

# Tumour volume determination and three dimensional visualisation of head and neck cancers from clinical magnetic resonance images

Zhou, Jiayin

2005

Zhou, J. (2005). Tumour volume determination and three dimensional visualisation of head and neck cancers from clinical magnetic resonance images. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/4061>

<https://doi.org/10.32657/10356/4061>

---

Nanyang Technological University

*Downloaded on 20 Mar 2024 19:41:48 SGT*

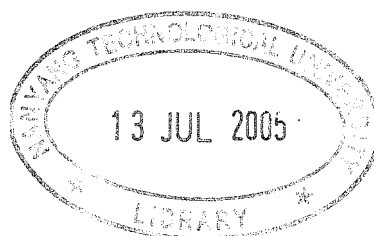
# **Tumour Volume Determination and Three Dimensional Visualisation of Head and Neck Cancers from Clinical Magnetic Resonance Images**

**Zhou Jiayin**

**School of Electrical & Electronic Engineering**

A thesis submitted to the Nanyang Technological University  
in fulfilment of the requirement for the degree of  
Doctor of Philosophy

**2005**



RC  
78.7  
.N83  
Z82  
2005

*To my dear wife and parents,  
for their love and encouragement.*

## Acknowledgments

I would like to express my deepest appreciation and gratitude to my research supervisors Dr Lim Tuan Kay, Dr Shankar Muthu Krishnan and Dr Chan Kap Luk of the Nanyang Technological University (NTU), and Dr Vincent Chong, Senior Consultant of the Singapore General Hospital (SGH), for their excellent scientific guidance, beneficial advice, kind and patient encouragement and support during the study of this research topic.

I would like to express my hearty thanks to the staff and other researchers at BMERC for their technical support and for providing a pleasant, productive, and academic environment throughout this research. Furthermore, I do appreciate the clinical support and the help on data retrieval by Dr James Khoo, Ms Chan Soh Hwa, and Mr Lim Fang Kiang from the National Cancer Centre of Singapore (NCCS). I also thank Mr Huang Jing from SGH for his valuable comments and suggestion.

Finally, I wish to acknowledge with thanks the financial support from the research scholarship, NTU and individual research grant of the Singapore National Medical Research Council (NMRC/0516/2001)



Table of Contents

<b>Acknowledgements</b>	<b>i</b>
<b>Summary</b>	<b>vii</b>
<b>List of Figures</b>	<b>ix</b>
<b>List of Tables</b>	<b>xv</b>
<b>Nomenclature</b>	<b>xvi</b>
<b>1. Introduction</b>	<b>1</b>
1.1 Motivation	1
1.2 Background	3
1.2.1 Cancers of Head and Neck	3
1.2.2 The State-of-the-Art in Imaging of Head and Neck Cancers	4
1.2.3 Impact of Volume on Tumour Staging and Prognosis	6
1.2.4 Determine the Volumes of Head and Neck Tumours	8
1.3 Objectives	9
1.4 Major Contributions of the Thesis	10
1.5 Organisation of the Thesis	12
<b>2. Literature Review</b>	<b>13</b>
2.1 Introduction	13
2.2 Imaging-based Tumour Volume Determination	14
2.3 MR Image Acquisition	18

Table of Contents	iii
2.3.1 Basic Principles of MRI	19
2.3.2 Image Acquisition	21
2.4 Methods of MRI Segmentation	23
2.4.1 Greyscale Single Image Segmentation	24
2.4.2 Multi-spectral Image Segmentation	26
2.4.3 Areas of Improvement in MRI Segmentation	28
2.5 Fundamentals of 3-D Visualisation in Medicine	29
2.5.1 3-D Visualisation Process and Algorithms	30
2.5.2 3-D Visualisation in Head and Neck Imaging	32
2.6 Concluding Remarks	34
<b>3. Semi-supervised MRI Segmentation and Tumour Volume Determination for Nasopharyngeal Carcinoma</b>	<b>36</b>
3.1 Introduction	36
3.2 Noise Removal and Initial Segmentation	40
3.2.1 Noise Removal	40
3.2.2 Initial Segmentation Using SFCM	41
3.3 Knowledge-based Image Analysis	47
3.3.1 Knowledge-based Image Analysis	47
3.3.2 Segmentation Results	50
3.4 NPC Tumour Volume Determination	52
3.4.1 Materials	53
3.4.2 Methodologies	53
3.4.3 Results	58
3.5 Discussion	63

Table of Contents	iv
3.5.1 Tumour Segmentation	63
3.5.2 NPC Tumour Volume Determination	66
3.6 Concluding Remarks	70
<b>4. Unsupervised MRI Segmentation and Tumour Volume Determination for Tongue Carcinoma</b>	<b>72</b>
4.1 Introduction	72
4.2 Initial Segmentation Using GAIFC	75
4.2.1 Determination of Primary Tissue Classes	75
4.2.2 Principle of Genetic Algorithm	76
4.2.3 GA-based Clustering for Coarse Classification	78
4.2.4 Fine Clustering Using FCM	83
4.3 ANN-based Symmetrical Detection and Refinement	84
4.3.1 Symmetrical Analysis for Mass Type Detection	84
4.3.2 Final Segmentation Results	90
4.4 Tongue Carcinoma Tumour Volume Determination	91
4.4.1 Materials	92
4.4.2 Methodologies	92
4.4.3 Results	94
4.5 Discussion	98
4.5.1 Unsupervised Segmentation	98
4.5.2 Potential Development in Segmentation	100
4.5.3 Tongue Carcinoma Tumour Volume Determination	102
4.5.4 Imaging Related Errors in Tumour Volume Determination	103

Table of Contents	v
4.6 Concluding Remarks	106
<b>5. Deformation Model-based MRI Segmentation</b>	<b>108</b>
5.1 Introduction	108
5.2 Region and Contour Deformation Models	111
5.2.1 Region Deformation Model	111
5.2.2 Contour Deformation Model	114
5.3 Application: NPC and Tongue Carcinoma Segmentation	117
5.3.1 NPC Segmentation	118
5.3.2 Tongue Carcinoma Segmentation	120
5.4 Application: Fast 3-D Segmentation	123
5.4.1 Segmentation Scheme	123
5.4.2 Experiments and Results	125
5.5 Discussion	127
5.6 Concluding Remarks	130
<b>6. 3-D Visualisation for Head and Neck Tumours</b>	<b>132</b>
6.1 Introduction	132
6.2 System Analysis of VTK	134
6.2.1 Overview	134
6.2.2 System Architecture of VTK	135
6.2.3 Object Models of VTK	136
6.2.4 Pipeline of VTK	139
6.3 Design of the Visualisation Platform	139
6.3.1 Overview	139

Table of Contents	vi
6.3.2 Module Descriptions	140
6.3.3 System Architecture	142
6.3.4 Implementation	143
6.4 Visualisation of Segmented Tumours	145
6.4.1 Inter-slice Interpolation	145
6.4.2 Selected Results	150
6.5 Visualisation of 3-D Medical Data	151
6.5.1 Visualisation of Brain Structure	152
6.5.2 Visualisation of NPC and Tongue Carcinoma	153
6.5.3 Visualisation of Brain Tumour	159
6.6 Discussion	162
6.6.1 Visualisation Platform and Techniques	162
6.6.2 3-D Medical Data Acquisition and Interpolation	165
6.7 Concluding Remarks	166
<b>7. Conclusions and Recommendations</b>	<b>168</b>
7.1 Conclusions	168
7.2 Recommendations for Future Work	171
<b>Author's Publications</b>	<b>173</b>
<b>Bibliography</b>	<b>175</b>
<b>Appendix: Medical Image File Format for the Study</b>	<b>193</b>

## Summary

Head and neck cancers are malignant tumours in the head and neck region. The accurate staging of such cancers is very important for prognosis and treatment. Recent findings show that tumour volume is an important factor for the diagnosis, treatment planning and prognosis of head and neck cancers. Imaging methods such as computed tomography (CT) and magnetic resonance imaging (MRI) play an important role in the evaluation of head and neck cancers. However, there is no quantitative imaging-based tumour volumetric analysis for head and neck tumours.

In this study, an analytical framework including novel MRI segmentation algorithms and a three-dimensional (3-D) visualisation platform is developed for the quantitative 3-D volumetric evaluation of head and neck tumours.

A semi-supervised, knowledge-based fuzzy clustering (KBFC) method is developed to segment nasopharyngeal carcinoma (NPC) from multispectral MR images. The algorithm integrates a fuzzy clustering and an image analysis procedure. The guidance of *a priori* knowledge in both the training set selection and refinement procedure gives this method additional power and flexibility. This method achieves successful lesion extraction for NPC and significantly higher inter-operator reproducibility than the manual tracing method.

In addition, an unsupervised hierarchical method is developed to segment tongue carcinoma from multispectral MR images. This segmentation has two steps. Step one is an initial segmentation using a genetic algorithm-induced fuzzy clustering (GAIFC). In step two, the segmented mass is fed into a neural network classifier for symmetry detection and necessary refinement is performed to reduce false-positives. Experimental results show that in comparison with the ground truth, satisfactory

## Summary

---

segmentation results are obtained using this method. The two algorithms presented are employed to clinically determine the tumour volume of NPC and tongue carcinoma.

Furthermore, deformation models are applied to the fast segmentation of lesion horn single MR images. In this two-step segmentation scheme, region deformation and active contour models are employed to locate the actual lesion boundary with the help of an initial outline. The method shows robustness when dealing with heterogeneous image, high noise and other contaminations. Moreover, a 3-D segmentation scheme is derived based on this method for the fast segmentation of tumour and lymph node from a 3-D data volume formed by a stack of two-dimensional (2-D) slices.

Besides these segmentation algorithms, a 3-D visualisation platform, “Voxur-3D”, is established for the generation of 3-D views for 3-D medical data. The visualisation platform consists of functions of surface rendering for segmented medical objects and volume rendering for 3-D raw medical data, respectively. The results show that the developed platform is able to produce 3-D views with a high quality for 3-D medical data on brain, head and neck tumours.

Based on the results obtained, it is concluded that the developed framework, consisting of several novel MRI segmentation algorithms and a 3-D visualisation platform, provides a feasible solution to the quantitative 3-D volumetric analysis on head and neck tumours. It is believed that this research effort paves the way to offer computer-assisted techniques for clinical management of head and neck tumour cases.

## List of Figures

1.1	Anatomic structures of head and neck region	3
1.2	Axial CT image shows the disappearance of pharyngeal recess caused by NPC (black arrow).	5
1.3	Axial MR image shows the high intensity lesion of tongue carcinoma (black arrow).	<b>5</b>
1.4	The visualisation of bone and nasopharyngeal angiofibroma (black arrow)	5
2.1	Four MR images with tongue carcinoma from the same patient	16
2.2	(a) 1-D diameter and 2-D diameters measurement; (b) manual tumour area tracing.	<b>16</b>
2.3	Diagram of an MRI system	19
2.4	(a) In the absence of a strong magnetic field, hydrogen nuclei are randomly aligned; (b) When the strong magnetic field $B_0$ is applied, the hydrogen nuclei precess about the direction of the field.	20
2.5	(a) The RF pulse, $B_{rf}$ , causes the net magnetic moment of the nuclei, $M$ , to tilt away from $B_0$ ; (b) When the RF pulse stops, the nuclei return to equilibrium such that $M$ is again parallel to $B_0$ . During realignment, the nuclei lose energy and a measurable RF signal.	21
2.6	Components of an image analysis system	24
2.7	Elementary process of 3-D visualisation	30
2.8	Shaded volume rendering of pelvis	32
2.9	Shaded surface rendering of pelvis	32



List of Figures	x
3.1 (a) Local spread of NPC in sagittal view; (b) Local spread of NPC in axial view.	37
3.2 (a) T1W image; (b) CET1W image; (c) Average joint histogram of 12 pairs of tumour-represented MRI slices.	38
3.3 The head mask (a) Before the morphological operation; (b) After the operation.	41
3.4 GUI for training data selection	46
3.5 (a) T1W image; (b) CET1W image; (c) GT; (d) Initial segmentation: 1. soft palate, 2. mucosa, 3. tumour, and the dash line is the symmetric axis; (e) Final result.	49
3.6 (a) T1W image; (b) CET1W image; (c) GT; (d) Initial segmentation: 1. soft palate, 2. mucosa, 3. tumour, and the dash line is the symmetric axis; (e) Final result.	49
3.7 (a) T1W image; (b) CET1W image; (c) GT; (d) Segmentation using KBFC; (e) Segmentation using MLM.	51
3.8 (a) T1W image; (b) CET1W image; (c) GT; (d) Segmentation using KBFC; (e) Segmentation using MLM.	52
3.9 MR images of normal brain (a) T1w image; (b) T2W image.	64
3.10 MR images of brain tumour (a) T1w image; (b) CET1W image; (c) T2W image.	64
3.11 GT volume and PM distribution	67
3.12 GT volume and CR distribution	67
4.1 Tongue carcinoma	73
4.2 Greatest diameter of tongue carcinoma	73

List of Figures	xi
4.3 Flowchart of a hierarchical MRI segmentation scheme	75
4.4 (a) T1W image; (b) T2W image; (c) Joint histogram (after normalisation) of the ROI.	77
4.5 Location map of the GA classified cluster centres	82
4.6 Initial segmentation result of the ROI shown in Figure 4.3	84
4.7 (a) T1W image; (b) T2W image; (c) Initial result.	84
4.8 (a) T1W image; (b) T2W image; (c) Extracted sublingual glands.	85
4.9 (a) T1W image; (b) T2W image; (c) Extracted tumour and involved sublingual gland.	85
4.10 The illustration of the structures of tumour-involved sublingual glands	87
4.11 Structure of a BPNN	88
4.12 The training process of the BPNN	89
4.13 (a) Final result of images in Figure 4.9; (b) GT.	90
4.14 (a) T1W image; (b) T2W image; (c) GT; (d) Initial segmentation result; (e) Final segmentation result.	90
4.15 (a) T1W image; (b) T2W image; (c) GT; (d) Initial segmentation result; (e) Final segmentation result.	91
4.16 (a) T1W MR image shows the NPC morphologically in the nasopharynx (white arrow); (b) However, in the corresponding T1W post-contrast MR image, a part of the tumour does not show high signal intensity, due to partial volume effect (white arrow). This may cause errors in tumour segmentation.	104
4.17 A T2W MR image with RF inhomogeneity artefact (white arrow)	105

List of Figures	xii
4.18 A T2W MR image of tongue carcinoma with typical motion artefact	106
5.1 Object described as a region with a closed curve as its boundary	113
5.2 The iteration of greedy algorithm	116
5.3 (a) Original CET1 W MR image; (b) Initial boundary; (c) Result after region deformation; (d) Result after contour deformation; (e) Boundary traced by radiologist; (f) Result using seeded region growing.	119
5.4 (a) Original CET1 W MR image; (b) Initial boundary; (c) Result after region deformation; (d) Result after contour deformation; (e) Boundary traced by radiologist; (f) Result using seeded region growing.	120
5.5 (a) Original T2W MR image with motion blur; (b) Initial boundary; (c) Result after region deformation; (d) Result after contour deformation; (e) Boundary traced by radiologist; (f) Result using the unsupervised hierarchical segmentation method.	121
5.6 (a) Original T2W MR image with Gaussian blur; (b) Initial boundary; (c) Result after region deformation; (d) Result after contour deformation; (e) Boundary traced by radiologist; (f) Result using JSBFC method.	122
5.7 Block diagram of fast 3-D segmentation	124
5.8 Illustration of applied fast 3-D segmentation scheme	125
5.9 (a) Initial plan on slice 15; (b) Final boundary extracted using region and contour deformations on slice 15; (c) Initial plan on slice 16.	125
5.10 Final results from slice 12 to slice 17 (a-f), respectively.	126

## List of Figures

---

5.11	Selected segmentation results from six slices (a)-(f) Slice 1, 4, 7, 10, 13, and 16 of 18 slices, respectively.	127
5.12	A CETIW MR image of brain tumour with two different initial plans	129
6.1	System frame of VTK-based applications	136
6.2	The relationship of graphical models in VTK	138
6.3	Data pipeline of VTK	139
6.4	System architecture of the visualisation platform	142
6.5	Clinical surface rendering pipeline	143
6.6	Clinical volume rendering pipeline	145
6.7	Illustration of inter-slice interpolation	146
6.8	Sketch of shape morphing	149
6.9	Interpolated slices between two original segmented tumour slices (white star)	149
6.10	Selected interpolated slices among five original segmented tumour slices (white star)	150
6.11	3-D views of tongue carcinoma (a) Without smoothing; (b) With Gaussian smoothing and standard deviation=2.0; (c) With Gaussian smoothing and standard deviation=5.0.	150
6.12	3-D views of NPC and brain tumour (a) Another NPC without interpolation; (b) NPC with interpolation; (c) Brain tumour without interpolation.	151
6.13	(a) 2-D slice; (b) 3-D left view; (c) 3-D back view; (d) 3-D overhead view.	152
6.14	(a) 3-D left view; (b) 3-D back view; (c) 3-D overhead view; (d)	153

List of Figures	xiv
Cropped data.	
6.15 (a) 2-D slice of NPC (black arrow); (b) 3-D view; (c) 3-D view of cropped data; (d) Overhead view of this data after the adjustment of intensity-opacity function.	154
6.16 (a) 3-D view; (b) and (c) 3-D views of cropped data from different viewpoints; (d) Overhead view of this data after the adjustment of intensity-opacity function.	155
6.17 (a) 2-D slice of NPC with bone erosion; (b)-(d) 3-D views from different viewpoints with different settings of intensity-opacity function.	156
6.18 (a) 3-D view of cropped data at skin-fat window; (b) Overhead view of cropped data at soft tissue window; (c) Front view at soft tissue-bone window; (d) Bottom view at bone window.	158
6.19 (a) 2-D slice of tongue carcinoma with enlarge cervical lymph node; (b) 3-D views of this data; (c) 3-D view after fence cropping; (d) 3-D view after the adjustment of intensity-opacity function.	159
6.20 (a) 2-D slice of brain aneurysm; (b)-(d) 3-D views from the same viewpoint with different settings of intensity-opacityfunction.	160
6.21 (a) 2-D slice of brain tumour; (b)-(d) 3-D views from the same viewpoint with different settings of intensity-opacity function and cropping.	161
6.22 3-D view of cervical lymphoma	164

## List of Tables

1.1	Imaging-based approaches for tumour volume determination	8
2.1	The calculation of volumes by different methods	17
2.2	Comparison of volume rendering and surface rendering	32
3.1	Volume determination in phantom model consisting of two components	59
3.2	Results of NPC tumour volume determination	60
3.3	Percentage matching (PM) and correspondence ratio (CR) calculation for patient 1	61
3.4	Percentage match and correspondence ratio comparing manual tracing and KBFC segmentation of two operators	61
3.5	Inter-operator variance (IOV) calculation for patient 1	62
3.6	Comparison of inter-operator reliabilities between manual tracing and KBFC segmentation	62
4.1	Volume determination in phantom model	95
4.2	Results of tongue carcinoma tumour volume determination	96
4.3	Percentage match and correspondenceratio comparing manual tracing and the unsupervised hierarchical segmentation of two operators	96
4.4	Comparison of inter-operator reliabilities between manual tracing and the unsupervised hierarchical segmentation	97

## Nomenclature

### Symbols

$A$ (2.2)	The tumour's maximum cross-sectional diameter
$A$ (5.2)	The number of pixels which form object $O$
$Abs$	The absolute operation
$A_i$	The measured tumour area of slice $i$
$A_M$	The manually traced tumour mass
$A_S$	The semi-automatically segmented tumour mass
$B$ (2.2)	The maximal perpendicular diameter to $A$
$B$ (5.2)	The boundary of the object for extraction
<b>B</b>	The circular structuring element
$B_0$	The strength of the applied magnetic field
$B_{rf}$	The radio-frequency pulse applied perpendicular to $B_0$
$c$	The significance level in KS test
<b>c</b>	The integer sequence in the determination of $c$
$C$	The number of slices making up the tumour
$d$ (3.3)	The Mahalanobis distance measure
$d$ (5.2)	The low case $d$ in Kolmogrov-Smirnov test
$d(x, y, z)$	The shortest distance from voxel $(x, y, z)$ to the boundary of object <b>O</b>
$D$ (2.2)	The diameter of maximum cross-sectional area of the tumour
$D$ (5.2)	The Kolmogrov-Smirnov distances
<b>D</b>	The distance map of <b>I</b>

---

$DC$	The distance-of-centres
$DCM$	The distance-of-center-to-midline
$DM$	The difference-of-masses
$D_{max}$	The shape difference between the two objects
$DS$	The difference-of-standard deviations
$N[.]$	The current voxel and its 4-neighbor voxels
$E_{int}$	The internal energy of a snake
$E_{ext}$	The external energy of a snake
$E_{snake}$	The energy of a snake
$f$	The fitness function in GA
$F_B$	The greyscale cumulative frequency distribution function of $B$
$F_O$	The greyscale cumulative frequency distribution function of $O$
$FSM$	The fuzzy symmetric measure
$F(P(g))$	The fitness function of gene generation $P(g)$
$g$	The generation in GA
$g_i$	The greyscale of pixel $i$
$G$	The gap between two consecutive slices
$G_\sigma$	The Gaussian function with standard deviation $\sigma$
$H_B$	The greyscale frequency distribution function of $B$
$H_O$	The greyscale frequency distribution function of $O$
$i$	A pixel in image $I$
$I$	The entire image
$I$	The image matrix
$I'$	The noise free image matrix



Nomenclature		xviii
$IA_M$	The value of index of agreement from manual tracing measurement	
$IA_S$	The value of index of agreement from semi-automated measurement	
$IOV_M$	The value of inter-operator variance from manual tracing method	
$IOV_S$	The value of inter-operator variance from semi-automated method	
$J_m$	The objective function for optimisation	
$L$	The number of pixels which form boundary $B$	
$M(2.3)$	The net magnetic moment parallel to $B_0$	
$M(4.2)$	The clustering metric	
$\mathbf{M}(3.2)$	The initial binary head mask	
$\mathbf{M}(3.3)$	The mean vector	
$\mathbf{M}'$	The noise free binary head mask	
$Max$	The maximum value operation	
$ME$	The tumour volume measurement error	
$n$	The number of slices of the object	
$O$	The object for extraction	
$\mathbf{O}$	The object region in $\mathbf{I}$	
$p$	The conditional probability of a feature vector if it is in a class	
$P$	The <i>a priori</i> probability for a class	
$P_c$	The crossover probability in GA	
$P_m$	The mutation probability in GA	
$P(g)$	The gene population at generation $g$	
$\mathbf{P}_k$	The variable weight vector in FCM	
$R^2$	The 2-D Euclidean space	
$s$	The arc length	

Nomenclature		xix
$S_j$	The image segment $j$	
$S_k$	The image segment $k$	
$T$	The thickness of a slice	
$u_{ik}$	The fuzzy membership of subset $k$ for tissue class $i$	
$\mathbf{U}$	The fuzzy membership matrix	
$volume$	The volume of a tumour	
$V_{GT}$	The tumour volume obtained by using manual tracing method	
$\mathbf{v}_s$	The first derivative of $\mathbf{v}$	
$\mathbf{v}_{ss}$	The second derivative of $\mathbf{v}$	
$\mathbf{v}(s)$	The parametric equation of the contour	
$\mathbf{v}_i$	The discrete format of the contour	
$\mathbf{V}_i$	The cluster centres vector for tissue class $i$	
$V_S$	The tumour volume obtained by semi-automated method	
$W_j$	The small window with pixel $j$ being the center	
$x_{kj}$	The $j^{th}$ feature of the $k^{th}$ subset	
$\mathbf{X}$	The finite object data set	
$\mathbf{X}_k$	The feature vector of the $k^{th}$ subset	
$\mathbf{y}$	The fuzzy label	
$\mathbf{Z}$	The matrix of cluster centres	
$\mathbf{Z}^*$	The matrix of new cluster centres	
$\gamma$	The gyromagnetic ratio	
$\delta$	The mutation factor in GA	
$\nabla$	The gradient operator	
$\nabla^2$	The second-order gradient operator	

Nomenclature		XX
$\epsilon$	The threshold value in FCM	
$\lambda_1$	The crossover factor 1 in GA	
$\lambda_2$	The crossover factor 2 in GA	
$\Sigma$	The covariance matrix	
$\omega_0$	Larmor frequency	

Abbreviations

1-D	One-dimensional
2-D	Two-dimensional
3-D	Three-dimensional
AFCM	Adaptive fuzzy <i>c</i> -means
AJCC	American Joint Committee on Cancer
ANN	Artificial neural network
ANOVA	Analysis-of-variance
AVW	A Visualisation Workbench
BPNN	Back propagation neural network
CET1W	Contrast enhanced T1-weighted
CFDF	Cumulative frequency distribution functions
CR	Corresponding ratio
CT	Computed tomography
DDP	Default display protocol
DICOM	Digital Imaging and Communicationsin Medicine
EM	Expectation-maximisation

Nomenclature		xxi
FCM	Fuzzy $c$ -means	
FDF	Frequency distribution functions	
FID	Free-induction decay	
FN	False negative	
FP	False positives	
GA	Genetic algorithm	
GAIFC	Genetic algorithm-included fuzzy clustering	
GT	Ground truth	
GS	Golden standard	
GUI	Graphic user interface	
HCM	Hard $c$ -means	
IA	Index of agreement	
IDL	Interactive Data Language	
IOV	Inter-operator variance	
KBFC	Knowledge-based fuzzy clustering	
KNN	$k$ nearest neighbours	
KS	Kolmogrov-Smirnov	
ME	Measurement error	
ML	Maximum likelihood	
MLM	Maximum likelihood method	
MR	Magnetic resonance	
MRA	Magnetic resonance angiography	
MRF	Markov random field	
MRI	Magnetic resonance imaging	

Nomenclature		xxii
MTL	Manual tracing and labelling	
NEMA	National Electrical Manufacturers Association	
NIH	National Institutes of Health	
NPC	Nasopharyngeal carcinoma	
PDW	Proton density-weighted	
PET	Positron emission tomography	
PM	Percent matching	
ROI	Region of interest	
RF	Radio-frequency	
SFCM	Semi-supervised fuzzy <i>c</i> -means	
SPGR	Spoiled gradient recalled echo	
SVDD	Support vector data description	
SVM	Support vector machine	
TNM	Tumour Node Metastasis	
TN	True negative	
TOF	Time-of-flight	
TP	True positive	
T1W	T1-weighted	
T2W	T2-weighted	
UHS	Unsupervised hierarchical segmentation	
UICC	International Union Against Cancer	
VTK	Visualization Toolkit	

# Chapter 1

## Introduction

### 1.1 Motivation

Malignant tumours of the head and neck can originate from the nasal cavity, paranasal, sinus, lip, salivary glands, oral cavity, oropharynx, nasopharynx and larynx. Compared to breast, prostate and colon cancers, head and neck cancers are not common in western countries. For example, in the United States, the overall incidence rate of head and neck cancers is 23.5 per 100,000 population and these cancers account for 4.9% of all cancers in the United States [1]. In Europe, the overall incidence rate of head and neck cancers is 24.3 per 100,000 population and these cancers account for 6.9% of all cancers in Europe [2]. However in Asian countries, these cancers are much more common and some are leading cancers in these regions. According to a report of Indian National Cancer Registry Programme, from six registry areas in India, the median percentage of head and neck cancers to all cancers is 14.7% [3]. Cancers of oral cavity, tongue and hypopharynx are among the 10 leading cancers in India [3]. In Malaysia, head and neck cancers account for 12.6% of all cancers and in addition, nasopharyngeal carcinoma (NPC) ranks second and

constitutes 8% of all cancers in males [4]. In Singapore NPC is the fifth most frequently encountered cancer [5].

The accurate prognostication and staging of head and neck cancers is very important in treatment planning. The Tumour Node Metastasis (TNM) classification system of the International Union Against Cancer (UICC) and the American Joint Committee on Cancer (AJCC) is widely adopted for the staging of malignant tumours. That is, malignant tumours are jointly staged by the evaluations of primary tumour (T-staging), local nodal metastasis (N-staging) and distant metastasis (M-staging). According to the TNM system, there are two methods of T-staging for head and neck tumours: 1. By measuring the tumour diameter (lip and oral cavity, oropharyngeal, hypopharyngeal, major salivary glands and thyroid gland cancers); 2. By analyzing the anatomical extent based on the structures involved (nasopharyngeal, paranasal sinuses and laryngeal cancers) [6]. Therefore, the quantitative estimation of tumour volume is important for tumour staging and prognosis.

The present study is motivated by the following: 1. Tumour volume is an important prognostic indicator for head and neck cancers, even for those staged by anatomic extension; 2. There is a paucity of reports on quantitative tumour volume determination methods for head and neck neoplasms using imaging-based techniques; 3. Tumours in main subsites of head and neck region are difficult to evaluate clinically. Although computed tomography (CT) and magnetic resonance imaging (MRI) provide important information of the internal structures, they are actually two-dimensional (2-D) methods. These images cannot give three-dimensional (3-D) vision information of tumour infiltration; 4. Computerised processing of medical images may enable physicians and radiologists to produce better imaging analysis; and 5. The progress of visualisation techniques in scientific computing and computer graphics make it possible to provide 3-D visual information of organ system and disease process.

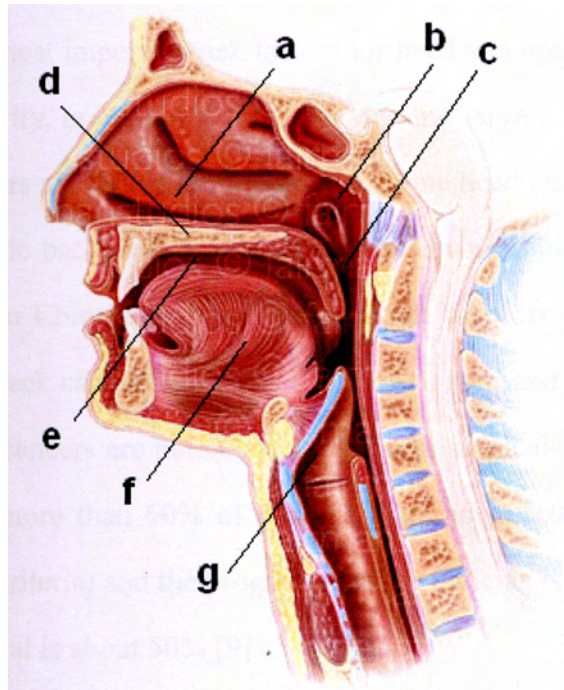
## 1.2 Background

---

## 1.2 Background

### 1.2.1 Cancers of Head and Neck

The sagittal view of head and neck region is shown in Figure 1.1.



**Figure 1.1** Anatomic structures of head and neck region

a. Nasal cavity; b. Nasopharynx; c. Pharynx; d. Palate; e. Oral cavity; f. Tongue; g. Larynx.

Head and neck cancers are often referred to as squamous cell carcinomas which occur in the following subsites:

- Oral cavity
- Salivary glands
- Paranasal sinuses and nasal cavity
- Pharynx—The pharynx has three parts: nasopharynx, oropharynx and hypopharynx.



- Larynx
- Lymph nodes in the upper part of the neck
- Thyroid

Two-thirds to three-quarters of the patients are male and one-quarter to one-third are female. The etiology of head and neck cancers is still not clear. However, tobacco and alcohol use are the most important risk factors for head and neck cancers, particularly those of the oral cavity, oropharynx, hypopharynx and larynx. Eighty five percent of head and neck cancers are linked to tobacco use. Some head and neck cancers depend strongly on the ethnic background of the patients. For example, NPC is much more common in Southern Chinese and their descendants in other parts of the world. In addition, head and neck cancers are more common in men and in people over age 50 [7]. Head and neck cancers are potentially curable by surgical-, chemo- and/or radio-therapy. However, more than 60% of patients have advanced diseases (T3 and T4 according to UICC criteria) and the prognosis for the patients is unfavourable [8]. The overall 5-year survival is about 50% [9].

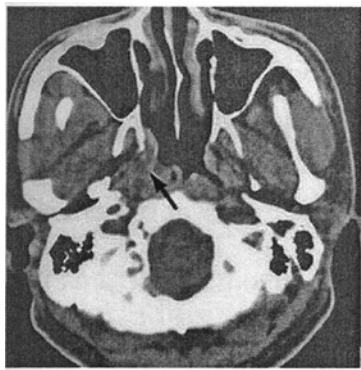
### **1.2.2 The State-of-the-Art in Imaging of Head and Neck Cancers**

Therapeutic decision is based on accurate assessment of primary tumour location and disease extent (T-staging) as well as the presence and extent of lymph node involvement (N-staging). Because of the anatomic characteristics of head and neck region, clinical methods like inspection, palpation and endoscopy are insufficient to assess the true disease extent. Therefore, CT and MRI are important complementary techniques in staging process [8].

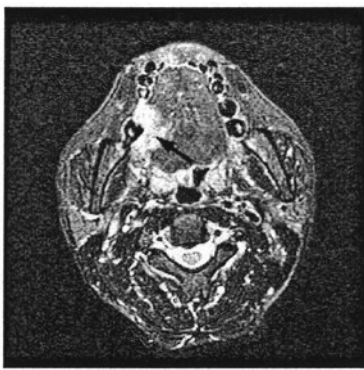
CT provides excellent delineation of the cortical bony structures of the skull base and facial skeleton. It is also very useful for demonstrating lymph node enlargement.

By selecting a bone or soft tissue window, the examiner can highlight CT image details that are relevant to a particular inquiry [10]. The CT characteristics of head and neck tumours include: 1. density changes of soft tissue; 2. the abnormality of structure morphology; 3. submucosa and deep soft tissue changes, and 4. the tumour invasion of adjacent organs [11]-[13]. An axial CT image of NPC is shown in Figure 1.2.

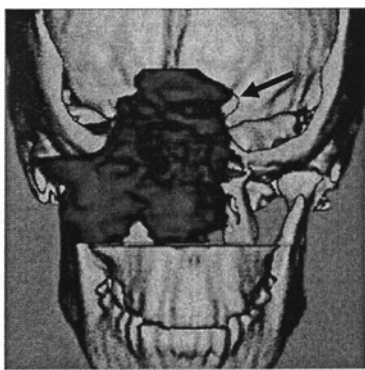
MRI is superior to CT because of better soft tissue delineation. Its multiplanar capabilities aids in the accurate demonstration of various lesions in relation to normal organ structures [10]. The MRI characteristics of head and neck tumours include: 1. Signal intensity changes of primary tumour; 2. the abnormality of structure morphology (This part is similar to that of CT.), 3. the tumour invasion of adjacent tissue, and 4. the metastatic involvement of lymph nodes [11]-[14]. Although MRI produces no bony details, the excellent soft tissue resolution makes MRI a more suitable modality (compared to CT) in the detection of head and neck tumours [12]-[18]. However, there were different reports on the detection of recurrent tumour and post-therapeutic changes with CT and MRI [11][ 19]. An axial MR image of tongue carcinoma is shown in Figure 1.3.



**Figure 1.2** Axial CT image shows the disappearance of pharyngeal recess caused by NPC (black arrow).



**Figure 1.3** Axial MR image shows the high intensity lesion of tongue carcinoma (black arrow).



**Figure 1.4** The visualisation of bone and nasopharyngeal angiofibroma (black arrow)

In clinical practice, CT and MRI provide 2-D images. Radiologists and Clinicians have to view a stack of consecutive slices and then “construct” the 3-D information in their mind. Visualisation is the processing that shows a volumetric data set on a display plane “directly”. It is helpful to visualise the complex anatomic and pathological structures in head and neck region, from a volumetric data set, for accurate diagnosis and treatment planning. The potential benefits include: 1. 3-D models allow the surgeon to anticipate problems during the procedure, such as the close proximity of blood vessels or nerves to tumour; 2. Surgical or radiotherapy planning, for example, the best patient position, incision, and method of resection; 3. Improved delineation of structures may lead to improved prosthetic devices, including implantable ones for the temporal bone [20]. An example of 3-D visualisation is shown in Figure 1.4

### 1.2.3 Impact of Volume on Tumour Staging and Prognosis

Accurate tumour staging is crucial because it (1) aids treatment planning; (2) gives some indication of prognosis; (3) assists the evaluation of treatment results; (4) facilitates the exchange of information between treatment centres; and (5) contributes to the continuing investigation of human cancer.

As mentioned in Section 1.1, there are currently two methods of T-staging for head and neck tumours: By tumour measurement and by the anatomic extension of tumour. Hence, the tumour size is an important staging parameter in many solid tumours and the largest axial tumour diameter has been used for many years in the TNM staging system [6]. Solid tumours are three-dimensional structures with unequal rates of tumour spread in different directions and in different planes. Hence the largest axial diameter may not reflect the total tumour volume and the total burden of malignant cells. For instance, superficial spreading tongue carcinomas frequently

exceed 4 cm in diameter without deep penetration. These tumours are classified as T3 lesions but have very low volumes. Sorensen, et al. have demonstrated that the differences in volumes derived from diameter measurement and computer assisted perimeter method were large enough to have an impact in gauging treatment response [21]. Pameijer, et al. found considerable variability of tumour volumes in T3-staged head and neck tumours. They suggested that the current TNM classification system is unable to group tumours with the same size into the same stage group and tumour volume analysis could be a useful parameter in the future research of head and neck tumours [22]. Furthermore, staging based on the judgment of tumour extension is also questionable. Chua, et al. found large tumour volume variation in different T stage disease of NPC, and that might be partly due to the limitation of using current staging system to segregate patients into large and small tumour bulk [23]. The study of Chang, et al. demonstrated the considerable variability in primary tumour volume of NPC and they suggested the further refinement of the 1997 UICC/AJCC staging system by incorporation of primary tumour volume [24].

Tumour volume has also been well recognised as one of the major prognostic factors in the treatment of malignancy, as increasing tumour bulk means increasing number of tumour clonogens that need to be sterilised [25]-[27]. This observation was also confirmed in studies relating tumour volume and treatment outcome in head and neck cancers. The study by Johnson, et al. demonstrated that the total tumour volume appears to be more predictive than AJCC clinical stage in advanced head and neck cancers and quantitative tumour volume determinations might prove to be a useful parameter in future analyses of head and neck cancers [28]. Kurek, et al. suggested that as a prognostic factor of survival, the knowledge of the initial tumour volume should be included in all future clinical trials regarding head and neck cancer patients [29]. Chua, et al. and Willner, et al. found that primary tumour volume is a very important and independent prognostic factor influencing the local control in NPC

[23][30]. The study of Chang, et al. suggested that volumetric measurement of primary tumours in early T-stage NPC would better refine the TNM classification system [31]. Sze, et al. found that the gross volume of primary tumour and involved retropharyngeal nodes is a strongly significant factor for predicting local control of NPC [32]. In addition, some studies showed that the depth of tumour (tumour thickness) is related to nodal metastasis in tongue carcinoma. Hence tumour thickness might be a better prognostic parameter compared to T-staging (largest diameter) for the prediction of nodal metastasis and treatment outcome [33]-[37]. Since tumour volume can vary considerably even with one tumour stage, tumour volume should always be considered as an independent prognostic parameter.

1.2.4 Determine the Volumes of Head and Neck Tumours

How to determine the volume of a solid tumour? Apart from the direct water displacement method using resected specimen, the most common way is by the imaging-based approaches. These approaches are summarised in Table 1.1 and their detailed descriptions and comparison are shown in Section 2.2.

Table 1.1 Imaging-based approaches for tumour volume determination

Approaches	Description
One-dimensional (1-D) largest diameter method	The largest tumour diameter in a cross-sectional slice is used to estimate the volume of a tumour.
Two-dimensional (2-D) largest diameters method	The largest tumour diameter and its maximal perpendicular diameter in a cross-sectional slice are used to estimate the volume of a tumour.
Three-dimensional (3-D) planimetry measurement method	The summation of the tumour areas from tumour bearing cross-sectional slices are used to calculate the volume of a tumour.

Little research work has been performed systemically during the last decade on imaging-based volumetric analysis for head and neck tumours. Among those studies involving volume measurement for head and neck tumours, planimetry measurement with manually delineated tumour outline, from either CT or MRI, is the most commonly employed method [22][24][28]-[31][33]-[37]. In some studies, tumour infiltration or extent was first manually traced by clinicians in diagnostic MRI or CT. The tumour outlines were then transcribed to the planning CT and the tumor volume was calculated by the radio-therapy planning system, taking into account the irregular tumour shape commonly seen in head and neck cancers [23][32][38]. It can be inferred that the 3-D planimetry measurement is considered as the method with the highest accuracy while the 1-D and 2-D methods are not satisfying for clinical research and application purposes. However currently, the tumour area is obtained by the clinician's manual tracing, hence there may be considerable subjectivities which result in intra- and inter-operator variances. Furthermore, the tracing work is really tedious and tough in a large-scale clinical investigation. Therefore, it has great significance to develop image segmentation algorithms and schemes to extract the tumour area from medical images, for determining the tumour volume of head and neck cancers.

### 1.3 Objectives

The major aims of the present study are to investigate, develop and validate effective segmentation algorithms and schemes to solve the key problems in the MRI-based 3-D tumour volume determination for head and neck cancers, and at the same time to develop an effective and PC-based system for 3-D tumour volume visualisation.

The potential benefits of the research project include the followings. A robust and reliable tumour volume determination method will serve as a quantitative tool for

tumour staging and radiation treatment evaluation. For tumours staged by measuring a single dimension (largest diameter), volume determination will provide more accurate information. In addition, a 3-D visualisation system will make the determination and assessment more directly perceivable.

In order to achieve these objectives of tumour volume determination and visualisation, the following aspects of MR image segmentation and 3-D visualisation are developed in the present work:

- A semi-supervised multi-spectral method for NPC segmentation from MRI,
- An unsupervised hierarchical method for tongue carcinoma segmentation from MRI,
- Boundary detection-based segmentation method and its application for 3-D MR segmentation, and
- 3-D visualisation for head and neck tumours including data interpolation and system design.

## 1.4 Major Contributions of the Thesis

Imaging-based tumour volumetric analysis is a popular research field in many universities, medical centres, and research institutes. Segmentation is a comparative problem because segmentation methods are highly dependent on: 1. tumour size, site and type; 2. stage, vascularity and related degree of MR contrast enhancement; 3. imaging sequence and parameters; 4. other factors that may affect MR relaxation parameters and the ability to serially differentiate tissues within the tumour bed [39]. It can be stated that a universally accepted segmentation method for tumour volumetric analysis is not presently available. A review of the recent literature showed



that most of the MR segmentation-based tumour volume determinations were reported in brain tumours and pelvic neoplasms, or for the evaluation of brain morphology. Currently there is a paucity of reports on MRI segmentation for head and neck tumours, as well as a quantitative tumour volumetric analysis for head and neck neoplasms.

In the present work, an intelligent MR image processing, analysis and visualisation framework was developed for image segmentation-based tumour volume determination and 3-D visualisation. Besides the necessary image pre-processing, three novel MRI segmentation methods for head and neck tumours were integrated in this framework:

- A semi-supervised multi-spectral MRI segmentation method using knowledge-based fuzzy clustering (KBFC) for NPC.
- An unsupervised hierarchical multi-spectral MRI segmentation method for tongue carcinoma. This algorithm includes two consecutive procedures: an initial segmentation using genetic algorithm-included fuzzy clustering (GAIFC) and an artificial neural network (ANN)-based symmetry detection/refinement.
- A boundary detection-based MRI segmentation for tumour and metastasised lymph node using region deformation and active contour model.

Furthermore, a shape-based interpolation method was applied to slice segmentation to reconstruct 3-D volume for visualisation. A visualisation platform including both surface and volume renderings, was developed as a part of the whole system with the support of the Visualization Toolkit (VTK).



## 1.5 Organisation of the Thesis

The organisation of this thesis is made in the following manner. Background aspects and literature review on imaging-based tumour volume determination, MR image acquisition, MRI segmentation and 3-D visualisation are provided in Chapter 2. Semi-supervised MRI segmentation for NPC using knowledge-based fuzzy clustering is described in Chapter 3. In this chapter, experiment of NPC tumour volume determination and results are presented and discussed. Unsupervised hierarchical MRI segmentation for tongue carcinoma is presented in Chapter 4. In this chapter, experiment of tongue carcinoma tumour volume determination and the corresponding results are also presented and discussed. Chapter 5 deals with the boundary detection-based MRI segmentation using region deformation and active contour model and its application on 3-D tumour segmentation. Chapter 6 covers the 3-D visualisation for head and neck tumours including theories, system implementations, and test results. In Chapter 7, conclusions and recommendations for future work are stated. The contributions made by the present study are also highlighted in Chapter 7. In the end, the author's publications as well as the bibliography are given and the file format of medical image is demonstrated in Appendix.

## Chapter 2

# Literature Review

## 2.1 Introduction

Tumour volume determination can be used for tumour staging and treatment planning. Besides this, it is also expected to improve clinical management in oncology and to facilitate the development and evaluation of anti-neoplasm drugs and other treatment modalities in cancer patients. The research on imaging-based tumour volumetric analysis and 3-D visualisation started in the early 1990s due to the rapid progress in image processing, pattern recognition, digital imaging, information technology, and computer graphics. In 1992, the National Institutes of Health (NIH), USA, released a call for proposals inviting research grant applications to advance the methods of imaging-based tumour volumetric analysis for optimisation of response assessment in cancer therapy [40]. The National Electrical Manufacturers Association (NEMA, USA) listed imaging-based quantitative analysis and 3-D visualisation as one of the future strategic directions of Digital Imaging and Communications in Medicine (DICOM). The Working Group 17 (3-D) of the DICOM Standards Committee is in charge of the representative applications of visualisation of volumetric and cine-

volumetric as well as diagnostic quantitative and assisted detection using automated segmentation and classification [41]. For the research topics concerning imaging-based 3-D tumour volumetric analysis, there are three closely intertwined basic scientific areas of highest priority, including automated segmentation or “tumour edge definition”, multimodality image registration/fusion (e.g., CT, positron emission tomography (PET), conventional and metabolic chemical shift MRI, and immunoimaging), and the design and implementation of visualisation algorithms and platforms.

The work in this study was based on MR images only. Therefore the thesis is focused on MRI segmentation and 3-D medical visualisation. In this chapter, the techniques and latest developments of tumour volume analysis, MR image acquisition, MR image segmentation and 3-D medical visualisation are reviewed and summarised.

## 2.2 Imaging-based Tumour Volume Determination

Before the digital imaging system was widely used in hospitals, some simple measurements such as one-dimensional (1-D) or 2-D greatest tumour diameters were performed on transilluminated films directly. In 1990, Mahaley, et al. reported a film-based tumour boundary tracing with the assistance of digitizing tablet and 16-button hand-held cursor connected to a microcomputer [42]. In 1993, interferon beta-1b received regulatory approval from FDA for the treatment of multiple sclerosis on the basis of disease activity monitored with MRI of the brain. From then on, more digital imaging devices were put into use and radiologic and digital evaluations of tumour treatment response during clinical testing have become increasingly important [43]. It is well recognised that imaging-based volumetric measurement represents the most accurate way in the assessment of tumour size. In addition, the relative immobility of solid tumour minimises the motion artefact during the process of scanning.

In general, based on CT and MR images, there are three methods used to quantify the tumour volume, as stated below [44].

a. Spherical geometric measurement (1-D)

In this method, the tumour is assumed to be a sphere with the volume given by

$$volume = \pi D^3/6, \quad (2-1)$$

where  $D$  is the diameter of maximum cross-sectional area of the tumour.

b. ABC ellipsoid geometric measurement (2-D)

In this method, the tumour is assumed to be an ellipsoid with the volume given by

$$volume = \pi ABC/6, \quad (2-2)$$

where  $A$  is the tumour's maximum cross-sectional diameter,  $B$  is the maximal perpendicular diameter to  $A$ , and  $C$  is the number of slices making up the tumour.

c. Planimetry measurement

In this method, the tumour volume is given by

$$volume = \sum_{i=1}^n A_i(T + G), \quad (2-3)$$

where  $n$  is the number of slices of the object,  $A_i$  is the measured tumour area of slice  $i$ ,  $T$  is the thickness of a slice, and  $G$  is the gap between two consecutive slices.

Obviously, for perfectly spherical lesions, the tumour diameter, cross-sectional area, and volume are mathematically related and either method a or b can be used for determining the tumour size. Unfortunately, biology does not produce perfect geometric shapes, and tumours commonly cause invaginations in the surrounding tissues, which makes tumour determination as much as an art rather than a science. Chua, et al. believed that 1-D or 2-D diameter measurement method may be less

applicable in tumours that tend to be infiltrative in growth pattern with an often highly irregular tumour contour such as NPC and other head and neck tumours [23]. An example is given to show the difference among the tumour volumes obtained using the three above mentioned methods by the same radiologist on the same data. Figure 2.1 shows the four MR images (slice thickness, 5 mm and slice gap, 1.5 mm) with tongue carcinoma. 1-D diameter and 2-D diameters measurements were performed on slice 3, as shown in Figure 2.2(a). Figure 2.2(b) demonstrates the manual tumour area tracing, also on slice 3. The calculations of corresponding volumes are summarised and compared in Table 2.1. The results show that the volumes produced by 1-D and 2-D methods have considerable variations to that produced by planimetry measurement.

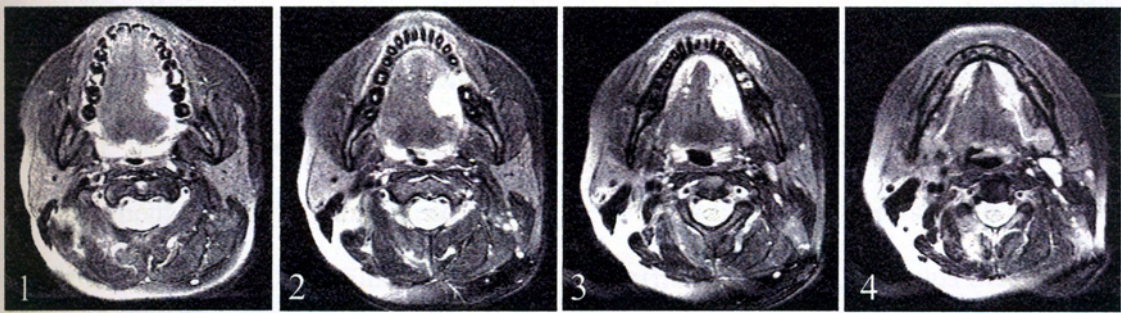


Figure 2.1 Four MR images with tongue carcinoma from the same patient

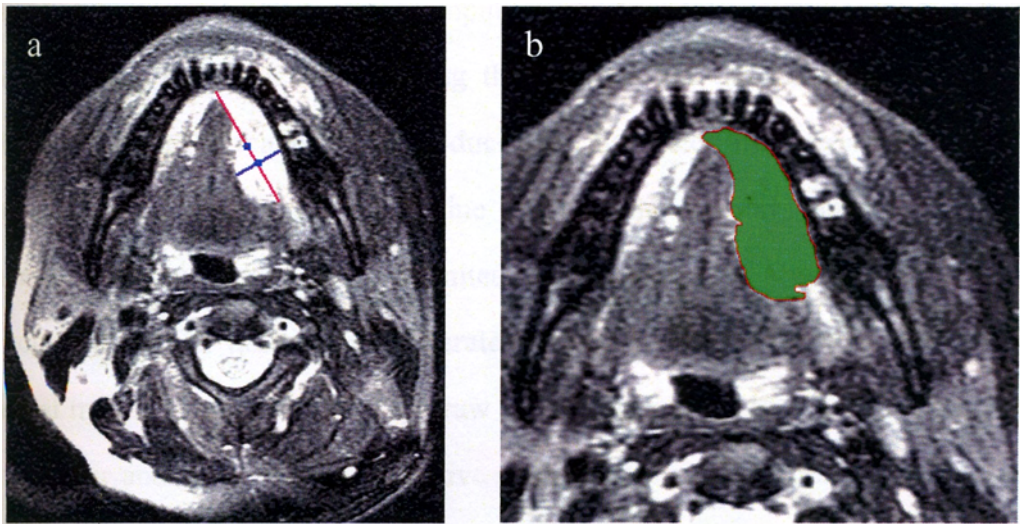


Figure 2.2 (a) 1-D diameter and 2-D diameters measurement; (b) manual tumour area tracing.

