regions to reduce noise (Fig. 6.2(d)). The value of window size is chosen so as to remove undesired high frequency components while avoiding excessive smoothing and merging of adjacent regions. The number of knots N_i in each distinct region is computed [160] as:

$$N_i = \left\lceil \frac{R_{i+1} - R_i}{\sqrt[4]{\frac{\varepsilon}{\chi R_i''''}}} \right\rceil \qquad 6.1$$

where R_i is the region under consideration, χ is a constant and r_i'''' is the fourth derivative value of a region under consideration. The error value ε is determined from the region with global maximum value of the fourth derivative. The total number of knots is obtained by summing the number of knots in each distinct region. The distance

between consecutive knots (spline distance) is subsequently calculated, as the number of elements in data divided by the total number of knots (Fig. 6.2(e)).

All knots have equal spacing between them. With knots spaced evenly across the slice, the estimated bias field value at each knot location is used for optimisation. The optimisation is carried out by minimising the local entropy of the slice, starting with a knot with the highest estimated bias field value, and merging areas of lower values in a sequential (iterative) fashion [10]. The process continues until all knots have been optimised. Fig. 6.1(c) shows a typical IIH reduced slice. Other example of IIH reduction using synthetic MRI slice is shown in Fig. 6.3.



(a) (b) (c)

Fig. 6.3: (a) Synthetic MRI slice (b) estimated bias field for Fig. 6.3(a) (c) IIH reduced slice.

6.2.1 Spline distance comparison: constant vs. adaptive

For synthetic MRI data, the spline distance ranged from 28 to 36 pixels for 27 slices. However, it was observed from the experiments that bias field estimation was not overly sensitive to the spline distance; a value of 30 pixels worked well for all synthetic MRI slices.

For real MRI dataset, the spline distances ranged from 11mm to 38mm for 102 axial slices obtained from three different MRI scanners. Experiments were performed by comparing set of constant spline distances (21mm [10], 30mm [72], and 60mm) against adaptive distance for each MRI axial slice for 10 patients. For quantitative measure of IIH reduction performance, we used a commonly accepted coefficient of variation (CoV) [46], [71], [72] of the tumour region segmented using manual consensus outline from corrected MRI slice. CoV is calculated as:

$$CoV = \frac{\varpi_{corrected}}{\mu_{corrected}} \qquad 6.2$$

where σ and ϖ are standard deviation and mean of a tumour region from IIH

corrected slice respectively. The underlying assumption of CoV is that increasing IIH leads to increase in standard deviation of image intensity within a desired region. Hence, a larger reduction in CoV would correspond to better IIH reduction. The advantage of using adaptive spline distance compared to the default setting of 21mm is illustrated in Fig. 6.4.

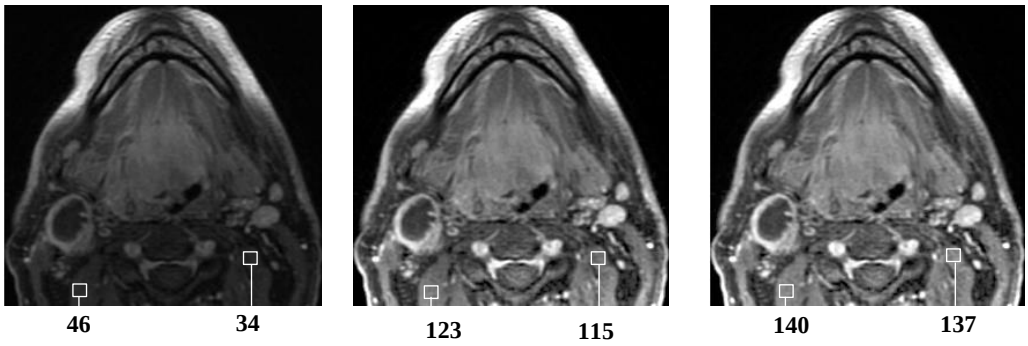


Fig. 6.4: (a) Original MRI slice (b) corrected slice with spline distance 21mm (c) corrected slice with spline distance 16mm (determined using technique described in Section 6.2.) Average intensity values of selected ROIs (white box) are indicated below each slice.

Further, from Fig. 6.5 it can be noticed that CoV value for adaptive distance is less compared to other constant values except for one patient (Patient 2). For Patient 2, due to large tumour region (Fig. 4.2), large knot spacing (60 mm) was required to correctly estimate the IIH compared to 38 mm obtained using proposed technique.

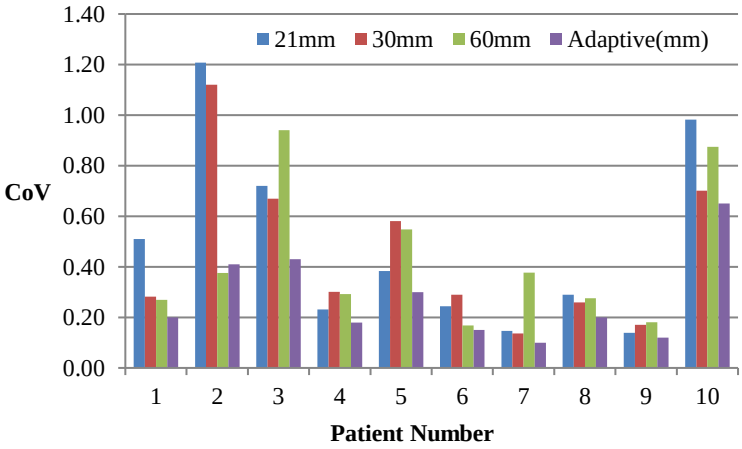


Fig. 6.6.5: Coefficient of variation (CoV) for different knot spacing (spline distance) parameter used to estimate the bias field. Adaptive knot spacing reduces intensity variation (CoV) in a region of interest (ROI) compared to constant knot spacing.

These results (Fig. 6.5) suggest that using adaptive spline distance produces better IIH reduction, which will translate into more accurate segmentation of a tumour region. The requirement for adaptive spline distance for the MRI dataset in this work can be attributed to different range of imaging parameters, variability in tumour areas (range: 89.97-3361.46 mm^2) and significant inter- and intra- intensity variations in MRI data across the patients.

6.3 Fuzzy rules-based throat region detection

In this work it is known that the oropharyngeal and laryngeal tumour regions will be adjacent to the throat region, as illustrated in Fig. 2.15 and Fig. 6.6(a). To utilise this spatial information of the expected tumour regions, to robustly extract the tumours, the throat region is detected. Detection of the throat region from each MRI slice is challenging as the size and intensities of the throat region varies for each slice and topology of the region changes frequently. Thus here, we use a combination of thresholding and fuzzy rule-based methods for the throat region detection. We refer this automatic throat detection algorithm as 'ThDA'.

In ThDA, at first the MRI slice $I(x,y)$ (Fig. 6.6(a)) is converted to a binary image $K(x,y)$ (Fig. 6.6(b)) using a threshold value obtained by Otsu method [155] as:

$$K(x,y)=\begin{cases}1, & if \quad I(x,y)<Th\\ 0, & otherwise\end{cases} \tag{6.3}$$

(a) (b) (c)

Fig. 6.6: Initial steps for a throat region detection (a) MRI slice $I(x,y)$ (b) binary image $K(x,y)$ (c) binary image $K(x,y)$ after removal of connected components with maximum signature radius greater than a certain value.

Further, a signature [54] of each connected component is calculated from the binary image. A signature is a 1-D representation of the boundary of an object,

calculated as distance between pixels on the boundary of an object and its centroid. Any connected component with a maximum signature radius greater than 10 percent (%) of number of rows of the binary image is removed from upper half of the image to reduce false positives (Fig. 6.6(c)). The value of 10% is decided based on empirical tests in which a range of values were systematically tested in turn to assess the effect on detection results.

Furthermore, the knowledge of the throat position in a binary image is incorporated in two fuzzy rules to classify pixels as the candidates for the throat region. These rules are created from the observation of a typical axial MRI slice. From the MRI slice, two typical characteristics of the throat region are observed:

1. The throat region is present in the central part of the binary image plane. Thus, pixels are associated with the throat region using a fuzzy vertical line membership function F_{VL}. F_{VL} is calculated from the column (Fig. 6.7(a)) centre (c_v) and it is given as:

$$F_{VL} = \{(x,y), \mu_{VL}(x,y)\} \quad x = 1...N_x, y = 1...N_y \tag{6.4}$$

with

$$u_{VL}(x,y) = f_{VL}(x) = \begin{cases} 1, & |c_v - x| \le Th_u \\ 1 - \dfrac{|(c_v - x) - Th_u|}{Th_u - Th_l}, & |c_v - x| > Th_u \;\&\&\; |c_v - x| \le Th_l \\ 0, & Otherwise \end{cases}$$

In this work for all MRI slices, Th_u is set to 5 and Th_l to 20. Thus, the highest membership value of 1 is given to c_v and its immediate neighbours gradually decreasing the membership values on either side as distance increases from the centre (Fig. 6.7(a)).

2. The throat region is close to the first row compared to the other connected components in the middle part of the binary image. Thus, fuzzy height membership rule F_H is used to associate the closeness of pixels to the first row of the binary image.

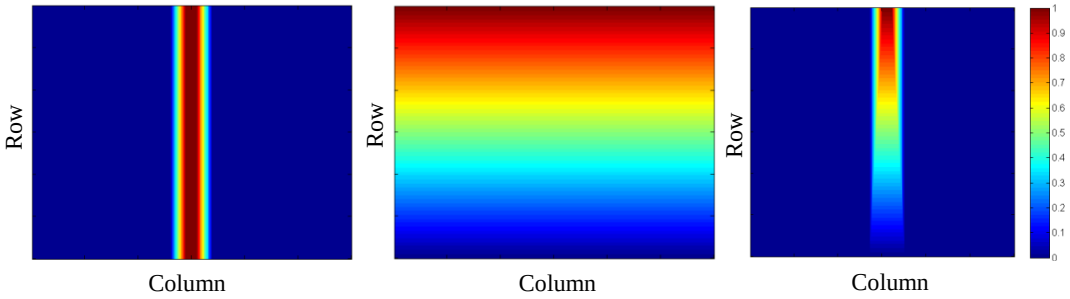
   Therefore, F_H is given as:

$$F_H = \{(x,y), \mu_H(x,y)\} \quad x = 1...N_x, y = 1...N_y \tag{6.5}$$

with

$$\mu_H(x,y) = f_H(y) = \begin{cases} 1, & y == Th_{u1} \\ \dfrac{|Th_{l1} - y|}{Th_{l1} - Th_{u1}}, & y > Th_{u1} \,\&\&\, y \le Th_{l1} \\ 0, & Otherwise \end{cases}$$

where Th_{u1} is set to 1 and Th_{l1} to r_n. Thus, in this the highest membership value of 1 is given to the first row of the binary image and gradually decreasing the membership values to other rows as distance increases from the first row (Fig. 6.7(b)).



(a) (b) (c)

Fig. 6.7: Fuzzy rules for throat region detection (a) Fuzzy vertical line membership function F_{VL} (b) Fuzzy height membership function F_H (c) F_{VLH} : Intersection of F_{VL} and F_H

The fuzzy 'min' operator is used to select the minimum of these two membership values for each element. It is given by:

$$F_{VLH} = F_{VL} \cap F_H \qquad 6.6$$

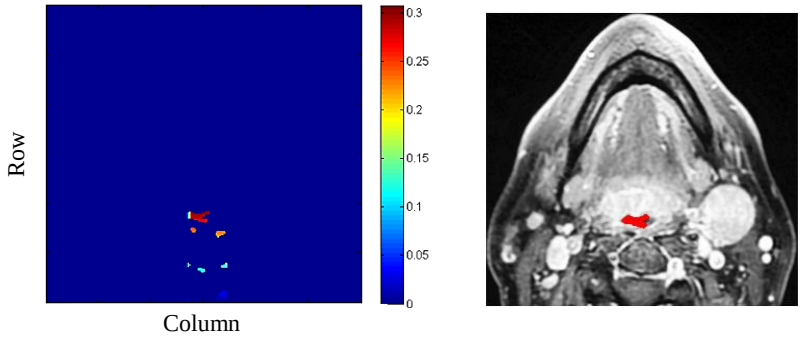
$$\mu_{F_{VLH}}(x,y) = min(w_1 \mu_{VL}(x,y), (1 - w_1)\mu_H(x,y))$$
$$x = 1...N_x, y = 1...N_y$$

The weights w_1 and $1 - w_1$ are assigned to the membership values to vary the relative importance between $\mu_H(x,y)$ and $\mu_{VL}(x,y)$. F_{VLH} obtained from these two fuzzy rules is demonstrated in Fig. 6.7(c). Finally, candidate pixels for the throat region are chosen as:

$$F_{Inter\,sect} = F_{VLH} \cap K \qquad 6.7$$

$$\mu_{F_{Inter\,sect}}(x,y) = min(\mu_{VLH}(x,y), K(x,y))$$

From $F_{Inter\,sect}$ (Fig. 6.8(a)), pixels with highest membership values are chosen as pixels belonging to the throat region. The detected throat region from representative MRI slice is illustrated in Fig. 6.8(b).



(a) (b)

Fig. 6.8: (a) $F_{Intersect}$: Intersection of F_{VL} and B_I (b) detected throat region pixels (in red) as highest membership of $F_{Intersect}$ from a MRI slice.

6.3.1 Experimental results

This section presents the performance analysis of the ThDA algorithm. It consists of two parts: in the first part the qualitative and quantitative results of the ThDA are assessed on real and synthetic MRI slices and in second part; the ThDA is compared with two existing algorithms [11], [12]. The reference segmentation for the throat region from the synthetic MRI slices is obtained by applying the threshold and selecting the smallest low intensity region.

For real MRI data, manual outlines from the expert are used as reference segmentation. For quantitative assessment, F-measure (Section 4.5) is used as an evaluation metric on slice-by-slice basis.

6.3.1.1 Qualitative and quantitative evaluation of ThDA on real MRI dataset

The ThDA successfully detected the desired (throat) region on all 27 synthetic and 102 real MRI slices. For comparison with reference segmentation, Canny edge detector [54] was applied to the detected region to obtain the outline of the region. The visual results of the algorithm are illustrated in Fig. 6.9. In all slices in Fig. 6.9, white outline is the reference outline and red outline is obtained from the ThDA. It can be seen that algorithm can correctly identify the throat region of any shape and size. The underestimation of detected region can be observed in some real MRI slices (Fig. 6.9(a), (b) and (c)) compared to the reference outline. This is because of the variation in image intensity from slice to slice and hence the variation in the threshold value and presence of high intensity tumour close to the region. Further, it can be observed that, if throat is split in two or more regions in the same slice (Fig. 6.9(d) and (e)), ThDA

can detect only one region which is close to the central part of the image. However, it can be noted that the aim of the (ThDA) algorithm development was the detection of the region (as a part or entire region), so that it can be used as a supplementary information in a tumour region extraction and not the segmentation or obtaining accurate outline of the throat region.

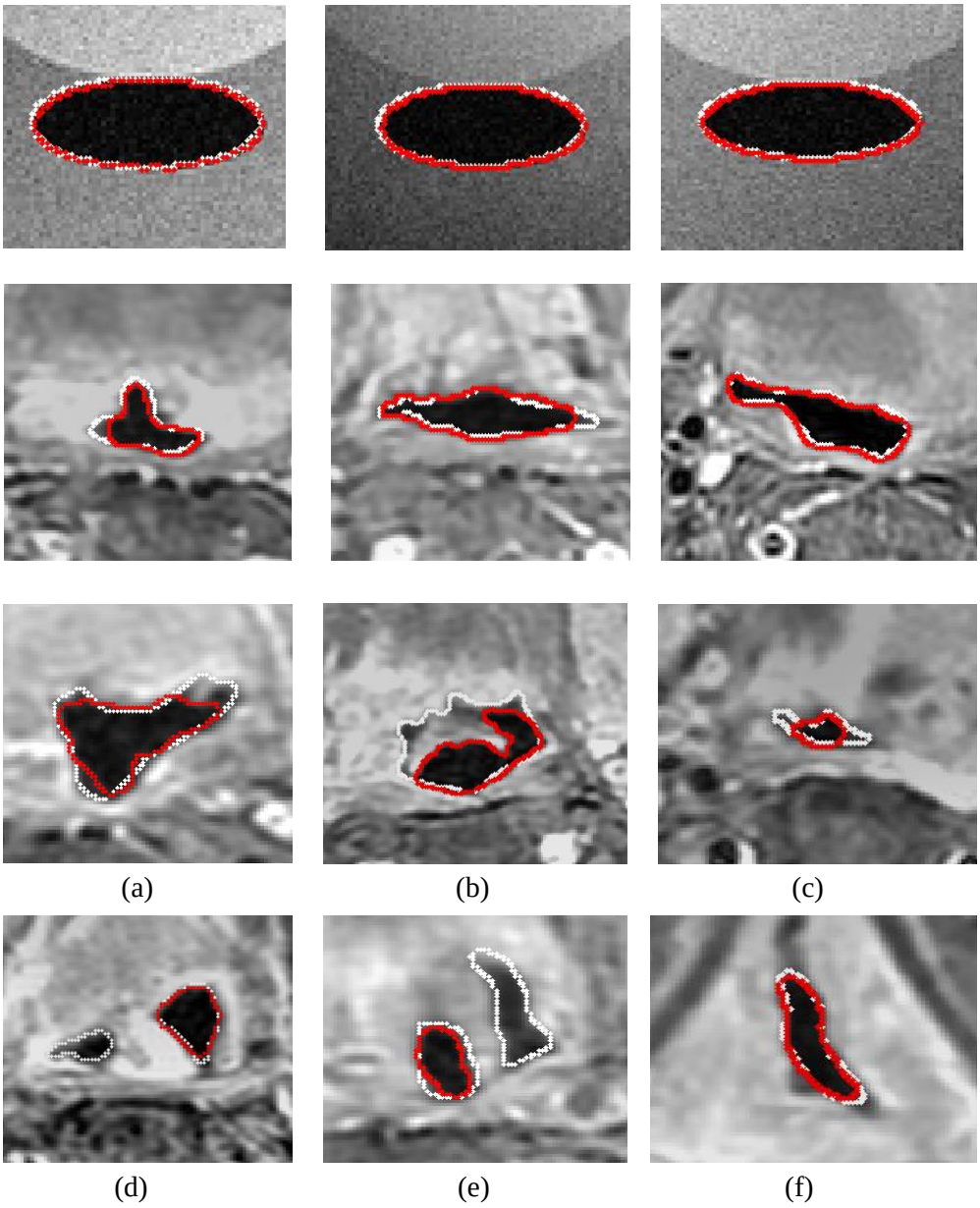


Fig. 6.9: Qualitative results of automatic throat detection algorithm (ThDA) on MRI slices from different patients (red outline: ThDA, white outline: reference outline) (first row-synthetic data). ThDA algorithm can successfully detect the throat region of various shapes, sizes and intensities from MRI images obtained using different imaging protocols.

The graph in Fig. 6.10 shows F-measure value for real MRI dataset. The F-measure value for a particular patient was obtained by averaging the F-measure values for all the slices in that patient. From Fig. 6.10, it can be seen that this value is greater than 0.6 for all patients with mean of 0.77 and standard deviation of 0.07 indicating acceptable agreement with manual results. The maximum value of F-measure observed was 0.94 for the slice shown in Fig. 5.9(f) and minimum value was 0.32 for the slice where throat was split into four regions and algorithm was able to detect only one region. Further, the low values of F-measure in some slices are due to large number of false negative pixels as ThDA underestimates the throat area.

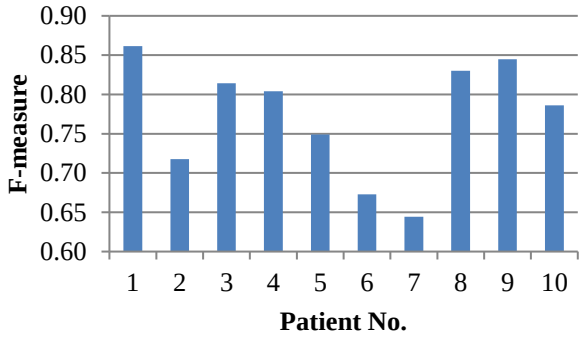


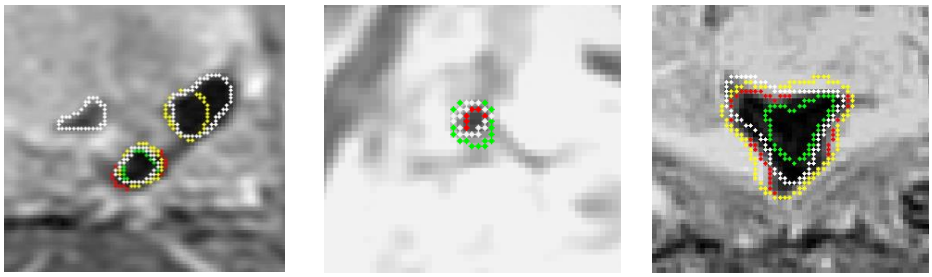
Fig. 6.10: Quantitative results of ThDA from real MRI dataset. The proposed ThDA successfully detected throat region from all MRI slices with mean F-measure of 0.77.

6.3.1.2 Comparison of ThDA with Hessian analysis and fuzzy connectedness

The first comparison algorithm was proposed in our paper [11]. In this algorithm, the low intensity pixels were analysed based on Hessian matrix to classify pixels as belonging to the throat region. The second comparison algorithm was proposed by Liu et al. in [12]. This algorithm employs fuzzy connectedness (FC) approach for upper airway segmentation from MRI. In this work, the seed for initialisation of fuzzy connectedness was provided manually by clicking inside the throat region on each MRI slice. All algorithms were tested on real MRI slices using manual outline as reference segmentation.

The comparison of the ThDA with other two algorithms can be seen in Fig. 6.11. It can be observed that although Hessian analysis method is able to detect more than one region of a throat on same slice compared to ThDA and FC (Fig. 6.11(a)), in other MRI slice (Fig. 6.11(b)), it was not able to detect the region correctly. This was observed for 3 MRI slices from 102. Compared to ThDA and Hessian analysis

algorithms, FC algorithm was sensitive to noise (Fig. 6.11(c)) and also manual initialisation. Further, FC is a segmentation method where throat region detection was performed manually.



(a) (b) (c)

Fig. 6.11: Qualitative comparison of algorithms for throat region detection (white outline: manual, red: ThDA, yellow: Hessian analysis, green: fuzzy connectedness (FC)).

F-measure graph in Fig. 6.12 demonstrates that the ThDA performed similar to manual outline compared to other two algorithms. The mean F-measure value for ThDA, Hessian analysis and FC are 0.77, 0.70 and 0.69 respectively and standard deviation values are 0.07, 0.10 and 0.09 respectively. The lowest value of F-measure (0.57) for Hessian analysis can be observed for Patient 4 where the algorithm fails to detect the desired region for two slices and thus F-measure value is zero for both slices which ultimately reduce the average value. The low values (between 0.6-0.65) of FC can be observed for some patients (Patient 1, 3, 4, 8 and 10), as due to noise FC algorithm was not able to segment the entire region.

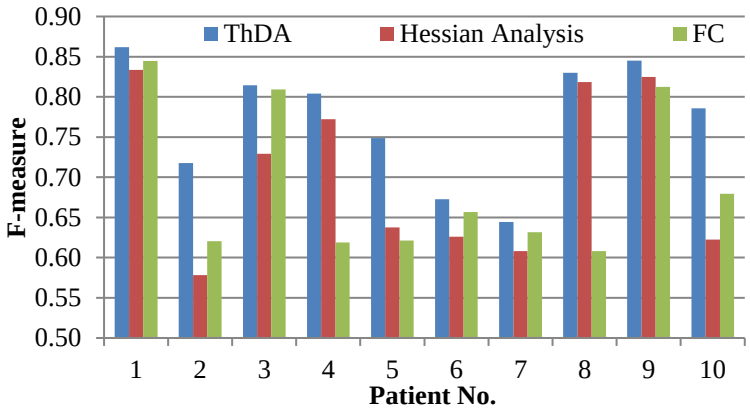


Fig. 6.12: Quantitative comparison of algorithms for throat region detection. ThDA performs better with mean F-measure of 0.77 on given MRI dataset compared to Hessian analysis (mean F-measure: 0.70) and fuzzy connectedness (FC) (mean F-measure: 0.69).

