

Facial Soft Tissue Segmentation

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Medizinischen Fakultät

der Universität Basel

von

Tahir Majeed

aus Karachi, Pakistan

Basel, 2014

Genehmigt von der Medizinischen Fakultät

auf Antrag von

Prof. Dr. Philippe C. Cattin, Basel

Dissertationsbetreuer

Prof. Dr. Jean-Philippe Thiran, Lausanne

Koreferent

Prof. Dr. med. Dr. med. dent. Dr. h. c. Hans-Florian Zeilhofer, Basel

Externer Gutachter

Basel, den February 21, 2014

Prof. Dr. med. Christoph Beglinger

Dekan

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Abstract

The importance of the face for socio-ecological interaction is the cause for a high demand on any surgical intervention on the facial musculo-skeletal system. Bones and soft-tissues are of major importance for any facial surgical treatment to guarantee an optimal, functional and aesthetical result. For this reason, surgeons want to pre-operatively plan, simulate and predict the outcome of the surgery allowing for shorter operation times and improved quality. Accurate simulation requires exact segmentation knowledge of the facial tissues. Thus semi-automatic segmentation techniques are required.

This thesis proposes semi-automatic methods for segmentation of the facial soft-tissues, such as muscles, skin and fat, from CT and MRI datasets, using a Markov Random Fields (MRF) framework. Due to image noise, artifacts, weak edges and multiple objects of similar appearance in close proximity, it is difficult to segment the object of interest by using image information alone. Segmentations would leak at weak edges into neighboring structures that have a similar intensity profile. To overcome this problem, additional shape knowledge is incorporated in the energy function which can then be minimized using Graph-Cuts (GC). Incremental approaches by incorporating additional prior shape knowledge are presented. The proposed approaches are not object specific and can be applied to segment any class of objects be that anatomical or non-anatomical from medical or non-medical image datasets, whenever a statistical model is present.

In the first approach a 3D mean shape template is used as shape prior, which is integrated into the MRF based energy function. Here, the shape knowledge is encoded into the data and the smoothness terms of the energy function that constrains the segmented parts to a reasonable shape.

In the second approach, to improve handling of shape variations naturally found in the population, the fixed shape template is replaced by a more robust 3D statistical shape model based on Probabilistic Principal Component Analysis (PPCA). The advantages of using the Probabilistic PCA are that it allows reconstructing the optimal shape and computing the remaining variance of the statistical model from partial information. By using an iterative method, the statistical shape model is then refined using image based cues to get a better fitting of the statistical model to the patient’s muscle anatomy. These image cues are based on the segmented muscle, edge information and intensity likelihood of the muscle. Here, a linear shape update mechanism is used to fit the statistical model to the image based cues.

In the third approach, the shape refinement step is further improved by using a non-linear shape update mechanism where vertices of the 3D mesh of the statistical model incur the non-linear penalty depending on the remaining variability of the vertex. The non-linear shape update mechanism provides a more accurate shape update and helps in a finer shape fitting of the statistical model to the image based cues in areas where the shape variability is high.

Finally, a unified approach is presented to segment the relevant facial muscles and the remaining facial soft-tissues (skin and fat). One soft-tissue layer is removed at a time such as the head and non-head regions followed by the skin. In the next step, bones are removed from the dataset, followed by the separation of the brain and non-brain regions as well as the removal of air cavities. Afterwards, facial fat is segmented using the standard Graph-Cuts approach. After separating the important anatomical structures, finally, a 3D fixed shape template mesh of the facial muscles is used to segment the relevant facial muscles.

The proposed methods are tested on the challenging example of segmenting the masseter muscle. The datasets were noisy with almost all possessing mild to severe imaging artifacts such as high-density artifacts caused by *e.g.* dental fillings and dental implants. Qualitative and quantitative experimen-

tal results show that by incorporating prior shape knowledge leaking can be effectively constrained to obtain better segmentation results.

To
my beloved wife
Annum Pathan
and
my dear son
Muhammad Ahsham Siddiqui

Acknowledgements

The work in this thesis has been carried out at the Medical Image Analysis Center, in the Medical Faculty of the University of Basel, Switzerland. It has been funded by a CO-ME/NCCR the research network of Swiss National Science Foundation (SNSF).

I would like to thank my supervisor Prof. Dr. Philippe C. Cattin for his guidance and supervision all through my Ph.D. It has been motivating and exciting to work under his supervision due to his pleasant and friendly personality. I have had the chance to learn a lot from him as he commands profound knowledge and deep insight into a vast variety of areas. I am deeply grateful for the confidence he has shown in me and the support he has given me at times when the picture looked bleak. I would specially like to thank Dr. Ketut Fundana for guiding me at each and every juncture of my PhD. His knowledge and command over the topics related to my work are remarkable, and I have had the chance to witness them time and again.

I am also grateful to all the project partners Prof. Dr. Thomas Vetter, Dr. Marcel Lüthi, Dr. Mauricio Reyes, Dr. Shamim Kamal, Dr. Hungmin Kim, Prof. Dr. Klaus Scheffler, Prof. Dr. Oliver Bieri and Ms. Jinxia Zhu for sharing their technological skills and knowledge, which helped me complete my work successfully. I would specially like to mention Mr. Jörg Bienemann for providing me the very important medical data needed for my work. Without it, I would not have been able to complete my PhD.

Special thanks and acknowledgements go to all the group members of Medical Image Analysis Center (MIAC) for all the help and support they have given me. I am grateful to them for creating a friendly, creative and wonderful environment at the workplace, which made learning all the more enjoyable. There was always someone to help me out when I needed it. It

was a pleasure working with them. I would like to show my gratitude to all the teachers who have taught me, as it was due to their efforts that enabled me to reach here.

I would like to thank my wife and adorable son for the patience, support and understanding they have shown all through my PhD. Finally, I would like to express my heartfelt gratitude for my parents, who dreamt the dream of my PhD, and without whom none of this would have ever been possible.

Contents

List of Figures	ix
List of Tables	xiii
Glossary	xv
1 Introduction	1
1.1 Background	1
1.2 Image Segmentation	8
1.2.1 Segmentation Based on Energy Minimization	11
1.2.2 Shape Prior Based Segmentation	15
1.3 Problem Statement	19
1.4 Contributions	20
1.5 Outline	21
2 Background	23
2.1 Medical Background	23
2.1.1 Facial Muscles	23
2.1.2 Imaging Artifacts	27
2.1.3 Challenges and Possible Solutions	27
2.1.4 Hounsfield Units of Different Soft and Hard Tissues	30
2.2 Markov Random Field	31
2.2.1 Image Segmentation Formulation as an MRF	33
2.2.2 Optimizing the MRF	37
2.2.3 Graph-Cut	38
2.2.4 Min-Cut Max-Flow Algorithm	40

CONTENTS

2.3	Statistical Model	41
2.3.1	Principal Component Analysis (PCA)	42
2.3.2	Probabilistic PCA	45
3	Approach A: A Shape Prior-Based MRF Model for 3D Masseter Muscle Segmentation	47
3.1	Introduction	47
3.2	Segmentation Model	48
3.2.1	Intensity Energy Function	49
3.2.2	Shape Prior Energy Function	50
3.3	Experimental Setting and Results	52
3.4	Conclusion	58
4	Approach B: Using a Flexibility Constrained 3D Statistical Shape Model for Robust MRF-Based Segmentation	61
4.1	Introduction	61
4.2	PPCA based Statistical Shape Model	62
4.2.1	Reconstruction from Partial Information	63
4.2.2	Remaining Variance	64
4.3	Segmentation Model	65
4.3.1	Integrating Shape Prior Term	66
4.3.2	Initial Shape from the Landmarks	67
4.3.3	Shape Optimization	68
4.3.4	Constraining the Variability	71
4.4	Algorithm	71
4.5	Experimental Setting and Results	72
4.6	Conclusion	80
5	Approach C: Graph Cut Segmentation using a Constrained Statistical Model with Non-Linear and Sparse Shape Optimization	83
5.1	Introduction	83
5.2	Segmentation Framework	84
5.3	Adaptive Shape Optimization	85
5.3.1	Shape Cost Function	85
5.3.2	Shape Optimization	86

5.4	Algorithm	88
5.5	Experimental Setting and Results	90
5.6	Conclusion	94
6	Facial Soft Tissue Segmentation	95
6.1	Introduction	95
6.2	Segmentation of Facial Soft-Tissues	95
6.2.1	Step 1 - Head Separation	97
6.2.2	Step 2 - Skin Layer Segmentation	99
6.2.3	Step 3 - Bone Segmentation	100
6.2.4	Step 4 - Brain Segmentation	100
6.2.5	Step 5 - Air Segmentation	102
6.2.6	Step 6 - Fat Segmentation	103
6.2.7	Step 7 - Facial Muscle Segmentation	103
6.3	Experimental Setting and Results	109
6.4	Conclusion	110
7	Conclusion	113
7.1	Discussion	113
7.2	Future Work	116
	Bibliography	119
	Curriculum Vitae	131

CONTENTS

List of Figures

1.1	Mimic muscles which are small, thin, overlapping and are difficult to identify.	5
1.2	Facial soft-tissues present in a human face; muscle (red), skin (light green), bone (gray), fat (blue), air (brownish-green).	6
1.3	Physical representation of voxels.	8
2.1	Facial muscles.	23
2.2	Individual facial muscles have been highlighted.	25
2.3	Imaging artifacts.	28
2.4	Challenges in facial muscle segmentation. (a,c) Original dataset. (b) Thin, adjacent muscles (Mentalis, Depressor Labii Inferioris, Depressor Anguli Oris) with inhomogeneous interior. (d) The red region shows the facial muscles, skin, tongue <i>etc.</i> all of which have an overlapping intensity profile. The skin has also been outlined in green and tongue in yellow. The image contains most of the muscles mentioned earlier and as can be seen the boundaries between the muscles and with the skin and tongue is invisible.	29
2.5	Hounsfield Unit values for different soft-tissues and bones.	30
2.6	MRF model (figure has been modified).	32
2.7	Neighborhood systems in 3D.	34
2.8	Graph-Cut model.	40

LIST OF FIGURES

2.9	Min-Cut Max-Flow Algorithm (the figure is inspired by Fig. 12.6 in [1]). (a) shows the original graph. The flow and the capacity of each edge is shown by $[0/0]$, (b,c,d,e,f) show the paths from the source to the sink. A path is shown in blue while the red color shows the saturated edges, (g) shows the edges in the cut set which separates the source nodes in green from the sink nodes in orange. The max-flow of the network is 22.	41
2.10	The first two main eigenvectors of the statistical model of the masseter muscle and the features of the shape they influence.	44
3.1	Masseter muscle shown in green.	48
3.2	Probability maps.	51
3.3	Graphs show comparison results between App. A (yelloish-green curve) which incorporates the shape prior in both the data and smoothness terms is more accurate compared to App. F which incorporates it in just the smoothness term (gray curve).	53
3.4	Qualitative segmentation results in 2D. The top row (a-d) shows the segmentation result in blue, while the bottom row (e-h) shows the shape prior in red. The overlaid green in both rows represents the ground truth.	55
3.5	Qualitative segmentation results in 3D of left muscle (a-d) of App. A shown in blue and overlaid in orange is the ground truth.	56
3.6	Qualitative comparison of the segmentation results in 3D of App. A with that of App. F over noisy datasets. The top row (a-d) shows segmentation results of App. A, while the bottom row (e-h) shows the results of App. F for the same noisy datasets.	58
3.7	Segmented masseter muscle in a clinical routine.	59
4.1	Normalized variances of the statistical model.	65
4.2	Probability maps.	67
4.3	Landmarks location: (a) in blue on the mean shape, (b) in red on the dataset.	68
4.4	Linear cost function.	69
4.5	In red are the vertices of the mesh in an axial slice which is a weighted combination of negative log-likelihood map of object C_{obj} , inverse edge map C_{edge} and unsigned distance map C_{seg}	70

4.6	Segmentation process.	73
4.7	The evolution (left to right) of the shape prior (green) to the target muscle (orange) on a corrupted dataset. (a-d) in 3D and their corresponding 2D slices (e-h).	74
4.8	The graphs show dice coefficient, sensitivity and specificity results. Incorporating the shape prior in the regional term (App. B1 shown with blue curve) is more accurate compared to incorporating it in the boundary term (App. B2 shown in green and App. F in red). The red curve shows the result of App. F where a fixed shape template is used.	75
4.9	Shape prior and the segmentation result in 2D. (a-c) show shape prior (green) overlaid with ground truth (orange), while (d-f) show the segmentation (blue).	77
4.10	Segmentation result of the left (a-c) and the right muscles (d-e) in 3D.	78
4.11	High density imaging artifact corrupted slice (a) with ground truth in orange overlaid in all the figures while the segmentation result is in white in (b,c). Segmentation result with the shape prior in boundary term (App. B2) in (b) is worse than the shape prior in regional term (App. B1)(c).	79
5.1	Generating non-linear cost function $\bar{C}(v)$	86
5.2	Normalized standard deviations and their corresponding penalty.	87
5.3	Normalized variance of the statistical model.	88
5.4	Segmentation process.	89
5.5	(a-d) Qualitative segmentation result in 2D where red is the ground truth and green is the segmentation boundary. (e,h) Qualitative segmentation result in 3D where ground truth is in gray and segmentation is in blue.	90
5.6	Quantitative segmentation results.	91
5.7	The shape update in different datasets. The top row (a...d) shows the initial shape contour in blue while the final shape contour in red. The ground truth is shown in green in the top row. The middle row (e...h) depicts the initial 3D mesh of the masseter muscle in blue while the final transparent masseter muscle mesh is shown in red. The bottom row (i...l) shows the same mesh overlaid in the dataset.	93

LIST OF FIGURES

6.1	Head region and background separation.	96
6.2	Workflow for head region and background separation.	97
6.3	Skin layer segmentation shown in green.	98
6.4	Skin layer segmentation workflow.	99
6.5	Bone segmentation in cream color.	100
6.6	Brain segmentation. (a) Original dataset, (b) anisotropic diffusion filtering, (c) edge detection and gradient thresholding, (d) morphological and arithmetic operation, (e) brain segmentation in 2D and (f) brain segmentation in 3D.	101
6.7	Brain segmentation workflow.	102
6.8	Air segmentation.	102
6.9	Fat background dataset.	104
6.10	Fat segmentation in blue.	105
6.11	Facial muscle background dataset.	106
6.12	(a) Facial muscle template mesh, (b) skull mesh, (c) warped template mesh.	107
6.13	Probability map for voxels belonging to the object.	107
6.14	Facial muscle segmentation in red.	108
6.15	Facial soft-tissues segmentation in comparison to each other as well as the skull.	111
6.16	Facial soft-tissues segmentation. Facial muscle in red, fat in blue, skin in green, and brownish orange is the background dataset.	112
7.1	Combined quantitative segmentation results.	114

List of Tables

2.1	Hounsfield units of different soft-and hard-tissues.	31
3.1	This table lists the mean & standard deviation, the median and the smallest as well as the largest value of the dice coefficient of App. A in comparison to App. F.	54
3.2	Quantitative comparison of App. A with that of App. F.	57
4.1	Quantitative comparison of App. B1 with the approaches App. B2 and App. F.	76
4.2	The table list the mean & standard deviation, the median and the smallest as well as the largest value of the dice coefficient of methods App. B1, App. B2 and App. F.	76
5.1	Quantitative comparison of App. C with App. B1 and App. F.	92
5.2	The table list the mean & the standard deviation, the median as well as the smallest and the largest dice coefficient of App. C, App. B1 and App. F.	92
7.1	Comparison of the proposed methods App. A, App. B1 and App. C with App. F.	113
7.2	Statistical comparison of the proposed methods (App. A, App. B1 and App. C) with the approach App. F.	116

GLOSSARY

Glossary

2D	two-dimensional	GVF	Gradient Vector Flow
3D	three-dimensional	HU	Hounsfield Unit
ACM	Active Contour Model	i.i.d.	Independent and Identically Distributed
AR	Augmented Reality	ICM	Iterated Conditional Modes
BP	Belief Propagation	MAP	Maximum-A-Posteriori
CE	Cross Entropy	MAP-MRF	Maximum-A-Posteriori-Markov Random Field
CMF	Cranio-Maxillofacial	MFA	Mean Field Annealing
CSF	Cerebrospinal Fluid	MRF	Markov Random Field
CT	Computed Tomography	MRI	Magnetic Resonance Imaging
DC	Dice Coefficient	MTM	Mass Tensor Model
DRG	Diagnosis-Related Group	n-links	Neighboring-Link
DTI	Diffusion Tensor Imaging	OR	Operation Room
EM	Expectation-Maximization	PCA	Principle Component Analysis
FEM	Finite Element Model	PPCA	Probabilistic Principal Component Analysis
GC	Graph-Cuts	RF	Radio Frequency
GM	Gray Matter	RL	Relaxation Labeling
GNC	Graduated Non-Convexity	ROI	Region of Interest
GRF	Gibbs Random Field	SA	Simulated Annealing
		SSM	Statistical Shape Model
		t-links	Terminal-Links
		US	Ultrasound
		WM	White Matter

GLOSSARY

1

Introduction

This chapter introduces the background material and the importance of why facial soft tissue segmentations is required. The chapter then delves into different segmentation methods listing their advantages and disadvantages and finally presenting state-of-the-art segmentation methods. The chapter concludes with the definition of the problem statement and contributions of the thesis.

1.1 Background

Face being the unique external perceivable identity of a human is of fundamental importance. It represents the individual personality and traits in daily contact with other humans. It is the area where some of the most important sense organs are present. Underneath the facial skin is the musculo-skeletal system consisting of bones and soft-tissues. Facial muscles are the main source for conveying our expressions to others. The importance of the face for socio-ecological interaction is the reason for a high demand for any surgical intervention on the facial musculo-skeletal system. Bones and soft-tissues are of major importance for any facial surgical treatment to guarantee an optimal, functional and aesthetical outcome. Cranio-Maxillofacial (CMF) surgery such as osteotomies, bone fragment relocation, restoration of bone defects and inserting implants, is used to remove the abnormalities in the facial hard-tissues. CMF surgery is needed in cases such as (1) inborn malformations like cleft lips and palate in new born babies, (2) treatment of severe facial injuries, (3) treatment of dysgnathia where the jaws are repositioned in the facial skeleton, or (4) resection of the facial soft and hard-

1. INTRODUCTION

tissues in ablative tumor surgery for treating defects and abnormalities in hard-tissues of the face, jaws and neck. Premature fusion of cranial structures also result in the craniofacial deformities.

Due to the importance of the face in interpersonal relationships, any minor changes to the facial anatomy that have negative impact on the facial appearance can portend disastrous consequences for the patient; therefore, surgeons want to pre-operatively predict the outcome of the CMF surgery [2]. If the preoperative prediction of the facial surgery is available, it would help the patient decide whether he/she wants to have the surgery or not. The surgeons can also optimally plan the surgery which will result in shorter operation time and improved surgical outcome.

Segmentation of the facial soft-tissues is an important step of the facial surgery planning and simulation [3, 4, 5]. Olszewski *et al.* [6] describe a prototype of a computer-assisted CMF planning system which can be used for preoperative planning and intra-operative navigation. The system provides 3D visualizing of the patient's specific mandible model, planning of the surgery and performing preoperative simulations. The planned data is then used in Operation Room (OR) using Augmented Reality (AR). AR provides real-time guidance on how the surgery is progressing or to check the location of the surgical instrument on the reconstructed image dataset of the patient. The authors have not incorporated any facial soft-tissue segmentation and simulation model except the facial skin and dental nerve. If only the facial hard-tissue data is considered during the CMF surgery planning, the result is functionally good but of poor aesthetic appearance [7]. Gladilin *et al.* [4] on the other hand incorporated soft-tissue simulation. The tissue deformations were calculated according to physics-based biomechanical models using Finite Elements to simulate facial expressions in craniofacial surgery planning. Their model assumes that different soft-tissues have similar properties. Mazza & Barbarino [8] presented one of the most accurate numerical 3D Finite Element Model (FEM) that provides a faithful representation of the facial anatomy for simulations. Gladilin *et al.* [4] and Olszewski *et al.* [9] have used the manual segmentation of the facial soft-tissues using specialized software. Mollemans *et al.* [10, 11] conducted studies to address the demands of the surgeons for patient-specific preoperative soft-tissue prediction models using Mass Tensor Model (MTM) [12, 13]. As it can be seen that CMF surgeries require accurate planning based on available pre-operative diagnostic datasets. Three-dimensional imaging, planning, simulation and

navigation tools are needed for the comparison of different surgical strategies and the prediction of the outcome.

The facial soft-tissue segmentation is also useful in a wide range of applications. Simulation of tissue deformations for facial expressions using FEM was also applied in gaming and animated character movies [14]. Sifakis & Selle [15] used 3D face models obtained from manually segmented MR scans and blended muscle activation signals with motion captured animations to simulate speech. Zhang *et al.* [16] used a 3D facial biomechanical model where skin motion was constrained using skull structures for animating expressions. Facial models of the bones and soft-tissues are also important in forensics and anthropology where De Greef & Willems [17] surveyed the existing techniques and found that the nose projection, eye protrusion and mouth width in addition to 3D craniofacial reconstructions can be used to identify an unknown person. Dental treatments also make use of facial models and reconstruction of the head. Zepa *et al.* [18] showed correlations between facial soft-tissues sizes and the dental patient groups by calculating the volume of the masseter muscle, medial pterygoid muscle and mandible which were manually segmented. The segmentation of the masseter muscle which is the thickest muscles in the human face is required as it plays a vital role in chewing and facial expressions [19, 20, 21, 22, 23, 24, 25]. The morphometric analysis of the segmented facial soft-tissues can help in diagnosing craniofacial neuromuscular disorders and subsequently devise strategies to treat patients [26, 27]. Farrugia *et al.* [26] performed manual measurements to quantify the length, width, area and volume of different facial soft-tissues such as tongue, masseter, pterygoid, orbicularis oris, temporalis and buccinator while Fischer *et al.* [27] performed measurements on the manually segmented masseter, temporal and medial & lateral pterygoid muscles.

All the aforementioned applications require the segmentation of the facial soft-and hard-tissues from which patient-specific anatomical models can be created. The current clinical practice is to manually segment the facial soft-and hard-tissues, which is a tedious and time consuming task, requires a medical expert (an anatomist or a radiologist) who uses his experience to identify and segment different anatomical structures. It results in increased overall costs for the surgery that according to today's Diagnosis-Related Group (DRG) system are not fully reimbursed. It means that no one pays for the cost of the time a doctor has spent in planning if he decides to perform a navigated surgery. This is why surgeons are looking for automatic segmentation

1. INTRODUCTION

methods that can replace the tedious and time consuming tasks. Thus, there is a dire need for automatic or semi-automatic algorithms to segment the facial soft-tissues.

Some software tools have been developed for CMF surgery planning. Meehan *et al.* [28] developed a software tool “Craniofacial Surgery Planner” to perform patient-specific virtual surgeries pre-operatively. The tool provides interactive computer-assisted craniofacial plastic surgery planning and visualization together with non-linear simulation of soft-tissue changes due to bone cutting and realignment. The creation of the patient-specific 3D models of the bone and soft-tissues was done manually. Many other commercial facial surgery planning products like SurgiCase from Materialise, Belgium, Maxilim from Medicim, Belgium or BrainLab, Germany/USA incorporate soft-tissue simulation but sacrifice accuracy over computational complexity.

The segmentation and simulation support of the soft-tissues is still in its infancy; therefore, the segmentation of the facial soft-tissues from preoperative diagnostic datasets such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) is an active research area. The automatic segmentation of the facial soft-tissues is not easy, as the facial region is one of the most complex parts of the human body. The facial muscles, like mimics muscles (Fig. 1.1^{1 2}), are small, thin and often overlapping each other, which is why they are difficult to identify. The segmentation task is further complicated by the presence of imaging noise, weak edges, inhomogeneous interior of the anatomical structure, high density metal artifacts and multiple anatomical structures in the close vicinity having an intensity profile which overlaps with that of the target structure. The mild or severe partial volume artifacts that depend on the image resolution which causes the voxels at the boundary of different tissue types to have intensities in between the tissue mean values [30], further add to the complication of the automatic facial soft-tissue segmentation.