**Table 2.1** The calculation of volumes by different methods

	Method a	Method b	Method c
<i>What is known</i>		$C = 4(0.5+0.15) = 2.6 \text{ cm}$	$T = 0.5 \text{ cm}$ $G = 0.15 \text{ cm}$
<i>What is measured</i>	$D = 3.95 \text{ cm}$	$A = 3.95 \text{ cm}$ $B = 1.63 \text{ cm}$	$A_1 = 2.82 \text{ cm}^2$ $A_2 = 4.18 \text{ cm}^2$ $A_3 = 5.62 \text{ cm}^2$ $A_4 = 2.99 \text{ cm}^2$
<i>Volume (cm<sup>3</sup>)</i>	32.3	8.8	10.1

In addition, for methods a and b, whether imaging software is used or not, the tumour outline needs to be traced and the diameter need to be measured manually, which is potentially subjective and prone to large variations in intra- and inter-operator performance. Thus, the reliability, reproducibility, and accuracy of the measurement cannot be guaranteed [45]. It is obvious that method c is more accurate than the other two. If the values of  $T$  and  $G$  are sufficiently small, it means the number of slices  $n$  making up the object is large and according to the principle of calculus, this result is an estimation of the actual volume.

The keys for getting satisfactory computing results include computing  $A_i$  accurately and acquiring more slices containing the tumour. For the latter, although current helical CT and MRI scanners can produce very thin slices, the z-axis resolution is not very high in routine examination due to cost and scanning time considerations. Therefore, the number of slices is limited. For computing  $A_i$  accurately, it is required to segment tumour tissue more accurately fi-om the images. In some early studies, manual tracing was widely used to draw the tumour outline to count the pixels inside the contour and a few semi-quantitative methods such as Cavalieri’s direct estimator [46] were used. In recent years, several semi-automated or automated methods were proposed for tumour volume determination using computer image segmentation.



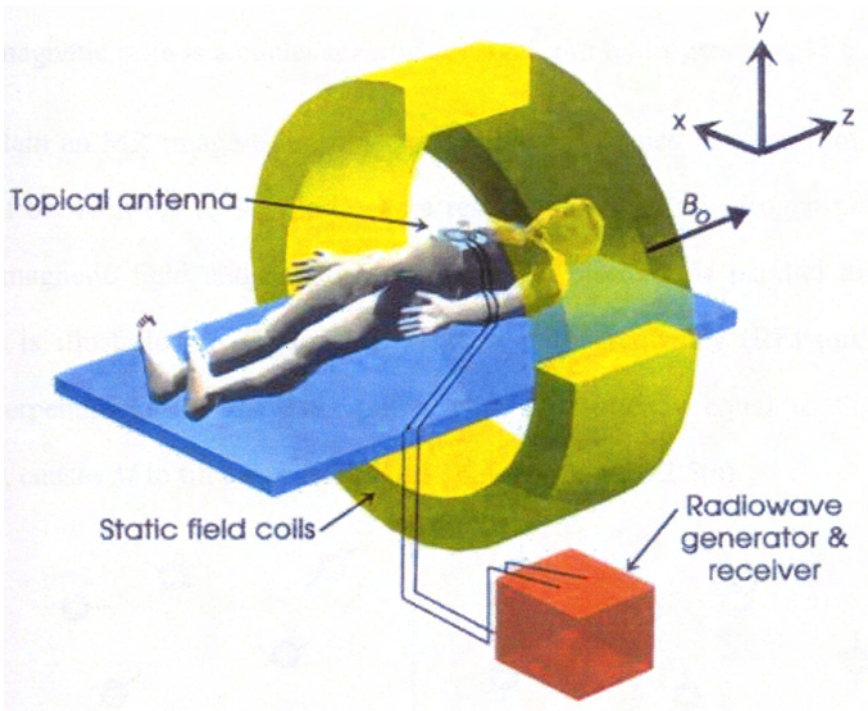
Some examinations on brain tumours, breast cancer and cervical cancer were performed to compare the results of manual and semi-automated/automated methods for tumour volume determination. The results showed that compared to manual methods, single semi-automated method has lower intra-operator and inter-operator errors and requires less operation time. On the other hand, for the same images, different semi-automated methods have variable results which may be partly due to different segmentation algorithms [45][47]-[57]. Fully automated methods almost have no intra-operator and inter-operator errors. In addition, the elimination of human supervision makes them suitable to process large image volumes. However, the measurement accuracy of automated methods is sometimes less than that of semi-automated methods. In addition, the automated methods often require long processing time [58]-[60].

Overall, the accuracy of  $A_i$  depends on image segmentation since it is the most critical step in imaging-based tumour volumetric analysis. The imaging-based tumour volume determination is actually the image segmentation-based tumour volume determination.

## 2.3 MR Image Acquisition

MRI technology has developed into a versatile and clinically useful diagnostic imaging modality since its inception in the early 1970s. It has become the diagnostic imaging modality of choice for many injuries and diseases. Dr Paul Lauterbur & Sir Peter Mansfield have been awarded the Nobel Prize in Physiology or Medicine 2003 for their discoveries concerning MRI. In contrast to X-ray CT, MRI is a non-invasive imaging technique that does not use ionizing radiation. Rather MRI is based on perturbing magnetic fields with radiowaves. MRI procedures can be manipulated in a number of ways to produce selected contrast between different anatomic structures.

MRI system provide mechanisms for intricate control of the signal being measured through modulation of the magnetic field and radiofrequency pulse sequence used to alter the spins of protons in the structure being imaged. MRI selectively images the distributions of protons, hence it is an excellent soft tissue imaging modality, providing highly detailed structural images [61]. Techniques have rapidly developed for fast MRI to capture functional characteristics, such as metabolism, flow, evoked neurofunctional response and biochemical composition of tissue (using MR spectroscopy) [62]. Figure 2.3 shows the diagram of an MRI system.



**Figure 2.3** Diagram of an MRI system

### 2.3.1 Basic Principles of MRI

The basis of MRI is the directional magnetic field, or moment, associated with charged particles in motion. Nuclei containing an odd number of protons and/or neutrons have a characteristic motion or precession. Because nuclei are charged



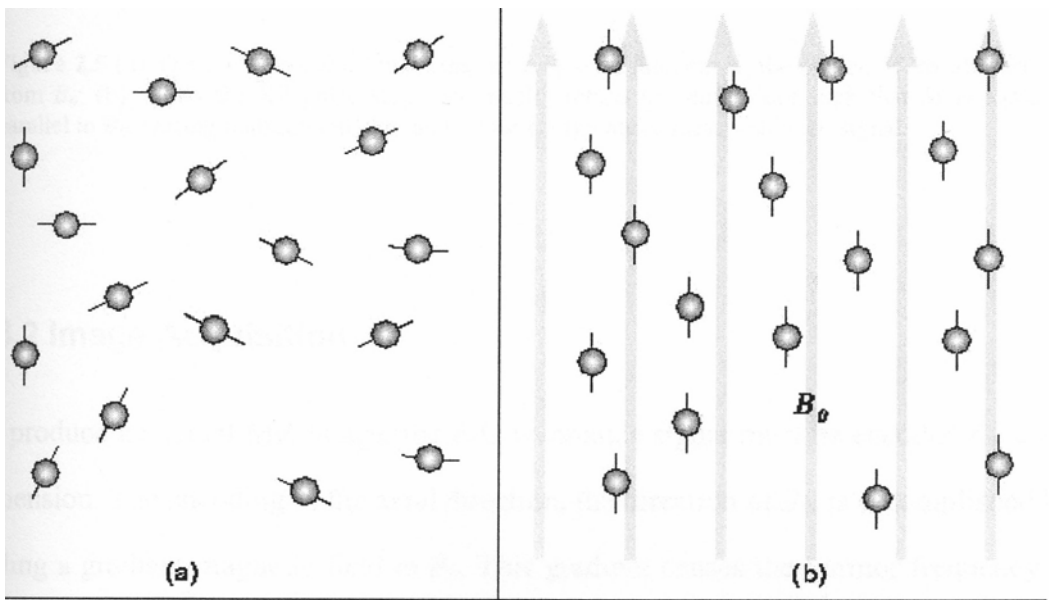
particles, this precession produces a small magnetic moment. When a human body is placed in a large magnetic field, many of the free hydrogen nuclei align themselves with the direction of the magnetic field. The nuclei precess about the magnetic field direction like gyroscopes. This behaviour is termed Larmor precession.

The frequency of Larmor precession is proportional to the applied magnetic field strength as defined by the Larmor frequency,  $\omega_0$

$$\omega_0 = \gamma B_0 \tag{2-4}$$

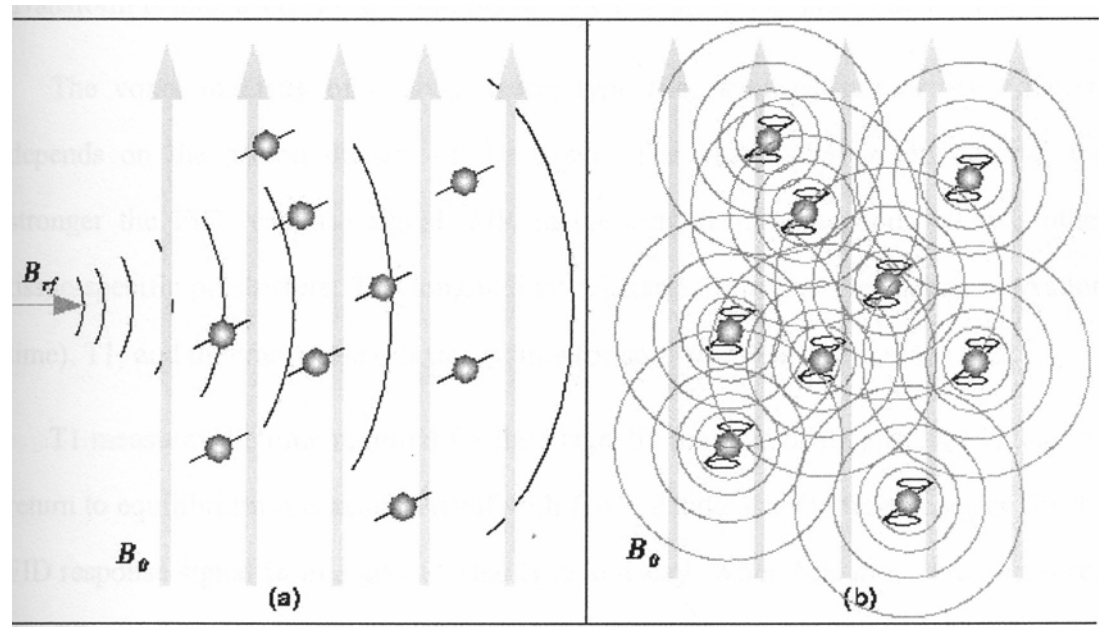
where  $\gamma$  is the gyromagnetic ratio and  $B_0$  is the strength of the applied magnetic field. The gyromagnetic ratio is a nuclei specific constant. For hydrogen,  $\gamma = 42.6 \text{ MHz/T}$ .

To obtain an MR image of an object, the object is placed in a uniform magnetic field  $B_0$  of between 0.5 to 1.5 Tesla. As a result, the object's hydrogen nuclei align with the magnetic field and create a net magnetic moment  $M$  parallel to  $B_0$ . This behaviour is illustrated in Figure 2.4. Next, a radio-frequency (RF) pulse,  $B_{rf}$ , is applied perpendicular to  $B_0$ . This pulse, with a frequency equal to the Larmor frequency, causes  $M$  to tilt away from  $B_0$  as shown in Figure 2.5(a).



**Figure 2.4** (a) In the absence of a strong magnetic field, hydrogen nuclei are randomly aligned; (b) When the strong magnetic field  $B_0$  is applied, the hydrogen nuclei precess about the direction of the field.

Once the RF signal is removed, the nuclei realign themselves such that their net magnetic moment  $M$  is again parallel with  $B_0$ . This return to equilibrium is referred to as relaxation. During relaxation, the nuclei lose energy by emitting their own RF signal, as shown in Figure 2.5(b). This signal is referred to as the free-induction decay (FID) response signal. The FID response signal is measured by a conductive field coil placed around the object being imaged. This measurement is processed or reconstructed to obtain 2-D grey-scale MR images.



**Figure 2.5** (a) The RF pulse,  $B_1$ , causes the net magnetic moment of the nuclei,  $M$ , to tilt away from  $B_0$ ; (b) When the RF pulse stops, the nuclei return to equilibrium such that  $M$  is again parallel to  $B_0$ . During realignment, the nuclei lose energy and a measurable RF signal.

2.3.2 Image Acquisition

To produce an actual MR image, the FID resonance signal must be encoded for each dimension. The encoding in the axial direction, the direction of  $B_0$ , is accomplished by adding a gradient magnetic field to  $B_0$ . This gradient causes the Larmor frequency to change linearly in the axial direction. Thus, an axial slice can be selected by choosing the frequency of  $B_1$  to correspond to the Larmor frequency of that slice. The 2-D

spatial reconstruction in each axial slice is accomplished using frequency and phase encoding. A “preparation” gradient  $G_y$  is applied causing the resonant frequencies of the nuclei to vary according to their position in the y-direction.  $G_y$  is then removed and another gradient,  $G_x$ , is applied perpendicular to  $G_y$ . As a result, the resonant frequencies of the nuclei vary in the x-direction due to  $G_x$  and have a phase variation in the y-direction due to the previously applied  $G_y$ . Thus, x-direction samples are encoded by frequency and y-direction samples are encoded by phase. A 2-D Fourier Transform is then used to transform the encoded image to the spatial domain [63].

The voxel intensity of a given tissue type (i.e. white matter or grey matter) depends on the proton density of the tissue: The higher the proton density, the stronger the FID response signal. MR image contrast also depends on two other tissue-specific parameters: The longitudinal relaxation time (or spin-lattice relaxation time), T1; and the transverse relaxation time (or spin-spin relaxation time), T2.

T1 measures the time required for the magnetic moment of the displaced nuclei to return to equilibrium (i.e. realign itself with  $B_0$ ). T2 indicates the time required for the FID response signal from a given tissue type to decay. When MR images are acquired, the RF pulse  $B_{rf}$  is repeated at a predetermined rate. The period of the RF pulse sequence is the repetition time TR. The FID response signals can be measured at various times within the TR interval. The time between which the RF pulse is applied and the response signal is measured is the echo delay time TE. By adjusting TR and TE, the acquired MR image can be made to contrast different tissue types.

The MR images used in clinical purpose were all acquired using a multiple echo spin echo pulse sequence in which two or more images are acquired simultaneously. TR and TE are adjusted such that tissues with different properties show different contrasts in images of different sequences. For an example, tissues with a high proton density appear bright in a proton density-weighted (PDW) image while tissues with a long T2 appear bright in a T2-weighted image. For an image of the brain, the T2 time

will provide contrast such that brightness of cerebral spinal fluid is greater than grey matter, which in turn is brighter than white matter, which is brighter than fat. On the contrary, the T1-weighted image would provide contrast just the reverse of T2-weighted images (fat > white matter > grey matter > cerebrospinal fluid). Moreover, image contrast can be further enhanced with the use of paramagnetic contrast materials like GdDTPA (gadolinium base) or newer superparamagnetic contrast agent pulse fat suppression technique [61]. It is particularly useful in the imaging of malignant tumours.

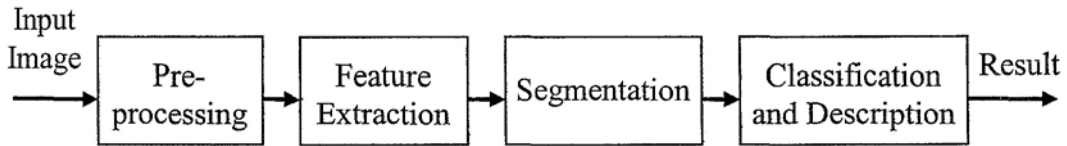
## 2.4 Methods of MRI Segmentation

The objective of image segmentation is to find the desirable object and isolate it from the rest of the scene [64]. If the domain of the image is given by  $I$ , then the segmentation problem is to determine the sets  $S_k \subset I$  whose union is the entire image  $I$ . Thus, the sets that make up the segmentation must satisfy

$$I = \bigcup_{k=1}^K S_k, \quad (2-5)$$

where  $S_k \cap S_j = \emptyset$  for  $k \neq j$ , and each  $S_k$  is connected. A representative diagram of the most common parts of a computer image analysis system is shown in Figure 2.6 [65]. Therefore, MRI segmentation is to find the desirable object from MFU by analyzing extracted image features, and hence define its boundary for further classification, description, and interpretation. Due to recent rapid advances in MRI system, image processing, pattern recognition and high performance computing, MRI segmentation has been proposed for a number of clinical investigations of varying complexity. MRI segmentation is becoming an increasingly important image processing step for a number of areas that include identifying anatomical areas of

interest for diagnosis, treatment or surgery planning paradigms, pre-processing for visualisation and multi-modality image registration, and improved correction of anatomical areas of interest with localised functional metrics [39].



**Figure 2.6** Components of an image analysis system

According to the number of MRI sequences used in segmentation, MRI segmentation can be broadly divided into two categories: a single image, or greyscale, segmentation where a single 2-D or 3-D MR image is used, and multi-spectral image segmentation where MR images of multiple sequences with different greyscale contrasts are available. The two categories are discussed in the following sections.

### 2.4.1 Greyscale Single Image Segmentation

Single greyscale image segmentation methods can be subdivided as thresholding-based methods, region-based methods, edge-based methods, morphological model methods, random field methods, etc.

- **Thresholding-based methods:** Global or local thresholding derived from image histogram is the most intuitive approach to segmentation. However, this method is of limited use despite the developments in knowledge-guided and adaptive thresholding methods. This method is available in segmentation software which was reported as not released for general use (Morph, Silicon Graphics, California, USA) [45].
- **Edge-based methods:** Edge-based methods try to find the places of rapid transition from one to the other region of different brightness or intensity. The main

principle is to apply some of the gradient operators and convolve them with the image. However, edge detection schemes often suffer from incorrect detection of edges due to noise, artefact, and variability in the threshold selection in the edge image. In recent years some investigators described a boundary tracing method using both region and contour deformation models [66]. In general, edge-based methods are likely to be restricted to segmentation of large, well defined structures such as brain tumour and parenchyma, but not to distinguish individual tissue types.

- **Region-based methods:** Region-based methods are complementary to the edge-based method. Here the point is to group neighbouring pixels to the region according to the given criteria of homogeneity. Seed-growing is a type of common but commercially available region-based method [39]. Results obtained with seed-growing are generally dependent on the operator settings and only well defined regions can be robustly identified. A “seeded region growing” method and the “improved seeded region growing” algorithm with the characteristics of rapidness, robustness, and without tuning parameters were reported by Adams, et al. in 1994 [67] and Mehnert, et al. in 1997 [68], respectively. Hojjatoleslami, et al. used a pixel aggregation-based region growing algorithm to segment large brain lesions and good results were obtained [69]. Furthermore, a similar method using the fuzzy-connectedness principles has also been developed by Udupa, et al [70]. These algorithms show that region-based methods still have potentials for MRI segmentation.
- **Morphological model methods:** Morphological models are often used with the combination of other methods. A common morphological model method is the template matching which is used for the classification of brain structures such as the white matter and ventricle [71]. Kaus, et al. used an adaptive template-moderated scheme to segment brain tumour from MRI automatically and

hierarchically [60]. Combining image segmentation based on statistical classification with *a priori* knowledge has been shown to significantly increase the robustness and reproducibility. This is known as the “knowledge-based method”, which utilises a probabilistic brain atlas [72] [73].

- Random field methods: Markov random field (MRF) model is a statistical model which can be used within segmentation methods. MRFs are often incorporated into clustering segmentation algorithms such as the *k*-means algorithm under a Bayesian prior model. The segmentation is then obtained by maximizing the *a posteriori* probability of the segmentation given the image data using iterative methods such as iterated conditional modes or simulated annealing [74]-[76]. A difficulty associated with the MRF models is the proper selection of the parameters controlling the strength of spatial interactions. In addition, MRF methods usually require computationally intensive algorithms. Despite these disadvantages, MRF methods are widely used not only to model segmentation classes, but also to model intensity heterogeneity that can occur in MR images

### 2.4.2 Multi-spectral Image Segmentation

Multi-spectral image segmentation uses two or more than two imaging sequences which contain more information than greyscale image segmentation, that is, vector-based features take the place of single intensity-based features. The most common approach for multi-spectral segmentation is pattern recognition. Multi-spectral segmentation can be divided into supervised segmentation and unsupervised segmentation.

- Supervised segmentation methods: Classifiers and algebraic methods are two kinds of supervised multi-spectral segmentation methods. For classifiers methods,

there are three most frequently used algorithms: the maximum likelihood (ML) or Bayes classifier,  $k$  nearest neighbours ( $k$ NN), and artificial neural networks (ANN). ML is a commonly used parametric classifier which assumes that the pixel intensities are independent samples from a mixture of probability distributions, usually Gaussian. By estimating the means and covariance matrices for each of the tissues from a user supplied training set, the remaining pixels are then classified by calculating the likelihood of each tissue class, picking the tissue type with the highest probability. On the other hand  $k$ NN, a non-parametric classifier, does not rely on predefined distributions, but on the actual distribution of the training samples themselves. Therefore,  $k$ NN has given superior results both in terms of accuracy and reproducibility compared to parametric classifier [77]. ANN represents a paradigm for machine learning and is widely applied in medical imaging as a classifier. First the weights of the neural networks are determined using training data, and the ANN is then used to segment the new data. Due to the multitude of interconnections used in a neural network, spatial information can easily be incorporated into its classification procedures [77]. For images with clearly identified signature vectors, algebraic methods which work with projections of feature vectors may provide an elegant solution to deal with the partial volume effect. However, these methods are optimal only for signature vectors that are more or less orthogonal, which may not be the case for pathological tissues that exhibit similar relaxation behaviour [39]. Algebraic methods are not common in multi-spectral segmentation and only a few schemes using vector decomposition have been reported [78] [79].

- Unsupervised segmentation methods: Unsupervised segmentation, also called “clustering”, automatically find the structure in the data. Three commonly used clustering algorithms are the  $k$ -means, the fuzzy  $c$ -means (FCM), and the expectation-maximisation (EM) [80]. The  $k$ -means algorithm 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 [80]. The FCM is a fuzzy generalisation of the commonly used  $k$ -means algorithm for unsupervised pattern recognition, which allows labels to be “fuzzy”. That is, a pixel can be partly in one class and partly in others. To overcome the limitations of FCM, some improved FCM algorithms such as semi-supervised fuzzy  $c$ -means (SFCM) and adaptive fuzzy  $c$ -means (AFCM) were proposed [81][82]. The EM algorithm applies the same clustering principles with the underlying assumption that the data follow a Gaussian mixture model. It iterates between computing the posterior probabilities and computing the maximum likelihood estimates of the means, covariances, and mixing coefficients of the mixture model, [80].

### 2.4.3 Areas of Improvement in MRI Segmentation

Each of the MRI segmentation methods described above has its own advantages, drawbacks and the objective for processing. The development goal is to increase the absolute accuracy and reproducibility as well as to decrease the observer variances and computational complexity. The potential research areas for improvement in MRI segmentation include: (1) the optimal selection of features in an image to maximise tissue contrast differentiation or segmentation in feature space while minimizing the computational complexity; (2) the optimisation of the level of supervision to best utilise the prior knowledge of the operator while increasing inter-operator reproducibility; (3) the development of atlas-guided approaches which make use of the standardised *a priori* knowledge about the image.

## 2.5 Fundamentals of 3-D Visualisation in Medicine

The complicated computing and computer graphics gave birth to the Visualisation in Scientific Computing which assists the investigators to better understand mass data from all kinds of perception and simulations. Visualisation in scientific computing is widely used in the medical areas and the modern imaging techniques provide necessary supports for the development of visualisation in scientific computing [83].

Generally, medical imaging modalities such as CT, MRI, PET and ultrasound produce a serial of 2-D tomographic images. The stack of these parallel 2-D tomographic images can describe the 3-D information of human body and is termed “volume data”. Nowadays, the high resolution of medical images in the x, y, and z axes leads to the mass volume data and a lot of physiological and pathological information is hidden inside. It is quite difficult for doctors to determine the pathological and spatial properties of lesions accurately by only observing a group of these 2-D images and “reconstructing the 3-D constructs in mind”. 3-D visualisation is just the technique that can present the volume data intuitively and realistically to assist in diagnosis and therapy.

The main task of 3-D visualisation in medicine is to acquire the rich information hidden in medical volume data by interactive image and graphic techniques. Mainly, there are a few primary research areas for 3-D visualisation in medicine that include volume graphics, volume rendering equation, transform coding of volume data, scattered data, enrichment of volumes with knowledge, segmentation and classification, real-time rendering and parallelism, and special-purpose hardware [84]. The discussions on 3-D visualisation algorithms and 3-D visualisation in head and neck imaging are given in Sections 2.5.1 and 2.5.2.

2.5.1 3-D Visualisation Process and Algorithms

The elementary process of 3-D visualisation is shown in Figure 2.7. The volume data are obtained from imaging devices in the first step. In the second step, for images with large amount of raw data, data refinement and selection are needed to reduce the amount of data while at the same time, prevent the loss of useful data. When the data attribution is sparse and thus insufficient for visualisation, data interpolation is needed. The core of the whole procedures is the visualisation mapping that transfers processed raw data into plotted geometric elements and attributes. In the fourth step, the geometric elements and attributes produced in the previous step will be transferred into displayed images using these fundamental techniques of computer graphics such as viewing transform, light calculation and scanning transform. In the implementation of 3-D visualisation, necessary steps include image acquisition, image merging, image segmentation, colour assignment, 3-D image rendering, and image display.

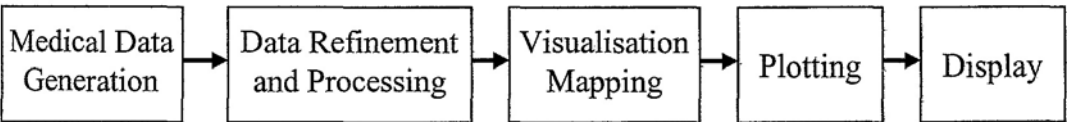


Figure 2.7 Elementary process of 3D medical visualisation

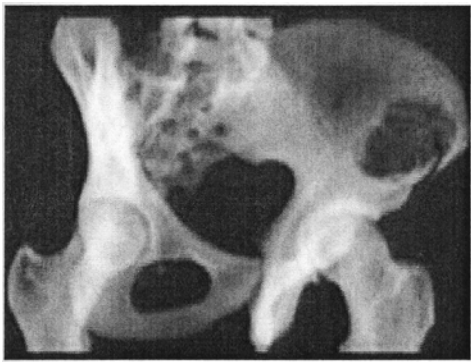
Most of the 3-D visualisation algorithms can be divided into surface rendering algorithms and volume rendering algorithms.

- Surface rendering: For surface rendering, intermediate geometric elements such as camber and plane are first constructed from 3-D spatial volume data. Frame rendering is then performed by the traditional computer graphics techniques. The most frequently used intermediate geometric element is the triangular patch, the same situation where a contour is extracted from the 3-D spatial volume data. The generation of the intermediate geometric elements can be regarded as a mapping that maps a part of the properties of original data into planes or cambers.

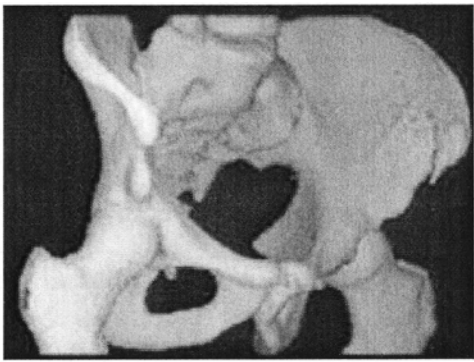
Therefore, the visualised graph generated by this method cannot reflect the full view and the details of the whole original volume data, but more distinct contour image can be generated [85]. In addition, the algorithm can use current graphic hardware that accelerates the speed of rendering and transformation. Typical algorithms of surface rendering include contour connecting, opaque cube, and matching cubes.

- Volume rendering: Volume rendering is different from surface rendering. Instead of constructing intermediate geometric elements, volume rendering generates 2-D images on the screen directly from 3-D volume data, skipping the step of mapping. The selected volume data are displayed on screen via the projections of corresponding voxels. During the projection, the transfer function which represents the different properties of each voxel such as greyscale, gradient, and coordinates, is introduced to compute the shading effect. Volume rendering can produce the entire image of 3-D data set, including the object surface and details inside with high quality. Its drawback is the large amount of computation and it is not supported by the traditional graphic hardware. To solve this problem, some special volume rendering-accelerated graphic hardware systems, such as VolumePro [86], were developed. The typical algorithms of volume rendering include ray casting, splatting, and shear warp factorisation.

The 3-D graphs of pelvis formed by volume rendering and surface rendering are shown in Figures 2.8 and 2.9, respectively [87]. The graph formed by the former actually provides the viewer with an inside view of the rendered model. A detailed comparison of these two algorithms is shown in Table 2.2.



**Figure 2.8** Shaded volume rendering of pelvis



**Figure 2.9** Shaded surface rendering of pelvis

**Table 2.2** Comparison of volume rendering and surface rendering

Volume rendering	Surface rendering
Uses entire data set	Only uses first defined voxel as the surface of the bone
Conveys more information than surface rendering	Displays gross 3-D relationships
Contributions of all voxels in the volume data are summed	Fails to display lesions hidden beneath the bone surface
A range of opacity from transparent to opaque is selected	Tends to demonstrate stair-step artefacts
The view angles from any plane or projection are generated	

2.5.2 3-D Visualisation in Head and Neck Imaging

Technical improvements in CT and MR imaging permit exquisite visualisation of bony and soft tissue anatomy, as well as various extra-cranial head and neck lesions. Recent advances in CT and MRI scanners have produced high-quality data suitable for 3-D visualisation. In addition, recent developments in computer graphics have enabled direct viewing of CT and MR imaging studies as 3-D volumes. These

advances provide surgeons or radiologists with reconstructed images and perspectives, which are not available with conventional scans, viewed from the surgical or radiotherapy treatment positions [88] .

In the early 1990s, only non-real time surface rendering was used in clinical visualisation application for head and neck imaging. 3-D view did give better understanding and a direct impression of the topographic relationship of lesion within normal tissue. However, it offered no significant advantage over 2-D and almost no diagnosis was substantially changed by the addition of 3-D over 2-D for temporal bone studies [89]. Darling, et al. found that the primary advantage of real time surface rendering for head and neck imaging was the additional information of depth perception, contours, volumes, and extent of abnormalities [90].

The changes from surface rendering to volume rendering under real time control further improve doctor's understanding of the complex anatomy of this region. In 1995, Johnson, et al. performed a pilot study on the use of volume visualisation in image-based treatment planning for head and neck cancer [91]. Lee, et al. studied the utility of volume rendering technique in the practice of CT-based radiotherapy planning for head and neck [92]. They all found that volume visualisation might be useful in target definition of head and neck cancers since the spatial relationship of critical normal structure to the gross target tumour and nodal areas could be visualised more clearly. Moreover, volume rendering had several advantages over previously segmentation-based 3-D display techniques [91][92]. Moharir, et al. developed a computer assisted 3-D reconstruction technique of head and neck tumour, which provided a very good 3-D interactive representation of the tumours and patient anatomy for potential applications in treatment, research, and medical education [20]. Greess, et al. demonstrated the value of 3-D visualisation from spiral CT on various studies of head and neck tumours [93]. The study results of Cavalcanti, et al. showed that by using commercial 3-D image workstation, 3-D CT reconstructed images of

head and neck tumours were of the greatest benefit to the clinicians in pre and postoperative assessment [94].

The major limitation of the studies cited here is that only CT data were used for visualisation and very little was done using MRI. This may be partly due to the fact that CT images are easier for structure and tissue classification and colour designation according to the following simple criteria: High CT level means high density corresponding to bone; medium CT level means medium density corresponding to soft tissue or fluid; and low CT level means low density corresponding to air or background. The various CT windows make it easy to demonstrate the serial removal of soft tissue structures from the same data on a fixed perspective [92]. MRI is better than CT in differentiating tumour margin from normal adjacent soft tissue, but it is also the high resolution for soft tissue of MRI that makes the classification of structures and tissues difficult for visualisation. Classification using multi-threshold in image intensity or lookup table is not applicable and fast tissue classification algorithm is needed.

## 2.6 Concluding Remarks

It is noted that imaging-based (especially image segmentation-based) tumour volumetric analysis offers very high accuracy in determining the volume of solid tumour for tumour staging, treatment planning and treatment evaluations. Currently, the quantitative method for head and neck tumour-volume analysis using MRI is still not well established. MRI segmentation plays the critical role in the whole processing procedure. Each of the MRI segmentation methods has its own advantages, drawbacks and objectives for processing. The potential directions for improvement include the selection of optimal features, the use of hierarchical segmentation methods and supervision for complex images, and the introduction of the standardised *a priori*

---

knowledge. It is also noted that 3-D visualisation, especially using volume rendering technique, has a lot of benefits to head and neck imaging for lesion diagnosis and treatment planning. Most of the work in 3-D visualisation was performed on CT data, it is therefore important to investigate 3-D head and neck visualisation using MRI data.

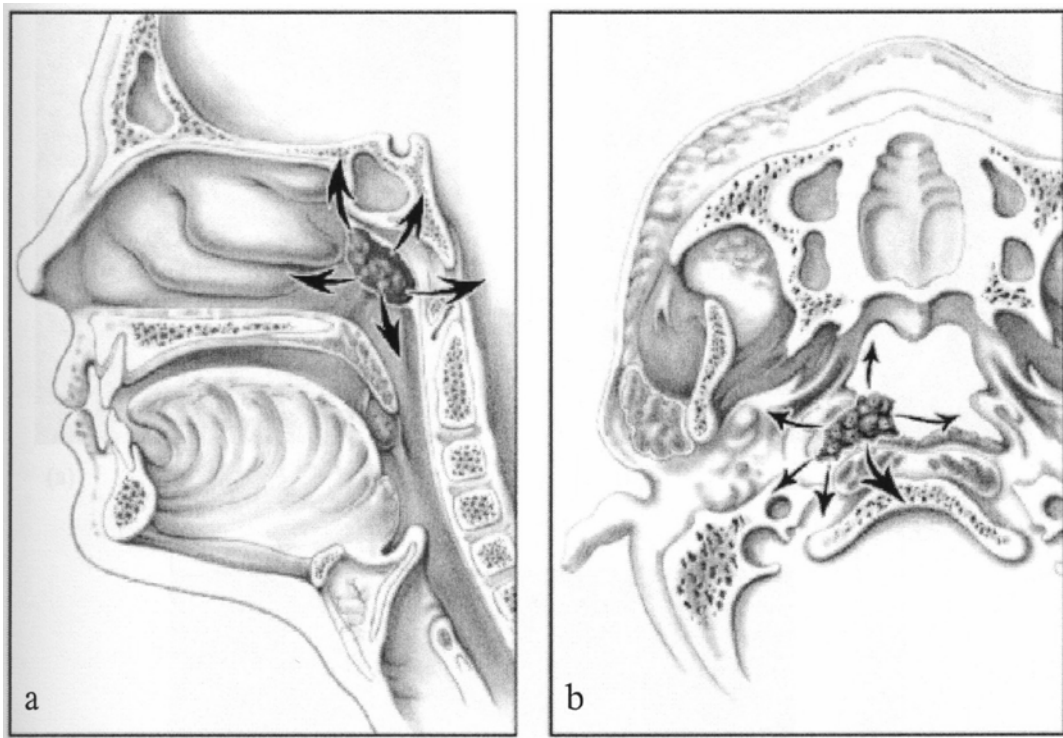


## **Chapter 3**

# **Semi-supervised MRI Segmentation and Tumour Volume Determination for Nasopharyngeal Carcinoma**

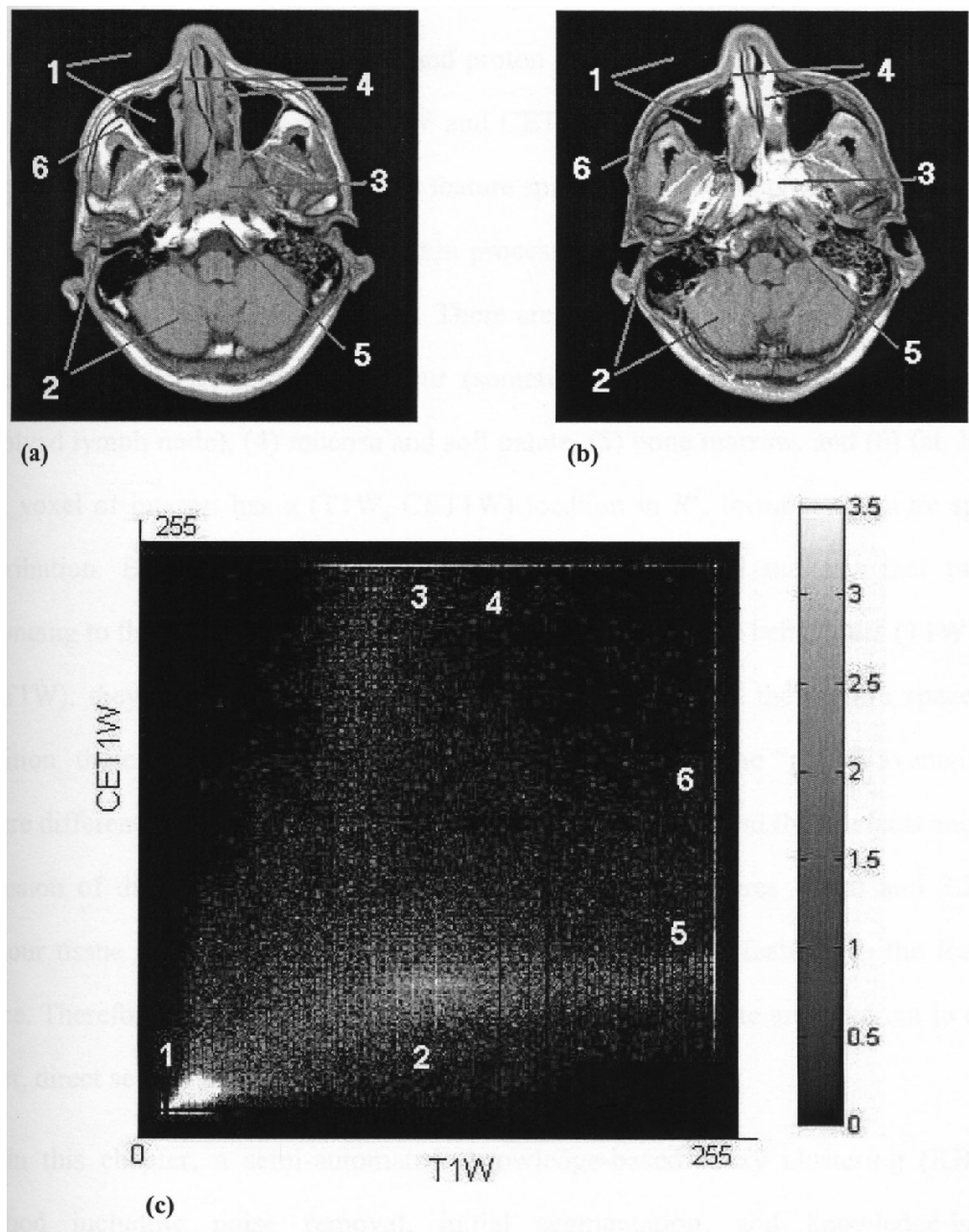
### **3.1 Introduction**

Nasopharyngeal carcinoma (NPC) is a malignant skull base tumour. High frequencies of occurrence are reported in Southeast Asia and South China. The risk does not diminish in Chinese migrants outside this region [11]. NPC shows an aggressive and infiltrative growth pattern. As shown in Figure 3.1, NPC has a high degree of region spread at presentation with the propensity to extend into the nasal cavity, oral pharynx, cervical spine, parapharyngeal space, skull base, and intra-cranial spaces. Radiation therapy is the mainstay of treatment and accurate tumour staging is crucial for effective treatment planning. According to the current Tumour Node Metastasis (TNM) classification of UICC/AJCC, NPC is not staged by volume determination but by anatomical extension. However, the determination of tumour volume could assist tumour staging and effective treatment planning.



**Figure 3.1** (a) Local spread of NPC in sagittal view; (b) Local spread of NPC in axial view [95].

There is a lack of quantitative methods for NPC tumour volume analysis due to the difficulties in obtaining accurate tumour segmentation from medical images such as CT and MRI. Literature review shows that there is also no specially developed algorithm for NPC tumour segmentation due to the infiltrative growth pattern and irregular tumour shape, the complex anatomic structures nearby, the heterogeneous MRI signals or blurry tumour contour due to partial volume effect, and the overlap of image feature distribution. Hsu, et al. developed an automatic medical diagnosis for segmenting NPC with dynamic gadolinium-enhanced MR imaging and pharmacokinetic analysis [96]. This method can identify the NPC regions effectively. However, it is indeed a type of “detection” rather than “segmentation” and only one section of the head can be examined.



**Figure 3.2** (a) T1W image; (b) CET1W image; (c) Average joint histogram of 12 pairs of tumour-represented MRI slices. The six primary tissue types shown here are: 1. Air; 2. Other normal soft tissue; 3. Tumour; 4. Mucosa and soft palate; 5. Bone marrow; 6. Fat.

Figures 3.2(a) and 3.2(b) show the fast spin echo sequenced T1-weighted (T1W) and contrast enhanced T1-weighted (CET1W, gadolinium-enhanced with fat suppression) images (512x512 pixels) of a pair of typical NPC slices respectively. Compared to brain images which have been extensively studied, images of head and neck region are more complex in both anatomy and tissue components. Although in

most of the previous brain tumour analyses, T1W (with or without contrast enhancement), T2-weighted (T2W), and proton density (PD) images were needed for multi-spectral segmentation, the T1W and CET1W sequences currently available in the present study can also provide the feature space for tumour-represented areas. The average joint histogram (after logarithm processing) of 12 pairs of NPC-represented MRI slices is shown in Figure 3.2(c). There are six primary tissue types: (1) Air, (2) other normal soft tissues, (3) tumour (sometimes includes oedema, necrosis and involved lymph node), (4) mucosa and soft palate, (5) bone marrow, and (6) fat. Each MR voxel of interest has a (T1W, CET1W) location in  $R^2$ , forming a feature space distribution. Based on the knowledge in Figure 3.2(c) and the fact that pixels belonging to the same tissue type will exhibit similar relaxation behaviours (T1W and CET1W), they will have approximately the same location in the feature space. In addition, there is some overlap between classes because of the “partial-averaging”, where different tissue types are quantised into the same voxel and the artefacts and the adhesion of different tissues are taken into account. In Figures 3.2(a) and 3.2(b), tumour tissue and mucosa show very similar signal intensity features in the feature space. Therefore, in cases where tumour, mucosa and soft palate are adherent to each other, direct segmentation becomes difficult.

In this chapter, a semi-automated, knowledge-based fuzzy clustering (KBFC) method including noise removal, initial segmentation, and knowledge-based refinement has been developed to segment NPC from MRI for tumour volume determination. The organisation of this chapter is as follows. Section 3.2 describes the noise removal and initial segmentation process using semi-supervised fuzzy c-means (SFCM), which are the first two steps of the KBFC method. Section 3.3 presents the knowledge-based image analysis for segmentation refinement, which is the third step of KBFC. In addition, some segmentation results are presented in this section. The experimental method of NPC tumour-volume determination using KBFC and its

validation are given in Section 3.4. Sections 3.5 and 3.6 are the discussions and concluding remarks, respectively.

## 3.2 Noise Removal and Initial Segmentation

### 3.2.1 Noise Removal

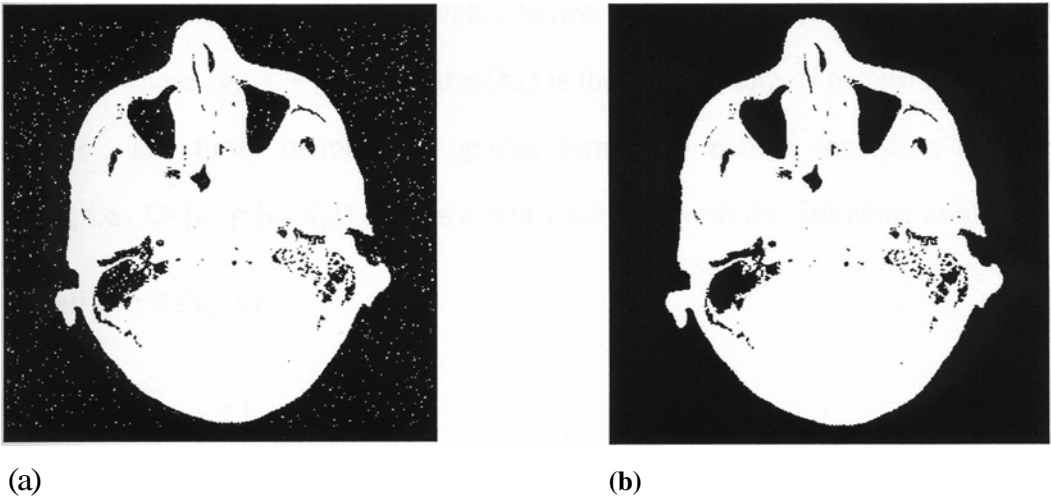
Median filter can eliminate the background grain noise in the original MR images. However, it also decreases the edge details. Hence in the present study, a head mask is used to remove the background noise but maintain the edge details of the original signal. First an initial binary head mask  $\mathbf{M}$  is generated from the original image  $\mathbf{I}$  using the minimising within-group variance method for the determination of the best threshold value [97]. As defined in Eq. (3-1), a binary morphological opening operation (“ $\circ$ ”, erosion followed by dilation) using a circular structuring element  $\mathbf{B}$  with the radius of 2, is then performed on  $\mathbf{M}$  to remove the background noise and the resulting new mask was named as  $\mathbf{M}'$ :

$$\mathbf{M}' = \mathbf{M} \circ \mathbf{B} = (\mathbf{M} \ominus \mathbf{B}) \oplus \mathbf{B} , \quad (3-1)$$

where “ $\ominus$ ” is the erosion operator of mathematical morphology and “ $\oplus$ ” is the dilation operator [97]. Figure 3.3 shows the head mask for the image in Figure 3.2. A background noise-free new image  $\mathbf{I}'$  is obtained using

$$\mathbf{I}' = \mathbf{M}' \bullet \mathbf{I} , \quad (3-2)$$

where “ $\bullet$ ” means the multiplication of corresponding pixels in the two matrices. In  $\mathbf{I}'$ , the intensities of almost all the background and air-content areas are set to zero. In the following procedures, the entire image is selected as the region of interest (ROI) for further processing.



**Figure 3.3** The head mask (a) Before the morphological operation; (b) After the operation.

### 3.2.2 Initial Segmentation Using SFCM

#### 1. Clustering using fuzzy $c$ means

SFCM is a hybrid, modified version of the fully unsupervised fuzzy  $c$  means (FCM) clustering method, while FCM is the fuzzy extension of  $k$ -means method, one of the classical clustering techniques [59]. FCM is also a clustering algorithm, but the resulting partition is fuzzy that the input feature vectors are not assigned exclusively to a single class, but partially to all classes. For example, a data point could be categorised into three classes with the fuzzy label  $\mathbf{y} = (0.1, 0.2, 0.7)^T$ , indicating that the data point has a probability of 0.1 belonging to the first class, 0.2 to the second class and 0.7 to the third class. If a single class must be chosen, this data point can be chosen to be in the class with the highest membership grade, the third class in this example. This process is called defuzzification which yields a crisp label.

Formally, FCM takes a finite object data set  $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \dots, \mathbf{X}_n)$  as an input. Here, each  $\mathbf{X}_k$  is a feature vector and  $\mathbf{X}_k = (x_{k1}, x_{k2}, \dots, x_{ks})$ , where  $x_{kj}$  is the  $j^{th}$  feature of the  $k^{th}$  subset,  $1 \leq k \leq n$ , and  $s$  is the dimensionality of  $\mathbf{X}_k$ . Each  $\mathbf{X}$  vector is labelled



with a fuzzy membership having a value between 0 and 1.  $\mathbf{X}$  is partitioned into  $c$  subsets of  $u$  where  $\{u_i: \mathbf{X} \rightarrow [0,1]\}$ ; and  $u_i(\mathbf{X}_k)$  is the membership of pattern  $\mathbf{X}_k$  in class  $i$ ,  $1 \leq i \leq c$ . The fuzzy membership grades form a  $\mathbf{U}$  matrix consisting of  $c \times n$  elements, i.e.,  $\mathbf{U}=[u_{ik}]=[u_i(\mathbf{X}_k)]$ ,  $1 \leq i \leq c$  and  $1 \leq k \leq n$ , with the following constraints:

(1) For all  $i, k$ :  $0 \leq u_{ik} \leq 1$ ,

(2) For all  $k$ :  $\sum_{i=1}^c u_{ik} = 1$ . (3-4)

FCM consists of an iterative optimisation of an objective function  $J_m$  [81] :

$$J_m(\mathbf{U}, \mathbf{V} : \mathbf{X}) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik})^m \|\mathbf{X}_k - \mathbf{V}_i\|^2, \quad (3-5)$$

where  $\mathbf{V}_i=\{\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_c\}$  are the cluster centres that are being sought,  $m \in [1, \infty)$  is a weighting exponent of each fuzzy membership, and  $\|\mathbf{X}_k - \mathbf{V}_i\| = \sqrt{(\mathbf{V}_k - \mathbf{V}_i)^T (\mathbf{X}_k - \mathbf{V}_i)}$ .

If  $m=1$  and

$$u_{ik} = \begin{cases} 1, & \|\mathbf{X}_k - \mathbf{V}_i\| < \|\mathbf{X}_k - \mathbf{V}_j\|, \quad j=1, \dots, c, j \neq i, \\ 0, & \text{otherwise,} \end{cases} \quad \text{for all } i \text{ and } k, \quad (3-6)$$

$$\mathbf{V}_i = \sum_{k=1}^n u_{ik} \mathbf{X}_k / \sum_{k=1}^n u_{ik} \quad \text{for all } i,$$

it is just the  $k$ -means problem, which is also referred to as hard  $c$ -means (HCM). For computational reasons,  $m=2$  is chosen in the present method. Therefore in FCM, the objective function  $J_2$ , as defined by

$$J_2(\mathbf{U}, \mathbf{V} : \mathbf{X}) = \sum_{i=1}^c \sum_{k=1}^n (u_{ik})^2 \|\mathbf{X}_k - \mathbf{V}_i\|^2, \quad (3-8)$$

is used to approximate the minimum of a sum-of-weighted-distance function. In other words, this objective function leads to the minimal square errors of the estimated membership matrix.

## 2. SFCM: An improvement for FCM

From a viewpoint of reproducibility, unsupervised FCM is clearly desirable. However, FCM does not necessarily arrive at meaningful segmentations and often requires long computation time [39][81]. In particular, FCM has the tendency to prefer equal size clusters which is partly due to the randomly generated initial membership matrix  $\mathbf{U}_0$  for the optimisation of  $J_2$ . To overcome the limitations of FCM and at the same time, to avoid the tedious, costly and impractical task of having a human assign a tissue class to each of the object data for initialisation, SFCM has been developed. In SFCM, a training set from each class is selected by the operator to initialise and guide the segmentation algorithm. This hybrid approach helps to minimise or eliminate the errors introduced in the FCM clustering method by selecting and validating the number of clusters in the image, correct labelling of the clusters that are anatomically relevant, and helping to overcome the tendency to prefer equal size clusters [51]. The following procedure is used to minimise  $J_2$  in SFCM [51].