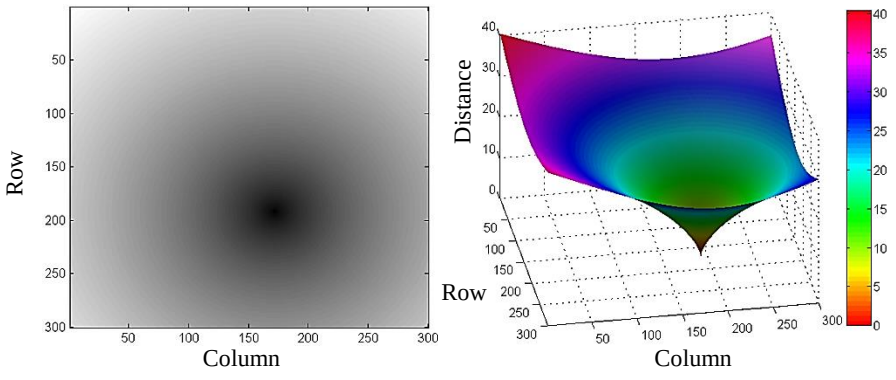
6.4 Modified fuzzy c-means (MFCM) clustering for a tumour region extraction

In this work, intensity and positional (spatial) information of the expected tumour region is utilised to separate the tumour region from rest of the anatomical structures present on a MRI slice. As discussed in Section 3.5.2, standard FCM [13] clustering is a widely used technique to segment a region of interest using pixel intensity (gray-value) information. Based on pixel intensity, FCM categorise pixels into different clusters. However, this technique does not use any other information of a pixel such as spatial or textural to assign a pixel to different clusters. Thus, to improve the capability of this method in order to extract the tumour region, information about the spatial location of the expected tumour is added in its objective function (Eq. 3.7).

As the expected tumour region is adjacent to the throat region, to embed this information in FCM, the average pixel of the throat region is calculated and it's x and y coordinates are saved as t_x and t_y respectively. Further, the squared Euclidean distance from the average pixel (t_x, t_y) of the throat region (Fig. 6.13) is calculated as follows:

$$d_s^2(x,y) = (x-t_x)^2 + (y-t_y)^2 \qquad x = 1..N_x, \quad y = 1..N_y \qquad 6.8$$

This distance matrix is converted into one dimensional column-wise vector $d_s(k)$. This vector is normalised and replicated number of cluster times and saved as $d_s(i,k)$. Thus, dimension of $d_s(i,k)$ is CxN.



(a) (b)

Fig. 6.13: Distance measure created from the average point (co-ordinate) of the throat region as additional information in MFCM clustering (a) 2D profile (b) 3D surface.

The rationale for using this additional distance measure is that the probability of being a tumour pixel is proportional to its distance to the throat region pixels. Thus, addition of this distance measure in the objective function of standard FCM leads to assigning a high membership values to the pixels when the distance to the throat region pixels is small and vice versa. Thus $d_s(i,k)$ is added to objective function

(Eq. 3.7) to form a new objective function of MFCM clustering [11] as follows:

$$J_{ST} = \sum_{i=1}^{C}\sum_{k=1}^{N} \mu_{ik}^{m} d^2(v_i, I_k) + \sum_{i=1}^{C}\sum_{k=1}^{N} \mu_{ik}^{m} d_s{}^2(i,k) \qquad 6.9$$

Here, membership function μ_{ik} is subject to same constraints as in Eq. 3.8.

The optimisation of this function (Eq. 6.9) is basically minimising this distance function as:

$$\min_{U,\{v_i\}_{i=1}^{C}} J_{ST} \qquad 6.10$$

The objective function J_{ST} can be minimised in a fashion similar to the standard FCM algorithm [13]. Thus, taking the first derivatives of J_{ST} with respect to μ_{ik} and v_i and setting them to zero, results in two necessary but not sufficient conditions for J_{ST} to be at local exterma. These two necessary conditions are derived as follows.

For each term (pixel) the optimisation will be solved using one Lagrange multiplier:

$$L = \sum_{i=1}^{C}\sum_{k=1}^{N} \mu_{ik}^{m} d^2(v_i, I_k) + \sum_{i=1}^{C}\sum_{k=1}^{N} \mu_{ik}^{m} d_s{}^2(i,k) + \sum_{k=1}^{N} \lambda_k \left(1 - \sum_{i=1}^{C} \mu_{ik}\right) \qquad 6.11$$

So taking the derivative of L with respect to μ_{ik} and setting the result equal to zero, we get for $m > 1$:

$$\frac{\partial L}{\partial u_{ik}} = m u_{ik}^{m-1} d^2(v_i, I_k) + m u_{ik}^{m-1} d_s^2(i,k) - \lambda = 0 \qquad 6.12$$

Solving for μ_{ik} we have:

$$u_{ik} = \left(\frac{\lambda}{m(d^2(v_i, I_k) + d_s^2\ (i,k))} \right)^{\frac{1}{m-1}} \tag{6.13}$$

Since $\sum_{i=1}^{C} \mu_{ik} = 1$

$$\sum_{i=1}^{C} \left(\frac{\lambda}{m(d^2(v_i, I_k) + d_s^2\ (i,k))} \right)^{\frac{1}{m-1}} = 1 \tag{6.14}$$

Or

$$\lambda = \frac{m}{\left(\sum_{i=1}^{C} \left(\frac{1}{d^2(v_i, I_k) + d_s^2(i,k)} \right)^{\frac{1}{m-1}} \right)^{m-1}} \tag{6.15}$$

Substituting this value of λ in Eq. 6.13, the membership function μ_{ik} can be rewritten as:

$$\mu_{ik} = \frac{(d^2(v_i, I_k) + d_s^2)^{-1/(m-1)}}{\sum_{j=1}^{C} (d^2(v_j, I_k) + d_s^2(j,k))^{-1/(m-1)}} \tag{6.16}$$

In similar fashion, taking the derivative of L with respect to v_i and setting the result equal to zero and solving for v_i we get:

$$v_i = \frac{\sum_{k=1}^{N} \mu_{ik}^m I_k}{\sum_{k=1}^{N} \mu_{ik}^m} \tag{6.17}$$

Thus membership functions μ_{ik} and cluster centres v_i are updated iteratively using Eq. 6.16 and Eq. 6.17 respectively in MFCM clustering until termination.

One disadvantage of this MFCM clustering is that is does not take into account local spatial information of the pixel and is therefore sensitive to noise. In [161] a spatial FCM algorithm which incorporates local spatial information directly into fuzzy membership functions was proposed as:

$$\mu'_{ik} = \frac{\mu_{ik}^{p} h_{ik}^{q}}{\sum_{j=1}^{C} \mu_{jk}^{p} h_{jk}^{q}} \qquad 6.18$$

where p and q are the control parameters. The spatial function h_{ik} which compensates membership value of noisy pixel by adding membership values of its neighbouring pixel; is defined as $\Lambda_{ik} = \Sigma_{l \in Wn} \mu_{il}$ where Wn is local window (5x5) centred at pixel k. More importance was given to spatial function Λ_{ik} to reduce the effect of noisy pixels and to get better smoothing effect. Thus, for each iteration after calculating the membership functions using Eq. 6.16, the membership functions are updated using Eq. 6.18 to reduce noisy pixels.

Thus, MFCM algorithm [11] for segmenting the MRI slice into different clusters can be summarised in the following steps:

1. Randomly initialise membership matrix U that satisfies the constraints in Eq. 3.8.
2. Compute the centroids of the clusters using Eq. 6.17.
3. Compute the objective function as in Eq. 6.9.
4. Update the membership matrix U using Eq. 6.16 and then Eq. 6.18.

Repeat Steps 2 to 4 till termination. The termination criterion is, $\left\| J_{ST}^{(t+1)} - J_{ST}^{t} \right\| < \varepsilon$ or the maximum number of iterations exceeds T.

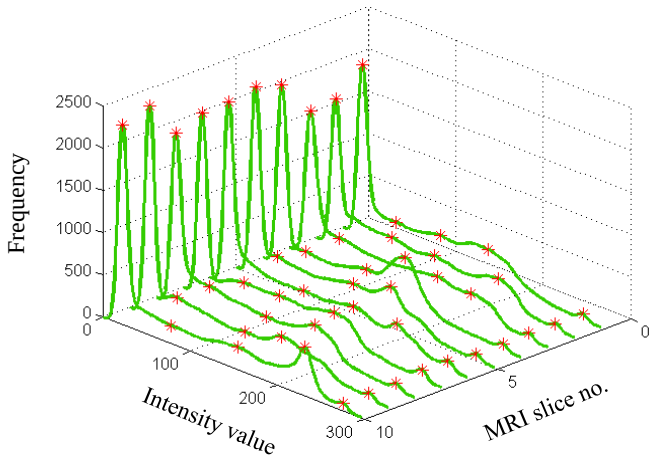


Fig. 6.14: Histogram of representative MRI slice from 10 patients showing 5 peaks (red stars) demonstrating five clusters for each slice.

Using this MFCM clustering technique, the MRI slice is divided into five clusters. The choice for five clusters as described in Section 2.7 and shown in Fig. 6.14 is due to the presence of five main tissue types on axial MRI slice with oropharynx or larynx tumour. The MRI slice and its five clusters are shown in Fig. 6.15. The Cluster 2 in Fig. 6.15 is the cluster with the tumour region and it can be verified that the tumour region have high membership values compared to other tissue type with similar intensities in the same cluster. Other clusters (Cluster 1, 3, 4 and 5) in Fig. 6.15 correspond to other tissue types (like fatty tissues, normal tissues, background) present in the slice. It is known that gadolinium-enhanced tumour pixels occupy the higher end of the histogram [114]. This is utilised to select the cluster with the expected tumour regions. To select cluster with expected tumour region, these five clusters are arranged in the descending order of cluster centres (intensities) and the second cluster is selected as the cluster with tumour region. From selected cluster, pixels with membership value ≥ 0.8 are considered (labelled) as the candidates for the tumour pixels.

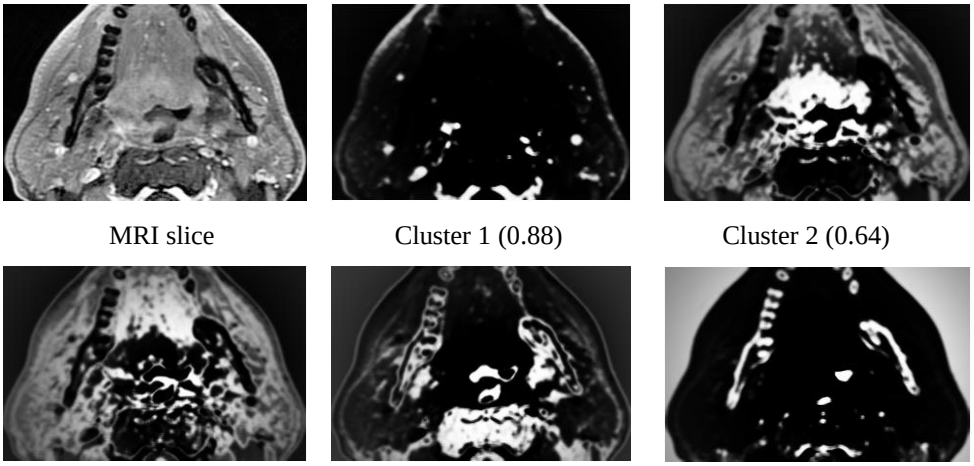


Fig. 6.15: Five clusters (with cluster centres) of MRI slice with visible tumour in Cluster 2. Different tissue types are included in different clusters based on grey-level values.

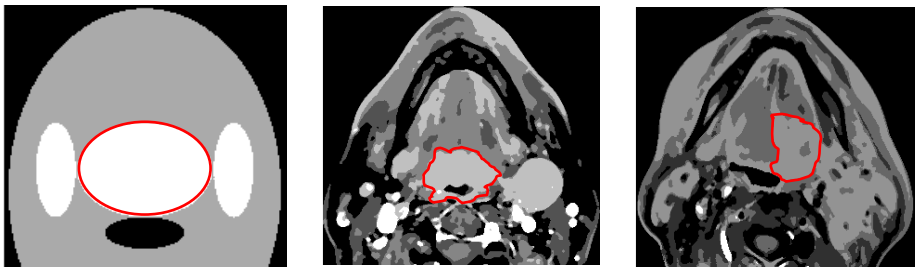
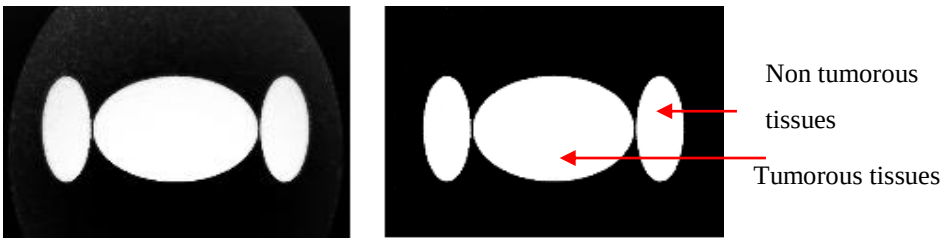


Fig. 6.16: Clusters demonstrated using different grey levels obtained using MFCM from axial MRI slices (first slice: synthetic MRI slice) (red outline: tumour outline). Entire tumour region/s (red outline) is included in a one cluster.

Additional examples of the clusters (represented using different grey levels) obtained using MFCM from synthetic and real MRI slices are shown in Fig. 6.16. From Fig. 6.16, it is clear that MFCM clustering algorithm efficiently divided the slice into appropriate clusters. Further, it can be observed that entire tumour region/s (red outline) is included in a one cluster, thus, selecting that appropriate cluster for tumour region extraction is sufficient. Other clusters can be discarded from further processing for computational efficiency.

6.4.1 Comparison of MFCM clustering with standard FCM clustering

The proposed MFCM clustering technique was compared with standard FCM [13] clustering. For comparison on real MRI data, manual consensus tumour outline was used as reference for MFCM and FCM extracted tumour outline and for quantitative assessment, DSC (Section 4.5) was used as an evaluation metric. For FCM, all the parameters such as the number of clusters, selection of cluster with expected tumour region, or termination criteria were similar to those used in the MFCM. Further, the membership values of noisy pixels in FCM were compensated using Eq. 6.18 as in MFCM. Fig. 6.17 show the cluster with tumour region obtained using MFCM (first column) and standard FCM (second column) clustering techniques.



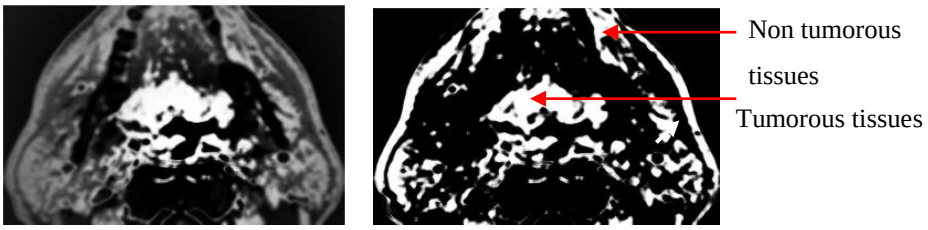


Fig. 6.17: Cluster with tumour region obtained using (a) MFCM (b) standard FCM. High membership values for the pixels near the throat region can be observed in MFCM clustering compared to pixels away from the throat region.

In standard FCM, non-tumorous tissues (or pixels) away from the throat region are assigned high membership values similar to the expected tumorous tissues (pixels) near the throat region as they have similar intensity range. Thus, it is difficult to separate the candidate pixels for the tumour regions from the cluster using simple technique like thresholding. However, in MFCM technique, due to the additional distance measure, non-tumorous tissues away from throat region are assigned low membership values. Thus, candidate pixels for the tumour regions can be obtained using a single threshold value.

The outlines of the expected tumour regions obtained from MFCM and FCM clustering is shown in Fig. 6.18. It can be noted that no post-processing steps such as morphological filtering or noise removal techniques were applied to the selected cluster. The candidate for the tumour pixels from the selected cluster were obtained only by considering pixels with membership values greater than certain threshold. From Fig. 6.18, in general MFCM extracted tumour outline, appear comparable to the manual outline than FCM extracted tumour outline. In particular, it was observed that if the contrast (intensity difference) between expected tumour region and background was good (>0.15), MFCM and FCM gave similar results (Fig. 6.18(a)). However, for the region with poor contrast (Fig. 6.18(b)) or where the region is fused into nearby normal structures (Fig. 6.18(c)), MFCM performed better than FCM, though, highly heterogeneous intensity distribution of the expected tumour region can be influence the accuracy of MFCM results.

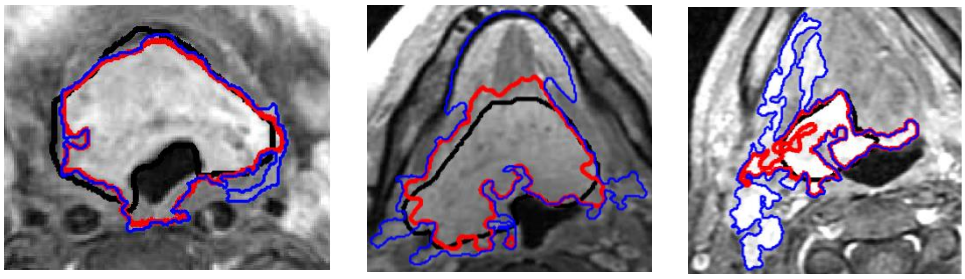


Fig. 6.18: Tumour extraction results on real MRI slices (black outline: reference (manual), red: MFCM, blue: standard FCM). MFCM clustering extracted results close to manual outline compared to standard FCM clustering.

Further analysis in terms of DSC between MFCM, FCM and manual tumour outline is shown in Fig. 6.19. The box plot in Fig. 6.19 demonstrates the distribution of DSC for MFCM with minimum value of 0.61, maximum of 0.79 and median of 0.69 and for FCM with minimum value of 0.57, maximum of 0.75 and median of 0.66 for 10 patients (102 axial slices). Thus, it was found that FCM clustering technique has demonstrated the ability to the extract the tumour region from all slices with mean DSC of 0.66. However, additional spatial information in terms of distance measure of the expected tumour region further increases the capability of FCM method to extract the tumour region and thus, increases the mean DSC to 0.70.

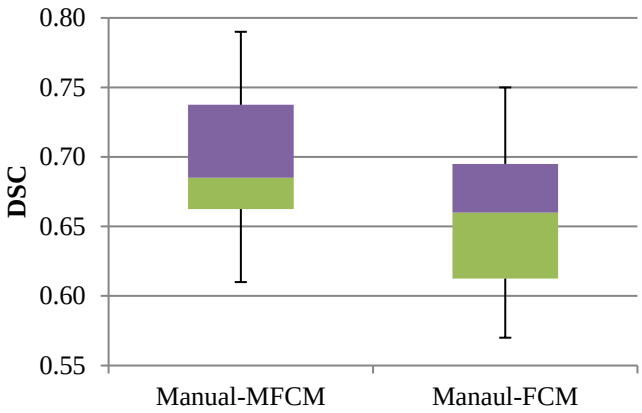


Fig. 6.19: Box plot showing distribution of DSC values for tumour extraction results comparing MFCM and standard FCM clustering techniques with manual tumour outlines.

6.5 Conclusion

In this chapter novel techniques, that are incorporated in the segmentation framework described in Chapter 5 have been presented to reduce IIH and to efficiently extract the

tumour/s (GTV) from T1+Gd axial MRI slice. Initially, a new technique has been presented to determine the placement of knots using derivative and filtering technique for the bias field (IIH) estimation using bicubic spline. This method allowed the reduction of complex bias field (IIH) present in MRI slices used in this work. Then a novel algorithm (ThDA) has been proposed that combines the spatial information in terms of fuzzy rules to detect the throat region with complex geometry and presence of pathology near the region. Finally, FCM clustering has been implemented with an additional spatial constraint (MFCM) to extract the tumour region. The performance of these algorithms has been demonstrated on the synthetic and real MRI dataset and efficiency of these algorithms is validated by comparing it with currently existing techniques. The next chapter (Chapter 7) presents the results of segmentation and quantification framework on real and synthetic MRI dataset.

Chapter 7
Qualitative and quantitative analyses of segmentation and quantification framework

7.1 Introduction

In the previous two chapters, the framework (PLCSF) and novel methodologies for obtaining the outlines of oropharynx and larynx tumours from MRI data were presented. In this chapter an empirical evaluation of the developed PLCSF on MRI data (Chapter 4) is described. This includes assessment of PLCSF on synthetic MRI data in presence of artefacts to validate the robustness of the algorithm; comparison of automatic results with manual outlines from different experts to evaluate the accuracy and reproducibility of the framework; and comparison of PLCSF outlines with MS and spectral clustering on real MRI dataset. These comparison algorithms were chosen because they are representative of current methods in tumour analysis [112]-[114], and implementations of the algorithms have been made available by their authors.

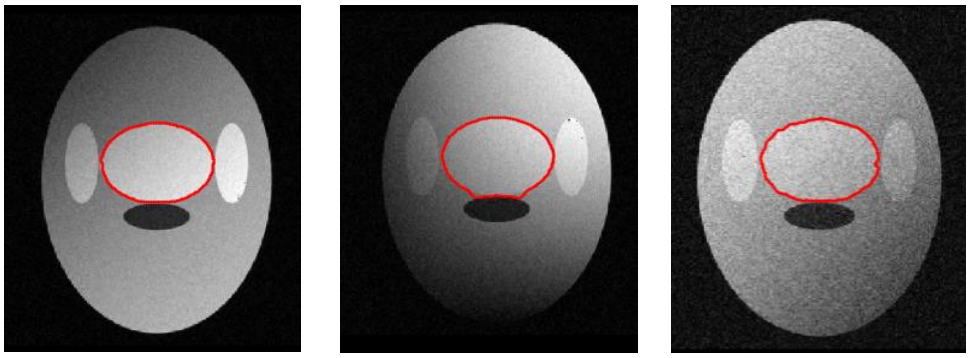
The remainder of this chapter is organised as follows. Section 7.2 describes PLCSF results on synthetic dataset with results on real dataset in Section 7.3 and 3D reconstruction and volume measurements results in Section 7.4. The discussion of the results in turn is presented in Section 7.5. Section 7.6 concludes the chapter.

7.2 PLCSF results on synthetic MRI dataset

The performance of the PLCSF was evaluated on 27 synthetic (Section 4.4) MRI slices. For synthetic data, the reference region of the tumour was obtained by applying the Otsu threshold [155] to the reference slice with no artefacts (Fig. 4.4(a)) and selecting the brightest intensity region near the throat. Canny edge detector [54] was employed to obtain the outline of the selected reference region. Variations in reference tumour outline and PLCSF tumour outlines were quantified using DSC and MHD (Section 4.5).

The visual results of the segmented tumour region (red outline) using PLCSF

on three synthetic MRI slices is illustrated in Fig. 7.1. It can be noticed that the PLCSF obtains accurate results for synthetic data and does not include healthy salivary glands (vertical ellipses near the tumour region ellipse) in the segmentation results even if they have similar intensities and are close to the tumour region (Fig. 7.1(a), (b) and (c)). However with increase in high noise (σ:15), PLCSF deviates from obtaining the smooth tumour boundary (Fig. 7.1(c)), thus under segmenting the tumour region.



(a) (b) (c)

Fig. 7.1: PLCSF results on synthetic MRI slices (red outline). Notice that even in presence of noise, IIH and low contrast between tumour and background, the PLCSF obtained accurate outline of a tumour region.

Table 7.1: Modified Hausdorff distance (MHD) values in mm for 27 synthetic MRI slices. Low values of MHD (mean: 0.16 mm) indicate high accuracy of PLCSF in terms of shape of the outline when compared with reference outline in presence of artefacts.

Noise / IIH	5	10	15
Sin	0.05	0.04	0.11
Cos	0.06	0.12	0.14
Sin+Cos	0.14	0.16	0.16

Contrast:0.5

Noise / IIH	5	10	15
Sin	0.05	0.12	0.19
Cos	0.13	0.11	0.19
Sin+Cos	0.16	0.19	0.20

Contrast:0.3

Noise / IIH	5	10	15
Sin	0.06	0.04	0.23
Cos	0.27	0.17	0.28
Sin+Cos	0.25	0.15	0.39

Contrast:0.2

Table 7.1 shows MHD values comparing PLCSF and reference outlines for all synthetic slices. This comparison yields a mean MHD value of 0.16mm that indicates high accuracy of PLCSF in terms of shape of the tumour outlines. Further high value of DSC (minimum: 0.96) demonstrates high spatial agreement of the PLCSF with

reference outline for low contrast (0.2), high noise (σ:15) and high IIH (sin+cos) synthetic MRI slices.

7.3 PLCSF results on real MRI dataset

The PLCSF tumour outlines obtained from real MRI data were compared to the manual outlines obtained from two CO (C1 and C2) and the consensus manual outlines (Section 4.3). For comparison, tumour outlines obtained using PLCSF was referred as C-auto. The consensus manual outline was used as a gold standard. The gold standard manual set of tumour outlines was referred as C-gold. The first manual set of tumour outlines contoured by C1 and C2 was referred as C1-A and C2-A respectively. The second set of outlines was referred as C1-B and C2-B respectively. The inter-variability was assessed between C1 and C2 and intra-variability between C1-A and C1-B and between C2-A and C2-B. The PLCSF were also compared to two other algorithms (MS clustering [82] and N-cut [83]) on real MRI data using C-gold as reference segmentation for all three algorithms.