The existing CMF surgery simulation models treat the space between skin and skull as homogeneous tissue during simulations. As mentioned earlier, the human face is one of the most complex parts of the human body which consists of muscles (red), blood vessels, nerves, glandular structures, skin (light green), bones (gray), bone marrow, fat (blue), ligaments, cartilages, tendon and air/sinus (brownish-green) as shown in

¹Fig. 1.1(a) printed with kind permission of Springer Science+Business Media. Source: Fig 1.1 of [29]

²Fig. 1.1(b) printed with permission. Copyright: CLIPAREA — Custom media/Shutterstock

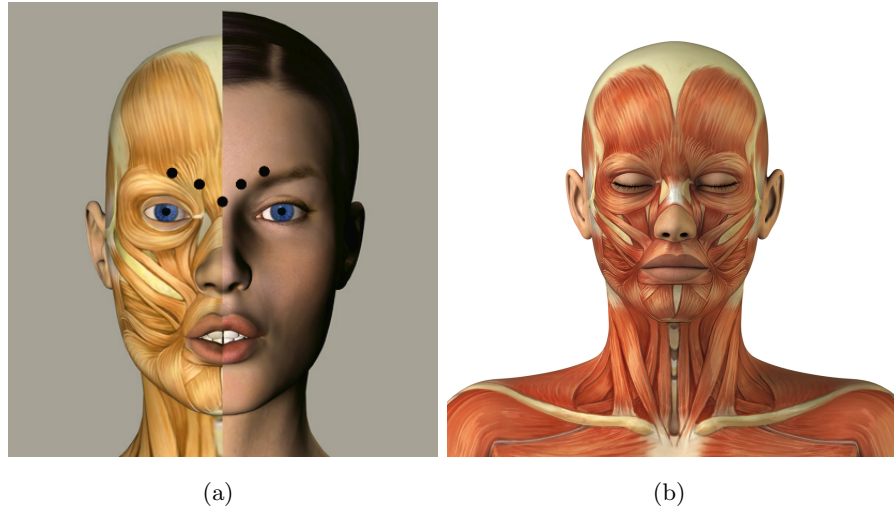


Figure 1.1: Mimic muscles which are small, thin, overlapping and are difficult to identify.

Fig. 1.2. The size, shape, location and orientation of these structures differ in the population with each anatomical structure having a different coefficient of elasticity. The over-simplification of the soft-tissue model by treating the space as homogeneous leads to inaccurate simulation results [9].

There have been only few studies trying to automatically segment the human head into different tissue classes [14, 30, 31, 32]. Shattuck *et al.* [31] were interested in the segmentation of the brain (Gray Matter (GM), White Matter (WM) and Cerebrospinal Fluid (CSF)), and prior to the segmentation of the brain they segmented the human face into brain and non-brain regions called *Skull Stripping* using a sequence of low-level operations such as anisotropic diffusion filtering, edge detection, and mathematical morphology. While Shattuck *et al.* concentrated mainly on segmenting the different structures in the brain and labeled the rest of the dataset as non-brain, Tasdizen *et al.* [32] went one step further and automatically classified the human head into nine tissue classes: GM, WM, CSF, blood vessels and sinuses, eyes, bone, bone marrow, muscle and fat tissue from multispectral MR scans. They used a combination of a scalar and multispectral Expectation-Maximization (EM) clustering algorithm to learn the parameters like mean, variance of the different tissue classes from the multispectral intensity histogram and then used a series of low-level image processing operations like flood-fill, thresholding and morphology to classify the different tissues. The drawback

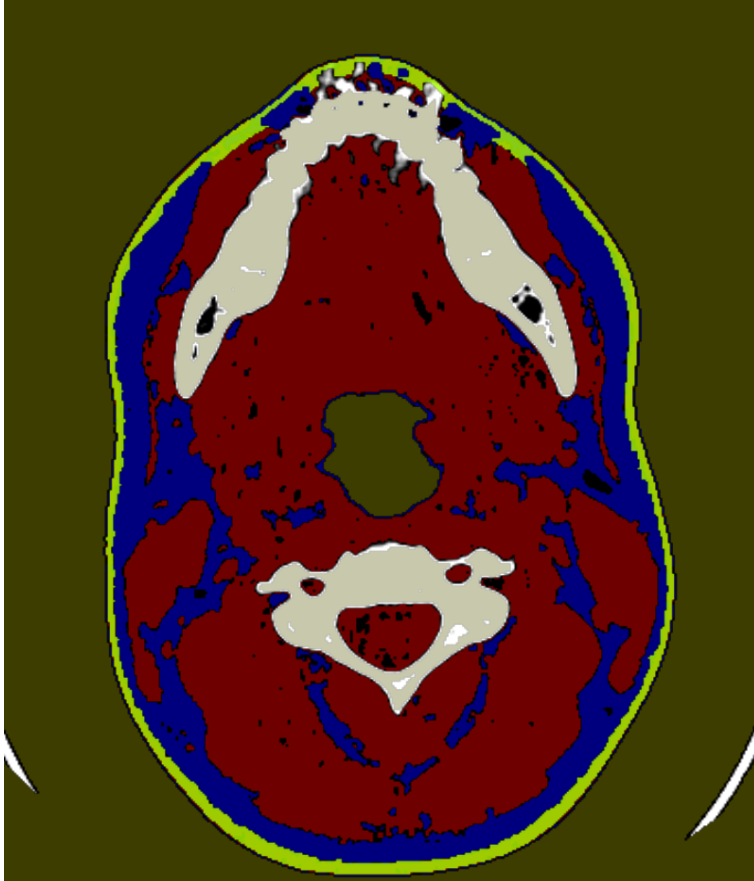


Figure 1.2: Facial soft-tissues present in a human face; muscle (red), skin (light green), bone (gray), fat (blue), air (brownish-green).

of this approach is that a large amount of manually segmented images are required to train the EM algorithm.

Rezaeitabar & Ulusoy [33] have proposed a specialized region growing based segmentation method to segment the two facial muscles; masseter and temporalis. They defined a local energy function using the intensity similarity with the neighboring pixels from within the dataset and other training datasets, that were rigidly registered to each other and the test dataset. The spatial interactions were modeled as Markov Random Field (MRF). Although they claimed the method to be completely automatic, it required a starting seed point, the selection of the starting and ending slices by the user and a threshold parameter.

Ng *et al.* have written a series of papers [22, 23, 24, 25, 34, 35, 36] for the segmenta-

tion of different facial muscles. Ng *et al.* [24, 25] segmented masseter, lateral & medial pterygoid from MRI datasets in a learned Region of Interest (ROI) in a two-dimensional (2D) slice. They have used low-level image processing algorithms to exclude bone and fat. An initial estimate of the muscle boundary was obtained using a muscle template and morphological operators. The final muscle segmentation was then refined using a Gradient Vector Flow (GVF) Snake based approach starting from the initial estimate of the muscle boundary. The first and last slices in the image stack had to be manually selected by the user. They segmented the muscles in a 2D slice one at a time and then combined the results of the individual segmentations to form a 3D volume. To segment a 3D structure by repeatedly segmenting it in 2D may result in poor segmentation as the target structure might have weak boundaries in some of the slices. This problem can be resolved if the complete structure is segmented in 3D as one object. In another semi-automatic segmentation approach, Ng *et al.* [34] required the user to manually segment the masseter muscle in the starting and ending slice as well as in the shape deterministic slices in between the ends. They then performed shape based interpolation to build the masseter model. The method was improved further with a 2D automatic segmentation refinement step [35]. Ng *et al.* [36] provided another coarse segmentation refinement approach by matching the intensity distribution and Ng *et al.* [22] proposed a similar approach to the previous one using fuzzy C-mean clustering. They proposed in [23] a further method for the segmentation of the temporalis muscle using a similarly learned ROI as in [24, 25] but using low-level image processing algorithms (thresholding and morphology) to remove fat, brain tissue and muscle tendon.

Kale *et al.* [30] proposed another automatic facial tissue segmentation method, including very thin anatomic structures, based on the Bayesian and Level-sets framework by fusing CT and MRI datasets. The CT and the MRI of the patients were registered to each other. The soft-tissues they have segmented include fat, muscle, skin and bone. The skin, air and bones were segmented using low-level image processing operations: thresholding, flood fill, morphology and image subtraction. They have mainly used variants of Level-set methods and Bayesian approaches to segment muscle and fat from the dataset. For the segmentation of thin muscles, Kale *et al.* have incorporated the modeling of the partial volume effect into the Level-set and Bayesian approaches. The inference over the underlying probabilistic model was done using Iterated Conditional Modes (ICM). They manually computed the mean intensity of the four classes and the

1. INTRODUCTION

standard deviation of the image noise for the Bayesian method and the MRI and CT resolution difference has been ignored.

Sener [14] has proposed extensions to the Bayesian approach of Kale *et al.* to obtain robust facial soft-tissue segmentation. In particular, he proposed the addition of a directional prior to get a better segmentation of thin muscles and the resolution deblurring model for CT and MRI. As it is hard to distinguish different facial muscles, the strategy used by Sener is to segment the facial muscles in an ROI which involves a lot of manual interaction. In addition, both CT and MRI are required for each patient.

The main purpose of fusing the CT and MRI datasets is to make use of different contrast information contained in both the modalities in the hope that it will help in better segmenting different facial soft-and hard-tissues. CT provides higher resolution images than MRI whereas MRI provides better soft-tissue contrast than CT but CT provides better contrast for bone and air than MRI. Fusing CT and MRI together involves the registration of both datasets. The registration of CT and MRI is prone to error and there is bound to be a mismatch between the modalities which will introduce error. Furthermore, to register two different modalities which have different inplane resolution/spacing and slice thickness one modality will have to be converted to have the same resolution/spacing as the other one, which will further introduce partial volume artifacts.

1.2 Image Segmentation

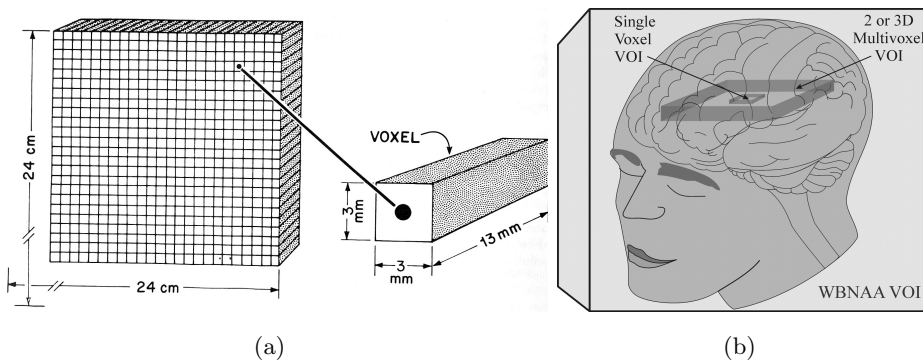


Figure 1.3: Physical representation of voxels.

Image segmentation is the process of separating an image into at least two non-intersecting mutually exclusive constituent parts such that each part is uniform & homogeneous and the union of two adjacent parts is non-homogeneous [37] with respect to some characteristic such as image intensity [38]. It assigns a label to every pixel of an image where pixels sharing the same label fulfill the homogeneity criterion. The aim is to simplify an image into a more meaningful and easier to analyze representation. The image segmentation algorithms consider the two basic features of image intensity; discontinuity and similarity [39] to identify regions that correspond to different anatomical structures. A medical image is made up of a collection of measurements or image intensities such as radiation absorption in X-Ray, acoustic pressure in ultrasound (US), or radio frequency (RF) signal amplitude in MRI [38]. These measurements are taken at discrete locations in 2D (Eq. 1.1) called pixels or in three-dimensions (3D) called voxels (see Fig. 1.3¹²). Let $I_{M \times N}$ be an image with size M and N , then a digital image can be represented as

$$I_{M \times N} = [f(x, y)]_{M \times N}, \quad (1.1)$$

where $f(x, y)$ is the intensity value at spatial coordinates (x, y) . Let Ω be the domain of the image, then the segmentation problem can be defined as finding the set $\{S_k \subset \Omega\}_{k=1, \dots, K}$ whose union covers the entire domain Ω as given by

$$\Omega = \bigcup_{k=1}^K S_k, \quad (1.2)$$

where $S_k \cap S_j = \emptyset$ for $k \neq j$, and each S_k is connected.

With the introduction of MRI, CT, US and other imaging modalities, it is now easier to view the internal anatomy of a patient and differentiate healthy tissue from a diseased one. As more and more data with larger size and finer resolution become available to the doctors, it is becoming costlier to analyze all the data. It is, therefore, imperative that computer algorithms and techniques be developed that can automatically/semi-automatically delineate anatomical structures of interest to the doctors to assist in specific radiological tasks. These algorithms are normally used for calculating tissue volumes, diagnosis, identifying tumorous regions, visualization of specific anatomical

¹Source Fig. 1.3(a): https://wiki.uiowa.edu/download/attachments/40534414/IMG_0093.jpg

²Printed with permission; Source Fig. 1.3(b) [40]: <http://www.ajnr.org/content/28/10/1843/F2.large.jpg>

1. INTRODUCTION

structures (alone or in comparison to other structures), treatment planning and computer assisted surgery. Over the last 20 years, numerous techniques have been proposed both in the continuous and discrete domains [37, 38] for medical image segmentation [41].

Thresholding is one of the simplest techniques, where each pixel in the image is considered alone and assigned to one or the other class based on intensity value [39, 42]. The drawback of thresholding algorithms is that they ignore the spatial information of an image [37], which makes them sensitive to noise and intensity inhomogeneities [38]. Region growing [43] considers both the intensity and the spatial information of a pixel, where the segmentation process starts from a user selected seed point and additional pixels are added to the segmented region if they are adjacent and are within the intensity threshold or meet some predefined homogeneity criterion [38, 39, 41]. Hojjatoleslami & Kruggel [44] applied region growing for the segmentation of brain lesions. Region split and merge [45] is another approach in which the image is first subdivided into arbitrary disjoint regions, which are then either merged together or split based on the homogeneity criteria [39]. Morphological Watershed [46] combines many of the concepts of the previous three approaches and thus produces more stable segmentation results as compared to previous approaches. Here the image is considered to be a 3D surface where x -, y -coordinates form the 2D plane and the intensity value at location x, y is considered as the heights of the surface. On such a surface, local minimum points can be identified and a hole is punched in each local minimum. From these holes, water is filled at a uniform rate which starts to fill up the regions. A dam is built at locations where the rising water from different catchments is about to meet. The algorithm terminates when everything has been submerged and only the tops of the dams are visible. These dam boundaries are those of different regions in the image [39]. Hahn & Peitgen [47] used the 3D Watershed Transform, for example, for skull stripping. Then, there are clustering techniques like K-Means, Fuzzy C-Means and Expectation Maximization (EM) clustering, which iteratively alternate between segmenting the image and characterizing the properties of the segmented classes. The clustering techniques assume that the number of classes and their initial parameters should be known before the first iteration [38].

These techniques and their variants are computationally efficient and are widely used for simple segmentation task. As these techniques are based mainly on the inten-

sity information, they do not provide robust segmentation; for example region growing is prone to leaking through thin boundaries between different structures, and thresholding can not be applied if the background has an overlapping intensity profile with the target object. They either require substantial human guidance to accurately segment anatomical structures from medical imaging datasets or are completely unusable as medical image datasets suffer from imaging artifacts [48].

Image segmentation is in general an ill-posed problem as the solution is not unique [49]; therefore, additional constraints are needed in order to achieve the desired solution. Mostly image intensity based constraints are used which are for example based on intensity or edge information, and sometimes regularization constraints are used. Prior knowledge about the shape and location of the target structure or any prior knowledge about the object of interest can also be used to achieve higher segmentation accuracy. In medical image analysis, a lot of prior information about the shape, appearance and location of the target anatomical structure is available, for example, in the CT volume dataset. The intensity profile of anatomical structures such as bones, muscles or blood vessels remain relatively constant. Also, the anatomical location of the target structure remains the same relative to the other anatomical landmark structures. All this prior information can be used to effectively constrain the solution space of the segmentation. Prior shape information can significantly help in avoiding ambiguities [50] that may arise due to noise, weak edges, an inhomogeneous interior of the object, artifacts and multiple anatomical structures in the close vicinity of the target object with overlapping intensity profiles to the target object while segmenting anatomical structures. There are segmentation techniques based on energy minimization that incorporate low-level image information (such as intensity and edge information) together with high-level information (shape, size, location of target structure) into the energy function. These techniques are discussed in the following section.

1.2.1 Segmentation Based on Energy Minimization

The previous approaches proposed for image segmentation were mainly based on heuristics. Recently image segmentation problems have been formulated in the form of optimizing an energy cost function which provides a principled method to segment objects [51]. This energy cost function is based on smoothness constraints, prior knowledge, image intensity and edge information or other regional and boundary criteria

1. INTRODUCTION

which are encoded in the form of different terms in the energy function. Minimizing the energy of the cost function then results in the segmentation of the object. This minimization provides the optimum balance between different terms of the energy function.

To minimize the energy of the cost function, energy optimization methods are needed. Depending on what optimization method is selected, the resulting segmentation could be trapped in a local minimum and then the optimization process does not provide the global optimal result. These energy minimization techniques can be differentiated on the basis of the spatial domain they are applied in: spatially continuous and spatially discrete [51]. Continuous domain techniques use variational optimization methods, which guarantee to find (local) minima if the energy function is non-convex. Discrete domain optimization techniques use combinatorial or stochastic optimization methods which guarantee to find the global minimum only if the energy function can be shown to be submodular [52]. Segmentation techniques based on energy minimization are better equipped to deal with image artifacts and image noise as they allow, in a mathematical framework, to seamlessly integrate prior knowledge about the target anatomical structure.

Spatially Continuous Minimization Techniques

Continuous domain minimization techniques, like Active Contour models [53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65] and Stable Mass Spring Models [66, 67], formulate the segmentation problem in the domain of continuous functions. These techniques define an energy cost function incorporating low-level image information (for example, intensity and edge information), together with high-level object knowledge (for example shape and size information). The final solution is obtained by optimizing the cost function using variational or gradient descent approaches. Variational and gradient descent approaches guarantee to find only local minima of the corresponding non-convex energy cost function. The local minimum obtained could also be a global minimum but achieving the global minimum relies solely upon the initialization value provided to the cost function. If the initialization is good and close to the global minimum, which is not a trivial task, these techniques can achieve the global minimum (although this is not guaranteed); otherwise, they are susceptible to getting stuck in a local minimum [68]. A more detailed classification of the Active Contour models can be found in [69].

Active Contour models, also called Deformable models, can be divided based on the form of the contour representation. The ones that use an explicit representation of the contour are called Snakes [53, 54, 55, 56], which are based on physical principles. They had an enormous impact on the segmentation community by formulating the segmentation as an energy minimization problem. These methods are defined by closed parametric curves or surfaces that deform to delineate an object boundary, under the influence of weighted internal and external forces. The internal forces are represented by the smoothing constraints imposed on the parametric curve to keep the curve smooth during deformation while the external forces are derived from the image so that the curve is pulled towards the desired feature of interest such as edge gradient [38]. The curve deformation stops when equilibrium has been reached between the internal and external forces. Their advantage over the low-level techniques is that they are robust to weak edges and can easily generate closed curves from the image [38]. A review of deformable model based techniques can be found in Jain *et al.* [70] and McInerney & Terzopoulos [71]. The drawback of deformable models is that they have to be initialized near the desired boundary and the parameters have to be perfectly tuned for the desired image. The explicit parameterization of the curve requires reparameterization to avoid self-intersection of control points and does not allow the segmentation of multiple objects in a straightforward manner. Generally, heuristics are used to allow topological changes in the curve. The Snake approach does not have a probabilistic basis; therefore, there is no straight forward way to extend it to segment objects based on color, texture or motion as mentioned by Cremers [51].

Techniques such as Level-set methods [57, 58, 59, 60, 61, 62, 63, 64, 65] use an implicit representation of the contour to circumvent the drawback associated with the Snakes formulation. In particular, they address the problems of topological changes, contour parameterization and control point regriding [51] and can be generalized to hypersurfaces in contrast to Snakes which are difficult to generalize to higher dimensions. They have been shown to be less sensitive to noise and tend to have less local minima as compared to Snakes making them less sensitive to initialization. The Bayesian formulation of the Level-set methods allow incorporating prior shape knowledge in a principled manner. The resulting energy function consisting of different terms corresponding to the boundary, region, shape or other information derived from the image can be minimized using a gradient descent strategy.

1. INTRODUCTION

The energy function of the Active Contour models in general is non-convex, which means initialization is critical. Recently, work has been done to convexify the energy function, such as [72, 73, 74, 75, 76]. Chan *et al.* [72], for instance, proposed a scheme which provides a global minimum of Chan and Vese [57] by relating together the image segmentation and image denoising problems. Bresson *et al.* [73] based on the work of Chan *et al.* [72] proposed global minimization models and developed fast numerical schemes for contour evolution. These methods could efficiently obtain global solutions, as are generally preferred to make the method independent of the initialization.

Spatially Discrete Minimization Techniques

Spatially discrete minimization techniques formulate the segmentation problem as a combinatorial or stochastic optimization in finite dimensional space R^n . These techniques define a discrete energy cost function and optimizing this cost function provide minima which are an approximate solution to the continuous problem. Spatially discrete minimization techniques like Graph-Cuts (GC) [77, 78] guarantee to find the global minima only if the corresponding energy functions can be shown to be submodular or cross entropy [79, 80], which does not have any restriction on the convexity or submodularity of energy function.

Techniques that guarantee a global optimum are superior and more attractive than techniques that guarantee only a local optimum as global solutions providing better segmentation are preferred over local solutions. Global methods are more reliable and robust than local methods, and global solutions are more stable than local solutions. Furthermore, the final segmentation result is also independent of the initialization value. If the solution obtained is not good, it is the direct cause of a bad energy function instead of a numerical problem during minimization. Tweaking the energy function is the most obvious and easiest approach to get better solutions [68].

The GC approach is simple and combines the best characteristics of vision based methods like global minimum, practical efficiency and numerical robustness. It blends together visual cues and constraints and provides freedom from topological constraints of the target object [68]. GC are also easily extendable to n -dimensional (higher than 2-dimensions) images, which is the normal domain for medical image segmentation. Before GC, there were techniques like active contours [53], shortest path algorithms [81, 82], ratio regions [83], and some other segmentation methods [84] that guaranteed a

global minimum, but they were restricted to 2D problems [68]. Although Stable 3D Mass Spring Models [66, 67, 85, 86] are easily extendable to n -dimension, they only provide locally optimum solutions.

MRF is a discrete domain statistical model framework which provides an elegant and harmonious way to unify diverse constraints and prior information from different domains by modeling spatial interactions of the neighboring pixels into the energy function [49]. The spatial interactions generally encoded in medical images assume that most pixels belong to the same class as their neighboring pixels. Thus the probability of a pixel belonging to a class different from the neighboring pixels is low. The underlying statistical model of the MRF framework is Bayesian probability; therefore, in order to segment a structure, the posteriori probability has to be maximized [38]. Many algorithms have been proposed during the last two decades to maximize the posterior probability such as; Simulated Annealing (SA), ICM, Belief Propagation (BP) [87, 88], GC and many others. The details of the MRF framework are discussed in Sec. 2.2.

Although MRFs hold a lot of promise due to their flexibility in combining information from different domains elegantly, they did not receive much attention due to the lack of efficient discrete optimization techniques. With the introduction of GC by Greig *et al.* [77] and Boykov & Jolly [89], which guarantees global minimum for certain class of functions [52]¹, MRF based models came back strongly. GC are a class of algorithms which are based on max-flow min-cut algorithms [90] to solve discrete optimization problems.

1.2.2 Shape Prior Based Segmentation

Traditional GC approaches are very successful in finding the global minimum in cases where the object of interest can be distinguished from its adjacent structures; however, they are vulnerable to misleading information that results from image noise, weak boundaries and the presence of adjacent structures having similar intensity profile as the target structure [91, 92]. This makes the segmentation of anatomical structures from medical image a challenging task. Thus, there is a dire need for adding additional information into the MRF based energy function which is challenging but can make

¹(see Kolmogorov *et al.* [52] for a description of what energy functions can be minimized through GC)

1. INTRODUCTION

the segmentation of anatomical structures robust. It is now common practice to incorporate prior shape information into MRF based energy functions to constrain the segmentation [92, 93, 94, 95, 96, 97, 98]. It is, however, very difficult to incorporate prior shape knowledge into the GC based approach. Techniques that can successfully incorporate prior shape knowledge into GC are promising because they can be used to effectively constrain the leaking together with achieving a global optimal solution.

There are two fundamental approaches to integrating the shape priors into GC. Either they are integrated into the smoothness term or into the data term of the MRF based energy function. Veksler [95], Freedman & Zhang [92] and Das *et al.* [99] have all incorporated shape prior in the smoothness term, while others like El-Zehiry & Elmaghraby [100], Freiman *et al.* [94], Ali *et al.* [97], Slabaugh & Unal [91], Vu & Manjunath [101] and Malcolm *et al.* [102] have incorporated it in the data term of the energy function. It is, however, not an exhaustive list of the methods incorporating shape priors into GC as it is not within the scope of this thesis.

Das *et al.* [99] focused on segmenting a particular class of objects by incorporating compact shape priors, assuming that the target object has a compact shape. A compact shape is one that fulfills the criteria where certain configuration of label assignments to neighboring pixels are prohibited. Foreground and background regional penalties were modeled using parametric distributions. An additional bias parameter was added to get larger object boundaries and to nullify the effects of standard GC where the smoothness term is biased towards shorter object boundaries. If the label assignments to neighboring pixels did not comply with the rules of the compact shape, the smoothness term was exceptionally increased to ensure that such label assignments were avoided. The class of shapes that can be segmented with such shape priors is, however, limited.