The initial  $\mathbf{U}_0$  matrix is formed with  $n_l$  columns ( $n_l$  is the number of elements in the training set,  $l$  means “labelled”) of labelled pixel vectors having crisp membership grade of 1 or 0, and the remaining  $n_u$  ( $n_u = n - n_l$ ,  $u$  means “unlabelled”) elements from the unlabeled pixel vectors of the object data to be classified. The initial cluster centres  $\mathbf{V}_i$  are computed from the first  $n_l$  column of  $\mathbf{U}_0$  using the equation:

$$\mathbf{V}_i = \sum_{k=1}^{n_l} (u_{ik}^l)^2 \mathbf{X}_k^l / \sum_{k=1}^{n_l} (u_{ik}^l)^2, \quad (3-9)$$

where the superscript  $l$  indicates the labelled training pixels. Then the remaining  $n_u$  unlabelled columns of  $\mathbf{U}$  are updated by calculating  $u_{ik}^u$  by the following equation:

$$u_{ik}^u = \left[ \sum_{j=1}^c \frac{\|\mathbf{X}_k^u - \mathbf{V}_i\|^2}{\|\mathbf{X}_k^u - \mathbf{V}_j\|^2} \right]^{-1}, \quad (3-10)$$



where the superscript  $u$  indicates the remaining unlabelled pixels. One variable weight vector  $\mathbf{P}_k=(p_1, p_2, \dots p_c)$  is introduced to make multiple copies of the training set in the  $\mathbf{U}$  matrix to increase or decrease the effect of training pixels on the new cluster centres. Then, the new cluster centres are found with all columns of the updated  $\mathbf{U}$  matrix using the following equation:

$$\mathbf{V}_i = \frac{\sum_{k=1}^{n_l} p_i (u_{ik}^l)^2 \mathbf{X}_k^l + \sum_{k=1}^{n_u} (u_{ik}^u)^2 \mathbf{X}_k^u}{\sum_{k=1}^{n_l} p_i (u_{ik}^l)^2 + \sum_{k=1}^{n_u} (u_{ik}^u)^2} . \quad (3-11)$$

Next, Eq. (3-10) is used to calculate the  $u_{ik}^u$  with the new cluster centres so that Eqs. (3-10) and (3-11) form an iterative process that converges to a local minimum of the objective function  $J_2$ . The iterative process is continued until the Euclidean distance between the new and old unlabelled columns of the  $\mathbf{U}$  matrices (i.e.,  $\|\mathbf{U}_{new}^u - \mathbf{U}_{old}^u\|$ ) is less than a threshold value  $\varepsilon$ . The final cluster centres  $\mathbf{V}_i$  corresponding to the final  $\mathbf{U}$  matrix are then computed.

### 3. The implementation of SFCM

The procedure for the implementation of the SFCM algorithm is as follows.

- 1) With a graphic user interface (GUI), select the initial training points of each tissue class which is labelled with different colours, as shown in Figure 3.4. For each class, four to ten points are selected. If no point is selected for a tissue class, this tissue class does not exist in this pair of images.
- 2) For every point  $(x, y)$  selected as class  $i$ , the memberships of class  $i$  in  $(x, y)$  and its 4-connective neighbourhood  $(x-1, y)$ ,  $(x+1, y)$ ,  $(x, y-1)$ , and  $(x, y+1)$  are set to 1. The memberships of other classes to the five pixels are set to 0. For example:

if  $(x, y)$  is selected as tumour tissue, then

tumour-membership $(x, y)$ =tumour-membership $(x \pm 1, y)$ =tumour-membership $(x, y \pm 1)$ =1;

other-membership $(x, y)$ =other-membership $(x \pm 1, y)$ =other-membership $(x, y \pm 1)$ =0;

end.

- 3) Compute the initial cluster centres  $V_i$  from the training points using Eq. (3-9).
- 4) Compute the memberships of every class for non-training data using Eq. (3-10).  
By this step, every pixel has a computed or assigned membership of every class and the matrix  $U$  is obtained.
- 5) Compute the new cluster centres  $V_i$  from all image pixels using Eq. (3-11). The weight vector  $P_k$  might be selected in proportion to cluster size, but here we use an equal weight vector for all classes that  $p_1=p_2=\dots=p_c$ . That means that the training set of every tissue class has the same contribution to the determination of the new cluster centres.
- 6) Repeat step 4) to get a new membership matrix  $U$  using the new cluster centres.
- 7) Compute Euclidean distance between the new and old non-training elements of the  $U$  matrices (i.e.,  $\|U_{new}^u - U_{old}^u\|$ ).

if  $\|U_{new}^u - U_{old}^u\| < \varepsilon$ , then

end;

else

repeat steps 5), 6) and 7);

end.

- 8) The pixels whose tumour memberships are more than the memberships of other tissue classes, are classified as tumour.

In the work of Vaidyanathan, et al., different values of  $P_k$  ( $P_k=1, 10, 20$  and  $100$ ) were tested and  $100$  was chosen as the proper  $P_k$  value [51]. In the present work,  $P_k$

value of 100 was adopted such that the medical experts selected training sets can make more contribution to the calculation of the desired cluster centres.  $\varepsilon$  is the threshold to determine the convergence of the membership matrix  $\mathbf{U}$  and in the present method,  $\varepsilon = 0.01$  was selected to ensure the good convergence of  $\mathbf{U}$ .

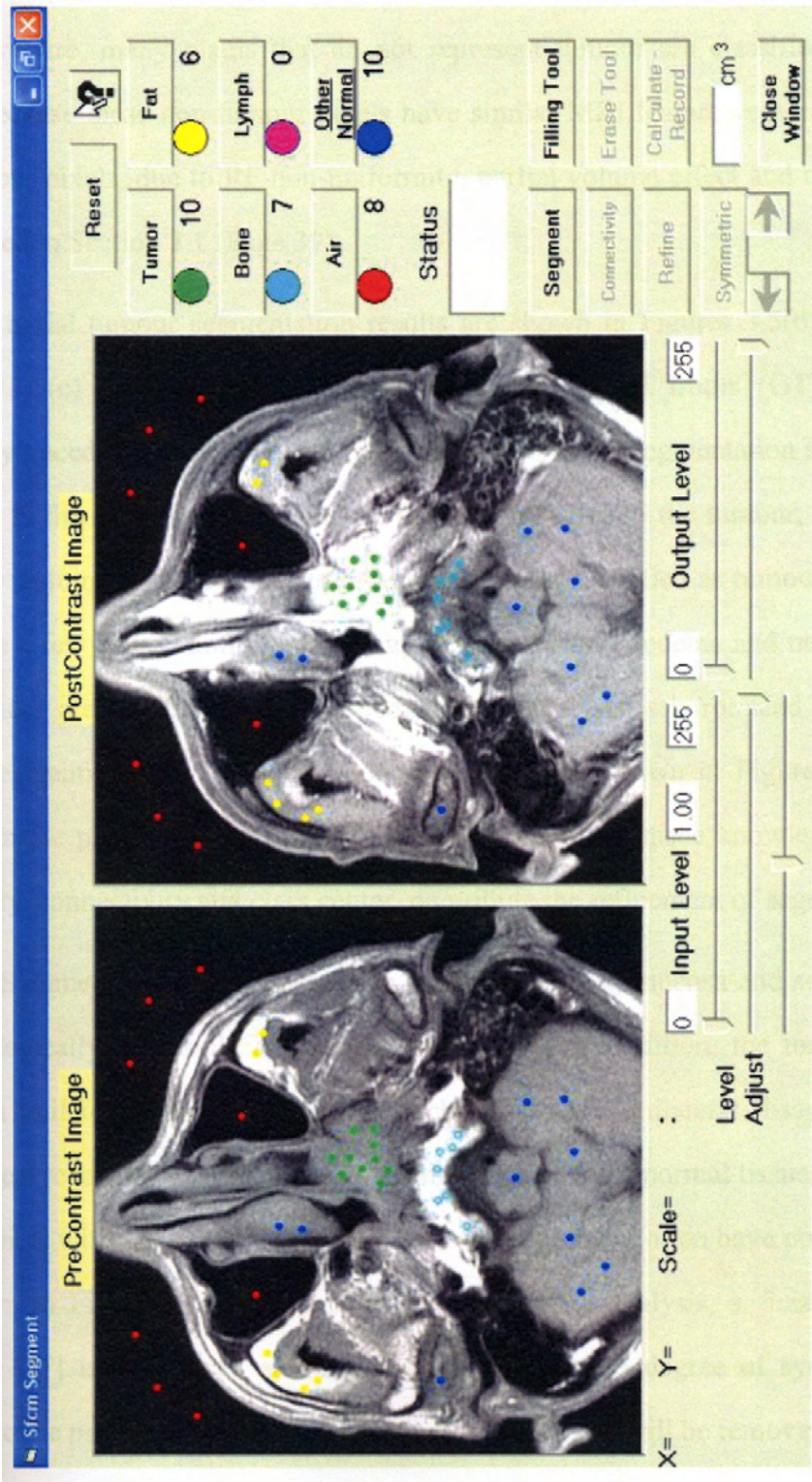


Figure 3.4 GUI for training data selection

### 3.3 Knowledge-based Image Analysis

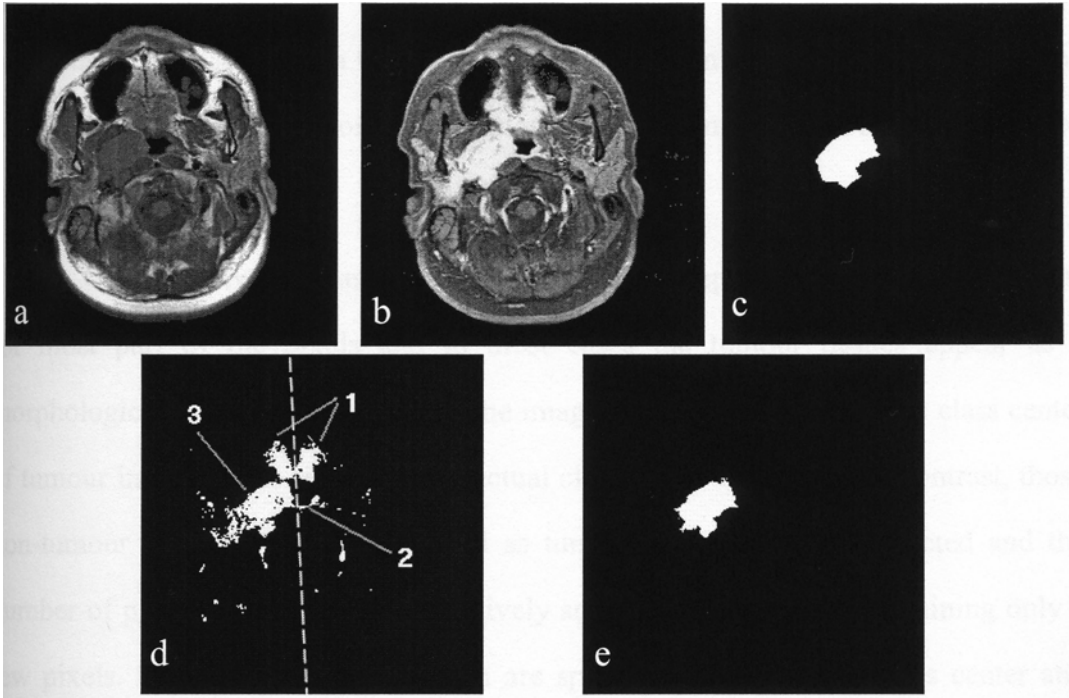
#### 3.3.1 Knowledge-based Image Analysis

The initial tumour segmentation using fuzzy clustering is a coarse procedure. After this procedure, many pixels that do not represent tumour are classified as tumour tissue because these non-tumour pixels have similar MRI feature vectors as those of the tumour pixels, due to RF non-uniformity, partial volume effect and other reasons mentioned in Section 3.1 (Page 39).

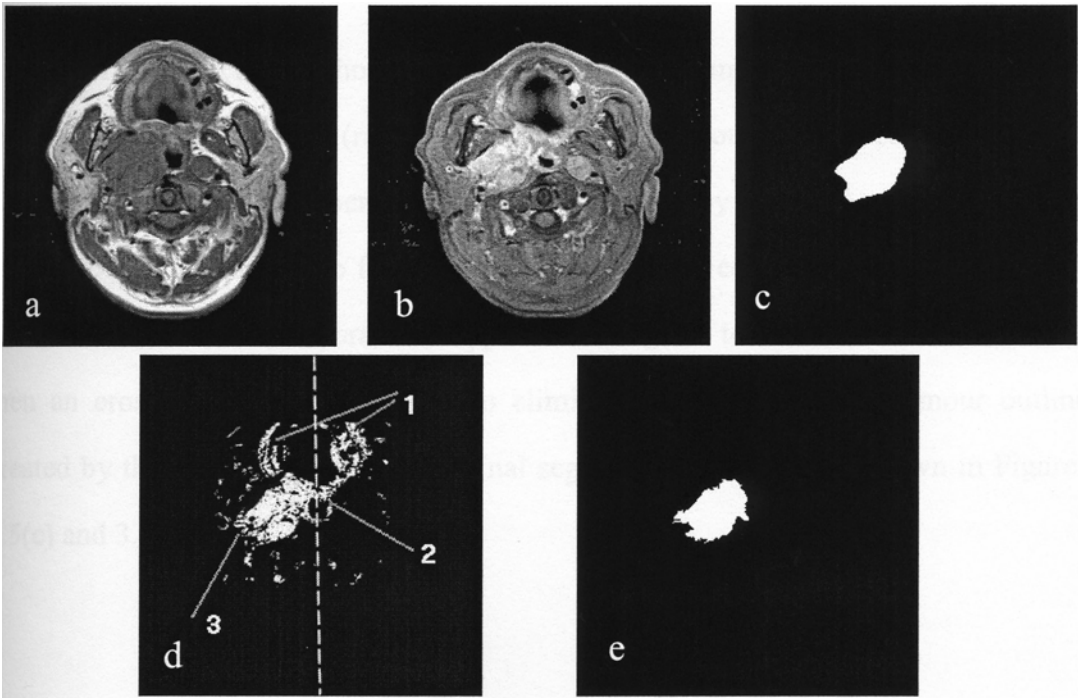
The initial tumour segmentation results are shown in Figures 3.5(d) and 3.6(d). Figures 3.5(c) and 3.6(c) are the corresponding ground truths (GT) which are manually traced by an experienced radiologist. The initial segmentation shows that in addition to the soft palate and mucosa that are adhesive to the tumour, a number of spatially disjoint small areas (scatter points) are also classified as tumour. Moreover, there are a few small “holes” which may represent small oedema and necrosis in the true tumour area. Therefore, a knowledge-based analysis is introduced to refine the initial segmentation and the corresponding results are shown in Figures 3.5(e) and 3.6(e). In the present study, three types of anatomic and space knowledge, namely, symmetry, connectivity and class center, contribute the refinement of segmentation.

(1). Symmetry: In normal condition the nasopharynx, mucosa and soft palate are morphologically bilaterally symmetric while in NPC condition, the tumour always shows a unilateral growing pattern, which causes the unilateral disappearance of pharyngeal recess [98]. Therefore in the image space, most normal tissues are roughly symmetrical with respect to the vertical axis while tumours often have poor symmetry, as shown in Figures 3.5(d) and 3.6(d). In the image analysis, a fuzzy symmetric measure [71] is introduced and those areas with certain degree of symmetry with respect to the previously determined vertical axis (Y axis) will be removed. Before the





**Figure 3.5** (a) T1W image; (b) CET1W image; (c) Ground truth; (d) Initial segmentation: 1. soft palate, 2. mucosa, 3. tumour, and the dash line is the symmetric axis; (e) Final result.



**Figure 3.6** (a) T1W image; (b) CET1W image; (c) Ground truth; (d) Initial segmentation: 1. soft palate, 2. mucosa, 3. tumour, and the dash line is the symmetric axis; (e) Final result.

initial segmentation, the center of the head is determined via the first moment of the binary head mask and then a Cartesian coordinate system is set up in the center of the head, according to the slant of head position. The resultant image of this step is named as image **I**.

(2). Connectivity and class center: After the first step, the tumour mass accounts for most part of the pixels and in most cases the tumour tissues appear as a morphologically continuous region in the image. Therefore the geometric class center of tumour in image **I** is similar to the actual class center of tumour. In contrast, those non-tumour areas which are classified as tumour are always disconnected and the number of pixels in those areas is relatively small with some areas containing only a few pixels. Therefore, the areas which are spatially disjoint to the class center and contain low pixel number are removed. For implementation, the first moment of image **I** is calculated as class center and an eight-wise connected components operation is applied from the class center. The resultant image of this step is named as image **II**.

(3). Mathematical morphology refinement: For the small “holes” which represent small oedema and necrosis (regarded as part of the tumour in most cases) inside the tumour area, a closing operator (dilation followed by erosion) of mathematical morphology is performed to fill these holes and get the complete presentation of the tumour. First, a dilation operator is applied to image **II** to fill these small holes and then an erosion operator is applied to eliminate the newly grown tumour outline created by the dilation operator. The final segmentation results are shown in Figures 3.5(e) and 3.6(e).

### 3.3.2 Segmentation Results

To visually evaluate the KBFC method, NPC images were segmented using both KBFC and maximum likelihood method (MLM, a commonly used supervised method), and the results were compared. A brief description of MLM is given below [77]:

MLM is based on Bayes decision rule which maximises *a posteriori* probability. This rule can be written as:

$$\mathbf{X} \text{ is class 2, if } P_1 p_1(\mathbf{X}) < P_2 p_2(\mathbf{X}), \quad (3-12)$$

where  $\mathbf{X}$  is a feature vector with its elements representing the intensities for a pixel in each of the MR images,  $P$  stands for *a priori* probability for a class, and the  $p$  is the conditional probability of  $\mathbf{X}$ , given that it is in that class. The conditional probabilities are given by the multivariate normal density function

$$p_i(\mathbf{X}) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2} d^2\right), \quad (3-13)$$

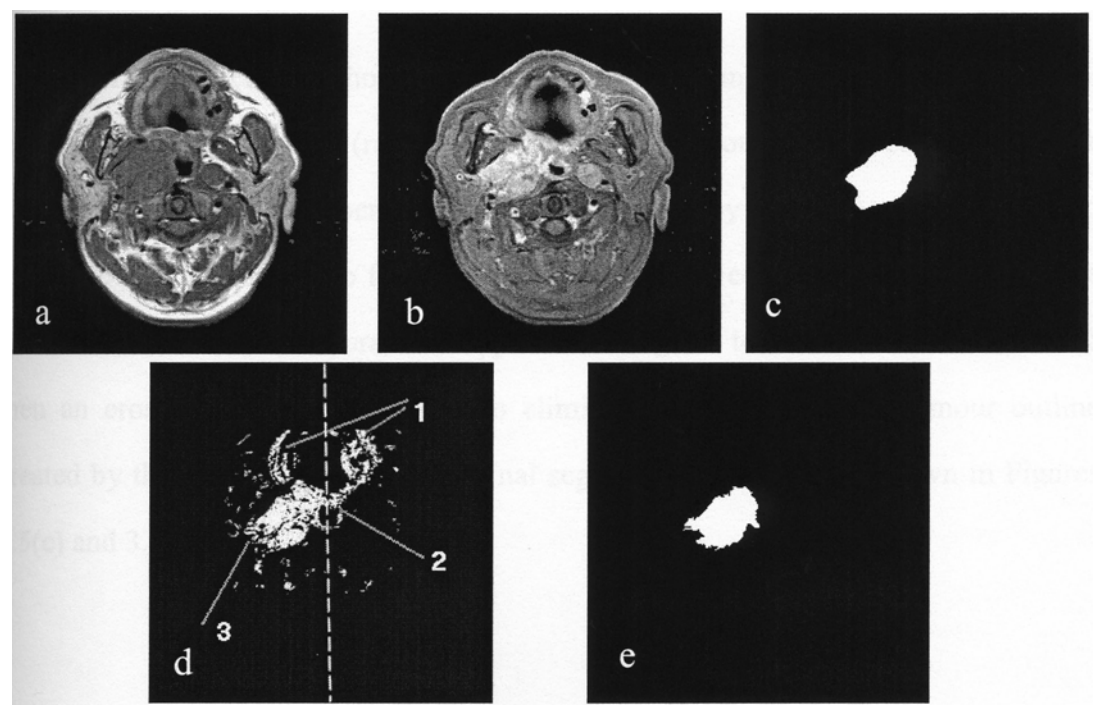
where  $n$  is the dimension of feature vector  $\mathbf{X}$  and  $d$  is a distance measure (the Mahalanobis distance) described by

$$d^2 = (\mathbf{X} - \mathbf{M})^T \Sigma^{-1} (\mathbf{X} - \mathbf{M}), \quad (3-14)$$

where  $\mathbf{M}$  is the mean vector and  $\Sigma$  is the covariance matrix. Both  $\mathbf{M}$  and  $\Sigma$  are estimated from the training set and the *a priori* probabilities are assumed to be equal. The algorithm first uses the training samples to calculate the mean vector and covariance matrix. Then for each pixel the probabilities for membership of each class is calculated using Eqs. (3-14) and (3-13). Finally a decision is made for the classification of each pixel using Eq. (3-12), choosing the class with the highest probability (maximum likelihood). For better comparison, the MLM segmented

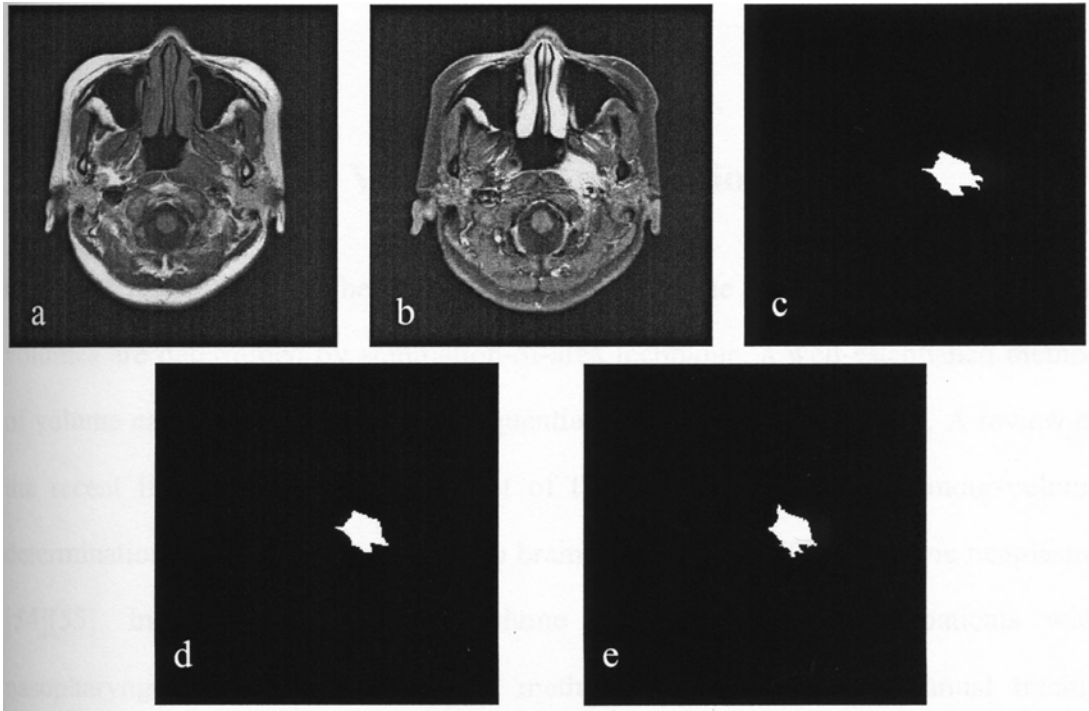
results in this study are refined using the knowledge-based image analysis described in Section 3.3.1.

Figures 3.7(a) and 3.7(b) show NPC with lateral extension in addition to posterior infiltration. The CET1W image shows a clear and homogenous enhanced tumour on the left side. Figures 3.7(d) and 3.7(e) show that both KBFC and MLM methods achieve fine segmentation results, compared with GT. Figures 3.8(a) and 3.8(b) show NPC with lateral and posterior extension as well as the invasion to tongue and the infiltration to pharynx. The CET1W image shows a large enhanced right tumour; however the tumour boundary is blurry and the pixel intensities inside the boundary are quite inhomogeneous. Figures 3.8(d) and 3.8(e) shows that KBFC achieve much better segmentation result than MLM, by comparing with GT.



**Figure 3.7** (a) T1W image; (b) CET1W image; (c) Ground truth;(d) Segmentation using KBFC; (e) Segmentation using MLM.





**Figure 3.8** (a) T1W image; (b) CET1W image; (c) Ground truth; (d) Segmentation using KBFC; (e) Segmentation using MLM.

### 3.4 NPC Tumour Volume Determination

Currently, most studies on head and neck tumour volume are CT-based where tumour volumes are determined by summation-of-area technique, a well-established method of volume calculation derived from sequential CT images [22]-[26][30]. A review of the recent literature showed that most of the MR imaging based tumour-volume determination methods were reported in brain tumours [45][60] and pelvic neoplasms [54][55]. In this study, tumour volume was measured in 10 patients with nasopharyngeal carcinoma using two methods: the conventional manual tracing method and semi-automated computer method aided by KBFC segmentation. The accuracy of KBFC segmentation method was quantitatively evaluated and the inter-operator variances of the two methods were compared.

### 3.4.1 Materials

This study consisted of 10 consecutive adult patients who had MR imaging performed for NPC staging before radiation therapy in the National Cancer Centre of Singapore. They were recruited over a 15-month period (Sept 2000 to Nov 2001). There were seven males, three females and the median age of all patients was 42.5 years (age range from 29 to 69 years). All patients had a histologically confirmed diagnosis and imaging was performed on an average of 15.2 days (range from 9 to 23 days) following biopsy. Imaging was performed using a 1.5T MR scanner (Signa, GE Medical Systems, Milwaukee, USA) using a standard clinical imaging protocol. Fast spin echo sequence (repetition time range from 500 to 800 msec, echo time range from 12.0 to 14.6 msec) was used to obtain T1W and CET1W (gadolinium enhanced with fat suppression) images. Images were acquired in both axial and coronal planes but only axial T1W and CET1W with fat suppression images (512x512 pixel, field of view range from 200 to 230 mm, slice thickness of 5 mm, inter-slice gap range from 1.8 to 2.2 mm) were used for the actual tumour volume determination in this study. Radiologists could however refer to all the other images acquired in the standard imaging protocol to aid in the accurate delineation of tumour boundary.

### 3.4.2 Methodologies

#### 1. Tumour volume determination

The actual tumour volume of NPC cannot be determined as this neoplasm is routinely treated with radiation therapy. Instead, the “ground truth” (GT) tumour volume against which measurement results from the semi-automated method could be compared, was determined by two experienced head and neck radiologists by manually tracing the tumour outline independently according to a common protocol as

described below. Both radiologists then measured the tumour volume independently using a semi-automated segmentation method. The time interval between manual tracing and semi-automated segmentation was more than two weeks to avoid any bias caused by the previous operation.

(a). Manual tracing method: Manual tracing was performed using a graphic user interface developed by the author on a personal computer (PC, 1.7GHz Pentium 4 CPU, Dell Computer Corp, Texas, USA). The area inside the outline was automatically labelled, calculated and multiplied by the MR slice thickness plus the inter-slice gap to calculate a per-slice tumour volume. NPC often involves the adjacent retropharyngeal lymph nodes and skull base. When these nodes are in contact with and cannot be separated from primary tumour, the entire mass is considered as primary tumour. When bone marrow changes are seen in contiguity with the tumour, the bone changes are included in the measurement. Similarly, when the tumour is in contact with the carotid sheath, the carotid sheath is included in the entire tumour volume measurement. It should be noted that NPC is a solid tumour and the problem of measuring or excluding cystic components does not arise. The total tumour volume was obtained by summing the volume calculations for all slices. The results are reported in Section 3.4.3.

(b). Semi-automated segmentation method: The semi-automated segmentation requires only axial images. Two radiologists independently performed KBFC to segment NPC using a graphic user interface developed by the author on the same PC mentioned above. The segmented tumour area was automatically calculated and multiplied by the MR slice thickness plus the inter-slice gap to calculate a per-slice tumour volume. The total tumour volume was obtained by summing over all the tumour-bearing slices. The results are reported in Section 3.4.3.

## 2. Validation of phantom volume measurement

KBFC segmentation algorithm was validated using a phantom model. The two compartments in the phantom were made by moulding rubber (modelling material) into structures with irregular shapes and different volumes. After the volumes of these structures were measured by water displacement, they were embedded by flour-water gel. The phantom model was imaged on a 1.5-T scanner (Siemens Medical System, Erlangen, Germany). Images were acquired by continuous scanning with a 256x256x128 matrix and a field of view of 16cmx16cmx16cm, giving a voxel resolution of 0.625mmx0.625mmx1.25mm.

Volume validation was carried out by summation of the area of continuous slices. The areas were calculated by pixel counting using (1) KBFC segmentation and (2) manual tracing and labelling (MTL). The final volume results of the structures were compared with the golden standard (GS) that was obtained by water displacement. The results are also reported in Section 3.4.3.

## 3. Validation data analysis

To quantitatively evaluate the segmentation quality, validations at the volume level and the pixel level were performed.

(a). Volume level: At the volume level, the measurement error (ME) was used to represent the accuracy of volume measurement [99][100], shown in Eq. (3-15).

$$ME(\%) = \frac{|V_s - V_{GT}|}{V_{GT}} \times 100\%, \quad (3-15)$$

where  $V_s$  is the volume obtained by semi-automated method and  $V_{GT}$  is the volume obtained by the same operator using manual tracing method on the same patient. In

addition, an index of agreement (IA) was used to represent the inter-operator agreement [45] [101] [102], with the indexes of agreement defined by

$$IA_M = 1 - \frac{2|V_{GT1} - V_{GT2}|}{V_{GT1} + V_{GT2}}, \quad (3-16)$$

and

$$IA_S = 1 - \frac{2|V_{S1} - V_{S2}|}{V_{S1} + V_{S2}}, \quad (3-17)$$

where  $IA_M$  is the value of IA from the manual tracing measurement, and  $V_{GT1}$  and  $V_{GT2}$  are the volumes obtained by operator 1 and operator 2 using manual tracing method on the same patient, respectively;  $IA_S$  is the value of IA from the semi-automated measurement, and  $V_{S1}$  and  $V_{S2}$  are the volumes obtained by operator 1 and operator 2 using semi-automated method on the same patient, respectively. An IA value of 1 indicates perfect agreement while the value of -1 indicates no agreement at all.

(b). Pixel level: Volume validation usually does not consider the location match between the manually traced tumour mass and the semi-automated segmented tumour mass. Therefore in this study, quantitative validation of segmentation results with GT at the pixel level was performed. The following were calculated: true positives (TPs, GT tumour pixels found algorithmically), false positives (FPs, pixels isolated as tumour but not within GT), false negatives (FNs, GT tumour pixels not found algorithmically) and true negatives (TNs, GT non-tumour pixels found algorithmically). The tumour regions identified by semi-automated segmentation were compared to the pixel level GT on a per slice basis. Two measures, percent matching (PM) and correspondence ratio (CR) were calculated [103][104], using the following equations:

$$PM(\%) = \frac{TPs}{GT} \times 100\%, \quad (3-18)$$

$$CR = \frac{TPs - 0.5 \times FPs}{GT}. \quad (3-19)$$

The PM, shown in Eq. (3-18), is calculated as the direct ratio of the TPs to the number of GT tumour pixels. An ideal PM is 100%, with a value of 0 indicating that there is a complete miss of any GT pixels for that slice. The CR shown in Eq. (3-19) allows us to discuss the way in which the isolated tumour corresponds in size and location to the GT tumour while weighting the importance of FPs. Although a CR of 1 is ideal, any value can be very descriptive since it negatively weighs the value of the TPs with 50% of the FPs, indicating that a FP is detrimental to a certain degree. Specifically, a negative CR indicates a greater than two to one rate of FPs with respect to TPs. Since we are interested in measuring tumour volume and location for the purpose of treatment, it is intuitive that the immediate area around the tumour will also be treated, hence the weighting mechanism in CR [103][104].

An inter-operator variance (IOV) was used to estimate the inter-operator reliability (manual tracing and KBFC segmentation algorithm) at the pixel level [105], with the IOVs defined by

$$IOV_M = 1 - \frac{A_{M1} \cap A_{M2}}{A_{M1} \cup A_{M2}} \times 100\%, \quad (3-20)$$

and

$$IOV_S = 1 - \frac{A_{S1} \cap A_{S2}}{A_{S1} \cup A_{S2}} \times 100\%, \quad (3-21)$$

where  $IOV_M$  is the value of IOV from the manual tracing method, and  $A_{M1}$  and  $A_{M2}$  are manually traced tumour masses obtained by operator 1 and operator 2 on the same image;  $IOV_S$  is the value of IOV from the semi-automated segmentation method, and

$A_{S1}$  and  $A_{S2}$  are semi-automatically segmented tumour masses obtained by operator 1 and operator 2 on the same image;  $X \cap Y$  is the overlap of  $X$  and  $Y$  while  $X \cup Y$  is the merging of  $X$  and  $Y$ . The IOV is generated from the set theory to evaluate the matching of two areas by pixel counting. A value of 0 shows the perfect reliability, while a value of 1 shows no reliability.

It is noted that in clinical practice a patient is regarded as a single unit. Therefore, the comparisons of PM, CR or IOV values are based on the total volume rather than on a slice-to-slice comparison between two observers.

#### 4. Statistical analysis

The values of  $V_{GT}$ ,  $V_S$ , ME, IA, PM, CR and IOV were all expressed by minimum, maximum, mean  $\pm$  SD format. Two-way analysis-of-variance (ANOVA) [106] was performed on  $V_{GT1}$ ,  $V_{S1}$ ,  $V_{GT2}$  and  $V_{S2}$  to evaluate the difference in volume estimation between the manually tracing and semi-automated segmentation method, and the inter-operator variance at volume level between operators 1 and 2. The inter-observer reliabilities of manual tracing and semi-automated methods were also compared, using two non-parametric Kruskal-Wallis tests [106] on  $IA_M$  and  $IA_S$  as well as  $IOV_M$  and  $IOV_S$ , respectively. Statistical significance was determined by a p value of less than 0.05. All statistical results were calculated using Matlab 6.5 (The Mathworks Inc., MA, USA).

### 3.4.3 Results

#### 1. Phantom validation

Sixty-two slices of the phantom were processed using the two methods mentioned in Section 3.4.2 (Page 54). The results are shown in Table 3.1. The errors of KBFC were

in the range of 4.2% to 5.0%, while the errors of MTL were in the range of 3.5% to 8.1%. Interestingly, the volumes derived from both manual tracing and KBFC methods were larger than the golden standards (GS) which were derived from water displacement method. It was probably due to the following reasons. For MTL, the operators may tend to draw the object margin with smooth curves so that the object volumes were somewhat overestimated. For KBFC, part of the object-background boundary may be classified as object so that the object volumes were overestimated as well.

**Table 3.1** Volume determination in phantom model consisting of two components

Unit: cm <sup>3</sup>	GS	KBFC	ME <sub>KBFC</sub> (%)	MTL	ME <sub>MTL</sub> (%)
Compartment 1	40.2	41.9	4.2	41.6	3.5
Compartment 2	26.1	27.4	5.0	28.3	8.1

Notes: GS, golden standard; KBFC, knowledge-based fuzzy clustering; MTL, manual tracing and labelling; ME<sub>KBFC</sub>, measurement error of KBFC method; ME<sub>MTL</sub>, measurement error of MTL.

2. NPC tumour volume

A total of 66 axial tumour-containing slices obtained from 10 patients were evaluated using manual tracing and KBFC segmentation methods. Tumour enhancement was seen on 3 to 10 (mean of 6.6 and SD of 2.3) slices per patient.

The results of tumour volume determination are presented in Tables 3.2. The two-way ANOVA test shows that there was no significant difference between the manually traced volume and KBFC segmented volume (p=0.906) and no significant difference existed between the volumes obtained from operator 1 and operator 2 (p=0.878). In addition, it shows that there was no interaction between operators and volume measurement methods (p=0.797).



**Table 3.2** Results of NPC tumour volume determination

	$V_{GT1} (cm^3)$	$V_{S1} (cm^3)$	$ME_1 (%)$	$V_{GT2} (cm^3)$	$V_{S2} (cm^3)$	$ME_2 (%)$
Minimum	10.40	9.51	0.44	10.15	10.38	0.01
Maximum	39.97	45.44	13.69	51.22	45.58	11.55
Mean	23.22	23.75	6.08	24.78	23.36	5.07
SD	10.45	11.89	4.77	13.44	11.82	4.06

Notes:  $V_{GT1}$  and  $V_{GT2}$ , the volume obtained by operators 1 and 2 using manual tracing;  $V_{S1}$  and  $V_{S2}$ , the volume obtained by operators 1 and 2 using KBFC method;  $ME_1$  and  $ME_2$ , the corresponding measurement error.

3. Quality of semi-automated segmentation

As an example, Table 3.3 shows the pixel numbers of TPs, FPs, and FNs as well as PM and CR, of seven different tumour-containing slices of a patient. The PM and CR of KBFC segmentation for all 10 patients compared with manual tracing segmentation are shown in Tables 3.4. These results show that the KBFC could achieve reasonable segmentation results for NPC, yet there is room for improvement.

4. Inter-operator variation

An example of the inter-operator variance (IOV) calculation is shown in Table 3.5. The  $IA$  and  $IOV$ , which represent the inter-operator variation of manual tracing method and KBFC segmentation method at the volume and pixel levels respectively, are shown in Table 3.6. Although the averaged  $IA_S$  is higher than the averaged  $IA_M$ , Kruskal-Wallis test on  $IA_M$  and  $IA_S$  shows that at the volume level there was no significant difference of inter-operator variance between manual tracing and KBFC segmentation ( $p=0.064$ ). However, Kruskal-Wallis test on  $IOV_M$  and  $IOV_S$  shows that at the pixel level the inter-operator variance of manual tracing was significantly

higher than that of KBFC segmentation ( $p=0.003$ ). Therefore from Table 3.6, the semi-automated segmentation method achieved significantly higher inter-operator consistency than the manual tracing method.

**Table 3.3** Percentage matching (PM) and correspondence ratio (CR) calculation for patient 1

Slice No.	GT	TPs	FPs	FNs	PM (%)	CR
5	3460	2718	993	742	78.56	0.64
6	4279	3744	1456	535	87.50	0.71
7	5283	4529	161	754	85.79	0.84
8	6682	4969	332	1713	74.36	0.72
9	4451	4087	592	364	91.82	0.85
10	3585	3278	655	307	91.44	0.82
11	2390	1804	332	586	75.48	0.69
Overall	27899	23796	4476	4103	83.40	0.76

Notes: GT, ground truth; TPs, true positives; FPs, false positives; FNs, false negatives; PM, percent matching; CR, correspondenceratio.

**Table 3.4** Percentage match and correspondence ratio comparing manual tracing and KBFC segmentation of two operators

	PM <sub>1</sub> (%)	CR <sub>1</sub>	PM <sub>2</sub> (%)	CR <sub>2</sub>
Minimum	83.40	0.76	77.60	0.70
Maximum	91.51	0.82	90.43	0.86
Mean	87.44	0.80	84.75	0.79
SD	2.73	0.02	4.21	0.05

Notes: PM<sub>1</sub> and PM<sub>2</sub>, percent matching of operators 1 and 2; CR<sub>1</sub> and CR<sub>2</sub>, correspondence ratio of operators 1 and 2.

**Table 3.5** Inter-operator variance (IOV) calculation for patient 1

Slice No.	Segmentation area overlap	Non-overlapping segmentation area		Inter-operator variance (%)
		Operator 1	Operator 2	
5	2892	819	527	31.76
6	4226	974	588	26.99
7	4682	8	267	5.59
8	5239	52	1390	21.58
9	4230	449	25	10.08
10	3268	695	0	17.67
11	2086	50	331	15.44
Overall	26593	3047	3128	18.85

**Table 3.6** Comparison of inter-operator reliabilities between manual tracing and KBFC segmentation

	$IA_M$	$IOV_M$ (%)	$IA_S$	$IOV_S$ (%)
Minimum	0.75	14.47	0.90	8.00
Maximum	0.98	32.01	1.00	23.20
Mean	0.92	23.70	0.96	15.61
SD	0.07	5.58	0.04	4.23

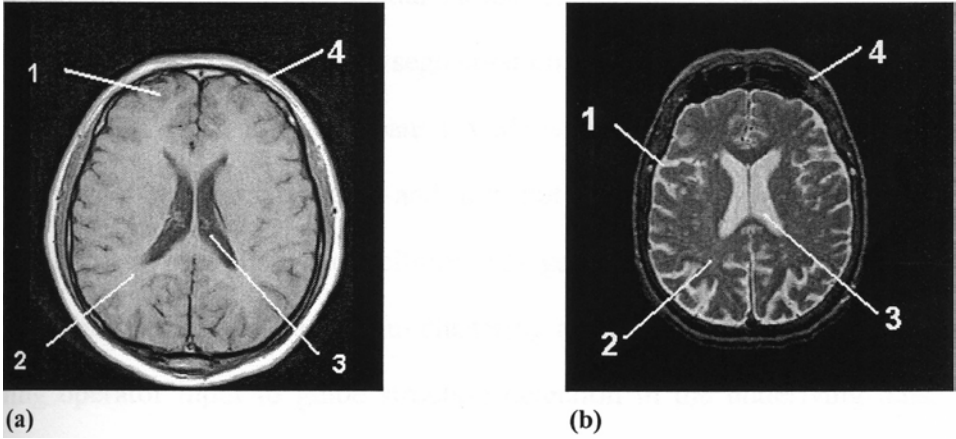
Notes:  $IA_M$  and  $IA_S$ , index of agreement of manual tracing measurement and KBFC measurement;  $IOV_M$  and  $IOV_S$ , inter-operator variance of manual tracing measurement and semi-automatic measurement.

## 3.5 Discussion

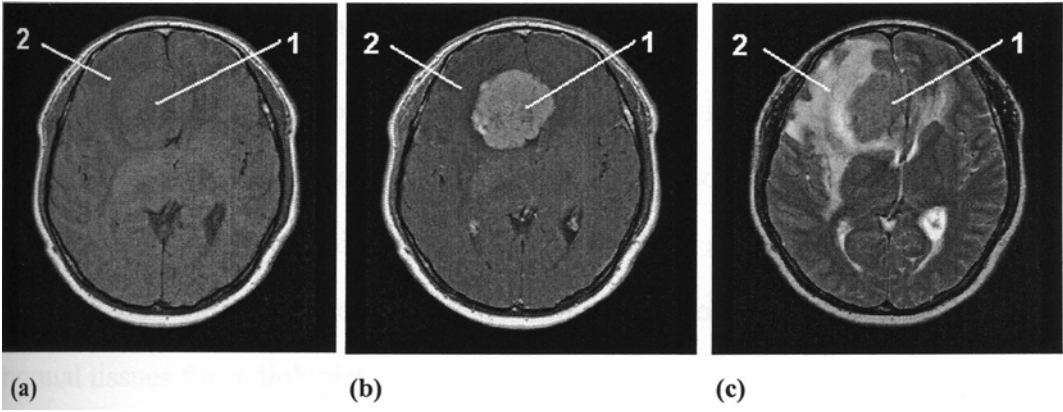
### 3.5.1 Tumour Segmentation

Currently in clinical practice, the determination of tumour volume involves tracing the tumour outline and deriving tumour volume by the summation of area techniques. Although the tracing can reflect the operator's comprehension and description for the lesion, there is always an important element of subjectivity that results in both intra- and inter-operator performance. In addition, whether this process is done by a radiologist or by a technologist, it is always tedious and time-consuming. Therefore, accurate segmentation results, good reproducibility, and easy implementation are necessary conditions for segmentation-based tumour volume determination to be a routine clinical examination for the staging and prognosis of solid tumour.

Most of the studies on MR segmentation theories, algorithms and applications were performed on the segmentation of brain structure and brain tumour. The T1W and T2W MR images of the normal brain structure are shown in Figure 3.9. It can be seen that there are four primary tissue classes: skull, grey matter, white matter, and ventricle (cerebrospinal fluid, CSF). The T1W, CET1W, and T2W MR images of brain tumour are shown in Figure 3.10. Besides the four tissue classes mentioned, there are two more tissue types: tumour and oedema (sometimes the oedema does not exist). Compared with the head and NPC images shown in Sections 3.2 and 3.3, the anatomic structures of brain are simpler than those of head and neck region. In addition, the feature distribution of each tissue type of brain image is distinct. Generally in brain images, there is no tissue type which is adjacent or adhesive to the tumour in image domain and at the same time, shows overlapping with the tumour in the feature space.



**Figure 3.9** MR images of normal brain (a) T1w image; (b) T2W image; 1. Grey matter; 2. White matter; 3. Ventricle; 4. Skull.



**Figure 3.10** MR images of brain tumour (a) T1w image; (b) CET1W image; (c) T2W image; 1. Tumour; 2. Oedema.

Extra difficulties did exist when we tried to segment NPC tumour from MRI. The main problem is how to differentiate some normal structures such as mucosa and soft palate from the tumour. The feature distributions of mucosa and soft palate are close to that of the tumour and they are often adhesive to NPC in image domain.

In this study, a knowledge-based fuzzy clustering (KBFC) method, which only requires two imaging sequences (T1W and CET1W), was developed. The guidance of the knowledge base gives this method additional power and flexibility by allowing semi-supervised segmentation and classification decisions to be made through iterative and successive refinement. It is different from most other multi-spectral efforts which attempt to segment tumour from the entire image in one step regardless

of the complex anatomic and spatial structure. The intervention of supervision is another important factor in tumour segmentation. In the works of Bensaid, et al. [39] and Clarke, et al. [81], fuzzy *c*-means (FCM) was shown to depend on subtleties in MRI performance characteristics and did not appear to be suitable for robust segmentation, although its reproducibility was good. In addition, initialisation is very important for FCM to get meaningful clustering results and reduced computation time. Enabling operator input to guide structure detection in the underlying data, SFCM utilises the operator's experience to achieve more meaningful clustering results than the unsupervised clustering and at the same time, reduce the operator dependency compared to those for fully supervised clustering.

Post-processing is also one of the important steps in this method and several anatomic and space knowledge were used. In this study, symmetry analysis was utilised to help in the removal of the normal mucosa and soft palate from the initially segmented mass since symmetry is a very important criterion to differentiate normal or abnormal tissues for radiologist.

Tests on 66 tumour-contained slices from the data sets of 10 NPC patients (Table 3.4) show that the segmentation result was satisfactory and our results show comparable values of PM and CR to the works of Fletcher-Heath, et al. [104] and Clark, et al. [103]. In other words, it could express the radiologists' interpretation for the location and extension of tumour in a high level. Two factors could affect PM considerably. One may be that the segmentation algorithm can provide a more detailed outline of the irregular tumour boundary while manual tracing on the other hand tends to approximate the tumour boundary with a smooth curve. The other is the noise and sharp changes around the tumour-normal tissue interface. In some cases, fuzzy clustering was sensitive to them. The typical example of the effects from these two factors is that in some cases, the algorithm-segmented tumour boundary is almost inside the GT boundary. Reducing FPs is also our target. In this study, mucosa, soft

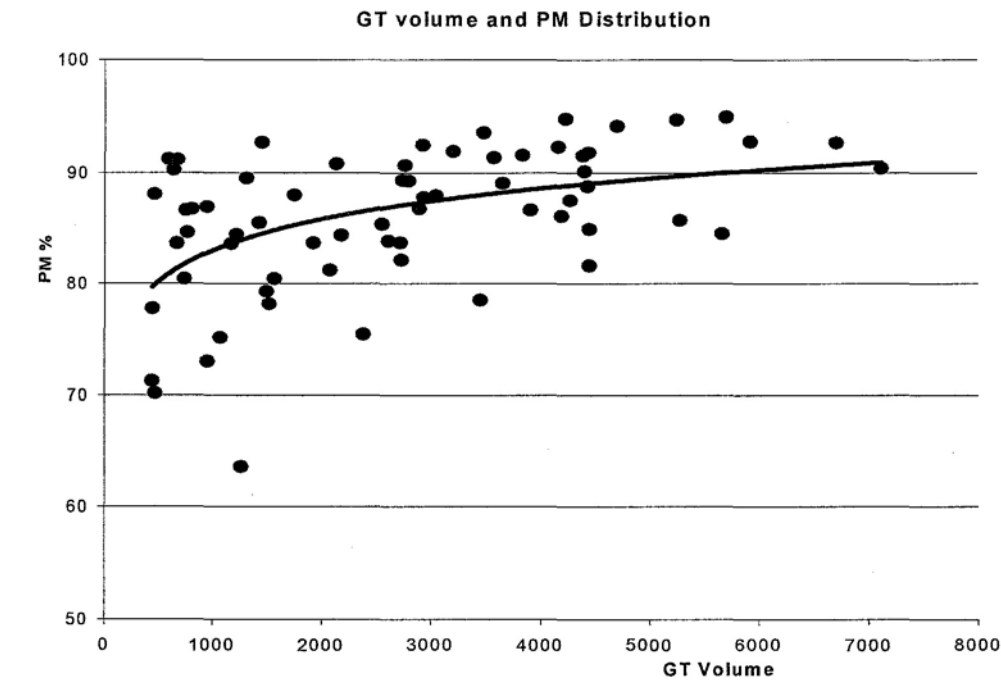
palate, inflammation, and lymph nodes near primary tumour were often segmented as FPs. In some cases, no definite dividing landmarks exist and tumour isolation for GT tracing is operator-dependent. Even for the same slice operated by the same radiologist, the judgment may change considerably on different occasions. Using the anatomic and space knowledge and without interaction, the successive image analysis procedure complied with the basic principles such as symmetry and connectivity which radiologists use, to refine the initial segmentation result. Although some FPs still did exist, the average CR value of 0.80 was acceptable.

Another interesting finding is GT volume-PM and GT volume-CR relationships. Figures 3.11 and 3.12 show the GT volume-PM and GT volume-CR distributions (volume unit: pixel; the values of GT, PM, and CR are the average of two measurements from Operators 1 and 2) from all these 66 slices segmented and two logarithmic curves were fitted. Positive correlations may exist between GT volume and PM and between GT volume and CR. This inference is reasonable because slices with small tumour volume are always those present in the inferior and superior parts of tumour, where segmentation is more difficult than those present in the intermediate part of tumour.

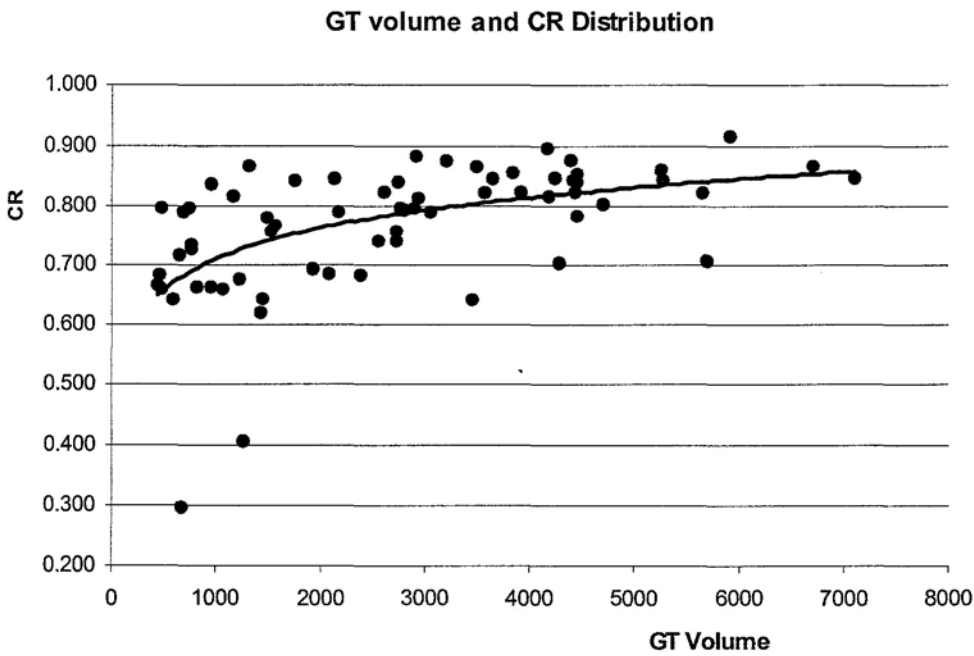
### 3.5.2 NPC Tumour Volume Determination

As mention in Section 3.5.1, technical considerations have prevented quantitative tumour volume determination from being routinely used. In the meantime, tumour diameter is routinely used as a surrogate measurement largely because of its simplicity. However staging and tumour volume estimation based on subjective and single dimensional measurement are often questionable. For instance, superficial spreading carcinomas frequently exceed 4 cm in diameter without deep penetration.

These tumours are classified as T3 lesions, but have very low volumes. In fact, there is considerable variation in tumour volume in T3 turnours.



**Figure 3.11** GT volume and PM distribution. The curve is the logarithmic curve fitted.



**Figure 3.12** GT volume and CR distribution. The curve is the logarithmic curve fitted.



Two-way analysis of variance (ANOVA) test on our results showed that there was no significant difference between manually traced and KBFC segmented volumes. Moreover, there was no significant variance between the volume results from the two operators. For a volume measurement tool to be clinically useful, it must show no significant difference among the end users. The present study has shown that semi-automated KBFC segmentation method meets this prerequisite.