7.3.1 Time comparison

The PLCSF was implemented in Matlab 2014a (Mathworks, Natick, MA), on Windows 7 system. The Matlab software was executed on a Dell U2412M machine with 8GB RAM. The average tumour segmenting time using PLCSF in one patient was 8 minutes (range 6-13 minutes, mean 45 seconds per slice) compared to 15 minutes (range 6-23mins; mean 88 seconds per axial slice) for C1 contours. It was observed that it takes less time to obtain C1 for the laryngeal tumours (mean 80 seconds per slice) than for the oropharyngeal tumours (mean 96 seconds per slice).

7.3.2 Inter- and Intra- variability calculations between manual outlines

Inter- and Intra- variability was calculated between the segmentations from experts (C1 and C2) in order to quantify the difficulty of the segmentation task. Further, results of this evaluation also serve as a baseline to compare PLCSF result against in the next section (Section 7.3.3). Inter-operator variation between C1 and C2 demonstrated the mean DSC of 0.80 and mean MHD of 2.24mm (Table 7.2). The maximum agreement (DSC: 0.87) between C1 and C2 was observed for Patient 8 (Fig. 4.1) where tumour region was clear and distinct and the minimum agreement for Patient 10 (DSC: 0.69)

as C1 included a throat region in the tumour outlines, while C2 avoided the inclusion of a throat region in tumour outlines. There is no particular guideline as to include or exclude the throat region from tumour outlines.

Table 7.2: Inter- and Intra- variability in manual (C1 and C2) tumour outlines (mean (standard deviation)). Inter- variability in manual DSC (mean: 0.80) indicate difficultly of a segmentation task.

Inter- and Intra-variability calculations	DSC	MHD (mm)
C1 vs. C2	0.80 (0.06)	2.24 (0.80)
C1-A vs. C1-B	0.85 (0.04)	1.68 (0.70)
C2-A vs. C2-B	0.86 (0.03)	1.39 (0.32)

Further, a PCC (Section 4.5) was obtained to assess the strength of linear relationship between the tumour areas obtained by C1 and C2. A PCC indicate a strong positive linear correlation between C1 and C2 tumour areas with r (100): 0.86 and ρ -value < 0.05. This ρ -value suggests that there is no significant difference in C1 and C2 tumour areas. However, C1 over-estimates the tumour region compared to C2 due the practice to draw the tumour border just outside the observed edges rather than drawing them on the edge.

The intra-variability between C1-A and C2-B in terms of mean DSC and MHD is 0.85 and 1.67mm respectively and between C2-A and C2-B, DSC is 0.86 and MHD is 1.39mm.

7.3.3 Comparison of PLCSF results with manual segmentations

The visual comparison of the segmentation results between C-auto and C-gold (consensus manual outline) for each patient for a single axial MRI slice is demonstrated in Fig. 7.2. More qualitative results between C-auto and C-gold tumour outlines (for 30 axial slices) are demonstrated in Appendix B.1. Further, visual comparison of C-auto with all five manual outlines (C-gold, C1-A, C1-B, C2-A and C2-B) for a single axial MRI slice from each of 10 patients is demonstrated in Appendix B.2.

In Fig. 7.2: Patients 2 and 3 demonstrate that the PLCSF system can effectively segment small tumour regions (area: 89.97-246.67mm^2), although slight over segmentation can be observed. Other examples in Fig. 7.2 such as Patients 1, 6 to 10 demonstrate variability in C-gold and C-auto outlines. For some MRI slices (Fig. 7.2: Patients 1 and 7), where the tumour region depicts intensity range close to the surrounding tissues and the non-distinct boundary between surrounding and tumourous tissues leads to a slight over segmentation by PLCSF. Inclusion or exclusion of the throat region in the outline may also influence the segmentation results as illustrated in Fig. 7.2: Patient 8. A variation in the segmentation results (DSC: 0.65) between C-gold and C-auto outline is shown in Fig. 7.2: Patient 6. This variation is mostly observed in MRI slices where the tumour region is small and often fused into nearby normal structures, thus exhibit low contrast between tumour and surrounding tissues (Table 4.1, Patient 6). For example in Fig. 7.2: Patient 6, PLCSF overestimated tumour region by including normal palatine tonsils in the outline. This low performance result obtained on the non-enhancing (low contrast) case, reveals the limitation of the framework by application to the tumours that are enhanced with the contrast agent.

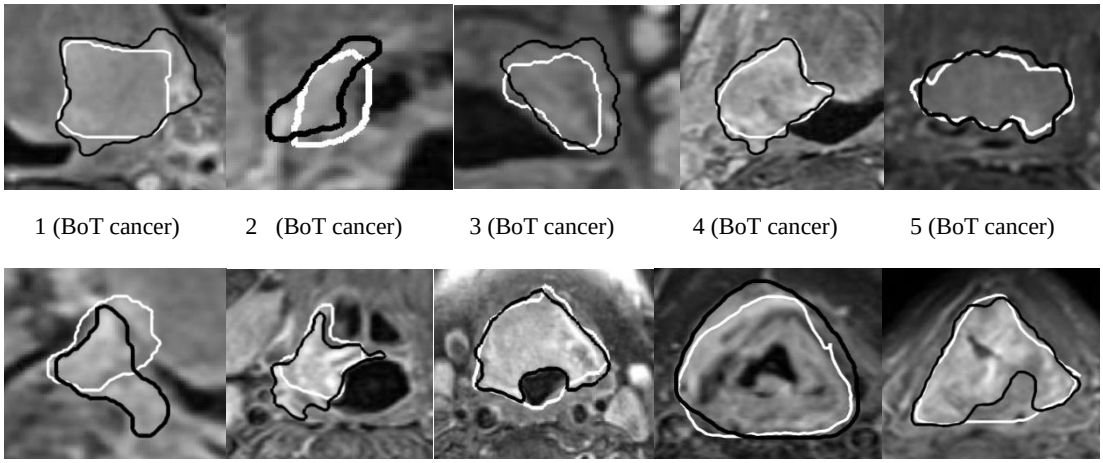


Fig. 7.2: Examples of C-auto (PLCSF) (black outline) on small tumour regions or challenging cases, superimposed with the C-gold (white outline) from real MRI dataset (Table 4.1 and Fig.

4.2) from each of 10 patients. PLCSF can effectively segment small and challenging tumour regions.

The results for laryngeal tumours, Fig. 7.2: Patients 9 and 10, demonstrate less agreement (DSC: 0.7-0.8) between C-gold and C-auto results as compared to BoT tumour, Fig. 7.2: Patient 4 and 5, which has a DSC: >0.9. The possible reason for this bias is the distinct characteristics of tumour tissues as compared to surrounding tissues in BoT tumours than laryngeal tumours. Additionally, large pixel spacing and thus low resolution (Table 4.1, Patient 7 and 9) can increase partial volume pixels and affect the segmentation (over-segment) of the tumour regions.

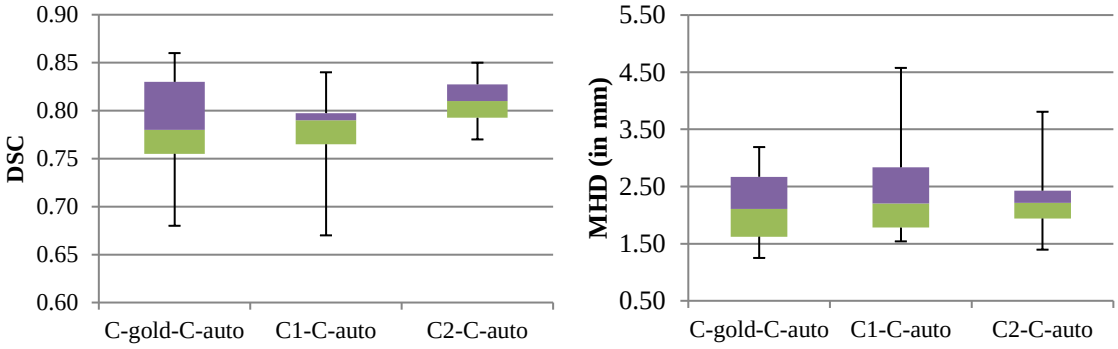
Table 7.3: Quantitative measures between manual outlines (C-gold, C1 and C2) and PLCSF outlines (C-auto) (mean (standard deviation))

Comparison	DSC	MHD (mm)
C-gold vs. C-auto	0.79 (0.05)	2.15 (0.64)
C1 vs. C-auto	0.77 (0.05)	2.48 (0.90)
C2 vs. C-auto	0.81 (0.03)	1.97 (0.46)

The performance measures calculated between the manual outlines from different experts and PLCSF are shown in Table 7.2. The mean DSC of 0.79(0.04) and MHD of 2.20(0.66) mm for all 10 patients indicate acceptable (Section 4.5) agreement between manual segmentation results and PLCSF. The standard deviations show the extent of performance of PLCSF for different patients. In some MRI slices, small DSC and large MHD values are due to the shape variation between manual and automatic results, as the PLCSF segments concavity region in detail than manual results (Fig. 7.2: Patient 10). These values are consistent with comparison of manual (C1 and C2) segmentation results (Section 7.3.2).

Further, a one-sample t-test was run to determine whether a mean DSC (0.79) obtained from comparison of C-gold-C-auto was different to a mean DSC obtained from C-gold-manual comparison defined as a reference DSC of 0.79 (Appendix B.3) to assess if PLCSF and manual segmentation from expert obtained similar results. The test reveal that mean C-auto-C-gold DSC (0.79, 95% CI: 0.75 to 0.82) was not different from the reference DSC of 0.79 with t (101): -0.23, ρ -value: 0.82.

The box plots demonstrating distribution (range) of DSC (Fig. 7.3(a)) and MHD (Fig. 7.3(b)) comparing C-auto and C-gold, C1 and C2 respectively shown in Fig. 7.3 indicate that there are no outliers in region and distance based evaluation metrics. Moreover, from Fig. 7.3 it also clear that C-auto performed worst compared to C1 than C-gold and C2 with long whiskers at the bottom of DSC and at the top of MHD indicating low DSC and high MHD values.



(a) (b)

Fig. 7.3: Box plots showing distribution range of (a) DSC (b) MHD between manual (C-gold, C1 and C2) and PLCSF (C-auto) results.

The PCC between C-auto and the C-gold, is 0.89 with ρ-value <0.05 (Fig. 7.4) verifying statistically significant results in terms of tumour areas. Similarly, PCC between C-auto tumour areas and C1 and C2 tumour areas is 0.81 and 0.90 respectively with ρ-value <0.05. The PCC between C-auto and C-gold for BoT tumours is 0.95 as compared to 0.83 for laryngeal tumours indicating that PLCSF performs better for BoT tumours than for larynx tumours.

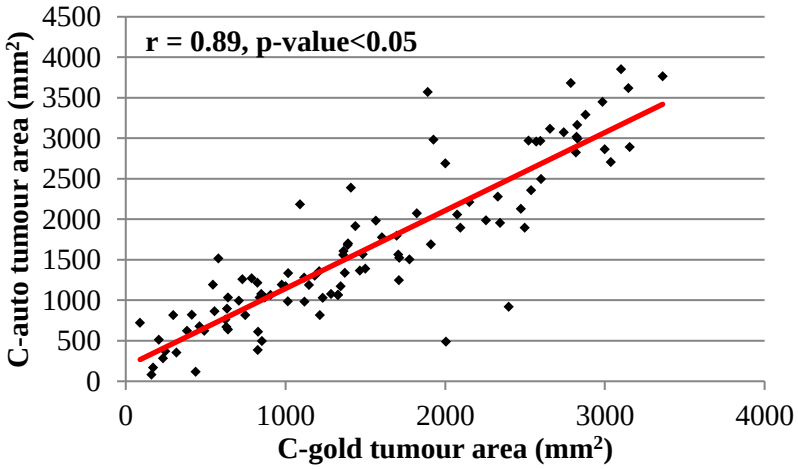


Fig. 7.4: Correlation analysis between PLCSF and C-gold cancer areas shows strong positive correlation for a given dataset

A one-sample t-test revealed that there was no statistical significant difference between mean C-auto tumour area (1633.22mm^2, 95% CI, 1444.3 to 1822.1 mm^2) and the mean reference (C-gold) area of 1507.37 mm^2 with t (101): 1.32, ρ -value 0.189.

7.3.4 Comparison of PLCSF with MS clustering and Ncut

The PLCSF was compared to two other algorithms in the literature used for tumour segmentation. One was mean-shift (MS) clustering [82] and other was spectral clustering (Ncut) [83] technique. Overview of MS clustering and NCut technique is provided in Section 3.5.2. These clustering (unsupervised) algorithms were chosen for comparison as they are recently used to segment contrast-enhanced tumour region from MRI data [112]-[114]. The hypothesis for this comparison was that FCM clustering is better compared to MS and spectral clustering for the dataset used in this work. Thus, to verify this hypothesis for both algorithms (MS and Ncut), other steps of the algorithm were similar to PLCSF (Fig. 5.1), only MFCM clustering was replaced by MS or Ncut clustering technique. Thus, the input for both algorithms was the pre-processed MRI slice obtained from Section 5.3.1. From the clusters (partition) obtained from both algorithms only the bright (intensity-wise) cluster near the throat region was further processed using the techniques described in Section 5.3.3. Parameters for both algorithms were chosen so as to give the best performances across the MRI dataset. For implementation of MS clustering and Ncut, we used publicly available Matlab codes [162], [163].

Qualitative results in Fig. 7.5 show comparable performance of PLCSF (C-auto) with C-gold (Fig. 7.5(e), (f)) compared to Ncut and MS clustering algorithms. For Ncut algorithm, it is well known that this clustering is biased towards partitioning the image into equal segments. This bias affects the segmentation results specifically in case of small tumour regions (Fig. 7.5(a)). For MRI cases with low variation between tumour and surrounding tissues, Ncut shows low performance due to segmentation leakage (Fig. 7.5(b)). Furthermore, this algorithm is computationally intensive. The average time required to obtain a tumour outline per axial slice using Ncut is 82 seconds compared to mean of 45 seconds for the PLCSF.

The results we observed with MS clustering shows that it tends to over-segment the MRI slice giving large number of clusters rather than finding the right object (Fig. 7.5(c), (d)). Furthermore, the varied number of clusters obtained by MS clustering for each MRI slice makes it challenging to select cluster with candidate tumour pixels. The MFCM clustering used in this work lack this issue as the number of clusters was fixed.

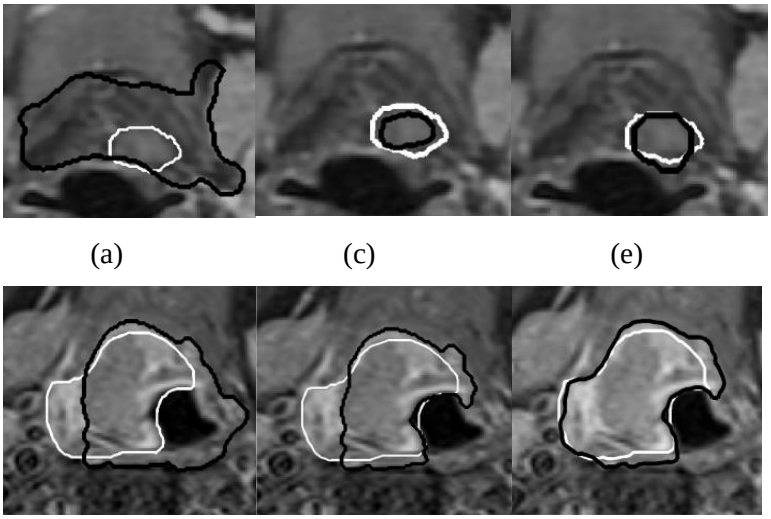


Fig. 7.5: Qualitative comparison of segmentation results. Results of proposed PLCSF framework (column 3 – black outline) are qualitatively more similar to the C-gold (consensus manual segmentation) (white outline) as compared to Ncut (column 1 – black outline), and mean-shift clustering (column 2 – black outline).

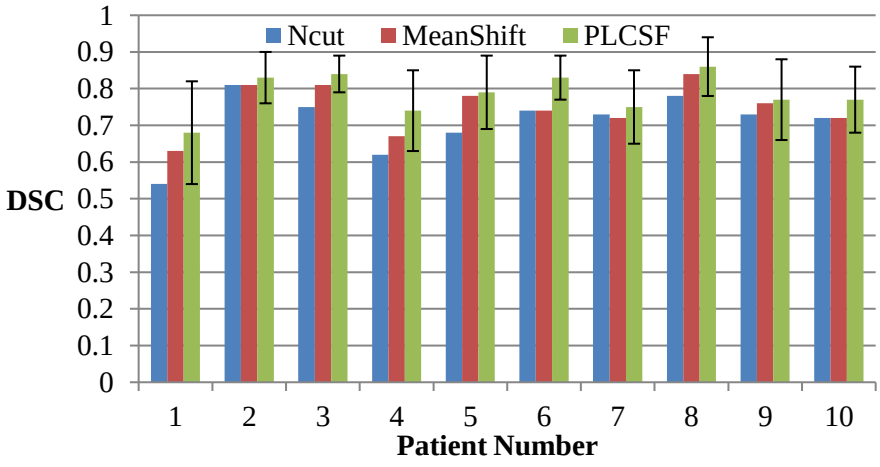


Fig. 7.6: Comparison of DSC for Ncut, Mean-shift and the proposed method (PLCSF) on real MRI data set, demonstrates improved spatial overlap with the PLCSF over this relatively more challenging tumour data. Black vertical bar indicates ± standard deviation for MRI slices for each patient for PLCSF.

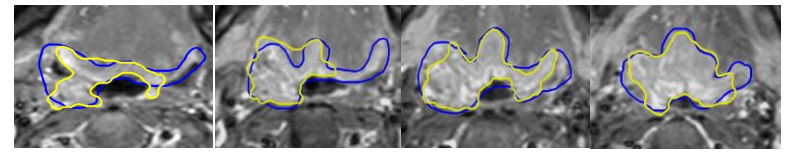
DSC comparison results for these three methods for all 10 patients to the C-gold (consensus manual outline) as a reference are shown in Fig. 7.6. This comparison yielded mean DSC of 0.71 for Ncut, 0.75 for Mean Shift and 0.79 for proposed PLCSF algorithm. The PLCSF algorithm exhibits a lower CoV (0.07) on average for DSC compared to Ncut (0.11) and MS (0.09) clustering. Paired t-test comparison of MS clustering and PLCSF demonstrates statistically significant improvement in DSC (ρ -value: 1.44 x 10^{-3}) with the PLCSF method.

7.4 3D reconstruction and volume calculation results

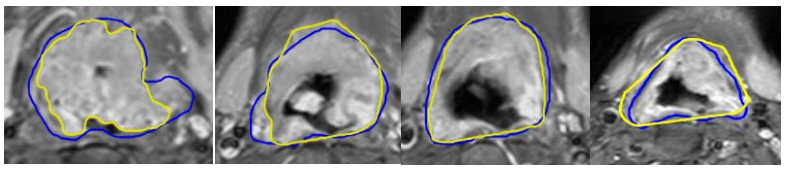
All tumour (GTV) outlines (C1-A, C1-B, C2-A, C2-B, C-gold and C-auto) were reconstructed into 3D and smoothing was applied to each reconstructed region. Further, volume (in cm^3) was calculated for each reconstructed tumour and volume concordance (agreement) C_{vol} was calculated between C-gold and C-auto, between C1 and C-auto and between C2 and C-auto as:

$$C_{vol}(\%) = (1 - \frac{|T_{volauto} - T_{volref}|}{T_{volref}}) * 100 \qquad 7.1$$

where $T_{volauto}$ was C-auto volume T_{volref}, was the reference (manual) volume either from C-gold, C1 and C2.

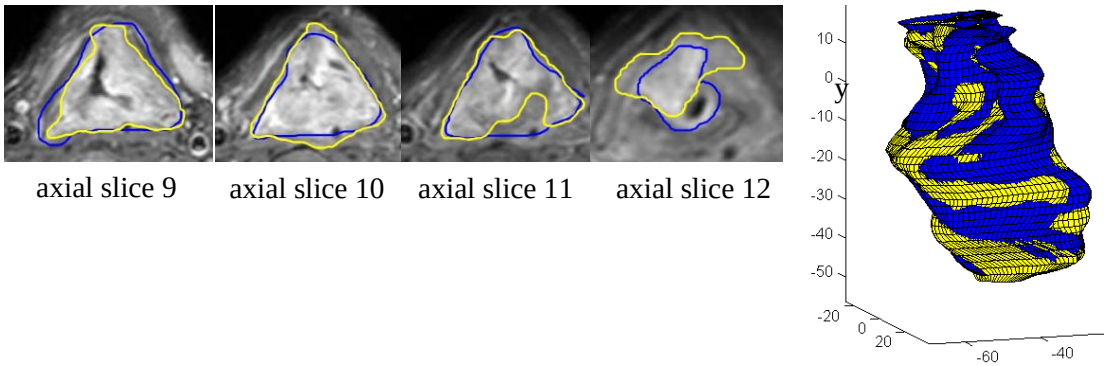


axial slice 1 axial slice 2 axial slice 3 axial slice 4



axial slice 5 axial slice 6 axial slice 7 axial slice 8

z



(a) (b)

Fig. 7.7: (a) Sequential axial MRI slices with C-auto (yellow) and C-gold (blue) tumour outline for one patient; (b) 3D reconstructed volume from tumour (GTV) outlines shown in Fig. 6.13(a) (C-auto reconstructed volume - yellow; C-gold reconstructed volume - blue).

An example of tumour outlines for the axial slices for one patient comparing C-gold and C-auto is shown in Fig. 7.7(a). Spatial discontinuity in between tumour outlines from neighbouring slices (for instance between axial slice 1 and axial slice 2) can be noticed from this sequential slices from a single patient. This limitation of PLCSF is due to the 2D approach employed in segmentation. This discontinuity is reduced to some extent by automatic smoothing of the 3D reconstructed volume shown in Fig. 7.7(b). A yellow colour in Fig. 7.7(b) represents a volume greater in C-auto than C-gold. It can be observed that automatic system underestimated the tumour region in initial slices (axial slice 1 and 2 Fig. 7.7 (a)) and overestimated in end slices (axial slices 10and 12 in Fig. 7.7(a)). This can be predicated from reconstructed volume (Fig. 7.7(b)) than verifying all 12 axial slices. Another example of 3D reconstructed tumour from C-gold and C-auto outlines are demonstrated in Appendix B.4.

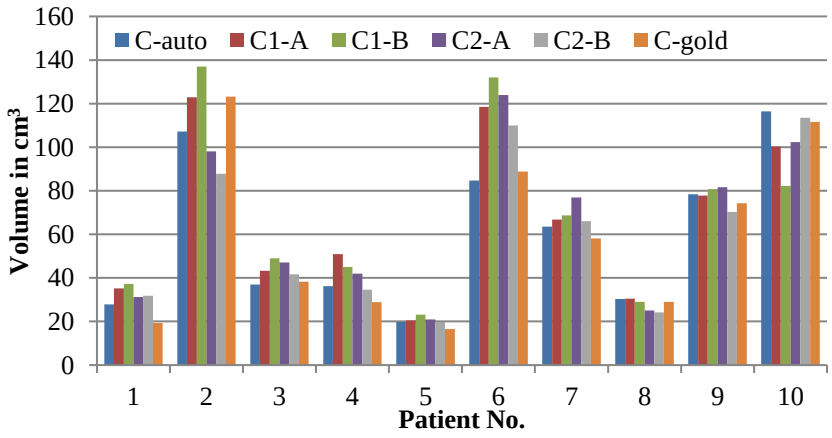


Fig. 7.8: Volume of reconstructed tumour (in cm^3) for all 10 patients. C-auto over estimated volume (mean: 60.15 cm^3) compared to C-auto (mean: 58.79 cm^3)

The volume (in cm^3) for each 3D reconstructed tumour is demonstrated in Fig. 7.8. Volumes ranged from 16.48 cm^3 to 136.93 cm^3. The mean volume concordance between C-gold (mean: 60.15 cm^3) and C-auto (mean: 58.79 cm^3) was 86.51%, however, C-auto slightly overestimated the tumour volume as C-auto over segmented the tumour region on 2D slices in case of non-enhancing tumour regions. The mean volume concordance between C1 and C-auto was 83.63% and between C2 and C-auto was 87.51%. This indicates that automatic volumes are comparable to the manual volumes.

7.5 Discussion

The validation of automatic segmentation results is a difficult task as there is no ground truth data available either due to variability in manual (reference) segmentation or no common database to test the algorithms. This chapter presented empirical evaluation of developed automatic framework (PLCSF) on MRI data with regard to time saving, robustness, accuracy and reproducibility by comparing the results to the different experts from the clinic and state of the art algorithms.