Veksler [95] used a star shape prior to segment a wider class of objects than that of Das *et al.* [99]. A shape prior was defined as star convex if every point inside the object could be connected to its center with a line that was completely inside the shape. This class of objects includes all convex shapes in addition to some other non-convex shapes. However, the shapes that can be segmented using the star-convex shape prior of Veksler is still limited. The idea was that for any point p , if assigned the label 1 then all the points q which lay on the line segment that connects p to c (the center of the

star shape) must also had the label 1, which was ensured by making the neighboring edge weights of all pixels lying on the line segment exceptionally large.

Freedman & Zhang [92] incorporated the shape prior as an aligned fixed shape template. Their method used an unsigned level-set distance map where the 0 level-set corresponded to the shape template. The use of the unsigned level-set distance map allowed them to express arbitrary shapes instead of the restricted class of shapes like the ones discussed earlier. They assumed that the shape template was perfectly aligned in areas where the object boundary was weak; the 0 level-set encoded the prior knowledge about the object boundary in such areas. Such an approach requires the use of a fixed shape template and may not work well in general for anatomical structures possessing large inter-patient shape variation. Furthermore, an accurate registration of the shape template with the target object is a difficult task [95]. They, however, claimed that their method is robust to slight misalignment or small shape variations. The use of the shape prior in the smoothness term has a further drawback. In medical imaging, where the datasets may be noisy and corrupted by artifacts, the use of the shape prior in the boundary term smooths out large chunks of the object which are corrupted by the artifacts.

Slabaugh & Unal [91] showed that the results of GC can be improved and made more robust by incorporating parametric models such as the elliptical shape priors. Elliptical shape priors helped in restricting the solution space of the segmentation result. The parametric shape was initialized around the object of interest and then the method tried to find the object boundary in a narrow band with an iterative refinement process. A weighted shape prior was incorporated in a hard constraint manner. The results showed that the method was able to accurately segment blood vessel even when other blood vessels were adjacent. It was shown that without shape prior knowledge, the segmentation leaks into the adjacent blood vessel. Others like Zhu-Jacquot [103] used geometric shape priors similar to the elliptical shape priors of Slabaugh & Unal [91] to incorporate prior shape knowledge. Zhu-Jacquot [103] applied an EM-style approach where he used the geometric shape prior to get better segmentation and then used the segmentation to update the shape parameters. In the case of the elliptical shape prior, an initial bad estimation could cause the segmentation to diverge instead of converging to a stable solution by incorrectly approximating the shape prior for the next iteration. Zhu-Jacquot [103] provided an elegant solution to this problem using the entropy of

1. INTRODUCTION

the segmented object, which should be as low as possible, making the segmented object more homogeneous. These approaches tried to spatially constrain the GC segmentation by incorporating parametric shape information. The parametric techniques have the drawback that only shapes that can be roughly represented parametrically or geometrically can be segmented, which is often not the case in medical image segmentation, particularly, when segmenting structures in 3D. The task is further complicated by the presence of different pathologies [104].

Statistically learned shape knowledge can be incorporated in the energy function as has been shown by El-Zehiry & Elmaghraby [100], Freiman *et al.* [94] and Malcolm *et al.* [102]. Freiman *et al.* [94] and El-Zehiry & Elmaghraby [100] used a non-parametric global shape prior. They used an EM-style approach, where the intensity and shape models were optimized in the E-step, and then performed the segmentation in the M-step making use of the optimized shape and intensity model. In an iterative refinement process, the shape and intensity models were corrected which finally converged to the target shape with increasing segmentation accuracy. The shape prior was constructed by aligning manual segmentations of the target object. Freiman *et al.* [94] created a probability map from weighted shapes in the training dataset while El-Zehiry & Elmaghraby [100] created an unweighted probability map, which defined what pixels were more likely to belong to the object as compared to the others. The weighting factor of the sample shapes depended on the dice coefficient of the sample shape and the current segmentation. These methods are prone to generating invalid shapes as there is no statistical dependence between the shapes. Malcolm *et al.* [102] incorporated non-linear shape priors learned from the Kernel Principle Component Analysis (PCA), which did not suffer from statistical independence and iteratively refined the shape prior in a Bayesian setting until it converged to the target shape. The shape prior was refined by fitting it in the feature space to the segmentation. The pre-image of the fitted shape prior in the input space was computed and then the updated shape prior was used to obtain better segmentation in the next iteration. Statistically learned shape knowledge can represent all sorts of shapes but requires a large amount of training samples. The proposed techniques have been shown on segmenting 2D shape. Although Malcolm *et al.* method uses non-linear shape priors, the computed pre-image at best is an approximation. A simpler model, such as statistical shape model based on PPCA, could also be used to achieve comparable results which will also rule out generation of invalid

shapes not represented by the statistical model. Vu & Manjunath [101] incorporated prior shape knowledge through shape distance in multi-phase GC to segment multiple objects from an image.

Ali *et al.* [97] incorporated a shape prior in the data term of the energy function. The shape prior was learned by aligning and registering the shapes in the training dataset with the target object. After registration of the sample shapes, three different regions were defined as common object, common background and variability region. The common object and common background regions were used as shape priors in a hard-constraints manner similar to that of Boykov *et al.* [78], while the shape prior, in the variability region, was incorporated as a distance probabilistic model. The distances were defined as iso-contours of the common object region. These distances were converted into a probability map. The probability map defined the probability of each iso-contour to belong to an object, which decreased exponentially from the common object region to the common background region. Poisson distribution was used to model the intensity distribution of the object and the background. The kidney was segmented slice-wise and then stacked together to get a 3D segmentation. The problem with such an approach is that there is a model of pixel probabilities at certain locations which are treated independently with respect to each other and do not guarantee reasonable shapes. The pixel probabilities are not connected to each other to guarantee a valid shape.

1.3 Problem Statement

In order to obtain an accurate facial surgery simulation, a patient-specific 3D soft-tissue model needs to be created. To create the patient-specific 3D soft-tissue model, the soft-tissues have to be segmented. This thesis, reports the development of a semi-automatic segmentation technique.

The aim of this study was to develop a semi-automatic facial soft-tissue segmentation method, using CT datasets. Facial soft-tissues addressed in this thesis include the segmentation of skin, fat and facial muscles. Since muscle, blood vessels, glandular tissues and tendons possess similar and overlapping intensity values in CT [30], they are treated as muscles class. Although the skin possesses a similar intensity value compared to that of muscles, it is segmented as a different class as it is important for facial

1. INTRODUCTION

surgery simulations. Skull, bones and teeth are treated as the “bone” class. The main classes for the segmentation are thus Skin, Fat and Facial Muscles.

Segmentation is an ill-posed problem as has been mentioned earlier. This is especially true for the face where the partial volume effect and multiple similar objects are responsible for many local minima in the optimization function. To overcome the problem with local minima and adjacent objects having a similar gray value distribution, GC is combined with statistical shape models. In this thesis, all the facial muscles are segmented as one object, as it is sufficient for the facial simulation model.

1.4 Contributions

The main contribution of this thesis is the incorporation of statistical shape knowledge in the MRF based framework for the segmentation of the facial soft-tissue, which has been demonstrated on the segmentation of the masseter muscle. As the standard GC cannot be used, because of the challenges outlined earlier, the standard GC energy function is augmented with shape knowledge to obtain the desired segmentation results. In the first approach, a fixed shape template is used and additional shape knowledge is incorporated in the data term of the energy function published in Majeed *et al.* [19].

As a fixed shape template is not versatile enough to capture the shape variability of the anatomical structures naturally found in the population, the fixed shape template is replaced in the second approach with a variability constrained Probabilistic Principal Component Analysis (PPCA) based statistical model. The model also allows to reconstruct the optimal shape and to compute the remaining variance of the statistical model from partial information. In an iterative framework, the statistical model is fitted to the segmentation and the fitted model is used as a shape prior in the next iteration. The process is repeated until the segmentation converges. The proposed method has been published in Majeed *et al.* [20].

The statistical model shape update mechanism in the previous approach is linear. In comparison to linear shape update mechanism, where all the vertices are weighted equally irrespective of their constrained variability, a better approach, that provides better fitting and more accurate shape updates, is to use a non-linear shape update mechanism. In a non-linear update mechanism, vertices that have higher variability incur less penalty compared to vertices that have lower variability if they move the

same distance from their mean positions. The proposed approach has been published in Majeed *et al.* [21].

1.5 Outline

The thesis is organized as follows: Chapter 2 outlines the background knowledge needed to understand the MRF framework shows and how GC is used to segment objects from image dataset as statistical shape prior knowledge is incorporated into GC. Thus the chapter also outlines how the statistical shape model is constructed. The chapter furthermore provides background knowledge about the facial soft-tissues. It discusses what facial muscles are segmented, the challenges that arise in segmenting the facial soft-tissues, specifically the imaging artifacts found in the datasets. Chapter 3 outlines the first approach to segmenting the facial soft-tissue, which is tested on segmenting the masseter muscle using a fixed shape template. In Chapter 4 the fixed shape template is replaced with a PPCA based statistical model incorporated into GC. This approach is also tested on segmenting the masseter muscle. The next chapter describes how the method is augmented by replacing the linear shape update mechanism of the statistical model with a non-linear shape update mechanism which provides a more robust and accurate shape fitting of the statistical model. Although the proposed approaches are tested on segmenting the masseter muscle, these approaches are general enough and can be applied to segment any object, whenever a statistical model is available. In Chapter 6, finally, a method is proposed to segment the facial soft-tissues (skin, fat and muscles) using GC which is incorporated with a fixed muscle shape template. Hereby, the facial skin is segmented using low-level image processing operations. To close, the conclusion of the thesis is presented in Chap. 7.

1. INTRODUCTION

2

Background

This chapter provides the background related to the facial anatomy, Markov Random Field and statistical model to understand the work proposed in this thesis.

2.1 Medical Background

2.1.1 Facial Muscles

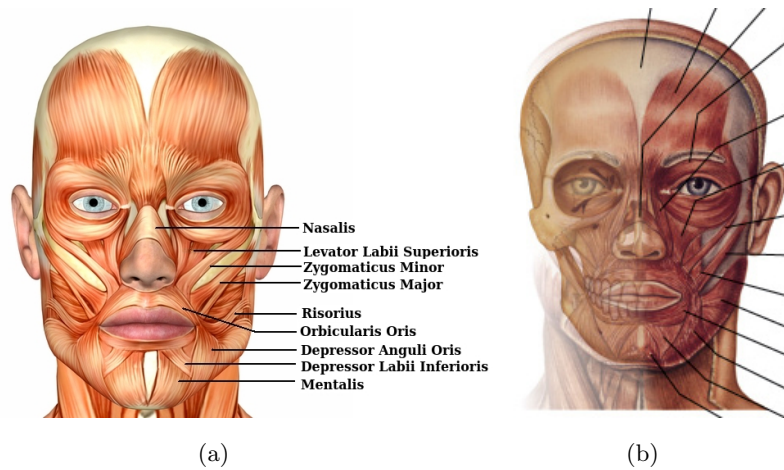


Figure 2.1: Facial muscles.

Facial muscles are a type of skeletal muscles. Skeletal muscles allow doing all the hard work that a human does all through the day like lifting weights and can be controlled voluntarily. These muscles make up the musculo-skeletal system which is a

2. BACKGROUND

combination of bones and muscles and is the main source of power and strength in a human. Figure 2.1¹² shows all facial muscles. The individual facial muscles have been highlighted in Fig. 2.2³. These muscles are attached to bones. Apart from helping in eating, speaking, breathing *etc.*, facial muscles are used to produce different kinds of expressions. Anatomical descriptions of the facial muscles that have been segmented are given below⁴.

- (a) **Mus. Mentalis** is located at the tip of the chin which is a paired central muscle of the lower lip used for raising and pushing the lower lip which in turn causes wrinkles in the chin. The muscle starts from the incisive fossa of the mandible and moves downward to connect to the skin of the chin.
- (b) **Mus. Depressor Labii Inferioris** starts from the oblique line of the mandible where it is continuous with the fibers of the platysma muscle and terminates by inserting into the skin of the lower lip where it joins with the orbicularis oris muscle. This muscle is intermingled with fat and is innervated by the mandibular branch of facial nerve. It also helps in lowering the lower lip.
- (c) **Mus. Depressor Anguli Oris** originates from the oblique line of the mandible, where it is continuous with the fibers of the platysma muscle and terminates at the angle of the mouth, where it is, in turn, continuous with orbicularis oris and risorius. This muscle is used for frowning and is innervated by the mandibular branch of the facial nerve.
- (d) **Mus. Levator Anguli Oris** originates from the canine fossa, immediately below the infraorbital foramen, whose fibers are intermingled with those of the Zygomaticus, Triangularis and Orbicularis oris. It is innervated by the buccal branches of the facial nerve.
- (e) **Mus. Orbicularis Oris** is a sphincter like muscle that encircles the mouth. It consists of fibers from other facial muscles that insert into the lips. It is used in

¹Fig. 2.1(a) printed with permission: Copyright: Sofia Santos/Shutterstock.

² Fig. 2.1(b) printed with permission. Source: http://img.sparknotes.com/figures/E/eb57ee3c0bbce61d887722fc5931002b/muscles_of_the_head.jpg

³The images were obtained from Google Body and were modified. Source Fig. 2.2: <http://www.zygotebody.com>

⁴Source: http://en.wikipedia.org/wiki/Facial_muscles

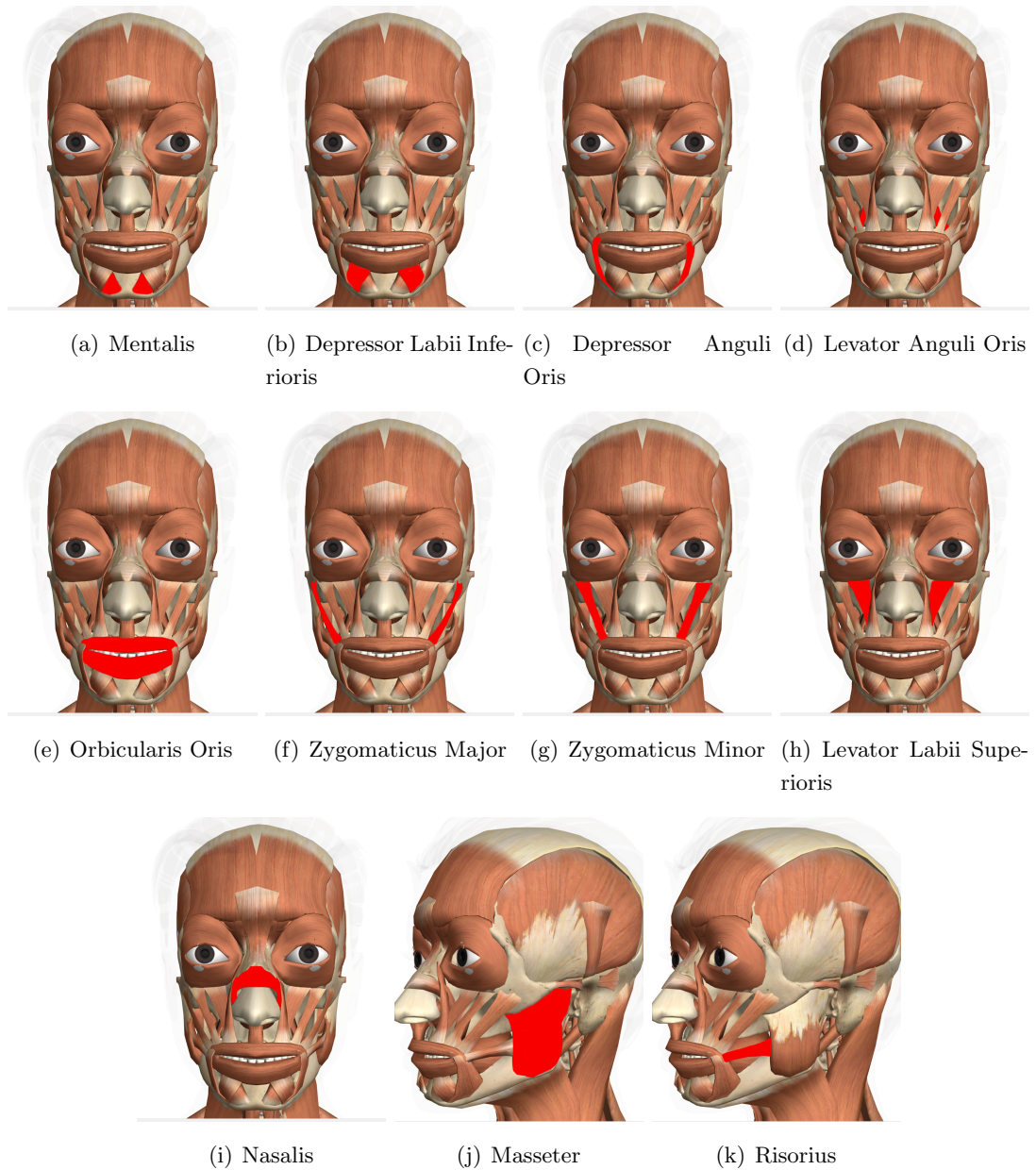


Figure 2.2: Individual facial muscles have been highlighted.

playing brass instruments and for kissing. Its function is to compress and protrude the lips closing the mouth when it contracts.

(f) **Mus. Zygomaticus Major** originates from the zygomatic arch (cheekbone) and

2. BACKGROUND

terminates at the corners of the mouth. It pulls the corner of the mouth upward and backward when smiling or laughing. The variations in the shape and size can cause dimples to appear. It is also innervated by the buccal and zygomatic branches of the facial nerve.

- (g) **Mus. Zygomaticus Minor** originates from the malar bone and moves along with the orbicularis oculi on the lateral face of the Levator labii superioris and terminates by inserting into the outer part of the upper lip. This muscle is used for making sad facial expressions. It is also innervated by the facial nerve.
- (h) **Mus. Levator Labii Superioris** originates from the side of the nose and terminates at the zygomatic bone. Its shape appears like a broad sheet and is used in forming facial expressions.
- (i) **Mus. Nasalis** is a sphincter-like muscle which consists of two parts; a transverse part that originates from the maxilla above and lateral to the incisive fossa, moves upward and medially, shaping into a thin aponeurosis and forms a bridge with the muscles on the opposite side. While the wing-shaped part is attached to the greater wing-shaped cartilage at one end and attached to the integument at the point of the nose at the other end.
- (j) **Mus. Masseter** is responsible for chewing by raising the lower jaw. It is thick, quadrilateral in shape and consists of two parts. The superficial part is large, thick, tendinous aponeurosis which originates from the zygomatic process of the maxilla, and from the anterior two-thirds of the lower border of the zygomatic arch and terminates by inserting into the angle as well as the lower half of the lateral surface of the ramus of the mandible. The deep part is smaller, more muscular and originates from the posterior third of the lower border and from the whole of the medial surface of the zygomatic arch, and terminates by inserting into the upper half of the ramus and the lateral surface of the coronoid process of the mandible. The deep part is hidden from the front by the superficial part and from behind covered by the parotid gland. It is innervated by the mandibular branch of the trigeminal nerve. It is one of the most important muscles in CMF surgery.

- (k) **Mus. Risorius** originates from the fascia over the parotid gland and moves along the platysma and terminates at the angle of the mouth. It is also used in smiling and is innervated by the buccal branch of the facial nerve.

2.1.2 Imaging Artifacts

There are many different types of artifacts that can be observed in CT or MRI datasets [48] which cause the degradation of the image quality, thus making the diagnosis or segmentation task difficult. The two main artifacts present in the datasets used throughout this thesis are the high density artifact and partial volume artifact which are discussed in the following paragraphs.

High Density Artifacts are caused by highly radiation absorbing materials such as metallic implants, dental fillings, dental implants, prosthetic devices and surgical clips. It causes severe streaking (see Fig. 2.3(a)) when such a material appears in the scan field. The streaking appear because of the incomplete attenuation due to the high density of metal which is outside the normal range that a computer can handle [48].

Partial Volume Artifacts are caused by sampling in discrete voxels when multiple tissues are within a voxel [38, 48]. The final voxel intensity is a combination of the intensities of the tissues present in the voxel [105]. This causes blurring of the boundary between the tissues (see Fig. 2.3(c)). Missegmentation can occur when the boundary between fat and bone causes the voxels to have intensities close to the ones of the muscle. The partial volume artifact increases with a decrease in the spatial resolution of the imaging dataset which can cause thin structures to completely blur out. The severity of partial volume artifact can be reduced by increasing the spatial resolution of the dataset.

2.1.3 Challenges and Possible Solutions

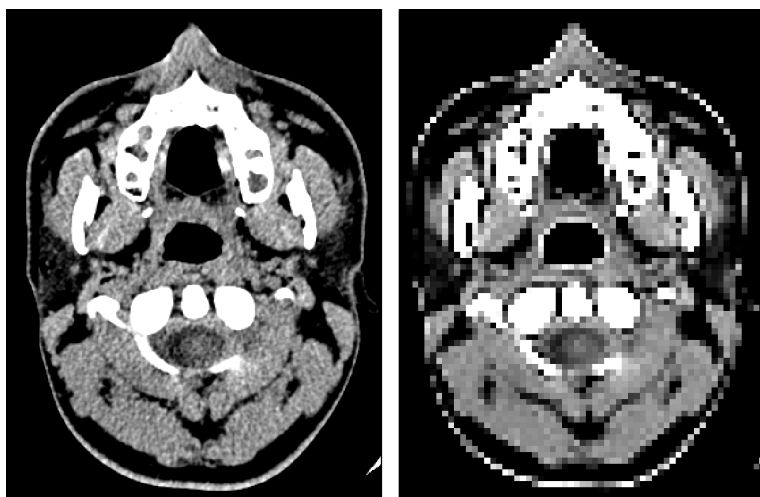
This section provides a short overview of the challenges one faces in facial muscle segmentation which are shown in Fig. 2.4 where some of the target muscles have been outlined.

As depicted in Fig. 2.4(b), the interior of the muscle is not homogeneous. This creates a major problem for techniques that do not guarantee to find the global minimum but can be trapped in local minima. To eliminate the problem, all sorts of ad-hoc mechanisms have been applied [53, 55, 106]. The solutions obtained are not robust

2. BACKGROUND



(a) Streaking caused by high density artifact.



(b) Without artifacts.

(c) With artifacts.

Figure 2.3: Imaging artifacts.

and depend heavily on the initialization value. Techniques that guarantee the global solution are more stable and independent of the initialization value [68].

Figure 2.4(b) shows a region where at least three different facial muscles are present, while Fig. 2.4(d) shows many muscles which are adjacent to each other and to other anatomical structure such as tongue (yellow) and skin (green) which have overlapping intensities with the target muscles. It is very difficult, even for a human, to detect the

boundary between the muscles and other anatomical structures as it is invisible. This is why anatomical shape knowledge is required to segment the individual muscles.

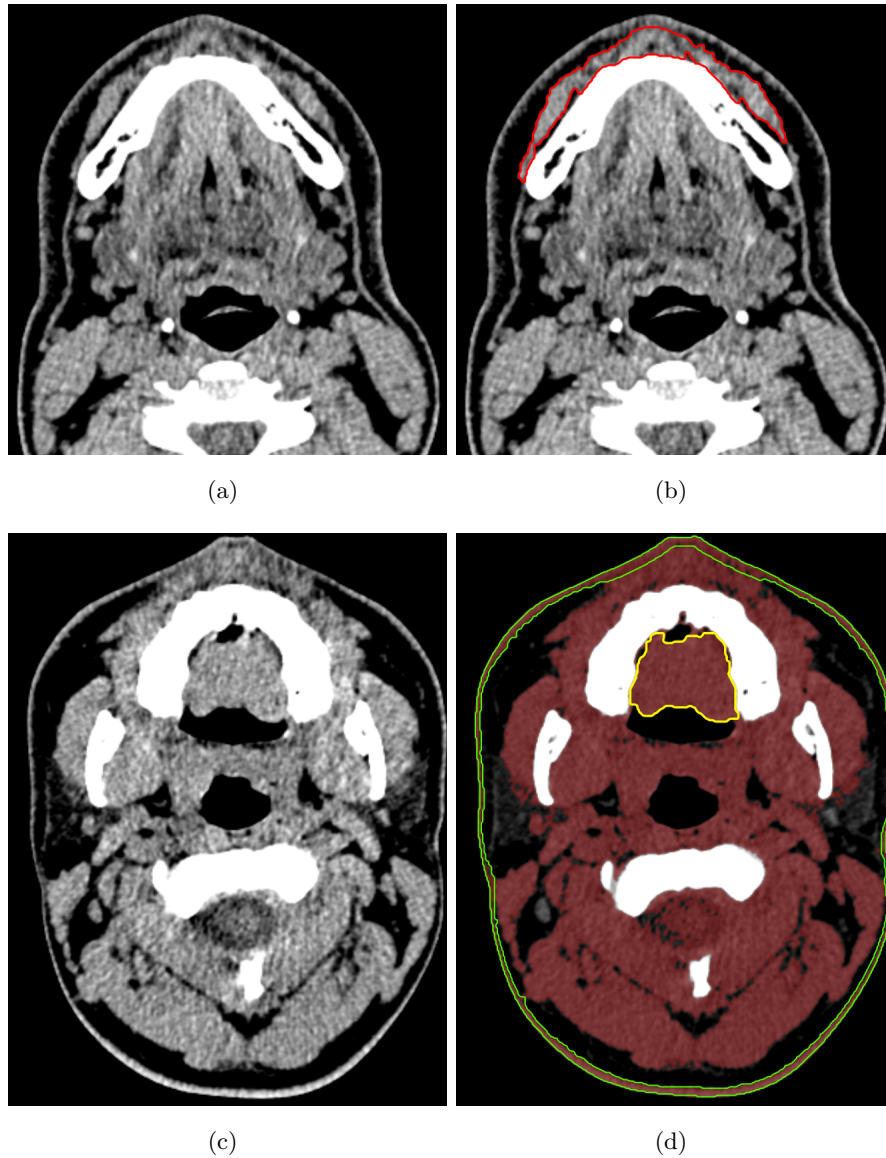


Figure 2.4: Challenges in facial muscle segmentation. (a,c) Original dataset. (b) Thin, adjacent muscles (Mentalis, Depressor Labii Inferioris, Depressor Anguli Oris) with inhomogeneous interior. (d) The red region shows the facial muscles, skin, tongue *etc.* all of which have an overlapping intensity profile. The skin has also been outlined in green and tongue in yellow. The image contains most of the muscles mentioned earlier and as can be seen the boundaries between the muscles and with the skin and tongue is invisible.