Most studies compared the results of manual tracing with volumes derived from various segmentation techniques. Volume validation does not usually encompass location match between the manually traced tumour and the semi-automated segmented tumour. In this study, we performed comparisons and derived statistical parameters such as true positives, false positives, false negatives and true negatives. Location match between manually traced tumour area and KBFC segmented area provides two useful parameters, namely, percent matching (PM) and correspondence ratio (CR). These parameters provide important information regarding how well the segmented tumour volume and tumour location match the results obtained through manual tracing. PM and CR provide a quantitative assessment of segmentation quality. Analysis of PM and CR may yield useful information that may help to increase accuracy resulting in improved inter-observer reliability. In the present study, inter-operator variance (IOV), a pixel matching-based indicator of inter-operator agreement at the pixel level, was also calculated. Although from the index of agreement (IA), the indicator of inter-operator agreement at the volume level, the inter-operator agreement of KBFC segmentation method was not significantly higher than that of manual tracing method, Kruskal-Wallis test on  $IOV_M$  and  $IOV_S$  shows that at the pixel level the inter-operator agreement of KBFC segmentation was significantly higher than that of manual tracing.

In practice, determining tumour volume in the skull base can be a very difficult exercise because tumour margins cannot be determined confidently and objectively.

Skull base tumour can provoke signal changes in the marrow containing bones. Bone enhancement could either signify actual tumour invasion or represent an inflammatory response. In our study, only bone signal changes in contiguity with enhanced tumour were considered to be tumour-infiltrated. The signal changes are of special clinical interest because of the growing use of conformal radiation therapy and the need to designate the correct T-stage. Furthermore, NPC often spreads to the first echelon retropharyngeal lymph nodes [107]. The tumour and enlarged nodes may merge into a single inseparable mass. Under such circumstances, the nodal component is included in the determination of tumour volume.

In a clinical investigation performed by our group, the tumour volume of 69 NPC patients were determined using manual tracing and KBFC segmentation methods, and the relationship between tumour volume and the corresponding T-staging was compared [108]. The average tumour volumes of T1 (12.25 cm<sup>3</sup>), T2 (20.63 cm<sup>3</sup>), T3 (32.39 cm<sup>3</sup>), and T4 (44.61 cm<sup>3</sup>) show a progressive increase with progressively higher T stages. This is as expected. However, within each T stage there was a wide range (T1, in the range of 5.00 cm<sup>3</sup> to 25.42 cm<sup>3</sup>; T2, in the range of 7.76 cm<sup>3</sup> to 51.66 cm<sup>3</sup>; T3, in the range of 6.91 cm<sup>3</sup> to 70.08 cm<sup>3</sup>; and T4, in the range of 13.42 cm<sup>3</sup> to 126.21 cm<sup>3</sup>). This result is consistent to the work of Pameijer, et al. [22] and Chua, et al. [23].

A major limitation of this study is the lack of a gold standard. NPC is treated with radiation therapy. Hence, there were no surgical specimens to determine the actual tumour volumes. This limitation can be partially addressed by using a phantom to validate the segmentation methods. In some early reports, the volumes obtained from imaging-based approaches were compared with volumes of surgical specimen from water immersion [109]-[111]. Ideally a surgical procedure would provide the true volume, but sometimes there is an over-excision and in many cases, the specimen will shrink considerable after removed from the body. Therefore it may prove to be not

logistically feasible as a method for verification of segmentation methods [39]. Manual traced volumes were employed as comparison reference in more reports [45][53][58][60][103][104]. The main advantage of choose manual tracing is that it truly mimics the radiologist's interpretation, which realistically is the only "valid truth" available for *in vivo* imaging. Moreover, the manually traced tumour region can be used for the validation of segmentation methods not only at the volume level but at the pixel level by position comparison. Hence, experts traced tumour volumes was chosen in evaluating segmentation performance in the presented study.

Another limitation is the use of relatively large slice thickness and inter-slice gaps. The use of thinner sections would improve the overall accuracy of tumour volume determination. However MRI scanning using thin sections always decreases the signal-noise ratio of the image, increases examination cost, and introduces partial volume effects. In addition the increased scanning time causes more motion artefacts that make the images blurry. Therefore a 5-mm slice thickness and a 2-mm inter-slice gap are used for routine examinations.

## 3.6 Concluding Remarks

In conclusion, a knowledge-based fuzzy clustering (KBFC) method that includes noise removal, initial segmentation, and knowledge-based image analysis was established to segment NPC from T1W and CET1W MR images. After noise removal, an initial segmentation was performed using a semi-supervised fuzzy *c*-means (SFCM). Then an image analysis procedure which utilised anatomic and geometric knowledge, as well as mathematical morphological operators, was applied to refine the initial result and the final segmentation results were obtained. Sixty-six pairs of tumour-contained MRI slices from 10NPC patients were segmented using KBFC and the tumour volumes were obtained by the summation-of-areastechnique. Quantitative

validations were performed at both the volume and pixel levels by comparing KBFC segmentation volumes with those from manual tracing method. Test results showed that at the volume level there was no significant difference between the manually traced volume and the KBFC segmented volume. In addition, no significant difference existed between the inter-operator agreements of the two methods at the volume level. On the other hand, the matching comparison at the pixel level showed that KBFC can not only segment NPC from MR images in a good way, but it also has significant higher inter-operator agreement than manual tracing method. Therefore, it is concluded that the developed semi-supervised KBFC method is suitable to segment NPC for tumour volume determination.

## Chapter 4

# Unsupervised MRI Segmentation and Tumour Volume Determination for Tongue Carcinoma

### 4.1 Introduction

Tongue carcinoma usually arises from the side of the tongue. The tumour often spreads to the lymph nodes on the same side of the neck. The lymph nodes on the opposite side of the neck may also be involved. A picture of tongue carcinoma is shown in Figure 4.1. Accurate tumour staging is crucial for effective treatment planning of surgery or radiotherapy for tongue carcinoma. According to the TNM tumour classification of UICC/AJCC, the largest tumour diameter is used as a surrogate measurement of tumour size for the staging of tongue carcinoma, as shown in Figure 4.2. Tongue carcinomas are classified according to the largest tumour diameter as follows: T1 tumours (under 2 cm), T2 tumours (between 2 cm and 4 cm), and T3 tumours (more than 4 cm). Tumours that infiltrate the extrinsic muscles of the tongue (genioglossus, hyoglossus, styloglossus and palatoglossus) and show involvement of the mandible are classified as T4 tumours [6].



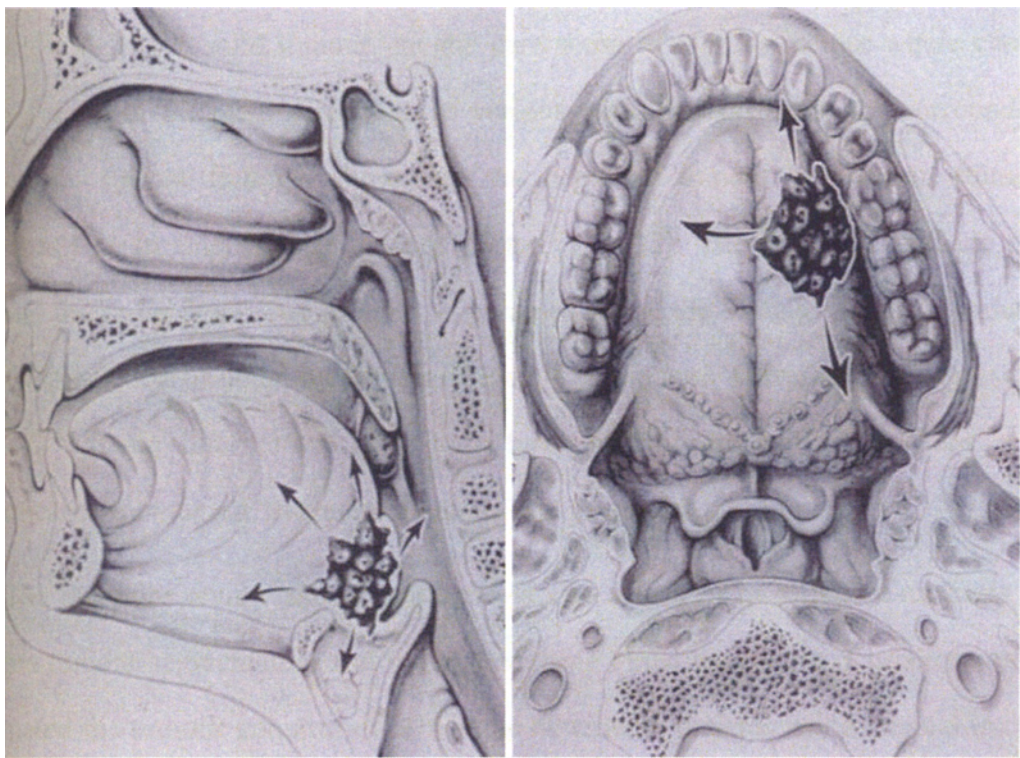


Figure 4.1 Tongue carcinoma [95]



Figure 4.2 Greatest diameter of tongue carcinoma

The greatest diameter cannot reflect accurately the total tumour volume. The main reason is that as a solid tumour, tongue carcinoma is a 3-D structure where cancer cells spread in different directions at various rates and infiltrate the surrounding structures. Recent findings show that the depth of tumour penetration (tumour thickness) is an important prognostic factor [33]-[37]. These studies indicate that tumour thickness is a better prognostic parameter than the T stage (largest diameter) for the prediction of nodal metastasis of tongue carcinoma. Therefore, the acquisition of 3-D volume information of tongue tumour from medical images has great significance for staging, treatment planning and prognosis.

The literature review shows that there is also a lack of quantitative methods for tongue tumour segmentation and volume analysis based on images. Kuriakose, et al. estimated the tumour volumes of 20 tongue carcinoma patients using manual tracing on CT images [112].

In Chapter 3, a semi-supervised method was used to segment NPC from MRI images. In this chapter, an unsupervised hierarchical segmentation scheme which is used to segment tongue carcinoma from T1W and T2W MR images is introduced. The kernel of this hierarchical scheme is from “coarse classification” to “fine clustering” based on image content, and then to “post-analysis and refinement” based on image morphology. The overall flowchart of this algorithm with multi-stage processes is shown in Figure 4.3. After image pre-processing, an initial segmentation is performed to get the initial extracted lesion regions. The procedure consists of two steps: step 1 is the unsupervised coarse classification using genetic algorithm (GA); and step 2 is the fine clustering using FCM, done on the basis of the results from coarse classification. The whole procedure combining steps 1 and 2 is also called a GA-induced fuzzy clustering (GAIFC). Due to the overlap of feature distribution, three kinds of masses may be extracted from the initial segmentation: tumour only, tumour in contact with normal structure, and normal structure only. To deal with this

problem, an automatic image analysis module using anatomic knowledge and artificial neural network (ANN) is performed to detect the actual type of extracted mass and make necessary refinement.

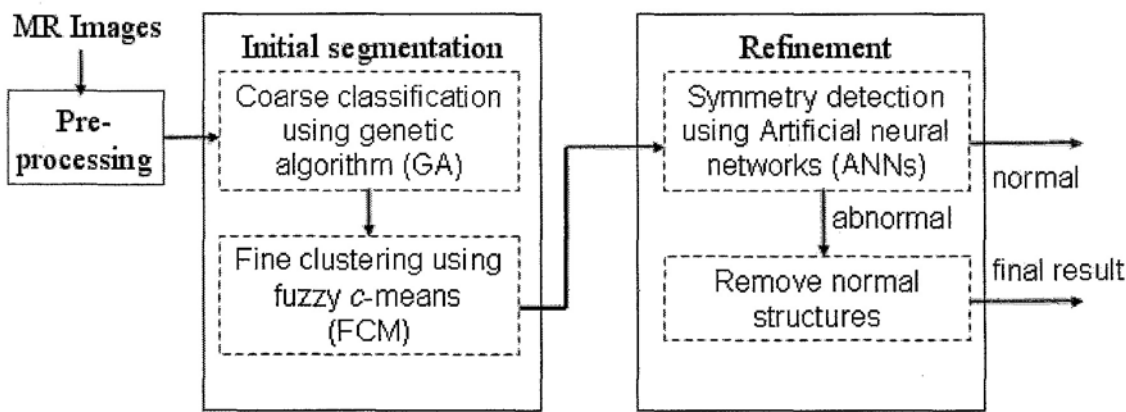


Figure 4.3 Flowchart of a hierarchical MRI segmentation scheme

Chapter 4 is organised as follows. The initial segmentation using GAIFC is presented in Section 4.2. Methodology and implementation for mass type detection using ANN and the corresponding refinement is described in Section 4.3. In addition, some visual results of tongue carcinoma segmentation are also shown in this section. Section 4.4 covers the experiment and results of tongue carcinoma tumour volume determination. The discussion and concluding remarks of this chapter are given in Sections 4.5 and 4.6, respectively.

## 4.2 Initial Segmentation Using GAIFC

### 4.2.1 Determination of Primary Tissue Classes

Figures 4.4(a) and 4.4(b) show the T1W and T2W images (512x512 pixels) of a pair of typical slices respectively after noise removal using the scheme as described in Section 3.3.1. The location of the tongue is generally fixed in the middle-anterior part



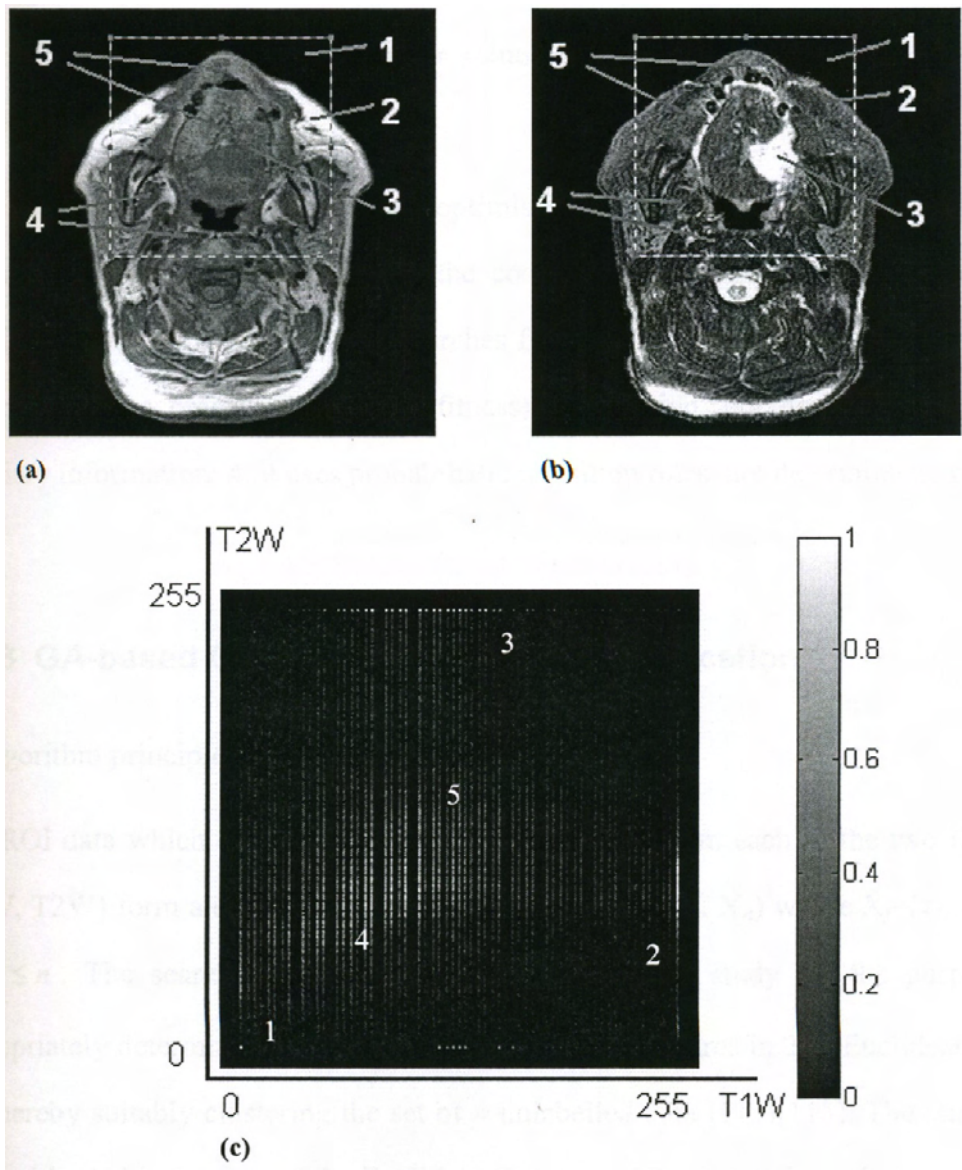
of the head, and hence, to decrease the computation cost, a rectangle region of 256x256 pixels is used as the region of interest (ROI), shown as the dashed square in the Figure. Figure 4.4(c) shows the joint histogram (after normalisation) of the ROI. According to the figures and prior knowledge, approximately five primary tissue types could be found in the feature space: (1) Air (low T1 and T2 signals); (2) fat (high T1 signal, intermediate T2 signal; (3) tumour (intermediate T1 signal, high T2 signal); (4) other normal soft tissue 1; and (5) other normal soft tissue 2. Each MR voxel of interest has a (T1W, T2W) location in  $R^2$ , forming a feature space distribution. Based on the knowledge and the fact that pixels belonging to the same tissue type will exhibit similar relaxation behaviours, they will also have approximately the same location in the feature space. In Sections 4.2.2 and 4.2.3, the initial classification using GA is described.

### 4.2.2 Principle of Genetic Algorithm

Having a large amount of implicit parallelism, GA is a randomised global search and optimisation technique guided by the principle of evolution and natural genetics. GA performs search in complex, large and multimodal landscape, and provides near-optimal solutions for the objective or fitness function of an optimisation problem [113][114]. It has found many useful applications in both the scientific and engineering areas such as pattern recognition, machine learning, microelectronics fabrication, VLSI design, control, logistics, etc.

In GA, the parameters of the search space are encoded in the form of strings (called chromosomes). A collection of such strings is called a population. Initially, a random population which represents different points in the search space is created. An objective and fitness function is associated with each string that represents the degree of goodness of the string. Based on the principle of survival of the fittest, a few of

these strings are selected and each of them is assigned a number of copies that go into the mating pool. Biologically-inspired operators like crossover and mutation are applied on these strings to yield a new generation of strings. In every generation, a new set of artificial creatures (strings) are created using parts of the fittest of the previous generations, and an occasional new part (mutation) is tried for good measure. The process of selection, crossover and mutation continues for a fixed number of generations or till a termination condition is satisfied.



**Figure 4.4** (a) T1W image; (b) T2W image; (c) Joint histogram (after normalisation) of the ROI. The five primary tissue types shown here are: 1. Air; 2. Fat; 3. Tumour; 4. Other normal soft tissue 1; 5. Other normal soft tissue 2.

GA works on the Darwinian principle of natural selection: “Survival of the fittest”. As a numerical optimiser, the solutions obtained by GA are not mathematically oriented. Instead, GA possesses an intrinsic flexibility and the freedom to choose the desirable optima according to the design specifications. Therefore, whether the criteria of concern are nonlinear, constrained, discrete, multimodal, or nonlinear programming hard, the GA is entirely equal to the challenge. In fact, because of the uniqueness of the evolutionary process and the gene structure of a chromosome, the GA processing mechanism can take the form of parallelism and multi-objective. These provide an extra dimension for solutions where other techniques may have failed completely [114].

GA differs from more traditional optimisation and search procedures in several fundamental ways: 1. it works with the coding of the decision variables, not the decision variables themselves; 2. it searches from a population of points, not a single point; 3. it uses objective function (fitness) information, not derivatives or other auxiliary information; 4. it uses probabilistic transition rules, not deterministic rules.

### 4.2.3 GA-based Clustering for Coarse Classification

#### 1. Algorithm principles

The ROI data which consist of  $n$  ( $n=256 \times 256$ ) pixels from each of the two features (T1W, T2W) form a set of pixel vectors  $\mathbf{X}=(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \dots \mathbf{X}_n)$  where  $\mathbf{X}_j=(x_{T1W}, x_{T2W})$ ,  $1 \leq j \leq n$ . The searching ability of GA is used in this study for the purpose of appropriately determining a fixed number  $K$  of cluster centres in 2-D Euclidean space  $R^2$ , thereby suitably clustering the set of  $n$  unlabelled data [113][115]. The clustering metric adopted is the sum of the Euclidean distance of the points from their respective cluster centres. The clustering metric  $M$  for the  $K$  clustering  $C_1, C_2, \dots, C_K$  is given by

$$M(C_1, C_2, \dots, C_K) = \sum_{i=1}^K \sum_{\mathbf{X}_j \in C_i} \|\mathbf{X}_j - \mathbf{Z}_i\|, \quad (4-1)$$

where  $\mathbf{Z}=(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_k)$  are the cluster centres we are seeking such that the clustering metric  $M$  is minimised, and  $\|\cdot\|$  is the Euclidean distance.

The basic steps of the GA-based clustering are as follows:

- 1) Generation  $g=0$ .
- 2) Initialise population  $P(g)$ .
- 3) Compute fitness function  $F(P(g))$ .
- 4) Set  $g=g+1$ .
- 5) If termination criterion is achieved, go to step 10).
- 6) Select  $P(g)$  from  $P(g-1)$ .
- 7) Crossover  $P(g)$ .
- 8) Mutate  $P(g)$ .
- 9) Go to step 3)
- 10) Output best results and stop.

## 2. Algorithm implementation

The detailed implementation of GA-based clustering is as follows:

- **String representation:** Each string (chromosome) was a sequence of real number representing the  $K$  cluster centres. For this study of a 2-D space with five classes, the length of a chromosome was  $2 \times 5 = 10$  words, where the first two words represented the 2-D coordinates of the first cluster center, the next two words represented those of the second cluster center, and so on. Each real number in the chromosome was an indivisible gene. For example, the chromosome

(13.5, 20.7, 81.2, 227.1, 172.2, 98.9, 71.1, 124.4, 56.5, 76.9)

represents the five cluster centres: (13.5, 20.7), (81.2, 227.1), (172.2, 98.9), (71.1, 124.4) and (56.5, 76.9).

- **Population initialisation:** There were 50 chromosomes in the population. The  $K$  cluster centres encoded in each chromosome were initialised to  $K$  randomly chosen points from the data set.
- **Fitness computation:** The clustering metric  $M$  was computed using Eq. (4-1) after the cluster designation and the replacement of cluster centres. The fitness function was defined as  $f = 1/M$ . There were two phases in this procedure. In the first phase, each point was assigned to one of the clusters  $C_i$  with center  $\mathbf{Z}_i$  such that  $\|\mathbf{X}_j - \mathbf{Z}_i\| < \|\mathbf{X}_j - \mathbf{Z}_p\|$ ,  $p=1, 2, \dots, K$ , and  $p \neq i$ . After the clustering was done, the cluster centres encoded in the chromosome were replaced by the mean points of the respective clusters. The new center  $\mathbf{Z}_i^*$  was computed as

$$\mathbf{Z}_i^* = \frac{1}{n_i} \sum_{\mathbf{X}_j \in C_i} \mathbf{X}_j, \quad (4-2)$$

where  $i=1, 2, \dots, K$ ,  $n_i$  is the number of points which belong to  $C_i$ . These  $\mathbf{Z}_i^*$ s then replaced the previous  $\mathbf{Z}_i$ s in the chromosome.

- **Selection (reproduction):** According to the survival of the fittest concept of natural genetic system, roulette wheel selection [114] was used to implement the proportional selection strategy. A chromosome was assigned a number of copies, which was proportional to its fitness in the population, and then went into the mating pool for further genetic operations. Before roulette wheel selection, a linear scaling [114] of the fitness values of chromosomes was performed to reduce the effort of genetic drift by producing an extraordinarily good chromosome in early generations and to avoid the random search behaviour in later generations.

- **Crossover:** Crossover is a probabilistic process that exchanges information between two parent chromosomes for generating two child chromosomes. In this study, a convex crossover [116] with a fixed crossover probability of  $P_c$  was used.

If  $A$  and  $B$  are two parents, then the convex crossover operator is of the form

$$\lambda_1 A + \lambda_2 B, \quad (4-3)$$

where  $\lambda_1 + \lambda_2 = 1$  and  $\lambda_1, \lambda_2 > 0$ . The parent  $A$  is replaced by  $A' = \lambda_1 A + \lambda_2 B$  and the parent  $B$  is replaced by  $B' = \lambda_1 B + \lambda_2 A$ . For example, assume that there are two parent chromosomes given by

$A = (13.5, 20.7, 81.2, 227.1, 172.2, 98.9, 71.1, 124.4, 56.5, 76.9)$  and

$B = (22.4, 23.6, 72.1, 203.5, 202.6, 79.6, 83.3, 145.8, 40.7, 87.1)$ .

Randomly, we get

$\lambda_1 = (0.10, 0.81, 0.47, 0.15, 0.31, 0.68, 0.80, 0.39, 0.56, 0.78)$ , and then

$\lambda_2 = 1 - \lambda_1 = (0.90, 0.19, 0.53, 0.85, 0.69, 0.32, 0.20, 0.61, 0.44, 0.22)$ .

Therefore, the two child chromosomes are given by

$A' = (21.5, 21.3, 76.4, 207.0, 193.2, 92.7, 73.5, 137.5, 49.5, 79.1)$ , and

$B' = (14.4, 23.0, 76.9, 223.6, 181.6, 85.8, 80.9, 132.7, 47.7, 84.9)$ .

- **Mutation:** Mutation is the occasional random alteration of the value of a gene position on a chromosome. In this study, each chromosome underwent mutation of real number presentation with a fixed probability  $P_m$  [116]. If at a gene position where mutation took place the value was  $v$ , after mutation the new value  $v'$  became

$$v' = v + \delta * v, \quad (4-4)$$

where  $\delta$  was randomly generated from -1 to 1.

- **Termination condition:** After one cycle of fitness computation, population selection, crossover and mutation, another cycle was executed on the basis of the last generation. These processes were executed for 50 iterations. Then the final resultant string with the maximal fitness value among the chromosome populations represented the final cluster centres.

3. Results of the example

The cluster centres of the ROI shown in Figures 4.4(a) and 4.4(b) are as follows: (110.0, 107.3), (220.5, 94.5), (89.0, 216.0), (74.2, 54.7) and (2.6, 4.0). The corresponding location map of these cluster centres is shown in Figure 4.5.

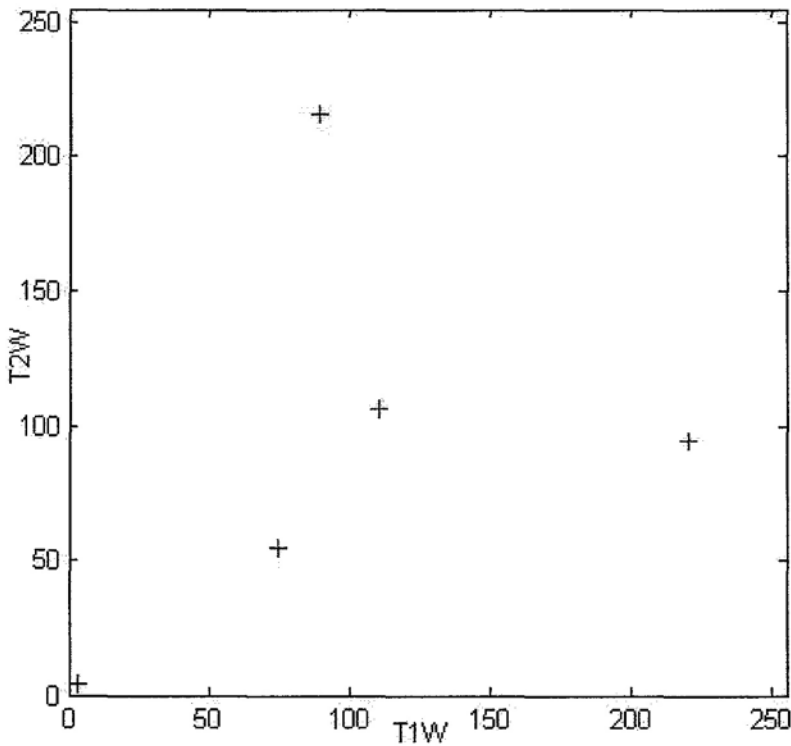


Figure 4.5 Location map of the GA classified cluster centers

### 4.2.4 Fine Clustering Using FCM

The principles of FCM were described in Section 3.2.2. The values of the  $\mathbf{U}$  matrix were found by minimizing the objective function:

$$J_2(\mathbf{U}, \mathbf{Z}; \mathbf{X}) = \sum_{i=1}^K \sum_{j=1}^n (u_{ij})^2 \|\mathbf{X}_j - \mathbf{Z}_i\|^2, \quad (4-5)$$

where  $\mathbf{X}_j$  is the objective pixel in ROI,  $\mathbf{Z}_i$  is the cluster center of class  $i$ , and  $u_{ij}$  is the membership of pixel  $j$  to tissue class  $i$ .  $\mathbf{U}=[u_{ij}]=[u_i(\mathbf{X}_j)]$ . An iteration process with the target of limiting the difference of Euclidean distance between the new and old  $\mathbf{U}$  matrices (i.e.,  $\|\mathbf{U}_{new} - \mathbf{U}_{old}\|$ ) in a threshold value was used to minimise Eq. (4-5). Here, the initial value of  $\mathbf{Z}_i$  came from the results of GA-based clustering described in Section 4.2.3 and the initial  $\mathbf{U}$  matrix was computed by

$$u_{ij} = \left[ \sum_{l=1}^K (\|\mathbf{X}_j - \mathbf{Z}_l\|^2 / \|\mathbf{X}_j - \mathbf{Z}_i\|^2) \right]^{-1}. \quad (4-6)$$

Then the cluster centres were updated by

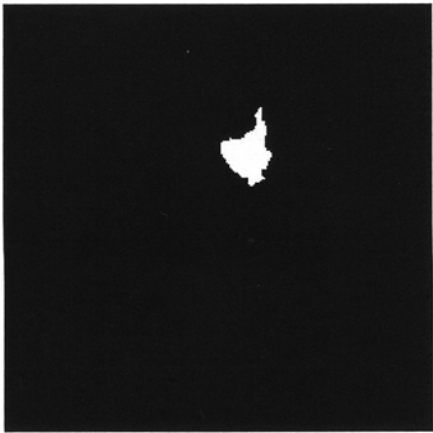
$$\mathbf{Z}_i = \sum_{j=1}^n (u_{ij})^2 \mathbf{X}_j / \sum_{j=1}^n (u_{ij})^2. \quad (4-7)$$

Therefore, Eq. (4-6) was used to calculate the new  $u_{ik}$  with the new cluster centres so that Eqs. (4-6) and (4-7) form an iterative process that converges to a local minimum of the objective function  $J_2$ . The iterative process was continued until the Euclidean distance between the new and old unlabeled columns of the  $\mathbf{U}$  matrices (i.e.,  $\|\mathbf{U}_{new} - \mathbf{U}_{old}\|$ ) was less than a threshold value  $\varepsilon = 0.01$ . This processing is similar to the processing of SFCM discussed in Section 3.2.2.

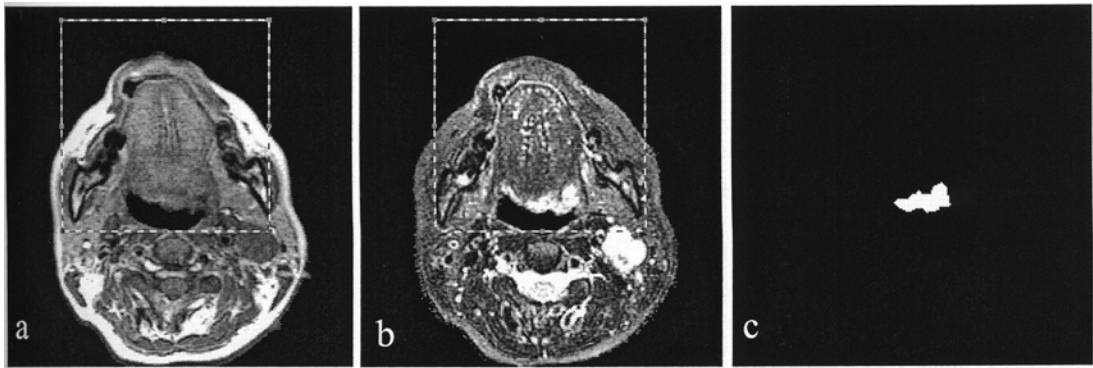
After the iteration ends, the tumour tissue was labelled according to the final  $\mathbf{U}$  matrix and some mathematical morphology refinements such as dilation and erosion



were applied to get the initial tumour segmentation. The initial segmentation result of the ROI shown in Figure 4.3 is shown in Figure 4.6. Another example of MR T1W and T2W images of posterior tongue carcinoma and the initial segmentation result are shown in Figure 4.7.



**Figure 4.6** Initial segmentation result of the ROI shown in Figure 4.3



**Figure 4.7** (a) T1W image; (b) T2W image; (c) Initial result.

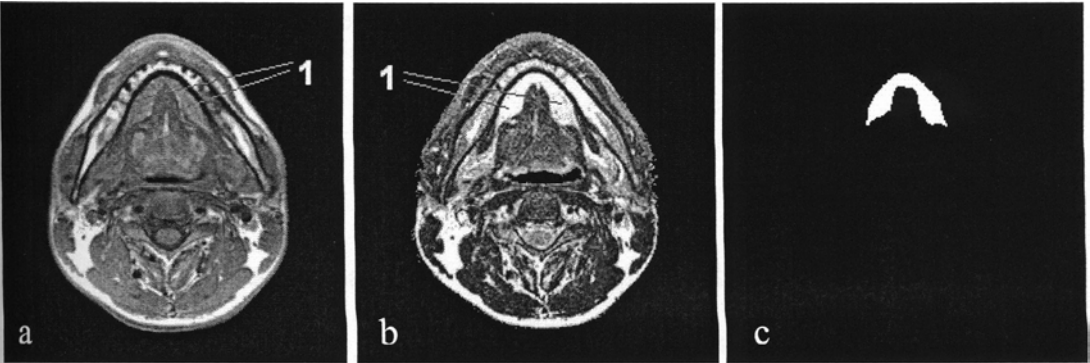
### 4.3 ANN-based Symmetrical Detection and Refinement

#### 4.3.1 Symmetrical Analysis for Mass Type Detection

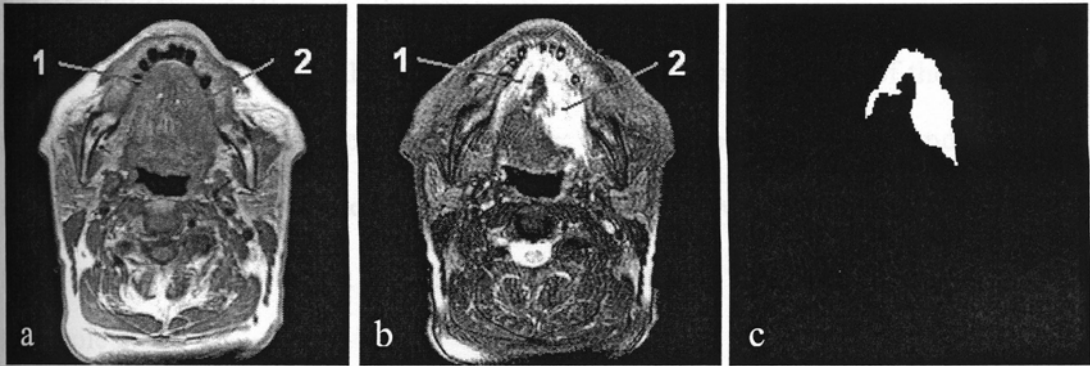
In MR signal feature space for head and neck cancer imaging protocols, the sublingual glands which are in the lower section of the tongue, always have similar

MR characteristics to tumour (i.e., intermediate intensity in T1 signal and high intensity in T2 signal). Furthermore, in some slices, the sublingual gland is in contact with the primary tumour and cannot be separated from the primary tumour. Therefore, the results from initial automatic segmentation of these slices containing sublingual glands have high false positives. Figure 4.8 shows the MR images of normal sublingual glands and the initially extracted sublingual glands. Figure 4.9 shows the MR images of a common type of tumour in contact with the sublingual glands and the corresponding initially extracted lesion mass.

Generally, there are three types of masses which may be automatically extracted during the procedure of initial segmentation. They are normal sublingual glands only, primary tumour only, and tumour with involved sublingual gland. In the following, their morphological behaviours are discussed and an ANN is used to detect the corresponding mass type.



**Figure 4.8** (a) T1W image; (b) T2W image; (c) Extracted sublingual glands; 1. Sublingual glands.



**Figure 4.9** (a) T1W image; (b) T2W image; (c) Extracted tumour and involved sublingual gland; 1. Normal sublingual gland; 2. Tumour-involved sublingual gland.

### 1. Analysis criterion

Under normal conditions, the tongue, palate and sublingual glands are morphologically bilaterally symmetric in the oral cavity. On the other hand, the tongue carcinoma always shows an asymmetric growth pattern. Therefore, in the image domain, most normal tissues are roughly symmetrical along the vertical axis, while tumours often have poor symmetry. In the present study, an ANN-based symmetry analysis procedure is developed to detect the Symmetry of the initially segmented masses, make normal/abnormal decisions, and refine the last result for cases which are detected as tumour with the involved sublingual gland. The refinement criterion is based on the tongue tumour manual tracing protocol: When the tongue carcinoma is in contact with the ipsilateral sublingual gland, the sublingual gland is considered as a part of the total tumour volume. The contralateral sublingual gland is not included in the volume because it is a separate structure and is rarely involved (on MR images, the sublingual glands often appear to be in contact).

### 2. Selection of the features

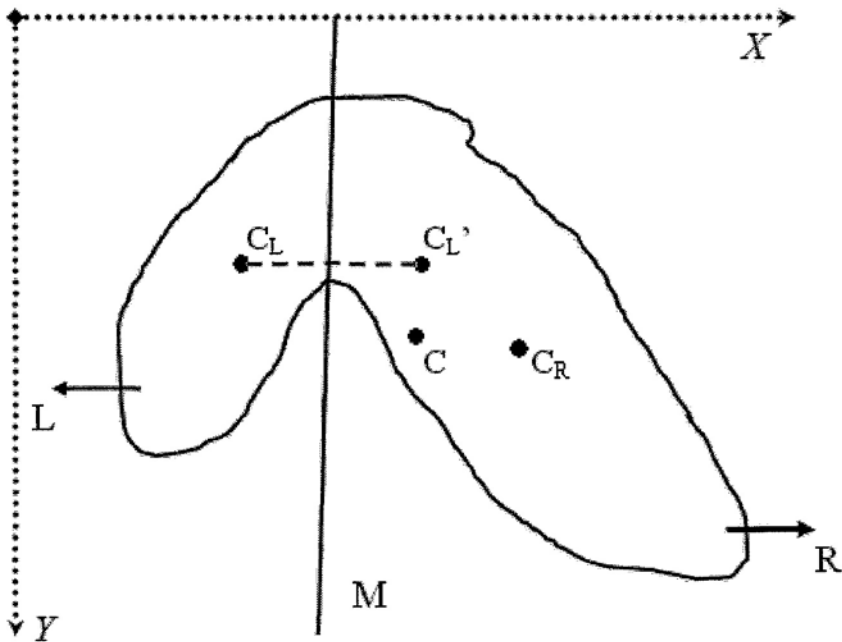
There are many different shapes of extracted normal sublingual glands and tumour-involved sublingual glands, hence one or two features are not enough to make the judgment. In this study, five features are selected as the inputs of an ANN to judge the symmetry. Figure 4.10 is an illustration of the extracted tumour-involved sublingual gland. In Figure 4.10, M is the tongue midline; L is the normal part of the sublingual gland in the left lateral of the tongue midline;  $C_L$  is its geometry center and  $C_L'$  is the mirror point of  $C_L$  in the other lateral; R is the tumour-involved part of the sublingual gland in the right lateral;  $C_R$  is its geometric center; and C is the geometric center of the whole mass.

Based on these definitions, the following five features are defined:

i) Difference-of-masses  $DM$ :  $DM = |N_L - N_R|$ , where  $N_L$  and  $N_R$  are the numbers of pixels of regions L and R, respectively.  $DM$  has a small value for structures of normal sublingual glands. In the case of tumor involving,  $DM$  has a large value.

ii) Fuzzy symmetric measure  $FSM$ :  $FSM = \left[ 1 + \left( \frac{N_L - N_R}{100} \right)^2 \right]^{-1}$ , where  $N_L$  and  $N_R$

are the numbers of pixels of regions L and R, respectively.  $FSM$  approximately equals to one for normal sublingual gland structures. In abnormal cases,  $FSM$  is much smaller than one [71].



**Figure 4.10** The illustration of the structures of tumour-involved sublingual glands

iii) Distance-of-centres  $DC$ :  $DC = |C_L' - C_R|$ , where  $| |$  denotes the distance.  $DC$  approximately equals to zero for normal sublingual gland structures. In abnormal cases,  $DC$  is much larger than zero.

iv) Difference-of-standard deviations  $DS$ :  $DS = \sqrt{(STD_X^L - STD_X^R)^2 + (STD_Y^L - STD_Y^R)^2}$ , where  $STD_X^L$  and  $STD_Y^L$  are the standard deviations of x and y coordinates of the

- pixels in L,  $STD_X^R$  and  $STD_Y^R$  are the corresponding quantities in R.  $DS$  approximately equals to zero for normal sublingual gland structures.  $DS$  is much larger than zero in the case of tumour involving.
- v) Distance-of-center-to-midline  $DCM$ :  $DCM$  is the distance from C to M.  $DCM$  approximately equals to zero for normal sublingual gland structures.  $DCM$  is much larger than zero in the case of tumour involving.

3. Neural networks implementation

The five features defined above were calculated and fed into a two-layer supervisory back propagation neural network (BPNN) as shown in Figure 4.11, where  $[x_1, x_2, x_3, x_4, x_5] = [DM, FSM, DC, DS, DCM]$  is the input and  $y$  is the output. Supervisory learning was performed on this classifier net to determine the network weights  $w_{ij}$  using the features from three normal masses and three abnormal masses before the actual detection. If a mass was detected as symmetrical, the original slice was regarded as a normal slice. Otherwise, the original slice was an abnormal one. Thus the normal part of sublingual gland was removed automatically by a smooth transition and the tumour-involved part was retained.

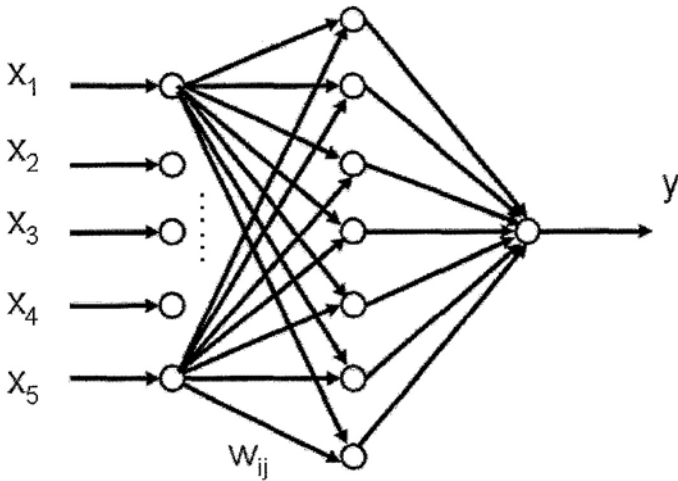


Figure 4.11 : Structure of a BPNN

The creation and training of this BPNN are described in the following Matlab 6.5 codes:

```
net = newff(minmax(training_input),[7,1],{'tansig', 'purelin'}, 'trainscg');  
  
% create a feed-forward BPNN, with 7 nodes in the 1st layer and 1 node in the 2nd layer  
  
% layer transfer function: hyperbolic tangent transfer function and liner transfer function  
  
% BPNN training function: scaled conjugate gradient back-propagation  
  
net.trainParam.epochs = 5000;  
  
% set parameter: the network is trained for 5000 epochs  
  
net.trainParam.goal = 1e-5;  
  
% set parameter: the training stops if the difference of weight matrix is less than 10-5  
  
net = train(net, training_input, training_output);  
  
% start training using the training data
```

In practice, the training stopped after 550 epochs and its process is shown in Figure 4.12.

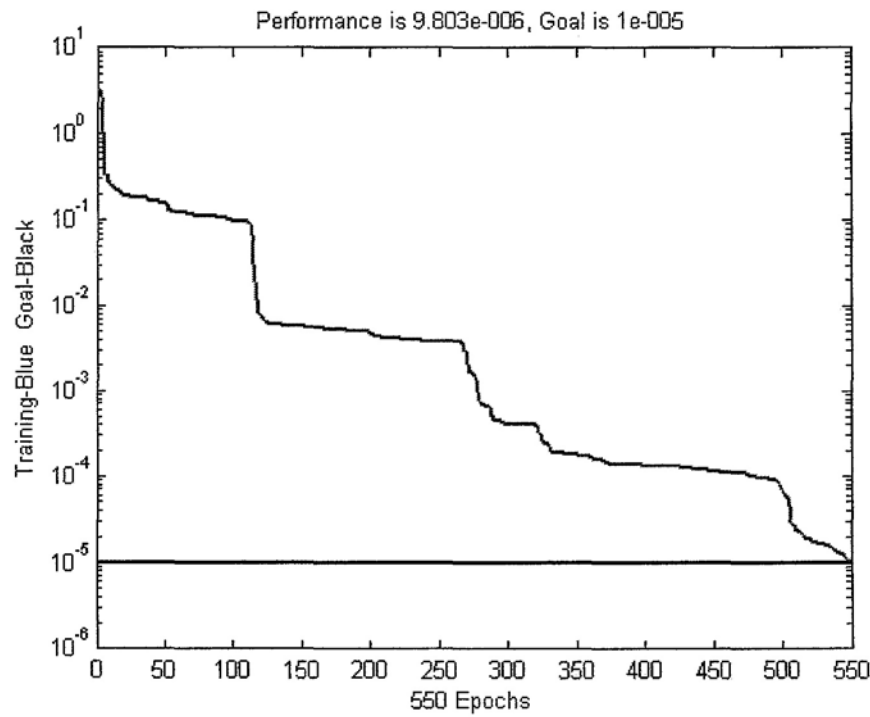


Figure 4.12: The training process of the BPNN



4.3.2 Final Segmentation Results

Figure 4.13 shows the final segmentation result of the images shown in Figure 4.9 and the corresponding GT traced by an experienced head and neck radiologist. Figure 4.14 shows the original T1W and T2W images of the tumour with the involved sublingual glands, the corresponding initial and final segmentation results as well as the GT. The algorithm successfully extracted the pathological lesion, made normal/abnormal decisions and performed the refinement for the final results.

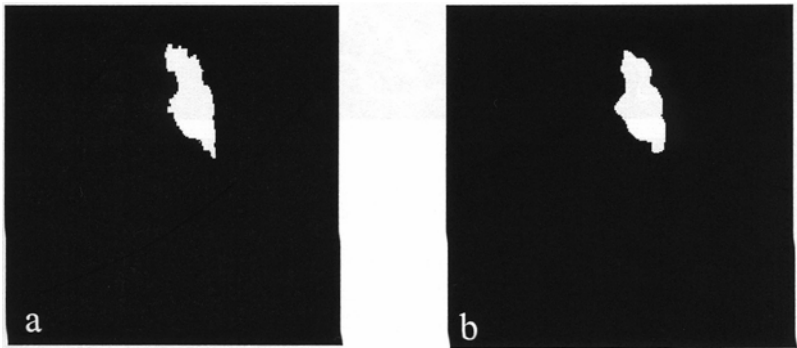


Figure 4.13 (a) Final result of images in Figure 4.9; (b) GT.

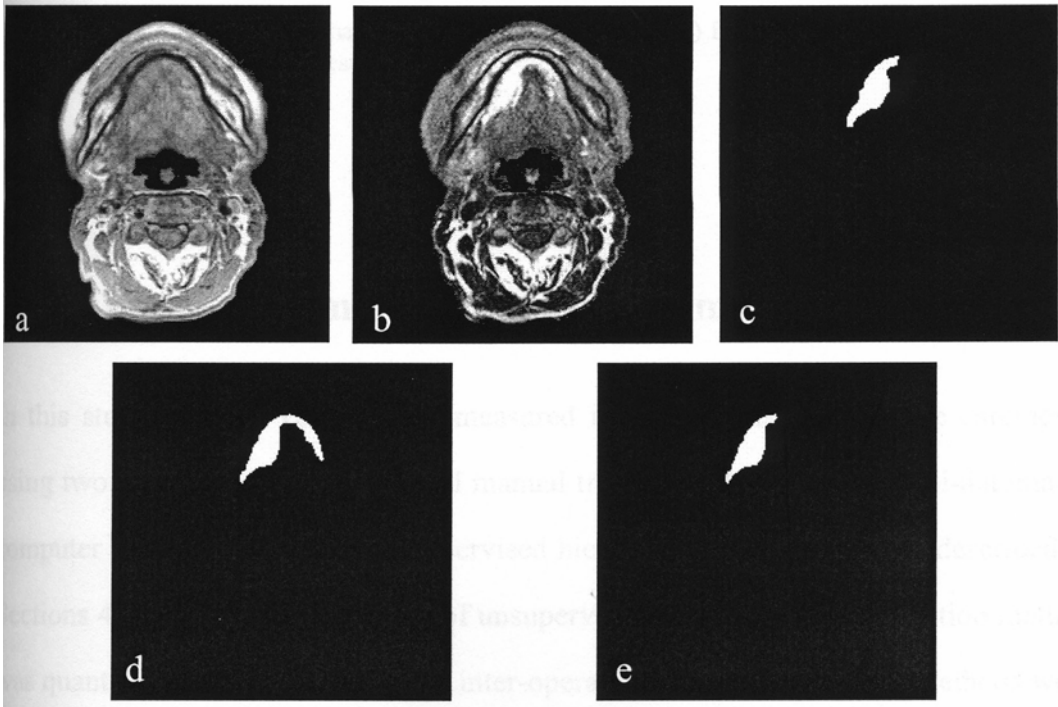
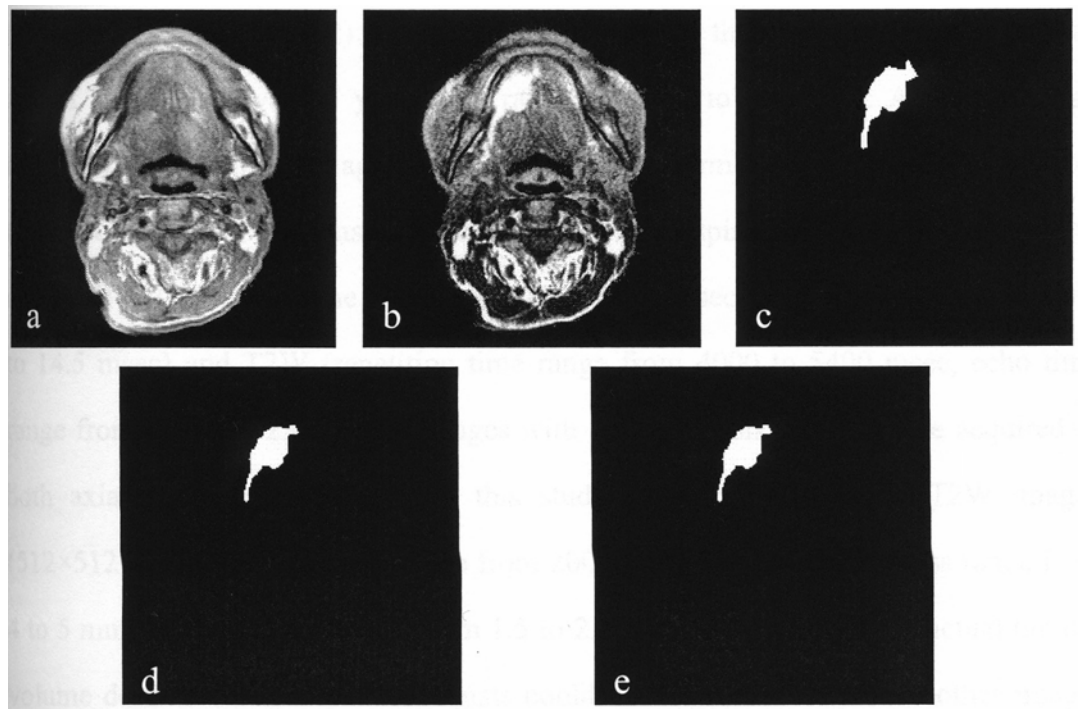


Figure 4.14 (a) T1W image; (b) T2W image; (c) GT; (d) Initial segmentation result; (e) Final segmentation result.