The synthetic data was segmented very well by the algorithm (PLCSF) (Section 7.3.1), with DSC values (mean: 0.99) that were much higher than those for real data (mean DSC: 0.79). As the synthetic datasets have a high variability in artefacts, but are less variable in tumour shape and location, the results suggest that the PLCSF is robust in the presence of artefacts provided the tumour location is consistent through the dataset.

On the real data the segmentation of a tumour region using PLCSF demonstrate time saving of 46% for a single patient compared to manual segmentation from C1 (Section 7.3.1). This time saving is comparable to the literature [6], [7], [30] that demonstrate the range of 30% to 40%. Further, quantitative comparison with the gold standard (C-gold) on 102 T1+Gd MRI axial slices from 10 patients, the PLCSF (C-auto) shows no significant difference in performance (r: 0.89, ρ -value <0.05) with the method used in a current clinical practice with regard to segmented tumour areas. Also, comparison with other experts (C1 and C2) indicate that DSC (C1-C-auto: 0.77 and C2-C-auto: 0.81) are within acceptable range. There is some discrepancy between the

values of DSC and MHD when comparing C-auto and C-gold outlines. This discrepancy can be contributed to the manual bias of including areas of high risk in the tumour outline (Fig. 7.7(a) axial slice 1-2) or processing each axial slice in a 2D approach by the PLCSF (Fig. 7.7(a) axial slice 6). Despite differences in the axial slices between the C-auto and C-gold results, there was a very strong volume concordance was observed. The limitation of the PLCSF is over-segmentation of a tumour region in case of similar characteristics of tumourous tissues as compared to surrounding tissues. However, this is also the same area where there is greatest clinical and radiological variability.

Using PLCSF, a mean DSC of 0.79, MHD of 2.2mm and volume concordance of 86.51% were obtained. The PLCSF result also demonstrated improved performance when compared to other algorithms (MS clustering (DSC: 0.75) and Ncut (DSC: 0.71)). A semi-automatic approach [103] validated on 16 patients (78 axial slices) with tongue tumours demonstrated a mean percentage match of 87.07% and a mean volume concordance of 92.74%. This technique [103] required manual-placing of seed points in the tumour region or drawing of close loop outside the tumour from expert and have no results to prove any validation on laryngeal tumours.

One of the main purposes of the automatic tumour region segmentation from T1+Gd MRI slices is the reproducibility of the segmentation results that contain intra- and inter- variability from manual segmentation results. For this framework, if the parameters values are unchanged, the system obtains similar results for repeated number of times, indicating the reproducibility of the system. Although our primary results are encouraging, before the PLCSF with 3D reconstruction and volume calculation can be established in clinical setting, further validation with more subjects and rigorous testing would be required. Clinical judgement, using all clinical and diagnostic information, will however always remain vital.

7.6 Conclusion

This chapter presented a validation of newly-developed framework for the analysis of MRI oropharynx and larynx tumours. This framework shows initial acceptable statistically-validated results for tumour segmentation and reconstruction. This

framework may help reduce inter- and intra- variability and can assist CO with time-consuming, complex radiotherapy planning procedure.

Chapter 8

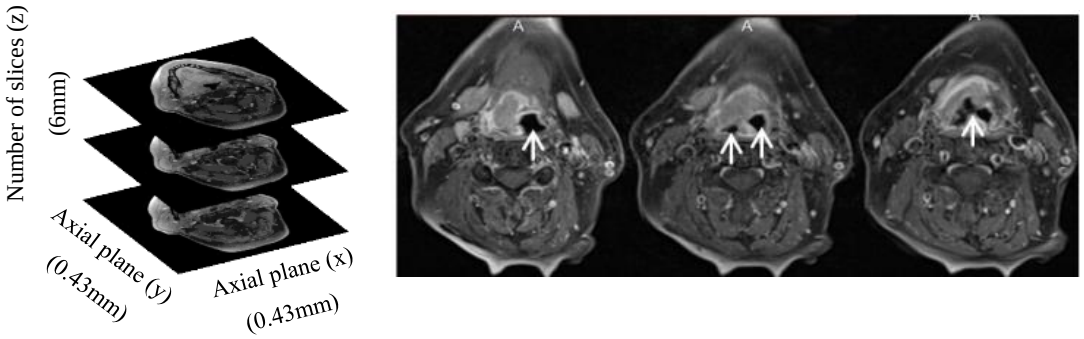
Novel 3D throat region segmentation algorithm based on FI and 3D LSM

8.1 Introduction

The PLCSF discussed in Chapter 5 for the automatic segmentation for oropharynx and larynx tumours is a 2D segmentation (Section 3.2) framework where the tumour region on each MRI slice is segmented separately and then reconstructed in 3D for quantification. However, to utilise more spatial information about a region of interest and to assess the feasibility of 3D segmentation for MRI dataset used in this work, a new 3D segmentation technique (Section 3.2), where all slices are considered together as a one volume for segmentation, is proposed in this chapter. In this chapter the region of interest is a throat region.

Automatic 3D segmentation of a throat region from a T1+Gd MRI volume involves addressing numerous challenges such as presence of artefacts, large spacing in between slices (anisotropic volume) (Fig. 8.1(a)), high variability in the shape due to presence of pathology near the region (Fig. 8.1(b)) and the connectivity of the region to other channels of air. Thus, a new robust technique (I+3DLSM) that reduces inherent MRI artefacts, reconstructs isotropic volume and obtains a smooth 3D structure of the throat region is proposed [77]. The novelty of this work lies in a more generalised approach that performs FI prior to the 3D LSM segmentation. This coupling of interpolation with LSM makes segmentation more reliable and accurate [164]. Furthermore, the automatic initialisation for LSM eliminates the need for manual interaction and utilises a simple data term in LSM due to high contrast between the intensity range of the throat region and its surrounding tissues.

The remainder of this chapter is organised as follows. Section 8.2 and Section 8.3 provide the overview of N-dimensional FI and 3D LSM respectively. 3D throat region segmentation algorithm (I+3DLSM) is discussed in Section 8.4, while Section 8.5 demonstrates the experimental results of I+3DLSM on synthetic and real MRI volumes. Finally, discussion is reported in Section 8.6 and conclusion in Section 8.7.



(a) (b)

Fig. 8.1: Axial slices from a MRI volume (a) demonstrating anisotropic voxel size in x-, y-, and z- dimensions (b) demonstrating change in topology (splitting and merging) of the throat region (white arrows) [164] © 2016 The Institution of Engineering and Technology.

8.2 m-dimensional (m-D) Fourier interpolation (I)

Section 3.4 discussed the FI in one dimension. This section describes FI in desired dimension/s for an $N-D$ array. As discussed in Section 3.4, FI [79] effectively employs the FFT version of the Discrete Fourier Transform (DFT) to interpolate data in spatial domain through sinc basis functions. The DFT and inverse DFT operations are separable meaning that a multi-dimensional transform can be decomposed into multiple one-dimensional transforms allowing the interpolation technique to be applied along specific dimension/s. Thus, the method for FI across specified dimension/s of $N-D$ array g of size $N_1 x N_2 x ... x N_m$ can be explained in three steps. Following are the steps to perform FI along the third dimension while no changes are made in first and second dimensions.

1. M_3-point FFT of input array g, given as G, is computed along the third dimension. The DFT across the third dimension of 3-D data ($m=3$) is calculated as:

$$G(n_1,n_2,K_3) = \sum_{n_3=0}^{N_3-1} a(n_1,n_2,n_3) e^{-j\frac{2\pi}{N_3} n_3 K_3} \qquad 8.1$$

   where $0 \le n_1 \le N_1 - 1,\ 0 \le n_2 \le N_2 - 1,\ 0 \le K_3 \le N_3 - 1$.

2. Zeros are appended to the centre of the data (G), to give G_0, with the Nyquist criterion being satisfied in the case of an even sized dimension, across FFT calculated dimension/s as:

$$G_0(n_1,n_2,K_{03}) = \begin{cases} A(n_1,n_2,K_{03}), & for \;\; K_{03} = 0,\ldots,\frac{N_3}{2}-1 \\ A(n_1,n_2,\frac{N_3}{2})/2, & for \;\; K_{03} = \frac{N_3}{2},M_3L_3-\frac{N_3}{2} \\ 0, & for \;\; K_{03} = \frac{N_3}{2}+1,\ldots,N_3L_3-\frac{N_3}{2}-1 \\ A(n_1,n_2,K_{03}+N_3-N_3L_3), & for \;\; K_{03} = N_3L_3-\frac{N_3}{2}+1,\ldots,N_3L_3-1 \end{cases} \quad 8.2$$

where $0 \le n_1 \le N_1 - 1,\ 0 \le n_2 \le N_2 - 1,$ L_3 is the interpolation factor along the third dimension.

3. N_3L_3-point inverse FFT of zero-appended array $G(n_1,n_2,K_{03})$ is calculated as:

$$g_0(n_1,n_2,n_{03}) = \frac{1}{N_3L_3} \sum_{K_{03}=0}^{N_3L_3-1} G(n_1,n_2,,K_{03}) e^{j\frac{2\pi}{N_3L_3}n_{03}K_{03}} \quad 8.3$$

where $0 \le n_{03} \le N_3L_3 - 1$ and g_0 is the interpolated version of g by factor L_3 in third dimension. The $1/L_3$ amplitude loss made by interpolation is compensated by multiplying it with L_3 [79]. Defining $G(n_1,n_2,K_{03})$ as in Eq. 8.2 maintains the complex-conjugate circular symmetry which characterises the FFT of real valued signals.

FI further allows one to increase the resolution in one particular dimension while decreasing it in another, preserving the original information content, and therefore allowing one to easily control the data size for segmentation. Down-sampling (decrease in resolution) in a particular dimension is attained by first increasing the resolution to integer multiple of the desired resolution and then extracting the relevant components for the preferred resolution.

8.3 3D Level set method (3DLSM)

3D LSM [88] is widely used for the segmentation of anatomical structures from medical imaging volumes with promising results [74], [76], [77]. 3D LSM compared to other segmentation techniques such as region growing [94] is particularly appealing due to the advantages of flexible framework and free topology change in higher dimension space.

Similar to 2D LSM described in Section 3.5.3, in 3D LSM, a closed 3D surface $v(t)$ is embedded as a zero level set of the 4D implicit function ϕ,

$\vartheta(t) = \{X \in \Re^3 \mid \phi(X,t) = 0\}$ and ϕ is evolved according to Eq. 3.14. The desirable feature of LSM is that numerical computations can be performed on a fixed Cartesian grid, where elements are unit cubes, without having to parameterise the points on a surface as in parametric active contour models [87]. Therefore, to be able to apply 3D LSM to anisotropic (non-cubic) medical imaging volumes, where the distance between consecutive slices along the z-dimension is significantly greater than the in-plane (x-y) pixel size (Fig. 8.1), and interpolation is performed in [74], [76], [77] to reconstruct isotropic (cubic) volumes before segmentation using 3D LSM.

8.4 3D throat region segmentation using I+3DLSM

The proposed 3D throat region segmentation process (I+3DLSM) as shown in Fig. 8.2 comprises three steps: 1. pre-processing of MRI slices, 2. Fourier interpolation (FI) of pre-processed slices to produce isotropic voxels and 3. 3D LSM segmentation of the throat region from a MRI volume.

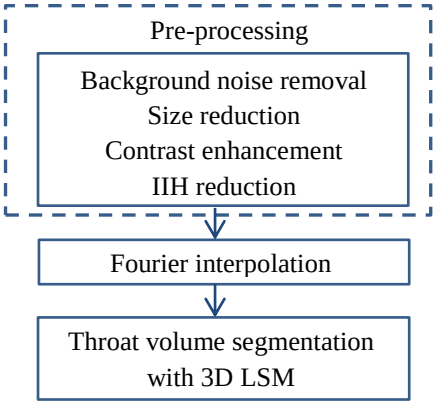


Fig. 8.2: 3D throat region segmentation algorithm flowchart (I+3DLSM) [77] © 2015 IEEE.

8.4.1 Pre-processing

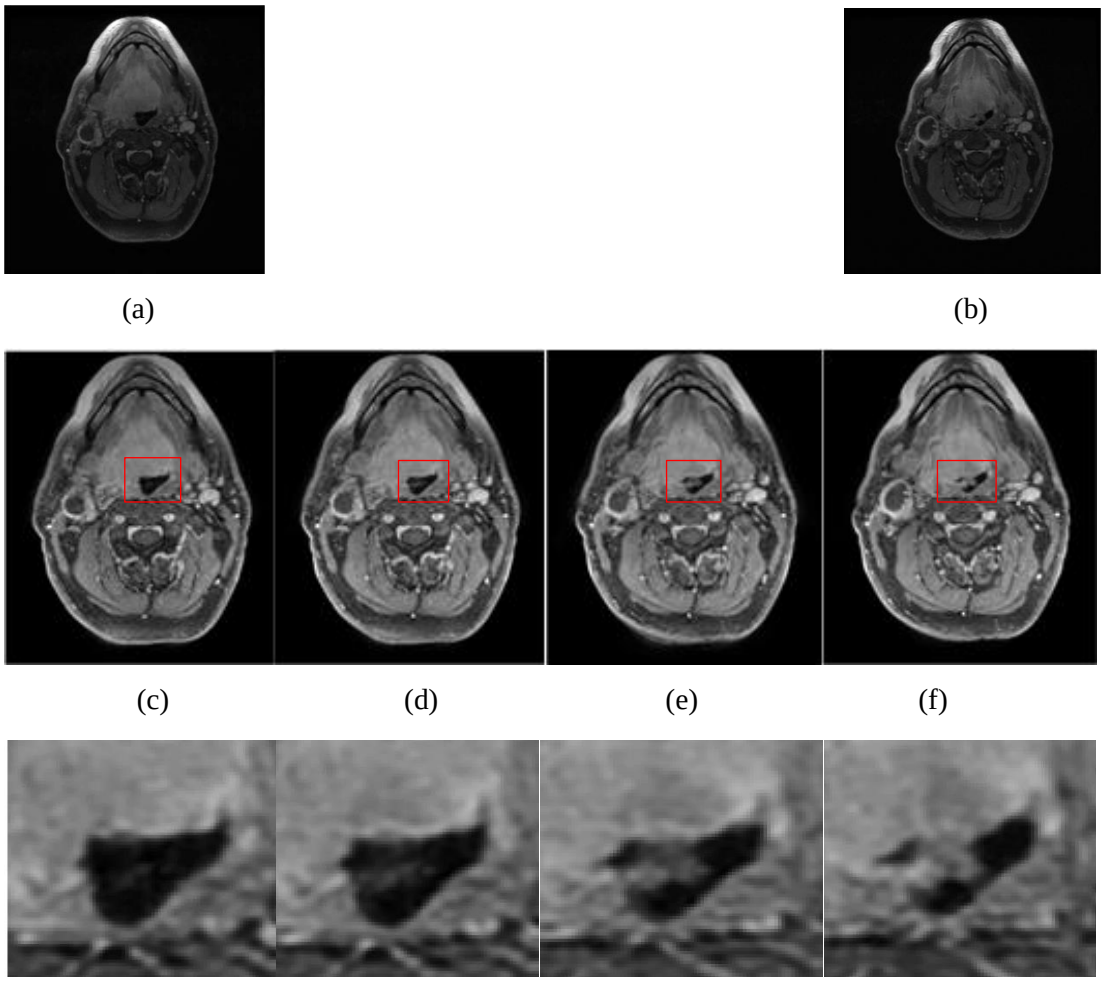
Prior to segmentation, the MRI data is pre-processed to reduce the data size, background noise and imaging artefacts and to enhance the contrast between a throat region and its neighbouring tissues. Initially, a morphological opening [54] and a majority operation [54] are applied for the removal of a significant proportion of background noise. After background noise removal, the overall size of the volume is reduced through detecting the first non-zero intensity value along each dimension and then performing cropping. In order to increase the contrast (intensity difference) a

background brightness preserving contrast enhancement technique [57] is applied. This technique improves contrast while preserving the throat region brightness. Finally, IIH observed in the MRI data used is reduced by using the technique described in Section 6.2. These pre-processing techniques, applied separately to each axial (2D) slices (Fig. 8.3(a) and (b)), shows considerable improvement in MRI slices while preserving the primary features of the slices (Fig.8.3(c) and (f)).

8.4.2 Fourier Interpolation (FI)

With imaging parameters of the real MRI dataset discussed in Section 4.2, it can be observed that the voxel size is highly anisotropic with a range of 0.43x0.43x3.3mm^3 to 0.94x0.94x6mm^3 (Table 4.1). However, LSM (Section 8.3) require a regular square mesh for evolution. Thus, N-D FI discussed in Section 8.2 is performed on pre-processed MRI slices along the required dimension/s to produce high resolution isotropic voxel of approximately size 0.47x0.47x0.47 mm^3. For example, for a dataset with voxel size 0.94x0.94x6mm^3, interpolation is performed along x- and y-dimensions by a factor of 2 to reduce pixel spacing from 0.94mm to 0.47mm and along z-dimension by a factor 13 to reduce slice spacing of 6mm to approximately 0.47mm. This means 12 additional MRI slices are added along z-dimension in between a pair of a real MRI slices. The visual result of FI along z-dimension is shown in Fig. 8.3. The Fig. 8.3(c) and (f) shows 7th and 8th slices from 11 real MRI slices respectively. Fig. 8.3(d) and (e) shows two interpolated slices obtained using FI technique between 7th and 8th slices which reduces the spacing in z-dimension from 3.5mm to approximately 1.17mm. Further, smooth expected change of topology across the throat region due to this interpolation is shown in Fig. 8.3(g).

However, a prominent drawback of the FI is the manifestation of Gibbs phenomenon, causing a ringing artefact (Section 2.6 and Section 3.3). In this work, to reduce a ringing artefact, extra slices of zero (null) intensity was appended at the end of real sequential MRI slices in the spatial domain before interpolation to increase the resolution in Fourier domain. These zero (null) intensity slices were discarded after interpolation. After interpolation and Gibbs ringing reduction, the volume is reconstructed in 3D using both the real and interpolated slices and throat region is segmented from the reconstructed volume using 3D LSM.

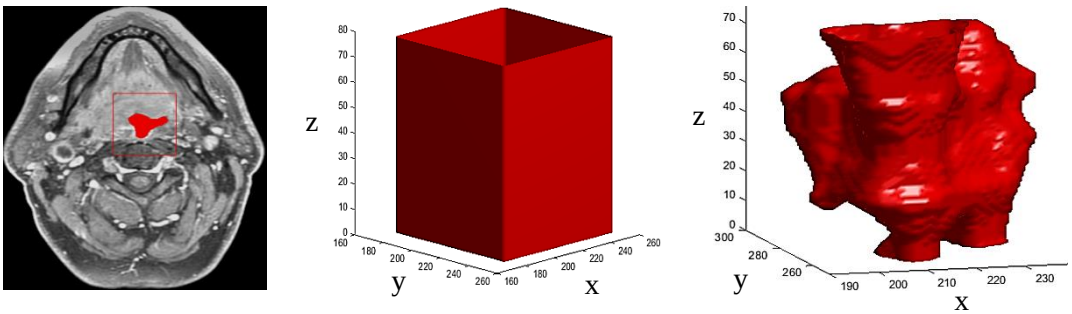


(a) (b)

(c) (d) (e) (f)

(g)

Fig. 8.3: (a) and (b) MRI slices before pre-processing, (c) and (f) MRI slices after pre-processing, (d) and (e) interpolated MRI slices using Fourier interpolation, (g) zoom-in version of MRI slices near the throat region (red region) to demonstrate expected change of topology (splitting) across the throat region. [77] © 2015 IEEE



(a) (b) (c)

Fig. 8.4: (a) detected throat region (red region) in 2D (b) Initialisation (cube) for 3D LSM (c) segmented throat region in 3D using I+3DLSM. Smooth 3D structure of the throat region can be observed in Fig. 8.4(c).

8.4.3 Throat volume segmentation with level sets

The initialisation of 3D LSM (Fig. 3.6) is obtained by automatically selecting the central slice from a MRI volume and detecting the throat region on this central 2D slice using the ThDA described in Section 6.3 (Fig. 8.4(a)). Further, the cube (Fig. 8.4(b)), of size approximately 10% of the total reconstructed volume encompassing all the slices in a volume, from the centre of the detected throat region is defined.

This cube is evolved using 3D LSM. In this work, the speed function V (Eq. 3.14) used for level set evolution was given as [165] (Fig. 8.5):

$$V = w(\sigma_R - |I(x,y,z) - \varpi_R|) + (1-w)\nabla.\frac{\nabla\phi}{|\nabla\phi|} \qquad 8.4$$

where $\nabla.\frac{\nabla\phi}{|\nabla\phi|}$ is the mean curvature of the surface and serves as a smoothness regulator term. In Eq. 8.4, $w \in [0\ \ 1]$ is a free parameter that controls the weighting between data term and smoothing term. In this work, a signed distance function, based on Euclidean distance, is utilised as the level set function ϕ, and a simple data term depending solely on the input intensity I at a point (x,y,z) is considered. This data term is suitable because there is high intensity contrast between the throat region and surrounding tissues. In Eq. 8.4, ϖ_R controls the intensity of the region to be segmented, σ_R controls the intensity (grey level) range around ϖ_R that could be considered inside the object (Fig. 8.5(a)).

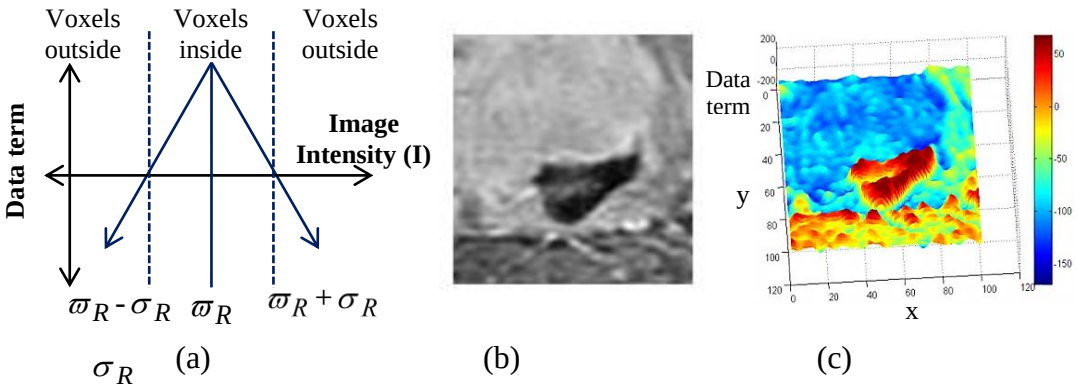


Fig. 8.5: (a) definition of parameters used in data term of speed function (F) in LSM (b) MRI slice with a throat region (c) data term of slice in (b). High values of data term for a throat region can be observed in Fig. 8.5(c).

In this work, ϖ_R is calculated as the mean and σ_R as the standard deviation of the detected throat region from the 2D central slice and is employed for the evolution of level set. The data term used in speed function V for the throat region (Fig. 8.5(b)) is shown in Fig. 8.5(c). It can be noted that value of data term is high for the throat region compared to the surrounding regions.

This integration of interpolation and 3D LSM (I+3DLSM) robustly handles splitting and merging and computes a smooth 3D structure of the throat region. The isolated noisy voxels segmented by the LSM technique are reduced only by considering the largest segmented volume as the throat region (Fig. 8.4(c)).

8.5 Experimental results

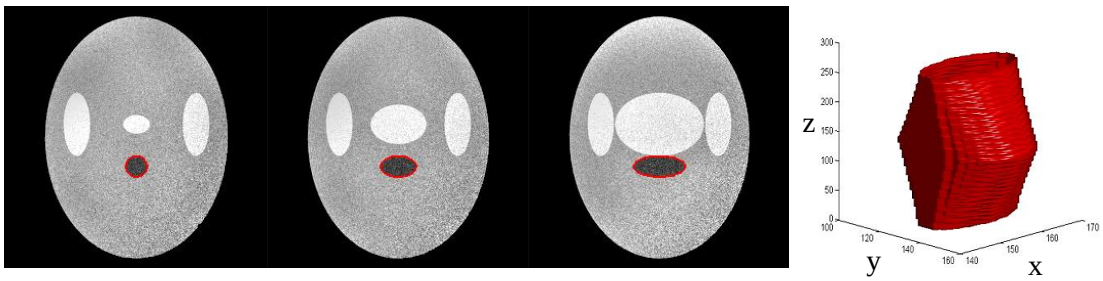
The proposed I+3DLSM algorithm was tested on 27 synthetic and 12 real MRI volumes to validate its robustness and accuracy. For quantitative comparison, manual segmentation of the throat regions on real MRI slices from a medical expert was used as reference segmentation. DSC and F-measure was calculated on slice-by-slice basis between reference and I+3DLSM segmentation results. F-measure was denoted as $F-measure_{I+3DLSM}$.