2. BACKGROUND

The masseter muscle is the largest muscle in the face and clearly visible in the CT dataset. The rest of the facial muscles are smaller, thinner and difficult to recognize [26]. Influenced by the scanning resolution of the datasets, the muscles boundaries are blurred due to the partial volume effect, which further complicates the segmentation of these muscles. Getting higher resolution images is not always feasible as it increases the amount of harmful radiation to the patient.

2.1.4 Hounsfield Units of Different Soft and Hard Tissues

In this section, Hounsfield units (HU) [105] in CT datasets of different soft-and hard-tissues found within the human body are presented (see Fig. 2.5 and Table 2.1¹). HU is a quantitative scale for describing radiodensity of materials. Water at standard pressure and temperature has a radiodensity of 0 HU while air at standard pressure and temperature has a radiodensity of -1000 HU. The HU scale linearly transforms the original linear attenuation coefficient measurement into the radiodensity where water and air has the earlier given radiodensities. If μ_{water} is the linear attenuation coefficients of water, then the HU of a material with average linear attenuation coefficient μ_X is given by:

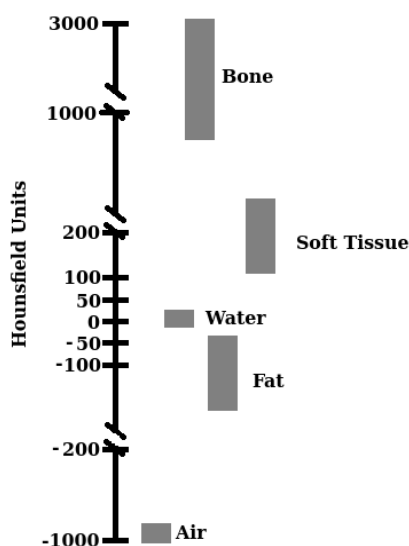


Figure 2.5: Hounsfield Unit values for different soft-tissues and bones.

¹The Table 2.1 is inspired by http://en.wikipedia.org/wiki/Hounsfield_scale

$$HU = 1000 \times \frac{\mu_X - \mu_{water}}{\mu_{water}}. \quad (2.1)$$

It is important to note that the HU scale is only applicable to CT scans and not to

Tissues	HU
Lung	-500
Fat	-50 to -100
GM	-37 to -45
WM	-20 to -30
Water	0
Muscle	+10 to +40
CSF	+15
Blood	+30 to +45
Soft Tissue	+100 to +300
Bone	+700 (spongy bone) to +3000 (dense bone)

Table 2.1: Hounsfield units of different soft-and hard-tissues.

Cone Beam CT scans as they generally do not scan the entire area.

2.2 Markov Random Field

There is a tendency in the computer vision community to develop algorithms based on sound mathematical frameworks. These make it easier to understand their properties as compared to the algorithms based on heuristics presented in the previous chapter. Optimization algorithms based on energy function minimization principle are now commonly used to solve vision problems such as segmentation, texture analysis, surface reconstruction, image restoration, edge detection [107]. Different constraints can be incorporated into the energy function that helps in suppressing the ambiguities and to reduce the search space. Obtaining the solution through optimization is desirable since it is difficult to find the perfect solution due to the inherent uncertainties and ambiguities in the problem [49].

The MRF framework provides a suitable and homogeneous way to model the problem where different types of constraints from a wide variety of sources can be easily incorporated into a unified framework. MRF provides probability distribution over the

2. BACKGROUND

interacting features within the MRF. It is an undirected graphical model that specifies both a factorization of the joint probability distribution of the random variables and a set of conditional independence relations between those random variables. It consists of a set of nodes each of which corresponds to a random variable and a set of undirected links or edges connecting a pair of nodes, Fig. 2.6¹. The MRF theory es-

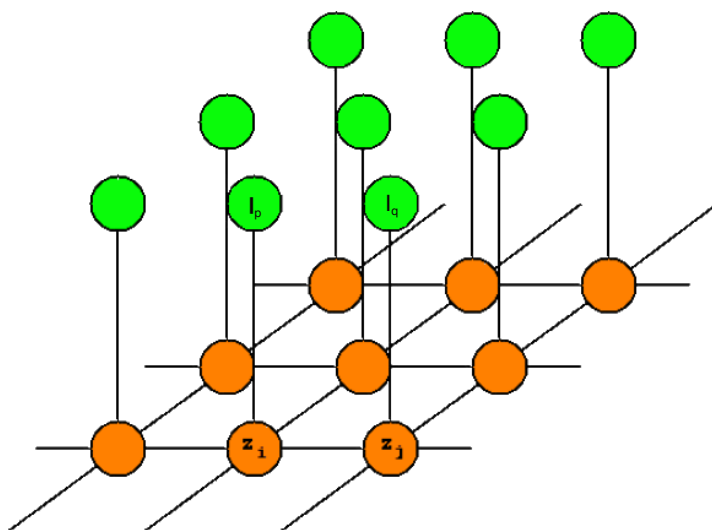


Figure 2.6: MRF model (figure has been modified).

tablishes the Maximum-A-Posteriori (MAP) optimality criteria utilizing methods from estimation and decision theory. The Maximum-A-Posteriori-Markov Random Field (MAP-MRF) framework [108] allows to solve a number of computer vision problems, using a systematic approach instead of ad-hoc heuristics. The objective function in the MAP-MRF framework is the joint posterior probability of the MRF labels whose form and parameters are determined according to Bayes formula. The sound mathematical and theoretical basis of the MRF together with development of efficient and accurate solvers like GC and BP has in recent years lead to the increased popularity of the MAP-MRF framework in modeling computer vision problems [49].

MRF based on the concept of Bayesian image analysis has also made great inroads in the field of medical image processing as it allows to seamlessly incorporate spatial interactions (contextual constraints). It transforms the medical image analysis problem into a mathematical optimization problem, thus making the problem computationally

¹Fig. 2.6 Source: http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/AV0809/ORCHARD/mrf.jpg

manageable [79]. It has been successfully used in many applications in the medical domain such as image enhancement [109, 110], object detection [111], data modeling [112], tissue classification (segmentation) [14, 30, 113].

One advantage of using the MRF frameworks is that if the solution obtained is not good enough then it is evident that it does not have to do with numerical stability but more to do with the objective function or the relative importance of the different terms of the objective function [49].

2.2.1 Image Segmentation Formulation as an MRF

Image segmentation is a labeling problem which can be cast as the MAP estimation of an MRF. Consider an image \mathbf{I} to be segmented. Let $\mathbf{P} = \{p_1, p_2, \dots, p_N\}$ be the set of N pixels/voxels of the volume datasets. Let $\mathbf{L} = \{l_1, l_2, \dots, l_M\}$ be the set of M labels from which a label l_m will be assigned to a specific pixel in the image. In the case of binary image segmentation (object/background), the set $\mathbf{L} = \{0, 1\}$ consists of only two labels 0 for background and 1 for object. The goal of the image segmentation problem is to find an optimal mapping $\phi : \mathbf{P} \mapsto \mathbf{L}$ that maps pixels to labels. In the MRF framework the mapping ϕ is defined as a field $\mathbf{Z} = (Z_1, Z_2, \dots, Z_N)$ of random variables associated with the pixel $p \in \mathbf{P}$. The random variable Z_p takes value from the set of labels \mathbf{L} . Assignment of labels \mathbf{L} to the random field variables is called a labeling which is denoted by the vector $\mathbf{z} = (z_1, z_2, \dots, z_N)$, where z_p is the label assigned to the random variable Z_p which in turn is equivalent to assigning label z_p to the corresponding pixel $p \in \mathbf{P}$. The set of all possible labelings is denoted by \mathbb{Z} and any instance of \mathbb{Z} denoted by \mathbf{z} defines a segmentation. A neighborhood system $\mathbb{N} = \{N_p | \forall p \in \mathbf{P}\}$ is defined over the set of pixels \mathbf{P} , where N_p consists of all the neighbors of the pixel $p \in \mathbf{P}$. Possible neighborhood systems could be 4- or 8-neighborhood system for 2D images and 6-, 18- or 26-neighborhood system for 3D datasets (see Fig. 2.7¹). For the random field \mathbf{Z} to be a valid MRF with respect to neighborhood system \mathbb{N} it must satisfy the following properties:

Positivity :

$$P(\mathbf{z}) > 0, \forall \mathbf{z} \in \mathbb{Z} \tag{2.2}$$

¹Fig. 2.7 Source: <http://www.cs.sunysb.edu/labs/projects/volume/Papers/Voxel/Images/FigureB.gif>

2. BACKGROUND

The positivity property states that the probability P of particular labeling \mathbf{z} should be greater than 0.

Markovianity :

$$P(z_p | \{z_q : q \in \mathbf{P} - \{p\}\}) = P(z_p | \{z_q : q \in N_p\}), \forall p \in \mathbf{P} \quad (2.3)$$

The Markovian property states that the assignment of label z_p to pixel p depends

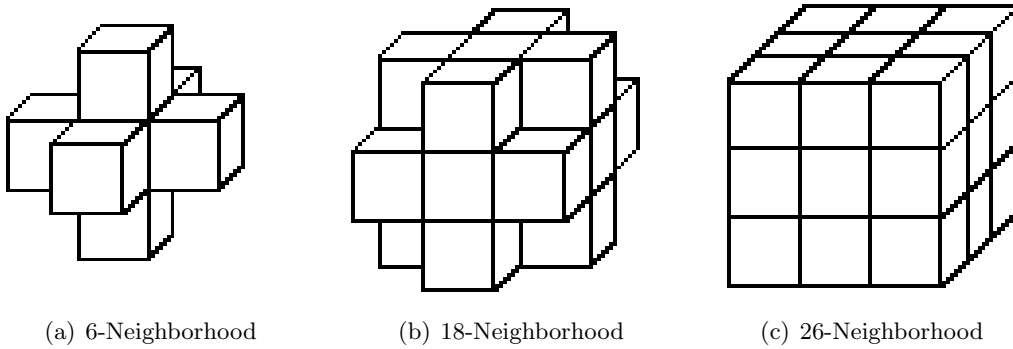


Figure 2.7: Neighborhood systems in 3D.

only on the labels assigned to q the neighbors of pixel p defined by the set N_p . This property generally holds for medical images as mentioned by Chen & Metaxas [114]. After defining the MRF, the posterior probability $P(\mathbf{z}|\mathbf{I})$ of a particular labeling \mathbf{z} given the observed image data \mathbf{I} needs to be defined. Using the Bayesian formulation an *A-Priori* distribution $P(\mathbf{z})$ can be defined over the set of all allowable images. Normally, $P(\mathbf{z})$ is defined as a locally dependent MRF which specifies the prior such that the underlying segmented image \mathbf{z} consists of either smoothly varying gray levels which sometimes change level discontinuously or large homogeneous patches. The likelihood distribution $P(\mathbf{I}|\mathbf{z})$ of the observed image data \mathbf{I} given the labeling \mathbf{z} together with *A-Priori* distribution $P(\mathbf{z})$ can be used in the Bayes formula [77] to define the posterior probability

$$P(\mathbf{z}|\mathbf{I}) \propto P(\mathbf{I}|\mathbf{z}) \times P(\mathbf{z}) \quad (2.4)$$

$$P(\mathbf{z}|\mathbf{I}) = \frac{P(\mathbf{I}|\mathbf{z}) \times P(\mathbf{z})}{P(\mathbf{I})} \quad (2.5)$$

$$P(\mathbf{z}|\mathbf{I}) = \frac{P(\mathbf{I}|\mathbf{z}) \times P(\mathbf{z})}{\sum_j P(\mathbf{I}|\mathbf{z}_j)P(\mathbf{z}_j)} \quad (2.6)$$

The optimal or MAP labeling z^* of the MRF can then be defined as the MAP probability of $P(\mathbf{z}|\mathbf{I})$ given as

$$z^* = \arg \max_{z \in \mathbb{Z}} P(\mathbf{z}|\mathbf{I}) \quad (2.7)$$

The direct optimization of z^* is not possible because the search space is huge, *e.g.* an image size of approx. $33 \times 33 \approx 1000$ pixels is thus not possible. To give the reader an idea of the search space, consider the following explanation of MacKay [115] in his own words: *"Even if each component x_n took only two discrete values, the number of evaluations of P^* would be 2^{1000} , a number that is still horribly huge. If every electron in the universe (there are about 2^{266} of them) were a 1000 gigahertz computer that could evaluate P^* for a trillion (2^{40}) states every second, and if we ran those 2^{266} computers for a time equal to the age of the universe (2^{58} seconds), they would still only visit 2^{364} states. We'd have to wait for more than $2^{636} \approx 10^{190}$ universe ages to elapse before all 2^{1000} states had been visited."*

Hammersley & Clifford in 1971 in their landmark theorem, the Hammersley-Clifford theorem [116], proved that a probability distribution with a positive mass or density can be represented as an MRF as it satisfies one of the Markov properties over an undirected graph \mathcal{G} if and only if it is a Gibbs Random Field (GRF) and derived the necessary and sufficient conditions for it. Thus, an MRF is equivalent to a GRF if the random field \mathbf{Z} obeys the Gibbs distribution. This results in the factorization of a joint probability distribution $P(\mathbf{Z} = \mathbf{z})$ of the random variables \mathbf{Z} over the set \mathbb{C} of all possible cliques¹ (complete subgraphs) of the graph \mathcal{G} . The joint probability distribution of the random variables \mathbf{Z} factorizes into the product of *potential functions* $\psi_c(\mathbf{z}_c) : \forall \psi_c(\mathbf{z}_c) > 0$ defined over the maximal cliques² of the graph [117] which can be written as

$$P(\mathbf{Z} = \mathbf{z}|\mathbf{I}) = \frac{1}{W} \prod_{c \in \mathbb{C}} \psi_c(\mathbf{z}_c|\mathbf{I}) \quad (2.8)$$

where W called the *partition function* is a normalization constant given by

$$W = \sum_{\mathbf{z}} \prod_{c \in \mathbb{C}} \psi_c(\mathbf{z}_c|\mathbf{I}) . \quad (2.9)$$

¹A clique is a graphical concept and is defined as subset of vertices in a graph with each vertex connected to the every other vertex in the subset (see *pg.* 385 of [117]).

²A maximal clique is defined as a clique such that it ceases to be a clique if a vertex from the graph is added to the clique (see *pg.* 385 of [117]).

2. BACKGROUND

As the potential functions $\psi_c(\mathbf{z}_c|\mathbf{I})$ are defined to be strictly positive they can be expressed as exponentials

$$\psi_c(\mathbf{z}_c|\mathbf{I}) = \exp[-E(\mathbf{z}_c|\mathbf{I})] , \quad (2.10)$$

$$E(\mathbf{z}_c|\mathbf{I}) = \sum_{c \in \mathcal{C}} V_c(z_c|\mathbf{I}) , \quad (2.11)$$

where $E(\mathbf{z}_c|\mathbf{I})$ is called a *Gibbs Energy function* and the exponential representation in Eq. 2.10 is called the *Boltzmann distribution* commonly known as *Gibbs distribution* in computer vision [117]. Since the joint probability distribution is defined as the product of the potential functions so the total energy is calculated by the summation of the energies of the maximal cliques potential $V_c(z_c|\mathbf{I})$ as written in Eq. 2.11. The joint probability distribution in Eq. 2.8 can then be written in terms of the Gibbs energy as

$$P(\mathbf{z}|\mathbf{I}) = \frac{1}{W} \cdot \exp[-E(\mathbf{z}_c|\mathbf{I})] . \quad (2.12)$$

Therefore, finding the MAP of $P(\mathbf{z}|\mathbf{I})$ in Eq. 2.7 is equivalent to minimizing the corresponding Gibbs energy function $E(\mathbf{z}_c|\mathbf{I})$ [108] which results in the optimal segmentation of the target structure.

A clique c is a subset of the set \mathbf{P} of pixels such that each element of the set is a neighbor of all other elements in the set. The local configuration \mathbf{z}_c of the clique c determines the value of $V_c(z_c|\mathbf{I})$. The clique potential and the order of the clique are determined by the number of pixels in a clique. If a 4-neighborhood system is used the maximal clique size is 2, therefore, only the first and the second order cliques can be used. The first and the second order cliques can be expressed as Gibbs energy by

$$E(\mathbf{z}|\mathbf{I}) = \sum_{p \in \mathbf{P}} V_p(z_p|\mathbf{I}) + \sum_{p \in \mathbf{P}} \sum_{q \in N_p} V_{p,q}(z_p, z_q|\mathbf{I}) , \quad (2.13)$$

where $V_p(z_p|\mathbf{I})$ and $V_{p,q}(z_p, z_q|\mathbf{I})$ are the first¹ and second² order clique potential functions respectively. The first order clique potential $V_p(z_p|\mathbf{I})$ also called the regional or the data term (as it encodes regional constraints), measures the cost or penalty pixel p incurs when the label z_p is assigned to it given the observed image data \mathbf{I} . The second order clique potential function $V_{p,q}(z_p, z_q|\mathbf{I})$, also called the boundary or the

¹First order clique potential is also called as likelihood, data, regional and unary term in literature.

²Second order clique potential is also called as prior, smoothness, boundary and pairwise term in literature.

smoothness term (as it encodes boundary constraints), measures the cost or penalty of pixels p, q incurs when they are jointly assigned labels z_p and z_q where $q \in N_p$ given the observed image data \mathbf{I} . The smoothness term assumes that the image is generally made up of homogeneous patches which often changes discontinuously, therefore, most nearby pixels are expected to have the same label. If the pixels have the same label then there is no penalty otherwise they incur a penalty.

The energy function $E(\mathbf{z}|\mathbf{I})$ in Eq. 2.13 is the main energy function that will be augmented with the shape information of the facial muscles to segment the facial muscles in Chapters 3, 4, 5 and 6. Minimizing the resulting energy function will result in the segmentation of the facial muscles. A complete overview of the energy function $E(\mathbf{z}|\mathbf{I})$ is given in the next sections as well as how GC is used to find the global minimum of this energy function.

2.2.2 Optimizing the MRF

In the previous section, it has been explained that direct optimization of the MAP estimation of the MRF is not possible. However, if the MRF is equivalent to the GRF then the corresponding Gibbs energy function Eq. 2.13 can be minimized. Many methods have been reported in literature that try to minimize the Gibbs energy function.

Simulated Annealing (SA) proposed by Kirkpatrick *et al.* [118] was the first optimizer of the MRF. Theoretically SA is guaranteed to converge to the global minimum only if the temperature is decreased slowly but in practice decreasing the temperature slowly is not possible as the computational time becomes prohibitive. Consequently, finding the global minimum is not guaranteed. Later Besag [119] proposed Iterated Conditional Modes (ICM) which allows to change the label of a single site at a time. ICM was computationally faster than SA but at the expense of obtaining only a local minimum solution. Few other solvers for optimizing MRFs were proposed such as Mean Field Annealing (MFA) of Peterson & Söderberg [120], Relaxation Labeling (RL) of Rosenfeld *et al.* [121], Graduated Non-Convexity (GNC) of Blake & Zisserman [122]. Wu & Chung [79] have proposed Cross Entropy (CE), a stochastic solver to simulate rare events in the field of operations research, for solving MRF. It has been claimed that due to the stochastic nature of their solver it tends to find more global optima as compared to BP [87, 88]. CE is a deterministic solver that can be applied to all sorts of energy functions and any kind of MRF formulation, irrespective of the initialization.

2. BACKGROUND

However, with the introduction of the two new efficient and fairly accurate solvers BP [87, 88] and GC [77, 78], interest in the MRF framework has renewed. Although these two solvers provide fairly accurate results, at best they just provide local minima only. Boykov *et al.* [123] has shown that global optimization, in general of an arbitrary MRF energy function is NP-hard. Kolmogorov & Zabih [52] specifies that there is a subset of energy functions or only a certain class of energy functions that can be globally minimized with GC. For a two class (object/background) labeling problem, if the energy in Eq. 2.13 is sub-modular, then GC are guaranteed to find the global minimum of the energy function [52]. GC has been used for inferencing on the MRF in this thesis. The basic details of the GC method are given in the next section.

2.2.3 Graph-Cut

The GC algorithm is a combinatorial optimization technique that is used to segment facial muscles. Consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ which consists of a set of vertices \mathcal{V} and a set of edges \mathcal{E} which connects two neighboring nodes. Let the graph \mathcal{G} be an undirected graph, where an edge $e = \{pq\} \in \mathcal{E}$ between a pair of nodes p, q is independent of the orientation of the edge *i.e.* $e = e_{pq} = e_{qp}$. There is a one-to-one correspondence between the voxels and the nodes in the graph. Let there be two additional nodes the source S and the sink T in the graph which act as terminal nodes. The source node represents the object label while the sink node represents the background label. Each voxel is connected to both the source and the sink with edges called the *t-links* or Terminal-Links. Medical imaging data (*e.g.* CT and MRI) is normally organized in a regular grid-like fashion which lays out the foundations of the neighborhood system, Fig. 2.7. Voxels are connected to their neighboring voxels by edges in a regular grid like fashion, see Fig. 2.8. These neighboring edges are called *n-links* or Neighboring-Links. A typical neighborhood system for a 3D volume dataset is 6-,18- or 26-neighborhood and 4- or 8-neighborhood for a 2D image. All the edges $\mathcal{E} = \{t-links \cup n-links\}$ of the graph consisting of all the *t-links* and *n-links* are assigned some non-negative cost c_e called the edge capacity. The cost assigned to the *t-links* is based on the likelihood information (data term) obtained from the object and the background intensity model such as the object and the background histograms while the cost assigned to the *n-links* is based on the prior information (smoothness term). *n-links* between voxels that have similar intensity tend to have higher cost and the cost decreases as the intensities

differ. In order to allow the segmentation boundary to follow the intensity edges, c_{pq} is a non-increasing function of $|I_p - I_q|$ where I_p is the intensity value of pixel p . This function proposed by Boykov & Jolly [78] avoids over-penalizing the object and the background discontinuity at the boundary and is given by

$$V_{p,q}(z_p, z_q | \mathbf{I}) = c_{pq} \cdot \mathcal{J}(z_p \neq z_q) \quad := \begin{cases} \mathcal{J}(\cdot) = 1 & \text{if } z_p \neq z_q \\ \mathcal{J}(\cdot) = 0 & \text{if } z_p = z_q. \end{cases} \quad (2.14)$$

$$c_{pq} = \exp\left(-\frac{(I_p - I_q)^2}{2\sigma_I^2}\right) \quad (2.15)$$

where σ_I is an average absolute intensity difference between neighboring pixels.

An s - t cut $\mathcal{C} : \mathcal{C} \subset \mathcal{E}$ is a subset of edges defined on the graph \mathcal{G} such that if \mathcal{C} is removed from \mathcal{G} , \mathcal{G} is divided into two separate disjoint sub-graphs S and T . The two terminals S and $T = \mathcal{V} - S$ are completely isolated and there is no path from one terminal to the other $\mathcal{G}(\mathcal{C}) : \{\mathcal{V}, \mathcal{E}/\mathcal{C}\}$, Fig. 2.8(c) [68]. This results in the segmentation of the image into object and background classes. The cost of the cut \mathcal{C} is defined as the sum of the cost of all edges on the cut $|\mathcal{C}| = \sum_{e \in \mathcal{C}} c_e$. The minimum cut is the cut that has the smallest cost. The goal of GC segmentation is to find the best cut that would give an optimal segmentation with respect to the terms of the energy function Eq. 2.13. A cut generally occurs at the segmentation boundary cutting the n -links; therefore, the total cost of the cut is the cost of the segmentation boundary. Cutting the t -links represents the regional properties of the segmented regions. Thus a minimum cost cut represents segmentation with a desirable balance between the regional and boundary properties of the object.