Figure 4.15 shows the original T1W and T2W images of unilateral tongue tumour, the corresponding initial and final segmentation results, and the GT. Since it is a unilateral primary tumour without involving the sublingual gland, no refinement was performed after the symmetrical detection.



**Figure 4.15** (a) T1W image; (b) T2W image; (c) GT; (d) Initial segmentationresult; (e) Final segmentationresult.

## 4.4 Tongue Carcinoma Tumour Volume Determination

In this study, tumour volume was measured in 10 patients with tongue carcinoma using two methods: the conventional manual tracing method and the semi-automated computer method aided by an unsupervised hierarchical segmentation as described in Sections 4.2 and 4.3. The accuracy of unsupervised hierarchical segmentation method was quantitatively evaluated and the inter-operator variances of the two methods were compared.



### 4.4.1 Materials

This study consisted of 10 consecutive adult patients who had MR imaging performed for tongue carcinoma staging in the National Cancer Centre of Singapore before radiation therapy or surgical resection. They were recruited over a 22-month period (June 2000 to March 2002). There were seven males, three females and the median age of all patients was 55 years (age range from 33 to 77 years). All patients had histologically confirmed diagnosis. Imaging was performed using a 1.5T MR scanner (Signa, GE Medical Systems, Milwaukee, USA). Fast spin echo sequence was used to obtain T1W (repetition time range from 420 to 720 msec, echo time range from 10.2 to 14.5 msec) and T2W (repetition time range from 4000 to 5400 msec, echo time range from 82.0 to 123.6 msec) images with fat saturation. Images were acquired in both axial and coronal planes. In this study only axial T1W and T2W images (512x512 pixels, field of view range from 200 to 220 mm, slice thickness range from 4 to 5 mm, inter-slice gap range from 1.5 to 2.2 mm) were used for the actual tumour volume determination. The radiologists could however refer to all the other images acquired in the imaging protocol to aid in the accurate delineation of tumour.

### 4.4.2 Methodologies

#### 1. Tumour volume determination

This was a retrospective study and we did not have access to surgical specimens for the actual measurement of specimen volumes. Therefore, the “ground truth” (GT) tumour volume, against which measurement result from semi-automated methods could be compared, was determined by two experienced head and neck radiologists by manually tracing the tumour outline independently. Two PhD students of biomedical

engineering independently measured the tumour volume using the unsupervised hierarchical segmentation methods.

a. Manual tracing method: Manual tracing was performed on T2W images using a graphic user interface developed by the author on a PC (1.7 GHz Pentium 4 CPU, Dell Computer Corp, Texas, USA). This choice was based on the good contrast between the high signal intensity tumour and the fat-suppressed low signal intensity tongue muscles. The area inside the outline was automatically labelled, calculated and multiplied by the MR slice thickness plus the inter-slice gap to calculate a per-slice tumour volume. The total tumour volume was obtained by summing the volume calculations for all slices.

b. Semi-automated method: The semi-automated method requires only axial images. Two students pursuing PhD degree in biomedical engineering independently performed the unsupervised hierarchical method to segment tongue carcinoma, using a graphic user interface developed by the author on the same PC mentioned above. The segmented tumour area was automatically calculated and multiplied by the MR slice thickness plus the inter-slice gap to calculate a per-slice tumour volume. The total tumour volume was also obtained by summing the tumour bearing slices. The segmentation itself is unsupervised, and hence the whole tumour volume determination procedure should be automated. However, before segmentation, the operators were required to decide the ROI of fixed size and determine the tongue midline interactively (the reason is discussed in Section 4.5.1). Hence, the whole procedure is not fully automated and the inter-operator variance was evaluated.

## 2. Validation of phantom volume measurement

The unsupervised hierarchical segmentation algorithm was also validated using the phantom model described in Section 3.4.2. Volume validation was carried out by

summation of the area of continuous slices. The areas were calculated by pixel counting using: (1) the unsupervised hierarchical segmentation (UHS) and (2) manual tracing and labelling (MTL). The final volume results of the structures were compared with the golden standard which was obtained by water displacement.

### 3. Validation data analysis

To quantitatively evaluate the qualities of unsupervised hierarchical segmentation and the corresponding semi-automated tumour volume measurement, validations at the volume level and pixel level were performed using the same analysis presented in Section 3.4.2.

### 4. Statistical analysis

In this study, the same statistical analysis was performed as described in Section 3.4.2.

## 4.4.3 Results

### 1. Phantom validation

Sixty-two slices of the phantom were processed using the two methods mentioned in Section 4.4.2. The results are shown in Table 4.1. The errors of unsupervised hierarchical segmentation were in the range of 4.7% to 6.9% while the errors of MTL were 3.5% to 8.1%. Interestingly, the volumes derived from manual tracing and unsupervised hierarchical methods were larger than the golden standards (GS) which were derived from water displacement.

**Table 4.1** Volume determination in phantom model

Unit: cm <sup>3</sup>	GS	UHS	ME <sub>UHS</sub> (%)	MTL	ME <sub>MTL</sub> (%)
Compartment 1	40.2	42.1	4.7	41.6	3.5
Compartment 2	26.1	27.9	6.9	28.3	8.1

Notes: GS, golden standard; UHS, unsupervised hierarchical segmentation; MTL, manual tracing and labelling; ME<sub>UHS</sub>, measurement error of UHS method; ME<sub>MTL</sub>, measurement error of MTL.

2. Tongue carcinoma tumour volume

A total of 52 axial tumour-containing slices obtained from 10 patients were evaluated using manual tracing. Tumour enhancement was seen on 2 to 8 (mean of 5.2 and SD of 1.8) slices per patient. While using the unsupervised hierarchical segmentation methods, lesion masses were extracted from 66 slices. Among these, 52 slices were classified as “abnormal” or “tumour-involved” and 14 were “normal sublingual glands”. Compared with GTs, only one slice of normal sublingual glands structure was classified as “abnormal” and one tumour-involved structure was wrongly classified as “normal”.

The results of tumour volume determination are presented in Table 4.2. The two-way ANOVA test shows that there was no significant difference between the manually traced volume and the unsupervised hierarchical segmented volume (p=0.989), and no significant difference existed between the volumes obtained from operator 1 and operator 2 (p=0.937). In addition, there did not appear to be any interaction between operators and volume measurement methods (p=0.946).

**Table 4.2** Results of tongue carcinoma tumour volume determination

	$V_{GT1} \text{ (cm}^3\text{)}$	$V_{S1} \text{ (cm}^3\text{)}$	$ME_1 \text{ (%)}$	$V_{GT2} \text{ (cm}^3\text{)}$	$V_{S2} \text{ (cm}^3\text{)}$	$ME_2 \text{ (%)}$
Minimum	2.34	3.13	1.03	2.82	3.10	0.32
Maximum	24.80	24.47	33.76	24.64	25.69	14.77
Mean	13.70	13.58	7.07	13.72	13.90	4.58
SD	6.77	6.57	10.02	7.01	6.98	4.29

Notes:  $V_{GT1}$  and  $V_{GT2}$ , the volume obtained by operators 1 and 2 using manual tracing;  $V_{S1}$  and  $V_{S2}$ , the volume obtained by operators 1 and 2 using the unsupervised hierarchical method;  $ME_1$  and  $ME_2$ , the corresponding measurement error.

3. Quality of the segmentation algorithm

The percent matching (PM) and corresponding ratio (CR) of the unsupervised hierarchical segmentation for all 10 patients compared with manual tracing segmentation are shown in Table 4.3. These results show that the unsupervised hierarchical method could achieve satisfactory segmentation results for tongue carcinoma.

**Table 4.3** Percentage match and correspondence ratio comparing manual tracing and the unsupervised hierarchical segmentation of two operators

	$PM_1 \text{ (%)}$	$CR_1$	$PM_2 \text{ (%)}$	$CR_2$
Minimum	81.23	0.71	82.81	0.75
Maximum	95.16	0.87	94.17	0.90
Mean	88.44	0.82	89.16	0.83
SD	4.04	0.05	4.21	0.05

Notes:  $PM_1$  and  $PM_2$ , percent matching of operators 1 and 2;  $CR_1$  and  $CR_2$ , correspondence ratio of operators 1 and 2.

4. Inter-operator variation

The index of agreement (IA) and inter-operator variance (IOV), which represent the inter-operator variation of manual tracing method and the unsupervised hierarchical segmentation method at both the volume and pixel levels respectively, are shown in Table 4.4. The averaged  $IA_S$  is higher than the averaged  $IA_M$ , however, Kruskal-Wallis test on  $IA_M$  and  $IA_S$  shows that at the volume level there was no significant difference in inter-operator variance between manual tracing and the unsupervised hierarchical segmentation ( $p=0.401$ ). Furthermore, Kruskal-Wallis test on  $IOV_M$  and  $IOV_S$  shows that at the pixel level, the inter-operator variance of manual tracing was significantly higher than that of the unsupervised hierarchical segmentation ( $p=0.001$ ). Therefore, it can be inferred that the unsupervised hierarchical segmentation method achieved significantly higher inter-operator consistency than the manual-tracing method.

**Table 4.4** Comparison of inter-operator reliabilities between manual tracing and the unsupervised hierarchical segmentation

	$IA_M$	$IOV_M$ (%)	$IA_S$	$IOV_S$ (%)
Minimum	0.78	12.77	0.92	7.71
Maximum	0.99	24.33	0.99	20.23
Mean $\pm$ SD	$0.94\pm0.06$	$18.92\pm3.69$	$0.96\pm0.03$	$11.04\pm3.63$

Notes:  $IA_M$  and  $IA_S$ : index of agreement of manual tracing measurement and semi-automated measurement;  $IOV_M$  and  $IOV_S$ : inter-operator variance of manual tracing and the unsupervised hierarchical segmentations.

## 4.5 Discussion

### 4.5.1 Unsupervised Segmentation

Tongue carcinoma is tumour arising from the mucosal surface of the tongue and frequently invades tongue muscles. Compared with NPC, the whole tongue carcinoma-affected domain is relatively simple and fixed, and only a few normal structures such as the sublingual gland, may be involved by the tumour. In this study, an unsupervised hierarchical method combining GA-induced fuzzy clustering (GAIFC) and artificial neural network (ANN)-based analysis procedure is developed.

Regarding image segmentation problems, our concern lies in the segmentation accuracy with acceptable operator dependency and computation complexity. Unsupervised FCM gave very good reproducibility. However it did not necessarily arrive at meaningful segmentations and often required long computation time [39]. SFCM, a method with partial supervision, yielded improved results and reduced computation time by the introduction of initial training sets. However, the selection of training sets was operator-dependent. In our proposed method, the coarse tissue types were classified using GA-based initial clustering which used crisp tissue membership. Following this, as to what the selected training sets in SFCM acted, the results were fed into the successive fuzzy clustering which used fuzzy tissue membership, to guide the final clustering results. As a result, GAIFC achieved improved and meaningful results of initial segmentation as SFCM did and at the same time, reducing operator dependency. The relative high computation complexity was the cost of this algorithm due to the introduction of GA. However, the long computation time could be reduced by algorithm optimisation and the rapid progress of computer hardware. For instance, the ROI used for GA-based clustering could be re-sampled by one half or one quarter of the original resolution. Furthermore the result of GA-based clustering from one

slice could be used as the initial value for succeeding FCM for another slice in the same image volume.

The ANN-based symmetrical analysis for segmentation refinement is also an important part in this method. Symmetry is an important factor that helps radiologists to read films and make diagnosis. A fuzzy symmetry measure was used as the sole factor to judge the symmetry in a previous research on brain tumour extraction [71]. In our study, because the shapes of extracted tumour-involved sublingual glands or normal sublingual glands vary significantly in shape and size, one or two hard criteria are not enough to make the correct judgment. Therefore, more features were selected or calculated to help the symmetry detection and a supervisory BPNN was selected as a classifier net. The final results show the flexibility and learning ability of the BPNN for this kind of fuzzy judgment problems.

Besides the initially segmented masses, another parameter that affects the symmetry detection is the tongue midline which is used as the symmetry axis. In the previous chapter of NPC segmentation, the symmetry axis of the axial head image was automatically fitted from three points. Calculated as the corresponding first moments, the three points were the geometric centres of the upper half, bottom half and the whole part of the binary head mask. In ideal conditions the head symmetrical axis is quite closed to the tongue midline, and yet this processing is not universally reliable for the generation of tongue midline. This is because that during the scanning process, the tongue midline can be significantly changed by the orientation of the head, the distortion of non-skull covered part of head, and the swelling or atrophy of local internal organs. In this study, an interactive method which uses distinct anatomical landmarks such as lingual septum was adopted to help in the determination of the tongue midline. Therefore, the segmentation process itself was unsupervised while the whole procedure was not fully automated and the inter-operator variance was introduced.



### 4.5.2 Potential Development in Segmentation

Pattern recognition is a classical and promising technique for multi-spectral image segmentation. An important problem in the segmentation of head and neck tumours from MRI is how to decrease the false positives (FPs) in the first stage of segmentation. According to the detailed technique considerations of this study, the potential developments are discussed as follows.

#### 1. The introduction of new features

Currently, signal intensities are the most important features used in the multi-spectral segmentation of MRI. Generally, for tissues like mucosa, soft palate and sublingual glands which have high density of micro-vessels, the MRI signals (intermediate signal intensity in T1W and high signal intensity in CET1 W or T2W) are similar to those of abnormal tissues such as tumour. MRI, however, has a very high resolution for soft tissue and sometimes there are subtle differences in the T1W images between the normal and abnormal tissue. The number of tissue classes used in the actual segmentation is limited due to the computation complexity and hence, it is difficult to detect the differences in the domain of intensity distribution. However, in the spatial frequency distribution domain, these differences may be detected using texture measure. Therefore, the introduction of texture feature into the original 2-D feature space formed by the intensities of two sequences of images may be helpful. In a preliminary study performed by our group, a texture combined multispectral fuzzy clustering was developed. A frequency spectrum-based texture measure of T1W image was introduced in this algorithm as a new feature for classification and promising results were acquired for some slices [117].

## 2. Machine learning technique

From the theory of pattern recognition, the problem of high FPs in initial segmentation results due to misclassification. is the challenge of how to deal with the nonlinearity of real medical data distribution in the feature space. For most of the statistical methods and fuzzy logic approaches such as  $k$  nearest neighbour ( $k$ NN), maximal likelihood method (MLM) and FCM, a specific parameterised distribution of the object data is usually defined. It means that such approaches have implicitly imposed some *a priori* assumption about the data distribution and they are only useful when the data distributions of different tissue classes are well-known [39]. However for real medical data, the actual distributions of tissue classes usually do not fit to these assumptions well and moreover, they are not linearly separable in some cases. A potential solution is the supervised or unsupervised machine learning which is used to learn the actual data distribution of the object data intelligently and optimally turn out a flexible decision boundary.

A simple application of learning-based technique is the change of feature weights. In most of the previous studies on multispectral MRI segmentation, the weights of all the features (signal intensities) used for classification were the same. In our studies, we found that the signal intensity distribution of lesion and its adjacent region in CET1W image or T2W images is quite homogenous, while it is often heterogeneous in T1W images. Therefore, the feature vector used for classification can be changed from  $(I_{T1W}, I_{CET1W})$  to  $(I_{T1W}^m, I_{CET1W}^n)$ , where  $I_{T1W}$  is the intensity of T1W image,  $I_{CET1W}$  is the intensity of CET1W,  $m + n = 2$ , and  $m/n = (\text{standard deviation of training data of tumour T1W image})/(\text{stand deviation of training data of tumour in CET1W image})$ . This process can nonlinearly stretch the feature space that makes the data distribution linearly separable.

In recent years, a new machine learning method, support vector machine (SVM), has been studied and showed encouraging results for nonlinear classification and clustering problems [118][119]. Our group has started the study on tumour segmentation based on multi-feature using machine learning via one-class SVM.

### 4.5.3 Tongue Carcinoma Tumour Volume Determination

The largest tumour diameter has been used for many years in the TNM classification to stage the tongue carcinoma. However, it cannot reflect the total tumour volume and the total burden of malignant cells. For instance, superficial spreading tongue carcinomas frequently exceed 4 cm in diameter without deep penetration. These tumours are classified as T3 lesions but have very low volumes. Recent findings showed the significance of tumour thickness (depth) and volume for the treatment planning and prognosis of tongue carcinoma [33]-[37].

Two-way ANOVA test on our results showed that there was no significant difference between manually traced and semi-automated segmented volume. Moreover, there was no significant variance between the volume results from the two operators. In this study of tongue carcinoma tumour volume determination, the percent matching (PM) and corresponding ratio (CR) were also used to quantitatively evaluate the accuracy of segmentation algorithm. These parameters showed that the segmented tumour areas were quite close to the manually traced areas. The averaged values of PM and CR in this study are higher than those of NPC tumour volume study. The inter-operator agreements of the manual tracing and unsupervised hierarchical method were evaluated at both the volume and pixel levels. The results are in good agreement with those obtained by the NPC tumour volume study: At the volume level, the inter-operator agreement of semi-automated segmentation was not significantly higher than that of the manual tracing method. At the pixel level, the inter-operator

variance of semi-automated segmentation was significantly lower than that of the manual tracing method. It should be noted that tumour volume is obtained by adding scan slices bearing tumour tissue. This summation process is likely to cancel out the over-estimated and under-estimated tumour bearing areas of the slices that make up the total tumour volume. This is the most likely reason for observing no significant differences between the two observers at the volume level. However, at the pixel level, direct comparison is made on the results of each scan slice. Hence there is no provision for the cancelling out of errors by the summation of slices. They indicate that the total tumour volume determination is more relevant to clinical practice than the comparison of location matching, however, the latter is very important to evaluate a computer-based segmentation method for tumour volume estimation efficiently and completely.

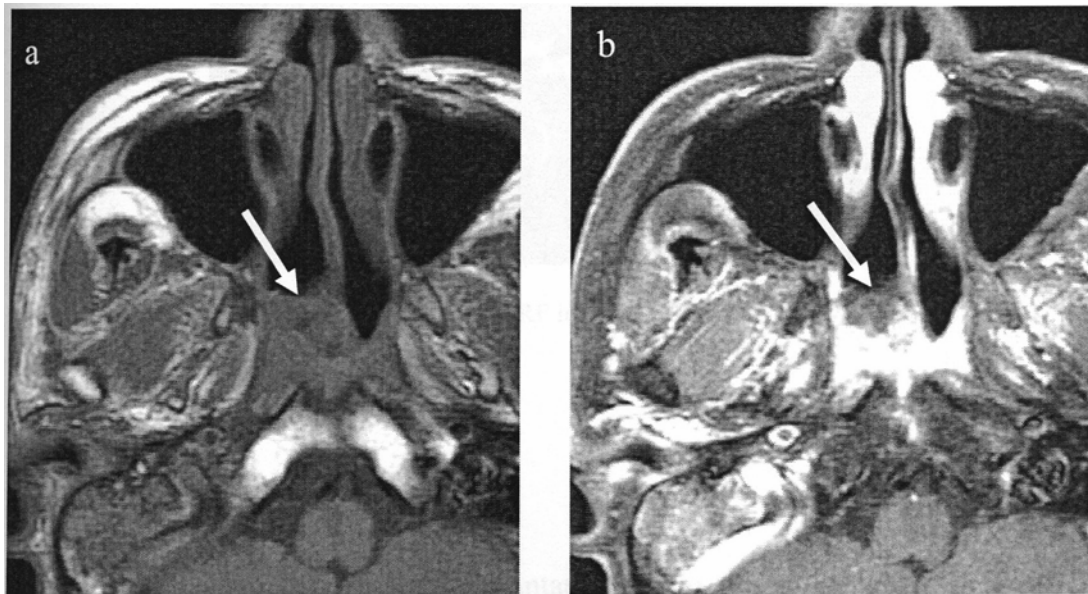
A major limitation of this study is the lack of a gold standard in comparing segmented tumour volume since this was a retrospective study and we did not have access to surgical specimens for the actual measurement of specimen volumes. Another limitation is the use of relatively large slice thickness and inter-slice gaps. The thinner slice without inter-slice gap will be used for further studies.

#### **4.5.4 Imaging Related Errors in Tumour Volume Determination**

Besides the errors existed in MR image post-processing and segmentation, there are also imaging related errors in tumour volume determination. Most of these errors come from imaging artefacts. Artefacts are those parts of an image that do not reproduce the actual tissue being imaged. They can vary from being relatively localised to a small part of an image to those that render the entire image useless. Sometimes artefacts are sufficiently subtle to make diagnosis difficult [63]. Artefacts can be caused by the patient or by extraneous effects.

### 1. Partial volume effect

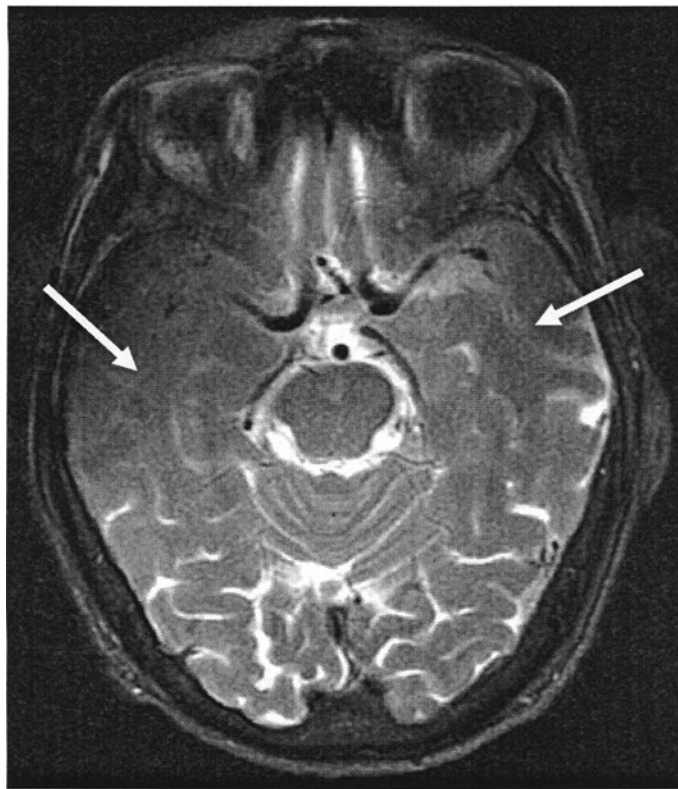
The clinical MR images are acquired as a stack of parallel slices in the axial, sagittal or coronal direction. In the presented study, each data set consists of 24–30 slices with the thickness of 5 mm and the pixel dimensions in each slice are about  $0.18 \text{ mm}^2$ . The images displayed represent the average of the free-induction decay (FID) response signals from all tissues in the 5mm slice. Thus, each voxel in the images could represent more than one tissue type. This phenomenon is referred to as the partial volume effect, or partial volume averaging. When a small structure is entirely contained within the slice thickness with other tissue of differing signal intensity then the resulting signal displayed on the image is a combination of these two intensities. This may cause the small structure to disappear and have large error in volume calculation [120]. The partial volume effect is particularly noticeable in the extreme slices of MRI volumes, i.e., the first and the last slices of a data set or the first and the last slices contain one organ (structure). Figure 4.16 shows the example of partial volume effect in NPC imaging.



**Figure 4.16** (a) T1W MR image shows the NPC morphologically in the nasopharynx (white arrow); (b) However, in the corresponding T1W post-contrast MR image, a part of the tumour does not show high signal intensity, due to partial volume effect (white arrow). This may cause errors in tumour segmentation.

## 2. RF inhomogeneity

Sometimes, MR images contain unwanted intensity variation due to inhomogeneity in the “uniform” magnetic field  $B_0$ , inhomogeneity in the applied RF pulse sequence  $B_{rf}$  and nonuniformity in RF field coil(s) used to measure the FID response signal. The overall intensity variation is referred to as RF inhomogeneity. For example, a metal object that prevents the RF field from passing into a tissue will cause a signal void in an image. Figure 4.17 shows a T2W MR image with RF inhomogeneity artefact.



**Figure 4.17** A T2W MR image with RF inhomogeneity artefact (white arrow)

## 3. Patient-related artefacts

People move both voluntarily and involuntarily during the course of scanning, hence motion is the most likely cause of image artefacts that can be contributed to the patient. Although patients are required to keep quite and breath-holding, head shifting,



coughing, swallowing and breathing are the common causes of artefacts for the scanning of brain, head and neck. Figure 4.18 shows an MR image with typical motion artefact. Sometimes the implants of orthopaedic surgery also produce false positive artefacts.



**Figure 4.18** A T2W MR image of tongue carcinoma with typical motion artefact

## 4.6 Concluding Remarks

In conclusion, an unsupervised hierarchical method that includes noise removal, initial segmentation, symmetry detection and refinement has been developed to segment tongue carcinoma from T1W and T2W MR images. After noise removal, an initial segmentation was performed using a GA-induced fuzzy clustering (GAIFC). Following this, an ANN-based symmetry detection procedure, which utilised anatomic and geometry knowledge, was applied to detect the initially extracted mass and then make necessary refinement to obtain the final segmentation results. Lesions were initially extracted from 66 MRI slices from 10 tongue carcinoma patients and 52

were detected as tumour slices using the developed method. The tumour volumes were obtained by summation-of-areas technique and quantitative validations were performed at both the volume and pixel levels, by comparing the computer segmented volumes with those from manual tracing method. Test results showed that at the volume level there was no significant difference between the manually traced volumes and the computer segmented volumes. In addition, at the volume level, there was no significant difference between the inter-operator agreements of the two methods. The matching comparison at the pixel level showed that the unsupervised hierarchical segmentation can not only well segment tongue carcinoma from MR images, but also have significant higher inter-operator agreement than the manual tracing method. Therefore, the developed unsupervised hierarchical method is found to be suitable to segment tongue carcinoma for tumour volume determination. Factors such as deformation are taken into account in the next chapter, titled “Deformation Model-based MRI Segmentation”.



## Chapter 5

# Deformation Model-based MRI Segmentation

## 5.1 Introduction

In Chapters 3 and 4, the semi-supervised and unsupervised MRI segmentation methods were presented. Both methods use classification/clustering-based techniques, which require at least two imaging sequences and explore the features derived from image contents. Besides pathological lesions such as tumour, lymph node, oedema, and inflammation, other irrelevant tissue types can be classified using these methods. However, in a lot of image segmentation cases, it is unnecessary to analyse every part of the image. The interest may be only in segmenting one or a few objects of the image. In addition, sometimes the area of interest only accounts for a small part of the whole image, therefore the clustering may not give meaningful results. One possible way in dealing with such problems is to use deformation models, by which the actual object edge and the boundary can be found and located in a local area by minimal interaction [121] [122].

As mentioned in Section 2.4.1 (page 25), the traditional edge-based methods try to find the places of rapid transition from one to other region of different brightness or

intensity by applying some gradient operators to the images. These methods do not perform well if the sampled data have typical shortcomings such as sampling artefacts, spatial aliasing and noise. For example, these methods only consider the local intensity information that makes incorrect assumptions during an integrated process and generate infeasible object boundaries. As a result, these model-free techniques usually require considerable amount of post-processing or expert interpretation. Therefore, the challenges are to extract the boundary elements belonging to the same structure or tissue type and integrate these elements into a coherent and consistent model of the structure of the tissue type [121].

In general, deformation models are physically motivated, model-based techniques for delineating region boundaries using closed parametric curves or surfaces that deform under the influence of internal and external forces [80]. The idea of using a model-based optimisation method to find the object boundary in images became popular after Kass, et al. presented the famous active contour model that solved the problem by minimising an energy function of the boundary [123]. Based on the basic idea of Kass, et al., deformable contour models with different energy functions have been proposed and processed by various optimising methods including the finite element method [124], dynamic programming [125] and greedy algorithm [126]. The above methods use image edge information, especially the local grey level gradient, in the energy function to attract the contour from an initially drawn plan to the real boundary. However, the high grey level gradient of the image may be caused by object boundaries as well as noise and object texture, and therefore the optimisation functions may have many local optima. In addition, these models require the initial plan to be “close enough” to the real boundary or the functions may not converge. A replacement is to use region information which utilises some statistic measures of region greyscale distribution to guide the region converge [122]. The region information can partly overcome the drawbacks of contour deformation because it

provides more constraints of shapes on the boundary and the region models can tolerate certain types of noises and textures.

Some pioneering investigations put forth the strategy of integrating both edge and region information into the deformation model for the segmentation of biomedical images and obtained meaningful results [66][122][127]-[130]. In this study, it is required to get the lymph node volume information, since the size of the node can be an indicator of tumour activity, and enlarged nodes are generally targeted for further study or treatment. In addition, some T2W images of tongue carcinoma were seriously contaminated by salt and pepper noise and motion artefacts resulting in the tumour not being well segmented using the fuzzy clustering method. Therefore, a boundary extraction method based on region and contour deformation models is presented in this chapter. Using this method, the T2W MR images of tongue carcinoma with severe salt and pepper noise and motion artefacts was successfully segmented, and a fast segmentation scheme was applied to a stack of 2-D MRI slices to segment the lymph node.

The organisation of this chapter is as follows. Section 5.2 describes the region and contour deformation models, respectively. Detailed implementation steps of this algorithm are presented in Section 5.3. Several applications such as the segmentation of NPC from CET1W images, the segmentation of tongue carcinoma from T2W MR images, and the fast segmentations of lymph node volume as well as intra-cranial aneurysm are shown in Section 5.4. Finally, discussion and concluding remarks are presented in Sections 5.5 and 5.6, respectively.

## 5.2 Region and Contour Deformation Models

### 5.2.1 Region Deformation Model

An object can be regarded as a region which is homogeneous in some properties. The property with some disparity between an object and its surrounding area should be chosen to identify the object. The boundary of an object can be considered as a sub-region of the object which borders other objects. The goal of region deformation is to find a region of maximum area whose boundary has the same properties as the region, based on an initial estimation of the object. Therefore, a property disparity between a region and its boundary, and the area of the region can be used to make the judgment. In this study, the greyscale statistical distribution is chosen as the discrimination parameter since both the tumour and the lymph node have a high contrast to the surrounding tissues.

#### 1. Greyscale distribution disparity of the region and its boundary

Let  $I$  be the pixel set of an image,  $i \in I$  be a pixel;  $g_i$  is the greyscale of pixel  $i$  with  $0 \leq g_i \leq 255$ . An object in the image can be described as a homogenous region  $O$ , and the boundary of  $O$  is a simple closed curve  $B$ . The Kolmogrov-Smirnov (KS) test is well known in nonparametric statistics and can be used to test if two sets of probability distribution belong to the same probability distribution [122][131]. In this study, KS test is used to test whether the boundary pixel set and the object pixel set have the same greyscale distribution.

First, the Kolmogrov-Smirnov distances  $D$  is defined as

$$D = \text{Max}|F_B(g_i) - F_O(g_i)|, \quad 0 \leq g_i \leq 255, \quad (5-1)$$

where  $F_B$  and  $F_O$  are the greyscale cumulative frequency distribution functions (CFDF) of boundary  $B$  and object  $O$ , respectively. Here,  $F_B$  and  $F_O$  can be calculated by

$$F_B(g_i) = \sum_{g_i^* \leq g_i} H_B(g_i^*), \quad (5-2)$$

and

$$F_O(g_i) = \sum_{g_i^* \leq g_i} H_O(g_i^*), \quad (5-3)$$

where  $H_B$  and  $H_O$  represent the greyscale frequency distribution functions (FDFs) of  $B$  and  $O$ , respectively and  $0 \leq g_i^* \leq 255$ . Then a lower case  $d$  which only depends on the pixel numbers of  $B$  and  $O$ , is defined as

$$d = c / \sqrt{\frac{AL}{(A+L)}}, \quad (5-4)$$

where  $c$  is the test significance level,  $A$  is the number of pixels which form  $O$ , and  $L$  is the number of pixels which form  $B$ . The hypothesis  $F_B = F_O$  is only accepted when  $D \leq d$ .

## 2. Two basic operations for region converge

After putting an initial closed contour and calculating the KS test, two basic morphological operations are used to deform the region for the final converge of the region to meet  $F_B = F_O$ , based on the different results of the KS test.

When  $D > d$ , the region plan covers a region somewhat different from the desired object, and a shrinking operation is used to deform the region. It means that the current contour will shrink inward for one pixel. In practice, the region boundary elements are deleted and their nearest inner neighbours are used as the new boundary elements. However, it is not necessary to delete all the boundary elements, but only

those which have a different property within the region. Figure 5.1 shows how to determine which boundary elements should be deleted. Here,  $O$  is the object,  $B$  is its boundary, and  $j$  is a pixel of  $B$ . A small window  $W_j$  is set with  $j$  being the center. If the greyscale distribution of  $W_j$  is close to that of  $O$ ,  $j$  is kept; otherwise it is deleted. In practice, the greyscale mean and standard deviation of  $W_j$  and object  $O$  are compared. The shrinking operation is performed till  $D < d$ .

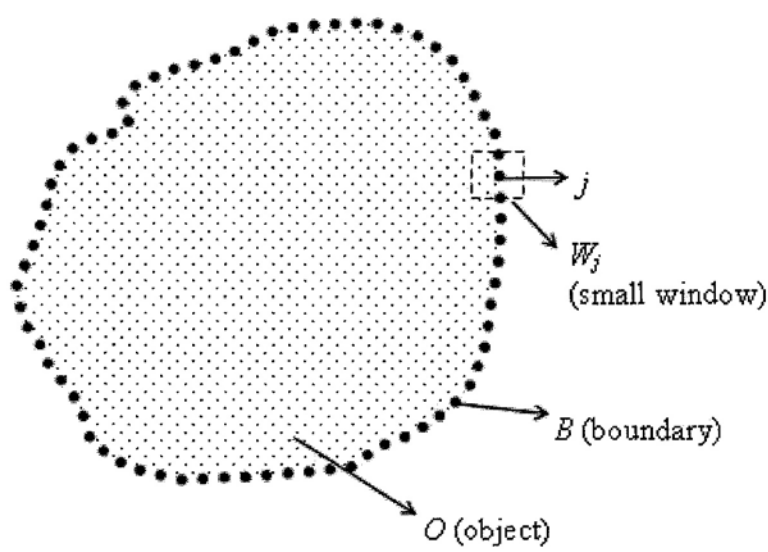


Figure 5.1 Object described as a region with a closed curve as its boundary

When  $D < d$ , the region plan may exactly cover the actual object or may be completely inside the object and not cover the whole object. Because of this kind of possibility, a growing operation is performed such that the current contour will grow outward for one pixel. In practice, the region boundary elements are assigned as the object pixels and their nearest outer neighbours were used as the new boundary elements. The growing operation is performed till  $D > d$ .

In practice, a coarse region is put totally inside or outside the actual object, or to partly cover the actual object, and the KS test is performed. There might be two results: (a) If initially  $D > d$ , the shrinking operation is performed till  $D < d$ . Then a cycle of growing operation is performed. After an iteration containing one growing

operation and one shrinking operation, the new region area is compared with the previous one. The process stops if the region area does not change. (b) If initially  $D < d$ , the growing operation is performed till  $D > d$  following the steps as described in (a).

### 5.2.2 Contour Deformation Model

The extracted boundary using region deformation is always in a zigzag form that it cannot well fit the actual tumour boundary which is fairly smooth. Therefore, a contour deformation model is used with this global geometric constrain to modify the contour extracted by region deformation [122].

#### 1. Contour deformation by Snake Method

The snake method [123] (active contour model) is an energy-minimisation spline technique which uses information on both the shape of the object of interest and the object's image properties to determine the segmentation of the object. In 2-D images, a snake is a sequentially connected set of pixels which form a closed curve with each of the pixels locates at one pixel in the image. The energy of the snake  $E_{snake}$  is guided by internal and external forces and can be defined by

$$E_{snake} = \int_0^1 [E_{int}(\mathbf{v}(s)) + E_{ext}(\mathbf{v}(s))] ds, \quad (5-5)$$

where  $\mathbf{v}(s) = (x(s), y(s))$  is the parametric equation of the contour, and  $s$  is the arc length;  $E_{int}$  is the internal energy, and  $E_{ext}$  is the external energy.

The internal energy  $E_{int}$  can be written as

$$E_{int} = \left( \alpha(s) |\mathbf{v}_s(s)|^2 + \beta(s) |\mathbf{v}_{ss}(s)|^2 \right) / 2, \quad (5-6)$$

where  $\mathbf{v}_s$  and  $\mathbf{v}_{ss}$  are the first and second derivatives of  $\mathbf{v}$ , which are used to control the smoothness of the contour; and  $\alpha$  and  $\beta$  are weights. The external energy  $E_{ext}$  is from the image edge information and can be written as

$$E_{ext} = -\gamma(s) |\nabla \mathbf{I}|^2, \quad (5-7)$$

where  $\nabla$  is the gradient operator,  $\mathbf{I}$  is the image, and  $\gamma$  is the weight. Since in actual images the transition of greyscale from background to object is not a sudden change, the second-order derivatives of  $\mathbf{I}$ ,  $\nabla^2 \mathbf{I}$ , was used to construct  $E_{ext}$ . Therefore Eq. (5-7) can be re-written as

$$E_{ext} = -\gamma(s) |\mathbf{G}_\sigma * \nabla^2 \mathbf{I}|^2, \quad (5-8)$$

where  $\mathbf{G}_\sigma$  is a Gaussian function with standard deviation  $\sigma$ , and “\*” is the convolution. In order to enhance the effect of zero-crossing places in the external energy distribution, the following equation was used instead:

$$E_{ext} = -\gamma(s) \log(1 + |\mathbf{G}_\sigma * \nabla^2 \mathbf{I}|). \quad (5-9)$$

In practice, the discrete forms of Eqs. (5-5), (5-6), and (5-9) can be written, respectively, as

$$E_{snake} = \sum_{i=1}^n [E_{int}(\mathbf{v}_i) + E_{ext}(\mathbf{v}_i)], \quad (5-10)$$

$$E_{int}(\mathbf{v}_i) = (\alpha(i) |\mathbf{v}_i - \mathbf{v}_{i-1}|^2 + \beta(i) |\mathbf{v}_{i+1} - 2\mathbf{v}_i + \mathbf{v}_{i-1}|^2) / 2, \quad (5-11)$$

and

$$E_{ext}(\mathbf{v}_i) = -\gamma(i) \log(1 + |\mathbf{G}_\sigma * \nabla^2 \mathbf{I}(\mathbf{v}_i)|), \quad (5-12)$$

where the contour  $\mathbf{v}$  is represented as a sequence of pixels  $(\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_i, \dots, \mathbf{v}_n)$  and  $\mathbf{v}_i = (x_i, y_i)$ . Substituting Eqs. (5-11) and (5-12) into Eq. (5-10), we obtain



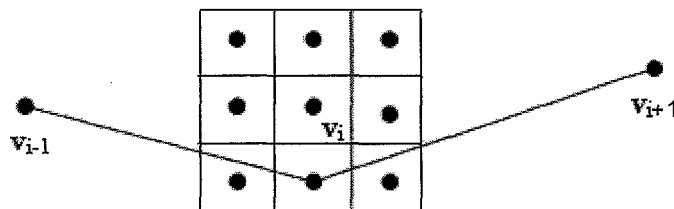
$$E_{snake} = \sum_{i=1}^n \left[ \left( \alpha(i) |\mathbf{v}_{i+1} - \mathbf{v}_i|^2 + \beta(i) |\mathbf{v}_{i+1} - 2\mathbf{v}_i + \mathbf{v}_{i-1}|^2 \right) / 2 \right] - \sum_{i=1}^n \gamma(i) \log(1 + |\mathbf{G}_\sigma * \nabla^2 \mathbf{I}(\mathbf{v}_i)|). \quad (5-13)$$

Therefore, minimising  $E_{snake}$  will smooth out a contour and move it to the nearby edge.

## 2. Minimisation of $E_{snake}$

A greedy algorithm proposed by William [126] is used to search the minimal energy contour. It is an iterative procedure that during each of the iteration, a neighbourhood of each contour pixel is examined and the pixel in the neighbourhood giving the smallest value of the energy term is chosen as the new location of the pixel.

Figure 5.2 shows how this algorithm works. The energy function is computed for the current location of  $\mathbf{v}_i$  and each of its neighbors. The location having the smallest value is chosen as the new position of  $\mathbf{v}_i$ . The location of  $\mathbf{v}_{i-1}$  has already been moved to its new position before the current iteration if necessary. Its location is used with that of each of the proposed locations for  $\mathbf{v}_i$  to compute the first order continuity term. The location of  $\mathbf{v}_{i+1}$  has not yet been moved. Its location, along with that of  $\mathbf{v}_{i-1}$ , is used to compute the second order constraint for each pixel in the neighbourhood of  $\mathbf{v}_i$ . For  $i = 0$ , only old values are used. For this reason,  $\mathbf{v}_0$  is processed twice.



**Figure 5.2** The iteration of greedy algorithm

In the prototype of greedy algorithm and the works of Chan, et al. [122] and Law, et al. [66][130], not all the pixels in the contour were examined for energy

minimisation. The initial contour for contour deformation is sampled to a point sequence with  $k$  pixels in intervals to use the fast snake model. After convergence, these points are connected sequentially to recover the actual contour. As a consequence, the recovered actual contour is not smooth enough. Hence, every pixel of the contour is examined in the present study.

### 5.3 Application: NPC and Tongue Carcinoma Segmentation

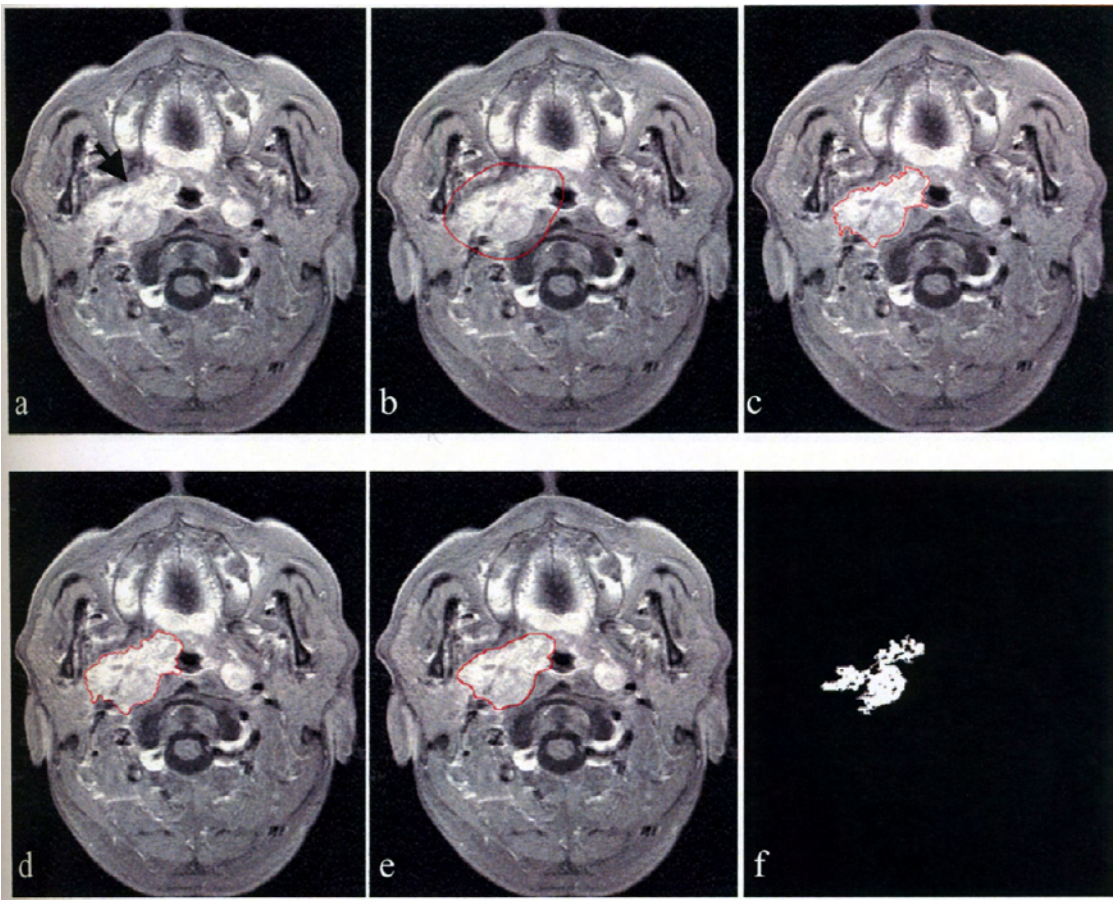
The scheme of segmentation of NPC and tongue carcinoma using deformation model is as follows. First a closed curve was manually drawn over the tumour region as the initial boundary. Then a region deformation was performed. A contour deformation was then applied using the result from region deformation as the initial contour. Since the process of region deformation can approximately locate the resulting contour to the actual boundary, the successive process of contour deformation can stop within a few cycles of iterations. In the implementation, the parameters  $\alpha(i)$ ,  $\beta(i)$ , and  $\gamma(i)$  were fixed at 1, 1, and 10 respectively [66]. This setting shows that in contour deformation, the contour location to the actual edge is more important than the smoothness of the contour. The value of the standard deviation  $\sigma$  of a Gaussian function was 4.

The test significance level  $c$  was determined by experiment. First, an experienced radiologist carefully drew the boundary of tumour from five MR images with NPC and tongue carcinoma. From the obtained closed contour and the included object, the corresponding KS distance  $D$  was calculated from Eq. (5-1). Let  $D = d$ , then the value of  $c$  was obtained from Eq. (5-4). From the five manually drawn closed contours and the included objects, five  $c$  values were obtained and they were nominated as  $c_1$ ,  $c_2$ ,  $c_3$ ,  $c_4$  and  $c_5$ . Let  $c_{\min} = \text{Int}[\text{Min}(c_1, c_2, \dots, c_5)]$  and

$c_{\max} = \text{Int}[\text{Max}(c_1, c_2, \dots, c_5)] + 1$ , hence an sequence of integers, “ $\mathbf{c}$ ”, was got:  $\mathbf{c} = (c_{\min}, c_{\min} + 1, \dots, c_{\max} - 1, c_{\max})$ , where  $\text{Int}$  is the integer operation,  $\text{Min}$  is the minimum value operation and  $\text{Max}$  is the maximum value operation. Then, every element in  $\mathbf{c}$  was used as the  $c$  value to perform the region deformation in the above mentioned five MR images, from the same initial plan. In every image, the corresponding ratio (CR) value was calculated from Eq. (3.19) by comparing the mass obtained from region deformation with the mass obtained from manual tracing at the pixel level, according to the different  $c$  value. The  $c$  value which achieved the highest CR value in an image was nominated as  $c^*$ . From the obtained five  $c^*$  values, the one with the highest occurrence was chosen. In the present study, the value of  $c$  was determined as 4, according to our experiment.

### 5.3.1 NPC Segmentation

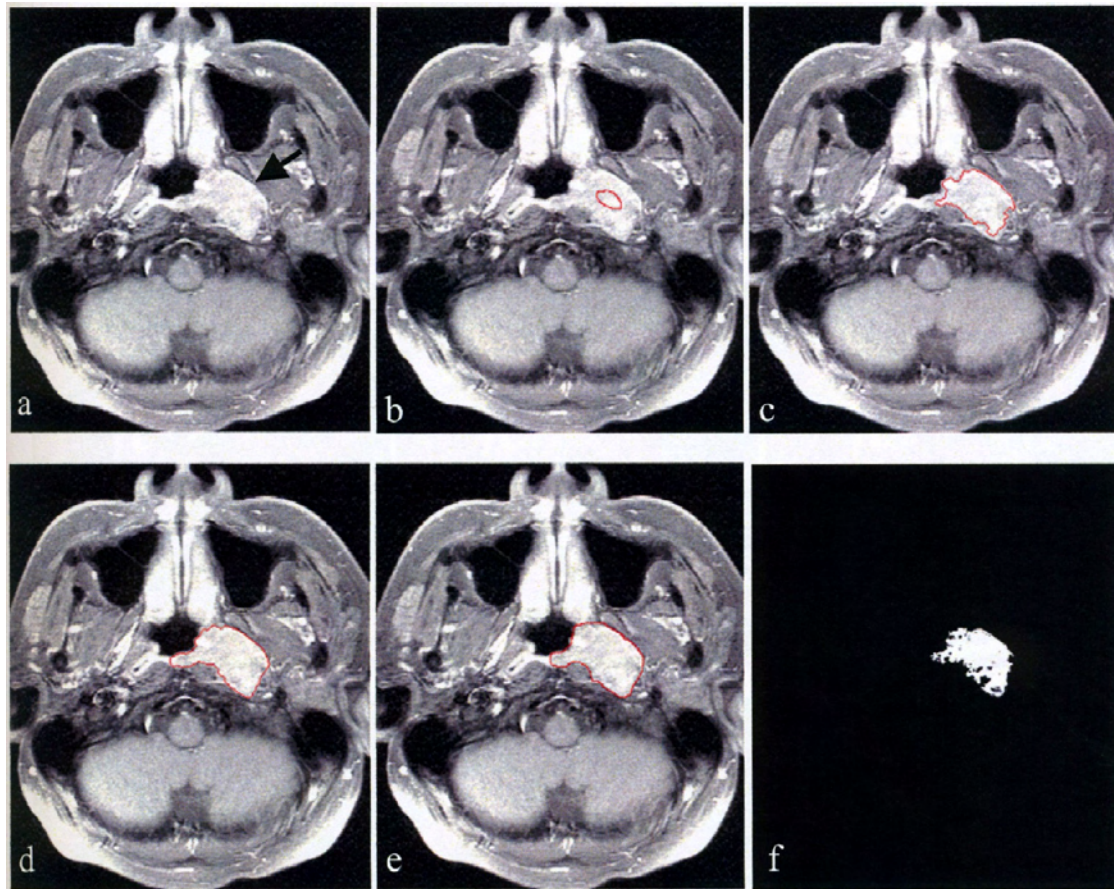
A contrast enhanced T1-weighted (CETIW) MR image of NPC is shown in Figure 5.3(a). The tumour enhancement can be seen on the right lateral section (arrow). However, the tumour boundary is quite blurry and the enhanced region is heterogeneous that even expert radiologists cannot trace the tumour boundary well. The initially drawn contour (totally outside), the result after region deformation, the result after successive contour deformation, the contour traced by a radiologist, and the result obtained using seeded region growing [67] are shown sequentially in Figures 5.3(b) to 5.3(f). Region deformation model can successfully find the actual tumour enhanced region, although its result is a little inside the actual tumour-enhanced region. After the successive contour deformation, the extracted contour fits the boundary of actual tumour enhanced region quite well. On the other hand, seeded region growing obtains poor result for this type of image with blurry boundary and heterogeneous tumour enhancement.



**Figure 5.3** (a) Original CET1W MR image; (b) Initial boundary; (c) Result after region deformation; (d) Result after contour deformation; (e) Boundary traced by radiologist; (f) Result using seeded region growing.

Figure 5.4(a) shows another CET1W MR image of NPC. The tumour enhancement can be seen on the left lateral section (arrow) and the boundary of enhanced region is more distinct than that of Figure 5.3(a). The initially drawn contour (totally inside), the result after region deformation, the result after successive contour deformation, the contour traced by a radiologist, and the result obtained using seeded region growing [67] are shown sequentially in Figures 5.4(b) to 5.4(f). They demonstrate that region and the successive contour deformation successfully find the actual tumour-enhanced region and extract the tumour boundary, whereas the seeded region growing algorithm is not so robust.