8.5.1 Time comparison

The I+3DLSM algorithm was implemented in Matlab 2014a (Mathworks, Natick, MA), on Windows 7 system. The Matlab software was executed on a Dell U2412M machine with 8GB RAM. I+3DLSM utilised parallel computing toolbox of Matlab (par for-Loops) to reduce the computational speed. The average time for a throat region segmentation from a MRI volume using I+3DLSM (including interpolated slices) per patient was 7 minutes (range 4 to 10 minutes) compared to 5 minutes (range 2 to 9 minutes) for manual segmentation on only real MRI slices (throat region on interpolated slices were not manually segmented). This time comparison demonstrates that automatic time for segmentation using real and interpolated slices is comparable to the manual time using only real MRI slices.

8.5.2 Synthetic dataset

For 27 synthetic volumes of size 256x256x32 (Section 4.4), the number of slices in each volume was interpolated to equal in-plane resolution, essentially creating a volume of 256x256x256 voxels. The ground truth volume for this evaluation was the reconstructed MRI volume without the addition of any artefacts. The initialisation for LSM for this dataset was determined using the technique described in Section 8.4.3 which acquired a cube of size 78x78x256 voxels. Fig. 8.6 (a) illustrates a number of individual slices with the throat section highlighted in red; despite some small rippling at the edges of each ellipse, the segmented volume (Fig. 8.6(b)) is predominantly representing a conical sphere. The LSM model for this data employed a low magnitude of α (0.02) for smoothness; however, instead of issuing a strict number of iterations for the LSM to be performed, a stopping criterion based on differences between the previous zero-order level set and the current one was given.



(a) (b)

Fig. 8.6: Synthetic MRI dataset (a) throat region boundary (red outline) from 3 slices (b) segmented throat volume from 256x256x256 voxels.

Table 8.1: F-measure values for I+3DLSM on 27 synthetic volumes

IIH \ Noise	5	10	15
Sin	0.98	0.98	0.98
Cos	0.98	0.98	0.97
Sin+Cos	0.98	0.98	0.97

Contrast:0.9

IIH \ Noise	5	10	15
Sin	0.97	0.97	0.97
Cos	0.97	0.97	0.96
Sin+Cos	0.97	0.97	0.97

Contrast:0.6

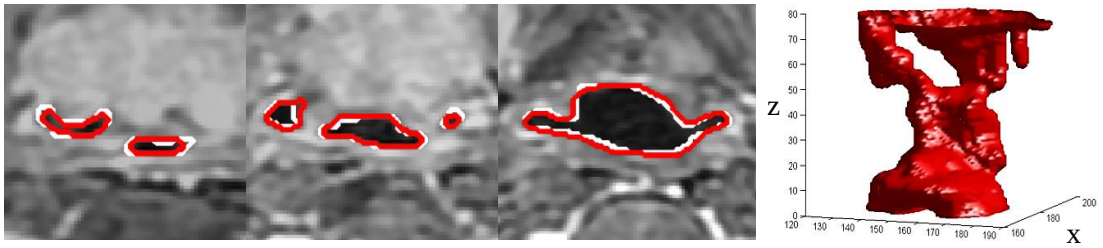
IIH \ Noise	5	10	15
Sin	0.96	0.98	0.97
Cos	0.95	0.96	0.96
Sin+Cos	0.96	0.95	0.96

Contrast:0.3

Table 8.1 shows F-measure values comparing I+3DLSM and ground truth results for all synthetic volumes. This comparison yields an average F-measure value of 0.97. Further, it can be observed that F-measure value is ≥0.95 for all volumes. The low value of F-measure (lowest 0.95) can be attributed to the rippling that occurs at the edges of ellipse. An increase in resolution would decrease the rippling effect and increase the algorithm accuracy, at the expense of additional computational time. For example, for 512x512x512 voxels data lowest F-measure was 0.97. However, for 512x512x512 algorithm execution time was approximately 26 minutes compared to approximately 4 minutes for 256x256x256 voxels. From Fig. 8.6 and Tabl6 8.1 it can be observed that even in presence of MRI artefacts, I+3DLSM can segment the throat region with high accuracy.

8.5.3 Real dataset

All real MRI volumes were reconstructed to create approximately isotropic voxel of size 0.47x0.47x0.47mm^3. Thus, the number of slices in volume after interpolation ranged from 70-259 with 5-12 interpolated slices in between consecutive real slices. Fig. 8.7(a) illustrates a number of slices with the throat section highlighted (red-automatic, white-manual) from a single volume with the individual segments being visibly accurate compared to the reference segmentation.



(a) (b)

Fig. 8.7: Real MRI dataset (a) representative 2D slices with throat region segmentation results (red outline) (b) final 3D structure of the throat region using I+3DLSM.

Further it can be noted that I+3DLSM was able to segment the entire throat region even when the change in topology of the region is observed on the MRI slice (Fig. 8.7(a)). In addition, Fig. 8.7(b) demonstrates the final automatic segmented volume. From Fig. 8.7(b) it is clear that the segmented volume is smooth and no further

processing is required. Furthermore the smooth transitions between all slices validate performing the segmentation in 3D.

Table 8.2: Quantitative values of I+3DLSM for throat region segmentation on real dataset

Dataset	1	2	3	4	5	6	7	8	9	10	11	12
DSC	**0.92**	**0.81**	**0.87**	**0.87**	**0.88**	**0.84**	**0.82**	**0.87**	**0.89**	**0.82**	**0.94**	**0.90**

The LSM algorithm which produced this segmentation employed a low magnitude for α (0.02) ensuring a smooth result, but with a slightly higher number of iterations (600) for convergence.

The quantitative result (DSC) comparing I+3DLSM and manual outlines is shown in Table 8.2. This comparison yielded a mean DSC of 0.87 with minimum of 0.81 and maximum of 0.94. These DSC values suggest that the obtained outlines are in agreement with the manual outlines. The high DSC values (0.96 and 0.93) belong to Fig. 8.8(a) and Fig. 8.8(b), where the throat is very accurately segmented, with minimal penetrations to the surrounding regions and the lowest DSC value (0.45) to Fig. 8.8(c) where due to high intensity range of the throat region, compared to other slices in a volume, I+3DLSM significantly underestimated the throat area.

(a) (b) (c)

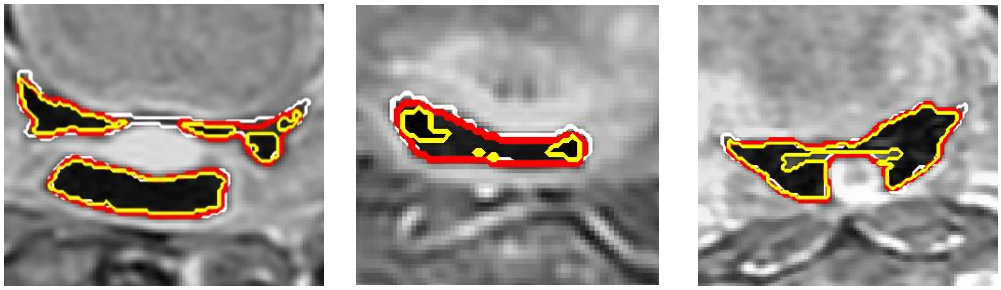
Fig. 8.8: DSC analysis (a) (DSC: 0.96) and (b) (DSC: 0.93) excellent (c) poor (DSC: 0.45) agreement between manual (white outline) and I+3DLSM (red outline) outlines.

8.5.4 Comparison of I+3DLSM with region growing

The results obtained from I+3DLSM for the throat region segmentation from real MRI dataset were compared to the 3D region growing algorithm [94] employed for upper airway or pharynx region segmentation. Region growing (Section 3.5.1) algorithm is an iterative process that merges nearby voxels based on connectivity and homogeneity properties among the voxels starting from initial points (seeds). For 3D region

growing, segmentation was performed using publicly available Matlab code of the algorithm [166]. For comparison, pre-processing (Section 8.4.1) and interpolation techniques (Section 8.4.2) developed in I+3DLSM were applied before segmenting the throat region using region growing technique [166]. The seed point for the region growing algorithm was considered as the centre of the throat region obtained from the central slice of the MRI volume as described in Section 8.4.3. The stopping criterion for the region growing is fulfilled if no voxels in the region's 8-neighbourhood have an intensity difference smaller than certain value to the region's mean intensity. For this dataset the stopping value was set to the variance of detected 2D throat region from a central MRI slice.

For most of the slices, both algorithms demonstrated results similar to the manual outlines. Further, like LSM, region growing algorithm was able to handle topological changes (Fig. 8.9(a)). However, for some slices (Fig. 8.9) where the variance of the throat region is larger than the set value, region growing algorithm was not able to segment the entire throat region (Fig. 7.9(b)) or underestimated the region (Fig. 8.9(c)) compared to I+3DLSM and manual region.



(a) (b) (c)

Fig. 8.9: Qualitative throat segmentation results comparing I+3DLSM (red outline) and 3D region growing (yellow outline) with manual outlines (white outline)

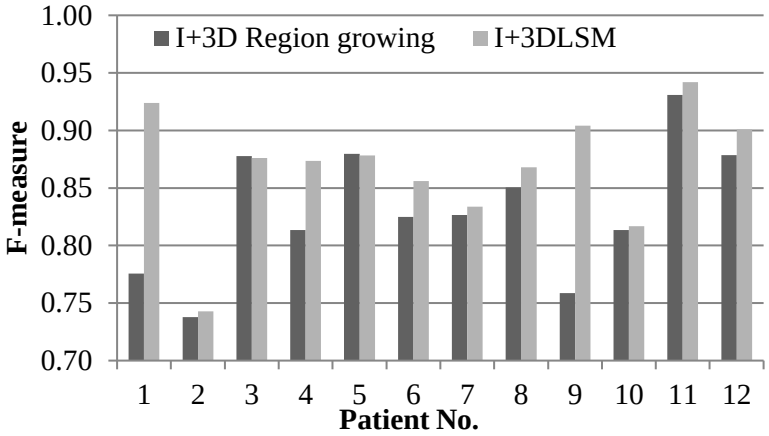


Fig. 8.10: F-measure values comparing I+3DLSM and 3D region growing

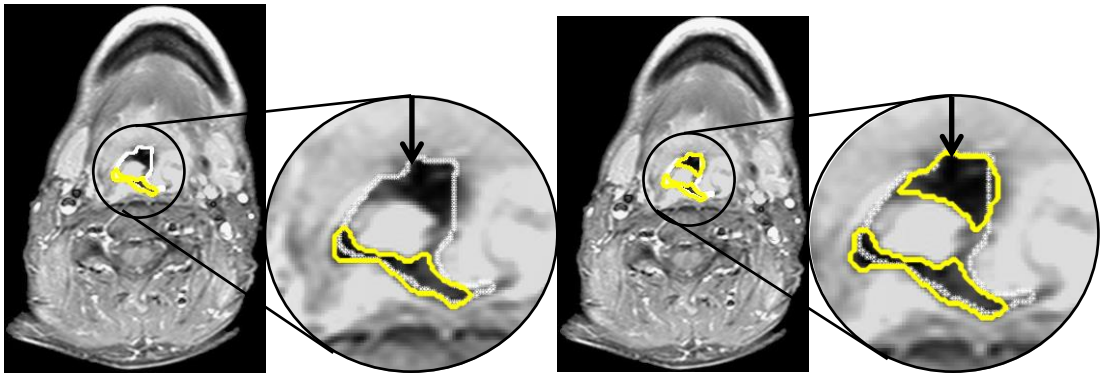
The quantitative (F-measure) comparison of both algorithms is shown in Fig. 8.10 and it can be seen that I+3DLSM performed better with average F-measure value of 0.87±0.05 compared to region growing (average F-measure: 0.83±0.05) in most of the patients except patient 3 and 5. For patient 3 and 5, noise was observed in the throat region. Thus, region growing performed better compared to I+3DLSM.

8.5.5 Comparison of I+3DLSM with 3DLSM and 3DLSM+I

To demonstrate the significance of interpolation before 3D LSM in I+3DLSM to accurately segment the throat regions from anisotropic volumes, comparisons between Matlab implementations of 3D LSM with and without interpolation on 12 MRI volumes are reported in this section. For comparison, the technique of using only 3D LSM is referred as '3DLSM' and technique of using interpolation after 3D LSM is referred as '3DLSM+I'.

Similar to $F-measure_{I+3DLSM}$, F-measure were also calculated between the reference and 3DLSM segmentation results denoted as $F-measure_{3DLSM}$. The F-measure values for 3D LSM+I are similar to $F-measure_{3DLSM}$ on real MRI slices. This is because in 3DLSM+I, the interpolation is carried out to add outlines in between the segmentation outlines obtained from real MRI slices. Thus, there is no change in the segmentation outlines on real MRI slices. The change in accuracy ($\nabla accuracy$) between $F-measure_{I+3DLSM}$ and $F-measure_{3DLSM}$ was calculated in percentage (%) as:

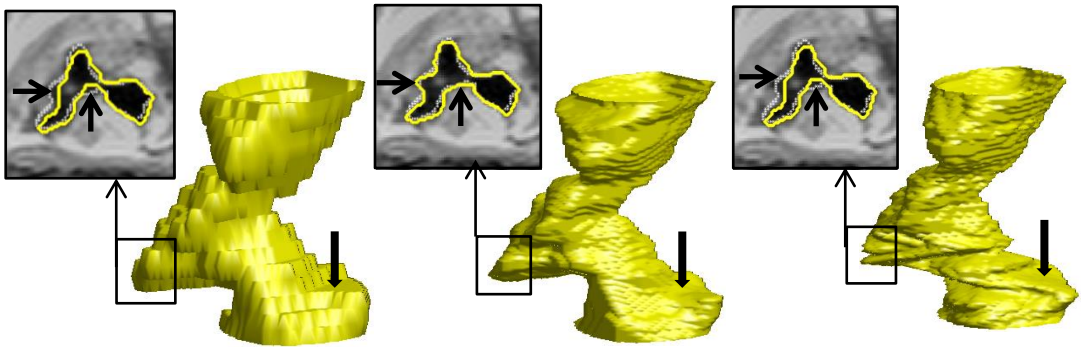
$$\nabla accuracy \quad (\%) = \frac{\left| F-measure_{I+3DLSM} - F-measure_{3DLSM} \right|}{F-measure_{3DLSM}} x100 \qquad 8.5$$



(a) (b)

Fig. 8.11: Illustration of segmented throat region (yellow outline) using (a) 3DLSM (b) I+3DLSM [164] © 2016 The Institution of Engineering and Technology.

Fig. 8.11 shows the segmented throat region (yellow outline) on an MRI slice, obtained using 3DLSM (Fig. 8.11 (a)) and the corresponding segmented throat regions (yellow outlines) from I+3DLSM (Fig. 8.11(b)). The outline obtained by 3DLSM+I is similar to the outline obtained by 3DLSM. Thus, 3DLSM+I outline is not included in Fig. 8.11. Visually comparing the results, it can be observed that I+3DLSM was able to segment both regions of the throat (yellow outlines Fig. 8.11 (b)) whereas the 3DLSM segmented only one part of the throat region (yellow outline Fig. 8.11(a)). This missing region (black arrow in Fig. 8.11(a)) can be attributed to the large spacing (6 mm) in between real MRI slices. The throat region in the previous slice to the slice in Fig. 8.11(a) and missing region in Fig. 8.11(a) are not connected components. This non-connectivity contributes to the collapsing of the level set function and missing the region in the slice in Fig. 8.11(a) and in subsequent slices. However, due to interpolation, the smooth variation in the throat region in I+3DLSM allowed segmentation of both regions which is comparable to the reference segmentation (white* outline) and increased accuracy for that particular slice by 81.84%.



(a) (b) (c)

Fig. 8.12: Illustration of throat volume (yellow outline) obtained using (a) 3DLSM (b) I+3DLSM (c) 3DLSM+I [164] © 2016 The Institution of Engineering and Technology.

Fig. 8.12 shows segmented throat volumes obtained using 3DLSM (Fig. 8.12 (a)), I+3DLSM (Fig. 8.12(b)) and 3DSLM+I (Fig. 8.12(c)). It can be seen that all three volumes appear visually similar; however, I+3DLSM eliminates the need for further smoothing of the throat volume and remove the stair-case effect observed (indicated by black open arrow in Fig. 8.12) in throat volumes segmented by 3DLSM and 3DLSM+I. Further, Fig. 8.12 also illustrates the performance accuracy in the segmented region (black box) using I+3DLSM, particularly in the concave regions (black arrows on 2D slice) as interpolation step extend the capture range of the 3DLSM and provide good convergence to boundary concavities. Thus, interpolation step before 3DLSM increased overall accuracy by 2.78% for this particular dataset with 6.86% increase for the slice shown in Fig. 8.12 compared to 3DLSM and 3DLSM+I.

The quantitative (F-measure) values for 3DLSM, I+3DLSM and 3DLSM+I in Fig. 8.13 show increase in F-measure values and thus, increase in accuracy, with the range from minimum of 0.27% to maximum of 14.86% and average of 4.80%, for all 12 MRI volumes due to interpolation before 3DLSM.

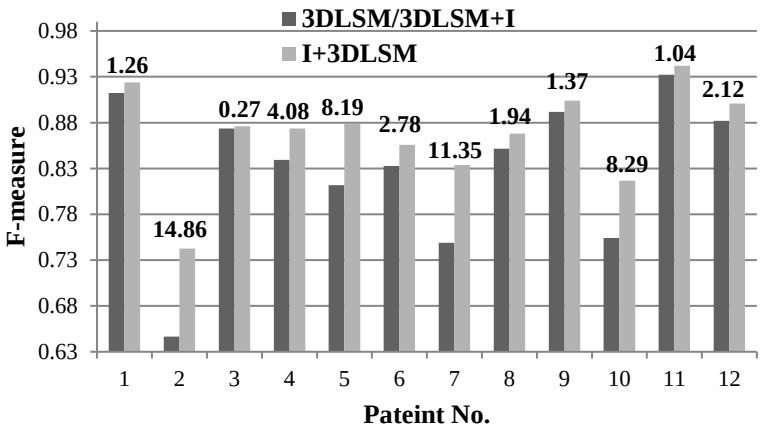


Fig. 8.13: F-measure values for 3DLSM, I+3DLSM and 3DLSM+I with $\nabla accuracy$ (%) values (data labels on each patient) [164] © 2016 The Institution of Engineering and Technology.

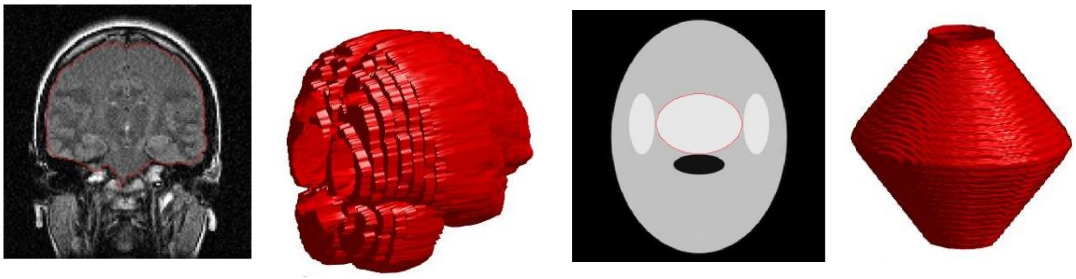
It was observed that the maximum increase in accuracy (14.86%) for I+3DLSM was observed for the dataset with maximum slice spacing (6mm). Further, low F-measure values for 3DLSM and 3DLSM+I compared to the reference segmentation is due to the large number of false negative pixels (e.g.: 353 pixels for particular slice) as uncertainties in between slices underestimate the throat area. The interpolation step before LSM segmentation reduced these uncertainties thus; reducing false negative pixels (to 76 pixels). The limitation of I+3DLSM, however, is that in some MRI slices it slightly overestimate (maximum false positive pixels: 59) the throat area when compared to reference segmentation, particularly on the slices where the throat region is small. From a clinical view point this overestimation, however, is preferable to an underestimation due to the dangers of recurrence of disease from undertreating the target. Overall, quantitative results agree with visual results that I+3DLSM segmentation results provide comparable result to the reference segmentation compared to 3DLSM and 3DLSM+I.

8.6 Discussion

A new method (I+3DLSM) for volumetrically segmenting the throat region from a number of MRI slices is proposed. The I+3DLSM method performs the segmentation in 3D, eliminating the need for post-processing reconstruction and smoothing algorithms that is required when performing the segmentation on a slice by slice basis as in Chapter 5. The pre-processing methods (Section 8.4.1) implemented in this work

allow reduction of false positives. The FI helps to increase the resolution along desired dimension/s, thus increasing the accuracy in the throat region segmentation. The LSM model discussed in this Chapter possesses a relatively simplistic data function in LSM, but the results demonstrate that it is sufficient for performing the segmentation process for relatively high contrasting regions.

The main advantage of the I+3DLSM method is that it makes no assumption regarding the individual shapes, size or areas. Furthermore using the method it becomes unnecessary to smooth the initial volume with Gaussian filters prior to segmentation, preserving data fidelity. The proposed algorithm is generally applicable in a number of areas and can be employed to segment object with good contrast between desired object and its neighbours (Fig. 8.14).



(a) (b)

Fig. 8.14: Application of I+3DLSM for the segmentation of different regions of interest (a) brain region (b) synthetic tumour region (representative slice with segmented region (red outline) and segmented volume is shown in the Figure)

The results of the proposed algorithm (I+3DLMS) on 12 real MRI datasets acquired using standard imaging protocols (average F-measure: 0.87) and 27 synthetic volumes generated with presence of MRI artefacts (average F-measure: 0.97) obtain consistent results when compared to the reference segmentation results. The proposed I+3DLSM algorithm performed better compared to 3D region growing algorithm (average F-measure: 0.83) and ThDA algorithm (average F-measure: 0.77) (Section 6.3.1). The DSC obtained for I+3DLSM with range 0.81 to 0.94 (average: 0.87) is also comparable to the DSC values in [94] with the range of 0.84 to 0.94. Experimental results also demonstrate that an interpolation before 3D LSM (I+3DLSM) produce more accurate results compared to 3DLSM and an interpolation after 3DLSM

(3DLSM+I). This was particularly observed in the presence of change of topology and concave regions of the throat region.

The segmented throat region volume can be employed as additional information for the segmentation of oropharynx and larynx tumours [10] or for the study or further analysis of pharyngeal airway in the patients with sleep related disorder as it can enable visualisation and detection of the obstructive region.

However, there are several areas in I+3DLSM algorithm which warrant further development. Computationally, the narrowband LSM algorithm has been constructed in a sequential manner and a performance boost could be obtained if the procedure were carried out in parallel. The current model requires thorough testing with large and varied datasets, particularly in cases where models exhibit weak-edges, to investigate how successfully the data function performs; however it certainly exhibits promise in the dataset used in this work. Further, algorithm can be improved to correct over or under segmentation by automatically tuning the parameters based on the underlying data.

8.7 Conclusion

This chapter presented a new automatic method (I+3DLSM) for segmenting the throat region from a MRI volume which reduces data size and MRI artefacts and combines Fourier interpolation with 3DLSM. The present results are encouraging with good overlap between automatic and manual results and demonstrate that proposed I+3DLSM have potential for clinical application.

Chapter 9
Discussion

9.1 Introduction

Accurate analysis and quantification of tumours from MRI data assist clinicians in appropriate tumour staging and precise RTP, successively eradicating tumourous tissues while reducing damage to the healthy tissues. The segmentation of tumour regions from MRI slices is essential, which must be accurate, reliable and reproducible, for subsequent quantitative measurements. However, obtaining such segmentation results for oropharynx and larynx tumours from MRI are difficult because of tumours pattern (such as overlapping intensities with nearby anatomical structures, no distinct edges, different shapes and sizes) and imaging challenges (such as variation in imaging protocols, low tissue contrast and IIH) (Chapter 4).

This thesis has presented novel algorithms and framework developed for automatic analysis of oropharynx and larynx tumours from MRI data. The qualitative and quantitative analyses with the current practice used in clinic and state-of-the art algorithms have also been discussed in Chapter 7. Chapter 8 presented the algorithm developed for 3D segmentation of throat regions from anisotropic MRI volumes. This chapter reviews key findings and limitations of these developed algorithms and framework.

The rest of chapter is organised as follows. Section 9.2 reviews and discuss contributions (Section 1.4) along with key findings of this work. Section 9.3 discuss generalisation of the developed framework and its sub-parts. Limitations of the work are presented in Section 9.4 and conclusion in Section 9.5.