A typical construction of the graph is shown in Fig. 2.8. Figure 2.8(a) shows a 3×3 pixel image. Figure 2.8(b) shows the corresponding graph. All the nodes are connected to the two terminal S and T and to their neighboring pixels. The thickness of the edges represents the edge capacity. There are many GC algorithms like Delong & Boykov [124] and Boykov & Kolmogorov [125] to compute the max-flow/min-cut of a graph separating the graph into two sub-graphs, one connected to the S and the other connected to the T , Fig. 2.8(c), while Fig. 2.8(d) shows the corresponding segmented image which is easy to obtain since the image pixels and the nodes in the graph are in one-to-one correspondence. The freely available implementations of both the algorithms for research purpose have been used in this work to compute the segmentation results.

2. BACKGROUND

All the nodes that are reachable from the source are labeled as object pixels while others are labeled as background pixels.

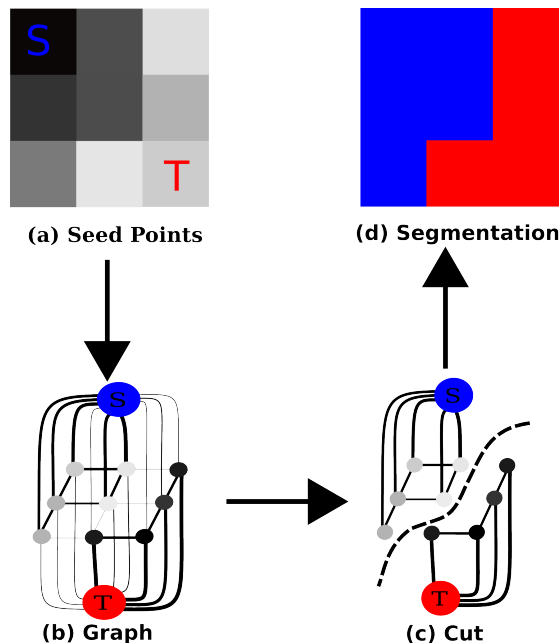


Figure 2.8: Graph-Cut model.

2.2.4 Min-Cut Max-Flow Algorithm

Consider the graph shown in Fig. 2.9. Max-flow algorithm can be considered as transferring a flow from S to T [126]. The goal is to send the maximum amount of flow from the source to the sink without exceeding any of the edge capacities. The algorithm works by finding a path from the source to the sink and sending the maximum amount of flow along this path. The maximum flow is restricted by the minimum edge capacity along the path. Once the flow has been transferred, the path will contain at least one saturated edge. The process is repeated until there is a path from the source to the sink. The algorithm terminates when there are no more paths connecting the source to the sink and the set of saturated edges defines the cuts set \mathcal{C} (see Fig. 2.9). The vertices that can be reached from the source are then labeled as the source vertices and the rest of the vertices are labeled as the sink vertices. There are many algorithms (such as augmenting path, push-relabel) to compute the min-cut max-flow algorithm, the complete details of which can be found in [126, 127].

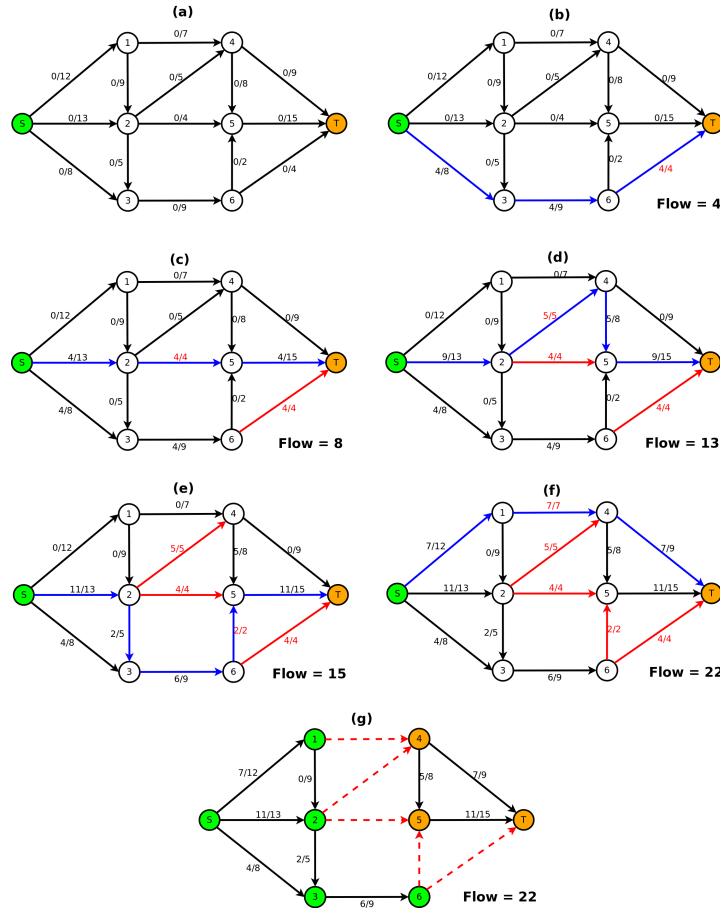


Figure 2.9: Min-Cut Max-Flow Algorithm (the figure is inspired by Fig. 12.6 in [1]). (a) shows the original graph. The flow and the capacity of each edge is shown by $[0/0]$, (b,c,d,e,f) show the paths from the source to the sink. A path is shown in blue while the red color shows the saturated edges, (g) shows the edges in the cut set which separates the source nodes in green from the sink nodes in orange. The max-flow of the network is 22.

2.3 Statistical Model

An anatomical structure shows considerable shape variability among the population which cannot be represented by a fixed shape template. Statistical shape models are widely used as a mathematical framework to capture the shape variability of a class of shapes in medical image analysis, computational anatomy and computer vision [61, 128, 129, 130, 131].

A set of shapes in the training dataset is used to capture the shape variability of a

2. BACKGROUND

particular structure. It is assumed that the samples are Independent and Identically Distributed (i.i.d.) and are sampled from the underlying unknown multivariate normal distribution $P \sim \mathcal{N}(\bar{\mathbf{x}}, \Sigma)$. If the training dataset contains good representatives of the class of shapes occurring in the test dataset, then the resulting statistical model is highly robust and can be fitted to any shape within this class. The goal is to build a statistical shape model of the muscle over this distribution of surface meshes using the PPCA based approach.

In this work a morphable model [132] representation of the sample shapes in the training dataset is chosen and then a probability density function over the morphable model representation is defined. Morphable models are used to restrict the solution space of ill-posed problems by penalizing the unlikely instances of the shape. This is not the case in the statistical model representation of El-Zehiry & Elmaghraby [100] where an unlikely instance of the object is also plausible. The shapes in the training dataset belong to the same class of object, therefore, they have some dependence on each other which introduces redundancy. This redundancy is minimized by using the extended statistical model approach based on the PPCA as suggested by Lüthi *et al.* [132] instead of using the traditional PCA based approach of [129, 130]. The difference is that PPCA provides an additional probabilistic interpretation of the PCA based model. PPCA provides best reconstruction and allows computing the remaining variability of a shape based on partial information. Here it will be used to reconstruct the muscle shape based on landmark information and simultaneously compute the statistical model with reduced variability.

2.3.1 Principal Component Analysis (PCA)

Let $\{\Gamma_i \subset \mathbb{R}^{3m} | i = 1, \dots, n\}$ be n 3D triangulated surface meshes in the training dataset each having m number of vertices. Shape Γ_1 is arbitrarily defined as the reference surface. All the surfaces are aligned with Γ_1 by aligning the manually labeled landmarks on the sample shapes Γ_i using the Procrustes. It is important to bring the shapes into correspondence when estimating shape variances of a class of object. If the corresponding parts of the sample shapes are misaligned, then the model will not capture the appropriate anatomical shape variance. Γ_i are then registered with Γ_1 by computing curvature sensitive deformation fields for the vertices of Γ_1 using the method of Dedner *et al.* [133], which resulted in all sample shapes having the same number of

vertices. The resulting shape vector $\mathbf{x}_i \in \mathbb{R}^M$ where $M = 3 \times m$ is defined as $\mathbf{x}_i = (v_x^{i,1}, v_y^{i,1}, v_z^{i,1}, \dots, v_x^{i,m}, v_y^{i,m}, v_z^{i,m})^T$ where the vector $v^{i,j} = (v_x^{i,j}, v_y^{i,j}, v_z^{i,j})$ represents the x, y, z coordinates of the j^{th} vertex of \mathbf{x}_i .

The PCA based shape model assumes that the training dataset \mathbf{x}_i and its linear combinations form a linear space of shapes having an underlying unknown distribution which can be estimated by the multivariate normal distribution $P \sim \mathcal{N}(\bar{\mathbf{x}}, \Sigma)$, where the mean shape vector $\bar{\mathbf{x}} \in \mathbb{R}^{M \times 1}$ is $\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i$. The latter is used to form the mean free shape data matrix $\mathbf{X} \in \mathbb{R}^{M \times n}$ as $\mathbf{X} = [\mathbf{x}_1 - \bar{\mathbf{x}}, \mathbf{x}_2 - \bar{\mathbf{x}}, \dots, \mathbf{x}_n - \bar{\mathbf{x}}]$. The covariance matrix $\Sigma \in \mathbb{R}^{M \times M}$ can be estimated from the mean free shape matrix \mathbf{X} as $\Sigma = \frac{1}{n} \mathbf{X} \mathbf{X}^T$.

As $\mathbf{X} \mathbf{X}^T$ is prohibitively large, SVD is applied on \mathbf{X} which decomposes into $\text{SVD}(\mathbf{X}) = \mathbf{U} \mathbf{D} \mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{M \times n}$ is a column-orthonormal matrix, $\mathbf{D} \in \mathbb{R}^{n \times n}$ is a diagonal matrix and $\mathbf{V} \in \mathbb{R}^{n \times n}$ is also an orthonormal matrix. \mathbf{V} gives the scores of the original shape vectors in mean free shape data matrix \mathbf{X} in the domain of the new shape space spanned by the shapes in the training set. The covariance matrix can now be written as

$$\Sigma = \frac{1}{n} \mathbf{X} \mathbf{X}^T (\because \mathbf{X} = \mathbf{U} \mathbf{D} \mathbf{V}^T) \quad (2.16)$$

$$\Sigma = \frac{1}{n} (\mathbf{U} \mathbf{D} \mathbf{V}^T) (\mathbf{U} \mathbf{D} \mathbf{V}^T)^T \quad (2.17)$$

$$\Sigma = \frac{1}{n} (\mathbf{U} \mathbf{D} \mathbf{V}^T \mathbf{V} \mathbf{D} \mathbf{U}^T) (\because \mathbf{V}^T \mathbf{V} = 1, \mathbf{D}^T = \mathbf{D}) \quad (2.18)$$

$$\Sigma = \frac{1}{n} (\mathbf{U} \mathbf{D}^2 \mathbf{U}^T) . \quad (2.19)$$

The columns \mathbf{u}_i of \mathbf{U} arranged in the descending order of the model variance they define, are the eigenvectors of Σ . These eigenvectors also called the principle components of the model, describe the main modes of variation of the training data. Figure 2.10 shows the first two main modes of variation of the statistical model. The eigenvalues σ_i^2 corresponding to the unit eigenvectors \mathbf{u}_i are given by $\sigma_i^2 = \frac{1}{n} \mathbf{D}_i^2$ where \mathbf{D}_i is the i -th diagonal entry of \mathbf{D} . The eigenvalues σ_i^2 describe the variance of the model in the direction of the corresponding eigenvectors. The mean shape and the main modes of variation provide a reasonable representation of shapes of the class of objects being learned. Any novel shape \mathbf{x} can now be generated from this statistical model using

2. BACKGROUND

the coefficients $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n)^T \in \mathbb{R}^n$ which have a multivariate normal distribution

$$P(\boldsymbol{\alpha}) = \mathcal{N}(0, \mathbf{I}_n)$$

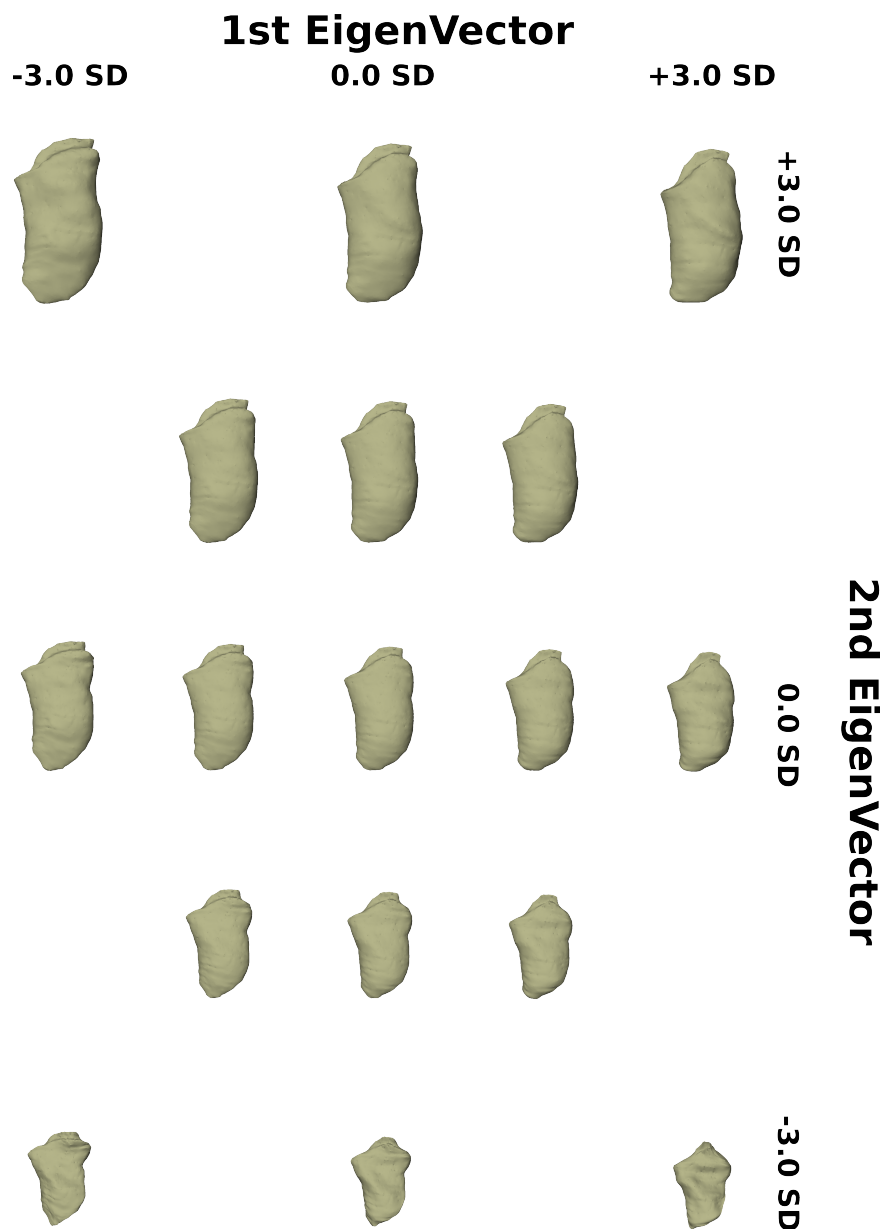


Figure 2.10: The first two main eigenvectors of the statistical model of the masseter muscle and the features of the shape they influence.

$$\mathbf{x} = \bar{\mathbf{x}} + \sum_{i=1}^n \alpha_i \sigma_i \mathbf{u}_i \quad (2.20)$$

$$\mathbf{x} = \bar{\mathbf{x}} + \frac{1}{\sqrt{n}} \mathbf{U} \mathbf{D} \boldsymbol{\alpha} . \quad (2.21)$$

A new shape \mathbf{y} belonging to the class of objects can be represented in the new eigen-space spanned by \mathbf{U} as

$$\boldsymbol{\alpha}_y = \mathbf{U}^T (\mathbf{y} - \bar{\mathbf{x}}) . \quad (2.22)$$

$\boldsymbol{\alpha}_y$ gives the coefficients of a shape \mathbf{y} in the eigen shape space.

2.3.2 Probabilistic PCA

The probabilistic statistical model used for shape modeling is created by applying standard PPCA [134] to the sample shapes in the training dataset. The main idea behind PPCA is that the sum of mapping of a low dimensional latent space variable $\boldsymbol{\alpha} = \mathcal{N}(0, \mathbf{I}_n)$ ¹ and some Gaussian noise can explain a high dimensional shape \mathbf{x} . The conditional distribution $P(\mathbf{x}|\boldsymbol{\alpha})$ of the shape vector \mathbf{x} is

$$P(\mathbf{x}|\boldsymbol{\alpha}) = \mathcal{N}(\mathbf{W}\boldsymbol{\alpha} + \bar{\mathbf{x}}, \sigma^2 \mathbf{I}_M) , \quad (2.23)$$

where $\mathbf{W} \in \mathbb{R}^{M \times n}$ are the scaled eigenvectors $\mathbf{W} = \mathbf{U} \mathbf{D}$. Thus $P(\mathbf{x}|\boldsymbol{\alpha})$ provides a generative model where a novel shape \mathbf{x} can be generated by mapping of the latent variables $\boldsymbol{\alpha}$ by \mathbf{W} and adding Gaussian noise $\epsilon \in \mathcal{N}(0, \sigma \mathbf{I}_M)$ which can be written in terms of a linear Gaussian model as

$$\mathbf{x} = \mathbf{W}\boldsymbol{\alpha} + \bar{\mathbf{x}} + \epsilon . \quad (2.24)$$

As $P(\boldsymbol{\alpha})$ and $P(\mathbf{x}|\boldsymbol{\alpha})$ are known and using the theorem of linear Gaussian model [117] it is easy to drive the expressions for $P(\mathbf{x})$ as in Eq. 2.25 and $P(\boldsymbol{\alpha}|\mathbf{x})$ as in Eq. 2.26.

$$P(\mathbf{x}) = \mathcal{N}(\bar{\mathbf{x}}, \mathbf{W} \mathbf{W}^T + \sigma^2 \mathbf{I}_M) , \quad (2.25)$$

¹ \mathbf{I}_n is the $(n \times n)$ identity matrix.

2. BACKGROUND

$$P(\boldsymbol{\alpha}|\mathbf{x}) = \mathcal{N}(\mathbf{M}^{-1}\mathbf{W}^T\sigma^{-2}(\mathbf{x} - \bar{\mathbf{x}}), \mathbf{M}^{-1}), \quad (2.26)$$

where $\mathbf{M} = \sigma^{-2}\mathbf{W}^T\mathbf{W} + \mathbf{I}_n$. The maximum likelihood estimates of $\bar{\mathbf{x}}_{ML}$, \mathbf{W}_{ML} , σ_{ML} as calculated by Tipping & Bishop [134] are given by

$$\bar{\mathbf{x}}_{ML} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.27)$$

$$\mathbf{W}_{ML} = \mathbf{U}(\mathbf{D}^2 - \sigma^2\mathbf{I}_n)^{\frac{1}{2}} \quad (2.28)$$

$$\sigma_{ML} = \frac{1}{M-n} \sum_{i=n+1}^M \mathbf{D}_{ii}^2. \quad (2.29)$$

\mathbf{U} and \mathbf{D} are obtained from the SVD of $\boldsymbol{\Sigma}$ as given in Eq. 2.19. The maximum likelihood estimates of $\bar{\mathbf{x}}_{ML}$, \mathbf{W}_{ML} and σ_{ML} characterize a complete generative model (Eq. 2.24) which is similar to the PCA model but in addition provide a fully probabilistic model, thus enabling the computation of a full posterior distribution and allow dealing with the missing data in a principled manner.

3

Approach A: A Shape Prior-Based MRF Model for 3D Masseter Muscle Segmentation

3.1 Introduction

The material presented in this chapter is an extended version of the paper Majeed *et al.* [19] published in Society of Photographic Instrumentation Engineers (SPIE): Medical Imaging.

This chapter describes an approach to segment the masseter muscle using GC by incorporating additional 3D shape priors in CT datasets, which is robust to noise, artifacts, and shape deformations. The main contribution is in translating the 3D shape knowledge into the first and second order clique potentials of the MRF. The segmentation task is casted as a MAP estimation of the MRF. The approach is evaluated on the difficult problem of segmenting the masseter muscle, Fig. 3.1. GC is then used to obtain the global minimum which results in the segmentation of the masseter muscle. The method is tested on 20 CT datasets of the masseter muscle, which are noisy with almost all possessing mild to severe imaging artifacts. The proposed technique produces clinically acceptable results. A quantitative and qualitative comparison of the proposed method with other methods has also been carried out. Statistical analysis is done to show that adding shape prior can increase the robustness of the proposed method in noisy datasets.

3. FIXED SHAPE TEMPLATE BASED SEGMENTATION

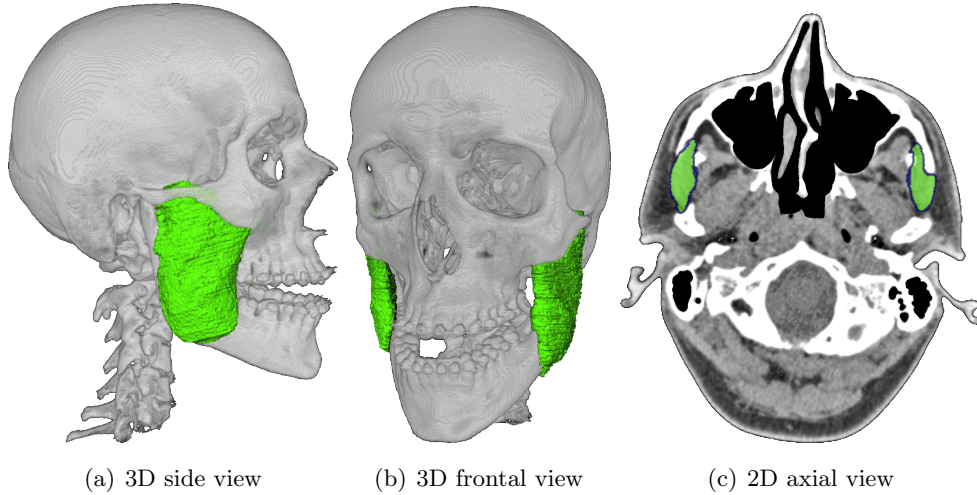


Figure 3.1: Masseter muscle shown in green.

This chapter is organized as follows; Sec. 3.2 describes the segmentation model where the two terms of the energy function based on intensity information and the prior shape information will be explained. Experimental results are presented in Sec. 3.3 and finally, Sec. 3.4 concludes the whole approach.

3.2 Segmentation Model

In Chap. 2 the segmentation problem is formulated in terms of the MAP-MRF which is equivalent to the minimization of the Gibbs energy function, Eq. 2.13, in order to segment the desired object. As described there, GC can successfully segment the desired object if its intensity profile is different from the surrounding. However, when this is not the case, and this is generally true for medical images, prior shape knowledge has to be incorporated into the energy function to successfully segment the desired object. Segmenting the masseter muscle from CT datasets using the traditional intensity based energy function is thus impossible because of the presence of adjacent anatomical structures having a similar intensity profile, weak edges, non-homogeneous interior of the muscle and medical imaging artifacts. The use of the shape prior is to constrain the segmentation based on the intensity information and force the resulting segmentation to have a shape, identical to that of the shape prior; therefore, the energy function,

Eq. 2.13, is augmented with additional shape knowledge. The masseter muscles segmentation problem is then formulated by minimizing the augmented energy functional

$$E(\mathbf{z}|\mathbf{I}, \bar{\mathbf{x}}) = E^I(\mathbf{z}|\mathbf{I}) + E^S(\mathbf{z}|\bar{\mathbf{x}}), \quad (3.1)$$

where \mathbf{z} is a binary variable which defines the segmentation, \mathbf{I} is the observed image data and $\bar{\mathbf{x}}$ is the shape prior. Equation 3.1 consists of an intensity term $E^I(\mathbf{z}|\mathbf{I})$ and a shape prior term $E^S(\mathbf{z}|\bar{\mathbf{x}})$ which are based on the intensity and shape information respectively. $E^I(\mathbf{z}|\mathbf{I})$ tries to segment the muscle based on the intensity information while $E^S(\mathbf{z}|\bar{\mathbf{x}})$ tries to segment the muscle based on the shape prior information. None of the energy terms alone are able to segment the muscle correctly. A fine balance between the two energy terms is required to get good segmentation of the muscle. Both the energy terms $E^I(\mathbf{z}|\mathbf{I})$ and $E^S(\mathbf{z}|\bar{\mathbf{x}})$ consists of the first and the second order clique potential functions of the Gibbs energy function. The energy terms $E^I(\mathbf{z}|\mathbf{I})$ and $E^S(\mathbf{z}|\bar{\mathbf{x}})$ are explained in detail in the following subsections. The main contribution here is in transforming the global shape information into both the local first and second order clique potentials.