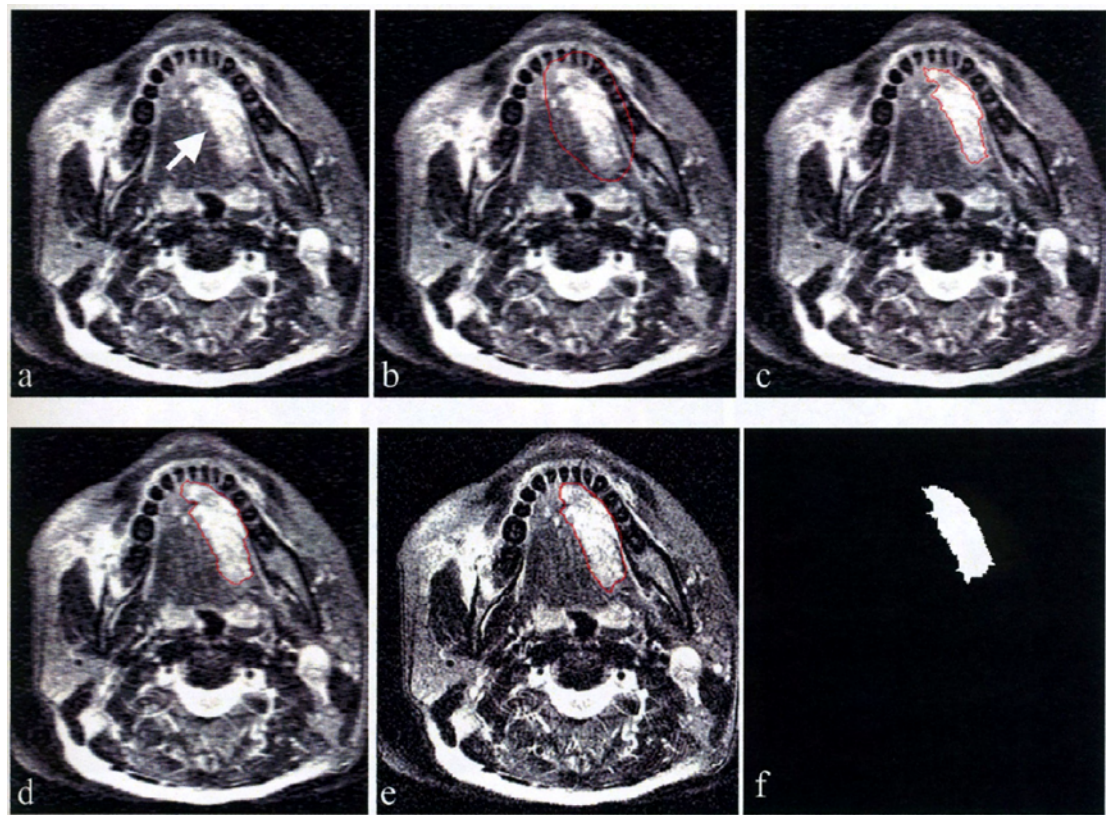


**Figure 5.4** (a) Original T1W MR image; (b) Initial boundary; (c) Result after region deformation; (d) Result after contour deformation; (e) Boundary traced by radiologist; (f) Result using seeded region growing.

### 5.3.2 Tongue Carcinoma Segmentation

A T2-weighted (T2W) MR image of tongue carcinoma with motion blur is shown in Figure 5.5(a). The tumour enhancement is on the left lateral section (arrow). The initially drawn contour (totally outside), the result after region deformation, the result after successive contour deformation, the contour traced by a radiologist on original image without motion blur, and the result obtained by the multispectral unsupervised hierarchical method (described in Chapter 4) using T1W and T2W images without motion blur are shown sequentially in Figures 5.5(b) to 5.5(f). These figures show that

both region plus contour deformation models and the unsupervised hierarchical method achieve fine segmentation results. However, the former utilised only one sequence of MR image with motion artefact, while the latter needed two sequences of MR images as well as post-processing such as mathematical morphological operations.

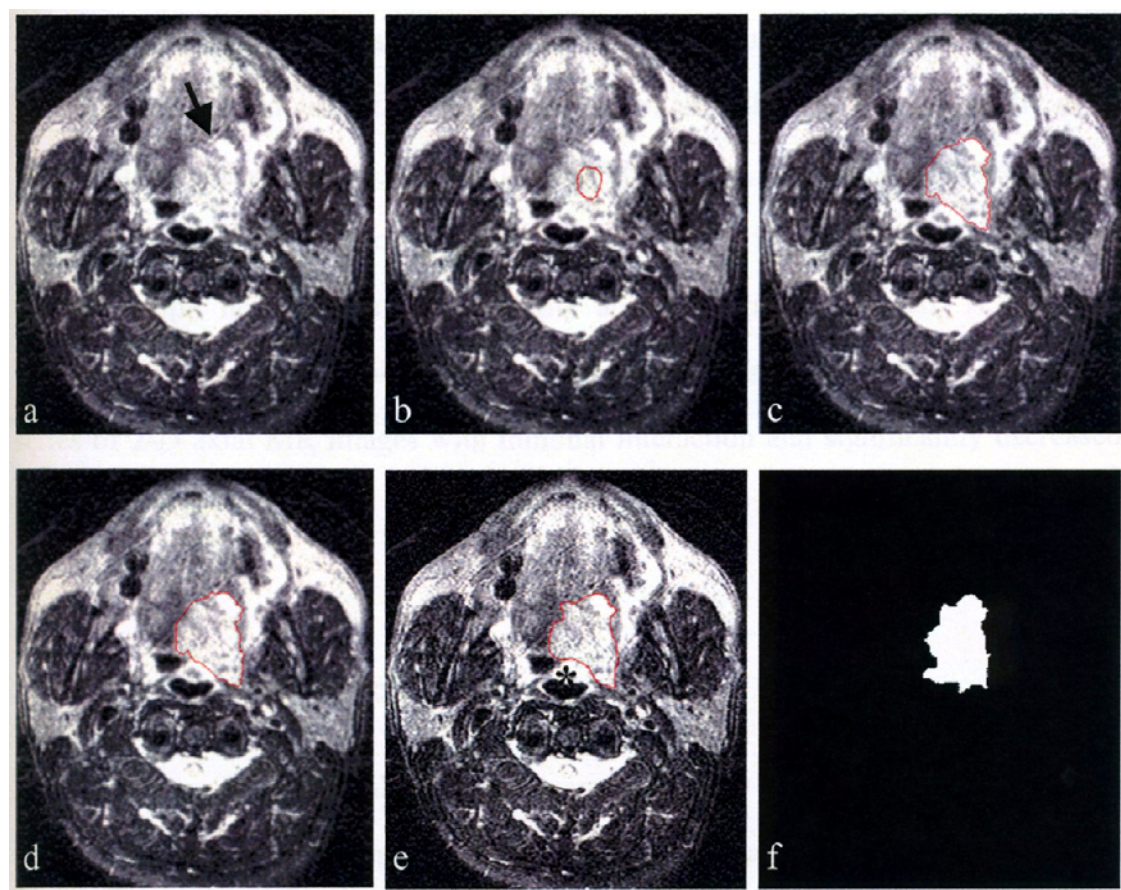


**Figure 5.5** (a) Original T2W MR image with motion blur; (b) Initial boundary; (c) Result after region deformation; (d) Result after contour deformation; (e) Boundary traced by radiologist; (f) Result using the unsupervised hierarchical segmentation method.

Another T2W MR image of tongue carcinoma with Gaussian blur is shown in Figure 5.6(a). The tumour enhancement is on the left lateral section (arrow) with very unclear tumour boundary. The initially drawn contour (totally inside), the result after region deformation, the result after successive contour deformation, the contour traced by a radiologist on original image without Gaussian blur, and the result obtained by the semi-supervised knowledge-based fuzzy clustering method (KBFC, described in Chapter 3) using T1W and T2W images without Gaussian blur are shown sequentially



in Figures 5.6(b) to 5.6(f). These figures show that by only exploring one sequence of image with Gaussian blur, the deformation-based method can find the actual tumour-enhanced region and extract the boundary which was quite close to the manually traced region. For the result from KBFC, although the tumour mass is almost segmented, a part of mucosa (black star in Figure 5.6(e)) which is enhanced and attached to tumour in the pharynx is also classified as tumour. Therefore, it is highly false positive. From these examples, it can be realized that the performance of deformation models on boundary detection from regions with blurry boundary and heterogeneous greyscales is fairly good.



**Figure 5.6** (a) Original T2W MR image with Gaussian blur; (b) Initial boundary; (c) Result after region deformation; (d) Result after contour deformation; (e) Boundary traced by radiologist; (f) Result using KBFC method.

## 5.4 Application: Fast 3-D Segmentation

The head and neck tumours frequently have accompanying swell of lymph nodes. According to the Tumour Node Metastasis (TNM) classification, the node volume is also an important diagnostic and prognostic indicator. A particularly important measure is the change in size of a node over time because this change can be used to determine treatment and evaluate patient response to a particular course of radiation- or chemotherapy. Therefore, during the treatment and the management of a patient with head and neck cancer, determination of lymph node volume is always required.

The lymph node is always of high contrast over the surrounding tissues in CET1 W and T2W images and is an isolated structure. Hence, deformation models can be used for accurate semi-automated segmentation. Although radiologists can manually draw the initial boundary for the extraction of actual boundary, it is really a tedious process to draw the initial contours on a stack of serial 2-D MR images, especially when the studied lymph node has a large volume and thin slices are used. A fast 3-D segmentation scheme is applied in the present study to segment lymph node from a series of 2-D axial MR images with minimal interaction and significantly decreased processing time.

### 5.4.1 Segmentation Scheme

A stack of parallel 2-D axial MR images produce a 3-D representation of the lymph node. If the inter-slice spacing is kept within a certain value, the successive 2-D MR images always have some similarities between each other in terms of size, shape, and greyscale distribution. Therefore the change in various properties of lymph node between one slice and its neighbouring one will be usually small.



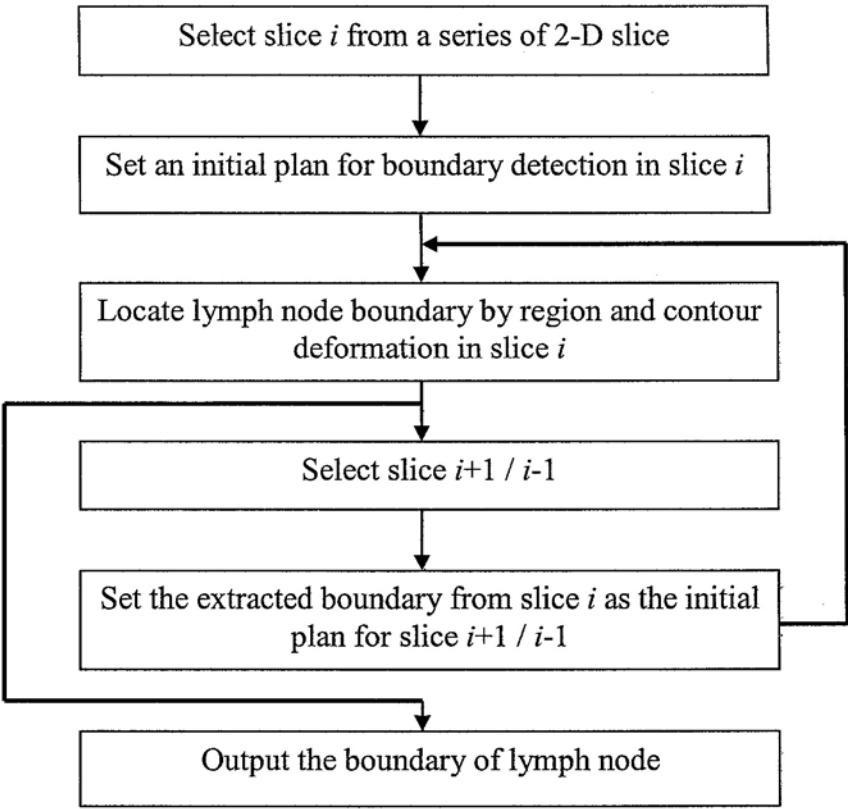


Figure 5.7 Block diagram of fast 3-D segmentation

The major steps of the proposed fast 3-D segmentation scheme are shown in Figure 5.7. Slice  $i$  is selected from a stack of  $n$  slices, where  $1 < i < n$ . After an initial plan was manually set on slice  $i$ , region and contour deformations were performed to extract the actual lymph node boundary. Then the extracted boundary was used as the initial plan on neighbouring slices  $i+1$  and  $i-1$ , and region and contour deformations were also performed on  $i+1/i-1$  to extract the actual boundary. This procedure was continued till all the slices were processed. By this method, the lymph node shape information between neighbouring MR images is utilised such that only one coarse manual initial plan is required for the whole stack of image. Moreover, the processing time is considerably decreased. An illustration of this scheme is also shown in Figure 5.8.

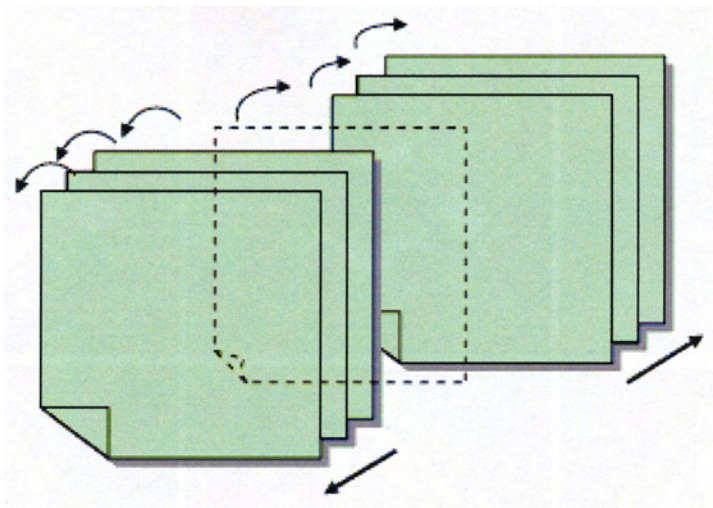


Figure 5.8 Illustration of applied fast 3-D segmentation scheme

5.4.2 Experiments and Results

1.3-D segmentation of lymph node

The MRI scanning of a tongue carcinoma patient was examined and the enhancement of the swollen posterior left neck node was found on T2W sequence from slice 12 to 17 (4 mm of slice thickness without gap). The initial plan was set on slice 15 and the 3-D segmentation was performed. The results are shown in Figures 5.9 and 5.10.

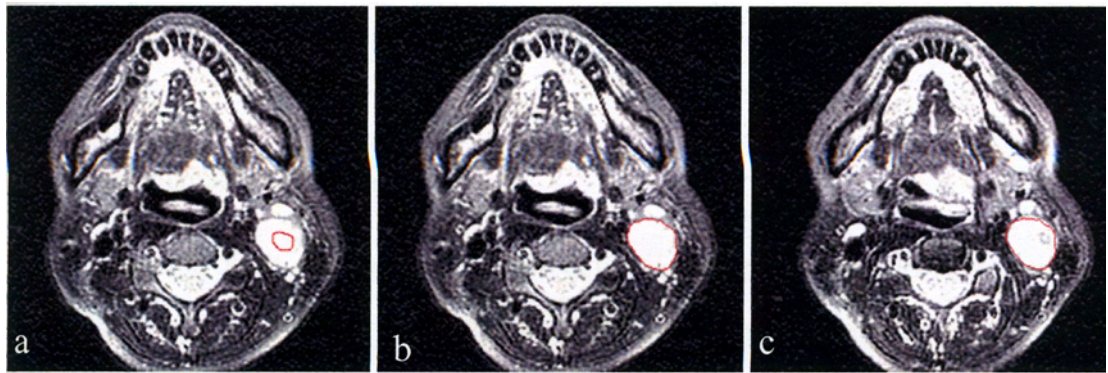
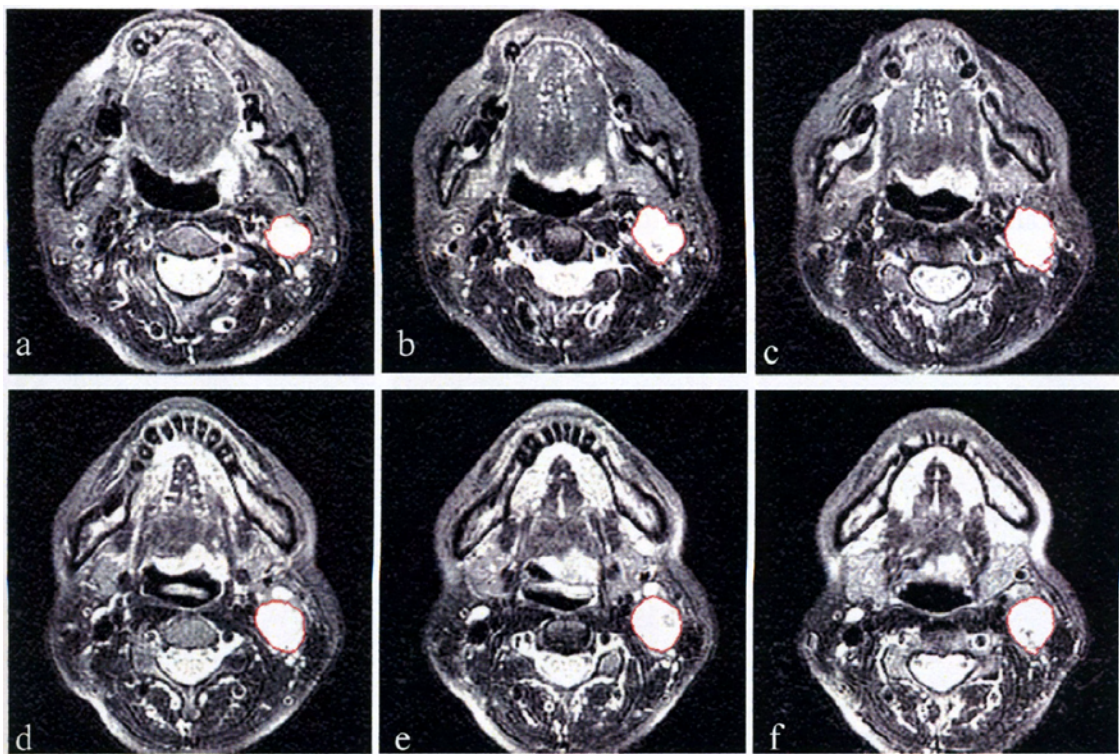


Figure 5.9 (a) Initial plan on slice 15; (b) Final boundary extracted using region and contour deformations on slice 15; (c) Initial plan on slice 16.



**Figure 5.10** Final results from slice 12 to slice 17 (a-f), respectively.

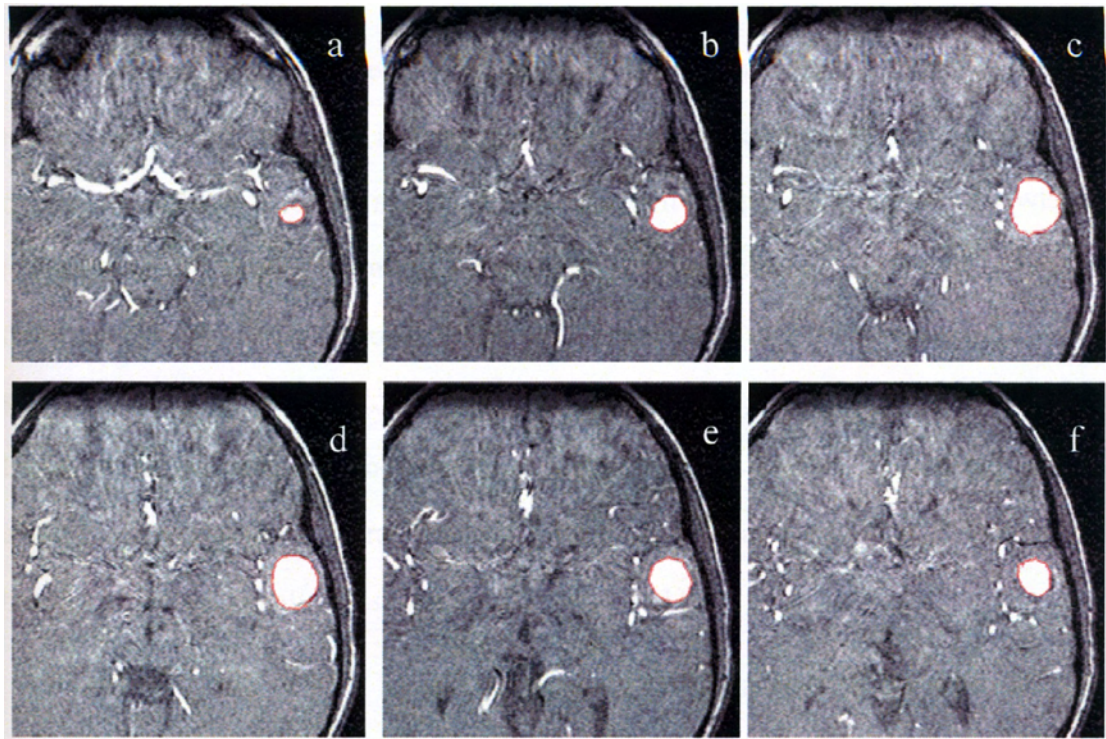
From Figure 5.9 it is observed that the initial plan on slice 16, which was the final boundary on slice 15, was quite close to the actual lymph node boundary on slice 16. Because of this, the boundary on slice 16 converged within a few cycles. The total processing time of the fast 3-D segmentation on this example was 29 s (including the time spent on initial plan drawing) while it took 113 s to complete the segmentation one slice by one slice (including the time spent on the drawing of six initial plans).

2.3-D segmentation of brain aneurysm

In this chapter, both region and contour deformations were used for the boundary extraction. Region deformation was used to keep the shape constraint while contour deformation was used to control the location of the curve to the actual edge and the smoothness of the curve. However, our experiments showed that if the object size is not large and thin slicing is used, good 3-D segmentation results can be achieved



using only region deformation. In this study, a time-of-flight (TOF) 3-D magnetic resonance angiography (MRA) volume data of brain aneurysm with 18 lesion-contained slices (1 mm slice thickness without gap) was segmented using the fast 3-D segmentation scheme with region deformation only. Segmentation results from selected six slices are shown in Figure 5.11.



**Figure 5.11** Selected segmentation results from six slices (a)-(f) Slice 1,4, 7, 10, 13, and 16 of 18 slices, respectively.

### 5.5 Discussion

Segmenting structures from medical images and reconstructing a compact geometric representation of these structures is not easy due to the sheer size of the dataset and the complexity and variability of the anatomic shapes of interest. Furthermore, the typical shortcomings of sampled data, such as sampling artefacts, spatial aliasing and noise, may cause the boundaries of structures to be indistinct and disconnected. The challenge is to extract boundary elements into a coherent and consistent model of the

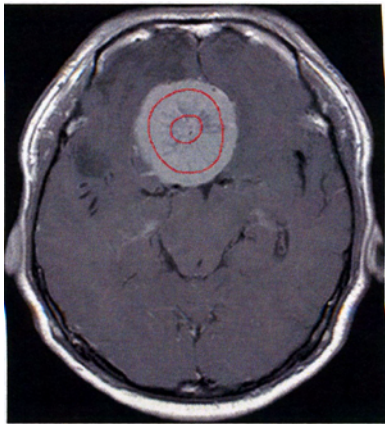
object. The traditional low-level image processing techniques which consider only local edge information can make incorrect assumptions during this integration process and generate infeasible object boundaries. As result, these model-free techniques usually require considerate amount of human intervention. Moreover, the subsequent analysis and interpretation of the segmented objects is hindered by pixel- or voxel-level structure representations generated by a lot of image processing operations.

The widely recognised potency of deformable models stemmed from their ability to segment, match and track images of anatomic structures by exploring (bottom-up) constraints derived from the image data together with (top-down) *a priori* knowledge (usually the initial input from an operator) about the locations, sizes and shapes of these structures. Deformation models are capable of accommodating the often significant variability of biological objects over time and across different individuals. Furthermore, deformation models support highly intuitive interaction mechanisms that allow medical experts and practitioners to bring their expertise to bear on the model-based image interpretation task when necessary [121].

Based on Kass, et al.'s basic idea, contour deformation models with different energy functions have been proposed and processed by various optimising methods. Most of these approaches utilised image edge information, typically the local grey level gradient, in the energy function to attract the contour to- the object boundary. However, the high grey level gradient of the image may be due to object boundaries as well as noise and object textures, and hence the optimisation functions may have many local optima. Consequently, such models require the initial plan to be "close enough" to the real boundary. One possible way is to integrate region and contour deformations to locate the boundary of an object. Because region information provides more constraints on the boundary of the object, the region model can tolerate certain noises and textures. The two-step method presented in this study gives an effective alternative to minimise the manual work. The region deformation can

provide a good enough initial contour plan for the next contour deformation because the region model presented is a constrained optimisation problem, in which maximising the region area is the objective with region and its boundary having the same greyscale distribution as the constraint. Some techniques also used the statistical information inside the object in their region models, however they assumed that the interior of the object follows a normal distribution and the parameters are kept unchanged during the deformation procedure [132] [133]. In our region model, there is no assumption on greyscale distribution of the object since we only tested whether the boundary and the whole object have the same greyscale distribution. In addition the greyscale distribution of the object keeps changing during the deformation process. Therefore this model is more applicable to various types of images.

Only a part of the image is analysed using the presented deformation model-based method, hence the processing speed of this method is 10 seconds or so (CPU, Intel Pentium-4 3.06 GHz; Memory, 512 MB), which is faster than the clustering-based methods described in Chapters 3 and 4. The factor that affects the speed of the algorithm is the difference of included area between the initial contour and the finally calculated contour. Figure 5.12 shows a CET1W MR image of brain tumour with two different initial plans. The times for tumour extraction were 7 seconds (started from the outer contour) and 12 seconds (started from the inner contour), respectively.



**Figure 5.12** A CET1W MR image of brain tumour with two different initial plans

The major drawback of the region deformation may be the operator dependency because the manual input of initial contour plan is needed to guide the deformation. A quantitative study on the inter-operator variance of the segmentation using region deformation is expected in the further investigation. However, the manual input of initial contour is the way how the medical experts contribute their prior knowledge to the abnormality extraction. It is based on the knowledge constraints in the image that the region deformation can achieve meaningful and flexible results, not affected by the local intensity shifts and the artefacts. For an instance, in some images the normal and abnormal objects have similar signal properties and are adjacent to or in contact with each other. The operator can draw the initial contour through the expected boundary between the normal and abnormal objects and start the region deformation. Then the contour will converge to the boundary of the desired abnormal object without spreading to the undesired object.

## 5.6 Concluding Remarks

In this chapter a two-step deformation model-based method has been used to locate the boundary of an object. The objective of this method is to segment not a whole image but some individual objects from an image. An initial plan of this boundary is required as in other deformation models. In the first step of region deformation, the KS test was used to detect if the boundary of a region and the enclosed region have the same greyscale statistical distribution. An iterative shrinking-growing operation was used by controlling the result of KS test to deform the rough initial contour to locate the boundary of an object. In the second step, the active contour model was used to refine the result from the first step according to the constraint of the minimisation of snake energy. Overall, the region deformation keeps the local shape constraint of an object from statistics. It depends only on the interior of the object and



makes no assumption about the greyscale distribution of the region. On the other hand, the contour deformation controls the location of the curve to the actual edge and the smoothness of the curve.

The experimental results show that the presented deformation model-based method accurately extracted the boundaries from those objects with very unclear object-background interface, inhomogeneous region contents, and motion as well as Gaussian blur. Based on the deformation models, a fast 3-D segmentation scheme, which uses the extracted object boundary as the initial plan for the next/previous slice, was applied to segment the boundaries of lymph node and intracranial aneurysm from a stack of 2-D slices.

In conclusion, although the deformation model-based method needs a certain degree of interaction, it is a suitable solution to segment a single medical structure even with unclear object boundary, inhomogeneous region contents and artefacts. The derived fast 3-D segmentation scheme can segment the single object volume from a stack of 2-D slices within a reasonable processing time.

## Chapter 6

# 3-D Visualisation for Head and Neck Tumours

### 6.1 Introduction

Visualisation is defined as the act or process of interpreting in visual terms or putting into a visual form. Schroeder, et al. gave a simple but clear concept defining visualisation as the transformation of data or information into pictures, engaging the primary human sensory apparatus, *vision*, as well as the processing power of the human mind [85].

The visual interpretation of traditional 2-D CT and MRI scans are usually sufficient for the diagnosis of head and neck tumours. However, physicians, surgeons, or radiologists have to mentally reconstruct a series of 2-D images into a 3-D picture. Head and neck tumours reside in an area of intricate anatomy. Hence, complex pathologic entities can sometimes be difficult to delineate. Moreover, modern imaging techniques can produce very thin slices from several imaging sequences (even isotropic 3-D imaging). It can be quite troublesome and time-consuming to browse the large volume data by 2-D slice format. Therefore, the important role of 3-D visualisation in medicine is to facilitate communication of anatomic and physiological

information to clinicians in a more natural and direct way of data exploration [134]. In addition, for efficient treatment planning procedures, 3-D visualisation becomes very useful [20]. Some therapeutic modalities, such as conformal radiation therapy and navigation-based neuro-surgery, require accurate knowledge of tumour volume and spatial location for maximum efficacy. A 3-D model of a patient's anatomy may help in such cases and can serve as a method for monitoring tumour size.

A short introduction of the two classes of fundamental visualisation techniques, surface rendering and volume rendering, has been given in Section 2.5.1. Currently, commercial solutions for head and neck tumour volume visualization packages are not available. Hence in the present study, research effort was undertaken to develop a novel visualisation packet, which can be used to show the 3-D view of segmented tumour and raw volume data (with the combination of segmentation results) with a visualisation software development toolkit.

Currently, there are several widely used visualisation software development toolkits with advanced design and rich functions including A Visualisation Workbench (AVW) Library developed by Mayo Clinic, Interactive Data Language (IDL) developed by Research System Inc., Matlab developed by MathsWorks Inc., the Visualization Toolkit (VTK) developed by Kitware Inc., and so on. AVW Library, IDL, and Matlab are all commercial visualisation application development kits which are not open-source for user's improvement, extension, and the integration of visualisation hardware acceleration system. On the contrary, VTK is an open-source and free visualisation development kit. Besides that the user can make improvement and extension based on its source code, the object-oriented design and implementation of VTK makes it convenient for the integration of visualisation hardware acceleration system. Hence, VTK was selected to facilitate the development of a 3-D medical visualisation scheme and platform in the present study.

The organisation of this chapter is as follows. Section 6.2 describes the system analysis of VTK. The design of a visualisation platform Voxur-3D, including the data format supported by VTK and the modules of the developed visualisation platform, is presented in Section 6.3. Section 6.4 presents the visualisation of segmented tumours using surface rendering. Since the slice thickness is quite large for a lot of data, an interpolation procedure which is used to reconstruct the tumour volume is also included in this section. The visualisation of whole volume data using volume rendering is given in Section 6.5. In this section, the different effects of volume renderings with and without the combinations of segmentation information are compared. The last two sections, Sections 6.6 and 6.7 are the discussion and the concluding remarks, respectively.

## 6.2 System Analysis of VTK

### 6.2.1 Overview

VTk is neither an executive program nor a fast graphic engine, but an object-oriented software development system for 3-D computer graphics, image processing, and visualisation [135]. It consists of a C++ class library using dynamic linking and device-independent techniques, and an encapsulation layer formed by a few interpreted languages such as Tcl/Tk, Java, and Python. A few main characteristics of VTK are described below [85].

- Open source code, which is convenient for the user's improvement and extension for the toolkit itself;
- Provide efficient interaction functions as well as several advanced and effective visualisation algorithms, such as matching cubes and dividing cubes for surface rendering, and ray casting and texture mapping for volume rendering;

- Object-oriented system design and implementation;
- Platform (Windows, Linux, or Unix) and low layer graphic library (OpenGL, Mesa) independent;
- The construction of system prototypes and even applications using interpretive languages attached such as Tcl/Tk, Java, and Python.

Relevant technical details of VTK [85][135]-[137] were made use of in the development of the proposed 3-D tumour visualisation scheme.

### 6.2.2 System Architecture of VTK

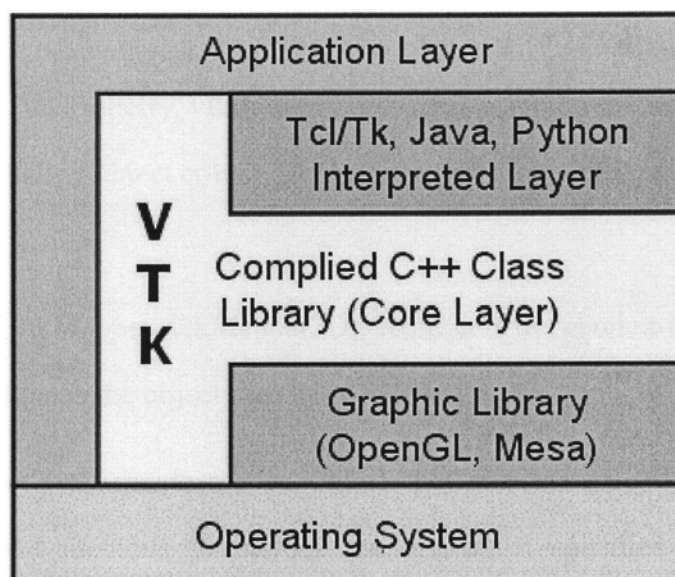
VTK consists of two major components: a compiled core (implemented in C++) and an automatically generated interpreted layer which currently supports Tcl/Tk, Java, and Python [136].

The data structure, algorithms, and time-critical system functions of VTK are implemented in the C++ core. Common design pattern such as object factories and virtual functions ensure portability and extensibility. Since VTK is independent of any graphical user interface (GUI), it does not depend on the windowing system. Hooks into the window ID and event loop let developers plug VTK into their own applications.

While the compiled C++ core provides speed and efficiency, the interpreted layer offers flexibility and extensibility. For example, professional applications can be built rapidly and efficiently using GUI prototyping tools such as Tcl/Tk, Python/Tk, or Java.

The system frame of VTK-based applications is shown in Figure 6.1. The VTK formed by the core layer and the interpreted layer is independent from the GUI (GUI

provides the encapsulations for different applications by abstract graphic model) and provides the application layer with visualised programming interface. In addition, the application layer has the abilities of not only operating the C++ class library via the interpreted layer and direct calling the programming interface of the C++ class library, but also dealing with the operating system in the bottom level. This architecture of VTK gives the users convenience and flexibility to the construction of VTK-based applications.



**Figure 6.1** System frame of VTK-based applications

### 6.2.3 Object Models of VTK

There are two distinct parts to the object design of VTK. The first is the graphic model which is an abstract model for 3-D graphics. The second is the visualisation model which is a data-flow of the visualisation process [137].

#### 1. The graphics model

The graphics model captures the essential features of a 3-D graphic system in a form that is easy to use. The abstraction is based on the movie-making industry with some

influence from GUI windowing systems. There are ten basic objects in the model and their relationships are shown in Figure 6.2 [136][137].

- Actor (`vtkActor`, `vtkActor2D`, `vtkVolume`): an object drawn by a renderer in a scene.
- Light (`vtkLight`): illuminates the actors in a scene.
- Camera (`vtkCamera`): defines the viewing position, focal point, and other camera characteristics.
- Property (`vtkProperty`, `vtkProperty2D`): represents the rendered attributes of an actor including object colour, lighting, texture map, drawing style, and shading style.
- Mapper (`vtkMapper`, `vtkMapper2D`): represents the geometric definitions of an actor and maps the object through a lookup table.
- Transform (`vtkTransform`): an object that consists of a 4x4 transformation matrix and methods to modify the matrix. It specifies the position and orientation of actors, cameras, and lights.
- Lookup Table (`vtkLookupTable`, `vtkColourTransformation`): implements the mapping from actor data value to real colour. It is one of the most important techniques of visualisation.
- Renderer (`vtkRenderer`): coordinates the rendering of light, cameras, and actors.
- Render Window (`vtkRenderWindow`): manages a window on display device. One or more renderers draw into a render window to generate a scene.
- Render Window Interactor (`vtkRenderWindowInteractor`): manages the interaction with the actor, such as manipulating the camera, invoking user-



defined methods, picking the objects, and changing some of the properties of actors.

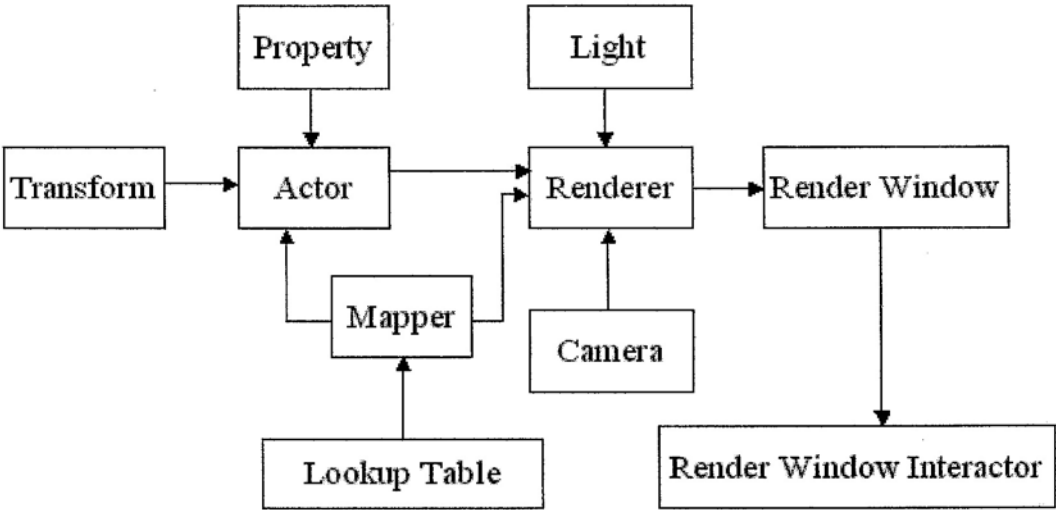


Figure 6.2 The relationship of graphical models in VTK

2. The visualisation model

The role of the graphics model is to transform graphical data pictures, while the role of the visualisation model is to transform various types of information into graphical data by means of data flow, and then put them into graphics model for rendering. The visualisation model consists of two basic types of objects: data object and process object. The data object performs the encapsulation for various data types. The data types supported by VTK include scalar (single data value), vector (3-D direction and magnitude), normal (3-D direction), texture coordinate (n-dimensional index into texture map), tensor (an nxn matrix), and field data (an array of arrays, each array can be of different data type). The process object is the data processing module in the whole visualisation process. According to the data input and output, the process object can be classified into three types: source (no data input and one or more outputs), filter (one or more data inputs and outputs), and mapper (one or more data inputs and no output).

6.2.4 Pipeline of VTK

The data pipeline of VTK is shown in Figure 6.3. The generation of data is from the start point of the pipeline to the end point. However, the requirement of data update is from the end point of the pipeline to the start point. There will be a requirement of data update when the rendering for an actor begins. This requirement is transferred following the reverse direction of data generation. The requirement transfer stops when the data in a node of the pipeline is the latest data and then from this node, the new data is generated step by step following the original direction of data generation. This procedure makes the guarantee that the rendered data are the latest according to the adjustment of the rendering conditions. By this inverse update algorithm, the update of data can commence from the node which is affected by exterior conditions instead of the data source. This processing significantly decreases the redundant computation and improves the real time performance for 3-D rendering and interaction.

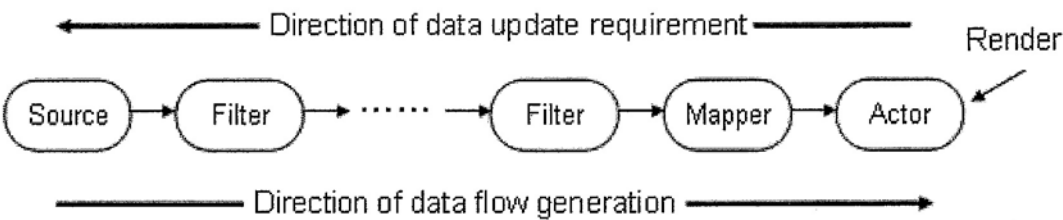


Figure 6.3 Data pipeline of VTK

6.3 Design of the Visualisation Platform

6.3.1 Overview

In this study, a PC and VTK-based 3-D visualisation platform, Voxur-3D, was developed to visualise both segmented tumour volumes and entire 3-D medical

volume data. Marching cubes method [138], an algorithm of surface rendering, was integrated into the platform for the rendering of the segmented tumour volume. Ray casting method [139], an algorithm of volume rendering, was integrated for the rendering of the entire 3-D medical volume data. Three main modules, i.e., data management, image display and rendering, and image control are included in this platform.

The development environments of Voxur-3D are as follows. Hardware configuration: CPU, Intel Pentium IV 3.06GHz; RAM, 512MB; Hard disk, 80GB; Monitor, 15-inch LCD (1024x768) or 17-inch CRT (1280x1024); Display adapter, MicroStar GeForce4 Ti 4200 with video memory of 128M. Operating system: MS-Windows XP. Development tools: MS-Visual C++ 6.0 and VTK 4.0.

### 6.3.2 Module Descriptions

#### 1. Data management module

The main function of the data management module is to read a stack of image files from hard disks or optical disks and then transform these files to a single volume data file for further processing. The image file types supported are as follows: BMP (bitmap) file, TIFF (tag image file format) file, and DICOM (digital imaging and communications in medicine) file. Both BMP and TIFF files are common 8-bit greyscale image formats while DICOM file are 16-bit image format specially designed for medical images only. The output volume data file for further processing is ANALYZE 7.5 format. ANALYZE file format developed by Mayo Clinic supports both single image and volume data and has been widely used in clinical and medical research fields. Hence, it is regarded as the standard volume data format for medical objects. A single ANALYZE formatted data consists of two files: a data file and a

header file. The couple of files have the same name and are distinguished by the extensions .img for data file and .hdr for the header file [145].

A brief introduction of ANALYZE file format for medical images is given in the appendix.

The VTK also provides its own file format. It supports not only image data but non-image data including the physiological signals (such as ECG, EEG, etc.), results of image processing (such as segmented contour, isosurface, etc.), external signals (such as the position of electrode, etc.), and so on [136]. However the VTK file format is not supported by many other tools. It was not included in this platform.

## 2. Image display and rendering

This module includes the display of 2-D images and the rendering of 3-D data. The rendering function of 3-D data is the most important part in this platform and the detailed implementations are presented in Section 6.3.4.

## 3. Image control module

The image control module includes the controls for 2-D and 3-D images. The operations for 2-D images include zoom out, zoom in, transition, and the adjustment of display window and level. Compared with 2-D images, more control functions are provided for 3-D images. In addition to zoom out, zoom in, transition, rotation, and cropping operations, the adjustments for colour/intensity opacity, gradient opacity, intensity-colour mapping, and shading methods are also included. These adjustments are important to achieve better rendering effects, especially for the volume rendering of 16-bit medical data.

6.3.3 System Architecture

The system architecture of the visualisation platform developed in the present study is shown in Figure 6.4.

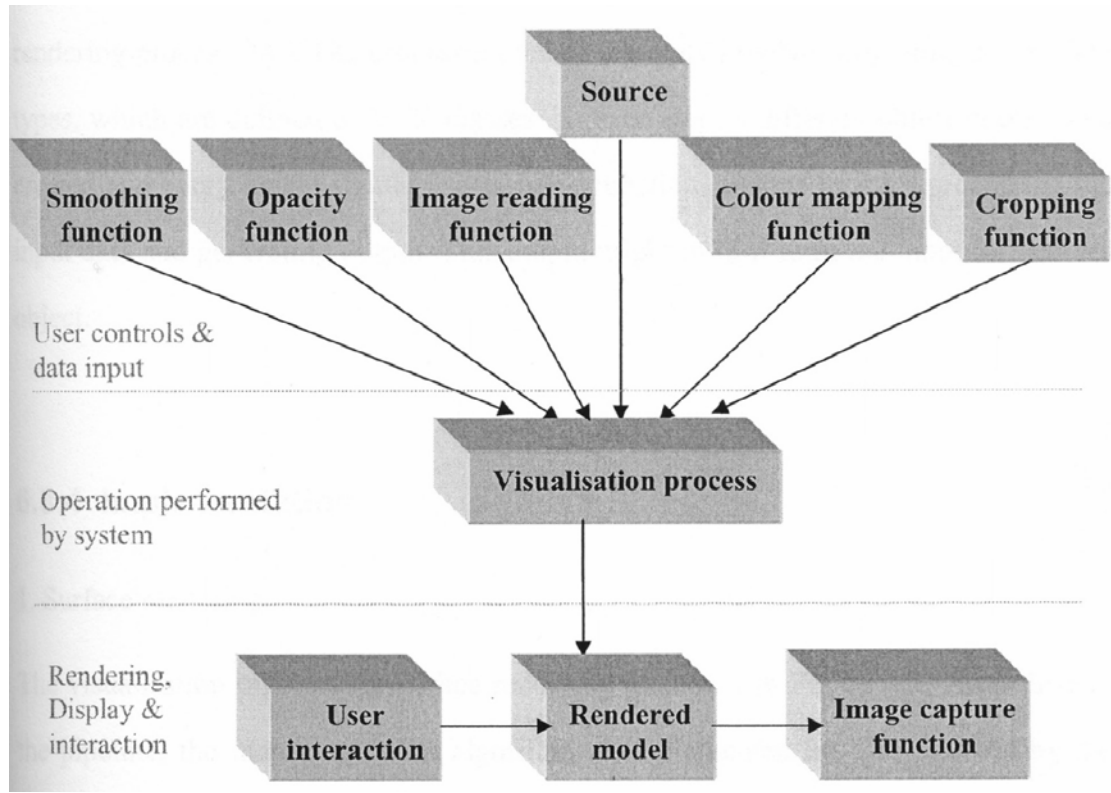


Figure 6.4 System architecture of the visualisation platform

The process of visualisation from reading the input to displaying the rendered image can be broken down into five main steps. They are:

- a. Reading input images to generate a single volume data;
- b. Smoothing the volume data with a Gaussian kernel;
- c. Creating and initializing a volume mapper (with colour and opacity functions);
- d. Defining cropping attributes; and,
- e. Rendering (volume or surface).

Following these, an image capturing function which has been made optional to the user may be called. In addition, if the user uses a mouse for interactions such as zoom in, zoom out, transition and rotation, a user interaction function will be called to transfer updated rendering parameters to the rendering function and trigger a new rendering process. In VTK, processing the data deeply involves migrating across data types, which are defined as VTK classes. At each step, a different object needs to be created and every object should use member functions unique to it for processing the input data and generating output. This output might in turn form the input for another object.

### 6.3.4 Implementation

#### 1. Surface rendering

The visualisation pipeline of surface rendering is shown in Figure 6.5. According to the pipeline, the marching cubes algorithm is implemented by VTK classes by the detailed code embedded within their member functions.

The surface rendering begins with the extraction of multi-contours, followed by the creation of a `vtkSurface` object. The `vtkSurface` object has to be assigned by a mapper and some properties before it can be used for rendering. Since marching cubes is the algorithm for surface rendering, a `vtkSurfaceReconstructionMapper` object is created which takes the `vtkImageLuminance` object (the 3-D matrix of intensity values obtained from the extracted 3-D isosurfaces) as the input. An instance of `vtkVoxelContoursToSurfaceFilter` is also created and assigned to the mapper. The `vtkSurfaceProperty` object is initialised with the colour and shading functions. The above mentioned objects are set as the input for the `vtkSurface` instance, which in turn is added to the render window. The final output displays the 3-D surface using the

marching cubes function, with the appropriate shading parameters and colours defined in the functions, for each scalar value in the 3-D isosurfaces.

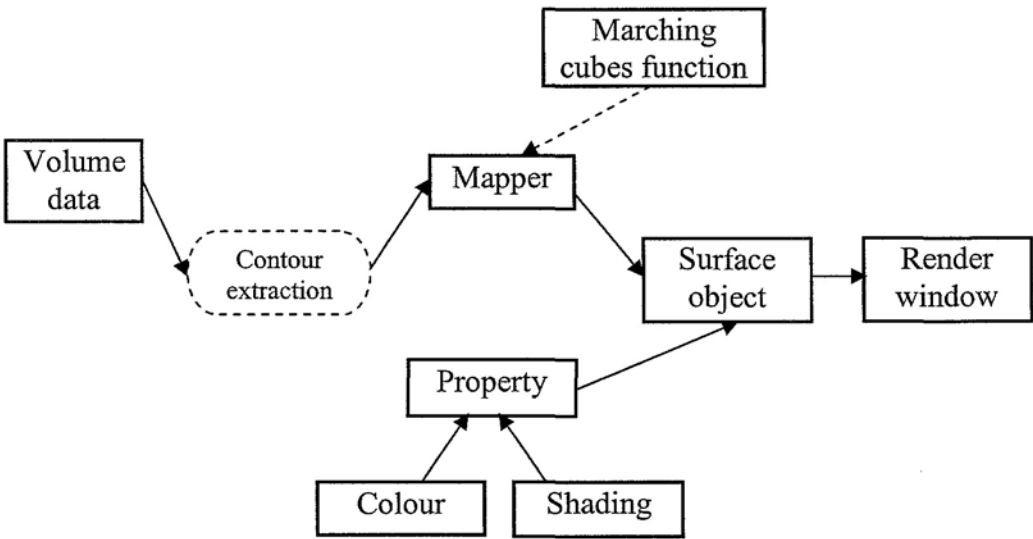


Figure 6.5 Clinical surface rendering pipeline

2. Volume rendering

The visualisation pipeline of volume rendering is shown in Figure 6.6. The implementation of volume rendering is different from that of surface rendering. Without contour extraction, the visualisation process begins with the direct creation of a `vtkVolume` object which is assigned by a mapper and some properties before it can be used for rendering. Since ray casting is the chosen algorithm for volume rendering, a `vtkVolumeRayCastMapper` object is created which takes the `vtkImageLuminance` object (the 3-D matrix of intensity values obtained from reading the volume data) as the input. An instance of `vtkVolumeRayCastCompositeFunction` is also created and assigned to the `vtkVolumeObject`. The `vtkVolumeProperty` object is initialised with the colour and opacity functions. These objects are then set as the input for the `vtkVolume` instance, which in turn is added to the render window. The final output



displays the volume data set using the ray casting function with the appropriate opacities and colours defined in the functions, for each scalar value.

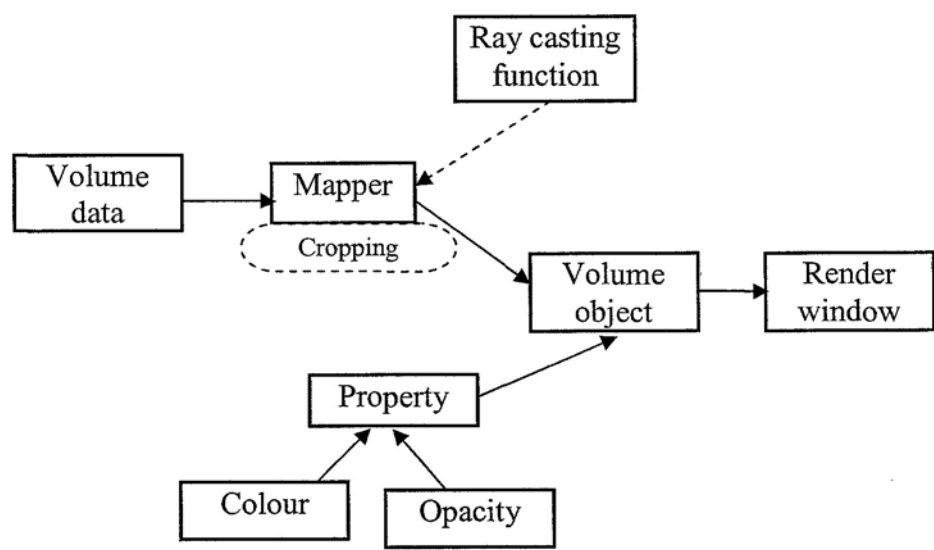


Figure 6.6 Clinical volume rendering pipeline

## 6.4 Visualisation of Segmented Tumours

### 6.4.1 Inter-slice Interpolation

Clinically, the inter-slice spacing of MRI images is much larger than intra-slice sampling resolution. Therefore, in order to obtain the accurate volume estimation and smooth 3-D reconstruction result with high quality, an interpolation procedure as shown in Figure 6.7 is always needed to get higher inter-slice resolution and decrease the uncertainty between slices. The techniques of inter-slice interpolation are divided into two categories: statistics-based greyscale interpolation and shape-based interpolation. Compared to statistics-based methods, shape-based interpolation methods can provide higher interpolation quality for fine detail and edge preservation with the high computation cost for complex geometry relationship, dynamic elastic and morphology model [140]-[143]. In the present study, a shape-based interpolation

method using a novel distance transform and morphing was applied with the supposition that under a certain resolution, the shape of objects within two neighbouring slices should change smoothly [144].