9.2 Key contributions and findings

The main contribution of this work is the development of the framework that demonstrates the clinical feasibility of the automatic segmentation and quantification of the oropharynx and larynx tumours from MRI data. To the best of our knowledge, this is the first framework developed for automatic analysis of such tumours from contrast-enhanced MRI data. The framework (PLCSF) was developed by integrating

novel image processing techniques for IIH reduction, throat region detection (ThDA), extraction and segmentation of a tumour region using MFCM clustering and LSM from 2D MRI slices with 3D reconstruction and volume calculation of a tumour region. The clinical feasibility of the framework was demonstrated by comparing obtained tumour outlines to the outlines from different CO and consensus manual outlines. Finally, the possibility of 3D segmentation for MRI dataset used in this work was verified by developing algorithm for 3D segmentation of a throat region from a MRI volume (minor contribution).

The key findings of the main contribution are:

1. Unsupervised 2D segmentation framework (PLCSF) (Section 5.3) can effectively segment geometrically variable tumour regions from T1+Gd MRI slices obtained using different scanners with different imaging protocols (Fig. 7.2) and in presence of artefacts. However, PLCSF (C-auto) over segments small and low contrast tumour regions compared to the manual (C-gold) approach (For example, Fig. 7.2: Patient 2,3 and 6). Further, PLCSF yield spatial discontinuity in between tumour outlines from neighbouring slices due to the 2D approach (Fig. 7.7).
2. The PLCSF (C-auto) obtained satisfactory quantitative results when compared to manual outlines from C-gold (mean DSC: 0.79), C1 (mean DSC: 0.77) and C2 (mean DSC: 0.81) with regard to spatial agreement. Further, there is no statistical significant difference (ρ -value: 0.82) between automatic-manual (C-auto-C-gold) and manual-manual (C1-C-gold and C2-C-gold) comparison indicating clinical acceptability of PLCSF results. Further, a strong positive correlation was observed between PLCSF and C-gold tumour areas (r: 0:89; ρ -value< 0.05) with better PCC for oropharynx (BoT) (r: 0.95) compared to larynx (r: 0.83) as tumours of oropharynx tumours are well-defined.
3. Automatic (PLCSF) segmentation of tumours allowed time saving (of 46%) compared to segmentation from an expert (C1) which would be favourable for potential clinical utility.
4. No significant variation was observed between manual segmentations (between C1 and C2) of a tumour region from different experts (inter-variability) in terms of tumour areas (ρ -value <0.05). However, C1 over segment tumour region in case

of weak edges. If parameters values of PLCSF are unchanged, PLCSF can obtain objective results, (without any manual bias) with the advantage of being fully reproducible.

5. PLCSF performed better than MS clustering [112] and N-cut algorithm [113]-[114] in terms of spatial agreement as MS clustering under-segment tumour regions by producing large number of clusters and N-cut over-segment tumour region as leaks through weak edges.
6. 3D reconstructed tumour can provide additional morphological information that can assist CO further in tumour stating. However, PLCSF over-estimates tumour volumes compared to manual (C-gold) as it over-segments the low contrast tumour region.

The key findings of the minor contributions are:

1. A new method developed to determine spacing between two consecutive knots for IIH estimation using bicubic spline (Section 6.2) proved to be effective in presence of cancer regions on MRI slice. The results indicate that IIH reduction with automatic knot spacing decrease intensity variation of a region better compared to constant knot spacing [10], [72] (mean decrease in CoV: 0.18). However in presence large tumour region, large knot spacing (60 mm) achieved better reduction in CoV than spacing determined by proposed method (38 mm) (Fig. 6.3: Patient 2)
2. Fuzzy rules based technique (ThDA), due to no statistical assumptions, was able to successfully detect the throat of different shapes and sizes (Fig. 6.9). ThDA also (mean F-measure: 0.77) outperformed Hessian analysis [11] (mean F-measure: 0.70) for the detection of small throat regions and FC [12] (mean F-measure: 0.69) in presence of noise. However, as Hessian analysis does not assume spatial location of the desired region performed better than ThDA in case of multiple throat regions on single MRI slice (Fig. 6.11).
3. Integration of additional information, in terms of spatial distance (MFCM clustering) in FCM clustering [13] increased the robustness of FCM for extraction of a tumour region from similar intensity surrounding tissues (Fig. 6.17). The results demonstrate that there was improvement in MFCM (mean DSC: 0.70) compared to FCM (mean DSC: 0.66) clustering with regard to spatial agreement.

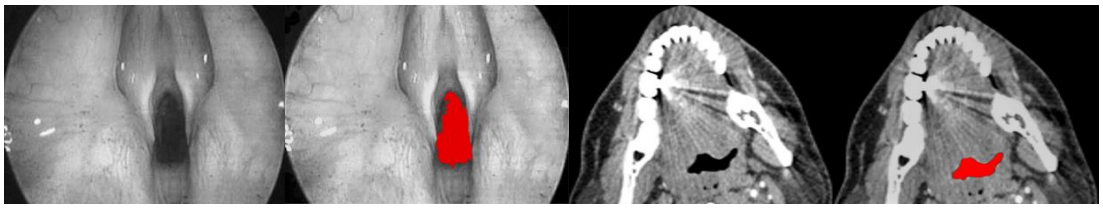
4. 3D segmentation (I+3DLSM) of a throat region, due to spatial continuity, achieved (mean F-measure: 0.87) better results compared to 2D algorithm (ThDA) (mean F-measure: 0.77). Interpolation before segmentation increases accuracy of the results (by 14.86%) compared to segmentation only and segmentation followed by interpolation for highly anisotropic MRI data.

9.3 Generalised application of the framework (PLCSF)

Although the general image processing techniques such as fuzzy-rules, FCM clustering and LSM are modified in this work to make them appropriate for segmentation of a desired region, due to variability in the dataset (Table 4.1), the proposed techniques should be applicable to other imaging modalities such CT or dataset from other imaging domains. The robustness of the framework in presence of artefacts and capability of this framework to segment cancer region from other type of data was demonstrated using synthetic dataset (Fig. 7.1).

The technique developed as a part of this framework such as IIH reduction does not use any information from MRI scanner such as the geometry of the head and neck coil or require a reference slice with no IIH. The technique [10] used as basis in this work for IIH reduction was applied to different MRI images. Thus, further improvement in this technique with respect to the automatic calculation of knot spacing make this technique applicable for varied dataset. Further, other technique developed for the throat region detection (ThDA) does not assume any morphological or appearance information, as is assumed in ASM or AAM technique, or statistics from underlying data. The only information used in the ThDA is the location (spatial position) and intensity of the region, thus ThDA can detect any region present in the middle of the slice with low intensity range. The general application of IIH reduction technique and ThDA is demonstrated in Fig. 9.1.

The final technique developed as a part of this framework, that integrates the spatial distance measure in FCM clustering can also be applied to other dataset. For instance, distance measure created from trachea (Fig. 9.1) can be included in FCM clustering to assign high membership value to the vocal cords region, so that vocal cords can be easily extracted or separated from other anatomical structures for further processing such as texture analysis.



(a) (b) (c) (d)

Fig. 9.1: Endoscopic image (a) before and (b) after IIH reduction and detection of a trachea. CT image (c) before and (d) after IIH reduction and detection of a throat region This Figure shows application of IIH reduction algorithm and ThDA for dataset other than MRI.

Finally for the 3D segmentation algorithm (I+3DLSM) developed, the main advantage is that it makes no assumption regarding the shapes, size or areas of the desired region. The interpolation technique (FI) used in this algorithm is general spatial interpolation (similar to linear or cubic interpolation techniques). Further, the external energy (data term) used in the level set segmentation only depend the intensity of the desired region. Thus, I+3DLSM algorithm can be employed to segment any object with good contrast between desired object and its neighbours. Examples of this algorithm employed to segment brain and cancer region from toy data is shown in Fig. 8.14.

9.4 Limitations of the proposed framework (PLCSF)

The developed framework has several limitations:

1. The proposed framework was developed for T1+Gd (contrast-enhanced) MRI slices with ‘bright’ tumour region compared to surrounding tissues. For other MR sequences such as T1W and T2W, where tumour region appears ‘dark’ or ‘similar’ to surrounding tissues, framework need to be further evaluated with certain adjustments to the segmentation algorithms.
2. The framework, in particular MFCM clustering, utilises spatial and intensity information to extract the expected tumour region, thus over segments the region in case of low contrast between tumour and neighbouring tissues (Fig. 7.3 Patient 6). Incorporating knowledge of morphology or texture in MFCM clustering would obtain more accurate results.
3. Algorithms developed in this thesis, in particular ThDA, MFCM and LSM, have parameters whose values were determined empirically. These include a threshold

on fuzzy membership values, number of clusters and number of iterations. While these empirical parameters worked on MRI dataset obtained from three different vendor's scanners (Table 4.1) and other dataset (Fig. 9.1), it does limit the applicability of the algorithm to new data sets without re-tuning of parameters. Thus, there is the potential that better results could be achieved with additional retuning.

4. Validation of framework results in terms of accuracy and reliability were obtained using different region based (DSC, F-measure) and surface based metrics (MHD). However, validation with metrics such as specificity, receiver operating curve (ROC) would provide different insight about the behaviour of the algorithms.
5. Although, the current framework (mean 8 minutes/patient) allowed time saving of 46 % when compared with a manual segmentation (mean: 15 minutes/patient) of a tumour region, no time required for manual correction of the obtained automatic results was considered. Thus, time comparison of automatic results plus manual correction with manual segmentation would be better to understand the time gain from clinical viewpoint.
6. Because of the time limit, 2D segmentation approach was employed to obtain segmentation results from MRI data with large variations in intensity and imaging protocols. While this 2D segmentation obtained comparable results with manual outlines (mean DSC: 0.79), standardising data or utilising information from neighbouring slices would increase consistency in the segmentation results.

9.5 Conclusion

This chapter discussed the key contributions, findings and limitations of this work. A framework developed for the analysis of oropharynx and larynx tumours from MRI data, using image processing techniques, can provide assistance to CO in RTP. Further the algorithms developed as sub-part of this framework can be applied to other dataset that satisfy assumptions made in the algorithms. However, framework need to be further evaluated in terms of prior information used, parameter values, and validation metrics. Higher level conclusions on the whole thesis are presented in the next chapter (Chapter 10).

Chapter 10

Conclusions and future work

10.1 Conclusions

This thesis investigated the various stages involved in the HNC analysis from MRI data. The main focus of this research was to develop an automatic (computer-aided) oropharynx and larynx tumours (GTV) segmentation and quantification system. The motivation for such work is due to the development in radiotherapy techniques (IMRT), that require accurate segmentation of tumour volumes from medical imaging in timely manner. Although the current manual method for a tumour analysis is used as the gold standard in RTP, limitations of this method lie in the subjectivity and time-demanding process. Thus, a new method that can improve the efficiency in terms of time and reproducibility of a tumor definition is essential. This thesis demonstrated that a new integrated system (framework) using image processing techniques is a viable solution to obtain objective and time efficient tumour segmentation and quantification results. This new automatic framework (PLCSF) (Chapter 5) includes techniques for artefacts reduction, the throat region detection, extraction and segmentation of a oropharynx or larynx tumour region with 3D reconstruction and volume calculation from T1+Gd MRI slices.

The initial artefacts, such noise and IIH, reduction techniques in the PLCSF simplified subsequent detection and segmentation of a tumour region. Further, the additional spatial information (in terms of distance measure created from the detected throat region) in clustering techniques such as FCM (Chapter 6) increased the capability of the clustering method to extract the expected tumour region of inherent geometric variations. Moreover, using local region statistics in the level set segmentation obtained smooth tumour outlines in case of weak edges and heterogeneous regions. Finally, automatic 3D reconstruction and tumour volume measurement provided additional morphological information that can contribute to further tumour staging and analysis. Thus, this proposed 2D automatic segmentation framework (PLCSF) with use of no training data, successfully handled intrinsically

geometric variations of the tumour region without biasing the solution towards a particular size or shape and obtained continuous outline of the tumour region even in the case of broken boundaries. Further, this 2D segmentation approach with 3D reconstruction achieved acceptable results (DSC: 0.79) even with large intra-intensity variations and spatial separation among MRI slices. Additionally, the techniques developed as a part of the framework have shown to be applicable to the images from other domains; given that region of interest in the image fulfil the assumptions made in the proposed techniques (Section 9.3).

The limitation of the PLCSF is the use of a single modality (T1+Gd MRI) that restricts the application of the PLCSF to the tumours that are enhanced with the contrast agent. However, for the clinical application such as RTP, using a single modality is an advantage due to the computational efficiency and ease of use. Other limitation of the current PLCSF is over segmentation of the tumour regions present at the end slices or small tumours that are geographically close and of similar intensity to healthy tissues. However, there is no statistically significant variation between PLCSF and expert outline in terms of tumour areas. Further the automatic segmentations results will be available to the experts and can be corrected (modified) if necessary by an expert (for example using "Modify GTV outline" push button from the GUI explained in Appendix A) before treatment planning.

The preliminary evaluation of the proposed PLCSF with 3D reconstruction and quantification system on real MRI datasets, obtained from MRI scanners from different manufacturers, using different quantitative metrics demonstrated that the system results are comparable to the manual results (average DSC: 0.79 and average volume concordance of 86.51% on real 102 axial MRI slices) across the cohort. Calculation of inter-variability demonstrated no significant variation between manual segmentation from 2 experts (C1 and C2), but there was subjective in the manual outlines from C1 to include a throat region or to over segment region in case non-distinct boundary. However, PLCSF demonstrated no such bias and obtained objective results which are clinically useful. Further, the clinical usefulness of this framework come from reduction in the time needed for segmentation, as the auto-segmented tumour may merely have to be modified, if needed, and validated, rather than segmented from scratch. Thus, this framework for oropharynx and larynx tumours

segmentation and quantification can assist clinician and increase the capability to analyse a tumour by not only providing spatial position and extent of tumour but also a quantitative value in an effort to improve RTP. This framework can be implemented to the software platform of MR systems for clinical utility without the use of any special equipment.

Finally, a novel automatic 3D throat region segmentation technique demonstrated the feasibility of 3D segmentation for a region with less intra-intensity variation from a MRI volume. Experimental results demonstrated that interpolation before 3D LSM improves the segmentation accuracy of the 3D LSM, by allowing it to successfully segment the concave boundaries and multiple regions of the throat on the same MRI slice. The 3D segmented throat region can be utilised as additional information for the 3D segmentation of the oropharynx and larynx tumours.

10.2 Future work

The following are the possible extensions for the work in this thesis:

1. Evaluation of PLCSF on larger dataset: The proposed PLCSF was evaluated on 102 axial MRI slices from 10 patients (Table 4.1). This current dataset is representative of the real clinical cases. However, it is desirable to re-evaluate the work with larger dataset that contain more challenging cases such as multiple primary tumour regions and diffusely infiltrating tumour regions. Also, the larger dataset can include tumour regions from different TNM stages (I to IV). Further, a through testing in a clinical environment will be required before the system can be used in the clinical settings.
2. 3D tumour segmentation: The proposed PLCSF segment the tumour regions separately on each axial MRI slice for a patient and then reconstructs into 3D for further quantification. In future, it would be desirable to perform a 3D segmentation of a tumour region from a MRI volume to utilise more spatial information from the neighbouring slices. 3D segmentation will also eliminate the need for the post-processing reconstruction and smoothing.
3. Initial step for oropharyngeal and laryngeal CTV and PTV segmentation: Current practice in RT requires segmentation of three different target volumes (GTV, CTV, and PTV) (Section 2.4) so that all affected regions (such as tumour

and lymph nodes) can receive some amount of radiation in RT treatment. The GTV includes primary tumour regions while CTV include GTV plus affected lymph nodes. PTV includes CTV plus margin for any motion or setup errors. This thesis considers the automatic segmentation of oropharyngeal and laryngeal GTV. Thus, the novel techniques developed in the PLCSF framework can be utilised as an initial steps for oropharyngeal and laryngeal CTV and PTV segmentation from T1+Gd MRI.

Appendix A

A.1 Graphical User Interface (GUI) for manual tumour (GTV) segmentation

As discussed in Section 4.3, a simple Graphical User Interface (GUI) was developed using Matlab "GUIDE" user interface tool to allow CO to draw an oropharynx or larynx tumour outline/s. A GUI is a set of techniques and mechanisms, used to create interactive communication between a program and a user. Fig. A.1 illustrates the screen shot of the built-in GUI.

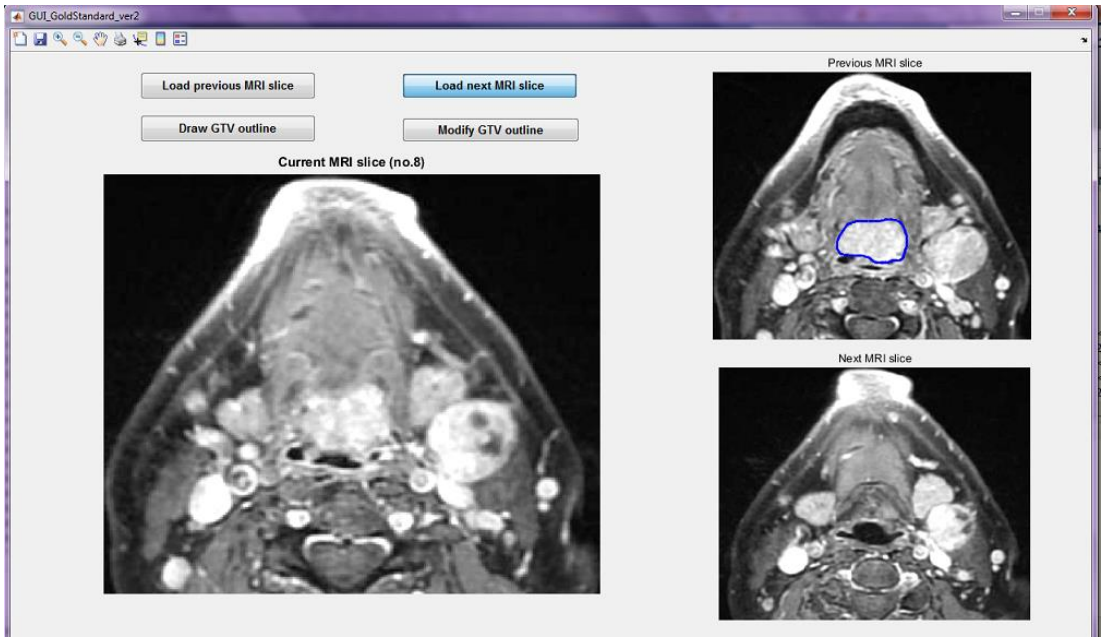


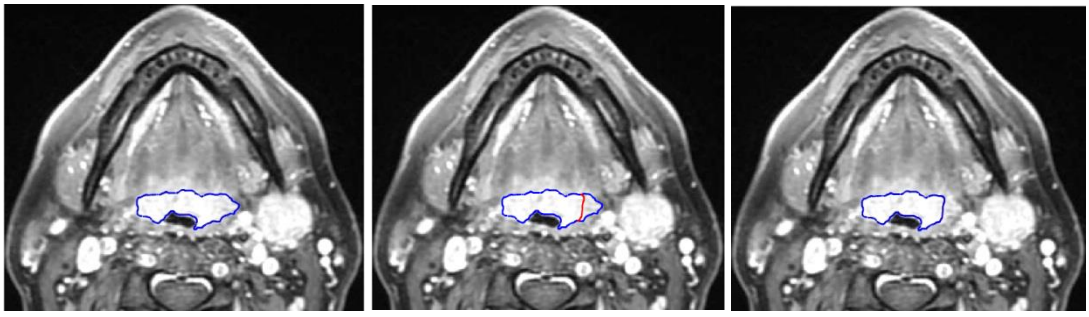
Fig. A.1: Simple Matlab GUI for CO to draw a tumour (GTV) outline

This GUI (Fig. A.1) consists of four push buttons.

1. 'Load previous MRI slice' push button permits to view a slice from 'Previous MRI slice' window in a 'Current MRI slice' window.
2. 'Load next MRI slice' push button permits to view a slice from 'Next MRI slice' window in a 'Current MRI slice' window.
3. 'Draw GTV outline' push button provides a freehand tool to draw a tumour outline on a MRI slice in a 'Current MRI slice' window. The freehand tool allows user to draw outline by holding a right mouse click button. Releasing the button

terminates the outline. User can draw more than one outline in case of splitting of a tumour is observed by pressing a 'Draw GTV outline' push button as many number of times as number of outlines a user wish to draw.

4. 'Modify GTV outline' push button allows a user to edit the tumour outline (Fig. A.2(a) blue outline) on a slice in a 'Current MRI slice' window. This button provides a freehand tool to draw an open segment (Fig. A.2(b) red segment (outline)). Further, the tool delete the smallest segment from the blue outline and joins the remaining segment of the blue outline to the red outline to form a new tumour outline (Fig. A.2(c) blue outline). User can modify the outline number of times until user is content with the tumour outline.



(a) (b) (c)

Fig. A.2: Process to modify tumour outline (a) Manual tumour outline (blue) (b) open segment (red) drawn near the outline in (a) (c) modified tumour outline.

Other options in GUI (Fig. A.1) in menu toolbar include zoom in, zoom out, pan buttons to aid user to observe clear outline of the tumour. While using this GUI, users (CO) were allowed to access any other clincial or radiological information to aid in manual segmentation if required.

Appendix B

B.1 Visual comparison of C-auto with C-gold tumour outlines

This section demonstrates 10 axial MRI slices comparing C-auto and C-gold tumour outlines for each excellent (DSC: >0.85), acceptable (DSC: 0.7-0.85) and worst (DSC: 0.6-0.7) cases.