3.2.1 Intensity Energy Function

The intensity information of the target anatomical structure is modeled by the energy function $E^I(\mathbf{z}|\mathbf{I})$, which is based on the traditional GC intensity based energy function of Boykov & Jolly [78] and is given by

$$E^I(\mathbf{z}|\mathbf{I}) = (1 - \lambda)(1 - \mu) \sum_{p \in \mathbf{P}} V_p^I(z_p) + \lambda(1 - \gamma) \sum_{p \in \mathbf{P}} \sum_{q \in N_p} V_{pq}^I(z_p, z_q), \quad (3.2)$$

$$V_p^I(z_p = \text{"1"}) = -\ln P(I_p | \text{"1"}), \quad (3.3)$$

$$V_p^I(z_p = \text{"0"}) = -\ln P(I_p | \text{"0"}). \quad (3.4)$$

where λ is the smoothness parameter, μ and γ are the shape prior parameters for data and smoothness terms respectively. N_p is a 6-neighborhood system for \mathbf{P} . The data term $V_p^I(z_p)$ encodes regional properties of the foreground, Eq. 3.3, and background, Eq. 3.4. It defines the individual penalties for assigning voxel p to foreground and background which are calculated from the respective intensity histograms. The smoothness

3. FIXED SHAPE TEMPLATE BASED SEGMENTATION

term $V_{pq}^I(z_p, z_q)$, Eq. 3.5, encodes the boundary properties of the foreground. It defines a discontinuity penalty between adjacent voxels p and q , when they are assigned different labels and is given by

$$V_{pq}^I(z_p, z_q) \propto \exp\left(-\frac{(I_p - I_q)^2}{2\sigma_I^2}\right) \quad (3.5)$$

where σ_I^2 is the average absolute intensity difference between neighboring voxels of the whole image in a 6-neighborhood system.

3.2.2 Shape Prior Energy Function

The prior shape information of the masseter muscle is encoded into the data and the smoothness terms of the energy function $E^S(\mathbf{z}|\bar{\mathbf{x}})$, Eq. 3.6. Such encoding of the global shape information into the first and the second order clique potential functions of the Gibbs energy transforms it into the local constraints:

$$E^S(\mathbf{z}|\bar{\mathbf{x}}) = (1 - \lambda)\mu \sum_{p \in \mathbf{P}} V_p^S(z_p) + \lambda\gamma \sum_{p \in \mathbf{P}} \sum_{q \in N_p} V_{pq}^S(z_p, z_q). \quad (3.6)$$

The data term $V_p^S(z_p)$ encodes how likely a particular voxel p is to belong to the foreground “1” and background “0”, given the shape prior $\bar{\mathbf{x}}$. An unsigned distance map [135] of the aligned shape prior is calculated. Based on the unsigned distance map, a probability map for the voxels is calculated, Fig. 3.2.

$$V_p^S(z_p = \text{“1”}) = -\ln P(\bar{\mathbf{x}}|\text{“1”}). \quad (3.7)$$

$$V_p^S(z_p = \text{“0”}) = -\ln P(\bar{\mathbf{x}}|\text{“0”}). \quad (3.8)$$

Equations 3.7 & 3.8 define the penalty voxel p incurs, given that it belongs to foreground or background respectively. These penalties are calculated by creating a probability map for the voxels as shown in Fig. 3.2. Two different probability maps, one for foreground and the other for background are created. The foreground probability map considers only the voxels enclosed by the shape prior’s contour. The voxels that are farthest away from the contour lie at the center of the shape and hence are likely to belong to foreground as compared to the voxels close to the contour, which is computed

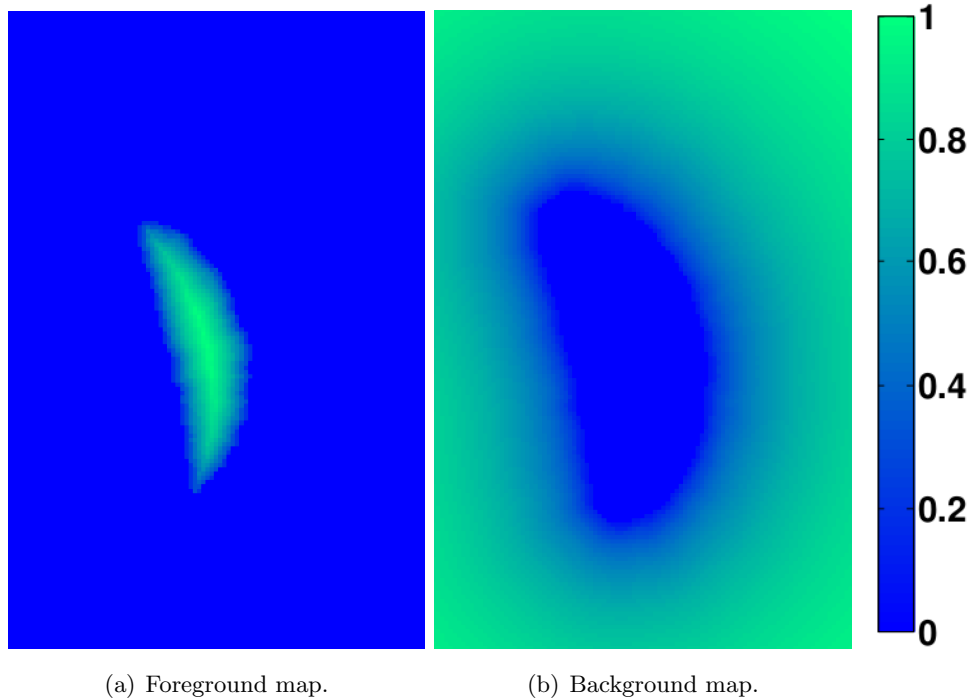


Figure 3.2: Probability maps.

by Eq. 3.7 as the negative log-likelihood of belonging to foreground. Thus, the voxels at the center of the shape prior, incur higher penalty for belonging to the background as compared to voxels close to the contour. Voxels outside the contour incur 0 penalty for belonging to the background. The opposite is true for the background probability map, where only the voxels outside the contour are considered. The voxels close to the contour are more likely to belong to the foreground as compared to the voxels farther away. Thus, the voxels farther away from the contour, incur a higher penalty for belonging to the foreground as compared to voxels close to the contour. Voxels inside the contour incur 0 penalty to belong to the foreground. A probability map has been used because adding the shape prior in a hard shape constraint manner (see Slabaugh & Unal [91]) is not applicable as the shape prior does not fit the target muscle well. The proposed approach called App. A, from here onwards uses only a fixed shape template in a single iteration for the segmentation.

Freiman *et al.* [94] also creates a probability map by registering binary image segmentations, where as the proposed probability map is created from the unsigned dis-

3. FIXED SHAPE TEMPLATE BASED SEGMENTATION

tance map of the template’s shape contour in 3D. The shape smoothness term $V_{pq}^S(z_p, z_q)$ is similar to that of Freedman and Zhang [92]. This work proposes to add shape prior in both the data $V_p^S(z_p)$ and smoothness $V_{pq}^S(z_p, z_q)$ terms of the energy function, which makes the segmentation robust to image artifacts and noise which is discussed in more detail in Sec. 3.3.

For a facial muscle, a mesh with a representative shape is used as the shape template $\bar{\mathbf{x}}$. The shape template is calculated by averaging manually segmented muscles as explained in Sec. 2.3.1. It is then semi-automatically initialized in the dataset using landmark registration. Algorithm 1 summarizes the segmentation method.

Algorithm 1

Require: An initialized shape template.

Ensure: GC segmentation of target muscle.

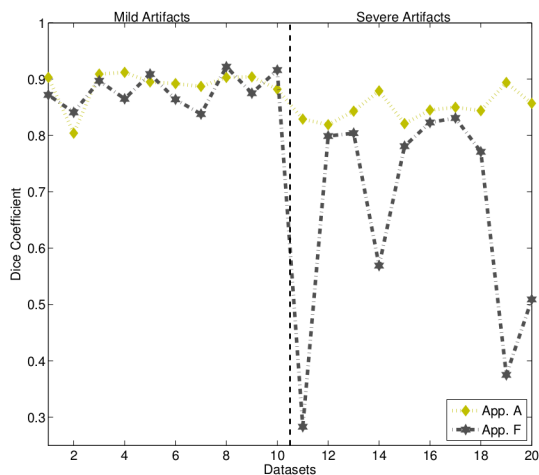
1. Compute: $V_p^S(z_p)$ as explained in Sec. 3.2.2.
 - (a) Compute: $V_p^S(z_p = \text{“1”})$.
 - (b) Compute: $V_p^S(z_p = \text{“0”})$.
 2. Compute: $V_{pq}^S(z_p, z_q)$.
 3. Create graph according to the energy function $E(\mathbf{z}|\mathbf{I}, \bar{\mathbf{x}})$.
 4. Compute the optimal segmentation using GC.
-

3.3 Experimental Setting and Results

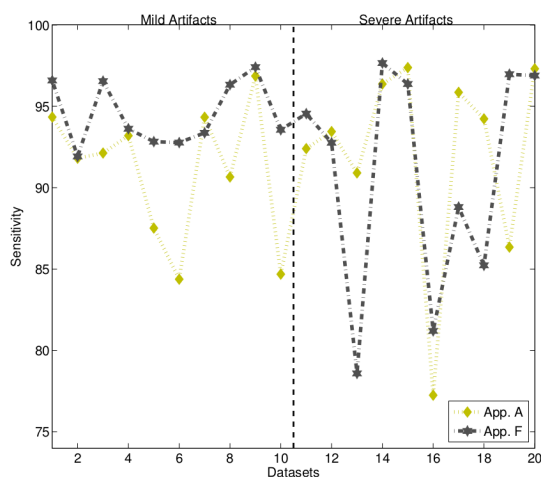
The above described segmentation method was tested on 20 CT datasets¹ of the masseter muscle with dataset dimensions $79\text{-}156 \times 148\text{-}214 \times 125\text{-}384$ voxels and spacing $0.3\text{-}0.5 \times 0.3\text{-}0.5 \times 0.3\text{-}1$ mm³. The datasets were noisy with almost all possessing mild to severe imaging artifacts and were chosen randomly from a CMF surgery department

¹The published paper lists 21 datasets. In the later approaches (Chap. 4 & 5), only 20 datasets were used, therefore, for consistency the results of the 20 datasets have been provided. When building the statistical model, it was observed, that one dataset was not completely segmented and thus was removed from building the statistical model and further segmentation.

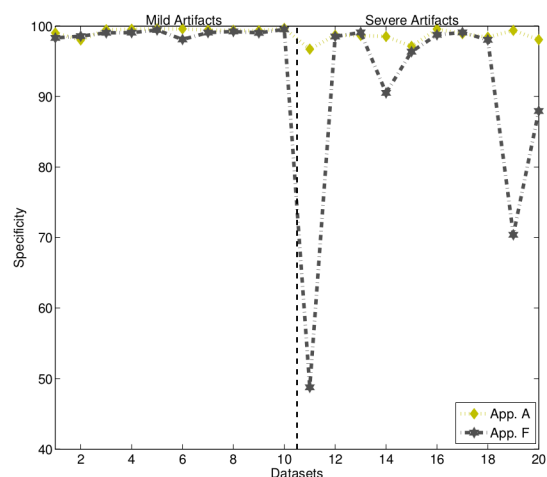
3.3 Experimental Setting and Results



(a) Dice coefficient



(b) Sensitivity



(c) Specificity

Figure 3.3: Graphs show comparison results between App. A (yelloish-green curve) which incorporates the shape prior in both the data and smoothness terms is more accurate compared to App. F which incorporates it in just the smoothness term (gray curve).

so that they contained representative anatomical variations. The ground truth manual segmentations were generated by a medical expert. Due to the limited number of datasets and the ground truth, the method was evaluated using a Leave-One-Out approach. The ground truth segmentations were used to calculate the mean shape, which was then used as the shape template. The parameters were manually optimized

3. FIXED SHAPE TEMPLATE BASED SEGMENTATION

on three datasets and then the same parameters were used on the rest of the datasets. The parameter values used were $\lambda = 0.3$, $\mu = 0.02$ and $\gamma = 0.004$. The Dice Coefficient (DC), sensitivity and specificity of the segmentation were calculated as similarity measures to ascertain the accuracy of the method.

	DC (Mean \pm Std)	DC (Median)	DC (Smallest - Largest)
App. A Mild	0.889 ± 0.031	0.899	(0.804 - 0.912)
App. F Mild	0.880 ± 0.030	0.874	(0.838 - 0.922)
App. A Severe	0.848 ± 0.024	0.845	(0.819 - 0.894)
App. F Severe	0.654 ± 0.205	0.776	(0.283 - 0.831)
App. A All	0.869 ± 0.034	0.881	(0.804 - 0.912)
App. F All	0.767 ± 0.183	0.835	(0.283 - 0.922)

Table 3.1: This table lists the mean & standard deviation, the median and the smallest as well as the largest value of the dice coefficient of App. A in comparison to App. F.

The graphs in Fig. 3.3 show the results of App. A (yelloish-green curve). App. A is also compared to the fixed shape template method of Freedman and Zhang [92] (gray curve) called App. F from here onwards, which just adds the shape prior in the smoothness term, to which App. A is closely related. The results clearly show that by adding the shape prior in both the terms (yelloish-green curve) the segmentation result is more accurate and robust to image artifacts than merely adding the shape prior in the smoothness term (gray curve). This is especially true for datasets severely corrupted with high-density artifacts. Table 3.2 gives a quantitative evaluation (dice coefficient, sensitivity and specificity) of the results of App. A and the results obtained by applying App. F to the same datasets for comparison. A concise table of the comparison of the dice coefficient values is presented in Table 3.1, which shows that App. A produces better results when compared to App. F with a higher dice coefficient and a lower standard deviation.

Figure 3.4 shows qualitative results in 2D, while Fig. 3.5 depicts the results in 3D for the datasets without noise. Furthermore, Fig. 3.6 (a-d) displays the results in 3D over noisy datasets. Figure 3.6 also shows a qualitative comparison of App. A with that of App. F.

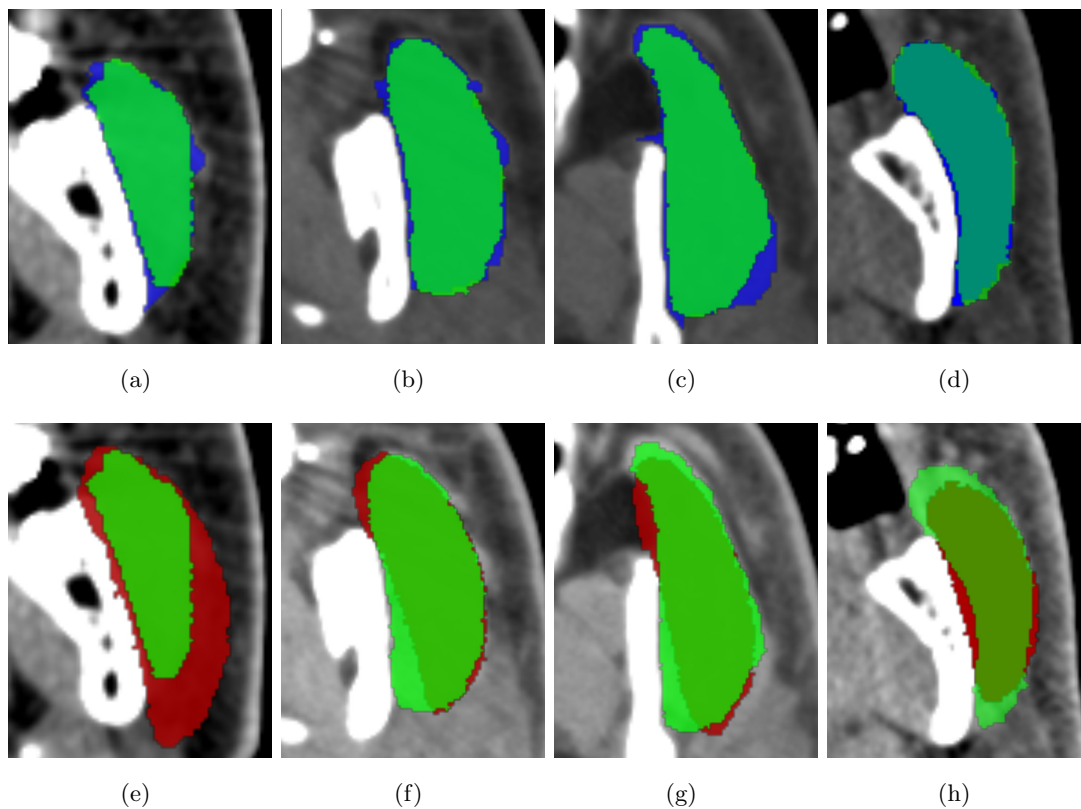


Figure 3.4: Qualitative segmentation results in 2D. The top row (a-d) shows the segmentation result in blue, while the bottom row (e-h) shows the shape prior in red. The overlaid green in both rows represents the ground truth.

A statistical comparison of App. A with App. F was also conducted and statistical analysis (t-test with significance level 0.05) was performed over the dice coefficient values to ascertain whether the two methods are statistically significantly different. For this, the available datasets were split depending on the severity of imaging artifacts. Statistical analysis showed that for the datasets without imaging artifacts both the methods, App. A as well as App. F, performed statistically equally well (p-value = 0.51). However, when comparing datasets with imaging artifacts, App. A performed statistically significantly better (p-value = 0.009) over the dice coefficients as shown in Fig. 3.6. This figure shows the results of App. A from (a-d) over noisy datasets, while (e-h) displays the results of App. F over the same noisy datasets. This clearly demonstrates the increased robustness gained by including a shape prior in both the

3. FIXED SHAPE TEMPLATE BASED SEGMENTATION

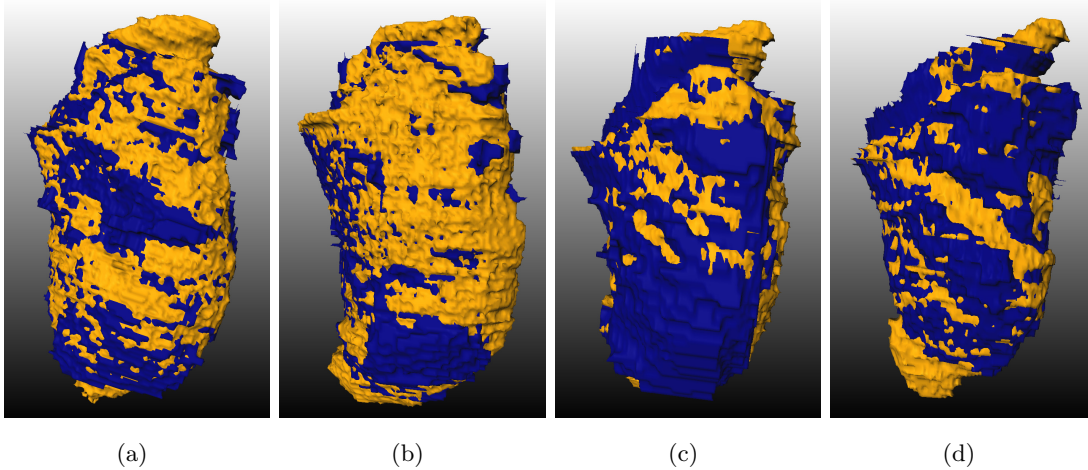


Figure 3.5: Qualitative segmentation results in 3D of left muscle (a-d) of App. A shown in blue and overlaid in orange is the ground truth.

data and the smoothness terms as proposed.

As stated earlier the main contribution is in adding shape prior in both the data and the smoothness terms, while App. F used the shape prior only in the smoothness term. This new shape prior in the energy functional makes the segmentation approach more robust against noisy and corrupted datasets. The shape prior in App. F has the effect of smoothing out regions, which results in a portion of the muscle completely smoothed out, Fig. 3.6 (e-h). This can be handled by App. A that incorporates the shape prior in the both terms. This also allows the flexibility to encode shape knowledge in areas where the intensity alone does not provide enough knowledge because of the artifacts. In such regions, through the shape prior in the data term, intensity likelihood can be counterbalanced by looking at the spatial location of the voxel with respect to the shape prior. That is, it can be computed, how likely is it for a voxel to belongs to the muscle.

As can be seen, the intensity energy function, Eq. 3.2, and shape prior energy function, Eq. 3.6, consist of two terms (data and smoothness terms), each of which are controlled by different parameters λ, μ, γ . If the value of μ is set to 0, which nullifies the shape prior in the data term, then the resulting energy function is the same as proposed by Freedman and Zhang [92]. If both μ, γ are set to 0, which completely nullifies the shape prior terms (the data and smoothness terms), then the resulting energy function turns into the one proposed by Boykov and Jolly [78].

3.3 Experimental Setting and Results

Data	Artifact	Similarity Measure					
		Dice Coefficient		Sensitivity%		Specificity%	
		App. A	App. F	App. A	App. F	App. A	App. F
1	Mild	0.903	0.872	94.34	96.59	99.01	98.30
2	Mild	0.804	0.841	91.81	91.92	98.02	98.56
3	Mild	0.909	0.897	92.13	96.55	99.46	99.03
4	Mild	0.912	0.865	93.20	93.62	99.53	99.05
5	Mild	0.895	0.909	87.52	92.83	99.59	99.43
6	Mild	0.892	0.864	84.38	92.77	99.58	98.08
7	Mild	0.887	0.838	94.33	93.37	99.41	99.06
8	Mild	0.903	0.922	90.66	96.33	99.37	99.22
9	Mild	0.904	0.875	96.86	97.41	99.32	99.02
10	Mild	0.882	0.916	84.69	93.56	99.65	99.51
11	Severe	0.829	0.283	92.40	94.55	96.71	48.76
12	Severe	0.819	0.799	93.45	92.76	98.68	98.50
13	Severe	0.843	0.804	90.91	78.59	98.65	99.08
14	Severe	0.879	0.569	96.37	97.66	98.50	90.51
15	Severe	0.821	0.781	97.38	96.37	97.11	96.35
16	Severe	0.845	0.823	77.25	81.20	99.55	98.74
17	Severe	0.850	0.831	95.86	88.81	98.91	99.08
18	Severe	0.844	0.771	94.23	85.23	98.40	98.03
19	Severe	0.894	0.375	86.35	96.97	99.38	70.39
20	Severe	0.857	0.509	97.31	96.89	98.05	87.94

Table 3.2: Quantitative comparison of App. A with that of App. F.

While the segmentations obtained through App. A were generally good, there were some areas, where they leaked because the shape prior was either not accurate enough or not optimally aligned. The average running time of the segmentation process was between 1-5 seconds. Although, the method was not as accurate as a human expert, the results are clinically acceptable as validated by medical experts from the department of CMF surgery of the University Hospital Basel. For a comparison of daily routine segmentation of the muscle structure at the University Hospital Basel, see Fig. 3.7. All results are computed using the freely available implementation of the Max-Flow Min-Cut algorithm of Boykov & Kolmogorov [125].

3. FIXED SHAPE TEMPLATE BASED SEGMENTATION

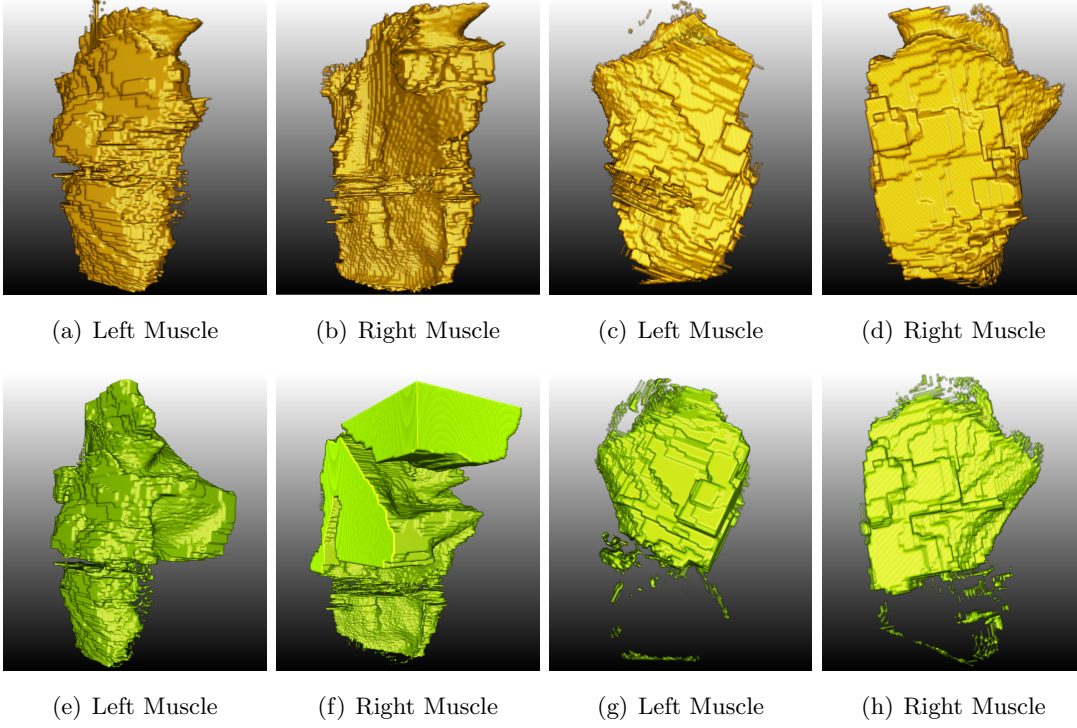


Figure 3.6: Qualitative comparison of the segmentation results in 3D of App. A with that of App. F over noisy datasets. The top row (a-d) shows segmentation results of App. A, while the bottom row (e-h) shows the results of App. F for the same noisy datasets.