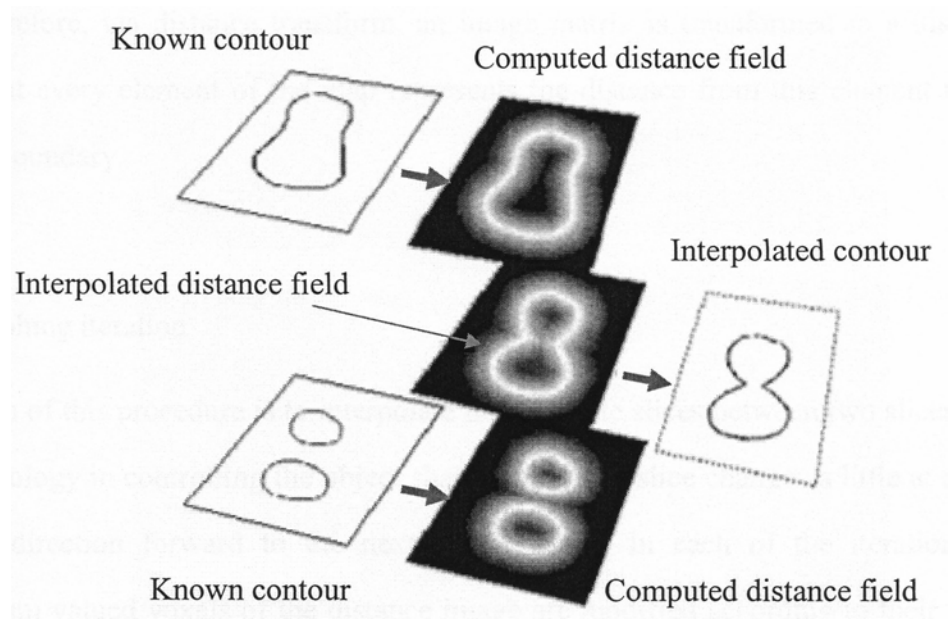


Figure 6-7 Illustration of inter-slice interpolation

1. Distance transform

The distance transform [64] is a powerful tool in image analysis. The distance transform used in this study is different from the traditional ones. It allows for both positive and negative distances: Distance within the area of the object from its boundary is positive and that out of the area of the object or in the background is negative. We assume that the boundary of the object is between the positive boundary and the negative boundary.

Definition:  $\mathbf{I} = \mathbf{I}(x, y, z)$  is the cross sectional image,  $\mathbf{O} = \mathbf{O}(x, y, z)$  is the object region in  $\mathbf{I}$  and  $\mathbf{D} = \mathbf{D}(x, y, z)$  is the distance map of  $\mathbf{I}$ , where  $(x, y, z)$  are the voxel coordinates ( $z$  is the slice serial number). The distance from the boundary of the object is defined by

$$\mathbf{D}(x, y, z) = \begin{cases} +d(x, y, z) & \text{if } \mathbf{I}(x, y, z) \in \mathbf{O} \\ -d(x, y, z) & \text{if } \mathbf{I}(x, y, z) \notin \mathbf{O} \end{cases} \quad (6-1)$$

where  $d(x, y, z)$  is the shortest distance from voxel  $(x, y, z)$  to the boundary of object

$\mathbf{O}$ . Therefore, via distance transform, an image matrix is transformed to a distance map that every element of the map represents the distance from this element to the object boundary.

## 2. Morphing iteration

The aim of this procedure is to interpolate intermediate slices between two slices. The methodology in controlling the object shape is that one slice changes a little at a time in the direction forward to the next object shape. In each of the iteration, the maximum valued voxels of the distance image are modified according to their signs. The positive maximum valued voxels are eroded and negative maximum valued voxels are dilated. From the result we can get an intermediate slice or morph.

Assume that the two neighbour slices are represented by  $\mathbf{I}_1(x, y, z_1)$  and  $\mathbf{I}_2(x, y, z_2)$ , and the distance transform of  $\mathbf{I}_1(x, y, z_1)$  is  $\mathbf{D}_1(x, y, z_1)$ . A masking step is performed using  $\mathbf{I}_2(x, y, z_2)$  as a mask to cover the distance map  $\mathbf{D}_1(x, y, z_1)$  to  $\mathbf{D}_2(x, y, z_2)$ , according to the following operation

$$\mathbf{D}_2(x, y, z_2) = \mathbf{I}_2 \bullet \mathbf{D}_1 = \begin{cases} \mathbf{D}_1(x, y, z) & \text{if } \mathbf{I}_2(x, y, z) \neq 0 \\ 0 & \text{otherwise} \end{cases}, \quad (6-2)$$

where “ $\bullet$ ” means the multiplication of corresponding voxels. Hence,  $\mathbf{D}_2(x, y, z_2)$  contains the shape information of both objects  $\mathbf{O}_1$  and  $\mathbf{O}_2$ . The non-zero voxels are the object voxels of  $\mathbf{O}_2$ . The positive voxels represent that they are in the corresponding object region of  $\mathbf{O}_1$ , while the negative voxels are out of the object

region of  $\mathbf{O}_1$ . The absolute values of non-zero voxels of  $\mathbf{D}_2$  are the distances from the boundary of  $\mathbf{O}_1$ . Therefore  $\mathbf{D}_2(x, y, z_2)$  is the link between the objects in the two slices. Let

$$D_{\max} = \text{Max}[\text{Abs}[\mathbf{D}_2(x, y, z_2)]], \quad (6-3)$$

where *Abs* is absolute value operation and *Max* is the maximum value operation. The value of  $D_{\max}$  represents the shape difference between the two objects.

To get the interpolated slices,  $\mathbf{D}_2(x, y, z_2)$  is iteratively morphed by eroding and dilating according to the value of  $D_{\max}$ , shown in Figure 6.8. At the end of iteration, we get the  $n$ -th intermediate slice or the  $n$ -th morph. The result of the  $n$ -th iteration is

$$\mathbf{D}_2^n(x, y, z_2) = \text{Erosion}[\text{Dilation}[\mathbf{D}_2^{n-1}(x, y, z_2)]] \quad (6-4)$$

where

$$\text{Erosion}[\mathbf{D}_2^n(x, y, z_2)] = \begin{cases} 0 & \text{if } \mathbf{D}_2^{n-1}(x, y, z_2) = -D_{\max} + n - 1, \\ \mathbf{D}_2^{n-1}(x, y, z_2) & \text{otherwise} \end{cases}, \quad (6-5)$$

$$\text{Dilation}[\mathbf{D}_2^n(x, y, z_2)] = \begin{cases} N[\mathbf{D}_2^{n-1}(x, y, z_2)] = N[\mathbf{D}_1(x, y, z_1)] & \text{if } \mathbf{D}_2^{n-1}(x, y, z_2) = D_{\max} - n + 1, \\ \mathbf{D}_2^{n-1}(x, y, z_2) & \text{otherwise} \end{cases} \quad (6-6)$$

where  $N[.]$  includes the current voxel and its 4-neighbor voxels.

Finally, the intermediate slices are computed by a simple comparison given by

$$\mathbf{I}^n(x, y, z_n) = \begin{cases} 1 & \text{if } \mathbf{D}_2^n(x, y, z_2) \neq 0 \\ 0 & \text{otherwise} \end{cases}. \quad (6-7)$$

After the distance transform and morphing, the morph  $\mathbf{D}_2(x, y, z_2)$  includes the shape information of both objects. It changes the shape of the first object towards the shape of the second object. Figure 6.9 shows the interpolated slices between two

original segmented tumour slices (isotropic, inter-slice spacing: 7 mm). The selected interpolated slices among five original segmented tumour slices are shown in Figure 6.10. The results show that complex interpolation problems such as irregular shape, branching, hollow regions, and invaginations can be processed properly.

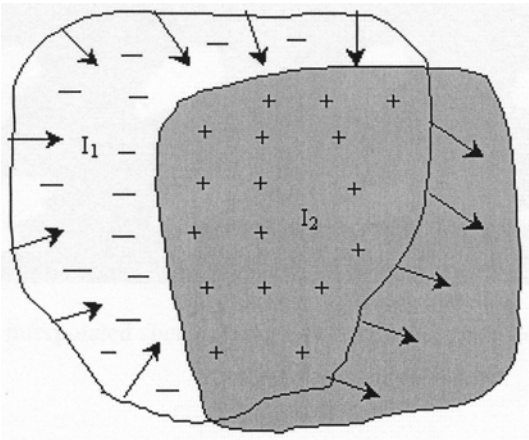


Figure 6.8 Sketch of shape morphing

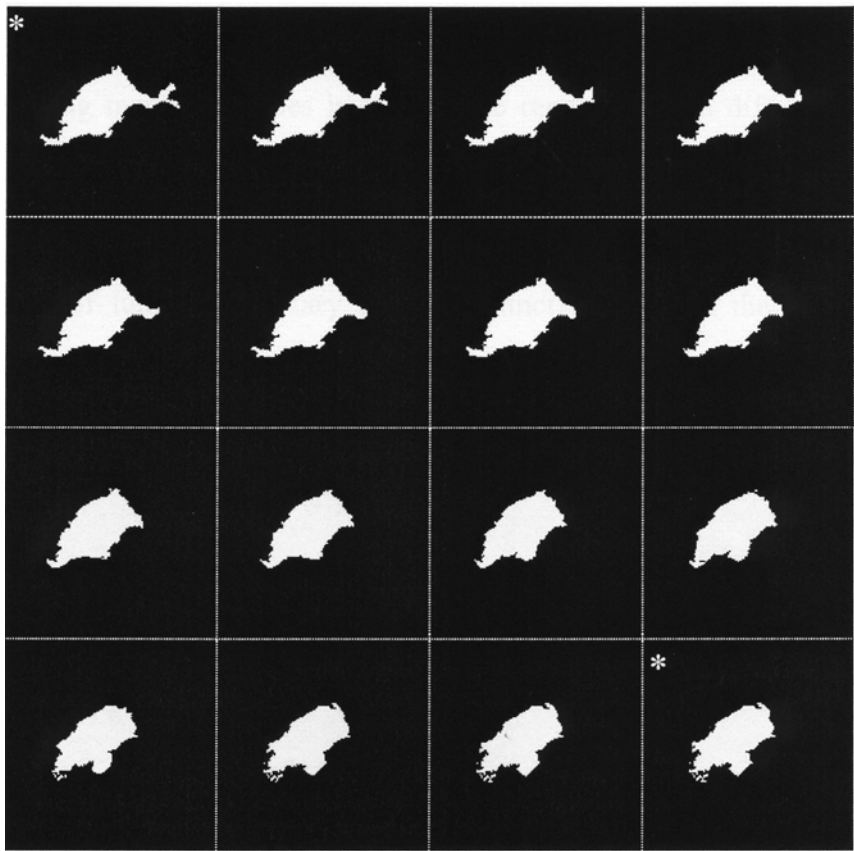
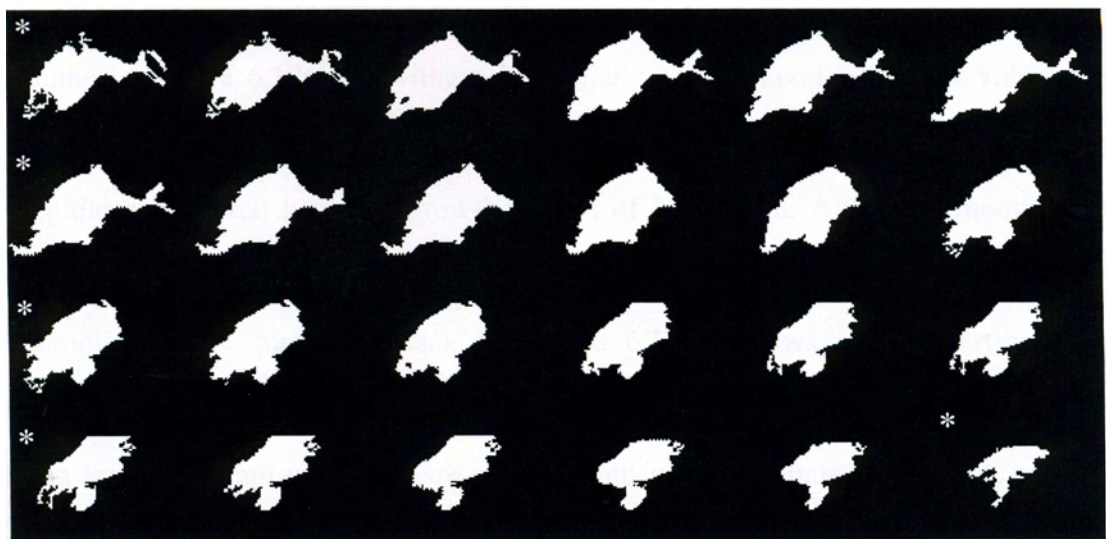


Figure 6.9 Interpolated slices between two original segmented tumour slices (white star)

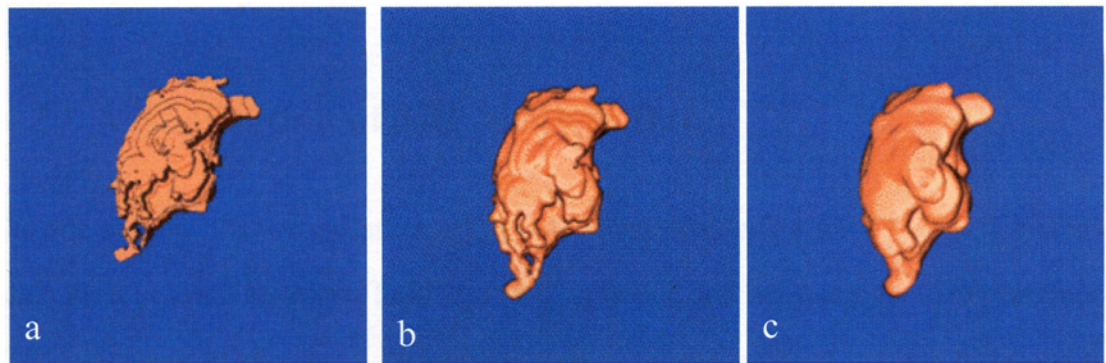




**Figure 6.10** Selected interpolated slices among five original segmented tumour slices (white star)

**6.4.2 Selected Results**

Figure 6.11 shows the 3-D views of visualised tumour of superficial spreading tongue carcinoma using marching cubes based surface rendering with different smoothing effects. Without smoothing, the visualised tumour has staircase like surface since the inter-slice resolution is lower than the intra-slice resolution. After Gaussian smoothing, the transition of tumour boundary is less distinct and shows the natural surface curvature.



**Figure 6.11** 3-D views of tongue carcinoma (a) Without smoothing; (b) With Gaussian smoothing and standard deviation=2.0; (c) With Gaussian smoothing and standard deviation=5.0.

Two 3-D views of visualised NPC tumours are shown in Figure 6.12. The tumour volume in Figure 6.12(a) is without interpolation while another tumour volume in Figure 6.12(b) is with interpolation. It is evident that NPC tumours are of very irregular shapes and have the growth pattern of infiltration. Although smoothing is used, the staircase-like tumour surface also appears in Figure 6.11(a). With interpolation, the tumour surface in Figure 6.12(b) shows a smooth transition, however it does not look natural. Figure 6.12(c) shows a 3-D view of a visualised brain tumour (1-mm of slice thickness without gap). Compared with NPC tumour, the shape of this brain tumour is more regular. Since the slice thickness is very thin and there is no inter-slice gap, the reconstructed tumour surface shows a very natural curvature and smooth transition.



**Figure 6.12** 3-D views of NPC and brain tumour (a) NPC without interpolation; (b) Another NPC with interpolation; (c) Brain tumour without interpolation.

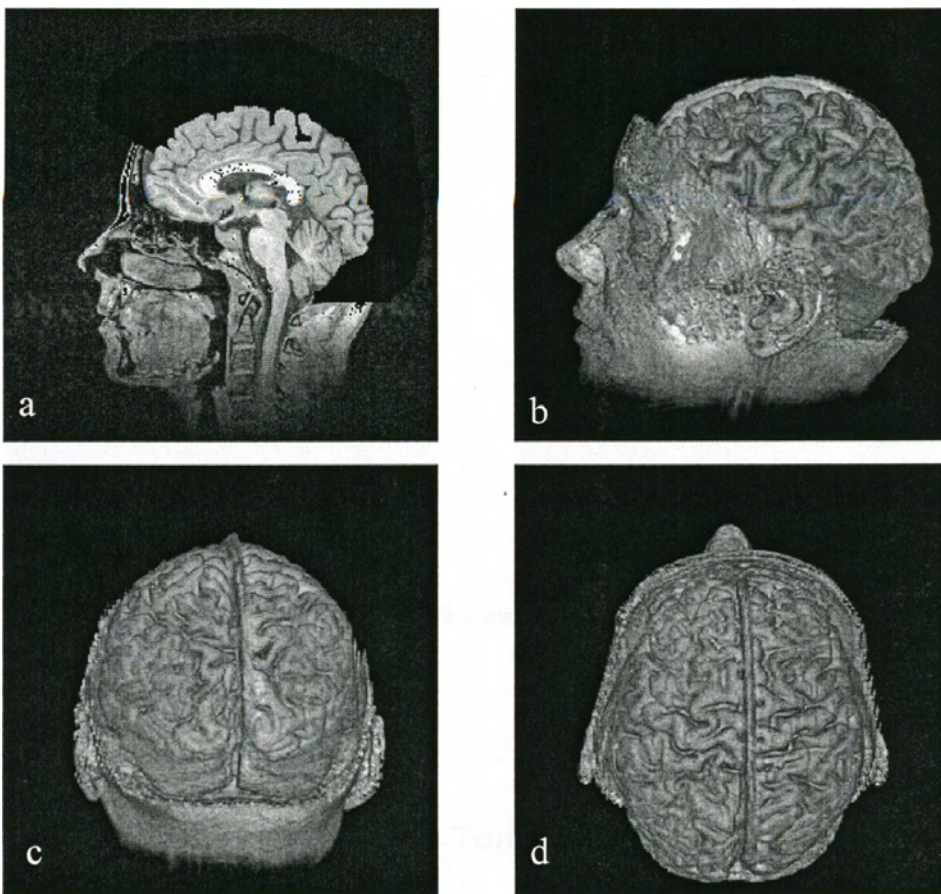
## 6.5 Visualisation of 3-D Medical Data

In this section, the visualisation results of volume rendering for a few 3-D medical data such as brain structure, head and neck tumours, and brain tumours are presented. The rendering effects for 3-D medical data with and without the combinations of segmentation results are also compared.



### 6.5.1 Visualisation of Brain Structure

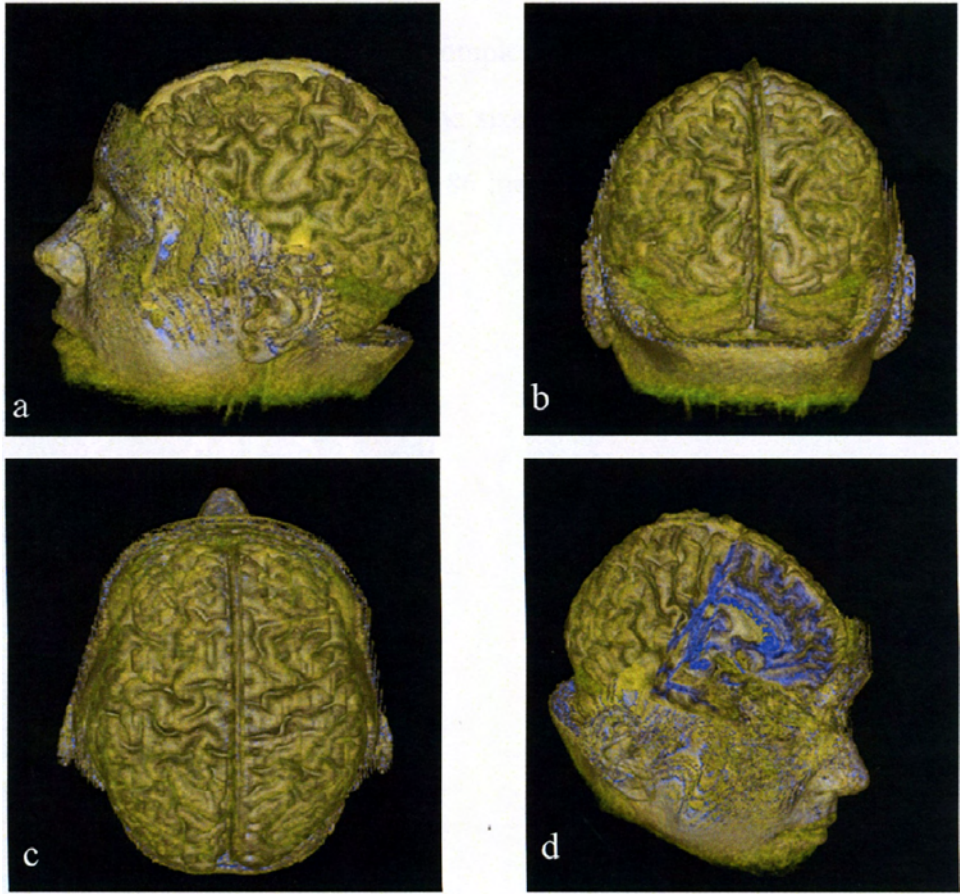
Figure 6.13 shows the 3-D views of a normal skull-removed 3-D MRI brain data (This data can be downloaded from [http://www.siggraph.org/education/materials/vol-viz/volume\\_visualisation\\_data\\_sets.htm](http://www.siggraph.org/education/materials/vol-viz/volume_visualisation_data_sets.htm)). Figure 6.13(a) is one of the 2-D slices and Figures 6.13(b)-6.13(d) are the greyscale 3-D views from three different viewpoints. The size of the data (sagittal slice) is 109x256x256(X-Y-Z) and the voxel resolution is 1.4 mm x 1 mm x 1 mm (X-Y-Z) with the voxel length of 16bits.



**Figure 6.13** (a) 2-D slice; (b) 3-D left view; (c) 3-D back view; (d) 3-D overhead view.

Figure 6.14 shows the colour shaded 3-D views of this 3-D MRI brain data. Figures 6.14(a)-6.14(c) are the 3-D views from three different viewpoints and Figure 6.14(d) is the 3-D view of the cropped data. From these figures, it can be seen that better visualisation effects and more elaborate anatomic details can be demonstrated

using colour shading. By cropping function, the internal structure of an opaque object can be observed.



**Figure 6.14** (a) 3-D left view; (b) 3-D back view; (c) 3-D overhead view; (d) Cropped data.

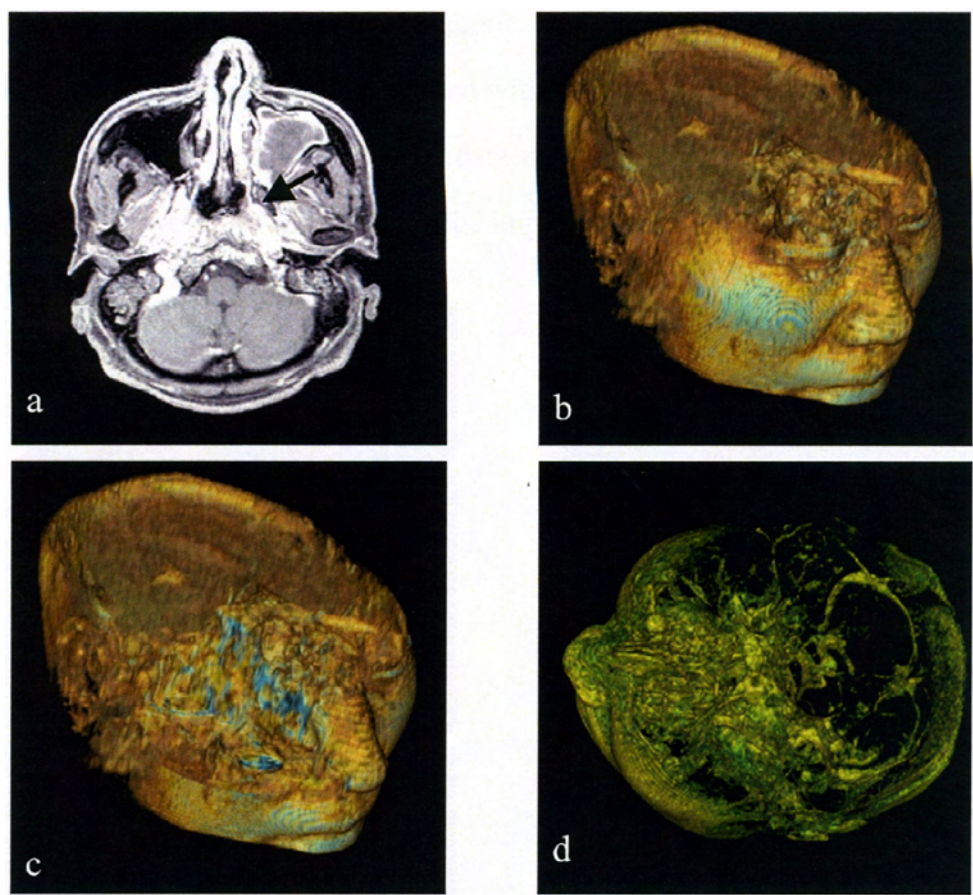
**6.5.2 Visualisation of NPC and Tongue Carcinoma**

1. Visualisation of NPC from 3-D MRI data

The 3-D views of a contrast-enhanced 3-D spoiled gradient recalled echo (SPGR) MRI head data obtained from the Singapore National Cancer Centre are shown in Figure 6.15. One of the 2-D slices is shown in Figure 6.15(a), the colour shaded 3-D view is shown in Figure 6.15(b), and Figure 6.15(c) is the colour shaded 3-D view of cropped data and the enhancement of nasal mucosa and inflammation can be clearly



observed in blue colour. However, the tumour enhancement cannot be observed clearly. In addition, Figure 6.15(d) is the overhead view of this data after the adjustment of intensity-opacity function for the removal of the brain tissue. From Figure 6.15(d), we can observe the complex bone-soft tissue structures around the skull base and nasopharynx region. The size of the data (axial slice) is 256x256x30 (X-Y-Z) and the voxel resolution is 0.86 mm x 0.86 mm x 3 mm (X-Y-Z) with the voxel length of 16bits.

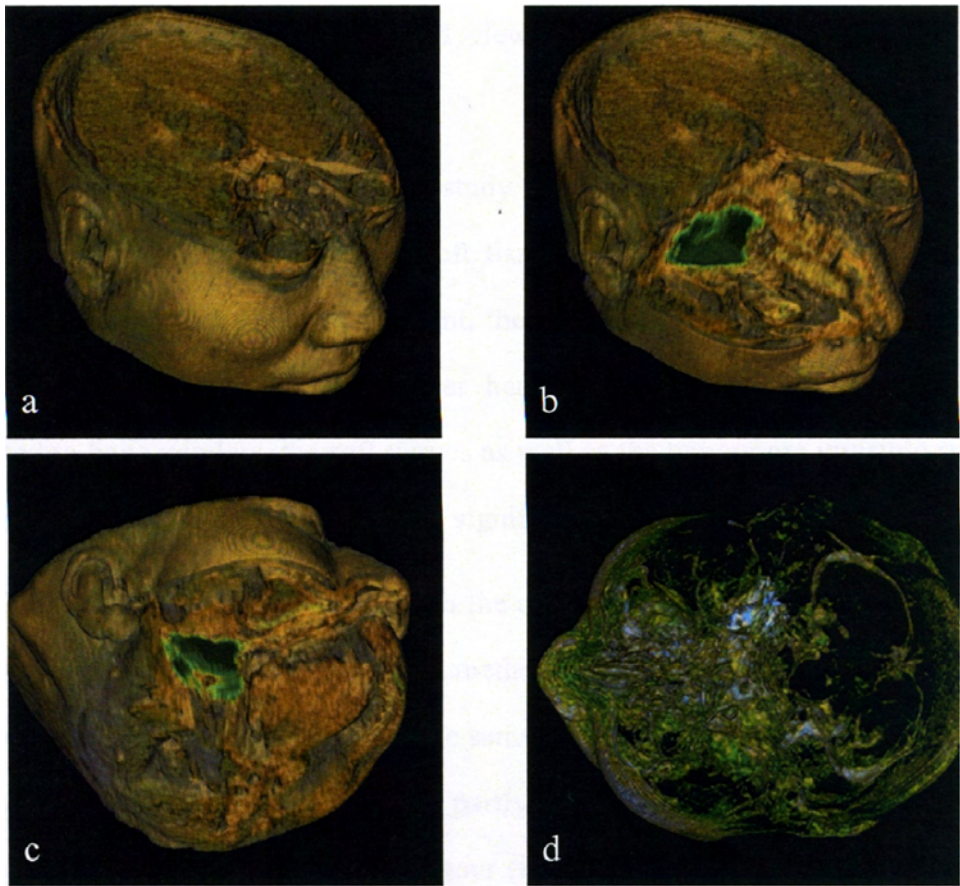


**Figure 6.15** (a) 2-D slice of NPC (black arrow); (b) 3-D view; (c) 3-D view of cropped data; (d) Overhead view of this data after the adjustment of intensity-opacity function.

The 3-D views of this data with the combinations of segmentation results are shown in Figure 6.16. The combination method is to use the segmentation information and remap the segmented tumour to the same locations of the original data by a high greyscale value. In this figure, Figure 6.16(a) is the colour shaded 3-D view which is

close to Figure 6.15(b), Figures 6.16(b) and 6.16(c) are the colour shaded 3-D view of cropped data from two different viewpoints and the tumour enhancement is clearly seen in green colour, and Figure 6.16(d) is the overhead view of this data after the adjustment of intensity-opacity function so that the brain tissue almost disappears. From Figure 6.16(d), a part of the NPC (in blue colour) is observed and the other part is covered by the structures around the skull base.

It can be observed that compared with traditional 2-D slices, 3-D views from volume rendering supplies better anatomical descriptions for both pathological and normal structures. In addition, compared with the examples shown in Figure 6.15, the segmentation combined 3-D medical data are introduced in the example shown in Figure 6.16 for volume rendering and the improved visual results are obtained.



**Figure 6.16** (a) 3-D view; (b) and (c) 3-D views of cropped data from different viewpoints; (d) Overhead view of this data after the adjustment of intensity-opacity function.

## 2. Visualisation of NPC from 3-D CT data

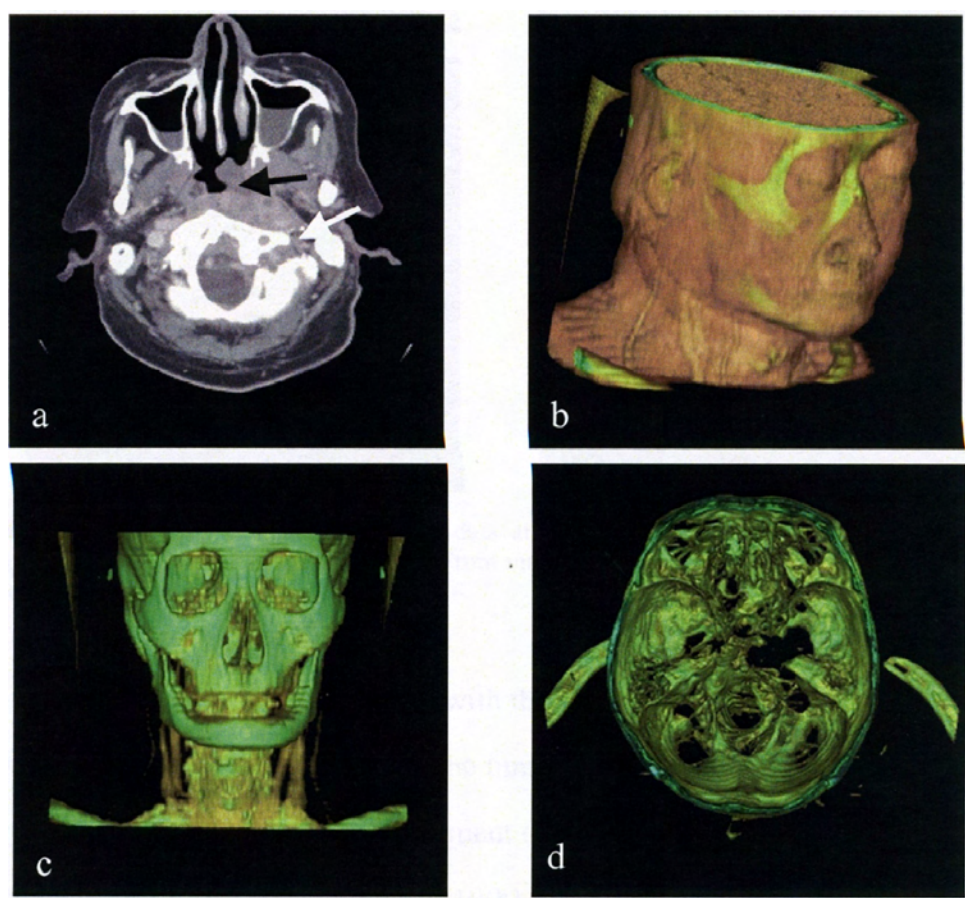
Figure 6.17 shows the 3-D views of a CT head data obtained from the Singapore National Cancer Centre. Where, Figure 6.17(a) is one of the 2-D slices, Figures 6.17(b)-6.17(d) are the 3-D views from different viewpoints with different settings of intensity-opacity function. The size of the data (axial slice) is 512x512x46 (X-Y-Z) and the voxel resolution is 0.43 mm x 0.43 mm x 4 mm (X-Y-Z) with the voxel length of 16bits. In Figure 6.17(a), the NPC extension to the pharyngeal recess (black arrow) and the NPC caused bone erosion (white arrow) can be clearly observed. In Figure 6.17(b), the opacity map is adjusted to render skin and fat transparent. As a result, superficial structures such as the external jugular veins and anterior jugular veins are visualised. In Figure 6.17(c), the presence of head and neck musculature is minimised and high contrast soft tissue structures such as lymph node and vessels are seen. In Figure 6.17(d), the overhead view at bone window clearly shows bone erosion caused by the tumour.

The limitation of this CT-based study is that the tumour cannot be directly differentiated from the surrounding soft tissue by visual observation at soft tissue window since without the contrast agent, the intensity of NPC is close to that of the surrounding soft tissues. On the other hand, if the intensity-opacity function is adjusted to bone window, the soft tissues as well as the tumour are invisible since the intensities of tumour and soft tissue are significantly less than those of bone structures.

The 3-D views of this CT data with the combinations of segmentation results are shown in Figure 6.18. The combination method is to use the segmentation information and remap the segmented tumour to the same locations of the original data by a high greyscale value. The 3-D view of the partly cropped data using skin-fat window is depicted in Figure 6.18(a) and the tumour (blue colour) is seen in the cropped part. The overhead view of this data after axial cropping using soft tissue window is shown in Figure 6.18(b) and the tumour (blue colour) is seen. The front view of this data



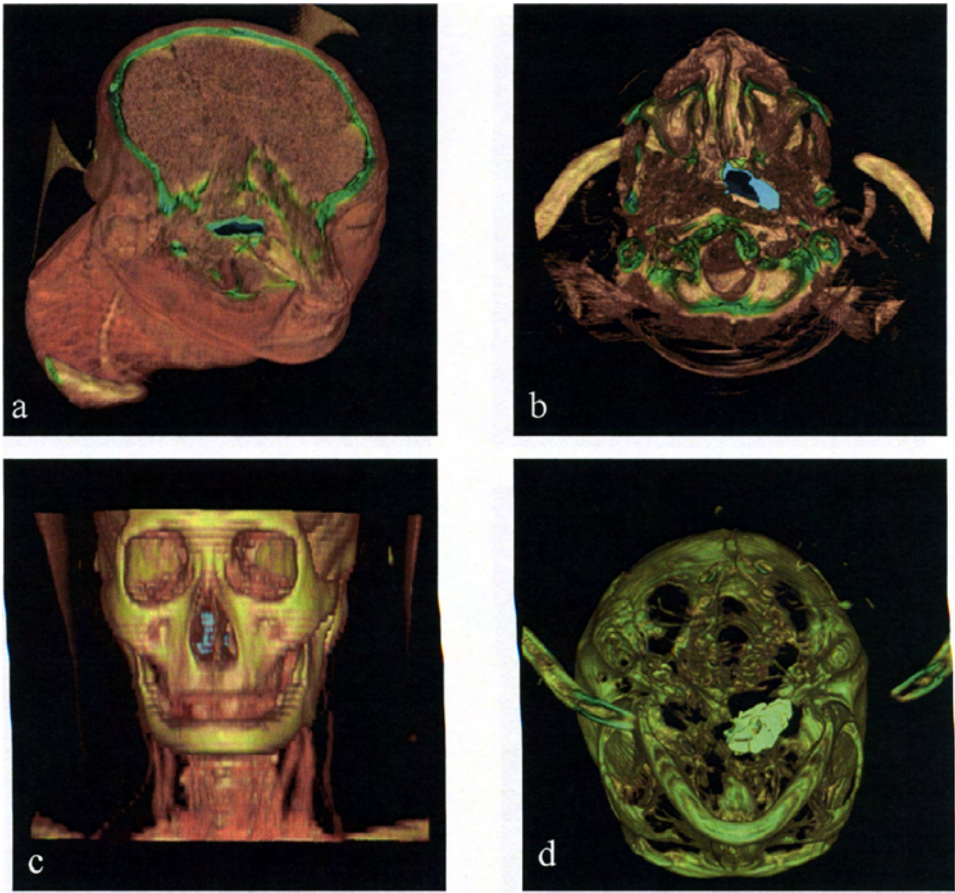
using soft tissue-bone window is presented in Figure 6.18(c) and the tumour (blue colour) hidden behind the nasal cavity is observed. The bottom view of this data using bone window is depicted in Figure 6.18(d) and the tumour (light green colour) can be clearly observed. The present work is very useful to the clinicians for clinical diagnosis and treatment planning. The 3-D views shown in Figures 6.15 to 6.18 had been presented in a lecture for fresh medical doctors.



**Figure 6.17** (a) 2-D slice of NPC with bone erosion; (b)-(d) 3-D views from different viewpoints with different settings of intensity-opacity function.

3. Visualisation of tongue carcinoma from 3-D MRI data

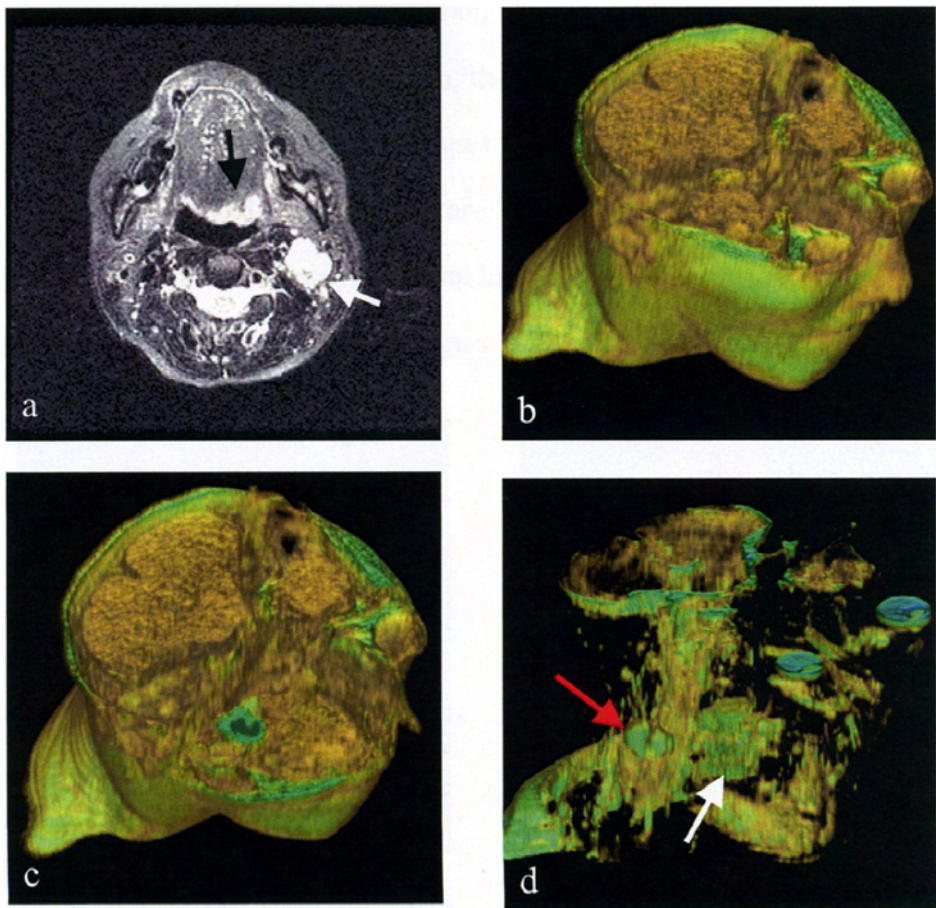
The 3-D view of a 3-D T2W MRI head data obtained from the Singapore National Cancer Centre with the combinations of segmentation results is shown in Figure 6.19. The size of the data (axial slice) is 256x256x24 (X-Y-Z) and the voxel resolution is



**Figure 6.18** (a) 3-D view of cropped data at skin-fat window; (b) Overhead view of cropped data at soft tissue window; (c) Front view at soft tissue-bone window; (d) Bottom view at bone window.

0.76 mm × 0.76 mm × 5 mm (X-Y-Z) with the voxel length of 16bits. One of the 2-D slices is shown in Figure 6.19(a) and the tumour enhancement (black arrow) with the enlarged cervical lymph node enhancement (white arrow) is observed. The 3-D view of this data is illustrated in Figure 6.19(b). The 3-D view of this data after fence cropping shown in Figure 6.19(c) displays a tumour mass at the posterior part of the tongue. The 3-D view of this data after the adjustment of intensity-opacity function is shown in Figure 6.19(d). In this figure, the enhanced tumour (white arrow) and the enlarged posterior cervical lymph node (red arrow) are visualised. This 3-D view corresponds to the 2-D slices well. The views in Figures 6.19(c) and 6.19(d) are clinically relevant and useful that they could be utilised for comparative evaluation after the surgical procedure.



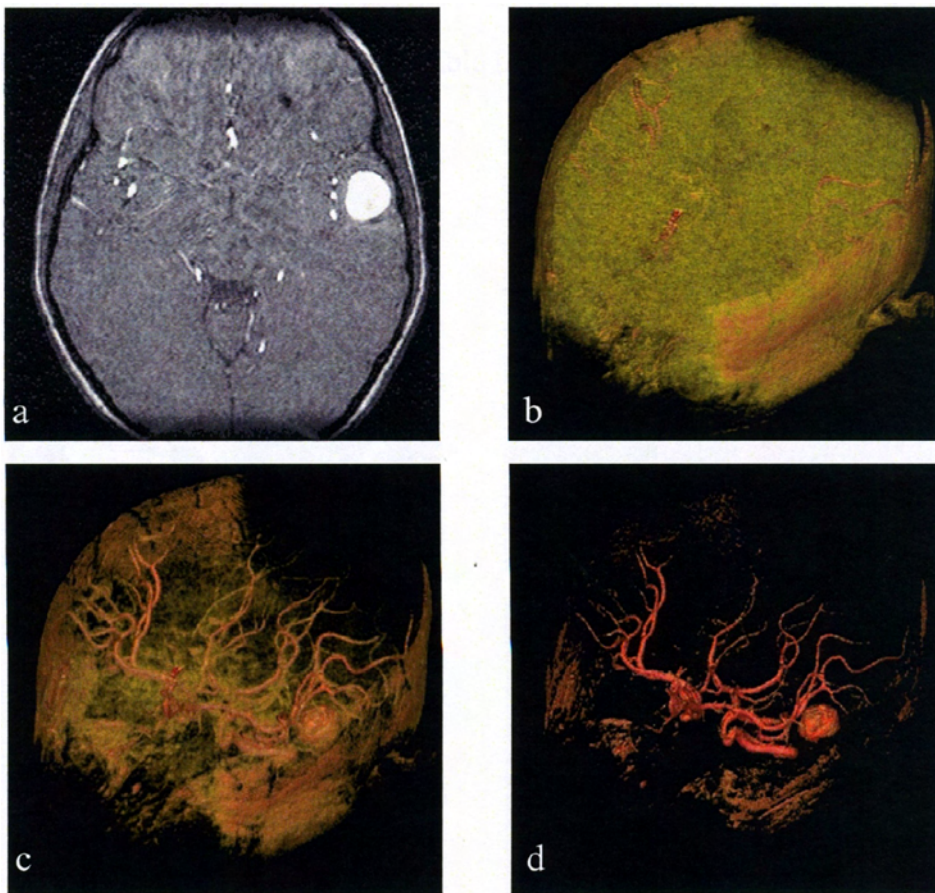


**Figure 6.19** (a) 2-D slice of tongue carcinoma with enlarge cervical lymph node; (b) 3-D views of this data; (c) 3-D view after fence cropping; (d) 3-D view after the adjustment of intensity-opacity function.

### 6.5.3 Visualisation of Brain Tumour

The 3-D views of a 3-D time-of-flight (TOF) MR angiography (MRA) brain data with brain aneurysm (acquired from the attached image database of Vitrea System, Vital Image Inc., MN, USA) are shown in Figure 6.20. Figure 6.20(a) corresponds to one of the 2-D slices and Figures 6.20(b)-6.20(d) display the 3-D views from the same viewpoint with different settings of intensity-opacity function. The size of the data (axial slice) is 512x512x74(X-Y-Z) and the voxel resolution is 0.31 mm x 1 mm x 1 mm (X-Y-Z) with the voxel length of 16 bits. In Figure 6.20(a), the aneurysm and the vessels have high contrast to the surrounding brain tissues. In Figure 6.20(b), the

visible brain tissues make most of the brain vessels and the aneurysm invisible. By the adjustment of intensity-opacity function, the brain vessel network is visible in Figure 6.20(c). In Figure 6.20(d), the brain tissues and the skull are almost removed, and the brain vessel network and the aneurysm are visualised clearly. Thus, the adjustment of intensity-opacity function enhances visualisation effect and enables better treatment planning. The different 3-D views in Figure 6.20 are the result of the present effort.

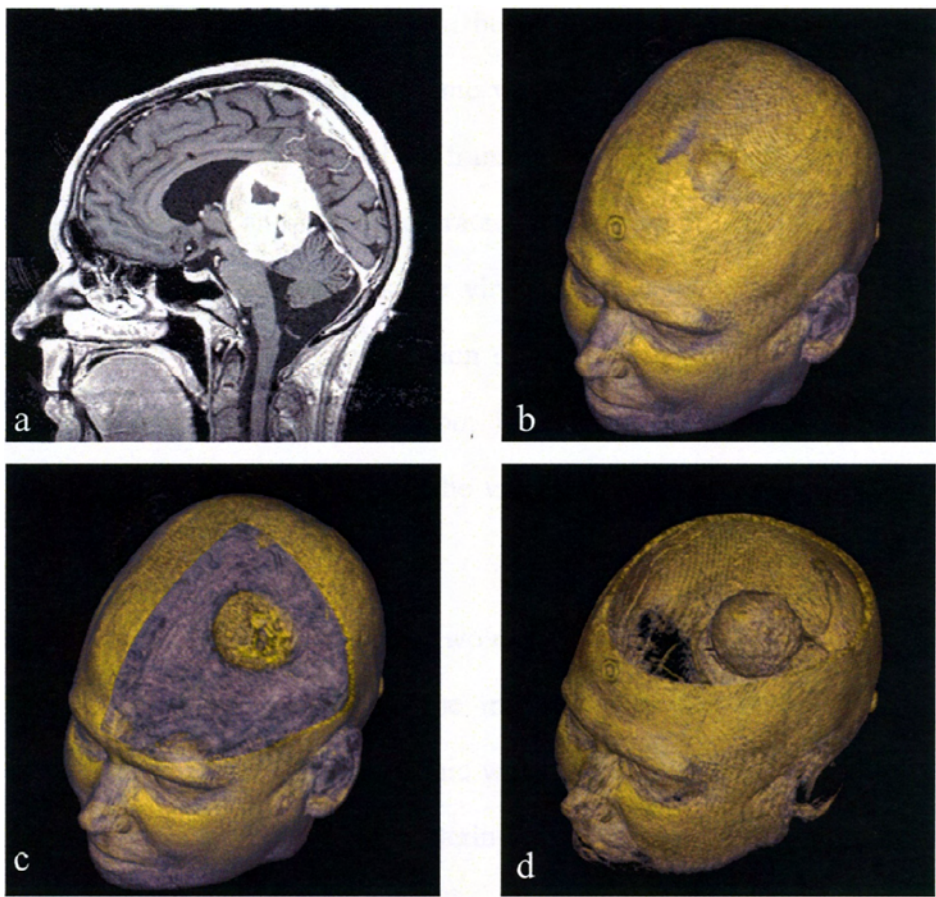


**Figure 6.20** (a) 2-D slice of brain aneurysm; (b)-(d) 3-D views from the same viewpoint with different settings of intensity-opacity function.

The 3-D views of a 3-D MR brain data with intracranial tumour (acquired from the attached image database of Vitrea System) are shown in Figure 6.21. One of the 2-D slices is displayed in Figure 6.21(a) and the 3-D views with different settings of intensity-opacity function and cropping are shown in Figures 6.21(b)-6.21(d). The size of the data (sagittal slice) is 119x256x256(X-Y-Z) and the voxel resolution is



1.67 mm × 0.90 mm × 0.90 mm (X-Y-Z), with the voxel length of 16 bits. In Figure 6.21(a), the enhanced tumour is clearly seen. Figure 6.21(b) is a 3-D view where the tumour is hidden inside the skull and brain tissues so that it is invisible. Figure 6.21(c) is the 3-D view of this data with the same intensity-opacity function setting but using fence cropping. In this figure, the tumour is visible at the cropping site. Figure 6.21(d) shows the 3-D view of this data with the adjusted intensity-opacity function and sub-volume cropping. In this figure brain tissue is almost invisible and part of the crania is removed, hence the whole tumour is visible from the viewpoint of the cropping site.



**Figure 6.21** (a) 2-D slice of brain tumour; (b)-(d) 3-D views from the same viewpoint with different settings of intensity-opacity function and cropping.

## 6.6 Discussion

### 6.6.1 Visualisation Platform and Techniques

#### 1. Hardware support

In this study, VTK, a visualisation software development kit, was used to build a visualisation platform and all the rendering algorithms integrated in VTK are implemented by software. Therefore, although the common 2-D/3-D graphic hardware make surface rendering perform well even on PC system, the real time volume rendering is still a big problem because of the huge computation. For an instance, in order to render a data volume with the size of 512x512x512 and voxel value length of 16-bit at the rate of 30 frame/sec, 256MB system memory, memory transfer rate at 8GB/sec, and instruction rate at 40GHz are needed [146]. In particular, when real time interaction such as the virtual endoscope is needed, the speed of software rendering is far from satisfaction using common computing system. When the interaction operations such as zoom in, zoom out, rotation or cropping are performed in our platform Voxur-3D, the view of the rendered object shall be very blurry.

Besides the algorithm optimisation, two methods are used to improve the real time performance of volume rendering. One method is to use the high performance computing system such as special graphic workstation or parallel processing system. The other approach is to use the rendering accelerator hardware on common PC system. For the first method, the limitations are the high cost, very little flexibility, and implementation difficulties.

Liu, et al. developed a visualisation framework “4DView” using the integration of VTK and the volume rendering accelerator VolumePro 500 on a common PC system [146]. VolumePro 500 developed by TeraRecon Inc. was the first PCI-based single

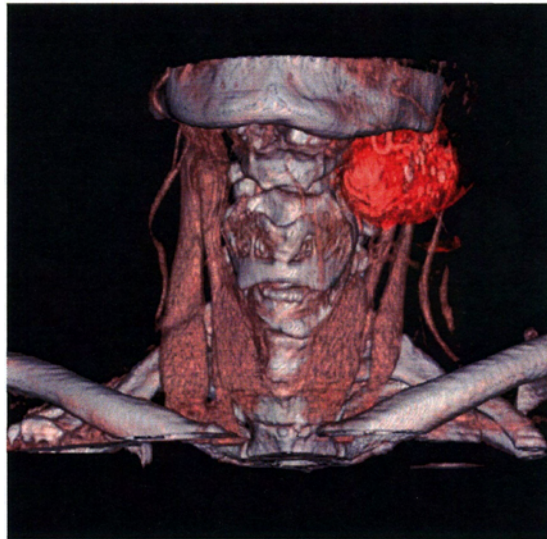
chip real time volume rendering accelerator. It achieved a rendering rate of 30 frame/sec for a data volume with the size of 256x256x256 using shear-warp ray casting algorithm. For VolumePro 1000, a rendering rate of 30 frame/sec was achieved for a data volume with the size of 512x512x512[147]. By the integration of VolumePro 500 and VTK, the 4DView achieved real time surface rendering and volume rendering for 3-D medical data on the common PC system at a high performance-to-cost ratio. Therefore in future development, the support from hardware can be employed to improve the performance of visualisation platform.

## 2. Visualisation techniques

One difficulty for common users to manipulate the medical visualisation system is the configure of visualisation control parameters, including window/level in 2-D image display, intensity-opacity function, intensity-colour function, gradient opacity function, shading properties, etc. If the user does not understand the relationship between voxel value and the corresponding tissue type, it is very hard to obtain idea of 3-D view to demonstrate potentially useful information hidden in the data via the adjustment of a set of control parameters. In addition, some parameters are related to each other. The frequent manual adjustment of control parameters is very inconvenient for clinicians and this process would tend to decrease the efficiency in diagnosis. In this study, a general set of control parameters provided by VTK was used to display the initial 3-D view. However, some further adjustments are needed according to different data.