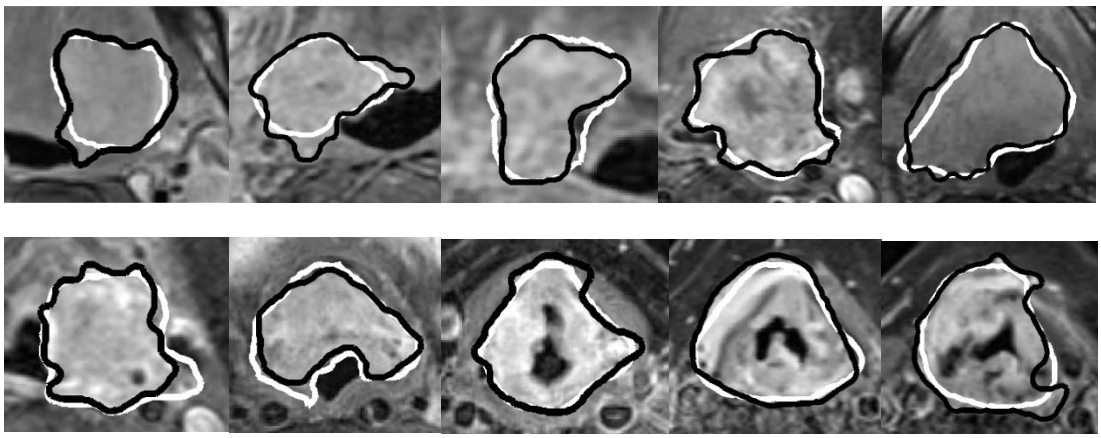


Fig. B.1: 10 axial MRI slices from different patients demonstrating excellent agreement (DSC: >0.85) between C-auto (black outline) and C-gold (white outline) tumour outlines

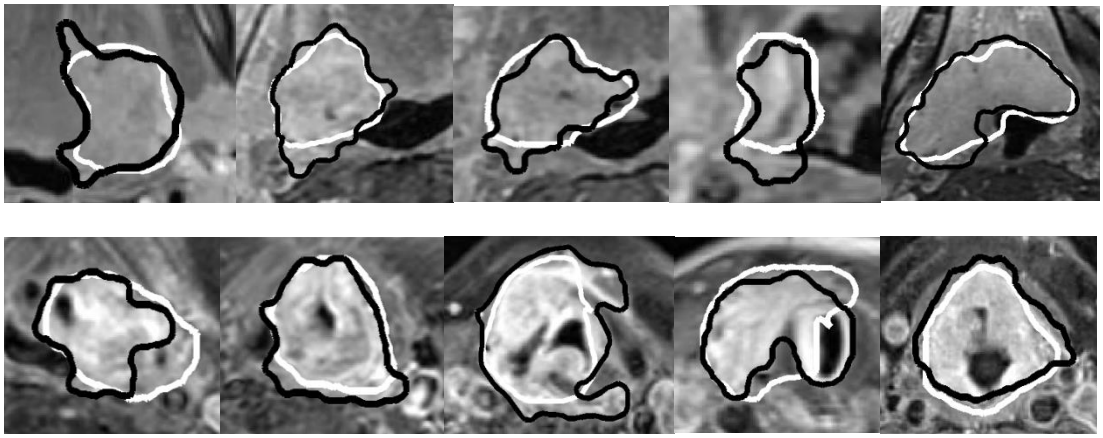


Fig. B.2: 10 axial MRI slices from different patients demonstrating acceptable agreement (DSC: 0.7-0.85) between C-auto (black outline) and C-gold (white outline) tumour outlines

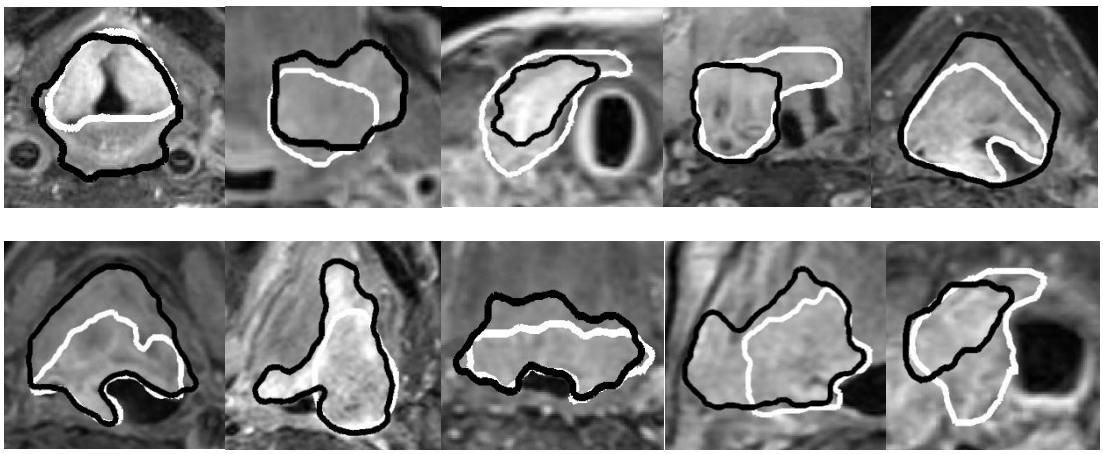


Fig. B.3: 10 axial MRI slices from different patients demonstrating worst agreement (DSC: 0.6-0.7) between C-auto (black outline) and C-gold (white outline) outlines

B.2 Visual comparison of C-auto with manual tumour outlines

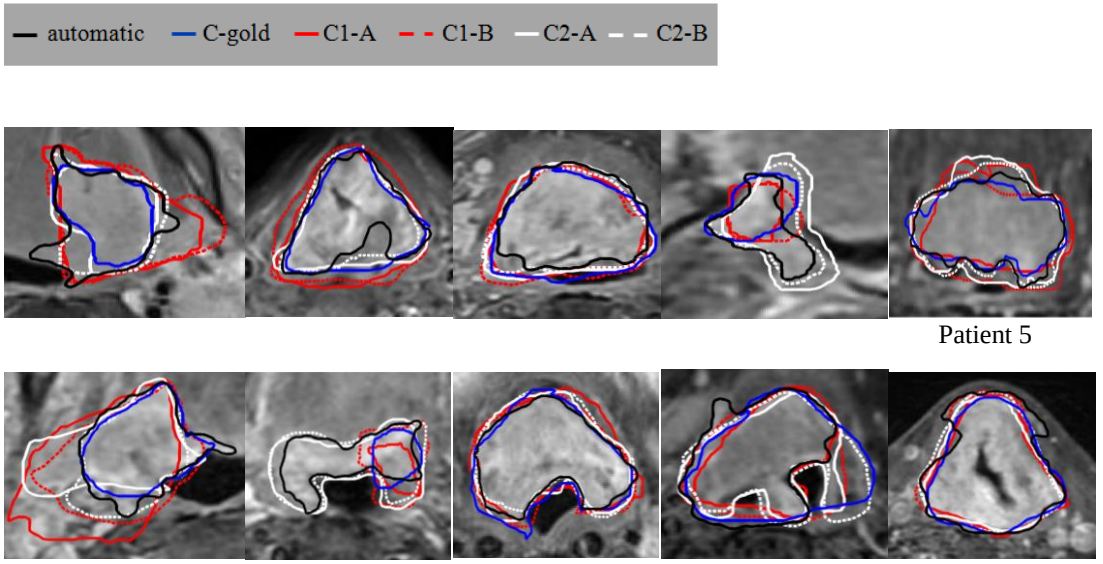


Fig. B.4: 10 axial MRI slices from 10 different patients comparing C-auto with manual (C-gold, C1-A, C2-A, C1-B and C2-B) tumour outlines

Patient 1, 3, 5, 8, 10 : examples of very strong correlation (>0.8).

Patient 2, 4, 6, 7, 9 : examples of variability between outlines (correlation 0.6-0.8).

B.3 Quantitative comparison of C1 and C2 with C-gold tumour outlines

Table B.1: Quantitative measures of C1 and C2 with C-gold outlines (mean (standard deviations))

Comparison	DSC	MHD (mm)
C-gold vs. C1-A	0.78 (0.08)	2.32 (0.75)
C-gold vs. C1-B	0.78 (0.08)	2.42 (0.84)
C-gold vs. C2-A	0.79 (0.06)	2.20 (0.63)
C-gold vs. C2-B	0.80 (0.05)	2.12 (0.64)

The performance measures calculated between the manual outlines from different experts and gold standard (C-gold) are shown in Table B.1. The mean DSC of 0.79 (0.07) and MHD of 2.27 (0.72) mm for all 10 patients indicate the results are similar with C-gold and C-auto comparison (Table 7.3). Further, an average area (C-manual) was calculated using C1-A, C1-B, C2-A and C2-B tumour areas and PCC was obtained between C-manual and C-gold area (mm^2) (Fig. B.5) which yielded a mean PCC of 0.88 with ρ-value<0.05.

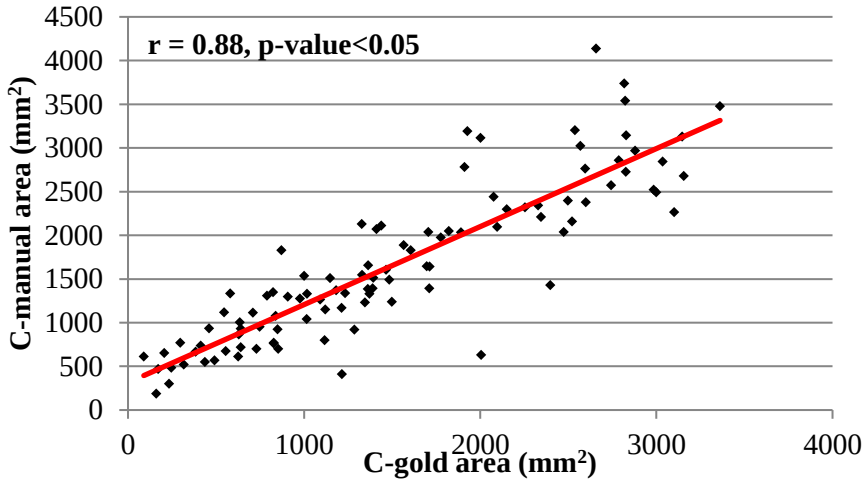


Fig. B.5: Correlation analysis between C-gold and C-manual area (mm^2)

B.4 Reconstructed volume comparison of C-auto with C-gold

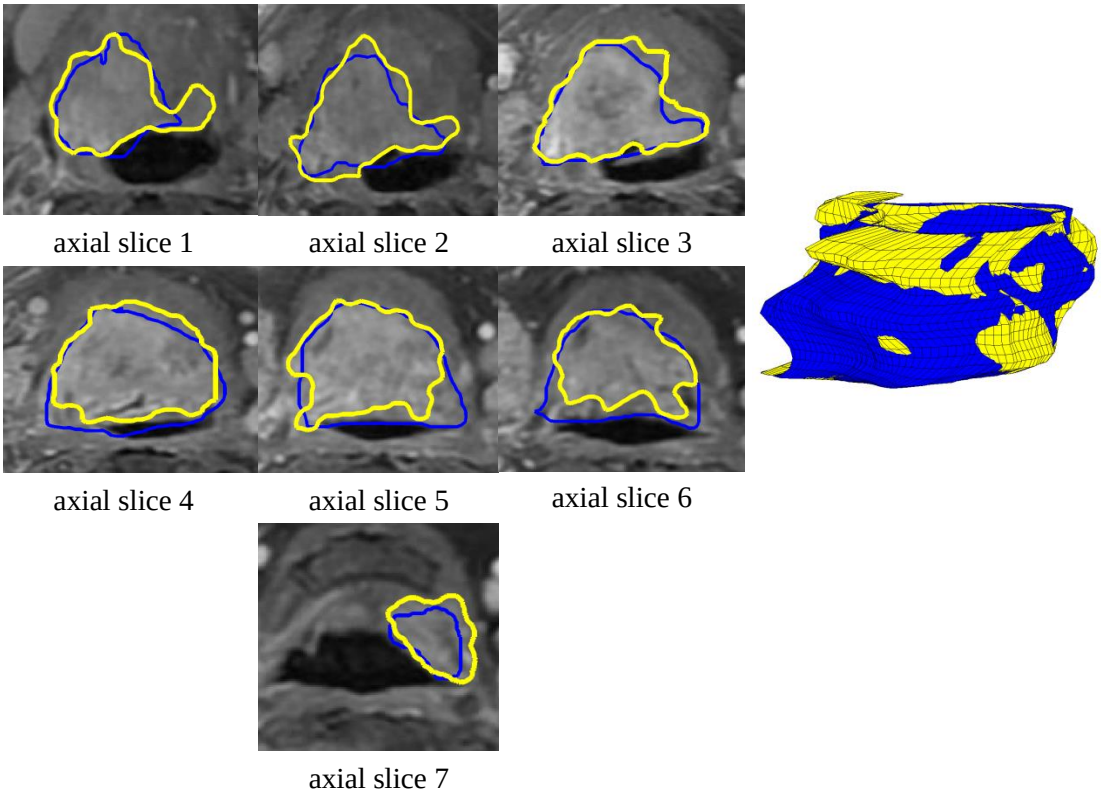


Fig. B.6: (a) Sequential axial MRI slices with C-auto (yellow) and C-gold (blue) tumour outlines for one patient; (b) 3D reconstructed volume from the tumour outlines shown in Fig. B.6(a) (C-auto reconstructed volume - yellow; C-gold reconstructed volume - blue)

References

[1] WHO Media Centre, 'WHO Cancer', Feb-2015. [Online]. Available: http://www.who.int/mediacentre/factsheets/fs297/en/. [Accessed: 24-Nov-2015].

[2] B. Menze, A. Jakab, S. Bauer, and et al., 'The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)', *IEEE Trans. Med. Imaging*, vol. 34, no. 10, pp. 1993–2024, Oct. 2015.

[3] S. Bauer, R. Wiest, L. Nolte, and M. Reyes, 'A survey of MRI-based medical image analysis for brain tumor studies', *Phys. Med. Biol.*, vol. 58, no. 13, pp. 97–129, Jul. 2013.

[4] J. Daisne and A. Blumhofer, 'Atlas-based automatic segmentation of head and neck organs at risk and nodal target volumes: a clinical validation', *Radiat. Oncol.*, vol. 8, no. 154, pp. 1–11, Jun. 2013.

[5] B. Fraass, K. Doppke, M. Hunt, and et al., 'American Association of Physicists in Medicine Radiation Therapy Committee Task Group 53: quality assurance for clinical radiotherapy treatment planning', *Med. Phys.*, vol. 25, no. 10, pp. 1773–1829, 1998.

[6] G. Walker, M. Awan, R. Tao, and et al., 'Prospective randomized double-blind study of atlas-based organ-at-risk autosegmentation-assisted radiation planning in head and neck cancer', *Radiother. Oncol.*, vol. 112, no. 3, pp. 321–325, Sep. 2014.

[7] I. S. Ramadaan, K. Peick, D. A. Hamilton, J. Evans, D. Iupati, A. Nicholson, L. Greig, and R. J. W. Louwe, 'Validation of Varian's SmartAdapt® deformable image registration algorithm for clinical application', *Radiat. Oncol.*, vol. 10, p. 73, 2015.

[8] X. Geets, J. Daisne, S. Arcangeli, and et al., 'Inter-observer variability in the delineation of pharyngo-laryngeal tumor, parotid glands and cervical spinal cord: comparison between CT-scan and MRI', *Radiother. Oncol.*, vol. 77, no. 1, pp. 25–31, Oct. 2005.

[9] G. Sharp, K. Fritscher, V. Pekar, and et al., 'Vision 20/20: perspectives on automated image segmentation for radiotherapy', *Med. Phys.*, vol. 41, no. 5, pp. 50902-1-50902–13, May 2014.

[10] O. Salvado, C. Hillenbrand, S. Zhang, and D. Wilson, 'Method to correct intensity inhomogeneity in MR images for atherosclerosis characterization', *IEEE Trans. Med. Imaging*, vol. 25, no. 5, pp. 539–552, May 2006.

[11] T. Doshi, J. Soraghan, D. Grose, K. MacKenzie, and L. Petropoulakis, 'Modified fuzzy c-means clustering for automatic tongue base tumour extraction from MRI data', in *European Signal Processing Conference*, Lisbon, Portugal, 2014, pp. 2460–2464.

[12] J. Liu, J. Udupa, D. Odhnera, J. McDonough, and R. Arens, 'System for Upper Airway Segmentation and Measurement with MR Imaging and Fuzzy Connectedness', *Acad. Radiol.*, vol. 10, no. 1, pp. 13–24, Jan. 2003.

[13] J. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*. Norwell, MA, USA: Kluwer Academic Publishers, 1981.

[14] J. Watkinson and R. Gilbert, *Stell & Maran's Textbook of Head and Neck Surgery and Oncology*, Fifth Edition. London: CRC Press, 2012.

[15] R. Corvo, 'Evidence-based radiation oncology in head and neck squamous cell carcinoma', *Radiother. Oncol.*, vol. 85, no. 1, pp. 156–170, Oct. 2007.

[16] J. Del Regato, 'The American Society of therapeutic radiologists. Introduction', *Cancer*, vol. 29, no. 6, pp. 1443–1445, Jun. 1972.

[17] P. Lauterbur, 'Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance', *Nature*, vol. 242, no. 5394, pp. 190–191, Mar. 1973.

[18] P. Mansfield, 'Multi-planar image formation using NMR spin echoes', *J. Phys. C Solid State Phys. Lett. Ed.*, vol. 10, no. 3, pp. L55–L58, 1977.

[19] D. McFarland, *Netter's Atlas of Anatomy for Speech, Swallowing, and Hearing*, Second Edition. Mosby, 2014.

[20] J. Ridge, B. Glisson, M. Lango, S. Feigenberg, and E. Horwitz, 'Head and neck tumors', *R Pazdur K Camphausen W Hoskins Ed.*, 2008.

[21] E. Ward and C. As-Brooks, *Head and Neck Cancer Treatment*, Second Edition. San Diego, CA: Plural Publishing Inc, 2014.

[22] V. Gregoire, J. Lefebvre, L. Licitra, E. Felip, and EHNS-ESMO-ESTRO Guidelines Working Group, 'Squamous cell carcinoma of the head and neck: EHNS-ESMO-ESTRO Clinical Practice Guidelines for diagnosis, treatment and follow-up', *Ann. Oncol.*, vol. 21, pp. 184–186, May 2010.

[23] A. Leslie, E. Fyfe, P. Guest, P. Goddard, and J. Kabala, 'Staging of squamous cell carcinoma of the oral cavity and oropharynx: a comparison of MRI and CT in T- and N-staging', *J. Comput. Assist. Tomogr.*, vol. 23, no. 1, pp. 43–49, Feb. 1999.

[24] M. Ahmed, M. Schmidt, A. Sohaib, and et al., 'The value of magnetic resonance imaging in target volume delineation of base of tongue tumours--a study using flexible surface coils', *Radiother. Oncol.*, vol. 94, no. 2, pp. 161–167, Feb. 2010.

[25] P. Zbaren, M. Becker, and H. Lang, 'Staging of laryngeal cancer: endoscopy, computed tomography and magnetic resonance versus histopathology', *Eur. Arch. Otorhinolaryngol.*, vol. 254, pp. 117–122, 1997.

[26] M. Donovan and M. Glackin, 'The lived experience of patients receiving radiotherapy for head and neck cancer: a literature review', *Int. J. Palliat. Nurs.*, vol. 18, no. 9, pp. 448–455, Sep. 2012.

[27] 'National Head and Neck Cancer Audit 2013, DAHNO Ninth Annual Report - NHS England Data Catalogue'. [Online]. Available: https://data.england.nhs.uk/dataset/national-head-and-neck-cancer-audit-open-data-2013/resource/289720df-2a34-41d1-8499-923baf54fb8b. [Accessed: 04-Oct-2015].

[28] E. Halperin, L. Brady, C. Perez, and D. Wazer, *Perez & Brady's Principles and Practice of Radiation Oncology*, Sixth Edition. Lippincott Williams and Wilkins, 2013.

[29] E. Paulson, B. Erickson, C. Schultz, and X. Allen Li, 'Comprehensive MRI simulation methodology using a dedicated MRI scanner in radiation oncology for external beam radiation treatment planning', *Med. Phys.*, vol. 42, no. 1, pp. 28–39, Jan. 2015.

[30] R. Sims, A. Isambert, V. Gregoire, and et al., 'A pre-clinical assessment of an atlas-based automatic segmentation tool for the head and neck', *Radiother. Oncol.*, vol. 93, no. 3, pp. 474–478, Dec. 2009.

[31] A. Gordon, L. Loevner, A. Shukla-Dave, and et al., 'Intraobserver variability in the MR determination of tumor volume in squamous cell carcinoma of the pharynx', *Am. J. Neuroradiol.*, vol. 25, no. 6, pp. 1092–1098, 2004.

[32] C. Clark, E. Miles, G. Urbano, and et al., 'Pre-trial quality assurance processes for an intensity-modulated radiation therapy (IMRT) trial: PARSPORT, a UK multicentre Phase III trial comparing conventional radiotherapy and parotid-sparing IMRT for locally advanced head and neck cancer', *Br. J. Radiol.*, vol. 82, no. 979, pp. 585–594, Jul. 2009.

[33] G. Urbano, C. Clark, C. Kong, and et al., 'Target volume definition for head and neck intensity modulated radiotherapy: pre-clinical evaluation of PARSPORT trial guidelines', *Clin. Oncol.*, vol. 19, no. 8, pp. 604–613, Oct. 2007.

[34] X. Liu, M. Niethammer, R. Kwitt, M. McCormick, and S. Aylward, 'Low-rank to the rescue - atlas-based analyses in the presence of pathologies', in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Boston, MA, USA, 2014, vol. 17, pp. 97–104.

[35] D. Schinagl, P. Span, F. van den Hoogen, and et al., 'Pathology-based validation of FDG PET segmentation tools for volume assessment of lymph node metastases from head and neck cancer', *Eur. J. Nucl. Med. Mol. Imaging*, vol. 40, no. 12, pp. 1828–1835, Dec. 2013.

[36] H. Veeraraghavan, M. Folkert, J. Deasy, and M. Traughber, 'Semi-Automated Probabilistic Segmentation of Head and Neck Anatomy Through Structure Specific Feature Selection From Multi-Sequence MRI', *Med. Phys.*, vol. 40, no. 6, pp. 536–536, Jun. 2013.

[37] H. Yu, C. Caldwell, K. Mah, and D. Mozeg, 'Coregistered FDG PET/CT-based textural characterization of head and neck cancer for radiation treatment planning', *IEEE Trans. Med. Imaging*, vol. 28, no. 3, pp. 374–383, Mar. 2009.

[38] E. Street, L. Hadjiiski, B. Sahiner, and et al., 'Automated volume analysis of head and neck lesions on CT scans using 3D level set segmentation', *Med. Phys.*, vol. 34, no. 11, pp. 4399–4408, Nov. 2007.

[39] D. Schinagl, W. Vogel, A. Hoffmann, J. van Dalen, W. Oyen, and J. Kaanders, 'Comparison of five segmentation tools for 18F-fluoro-deoxy-glucose-positron emission tomography-based target volume definition in head and neck cancer', *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 69, no. 4, pp. 1282–1289, Nov. 2007.

[40] D. Han, J. Bayouth, Q. Song, and et al., 'Globally optimal tumor segmentation in PET-CT images: a graph-based co-segmentation method', in *Information Processing in Medical Imaging Conference*, Kloster Irsee, Germany, 2011, vol. 22, pp. 245–256.

[41] J. Prince and J. Links, *Medical Imaging Signals and Systems*, Second Edition. Boston: Prentice Hall, 2014.

[42] J. Hornak, 'The Basics of MRI', 1996. [Online].
Available: https://www.cis.rit.edu/htbooks/mri/index.html. [Accessed: 06-Oct-2015].

[43] G. Liney, *MRI in clinical practice*. London: Springer-Verlag, 2006.

[44] Z. Liang and P. Lauterbur, *Principles of Magnetic Resonance Imaging: a Signal Processing Perspective*, First Edition. Bellingham, Wash. : New York: John Wiley & Sons, 1999.

[45] M. Graves and D. Mitchell, 'Body MRI artifacts in clinical practice: A physicist's and radiologist's perspective', *J. Magn. Reson. Imaging*, vol. 38, no. 2, pp. 269–287, Aug. 2013.

[46] U. Vovk, F. Pernus, and B. Likar, 'A Review of Methods for Correction of Intensity Inhomogeneity in MRI', *IEEE Trans. Med. Imaging*, vol. 26, no. 3, pp. 405–421, Mar. 2007.

[47] A. Simmons, P. Tofts, G. Barker, and S. Arridge, 'Sources of intensity nonuniformity in spin echo images at 1.5 T', *Magn. Reson. Med.*, vol. 32, no. 1, pp. 121–128, 1994.

[48] M. Ballester, A. Zisserman, and M. Brady, 'Estimation of the partial volume effect in MRI', *Med. Image Anal.*, vol. 6, no. 4, pp. 389–405, 2002.

[49] L. Czervionke, J. Czervionke, D. Daniels, and V. Haughton, 'Characteristic features of MR truncation artifacts', *Am. J. Neuroradiol.*, vol. 9, no. 5, pp. 815–824, 1988.

[50] R. Hermans, F. D. Keyzer, and V. Vandecaveye, 'Imaging Techniques', in *Head and Neck Cancer Imaging*, Springer Berlin Heidelberg, 2008, pp. 31–42.

[51] H. Lewis-Jones, R. Hanlon, and N. Roland, 'Head And Neck Imaging Protocols', NHS England Cheshire and Merseyside Strategic Clinical Network, Report, Aug. 2013.

[52] D. Pham, C. Xu, and J. Prince, 'Current methods in medical image segmentation', *Annu. Rev. Biomed. Eng.*, vol. 2, pp. 315–337, 2000.

[53] A. Trifas, 'Medical image enhancement', PhD Dissertation, University of Bucharest, Bucharest, Romania, 2005.

[54] R. Gonzalez and R. Woods, *Digital Image Processing*, Third Edition. Upper Saddle River, N.J: Prentice Hall, 2007.

[55] A. Al-Manea and A. El-Zaart, 'Contrast enhancement of MRI images', in *International Conference on Biomedical Engineering*, Kuala Lumpur, Malayasia, 2007, pp. 255–258.

[56] N. Senthilkumaran and J. Thimmiaraja, 'Histogram Equalization for Image Enhancement Using MRI Brain Images', in *World Congress on Computing and Communication Technologies*, Tiruchirappalli, India., 2014, pp. 80–83.

[57] T. Tan, K. Sim, and C. Tso, 'Image enhancement using background brightness preserving histogram equalisation', *Electron. Lett.*, vol. 48, no. 3, pp. 155–157, Feb. 2012.

[58] J. Paul, J. Mathew, and C. Kesavadas, 'MR image enhancement using an extended neighborhood filter', *J. Vis. Commun. Image Represent.*, vol. 25, no. 7, pp. 1604–1615, Oct. 2014.

[59] A. Khademi, A. Venetsanopoulos, and A. Moody, 'Image Enhancement and Noise Suppression for FLAIR MRIs With White Matter Lesions', *IEEE Signal Process. Lett.*, vol. 17, no. 12, pp. 989–992, Dec. 2010.

[60] R. Haralick and L. Shapiro, *Computer and Robot Vision*, First Edition. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1992.

[61] S. Smith and J. Brady, 'SUSAN—A new approach to low level image processing', *Int. J. Comput. Vis.*, vol. 23, no. 1, pp. 45–78, 1997.

[62] X. Yan, M. Zhou, L. Xu, W. Liu, and G. Yang, 'Noise Removal of MRI Data with Edge Enhancing', in *International Conference on Bioinformatics and Biomedical Engineering*, Wuhan, China, 2011, pp. 1–4.