3.4 Conclusion

A new way to incorporate shape priors in the MRF framework and utilize GC to obtain the global solution to segment masseter muscle from CT datasets has been presented. Segmenting the muscle is a challenging task because of the presence of soft-tissues in close proximity, weak edges and artifacts. However, we demonstrated that by incorporating prior shape knowledge, clinically acceptable results can be achieved. The contribution of App. A is in transforming global shape knowledge into the data and the smoothness energy terms.

Since a mean shape is not a good representation of the natural muscle size and shape variations found in the population, the mean shape template is replaced with a statistical shape model in Chap. 4 to better capture the shape variations in the population.

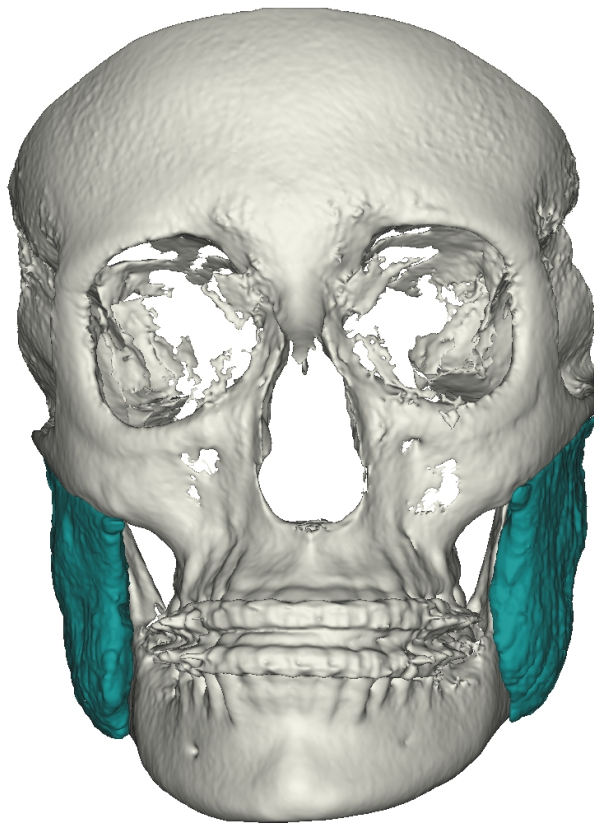


Figure 3.7: Segmented masseter muscle in a clinical routine.

3. FIXED SHAPE TEMPLATE BASED SEGMENTATION

4

Approach B: Using a Flexibility Constrained 3D Statistical Shape Model for Robust MRF-Based Segmentation

4.1 Introduction

The material presented in this chapter is an extended version of the paper Majeed *et al.* [20] that was published in *Mathematical Methods in Biomedical Image Analysis (MMBIA)*.

In this approach, the previous method is extended by replacing the mean shape template $\bar{\mathbf{x}}$ with the shape knowledge obtained from a 3D statistical model \mathbf{x}^* , which is integrated into the MRF segmentation framework. As automatic segmentation with MRF is prone to leaking due to weak edges, noise, imaging artifacts, missing shape knowledge, and shape deformations, this approach combines MRF-based segmentation with a 3D statistical shape model as prior anatomical knowledge. The statistical model is learned using a PPCA, which is constructed from the manually segmented CT datasets to capture the shape variability naturally found in the population. It also allows reconstructing the optimal shape and to compute the remaining variance of the statistical model from partial information. Instead of constraining the variability of the model through regularization [117, 130], where finding a good regularization parameter

4. STATISTICAL MODEL BASED SEGMENTATION

for all datasets is difficult, the constrained variability approach of Lüthi *et al.* [132] is used, which results in a more efficient shape update as compared to a regularization-based shape model reconstruction. The variability can be constrained by additional observed knowledge, such as input landmarks. In segmenting the masseter muscle (see Fig. 3.1), for instance, the muscle attachments to the bone, from a preceding skull segmentation step, can be used as landmarks.

The segmentation process is as follows. First, an initial shape generated using the statistical model that best fits the landmarks is used as the starting shape prior to segment the muscle. Second, the segmentation result is used to reconstruct a new optimal shape prior using the statistical model for the next iteration. The reconstructed shape is optimized over an edge weighted unsigned distance map calculated from the current segmentation, and the optimized shape is then used as a shape prior for the next iteration of the segmentation. Finally, this process is repeated until convergence, Fig. 4.6. The proposed method is robust to high-density imaging artifacts as quantitatively and qualitatively evaluated on the challenging problem of 3D masseter muscles segmentation from CT datasets. The main advantages of the proposed method are efficient ways: 1) to integrate a 3D shape model into the regional term of the GC; 2) to constrain the shape model to maximize the robustness with respect to noise, artifacts and shape degeneration; 3) to iteratively optimize the shape prior and intensity profile based on the evidence from the previous segmentation which increases the segmentation accuracy.

This chapter is organized as follows; Section 4.2 describes how a PPCA based statistical shape model is used to reconstruct a complete shape based on partial information, and how a constrained variability statistical model can be computed. Section 4.3 describes the segmentation model and how the shape prior knowledge is added into the MRF framework and how a shape prior is updated in an iterative segmentation process. Section 4.5 provides the evaluation of the proposed approach on the segmentation of masseter muscle. Finally, Sec. 4.6 concludes the chapter.

4.2 PPCA based Statistical Shape Model

The details of the Probabilistic PCA have already been explained in Sec. 2.3.2. Here the PPCA has been used to reconstruct a complete shape based on partial information

in the form of manually labeled landmarks (see Sec. 4.2.1). The theory of PPCA can also be used to compute the remaining variability of the statistical model once the information from the landmarks is available (see Sec. 4.2.2) [132].

4.2.1 Reconstruction from Partial Information

A surface mesh \mathbf{x} can be partitioned into $\mathbf{x} := (\mathbf{x}_a, \mathbf{x}_b)^T$, where $\mathbf{x}_b \in \mathbb{R}^{3lp}$ is the known information based on lp -landmarks, while \mathbf{x}_a is unknown that is estimated from the known information. It is assumed that this additional knowledge is provided in the form of lp -landmark points. These landmarks are obtained from manually labeled landmarks, which in this work are the muscle attachments at the facial bones. The partial information \mathbf{x}_b can be used to reconstruct the full shape \mathbf{x} , which is discussed in this section. Once \mathbf{x}_b is known, it is natural to assume that the variability or uncertainty of the statistical model will reduce as further evidence is obtained. The next Sec. 4.2.2 deals with how to compute the remaining variance of a statistical model from the partial information \mathbf{x}_b . The probability distribution function of shape \mathbf{x} can be written as

$$P(\mathbf{x}) = P(\mathbf{x}_a, \mathbf{x}_b) = \mathcal{N}\left(\begin{bmatrix} \bar{\mathbf{x}}_a \\ \bar{\mathbf{x}}_b \end{bmatrix}, \begin{bmatrix} \mathbf{W}_a \mathbf{W}_a^T & \mathbf{W}_a \mathbf{W}_b^T \\ \mathbf{W}_b \mathbf{W}_a^T & \mathbf{W}_b \mathbf{W}_b^T \end{bmatrix} + \sigma \mathbf{I}_{3lp}\right), \quad (4.1)$$

where, \mathbf{I}_{3lp} is a $3lp \times 3lp$ identity matrix, $\mathbf{W} = [\mathbf{W}_a \mathbf{W}_b]^T \in \mathbb{R}^{3m \times n}$ is the scaled eigenvectors and σ is a parameter that controls the remaining variance of the statistical model. It is assumed that \mathbf{x} has a multivariate normal distribution, therefore, the conditional distribution $P(\mathbf{x}_a | \mathbf{x}_b) \sim \mathcal{N}(\bar{\mathbf{x}}_{\mathbf{x}_a | \mathbf{x}_b}, \Sigma_{\mathbf{x}_a | \mathbf{x}_b})$ is also a multivariate normal distribution with mean $\bar{\mathbf{x}}_{\mathbf{x}_a | \mathbf{x}_b}$ and covariance $\Sigma_{\mathbf{x}_a | \mathbf{x}_b}$, which can be given by Eq. 4.2 and Eq. 4.3 respectively [132]

$$\bar{\mathbf{x}}_{\mathbf{x}_a | \mathbf{x}_b} = \mathbf{x}_a + \mathbf{W}_a \mathbf{W}_b^T (\mathbf{W}_b \mathbf{W}_b^T + \sigma \mathbf{I}_{3lp})^{-1} (\mathbf{x}_b - \bar{\mathbf{x}}_b), \quad (4.2)$$

$$\Sigma_{\mathbf{x}_a | \mathbf{x}_b} = (\mathbf{W}_a \mathbf{W}_a^T + \sigma \mathbf{I}_{3lp}) - \mathbf{W}_a \mathbf{W}_b^T (\mathbf{W}_b \mathbf{W}_b^T + \sigma \mathbf{I}_{3lp})^{-1} \mathbf{W}_b \mathbf{W}_a^T. \quad (4.3)$$

It is not possible to compute $\Sigma_{\mathbf{x}_a | \mathbf{x}_b}$ directly for it is potentially huge. As the shape \mathbf{x} can be determined from $\boldsymbol{\alpha}$, the coefficients of the main modes of variation of the statistical model, and the distribution of $\boldsymbol{\alpha}$ in turn can be estimated from the partial information \mathbf{x}_b as $p(\boldsymbol{\alpha} | \mathbf{x}_b)$. It is known that

$$P(\mathbf{x}_b | \boldsymbol{\alpha}) = \mathcal{N}(\mathbf{W}_b \boldsymbol{\alpha} + \bar{\mathbf{x}}_b, \sigma^2 \mathbf{I}_{3lp}), \quad (4.4)$$

4. STATISTICAL MODEL BASED SEGMENTATION

and since $P(\boldsymbol{\alpha}) \sim \mathcal{N}(0, \mathbf{I}_n)$, then using the Theorem of the Linear Gaussian Model [117]

$$P(\boldsymbol{\alpha}|\mathbf{x}_b) = \mathcal{N}(\mathbf{M}^{-1}\mathbf{W}_b^T\sigma^{-2}(\mathbf{x}_b - \bar{\mathbf{x}}_b), \mathbf{M}^{-1}), \quad (4.5)$$

where, $\mathbf{M} = \sigma^{-2}\mathbf{W}_b^T\mathbf{W}_b + \mathbf{I}_n$. $\mathbf{W}_b^T\mathbf{W}_b$ is generally small and can be easily computed. Once $\boldsymbol{\alpha}$ has been determined from the partial information \mathbf{x}_b , the most likely reconstructed shape $\tilde{\mathbf{x}}$ can be calculated using

$$\mathbf{x}_{\mathbf{x}_a|\mathbf{x}_b} = \tilde{\mathbf{x}} = \arg \max_x p(\mathbf{x}|\boldsymbol{\alpha}) = \mathbf{W}\boldsymbol{\alpha} + \bar{\mathbf{x}}. \quad (4.6)$$

Here, $\tilde{\mathbf{x}}$ also happens to be the mean shape $\bar{\mathbf{x}}_{\mathbf{x}_a|\mathbf{x}_b}$ of the statistical model with constrained variability. If $\sigma > 0$, then \mathbf{x}_b is allowed to move. Larger σ means higher remaining variance and the landmarks are allowed to deviate more from the marked locations. Smaller values of σ mean lower variance and the landmarks stay closer to the marked locations but then the model undergoes over fitting.

4.2.2 Remaining Variance

The known lp -landmarks can be further utilized to constrain the variability of the statistical model. Since \mathbf{x}_b provides additional information, in a probabilistic setting it is natural to assume that the uncertainty of the statistical model will reduce as further evidence is obtained. $P(\boldsymbol{\alpha}|\mathbf{x}_b)$ (see Eq. 4.5) also models the remaining variability of the statistical model given the partial information \mathbf{x}_b . Reconstructing a complete shape from partial information does not require the computation of the full covariance matrix $\Sigma_{\mathbf{x}_a|\mathbf{x}_b}$. However, the full covariance matrix $\Sigma_{\mathbf{x}_a|\mathbf{x}_b}$ has to be computed if the remaining variability of the statistical model needs to be calculated which can be written using the standard formula as

$$\Sigma_{\mathbf{x}_a|\mathbf{x}_b} = \mathbf{W}_{ML}\mathbf{M}^{-1}\mathbf{W}_{ML}^T, \quad (4.7)$$

if $\mathbf{W} = \mathbf{W}_{ML}$. The covariance matrix $\Sigma_{\mathbf{x}_a|\mathbf{x}_b}$ then decomposes as

$$\mathbf{W}_{ML}\mathbf{M}^{-1}\mathbf{W}_{ML}^T = \mathbf{U}\underbrace{\{(\mathbf{D}^2 - \sigma^2\mathbf{I}_n)^{\frac{1}{2}}\mathbf{M}^{-1}(\mathbf{D}^2 - \sigma^2\mathbf{I}_n)^{\frac{1}{2}}\}}_{\mathbf{A}}\mathbf{U}^T. \quad (4.8)$$

As it is prohibitively large, Lüthi *et al.* [132] suggested applying SVD on \mathbf{A} , that is,

$$\text{SVD}(\mathbf{A}) := \tilde{\mathbf{U}}\mathbf{D}_{\mathbf{x}_a|\mathbf{x}_b}^2\tilde{\mathbf{U}}^T. \quad (4.9)$$

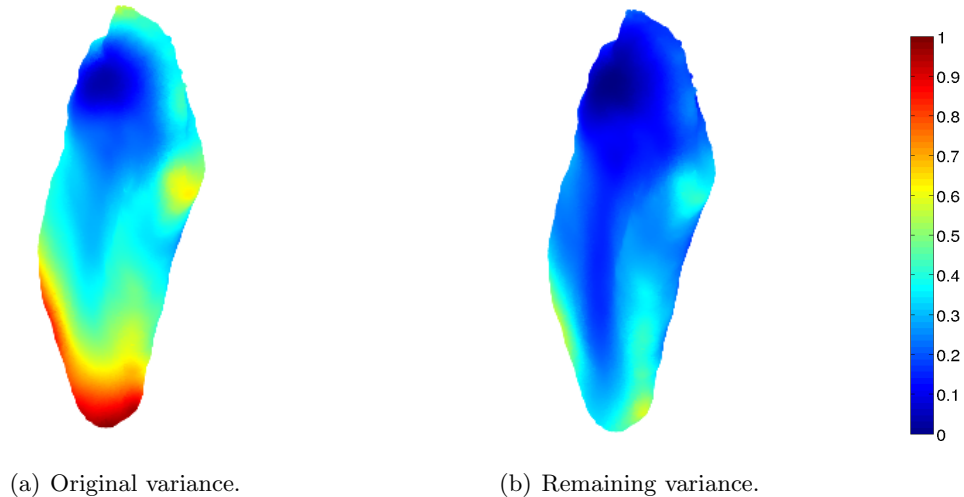


Figure 4.1: Normalized variances of the statistical model.

Thus we have an eigenvalue decomposition

$$\mathbf{W}\mathbf{M}^{-1}\mathbf{W}^T = \underbrace{(\mathbf{U}\tilde{\mathbf{U}})}_{\text{Eigenvectors}} \underbrace{\mathbf{D}_{\mathbf{x}_a|\mathbf{x}_b}^2}_{\text{Eigenvalues}} (\mathbf{U}^T\tilde{\mathbf{U}}^T), \quad (4.10)$$

where $\mathbf{U}\tilde{\mathbf{U}}$ are the orthogonal principal components and $\mathbf{D}_{\mathbf{x}_a|\mathbf{x}_b}^2$ is the diagonal matrix of the corresponding variances. $\mathbf{U}\tilde{\mathbf{U}}$, $\mathbf{D}_{\mathbf{x}_a|\mathbf{x}_b}^2$ and $\tilde{\mathbf{x}}$ can now be used to generate a new shape \mathbf{x}^* with the remaining flexibility of the model using

$$\mathbf{x}^* = \tilde{\mathbf{x}} + (\mathbf{U}\tilde{\mathbf{U}})\mathbf{D}_{\mathbf{x}_a|\mathbf{x}_b}\boldsymbol{\alpha}. \quad (4.11)$$

As an illustration of the concept of remaining variability, we show in Fig. 4.1 the original variance of the model (a) and the remaining variance of the model after being fit to the muscle attachments (b). For an in-depth analysis of the reconstruction of a complete shape given partial information and calculating the remaining variance see Lüthi *et al.* [132].

4.3 Segmentation Model

The segmentation problem is casted as a binary labeling problem in the MRF framework (see Sec. 2.2). Let $\mathbf{P} = \{p_1, p_2, \dots, p_N\}$ be the set of N voxels of the volume dataset and

4. STATISTICAL MODEL BASED SEGMENTATION

let a neighborhood system $\mathbb{N} = \{N_p | \forall p \in \mathbf{P}\}$ defined over the set of voxels \mathbf{P} , where N_p consists of all unordered 26 neighbors of the voxel $p \in \mathbf{P}$. Let $\mathbf{Z} = (z_1, z_2, \dots, z_N)$ be a field of random variables that take values from the set of binary labels $\mathbf{L} = \{0, 1\}$, where 0 stands for background and 1 for object. The goal of the image segmentation problem is to find an optimal mapping $\phi : \mathbf{P} \mapsto \mathbf{L}$ that assigns a label from \mathbf{L} to the random variable Z_p which in turn is equivalent to assigning label z_p to the corresponding pixel $p \in \mathbf{P}$ to each voxel $p \in \mathbf{P}$. The general energy function as explained in Sec. 2.2.1 is given by

$$E(\mathbf{z}|\mathbf{I}) = \sum_{p \in \mathbf{P}} V_p(z_p|\mathbf{I}) + \sum_{p \in \mathbf{P}} \sum_{q \in N_p} V_{p,q}(z_p, z_q|\mathbf{I}), \quad (4.12)$$

where \mathbf{z} is a binary variable which defines the segmentation, \mathbf{I} is the observed intensity data, $V_p(z_p|\mathbf{I})$ and $V_{p,q}(z_p, z_q|\mathbf{I})$ are the regional and the boundary terms respectively. The regional term encodes the individual penalties for assigning voxel p to object and background given the observed image data \mathbf{I} . The boundary term encodes a discontinuity penalty between adjacent voxels p, q incurred when they are jointly assigned labels z_p and z_q where $q \in N_p$ given the observed image data \mathbf{I} . They are based on the traditional GC intensity based energy function of Boykov and Jolly [78].

4.3.1 Integrating Shape Prior Term

The shape knowledge is incorporated by adding an additional shape prior term $V_p(z_p|\mathbf{x}^*)$ in the regional term of the energy functional $E(\mathbf{z}|\mathbf{I}, \mathbf{x}^*)$. The muscle segmentation problem is then solved by minimizing the modified energy functional given by

$$E(\mathbf{z}|\mathbf{I}, \mathbf{x}^*) = \sum_{p \in \mathbf{P}} \left\{ V_p(z_p|\mathbf{I}) + \mu V_p(z_p|\mathbf{x}^*) \right\} + \lambda \sum_{p \in \mathbf{P}} \sum_{q \in N_p} V_{p,q}(z_p, z_q|\mathbf{I}), \quad (4.13)$$

where \mathbf{x}^* is the optimal reconstructed shape as explained in Sec. 4.2 which is used as the shape prior, λ is the smoothness parameter and μ is the shape prior parameter. The shape term encodes how likely a particular voxel p is to belong to the object “1” and the background “0”, given the shape prior \mathbf{x}^* .

A probability map for the voxels to belong to the object and the background is created using the unsigned distance map of the shape prior’s contour. All the voxels enclosed by the shape prior are used to create the object probability map. The voxels at the center of the shape are more likely to belong to the object as compared to the voxels close to the contour (see Fig. 4.2(a)). Thus the voxels at the center incur a

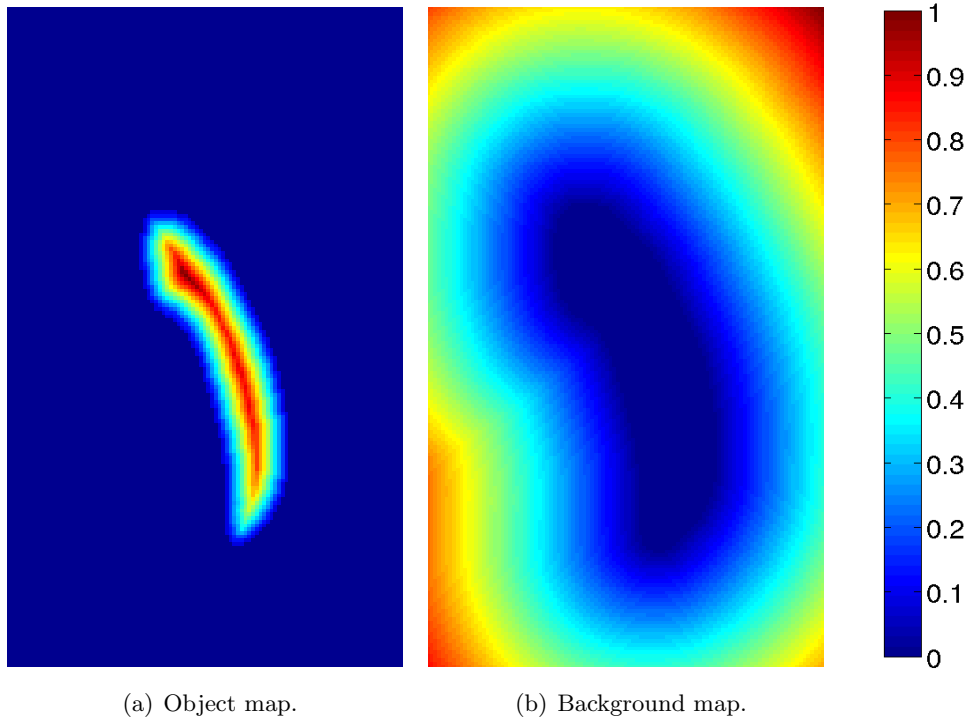


Figure 4.2: Probability maps.

higher penalty to belong to the background as compared to voxels close to the contour. Voxels outside the contour incur 0 penalties for belonging to background. The opposite is true for the background probability map, where only the voxels outside the contour are considered. The voxels close to the contour are more likely to belong to object as compared to the voxels farther away (see Fig. 4.2(b)). The penalty is computed as the negative log-likelihood of belonging to object $V_p(z_p = \text{“1”}) = -\ln P(\mathbf{x}_p | \text{“1”})$ and background $V_p(z_p = \text{“0”}) = -\ln P(\mathbf{x}_p | \text{“0”})$. Such encoding of the global shape information transforms it into local constraints.

4.3.2 Initial Shape from the Landmarks

To get an initial estimation of the masseter muscle, procrustes method is used to align the statistical model with the given 6-landmarks (red in Fig. 4.3(b)) which were obtained manually while the corresponding landmarks (blue) on the mean shape are shown in Fig. 4.3(a). A reconstruction of the statistical model given the landmarks, as ex-

4. STATISTICAL MODEL BASED SEGMENTATION

plained in Sec. 4.2.1, is then calculated to get an initial shape prior for the segmentation.

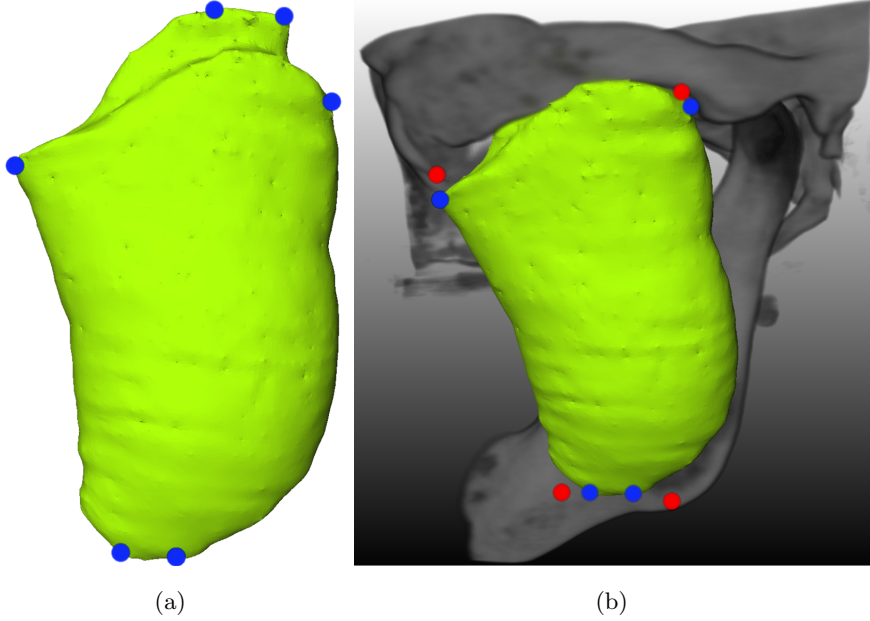


Figure 4.3: Landmarks location: (a) in blue on the mean shape, (b) in red on the dataset.