The concept of default display protocol (DDP) was proposed for not only the display of 2-D images, but also the generated 3-D view [148]. In other words, a set of recommended rendering parameters are integrated and provided to achieve the ideal rendering result according to the particular imaging modality (CT or MRI), imaged

organ (brain, head, bone, spine, abdomen, cardiac, etc.), and imaging aim (morphology, perfusion, or function evaluation). In the commercial visualisation software Vitrea System, a set of DDPs were provided for users. The 3-D view of a cervical lymphoma using VolumePro and the DDP of head and neck CT is shown in Figure 6.22 [149]. In every DDP, the recommended scanning/acquisition parameters are also attached. However, currently there is no known published document on how to configure, combine, and test DDP, especially the rendering control parameters corresponding to the different object data.



**Figure 6.22** 3-D view of cervical lymphoma

Another challenge is how to distinctly visualise the abnormal part in head and neck region. From the experiment results shown in Sections 6.5.2 and 6.5.3, it is easier to view and recognise brain tumours than to recognise head and neck tumours. That may be due to some head and neck tumours being located in areas often obscured by surrounding tissues and bone structures, head and neck tumours often being adjacent to surrounding soft tissues so that the visual differentiation is difficult, and sometimes, the closeness of intensity of tumour tissue to those of surround tissues, therefore making it difficult to differentiate the tumour from the surrounding tissues by the generated colour 3-D views even if there are morphological changes.

Some potential methods for improvement include the introduction of cropping function, the integration of tumour segmentation results into the raw 3-D medical data, etc. Another promising enhancement can be the employment of virtual endoscopy and navigation. The oral cavity, nasal cavity, nasopharynx, pharynx, and larynx are either open structures or hollow alleyways for food, drink or air. Therefore it is suitable to perform the examination and diagnosis using virtual endoscopy and navigation from the original 3-D data. Chen, et al. developed an interactive virtual laryngoscopy based on a self-adaptive automated image segmentation, feature extraction, and efficient visualisation system [150]. The studies of Fried, et al. and Wang, et al. showed that virtual laryngoscopy might well provide clinical benefit in preoperative planning, staging, and intra-procedural guidance for head and neck disease [151][152]. Currently there is no report on virtual nasopharyngoscopy and hence it merits a further study. As mentioned above, virtual endoscopy needs frequent interaction and hence, hardware support is necessary.

### 6.6.2 3-D Medical Data Acquisition and Interpolation

The quality of 3-D medical data is also very important for visualisation. From the visualisation results presented in Sections 6.4 and 6.5, the 3-D data formed by thin slices produce elaborate and realistic effects of visualisation, such as Figures 6.12(c), 6.13, 6.14, 6.20, and 6.21. On the other hand, the 3-D views generated using thick slices look not so smooth and the serrate boundaries are always observed. In the embedded rendering algorithms of VTK, nearest neighbour interpolation is used for the dimension whose resolution is less than others to avoid the distortion of actual geometric shape. In the commercially available 3-D medical visualisation software Vitrea System, interpolation on original slices is performed after the data is imported



from DICOM server to produce the 3-D isotropic data for further processing and visualisation.

Thin slices with thickness of 1-2 mm and no inter-slice gap are preferred for visualisation. However, slices with thickness of 3-5 mm and inter-slice gap of 0-2 mm are clinically widely applicable for routine diagnostics because of the high signal-to-noise ratio, low examination cost and short scanning time. In the present study, distance transform and morphing were applied in a shape-based interpolation procedure for the segmented tumour slices in surface rendering. By this method, smooth intermediate slices were obtained. However, some questions surface, such as (a) the long computing time for image distance transform, (b) the transition of a part of object boundary in some interpolated slices being rigid due to the discrete changes of object difference for morphing control, and (c) the interpolated slices are estimations and the quantitative evaluation of interpolation quality is difficult because to obtain the GT of intermediate slices is hard, unless thin-sliced imaging is performed. In the present study, no interpolation was performed as the pre-processing for the 3-D medical data formed by original stacked slices before volume rendering. If the resolution of one dimension is about  $1/5$  to  $1/6$  or more of the other dimensions, the quality of generated 3-D views processed by nearest neighbour is acceptable. If the resolution of one dimension is less than  $1/8$  of the other dimensions, the serration effect of generated 3-D views is obvious as shown in Figures 6.17(c) and 6.18(c). In this situation, linear or bilinear interpolations on original stacked slices can be utilised to improve the rendering effects.

## 6.7 Concluding Remarks

In conclusion, a 3-D visualisation platform Voxur-3D that integrates marching cubes for surface rendering and ray casting for volume rendering was established to generate

the 3-D views of segmented medical objects and 3-D entire medical data. The development of this platform was based on the Visualization Toolkit (VTK) and Microsoft Visual C++. In summary, VTK is an object-oriented software development system for 3-D computer graphics, image processing, and visualisation. It consists of a C++ class library using dynamic linking and device-independent techniques, and an encapsulation layer formed by a few interpreted languages. In addition, the design of Voxur-3D including module description, system architecture, and the implementation was introduced. This platform was used to visualise segmentation tumours and 3-D medical data of head and neck tumours, brain tumours, and brain structures. An interpolation procedure based on distance transform and morphing was developed and applied to get higher inter-slice resolution for surface rendering. The results showed that the developed platform was able to produce 3-D views for 3-D medical data. In particular, the segmentation information was combined into the original 3-D clinical data of head and neck tumours. Consequently, the visualisation effect of the most pathologically significant object, tumour, was enhanced. For the purpose of accurate diagnosis and treatment planning, further investigations of certain specific visualisation techniques such as default display protocol (DDP), virtual laryngoscopy and nasopharyngoscopy are needed to better visualise the tumours and the anatomic structures in head and neck region, illustrating the relative positions and connections of vessels, nerves, lymph nodes, muscles, and bones.

## Chapter 7

# Conclusions and Recommendations

### 7.1 Conclusions

This thesis presents an in-depth research on the imaging-based tumour volumetric analysis of head and neck cancers. A novel image processing and analysis solution module on MRI segmentation, tumour volume estimation, and 3-D visualisation has been developed for fulfilling the increasing requirements on tumour volumetric analysis tool in clinical practice. The following contributions are the newly developed techniques integrated in the solution module:

- (1). A knowledge-based fuzzy clustering (KBFC) method that includes noise removal, initial segmentation, and knowledge-based image analysis was established to segment NPC from T1W and CET1W MR images. It first segments the MR images using semi-supervised fuzzy c-means and then, a knowledge-based image analysis procedure that integrates three types of anatomic and space knowledge contributes the refinement of segmentation. KBFC is a suitable segmentation method which not only well segments

NPC from MR images, but also achieves significant higher inter-operator agreement than manual tracing methods.

- (2). An unsupervised hierarchical method that includes noise removal, initial segmentation, symmetry detection and refinement has been developed to segment tongue carcinoma from T1W and T2W MR images. The initial segmentation is performed using a genetic algorithm-induced fuzzy clustering. An artificial neural network-based symmetry detection procedure, which utilises anatomic and geometry knowledge, is then applied to detect the initially extracted mass and make necessary refinement to obtain the final segmentation results. This method not only well segments tongue carcinoma from MR images, but also has high inter-operator agreement by using minimal interaction.
- (3). A two-step deformation model-based method has been applied to locate the boundary of an object for the segmentation of certain individual objects from an image. The region deformation keeps the local shape constraint of an object from statistics. On the other hand, the contour deformation controls the location of the curve to the actual edge and at the same time, the smoothness of the curve. This method accurately extracts the boundaries from those objects with a highly unclear object-background interface, inhomogeneous region contents, and motion as well as Gaussian blur. Based on the deformation models, a fast 3-D segmentation scheme using the extracted object boundary as the initial plan for the next/previous slice is applied to segment the boundary of lymph node and intracranial aneurysm from a stack of 2-D slices.
- (4). Finally, a 3-D visualisation platform Voxur-3D that integrates marching cubes for surface rendering and ray casting for volume rendering is established to generate the 3-D views of segmented medical objects and

whole 3-D medical data. The development of this platform is based on the Visualization Toolkit (VTK) and Microsoft Visual C++. The results show that the developed platform is able to produce 3-D views for 3-D medical data. In particular, we combined the segmentation information into the original 3-D clinical data of head and neck tumours. As a result, the visualisation effect of the most pathologically interested object, the tumour, is enhanced.

In the present work, an in-depth study of two main problems has been carried out for the development of an intelligent MR image analysis framework for the quantitative evaluation of tumour 3-D volume. This includes 1) the development of novel, suitable and accurate image segmentation methods for tumour volume evaluation, and 2) the validations for these methods by measuring the tumour volumes using a large amount of clinical MR images of NPC and tongue carcinoma. They form the major contributions of this thesis. To the best of our knowledge, such an intensive research work on quantitative volumetric analysis and visualisation for head and neck tumours is not found in literature.

The most significant contributions of the solutions developed in the present research are:

- (a). The solution module developed in the present study provides a novel intelligent image processing and analysis approach to quantitate and visualise tumour volume from MRI with minimal interaction.
- (b). This developed segmentation and visualisation tool can be easily extended to analyse other solid tumours as well as the organs for the purpose of volume and deformation evaluations.

## 7.2 Recommendations for Future Work

The task of tumour volumetric analysis using image processing and analysis on MRI can be broadly divided into image segmentation and 3-D visualisation. A high quality of image segmentation is undoubtedly the first important factor for good performance of tumour volumetric analysis. Also, a high quality of 3-D visualisation provides clinicians the most natural and direct method to explore the plentiful information hidden in medical data volume. Based on these and other factors related to this study, recommendations for further research are listed below.

- (a). Feature definition and extraction based on the information obtained from multi-sequence (three or more) MRI can be developed to provide better discrimination of different kinds of abnormalities such as tumour, oedema, necrosis, and post-treatment impairment. In particular, the discrimination of radiation impairment and recurrent tumour is very important for the follow-up management of cancer patients.
- (b). Work can be done on the quantitative description of knowledge model to better introduce the prior knowledge to assist the segmentation and decrease the false positives and false negatives.
- (c). Severe noise and imaging artefacts can affect the imaging quality as well as the image segmentation for volumetric analysis. The systematic and quantitative investigation is desirable on the mathematical description of the image noise, the estimation and modelling of the “true” observed noise, noise reduction and the noise tolerance of the segmentation methods.
- (d). The use of machine learning and soft computing-based techniques such as support vector machine (SVM) and support vector data description (SVDD)



on image segmentation can be studied to overcome the limitations in statistical model-based clustering methods.

- (e). New visualisation techniques such as hardware supported real time volume rendering, default display protocol, and virtual laryngoscopy and nasopharyngoscopy merit further studies to provide better visualisation result for tumour examination, detection, and diagnosis.
- (f). In the present study, the imaging-based volumetric analysis including tumour segmentation, volume estimation, and 3-D visualisation was performed on MR images of pre-treated head and neck tumours only. The analysis on images of post-treated tumours should be carried out in the further study to evaluate the treatment outcome and manage the follow-up. By these developed tools, the systemic investigation of the relationship between the tumour volume and TNM staging result can be carried out on head and neck cancers.

In closing, research effort was undertaken on applying engineering principles and technologies to solve challenging medical problems, particularly for the tumour volumetric analysis of head and neck cancers by the close collaboration with the physicians. It is believed that significant contributions have been made in the present research work, paving the way for intelligent and quantitative diagnosis for the head and neck cancer patients while attempting to render efficient and cost-effective services in healthcare delivery.

## Author's Publications

### *Papers in Peer Review Journals*

1. J. Zhou, T. K. Lim, V. F. H. Chong, et al. Segmentation and visualization of nasopharyngeal carcinoma using MFU. *Computers in Biology and Medicine*, 33:407-424, 2003.
2. J. Zhou, T. K. Lim, V. F. H. Chong, et al. A texture combined multispectral MRI segmentation for nasopharyngeal carcinoma. *Optical Review*, 10:405-410, 2003.
3. V. F. H. Chong, J. Zhou, J. B. K. Khoo, et al. Nasopharyngeal carcinoma tumor-volume measurement. *Radiology*, 231:914-921, 2004.
4. V. F. H. Chong, J. Zhou, J. B. K. Khoo, et al. Tongue Carcinoma: Tumor volume measurement. *International Journal of Radiation Oncology, Biology and Physics*, 59:59-66, 2004.
5. V. F. H. Chong, J. Zhou, J. B. K. Khoo, et al. Imaging-based metastatic lymph node volume measurement. Submitted to *European Radiology*, 2004.

### *Papers Presented at Conferences*

1. J. Zhou, T. K. Lim, and V. F. H. Chong. Tumor volume measurement for nasopharyngeal carcinoma using knowledge-based fuzzy clustering MRI segmentation. Presented at SPIE Medical Imaging, 24-28 Feb 2002, San Diego, CA, USA.

2. J. Zhou, T. K. Lim, V. F. H. Chong, et al. A texture combined multispectral MRI segmentation for nasopharyngeal carcinoma. Presented at Asian Symposium on Biomedical Optics and Photomedicine, 21-23 Oct 2002, Sapporo, Japan.
3. J. Zhou, T. K. Lim, V. F. H. Chong, et al. MRI tumor segmentation for nasopharyngeal carcinoma using genetic algorithm-based clustering techniques. Presented at International Congress on Biomedical Engineering, 4-7 Dec 2002, Singapore.
4. J. Zhou, F. Shao, T. K. Lim, et al. Fast segmentation of brain tumor from 3D MR angiography. Presented at World Congress on Medical. Physics and Biomedical Engineering, 24-29 Aug 2003, Sydney, Australia.
5. J. Zhou, T. K. Lim, V. F. H. Chong, et al. Tumor-volume measurement of tongue carcinoma using region deformation-based MRI segmentation. Presented at the 87th OSA Annual Meeting, 5-9 Oct 2003, Tucson, AZ, USA.
6. J. Zhou, S. M. Krishnan, V. F. H. Chong, et al. Extraction of tongue carcinoma using genetic algorithm-induced fuzzy clustering and artificial neural network from *MR* images. Presented at the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1-5 Sept 2004, San Francisco, CA, USA.

## Bibliography

- [1] “United States Cancer Statistics: 2000 Incidence”, Department of Health and Human Services, USA, 2003.  
URL: <http://www.cdc.gov/cancer/npcr/uscs/2000/index.htm>.
- [2] F. Bray, R. Sankila, J. Ferlay, et al. Estimates of cancer incidence and mortality in Europe in 1995. *European Journal of Cancer*, 38:99-166, 2002.
- [3] “Consolidated Report of the Population Based Cancer Registries: Incidence and Distribution of Cancer 1990-96”, Indian Council of Medical Research, 2001.  
URL: [http://icmr.nic.in/ncrp/ncrp\\_p/cancer\\_p\\_based.htm](http://icmr.nic.in/ncrp/ncrp_p/cancer_p_based.htm).
- [4] “The First Report of the National Cancer Registry: Cancer Incidence in Malaysia 2002”, National Cancer Registry, Ministry of Health, Malaysia, 2003.  
URL: [http://www.crc.gov.my/ncr/ncrReport\\_2002.htm](http://www.crc.gov.my/ncr/ncrReport_2002.htm).
- [5] “State of Health 2001: The Report of the Director of Medical Services”, Ministry of Health, Singapore, 2002. URL: <http://app.moh.gov.sg/pub/pub01.asp>.
- [6] L. H. Sobin and C. Wittekind (eds). *UICC TNM Classification of Malignant Tumors*, 6<sup>th</sup> edition. Wiley-Liss, New York, 2002.
- [7] “Head and neck cancer: Questions and answers”, Head and neck cancer home page, National Cancer Institute, USA. URL: [http://ilcis.nci.nih.gov/fact/6\\_37.htm](http://ilcis.nci.nih.gov/fact/6_37.htm).

- [8] M. Lenz. Editorial: Imaging of head and neck tumors. *European Journal of Radiology*, 33:151-152,2000.
- [9] “Molecular anatomy of head and neck cancer: A genomic/proteomic approach”, Proposal for research funds applications, National Institutes of Health, USA, 2003. URL: <http://grants1.nih.gov/grants/guide/rfa-files/RFA-DE-04-003.html>.
- [10] T.J. Vogl, J. Balzer, M. Mark, et al. *Differential Diagnosis in Head and Neck Imaging*. Thieme, New York, 1999.
- [11] V. F. H. Chong and Y. F. Fan. Detection of recurrent nasopharyngeal carcinoma: MR Imaging versus CT. *Radiology*, 202:463-470, 1997.
- [12] S. H. Ng, T. C. Chang, S. F. Ko, et al. Nasopharyngeal carcinoma: MRI and CT assessment. *Neuroradiology*, 39:741-746, 1997.
- [13] K. W. Sievers, H. Greess, U. Baum, et al. Paranasal sinuses and nasopharynx CT and MRI. *European Journal of Radiology*, 33:185-202, 2000.
- [14] M. Wakisaka, H. Mori, N. Fuwa, et al. MR analysis of nasopharyngeal carcinoma: correlation of the pattern of tumor extent at the primary site with the distribution of metastasized cervical lymph nodes: Preliminary results. *European Radiology*, 10:970-977, 2000.
- [15] M. Lenz, H. Greess, U. Baum, et al. Oropharynx, oral cavity, floor of the mouth: CT and MRI. *European Journal of Radiology*, 33:203-215, 2000.
- [16] W. M. Brekel. Lymph node metastases: CT and MRI. *European Journal of Radiology*, 33:230-238, 2000.
- [17] A. Tetsumura, N. Yoshino<sup>1</sup>, T. Amagasa, et al. High-resolution magnetic resonance imaging of squamous cell carcinoma of the tongue: an in vitro study. *Dentomaxillofacial Radiology*, 30:14-21,2001.

- [18] F. R. Miller, J. R. Wanamaker, P. Lavertu, et al. Magnetic resonance imaging and the management of parapharyngeal space tumors. *Head and Neck*, 18:67-77, 1996.
- [19] M. Lell, U. Baum, H. Greess, et al. Head and neck tumors: imaging recurrent tumor and post-therapeutic changes with CT and MRI. *European Journal of Radiology* 33:239-247, 2000.
- [20] V. M. Moharir, M. F. Fried, D. M. Vernick, et al. Computer-assisted three-dimensional reconstruction of head and neck tumors. *Laryngoscope*, 108:1592-1598, 1998.
- [21] A. G. Sorensen, S. Patel, C. Harmath, et al. Comparison of diameter and perimeter methods for tumor volume calculation. *Journal of Clinical Oncology*, 19:551-557, 2001.
- [22] F. A. Pameijer, A. J. M. Balm, F. J. M. Hilgers, et al. Variability of tumor volumes in T3-staged head and neck tumors. *Head and Neck*, 19:6-13, 1997.
- [23] D. T. T. Chua, J. S. T. Sham, D. L. W. Kwong, et al. Volumetric analysis of tumor extent in nasopharyngeal carcinoma and correlation with treatment outcome. *International Journal of Radiation Oncology Biology and Physics*, 39:711-719, 1997.
- [24] C. C. Chang, M. K. Chen, M. T. Liu, et al. Primary tumor volume delineation in nasopharyngeal carcinoma and correlation with 1997 AJCC tumor stage classification. *Journal of Otolaryngology*, 30:231-234, 2001.
- [25] D. J. Brenner. Dose, volume and tumor control predictions in radiotherapy. *International Journal of Radiation Oncology Biology and Physics*, 26:171-179, 1993.

- [26] C. R. Johnson, H. D. Thames, D. T. Huang, et al. The tumor volume and clonogen number relationship: tumor control predictions based upon tumor volume estimates derived from computed tomography. *International Journal of Radiation Oncology, Biology and Physics*, 33:281-287, 1995.
- [27] H. H. Dubben, H. D. Thames, and H. P. Beck-Bomholdt. Tumor volume: a basic and specific response predictor in radiotherapy. *Radiotherapy and Oncology*, 47:167-174, 1998.
- [28] C. R. Johnson, S. R. Khandelwal, R. K. Schmidt-Ullrich, et al. The influence of quantitative tumor volume measurements on local control in advanced head and neck cancer using concomitant boost accelerated superfractionated irradiation. *International Journal of Radiation Oncology, Biology and Physics*, 32:635-641, 1995.
- [29] R. Kurek, A. Kalogera-Fountzila, K. Muskalla, et al. Usefulness of tumor volumetry as a prognostic factor of survival in head and neck cancer. *Strahlentherapie und Onkologie*, 179:292-297, 2003.
- [30] J. Willner, K. Baier, L. Pfreunder, et al. Tumor volume and local control in primary radiotherapy of nasopharyngeal carcinoma. *Acta Oncologica*, **38**:1025-1030, 1999.
- [31] C. C. Chang, M. K. Chen, M-T. Liu, et al. Effect of Primary Tumour volumes in early T-stage nasopharyngeal carcinoma. *Journal of Otolaryngology*, 32:87-92, 2003.
- [32] W. M. Sze, A. W. M. Lee, T. K. Lau, et al. Primary tumor volume of nasopharyngeal carcinoma: Prognostic factor for local control. *International Journal of Radiation Oncology, Biology and Physics*, 59:21-27, 2004.



- [33] R. H. Spiro, A. G. Huvos, G. Y. Wong, et al. Predictive value of tumor thickness in squamous carcinoma confined to the tongue and floor of the mouth. *American Journal of Surgery*, 152:345-350, 1986.
- [34] H. Fukukano, H. Matsuura, Y. Hasegawa, et al. Depth of invasion as a predictive factor for cervical node metastasis in tongue carcinoma. *Head and Neck*, 19:205-210, 1997.
- [35] T. Asakage, T. Yokose, K. Mukai, et al. Tumor thickness predicts cervical metastasis in patients with stage I/II carcinoma of tongue. *Cancer*, 82:1443-1448,1998.
- [36] A. P. W. Yuen, K. Y. Lam, L. K. Lam, et al. Prognostic factors of clinically stage I and II oral tongue carcinoma-A comparative study of stage, thickness, shape, growth pattern, invasive front malignancy grading, Martinez-Gimeno score, and pathologic features. *Head and Neck*, 24:513-520,2002.
- [37] A. P. W. Yuen, K. Y. Lam, W. I. Wei, et al. A comparison of the prognostic significance of tumor diameter, length, width, thickness, area, volume, and clinicopathological features of oral tongue carcinoma. *American Journal of Surgery*, 180:139-143,2000.
- [38] P. M. Wu, D. T. T. Chua, J. S. T. Sham, et al. Tumor control probability of nasopharyngeal carcinoma: A comparison of different mathematics models. *International Journal of Radiation Oncology, Biology and Physics*, 37:613-620, 1997.
- [39] L. P. Clarke, R. P. Velthuizen, M. A. Camacho, et al. MRI segmentation: methods and applications. *Magnetic Resonance Imaging*, 13:343-368, 1995.

- [40] “NIH guide: Quantitation of tumor response to treatment: A three-dimensional approach”, Grants guide, National Institutes of Health, USA, 1992. URL: <http://grants.nih.gov/grants/guide/rfa-files/RFA-CA-92-008.html>.
- [41] “Strategic Document of DICOM”, DICOM Standards Committee, National Electrical Manufacturers Association, USA, 2003. URL: [http://medical.nema.org/dicom/geninfo/dicom\\_strategy/Strategy\\_2003-02-07.htm](http://medical.nema.org/dicom/geninfo/dicom_strategy/Strategy_2003-02-07.htm)
- [42] M. S. Mahaley, G. Y. Gillespie, and R. Hammett. Computerized tomography brain scan tumor volume determinations. *Journal of Neurosurgery*, 72:872-878, 1990.
- [43] S. Saini. Radiologic measurement of tumor size in clinical trials: Past, present, and future. *American Journal of Roentgenology*, 176:333-334, 2001.
- [44] D. X. Xue, E. Robert, and J. Albright. Microcomputer-based technique for 3-D reconstruction and volume measurement of computed tomographic images. Comparison of geometric and planimetry post-operative tumor volume effects on patient survival. *Computers in Biology and Medicine*, 29:377-392, 1999.
- [45] B. N. Joe, M. B. Fukui, C. C. Meltzer, et al. Brain tumor volume measurement: comparison of manual and semiautomated methods. *Radiology*, 212:811-816, 1999.
- [46] R. E. Clatterbuck, E. P. Sipos. The efficient calculation of neurosurgically relevant volumes from computed tomographic scans using Cavalieri’s direct estimator. *Neurosurgery*, 40:339-343, 1997.
- [47] P. Gibbs, D. L. Buckley, S. J. Blackband, et al. Tumor volume determination from MR images by morphological segmentation. *Physics in Medicine and Biology*, 41:2437-2446, 1996.

- [48] M. Vaidyanathan, L. P. Clarke, R. P. Velthuisen RP, et al. Comparison of supervised MRI segmentation methods for tumor volume determination during therapy. *Magnetic Resonance Imaging*, 13:719-728, 1995.
- [49] S. Partridge, L. Esserman, E. Kaplan, et al. Method for quantitative assessment of tumor volume response to neoadjuvant chemotherapy using MRI. In *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 3, pages 1645–1647, Chicago, USA, July 2000.
- [50] W. M. Shi, D. M. Wildrick, and R. Sawaya. Volumetric measurement of brain tumors from MR imaging. *Journal of Neuro-Oncology*, 37:87-93, 1998.
- [51] M. Vaidyanathan, R. P. Velthuisen, P. Venugopal, et al. Tumor volume measurements using supervised and semi-supervised MRI segmentation methods. In *Proceedings of Artificial Neural Networks in Engineering Conference*, pages 629–637, St. Louis, USA, Nov 1994.
- [52] S. M. Haney, P. M. Thompson, T. F. Cloughesy, et al. Tracking tumor growth rates in patients with malignant gliomas: a test of two algorithms. *American Journal of Neuroradiology*, 22:73-82, 2001.
- [53] M. Vaidyanathan, L. P. Clarke, L. O. Hall, et al. Monitoring brain tumor response to therapy using MRI segmentation. *Magnetic Resonance Imaging*, 15:323-34, 1997.
- [54] N. A. Mayr, V. A. Magnotta, J. C. Ehrhardt, et al. Usefulness of tumor volumetry by magnetic resonance imaging in assessing response to radiation therapy in carcinoma of the uterine cervix. *International Journal of Radiation Oncology Biology and Physics*, 35:915-924, 1996.

- [55] N. A. Mayr, W. T. C. Yuh, J. Zheng, et al. Tumor size evaluated by pelvic examination compared with 3-D MR quantitative analysis in the prediction of outcome for cervical cancer. *International Journal of Radiation Oncology Biology and Physics*, 39:395-404, 1997.
- [56] G. Moonis, J. Liu, J. K. Udupa, et al. Estimation of tumor volume with fuzzy-connectedness segmentation of MR images. *American Journal of Neuroradiology*, 23:356-363, 2002.
- [57] E. A. Vokurka, A. Henvadkar, N. A. Thacker, et al. Using Bayesian tissue classification to improve the accuracy of vestibular schwannoma volume and growth measurement. *American Journal Neuroradiology*, 23:459-467, 2002.
- [58] L. P. Clarke, R. P. Velthuizen, M. Clark, et al. MRI measurement of brain tumor response: comparison of visual metric and automatic segmentation. *Magnetic Resonance Imaging*, 16:271-279, 1998.
- [59] R. P. Velthuizen, L. P. Clarke, S. Phuphanich, et al. Unsupervised measurement of brain tumor volume on MR images. *Journal of Magnetic Resonance Imaging*, 5:594-604, 1995.
- [60] M. R. Kaus, S. K. Warfield, A. Nabavi, et al. Automated segmentation of MR images of brain tumors. *Radiology*, 218:586-591, 2001.
- [61] R. A. Robb. *Biomedical Imaging, Visualization, and Analysis*. Wiley-Liss, New York, 2000.
- [62] S. J. Riederer. The future technical development of MRI. *Journal of Magnetic Resonance Imaging*, 6:52-56, 1996.
- [63] J. Beutel, H. L. Kundel, R. L. Van Metter (eds). *Handbook of Medical Imaging*, volume 1: Physics and Psychophysics. SPIE Press, Bellingham, 2000.
- [64] K. R. Castleman. *Digital Image Processing*. Prentice-Hall, New Jersey, 1996.

- [65] A. K. Jain. *Fundamental of Digital Image Processing*. Prentice-Hall, New Jersey, 1989.
- [66] A. K. W. Law, H. Zhu, B. C. B. Chan, et al. Semi-automatic tumor boundary detection in MR image sequences. In *Proceedings of 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing*, pages 28–31, Hong Kong, May 2001.
- [67] R. Adams and L. Bischof. Seeded region growing. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 16:641-647, 1994.
- [68] A. Mehnert and P. Jackway. An improved seeded region growing algorithm. *Pattern Recognition Letters*, 18:1065-1071, 1997.
- [69] S. A. Hojjatoleslami and F. Kruggel. Segmentation of large brain lesions. *IEEE Transaction on Medical Imaging*, 20:666-669, 2001.
- [70] J. K. Udupa, L. Wei, S. Samarasekera, et al. Multiple sclerosis lesion quantification using fuzzy-connectedness principles. *IEEE Transaction on Medical Imaging*, 16:598-609, 1997.
- [71] C. Li, D. B. Goldgof, and L. O. Hall. Knowledge-based classification and tissue labeling of MR images of human brain. *IEEE Transaction on Medical Imaging*, 12:740-750, 1993.
- [72] N. Moon, E. Bullitt, K. V. Leemput, et al. Model-based brain and tumor segmentation. In *Proceedings of 16th International Conference on Pattern Recognition*, volume 1, pages 528-531, Quebec, Canada, Aug 2002.
- [73] R. Beichel, S. Mitchell, E. Sorantin, et al. Shape- and appearance-based segmentation of volumetric medical images, In *Proceedings of 2001 IEEE International Conference on Image Processing*, volume 2, pages 589-592, Thessaloniki, Greece, Oct 2001.

- [74] C. J. Rajapakse, N. J. Giedd, and J. L. Rapoport. Statistical approach to segmentation of single-channel cerebral MR images. *IEEE Transaction on Medical Imaging*, 16:176-186, 1997.
- [75] K. Held, E. R. Kops, B. J. Krause, et al. Markov random field segmentation of brain MR images. *IEEE Transaction on Medical Imaging*, 16:878-886, 1997.
- [76] Y. Zhang, M. Brady, and S. Smith. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging*, 20:45-57, 2001.
- [77] L. P. Clarke, R. P. Velthuizen, S. Phuphanich, et al. MRI: stability of three supervised segmentation techniques. *Magnetic Resonance Imaging*, 11:95-106, 1993.
- [78] Y. H. Kao, J. A. Sorenson, and S. S. Winkler. MR image segmentation using vector decomposition and probability techniques: a general model and its application to dual-echo images. *Magnetic Resonance in Medicine*, 35:114-125, 1996.
- [79] Y. H. Kao, J. A. Sorenson, and S. S. Winkler. Dual-echo MRI segmentation using vector decomposition and probability techniques: a two-tissue model. *Magnetic Resonance in Medicine*, 32:342-357, 1994.
- [80] D. L. Pham, C. Xu, and J. L. Prince. Current methods in medical image segmentation. *Annual Review of Biomedical Engineering*, 2:315-337, 2000.
- [81] A. M. Bensaid, L. O. Hall, J. C. Bezdek, et al. Partially supervised clustering for image segmentation. *Pattern Recognition*, 29:859-871, 1996.
- [82] D. L. Pham and J. L. Prince. Adaptive fuzzy segmentation of magnetic resonance images. *IEEE Transaction on Medical Imaging*, 18:737-752, 1997.

- [83] M. L. Rhodes. Computer graphics in Medicine: the past decade. *IEEE Computer Graphics and Applications*, 11:52-54, 1991.
- [84] A. Kaufman, K. Hoehne, W. Krueger, et al. Research issues in volume visualization. *IEEE Computer Graphics and Applications*, 14:63-67, 1994.
- [85] W. Schroeder, K. Martin, and B. Lorensen. *The Visualization Toolkit: An Object-Oriented Approach to 3D Graphics*. Prentice Hall, New Jersey, 1998.
- [86] [http://www.terarecon.com/products/volume\\_pro\\_prod.html](http://www.terarecon.com/products/volume_pro_prod.html)
- [87] L. K. Ng and N. Madhusudanan. *Tumor volume measurement and 3D visualization of nasopharyngeal carcinoma using magnetic resonance imaging*. Final Year Project Report, Nanyang Technological University, Singapore, 2003.
- [88] S. K. Mukherji and J. A. Castelijns (eds). *Modern Head and Neck Imaging*. Springer-Verlag, Berlin Heidelberg, 1999.
- [89] J. D. Howard, A. D. Elster, and J. S. May. Temporal bone: three-dimensional CT. I. Normal anatomy, techniques and limitations. *Radiology*, 177:421-425, 1990.
- [90] C. F. Darling, S. E. Byrd, E. D. Allen, et al. Three-dimensional computed tomography imaging in the evaluation of craniofacial abnormalities. *Journal of the National Medical Association*, 86:676-680, 1994.
- [91] L. S. Johnson, C. A. Pelizzari, R. Grzeszczuk, et al. A pilot study on the use of volume visualization in image-based treatment planning for head and neck cancer. *International Journal of Radiation Oncology, Biology and Physics*, 32(supplement 1):302, 1995.
- [92] J. S. Lee, A. B. Jani, C. A. Pelizzari, et al. Volumetric visualization of head and neck CT data for treatment planning. *International Journal of Radiation Oncology, Biology and Physics*, 44:693-703, 1999.



- 
- [93] H. Greess, A. Nomayr, B. Tomandl, et al. 2D and 3D visualization of head and neck tumours from spiral-CT data. *European Journal of Radiology*, 33:170-177, 2000.
  - [94] M. G. Cavalcanti and M. W. Vannier. Measurement of the volume of oral tumors by three-dimensional spiral computed tomography. *Dentomaxillofacial Radiology*, 29:35-40, 2000.
  - [95] D. G. Bragg, P. Rubin, and H. Hricak (eds). *Oncologic Imaging 2<sup>nd</sup> Ed*, WB Saunders, Philadelphia, 2002.
  - [96] C. C. Hsu, P. H. Lai, C. Lee, et al. Automated nasopharyngeal carcinoma detection with dynamic gadolinium-enhanced MR imaging. *Methods of Information in Medicine*, 40:331-337, 2001.
  - [97] R. M. Haralick and L. G. Shapiro. *Computer and Robot Vision*, volume 1. Addison-Wesley, New York, 1992.
  - [98] V. F. H. Chong and S. Y. Tsao (eds). *Nasopharyngeal Carcinoma*. Armour, Singapore, 1997.
  - [99] S. Nawaratne, R. Fabiny, J. E. Staquet, et al. Accuracy of volume measurement using helical CT. *Journal of Computer Assisted Tomography*, 21:481-486, 1997.
  - [100] S. A. Sohaib, B. Turner, J. A. Hanson, et al. Ct assessment of tumour response to treatment: comparison of linear, cross-sectional and volumetric measures of tumour size. *British Journal of Radiology*, 73:1178-1184, 2000.
  - [101] J. M. Bland and D. G. Altman. Statistical methods of assessing agreement between two methods of clinical measurement. *Lancet*, 1:307-310, 1986.
  - [102] M. Filippi, M. A. Horsfield, S. Bressi, et al. Intra- and inter-observer agreement of brain MRI lesion volume measurement in multiple sclerosis: A comparison of techniques. *Brain*, 118:1593-1600, 1995.

- [103] M. C. Clark, L. O. Hall, D. B. Goldgof, et al. Automatic tumor segmentation using knowledge-based techniques. *IEEE Transaction on Medical Imaging*, 17:187-201, 1998.
- [104] L. M. Fletcher-Heath, L. O. Hall, D. B. Goldgof, et al. Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. *Artificial Intelligence in Medicine*, 21:43-63, 2001.
- [105] A. Mahr, S. Levegrün, M. L. Bahner, et al. Usability of semiautomatic segmentation algorithms for tumor volume determination. *Investigative Radiology*, 34:143-150, 1999.
- [106] R. V. Hogg and J. Ledolter. *Engineering Statistics*. MacMillan, New York, 1987.
- [107] V. F. H. Chong, Y. F. Fan, and J. B. K. Khoo. Retropharyngeal lymphadenopathy in nasopharyngeal carcinoma. *European Journal of Radiology*, 21:100-105, 1995.
- [108] V. F. H. Chong, J. Y. Zhou, J. B. K. Khoo, et al. Nasopharyngeal carcinoma tumor-volume measurement. *Radiology*, 231:914-921, 2004.
- [109] T. Taxt, A. Lundervold, B. Fuglaas, et al. Multispectral analysis of uterine corpus tumors in magnetic resonance imaging. *Magnetic Resonance in Medicine*, 23:55-76, 1992.
- [110] R. L. Galloway, R. J. Maciunas, and A. L. Failing. Factors affecting perceived tumor volumes in magnetic resonance imaging. *Annals of Biomedical Engineering*, 21:367-375, 1993.
- [111] M. Tann, V. Sopov, S. Croitoru, et al. How accurate is helical CT volumetric assessment in renal tumors? *European Radiology*, 11:1435-1438, 2001.

- [112] M. A. Kuriakose, T. R. Loree, W. L. Hicks, et al. Tumor volume estimated by computed tomography as a predictive factor in carcinoma of the tongue. *British Journal of Oral and Maxillofacial Surgery*, 38:460-465, 2000.
- [113] U. Maulik and S. Bandyopadhyay. Genetic algorithm-based clustering technique. *Pattern Recognition*, 33:1455-1465, 2000.
- [114] K. F. Man, K. S. Tang, and S. Kwong. *Genetic Algorithms Concepts and Design*. Springer-Verlag, London, 1999.
- [115] L. Y. Tseng and S. B. Yang. A genetic clustering algorithm for data with non-spherical-shape clusters. *Pattern Recognition*, 33:1251-1259, 2000.
- [116] N. Sundararajan. *Genetic Algorithms*. Lecture Notes, Nanyang Technological University, Singapore, 2002.
- [117] J. Zhou, T. K. Lim, V. F. H. Chong, et al. A texture combined multispectral magnetic resonance imaging segmentation for nasopharyngeal carcinoma. *Optical Review*, 10:405-410, 2003.
- [118] K. I. Kim, K. Jung, S. H. Park, et al. Support vector machines for texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:1542-1550, 2002.
- [119] K. Chan, T. W. Lee, P. A. Sample, et al. Comparison of machine learning and traditional classifiers in glaucoma diagnosis. *IEEE Transactions on Biomedical Engineering*, 49:963-974, 2002.
- [120] D. P. Peck, E. M. Spickler, D. D. Cody, et al. Comparison of tumor volume measurement and partial volume effects. Presented at *the 7th Scientific Meeting of the International Society for Magnetic Resonance in Medicine*, Philadelphia, USA, May, 1999.

- [121] T. McInerney and D. Terzopoulos. Deformable models in medical image analysis: a survey. *Medical Image Analysis*, 1:91-108, 1996.
- [122] F. H. Y. Chan, F. K. Lam, P. W. F. Poon, et al. Object boundary location by region and contour deformation. *IEE Proceedings of Vision, Image and Signal Processing*, 143:353-360, 1996.
- [123] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour Model. *International Journal of Computer Vision*, 1:321-331, 1988.
- [124] L. D. Cohen and I Cohen. Finite-element methods for active contour models and balloons for 2-D and 3-D images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15:1131-1147, 1993.
- [125] A. A. Amini, T. E. Weymouth, and A. K. Jain. Using dynamic programming for solving variational problems in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12:855-867, 1990.
- [126] D. J. Williams and M. Shah. A fast algorithm for active contours and curvature estimation. *Computer Vision, Graphics and Image Processing: Image Understanding*, 55:14-26, 1992.
- [127] J. M. Gauch, H. H. Pien, and J. Shah. Hybrid boundary-based and region-based deformable models for biomedical image segmentation. *Proceedings of SPIE*, 2299:72-83, 1994.
- [128] D. M. Honea, Y. Ge, W. E. Snyder, et al. Lymph-node segmentation using active contours. *Proceedings of SPIE*, 3034:265-273, 1997.
- [129] X. M. Pardo, M. J. Carrerira, A. Mosuquera, et al. A snake for CT image segmentation integrating region and edge information. *Image and Version Computing*, 19:461-475, 2001.

- [130] A. K. W. Law, H. Zhu, F. K. Lam, et al. Tumor boundary extraction in multislice MR brain images using region and contour deformation. In *Proceedings of 2001 International Workshop on Medical Imaging and Augmented Reality*, pages 183-187, Hong Kong, June 2001.
- [131] W. J. Conover. *Practical Nonparametric Statistics*, 2<sup>nd</sup> edition, John Wiley, New York, 1980.
- [132] J. Ivins and J. Porrill. A semiautomatic tool for 3D medical image analysis using active contour models. *Medical Informatics*, 19:81-90, 1994.
- [133] I. L. Herlin, C. Nguyen and C. Graffigne. Stochastic segmentation of ultrasound images. In *Proceedings of 11th IEEE International Conference on Pattern Recognition*, volume 1, pages 289-292, Hague, Netherlands, August 1992.
- [134] B. S. Kuszyk, D. R. Ney, and E. K. Fishman. The current state of the art in three dimensional oncologic imaging: an overview. *International Journal of Radiation Oncology, Biology and Physics*, 33:1029-1039, 1995.
- [135] W. J. Schroeduer, L. S. Avila, and W. Hoffman. Visualizing with VTK: a tutorial. *IEEE Computer Graphics and Applications*, 20:20-27, 2000.
- [136] W. J. Schroeder (ed). *The Visualization User's Guide*, Kitware Inc., New York, 2001.
- [137] W. J. Schroeder, K. M. Martin, and W. E. Lorensen. The design and implementation of an object-oriented for 3D graphics and visualization. In *Proceedings of Visualization '96*, pages 93-100, 472, San Francisco, USA, Oct 1996.
- [138] W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. *Computer Graphics*, 21:163-169, 1987.

- [139] M. Levoy. Efficient ray tracing of volume data. *ACM Transaction on Graphics*, 9:245-261, 1990.
- [140] A. Goshtasby, D. A. Turner, and L. V. Ackerman. Matching of tomographic slices for interpolation. *IEEE Transactions on Medical Imaging*, 11:507-516, 1992.
- [141] P. N. Werahera, G. J. Miller, G. D. Taylor, et al. A 3-D reconstruction algorithm for interpolation and extrapolation of planar cross sectional data. *IEEE Transactions on Medical Imaging*, 14:765-771, 1995.
- [142] J.F. Guo, Y. L. Cai, Y. P. Wang, et al. Morphology-based interpolation for 3D medical image reconstruction. *Computerized Medical Imaging and Graphics*, 19:267-279, 1995.
- [143] Y. H. Liu, Y. N. Sun, C.W. Mao, et al. Edge-shrinking interpolation for medical images. *Computerized Medical Imaging and Graphics*, 21:91-101, 1997.
- [144] B. Luo and E. R. Hancock. Slice interpolation using the distance transform and morphing. In *Proceedings of 13th International Conference on Digital Signal Processing*, volume 2, pages 1083-1086, Santorini, Greece, July 1997.
- [145] "ANALYZE 7.5 File Format", Mayo Clinic, USA, 2003. URL: <http://www.mayo.edu/bir/PDF/ANALYZE75.pdf>
- [146] J. Liu, C. Jia, and H. Duan. A real-time medical image visualization framework with VTK and VolumePro. In *Proceedings of Computer Assisted Radiology and Surgery*, pages 1-5, Paris, France, June 2002.
- [147] [http://www.terarecon.com/downloads/products/datasheet\\_vp1000.pdf](http://www.terarecon.com/downloads/products/datasheet_vp1000.pdf)
- [148] <http://www.dkfz-heidelberg.de/biostatistics/iscb-gmds-99/abstracts/20035.pdf>
- [149] [http://www.terarecon.com/gallery/med\\_gallery.html](http://www.terarecon.com/gallery/med_gallery.html)

- 
- [150] D. Chen, B. Li, P. Roche, et al. Feasibility studies of virtual laryngoscopy by CT and MRI-from data acquisition, image segmentation, to interactive visualization. *IEEE Transactions on Nuclear Science*, 48:51-57, 2001.
- [151] M. P. Fried, V. M. Moharir, H. Shinmoto, et al. Virtual laryngoscopy. *Annals of Otolaryngology, Rhinology and Laryngology*, 108:221-226, 1999.
- [152] D. Wang, W. Zhang, M. Xiong, et al. Laryngeal and hypopharyngeal carcinoma: comparison of helical CT multiplanar reformation, three-dimensional reconstruction and virtual laryngoscopy. *Chinese Medical Journal (English)*, 114:54-58, 2001.



# Appendix

## Medical Image File Format for the Study

There are a few commonly used file formats for medical image, such as DICOM (Digital Imaging and Communication in Medicine) developed by the American College of Radiology (ACR) and the National Electrical Manufacturers Association (NEMA), ANALYZE developed by Mayo Clinic, HUGO developed by Medical Virtual Reality Studio GmbH in Germany, etc. Here, the ANALYZE 7.5 file format used in the present study is given.

### ANALYZE 7.5 File Format

The image database is the system of files that the ANALYZE™ package uses to organise and access image data on the disk. Facilities are provided for converting data from a number of sources for use with the package. A description of the database format is provided to aid developers in porting images from other sources for use with the ANALYZE™ system. An ANALYZE™ image database consists of at least two files:

- an image file
- aheaderfile

The files have the same name being distinguished by the extensions .img for the image file and .hdr for the header file. Thus, for the image database heart, there are the UNIX files heart.img and heart.hdr. The ANALYZE™ programs all refer to this pair of files as a single entity named heart.

### Image File

The format of the image file is very simple containing usually uncompressed pixel data for the images in one of several possible pixel formats:

- **1 bit**                      packed binary (slices must begin on byte boundaries)
- **8 bit**                      8 bits per pixel (unsigned char)
- **16 bit**                     16bits per pixel (signed short)
- **32 bit**                     32 bits per pixel signed integers, or floating point
- **64 bit**                     64 bits per pixel; double precision, floating point, or complex
- **24 bit**                     RGB, 8-bits per channel Red, Green, Blue.

### Head File

The header file is represented here as a 'C' structure which describes the dimensions and history of the pixel data. The header structure consists of three substructures:

**header\_key**                      describes the header

<b>image_dimension</b>	describes image sizes
<b>data_history</b>	optional.

The description of data structure of ANALYZE™ 7.5 file format is as follows.

```
/* ANALYZE™ 7.5 Header File Format
*
*
* Biomedical Imaging Resource
* Mayo Foundation
*
* dbh.h
*
*
* databse sub-definitions
*/

struct header_key /* header key */
{ /* off + size */
    int sizeof_hdr; /* 0 + 4 */
    char data_type[10]; /* 4 + 10 */
    char db_name[18]; /* 14 + 18 */
    int extents; /* 32 + 4 */
    short int session_error; /* 36 + 2 */
    char regular; /* 38 + 1 */
    char hkey_un0; /* 39 + 1 */
}; /* total=40 bytes */

struct image_dimension /* off + size */
{ /* 0 + 16 */
    short int dim[8]; /* 16 + 2 */
    short int unused8; /* 18 + 2 */
    short int unused9; /* 20 + 2 */
    short int unused10; /* 22 + 2 */
    short int unused11; /* 24 + 2 */
    short int unused12; /* 26 + 2 */
    short int unused13; /* 28 + 2 */
    short int unused14; /* 30 + 2 */
    short int datatype; /* 32 + 2 */
    short int bitpix; /* 34 + 2 */
    short int dim_un0; /* 36 + 2 */
    float pixdim[8]; /* 36 + 32 */
    /*
        pixdim[] specifies the voxel dimensions:
        pixdim[1] - voxel width
        pixdim[2] - voxel height
        pixdim[3] - interslice distance
        ...etc
    */
    float vox_offset; /* 68 + 4 */
    float funused1; /* 72 + 4 */
    float funused2; /* 76 + 4 */
    float funused3; /* 80 + 4 */
    float cal_max; /* 84 + 4 */
    float cal_min; /* 88 + 4 */
}
```

```
float compressed;          /* 92 + 4      */
float verified;            /* 96 + 4      */
int glmax,glmin;           /* 100 + 8     */
};                          /* total=108 bytes */

struct data_history
{
    /* off + size */
    char descrip[80];       /* 0 + 80      */
    char aux_file[24];      /* 80 + 24     */
    char orient;            /* 104 + 1     */
    char originator[10];    /* 105 + 10    */
    char generated[10];     /* 115 + 10    */
    char scannum[10];       /* 125 + 10    */
    char patient_id[10];    /* 135 + 10    */
    char exp_date[10];      /* 145 + 10    */
    char exp_time[10];      /* 155 + 10    */
    char hist_un0[3];       /* 165 + 3     */
    int views               /* 168 + 4     */
    int vols_added;         /* 172 + 4     */
    int start_field;        /* 176 + 4     */
    int field_skip;         /* 180 + 4     */
    int omax, omin;         /* 184 + 8     */
    int smax, smin;         /* 192 + 8     */
};

struct dsr
{
    struct header_key hk;   /* 0 + 40      */
    struct image_dimension dime; /* 40 + 108   */
    struct data_history hist; /* 148 + 200  */
};                          /* total= 348 bytes*/

/* Acceptable values for datatype */

#define DT_NONE            0
#define DT_UNKNOWN        0
#define DT_BINARY         1
#define DT_UNSIGNED_CHAR  2
#define DT_SIGNED_SHORT   4
#define DT_SIGNED_INT     8
#define DT_FLOAT          16
#define DT_COMPLEX        32
#define DT_DOUBLE         64
#define DT_RGB            128
#define DT_ALL            255

typedef struct
{
    float real;
    float imag;
} COMPLEX;
```

Comments

The header format is flexible and can be extended for new user-defined data types. The essential structures of the header are the **header\_key** and the **image\_dimension**. The required elements in the **header\_key** substructure are:

**int sizeof\_header**     Must indicate the byte size of the header file.

<b>int extents</b>	Should be 16384, the image file is created as contiguous with a minimum extent size.
<b>char regular</b>	Must be 'r' to indicate that all images and volumes are the same size.

The **image\_dimension** substructure describes the organisation and size of the images. These elements enable the database to reference images by volume and slice number. Explanation of each element follows:

```

short int dim[]; /* array of the image dimensions */
    dim[0]  Number of dimensions in database; usually 4
    dim[1]  Image X dimension; number of pixels in an image row
    dim[2]  Image Y dimension; number of pixel rows in slice
    dim[3]  Volume Z dimension; number of slices in a volume
    dim[4]  Time points, number of volumes in database.
char vox_units[4]  specifies the spatial units of measure for a voxel
char cal_units[4]  specifies the name of the calibration unit
short int datatype /* datatype for this image set */
/* Acceptable values for datatype are */
#define DT_NONE          0
#define DT_UNKNOWN       0 /*Unknown data type*/
#define DT_BINARY        1 /*Binary (1 bit per voxel)*/
#define DT_UNSIGNED_CHAR 2 /*Unsigned character (8 bits per
voxel)*/
#define DT_SIGNED_SHORT  4 /*Signed short (16 bits per voxel)*/
#define DT_SIGNED_INT    8 /*Signed integer (32 bits per voxel)*/
#define DT_FLOAT         16 /*Floating point (32 bits per voxel)*/
#define DT_COMPLEX       32 /*Complex (64 bits per voxel; 2
floating point numbers)
#define DT_DOUBLE        64 /*Double precision (64 bits per
voxel)*/
#define DT_RGB           128 /* */
#define DT_ALL           255 /* */
short int bitpix; /* number of bits per pixel; 1,8, 16,32, or 64. */
short int dim_un0; /* unused */
float pixdim[];   Parallel array to dim[], giving real world measurements in mm.
                  and ms.
    pixdim[1];    voxel width in mm.
    pixdim[2];    voxel height in mm.
    pixdim[3];    slice thickness in mm.
float vox_offset; byte offset in the .img file at which voxels start. This value can
                  be negative to specify that the absolute value is applied for
                  every image in the file.
float calibrated Max, Min specify the range of calibration values
int glmax, glmin;  The maximum and minimum pixel values for the entire database.

```

The **data\_history** substructure is not required, but the **orient** field is used to indicate individual slice orientation and determines whether the Movie program will attempt to flip the images before displaying a movie sequence.

- orient:** slice orientation for this dataset.
- 0 transverse unflipped
  - 1 coronal unflipped
  - 2 sagittal unflipped
  - 3 transverse flipped
  - 4 coronal flipped
  - 5 sagittal flipped