[63] P. Hastreiter, R. Naraghi, B. Tomandl, M. Bauer, and R. Fahlbusch, '3D-visualization and registration for neurovascular compression syndrome analysis', in *Medical Image Computing and Computer-Assisted Intervention*, Tokyo, Japan, 2002, pp. 396–403.

[64] A. Srivastava, A. Raj, V. Bhateja, and et al., 'Combination of wavelet transform and morphological filtering for enhancement of magnetic resonance images', in *Digital Information Processing and Communications*, Ostrava, Czech Republic, 2011, pp. 460–474.

[65] A. Koivula, J. Alakuijala, and O. Tervonen, 'Image feature based automatic correction of low-frequency spatial intensity variations in MR images', *Magn. Reson. Imaging*, vol. 15, no. 10, pp. 1167–1175, 1997.

[66] E. Lewis and N. Fox, 'Correction of differential intensity inhomogeneity in longitudinal MR images', *NeuroImage*, vol. 23, no. 1, pp. 75–83, Sep. 2004.

[67] M. Styner, C. Brechbühler, G. Szckely, and G. Gerig, 'Parametric estimate of intensity inhomogeneities applied to MRI', *IEEE Trans. Med. Imaging*, vol. 19, no. 3, pp. 153–165, 2000.

[68] J. Mangin, 'Entropy minimization for automatic correction of intensity nonuniformity', in *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, Hilton Head Island, SC, USA, 2000, pp. 162–169.

[69] S. Sarra, 'Algorithm 899: The Matlab Postprocessing Toolkit', *ACM Trans Math Softw*, vol. 37, no. 1, p. 10:1–10:15, Jan. 2010.

[70] Y. Sato, C. Westin, A. Bhalerao, and et al., 'Tissue classification based on 3D local intensity structures for volume rendering', *IEEE Trans. Vis. Comput. Graph.*, vol. 6, no. 2, pp. 160–180, Apr. 2000.

[71] R. Boyes, J. Gunter, C. Frost, and et al., 'Intensity non-uniformity correction using N3 on 3-T scanners with multichannel phased array coils', *NeuroImage*, vol. 39, no. 4, pp. 1752–1762, Feb. 2008.

[72] W. Zheng, M. Chee, and V. Zagorodnov, 'Improvement of brain segmentation accuracy by optimizing non-uniformity correction using N3', *NeuroImage*, vol. 48, no. 1, pp. 73–83, Oct. 2009.

[73] T. Lehmann, C. Gonner, and K. Spitzer, 'Survey: interpolation methods in medical image processing', *IEEE Trans. Med. Imaging*, vol. 18, no. 11, pp. 1049–1075, Nov. 1999.

[74] S. Soleimanifard, M. Schar, A. Hays, R. Weiss, M. Stuber, and J. Prince, 'Vessel Centreline Tracking and Boundary Segmentation In Coronary MRA with Minimal Manual Interaction', in *IEEE International Symposium on Biomedical Imaging*, Barcelona, Spain, 2012, pp. 1417–1420.

[75] T. Schanze, 'Sinc interpolation of discrete periodic signals', *IEEE Trans. Signal Process.*, vol. 43, no. 6, pp. 1502–1503, Jun. 1995.

[76] P. Yushkevich, J. Piven, H. Hazlett, R. G. Smith, S. Ho, J. C. Gee, and G. Gerig, 'User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability', *NeuroImage*, vol. 31, no. 3, pp. 1116–1128, Jul. 2006.

[77] S. Campbell, T. Doshi, J. Soraghan, and et al., 'FFT based interpolation and 3D level set methods for 3D throat region segmentation from MRI data', presented at the Engineering in Medicine and Biology Conference, Milan, Italy, 2015, pp. 2419–2422.

[78] R. Zoroofi, Y. Sato, T. Sasama, and et al., 'Automated segmentation of acetabulum and femoral head from 3-d CT images', *IEEE Trans. Inf. Technol. Biomed.*, vol. 7, no. 4, pp. 329–343, Dec. 2003.

[79] J. Adams, ‘A subsequence approach to interpolation using the FFT’, *IEEE Trans. Circuits Syst.*, vol. 34, no. 5, pp. 568–570, May 1987.

[80] W. Hawkins, ‘FFT interpolation for arbitrary factors: a comparison to cubic spline interpolation and linear interpolation’, in *Nuclear Science Symposium and Medical Imaging Conference*, Norfolk, VA, 1994, vol. 3, pp. 1433–1437.

[81] J. Udupa and S. Samarasekera, ‘Fuzzy Connectedness and Object Definition: Theory, Algorithms, and Applications in Image Segmentation’, *Graph. Models Image Process.*, vol. 58, no. 3, pp. 246–261, May 1996.

[82] D. Comaniciu and P. Meer, ‘Mean shift: a robust approach toward feature space analysis’, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.

[83] J. Shi and J. Malik, ‘Normalized cuts and image segmentation’, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 888–905, 2000.

[84] C. Cortes and V. Vapnik, ‘Support-vector networks’, *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[85] R. Kindermann, J. Snell, and American Mathematical Society, *Markov random fields and their applications*. Providence, R.I.: American Mathematical Society, 1980.

[86] C. Xu, D. Pham, and J. Prince, ‘Image segmentation using deformable models’, in *Handbook of medical imaging*, vol. 2, Bellingham, WA: SPIE, 2000, pp. 129–174.

[87] M. Kass, A. Witkin, and D. Terzopoulos, ‘Snakes: Active contour models’, *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.

[88] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, 2003 Edition. New York: Springer, 2002.

[89] T. Cootes, C. Taylor, D. Cooper, and J. Graham, ‘Active Shape Models-Their Training and Application’, *Comput. Vis. Image Underst.*, vol. 61, no. 1, pp. 38–59, Jan. 1995.

[90] T. Cootes, G. Edwards, and C. Taylor, ‘Active appearance models’, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.

[91] Z. Raeesy, S. Rueda, J. Udupa, and J. Coleman, ‘Automatic segmentation of vocal tract MR images’, in *International Symposium on Biomedical Imaging*, San Francisco, CA, 2013, pp. 1328–1331.

[92] S. Silva and A. Teixeira, ‘Unsupervised segmentation of the vocal tract from real-time MRI sequences’, *Comput. Speech Lang.*, vol. 33, no. 1, pp. 25–46, Sep. 2015.

[93] E. Bresch and S. Narayanan, ‘Region Segmentation in the Frequency Domain Applied to Upper Airway Real-Time Magnetic Resonance Images’, *IEEE Trans. Med. Imaging*, vol. 28, no. 3, pp. 323–338, Mar. 2009.

[94] A. Javed, Y. Kim, M. Khoo, S. Ward, and K. Nayak, ‘Dynamic 3D MR Visualization and Detection of Upper Airway Obstruction during Sleep using Region Growing Segmentation’, *IEEE Trans. Biomed. Eng.*, no. 99, pp. 1–7, 2015.

[95] T. Ivanovska, E. Buttke, R. Laqua, H. Volzke, and A. Beule, ‘Automatic trachea segmentation and evaluation from MRI data using intensity pre-clustering and graph cuts’, in *International Symposium on Image and Signal Processing and Analysis*, Dubrovnik, Croatia, 2011, pp. 513–518.

[96] T. Ivanovska, J. Dober, R. Laqua, K. Hegenscheid, and H. Völzke, 'Pharynx Segmentation from MRI Data for Analysis of Sleep Related Disoders', in *Advances in Visual Computing*, Springer, 2013, pp. 20–29.

[97] R. Amendola, J. Reinhardt, Y. Sato, M. Zimmerman, H. Diggelmann, and D. Kacmarynski, 'Graph-based segmentation of the pediatric trachea in MR images to model growth', in *SPIE Medical Imaging*, Lake Buena Vista, Florida, USA, 2013, vol. 8672, pp. 867210-1-867210–8.

[98] M. Daneshzand, R. Zoroofi, and M. Faezipour, 'MR image assisted drug delivery in respiratory tract and trachea tissues based on an enhanced level set method', in *American Society for Engineering Education*, Bridgeport, Connecticut, US, 2014, pp. 1–7.

[99] S. Seifert, I. Wachter, G. Schmelzle, and R. Dillmann, 'A Knowledge-Based Approach to Soft Tissue Reconstruction of the Cervical Spine', *IEEE Trans. Med. Imaging*, vol. 28, no. 4, pp. 494–507, Apr. 2009.

[100] N. Gordillo, E. Montseny, and P. Sobrevilla, 'State of the art survey on MRI brain tumor segmentation', *Magn. Reson. Imaging*, vol. 31, no. 8, pp. 1426–1438, Oct. 2013.

[101] P. Gibbs, D. Buckley, S. Blackband, and A. Horsman, 'Tumour volume determination from MR images by morphological segmentation', *Phys. Med. Biol.*, vol. 41, no. 11, p. 2437, 1996.

[102] L. A. Meinel, T. Buelow, D. Huo, and et al., 'Robust segmentation of mass-lesions in contrast-enhanced dynamic breast MR images', *J. Magn. Reson. Imaging*, vol. 32, no. 1, pp. 110–119, Jul. 2010.

[103] V. Chong, J. Zhou, J. Khoo, J. Huang, and T. Lim, 'Tongue carcinoma: tumor volume measurement', *Int. J. Radiat. Oncol.*, vol. 59, no. 1, pp. 59–66, May 2004.

[104] V. Harati, R. Khayati, and A. Farzan, 'Fully automated tumor segmentation based on improved fuzzy connectedness algorithm in brain MR images', *Comput. Biol. Med.*, vol. 41, no. 7, pp. 483–492, Jul. 2011.

[105] H. Moftah, A. Hassanien, and M. Shoman, '3D brain tumor segmentation scheme using K-mean clustering and connected component labeling algorithms', in *International Conference on Intelligent Systems Design and Applications*, Cairo, Egypt, 2010, pp. 320–324.

[106] M. Clark, L. Hall, D. Goldgof, R. Velthuizen, F. Murtagh, and M. Silbiger, 'Automatic tumor segmentation using knowledge-based techniques', *IEEE Trans. Med. Imaging*, vol. 17, no. 2, pp. 187–201, Apr. 1998.

[107] M. Ahmed, S. Yamany, N. Mohamed, A. Farag, and T. Moriarty, 'A modified fuzzy c-means algorithm for bias field estimation and segmentation of MRI data', *IEEE Trans. Med. Imaging*, vol. 21, no. 3, pp. 193–199, Mar. 2002.

[108] K. Sikka, N. Sinha, P. Singh, and A. Mishra, 'A fully automated algorithm under modified FCM framework for improved brain MR image segmentation', *Magn. Reson. Imaging*, vol. 27, no. 7, pp. 994–1004, Sep. 2009.

[109] M. El-Melegy and H. Mokhtar, 'Tumor segmentation in brain MRI using a fuzzy approach with class center priors', *EURASIP J. Image Video Process.*, vol. 2014, no. 1, pp. 1–14, 2014.

[110] J. Zhou, S. Krishnan, V. Chong, and J. Huang, 'Extraction of tongue carcinoma using genetic algorithm-induced fuzzy clustering and artificial neural network from MR

images', in *IEEE Engineering in Medicine and Biology Conference*, San Francisco, CA, 2004, vol. 3, pp. 1790–1793.

[111] H. Khotanlou, O. Colliot, J. Atif, and I. Bloch, '3D brain tumor segmentation in MRI using fuzzy classification, symmetry analysis and spatially constrained deformable models', *Fuzzy Sets Syst.*, vol. 160, no. 10, pp. 1457–1473, May 2009.

[112] D. McClymont, A. Mehnert, A. Trakic, D. Kennedy, and S. Crozier, 'Fully automatic lesion segmentation in breast MRI using mean-shift and graph-cuts on a region adjacency graph', *J. Magn. Reson. Imaging*, vol. 39, no. 4, pp. 795–804, Apr. 2014.

[113] N. Birkbeck, D. Cobzas, M. Jagersand, A. Murtha, and T. Kesztyues, 'An interactive graph cut method for brain tumor segmentation', in *Workshop on Applications of Computer Vision*, Snowbird, Utah, 2009, pp. 1–7.

[114] A. Hamamci, N. Kucuk, K. Karaman, K. Engin, and G. Unal, 'Tumor-Cut: Segmentation of Brain Tumors on Contrast Enhanced MR Images for Radiosurgery Applications', *IEEE Trans. Med. Imaging*, vol. 31, no. 3, pp. 790–804, Mar. 2012.

[115] J. Zhou, K. L. Chan, P. Xu, and V. F. H. Chong, 'Nasopharyngeal carcinoma lesion segmentation from MR images by support vector machine', in *IEEE International Symposium on Biomedical Imaging*, Arlington, Virginia, US, 2006, pp. 1364–1367.

[116] J. Zhang, K. Ma, M. Er, and V. Chong, 'Tumor segmentation from magnetic resonance imaging by learning via one-class support vector machine', in *International Workshop on Advanced Image Technology*, Singapore, 2004, pp. 207–211.

[117] A. Ashraf, S. Gavenonis, D. Daye, and et al., 'A multichannel Markov random field approach for automated segmentation of breast cancer tumor in DCE-MRI data using kinetic observation model', in *Medical Image Computing and Computer-Assisted Intervention*, Toronto, Ontario, Canada, 2011, pp. 546–553.

[118] Y. Zhu and H. Yan, 'Computerized tumor boundary detection using a Hopfield neural network', *IEEE Trans. Med. Imaging*, vol. 16, no. 1, pp. 55–67, Feb. 1997.

[119] T. Wang, I. Cheng, and A. Basu, 'Fluid Vector Flow and Applications in Brain Tumor Segmentation', *IEEE Trans. Biomed. Eng.*, vol. 56, no. 3, pp. 781–789, Mar. 2009.

[120] S. Agner, J. Xu, and A. Madabhushi, 'Spectral embedding based active contour (SEAC) for lesion segmentation on breast dynamic contrast enhanced magnetic resonance imaging', *Med. Phys.*, vol. 40, no. 3, pp. 32305-1-32305–12, Mar. 2013.

[121] K. Xie, J. Yang, Z. Zhang, and Y. Zhu, 'Semi-automated brain tumor and edema segmentation using MRI', *Eur. J. Radiol.*, vol. 56, no. 1, pp. 12–19, Oct. 2005.

[122] T. Doshi, J. Soraghan, L. Petropoulakis, D. Grose, and K. MacKenzie, 'Semi-automatic segmentation of tongue tumors from magnetic resonance imaging', in *International Conference on Systems, Signals and Image Processing*, Bucharest, Romania, 2013, pp. 143–146.

[123] S. Ho, E. Bullitt, and G. Gerig, 'Level-set evolution with region competition: automatic 3-D segmentation of brain tumors', in *International Conference on Pattern Recognition*, Quebec City, Canada, 2002, vol. 1, pp. 532–535.

[124] K. Thapaliya, J. Pyun, C. Park, and G. Kwon, 'Level set method with automatic selective local statistics for brain tumor segmentation in MR images', *Comput. Med. Imaging Graph.*, vol. 37, no. 7–8, pp. 522–537, Oct. 2013.

[125] D. Cobzas, N. Birkbeck, M. Schmidt, M. Jagersand, and A. Murtha, '3D Variational Brain Tumor Segmentation using a High Dimensional Feature Set', in *International Conference on Computer Vision,* Rio de Janeiro, Brazil, 2007, pp. 1–8.

[126] T. Shepherd, S. Prince, and D. Alexander, 'Interactive lesion segmentation with shape priors from offline and online learning', *IEEE Trans. Med. Imaging,* vol. 31, no. 9, pp. 1698–1712, Sep. 2012.

[127] S. Lankton and A. Tannenbaum, 'Localizing Region-Based Active Contours', *IEEE Trans. Image Process.*, vol. 17, no. 11, pp. 2029–2039, Nov. 2008.

[128] M. Cuadra, C. Pollo, A. Bardera, O. Cuisenaire, J. Villemure, and J. Thiran, 'Atlas-based segmentation of pathological MR brain images using a model of lesion growth', *IEEE Trans. Med. Imaging,* vol. 23, no. 10, pp. 1301–1314, Oct. 2004.

[129] S. Bauer, C. Seiler, T. Bardyn, P. Buechler, and M. Reyes, 'Atlas-based segmentation of brain tumor images using a Markov random field-based tumor growth model and non-rigid registration', in *IEEE Engineering in Medicine and Biology Society Conference,* Buenos Aires, Argentina, 2010, vol. 1, pp. 4080–4083.

[130] M. Prastawa, 'An MRI segmentation framework for brains with anatomical deviations', PhD Dissertation, University of North Carolina, Chapel Hill, NC, USA, 2007.

[131] A. Asman and B. Landman, 'Out-of-atlas labeling: A multi-atlas approach to cancer segmentation', in *IEEE International Symposium on Biomedical Imaging,* Barcelona, Spain, 2012, pp. 1236–1239.

[132] M. Levoy, H. Fuchs, S. Pizer, and et al., 'Volume rendering in radiation treatment planning', in *Visualization in Biomedical Computing Conference,* Atlanta, GA, 1990, pp. 4–10.

[133] J. Marker, I. Braude, K. Museth, and D. Breen, 'Contour-based surface reconstruction using implicit curve fitting, and distance field filtering and interpolation', in *International Workshop on Volume Graphics,* Boston, Massachusetts, US, 2006, pp. 95–102.

[134] G. Barequet and M. Sharir, 'Piecewise-Linear Interpolation between Polygonal Slices', *Comput. Vis. Image Underst.*, vol. 63, no. 2, pp. 251–272, Mar. 1996.

[135] M. Arakeri and G. Reddy, 'An Effective and Efficient Approach to 3D Reconstruction and Quantification of Brain Tumor on Magnetic Resonance Images', *Int. J. Signal Process. Image Process. Pattern Recognit.*, vol. 6, no. 3, pp. 111–128, 2013.

[136] D. Wang, O. Hassan, K. Morgan, and N. Weatherill, 'Efficient surface reconstruction from contours based on two-dimensional Delaunay triangulation', *Int. J. Numer. Methods Eng.*, vol. 65, no. 5, pp. 734–751, Jan. 2006.

[137] D. Levin, 'Multidimensional reconstruction by set-valued approximations', *J. Numer. Anal.*, vol. 6, no. 2, pp. 173–184, 1986.

[138] V. Chong, 'Tumour volume measurement in head and neck cancer', *Cancer Imaging,* vol. 7, pp. S47–S49, Oct. 2010.

[139] H. Kim and W. Kim, 'Method of tumor volume evaluation using magnetic resonance imaging for outcome prediction in cervical cancer treated with concurrent chemotherapy and radiotherapy', *Radiat. Oncol.*, vol. 30, no. 2, pp. 70–77, Jun. 2012.

[140] A. Sorensen, S. Patel, C. Harmath, and et al., ‘Comparison of diameter and perimeter methods for tumor volume calculation’, *J. Clin. Oncol.*, vol. 19, no. 2, pp. 551–557, Jan. 2001.

[141] N. Mayr, T. Taoka, J. Buatti, and et al., ‘Method and timing of tumor volume measurement for outcome prediction in cervical cancer using magnetic resonance imaging’, *Int. J. Radiat. Oncol. Biol. Phys.*, vol. 52, no. 1, pp. 14–22, Jan. 2002.

[142] H. Gach, C. Tanase, and F. Boada, ‘2D and 3D Shepp-Logan Phantom Standards for MRI’, in *International Conference on Systems Engineering*, Las Vegas, NV, US, 2008, pp. 521–526.

[143] V. Magnotta and L. Friedman, ‘Measurement of Signal-to-Noise and Contrast-to-Noise in the fBIRN Multicenter Imaging Study’, *J. Digit. Imaging*, vol. 19, no. 2, pp. 140–147, Jun. 2006.

[144] J. Park, S. Jung, Y. Joo, C. Jung, K. Cho, and M. Kim, ‘Diagnostic accuracy of magnetic resonance imaging (MRI) in the assessment of tumor invasion depth in oral/oropharyngeal cancer’, *Oral Oncol.*, vol. 47, no. 5, pp. 381–386, May 2011.

[145] F. Sakai, S. Sone, K. Kiyono, and et al., ‘MR evaluation of laryngohypopharyngeal cancer: value of gadopentetate dimeglumine enhancement’, *AJNR Am. J. Neuroradiol.*, vol. 14, no. 5, pp. 1059–1069, Oct. 1993.

[146] L. Shepp and B. Logan, ‘The Fourier reconstruction of a head section’, *IEEE Trans. Nucl. Sci.*, vol. 21, no. 3, pp. 21–43, Jun. 1974.

[147] R. Van de Walle, H. Barrett, K. Myers, and et al., ‘Reconstruction of MR images from data acquired on a general nonregular grid by pseudoinverse calculation’, *IEEE Trans. Med. Imaging*, vol. 19, no. 12, pp. 1160–1167, Dec. 2000.

[148] L. Dice, ‘Measures of the Amount of Ecologic Association Between Species’, *Ecology*, vol. 26, no. 3, pp. 297–302, Jul. 1945.

[149] M. Dubuisson and A. Jain, ‘A modified Hausdorff distance for object matching’, in *International Conference on Pattern Recognition*, Jerusalem, Israel, 1994, vol. 1, pp. 566–568.

[150] F. Estrada and A. Jepson, ‘Benchmarking Image Segmentation Algorithms’, *Int. J. Comput. Vis.*, vol. 85, no. 2, pp. 167–181, Nov. 2009.

[151] R. McGill, J. Tukey, and W. Larsen, ‘Variations of Box Plots’, *Am. Stat.*, vol. 32, no. 1, pp. 12–16, 1978.

[152] T. Lumley, P. Diehr, S. Emerson, and L. Chen, ‘The importance of the normality assumption in large public health data sets’, *Annu. Rev. Public Health*, vol. 23, no. 1, pp. 151–169, 2002.

[153] H. Kraemer, C. Kraemer, and S. Thiemann, *How Many Subjects?: Statistical Power Analysis in Research*. Newbury Park: Sage Publications, 1987.

[154] T. Doshi, J. Soraghan, D. Grose, K. MacKenzie, and L. Petropoulakis, ‘Automatic detection of larynx cancer from contrast-enhanced magnetic resonance images’, presented at the SPIE Medical Imaging, Orlando, Florida, US, 2015, vol. 9414, p. 94142N–1–94142N–9.

[155] N. Otsu, ‘A Threshold Selection Method from Gray-Level Histograms’, *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.

[156] H. Freeman and R. Shapira, 'Determining the Minimum-area Encasing Rectangle for an Arbitrary Closed Curve', *Commun. ACM*, vol. 18, no. 7, pp. 409–413, Jul. 1975.

[157] A. Yezzi, A. Tsai, and A. Willsky, 'A Fully Global Approach to Image Segmentation via Coupled Curve Evolution Equations', *J. Vis. Commun. Image Represent.*, vol. 13, no. 1–2, pp. 195–216, Mar. 2002.

[158] R. Bartels, J. Beatty, and B. Barsky, *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling*. Morgan Kaufmann, 1995.

[159] K. Kadir, H. Gao, A. Payne, J. Soraghan, and C. Berry, 'Automatic quantification and 3D visualisation of edema in cardiac MRI', in *IEEE Engineering in Medicine and Biology Society Conference*, Boston, MA, USA, 2011, vol. 1, pp. 8021–8024.

[160] J. Mason, K. Hlavackova, and K. Warwick, 'Approximation Using Cubic B-Splines with Improved Training Speed and Accuracy', in *Computer Intensive Methods in Control and Signal Processing*, M. Karny and K. Warwick, Eds. Birkhauser Boston, 1997, pp. 295–303.

[161] K. Chuang, H. Tzeng, S. Chen, J. Wu, and T. Chen, 'Fuzzy c-means clustering with spatial information for image segmentation', *Comput. Med. Imaging Graph.*, vol. 30, no. 1, pp. 9–15, Jan. 2006.

[162] 'Shai Bagon's Matlab Code'. [Online]. Available: http://www.wisdom.weizmann.ac.il/~bagon/matlab.html. [Accessed: 17-Dec-2015].

[163] 'MATLAB Normalized Cuts Segmentation Code'. [Online]. Available: http://www.cis.upenn.edu/~jshi/software/. [Accessed: 17-Dec-2015].

[164] T. Doshi, G. Di Caterina, J. Soraghan, and et al., 'Combining interpolation and 3D level set method (I+3DLSM) for medical image segmentation', *Electron. Lett.*, vol. 52, no. 8, pp. 592–594, Apr. 2016.

[165] A. Lefohn, J. Kniss, C. Hansen, and R. Whitaker, 'Interactive deformation and visualization of level set surfaces using graphics hardware', *IEEE Trans Vis. Comput. Graph.*, vol. 10, no. 4, pp. 422–433, 2004.

[166] 'File Exchange - MATLAB Central'. [Online]. Available: http://uk.mathworks.com/matlabcentral/fileexchange/41666-fast-3d-2d-region-growing--mex-/content/RegionGrowing.m. [Accessed: 23-Dec-2015].