4.3.3 Shape Optimization

To optimize the statistical model shape prior to fit the current segmentation, an energy cost function, Eq. 4.14, is created (see Fig. 4.4). The weighted unsigned distance map $C(v)$ is a linear combination of the negative log-likelihood map of object C_{obj} , the inverse edge map C_{edge} and an unsigned distance map C_{seg} obtained from the segmentation. The object intensity negative log-likelihood map C_{obj} is calculated using the parzen window. An inverse edge map C_{edge} provides low values where there is an edge. The third term in Eq. 4.14 is the unsigned distance map C_{seg} which is calculated from the segmentation boundary. All the maps are then linearly combined as in Eq. 4.14, Fig. 4.4, to generate the linear cost function $C(v)$. The weighted unsigned distance map $C(v)$ is used as a cost function to fit the statistical shape model. Since the statistical shape model (controlled by the shape coefficients α) is represented by a very dense triangulated surface mesh consisting of 39156 vertices, adjacent vertices are close

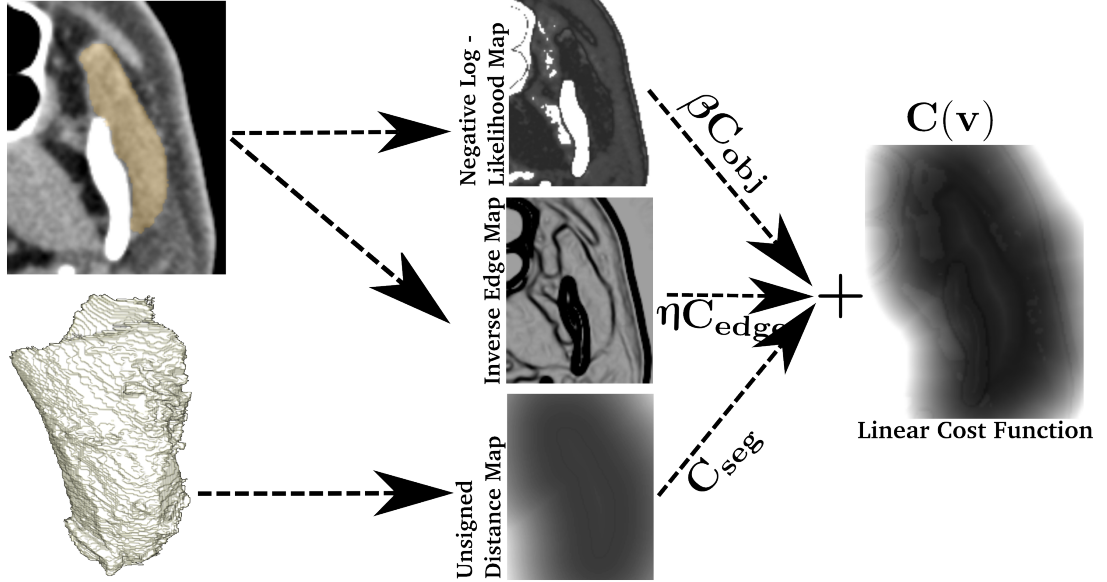


Figure 4.4: Linear cost function.

enough to occupy adjacent voxels. The morphable model of Blanz and Vetter [130] has been used which is very dense (around 76000 vertices) as compared to the active shape model of Cootes *et al.* [129] which are not dense (only 72 vertices). On average for all datasets there are approximately 1.75 vertices per voxel with a density of approximately 7.66 vertices per mm^2 . Once a vertex is in a voxel, the cost of the vertex can be directly read out of the voxel, defined by

$$C(v) = \beta C_{obj}(v) + \eta C_{edge}(v) + C_{seg}(v), \quad (4.14)$$

where η and β are weighting parameters, $v = (v_x, v_y, v_z)$ represents the x , y , z coordinates of a vertex of \mathbf{x} . The objective is to optimize the statistical shape model by minimizing the sum of the cost of vertices constrained by generating only statistically valid shapes, Eq. 4.15. If the statistical model fits perfectly to the segmentation boundary, the cost of such a shape would be 0. Figure 4.5 shows a particular setting of the cost function in an axial slice along with the vertices of the mesh. The sum of the cost of all the vertices for a particular setting of shape coefficients α gives the cost of the

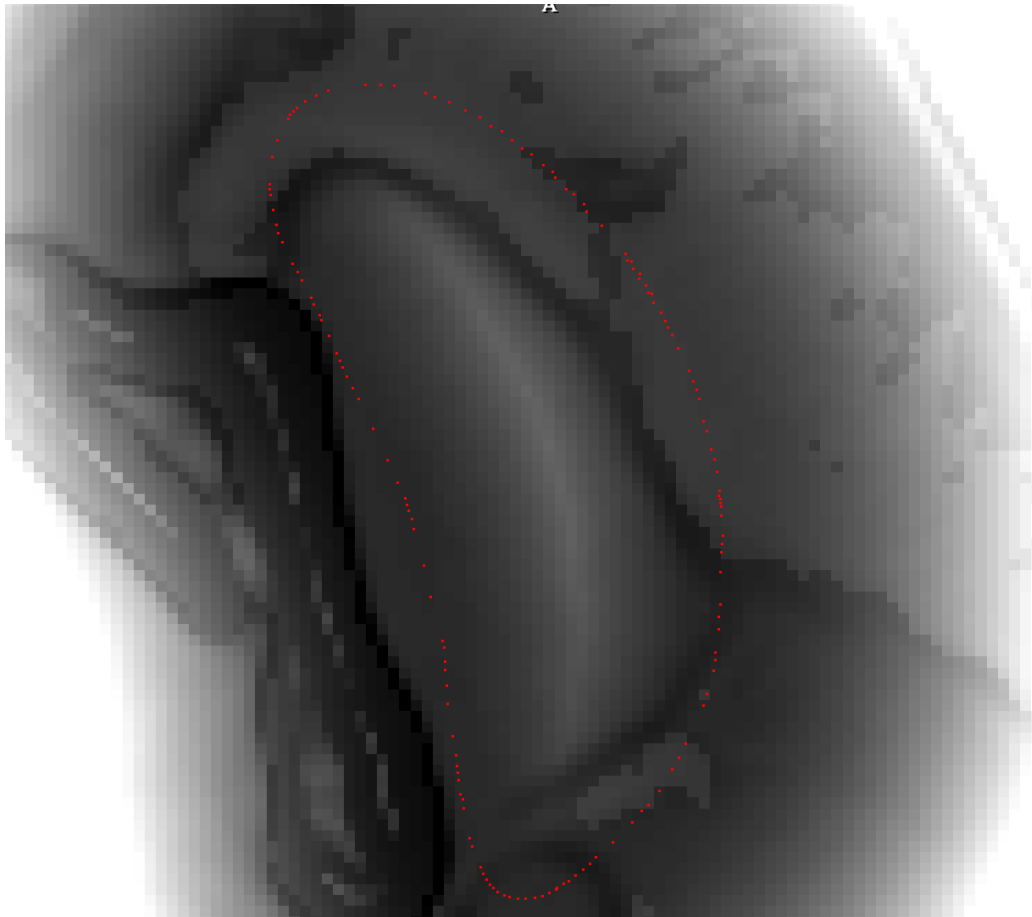


Figure 4.5: In red are the vertices of the mesh in an axial slice which is a weighted combination of negative log-likelihood map of object C_{obj} , inverse edge map C_{edge} and unsigned distance map C_{seg} .

shape as follows

$$C(\mathbf{x}^*) = C(\tilde{\mathbf{x}} + (\mathbf{U}\tilde{\mathbf{U}})\mathbf{D}_{\mathbf{x}_a|\mathbf{x}_b}\boldsymbol{\alpha}) \quad (4.15)$$

$$= \sum_{v=1}^m C(v) . \quad (4.16)$$

The statistical model is optimized by minimizing the sum of the cost of vertices over the cost function $C(v)$. The coefficients $\boldsymbol{\alpha}$ corresponding to the main modes of variation are obtained by solving the minimization problem using a gradient descent based

optimization approach

$$\min_{\alpha} \left\{ C(\tilde{\mathbf{x}} + (\mathbf{U}\tilde{\mathbf{U}})\mathbf{D}_{\mathbf{x}_a|\mathbf{x}_b}\alpha) \right\}. \quad (4.17)$$

Once the optimal α are found, the optimized shape is then constructed using Eq. 4.11 and used as a shape prior for the next iteration.

4.3.4 Constraining the Variability

The theory of variability constrained statistical models (see Sec. 4.2.2) is used during shape optimization as it is more stable as compared to the unconstrained statistical models. The vertices which are fixed at the known landmarks stay fixed or undergo only slight movement from their position; this avoids shrinking of the model. The generally used regularization scheme to constrain the model [136], forces the shape to be as close as possible to the mean shape $\bar{\mathbf{x}}$. Forcing a shape to be close to mean shape, however, decreases the flexibility of the statistical model to fully represent the shape variability of the anatomical structures. The benefit of smooth shapes comes at the cost of a reduced flexibility of the statistical model. Furthermore, it is also difficult to find good regularization parameters for each datasets as the optimal settings might change between datasets. Using the statistical model with constrained variability in contrast provides a model which is stable, generates smooth shapes and has reduced variability but no reduction in shape representational power of the model, because only the unnecessary variability has been removed. It is also important to note that the reduction in variability of the model is patient-specific which is achieved through the use of the landmarks. This constrained statistical model can still be regularized which will force the shape to be close to the optimal shape, as the optimal shape is also the mean shape of the constrained statistical model.

4.4 Algorithm

This section presents the Algorithm 2 along with Fig. 4.6 and summarizes the steps of the proposed scheme for segmenting the masseter muscle. The algorithm starts with an initial shape prior obtained from reconstructing the complete shape from partial information in the form of manually labeled 6-landmark points (see Sec. 4.2.1). It is used to create the shape probability maps for the object (see Fig. 4.2(a)) and the background

4. STATISTICAL MODEL BASED SEGMENTATION

Algorithm 2

Input: Reconstruct optimal shape from landmarks.

Output: GC segmentation of target muscle.

Repeat

- Compute $V_p(z_p|\mathbf{x}^*)$ as explained in Section 4.3.1.
- Create graph according to the energy function $E(\mathbf{z}|\mathbf{I}, \mathbf{x}^*)$.
- Segment using GC.
- Create weighted unsigned distance map from segmentation.
- Optimize the shape model over this unsigned distance map.

Until Convergence

(see Fig. 4.2(b)) as explained in Sec. 4.3.1. A graph \mathcal{G} corresponding to the energy function $E(\mathbf{z}|\mathbf{I}, \mathbf{x}^*)$, Eq. 4.13, is created and the GC segmentation algorithm of Boykov & Kolmogorov [125] is used to obtain the muscle segmentation. If the segmentation has not yet fully converged, then the boundary of the resulting segmentation is used to create a weighted unsigned distance map (see Fig. 4.4 and Sec. 4.3.3), over which the statistical shape model is optimized. The optimized shape is then used as a shape prior for the next iteration to generate probability maps for the object/background and the same process is iteratively applied until the convergence is reached (see Fig. 4.7). Comparison with the previous segmentation is used as the criteria to evaluate the segmentation convergence. If no voxels have changed label then the convergence has been reached or the algorithm run for a maximum of 15 iterations.

4.5 Experimental Setting and Results

The above described segmentation method was tested on 20 CT datasets of the masseter muscle with dataset dimensions $79-156 \times 148-214 \times 125-384$ voxels and spacing of $0.3-0.5 \times 0.3-0.5 \times 0.3-1$ mm³. All datasets were noisy, possessing mild to severe imaging artifacts. The datasets were chosen randomly from the hospital repository so that they were representative for anatomical variations. A medical expert provided the ground truth segmentations that were also used to create the statistical model with constrained

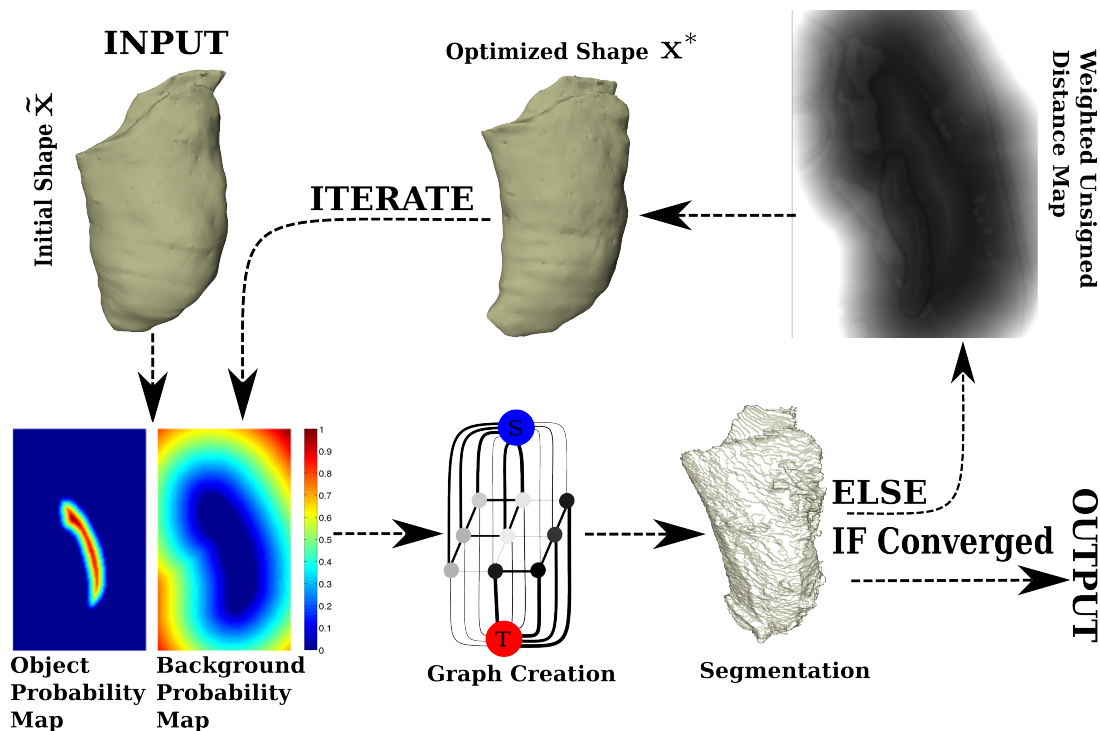


Figure 4.6: Segmentation process.

variances. Leave-One-Out cross-validation was then used to evaluate the method.

As it is difficult to pin-point the muscle attachments on the bone, the landmarks only provide a rough estimation of the location and the shape of the specific patient’s muscle. The proposed technique, however, is unaffected by landmark placement and the model behaves nicely even when the landmarks have not been correctly placed. It was found that the statistical shape model converges to the target muscle shape in 4 – 7 iterations. The segmentation time was around 5 minutes per dataset on an Ubuntu 11.04 PC (Intel(R) Core(TM) i7 CPU (8) 4 cores hyper-threading \times 2.9 GHz, 16 GB RAM). The parameters $\sigma = 2.8$, $\lambda = 0.1$ and $\mu = 0.15$ were optimized on three different datasets and used throughout the entire segmentation experiments. The dice coefficient, sensitivity and specificity of the segmentation were calculated as similarity measures to determine the accuracy of the proposed method called App. B1 here onwards.

Figure 4.7 shows a shape update on one of the corrupted datasets. The ground truth is shown in orange while the statistical shape model is shown in green. It can be seen in Fig. 4.7(e-h) that the statistical model shape update is robust to the high-density

4. STATISTICAL MODEL BASED SEGMENTATION

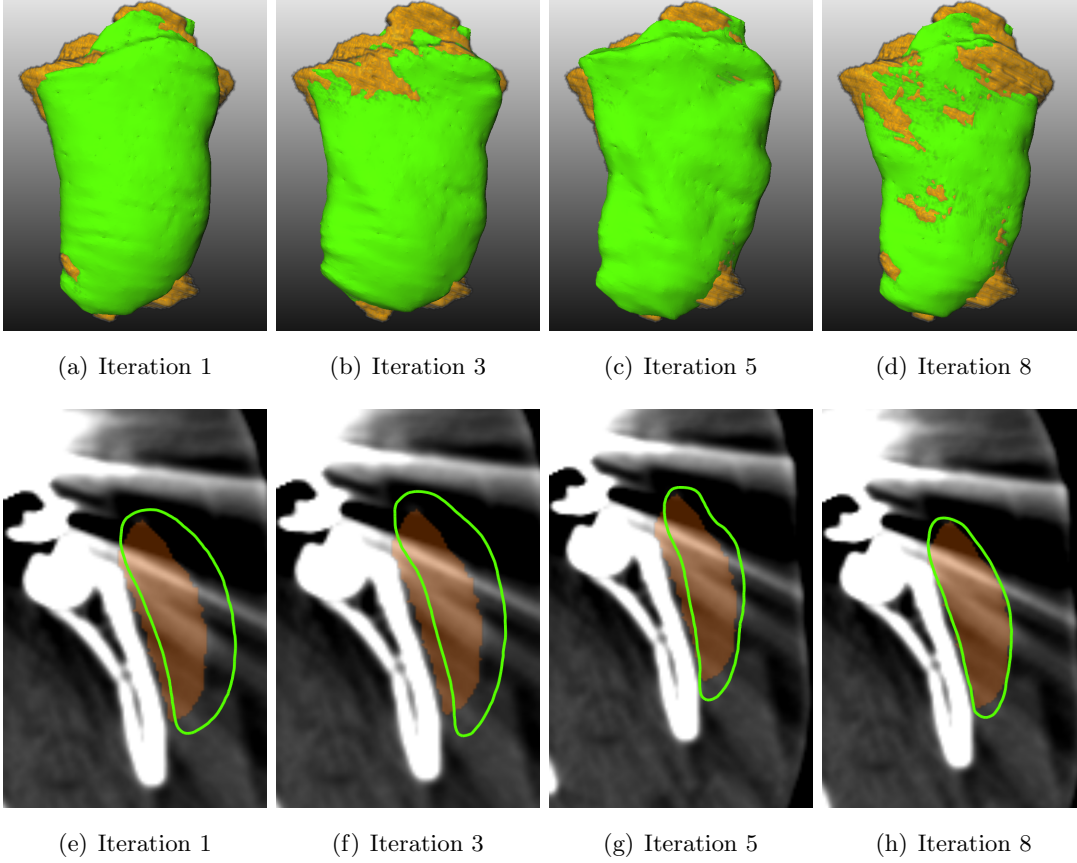
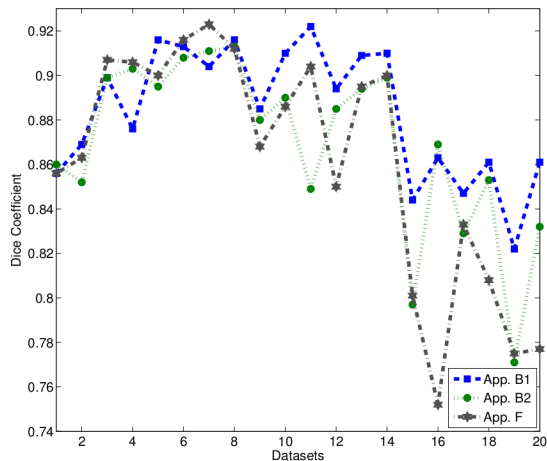


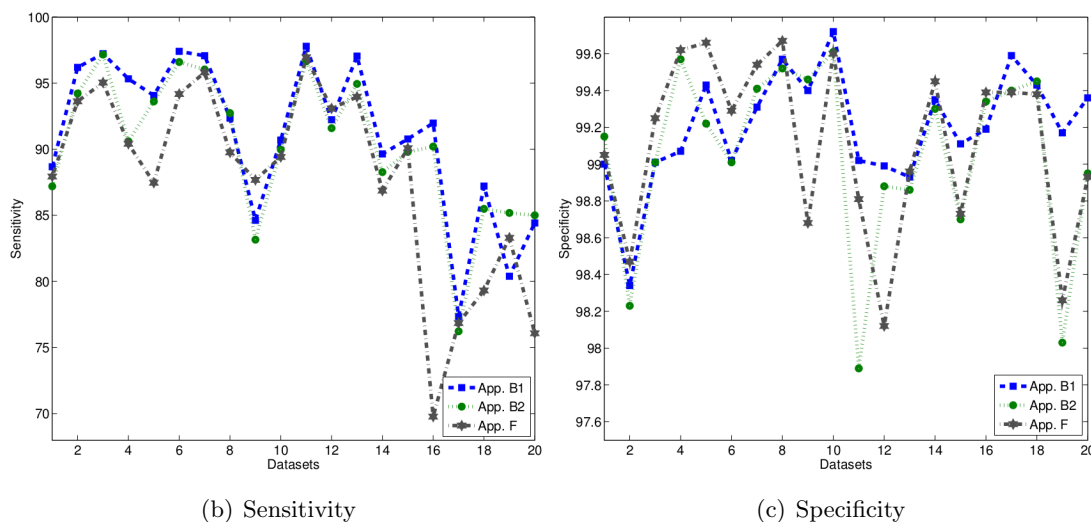
Figure 4.7: The evolution (left to right) of the shape prior (green) to the target muscle (orange) on a corrupted dataset. (a-d) in 3D and their corresponding 2D slices (e-h).

artifacts that corrupted the image data.

The graphs in Fig. 4.8 show the results of App. B1 when the shape prior is added in the regional term (blue curve) and the approach App. B2 which adds shape prior in the boundary term (green curve). The proposed method App. B1 was also compared to the fixed shape template of Freedman & Zhang [92] (App. F - gray curve) which also adds the shape prior in the boundary term. The results clearly demonstrate that by adding the shape prior in the regional term (blue curve) the segmentation result was more accurate than adding the shape prior in the boundary term (green curve). This was especially true for datasets (Datasets 11-20 in Fig. 4.8) with severe high-density artifacts (see Fig. 4.11). Table 4.1 lists the dice coefficients, sensitivity and specificity



(a) Dice coefficient



(b) Sensitivity

(c) Specificity

Figure 4.8: The graphs show dice coefficient, sensitivity and specificity results. Incorporating the shape prior in the regional term (App. B1 shown with blue curve) is more accurate compared to incorporating it in the boundary term (App. B2 shown in green and App. F in red). The red curve shows the result of App. F where a fixed shape template is used.

values of the three approaches; adding shape prior in regional terms (App. B1), adding shape prior in boundary term (App. B2) and adding fixed shape template of Freedman & Zhang [92] (App. F) in boundary term. In addition, Table 4.2 summarizes the values of the dice coefficient encompassing the mean, the standard deviation, the median as

4. STATISTICAL MODEL BASED SEGMENTATION

Data	Similarity Measure								
	Dice Coefficient			Sensitivity%			Specificity%		
	App. B1	App. B2	App. F	App. B1	App. B2	App. F	App. B1	App. B2	App. F
1	0.856	0.860	0.856	88.67	87.19	87.93	99.00	99.15	99.05
2	0.869	0.852	0.863	96.18	94.21	93.62	98.34	98.23	98.47
3	0.899	0.899	0.907	97.22	97.16	95.04	99.01	99.01	99.25
4	0.876	0.903	0.906	95.32	90.60	90.41	99.07	99.57	99.62
5	0.916	0.895	0.900	94.10	93.61	87.46	99.43	99.22	99.66
6	0.913	0.908	0.916	97.40	96.59	94.17	99.02	99.01	99.29
7	0.904	0.911	0.923	97.07	96.03	95.81	99.31	99.41	99.54
8	0.916	0.913	0.912	92.35	92.73	89.76	99.57	99.52	99.67
9	0.885	0.880	0.868	84.62	83.16	87.68	99.40	99.46	98.68
10	0.910	0.890	0.886	90.69	89.97	89.41	99.72	99.61	99.60
11	0.922	0.849	0.904	97.78	96.64	96.95	99.02	97.89	98.81
12	0.894	0.885	0.850	92.23	91.59	93.08	98.99	98.88	98.12
13	0.909	0.894	0.895	97.04	94.94	93.96	98.93	98.86	98.96
14	0.910	0.899	0.900	89.64	88.27	86.87	99.35	99.30	99.45
15	0.844	0.797	0.802	90.77	89.80	90.07	99.11	98.70	98.73
16	0.863	0.869	0.752	91.97	90.20	69.78	99.19	99.34	99.39
17	0.847	0.829	0.833	77.34	76.23	76.85	99.59	99.40	99.39
18	0.861	0.853	0.808	87.21	85.49	79.32	99.43	99.45	99.38
19	0.822	0.771	0.775	80.39	85.18	83.30	99.17	98.03	98.26
20	0.861	0.832	0.777	84.42	85.00	76.08	99.36	98.95	98.93

Table 4.1: Quantitative comparison of App. B1 with the approaches App. B2 and App. F.

	DC (Mean \pm Std)	DC (Median)	DC (Smallest - Largest)
App. B1	0.884 \pm 0.029	0.890	(0.822 - 0.922)
App. B2	0.869 \pm 0.039	0.883	(0.771 - 0.913)
App. F	0.862 \pm 0.053	0.877	(0.752 - 0.923)

Table 4.2: The table list the mean & standard deviation, the median and the smallest as well as the largest value of the dice coefficient of methods App. B1, App. B2 and App. F.

well as the smallest and the largest values. From the Table 4.2 it can be seen when

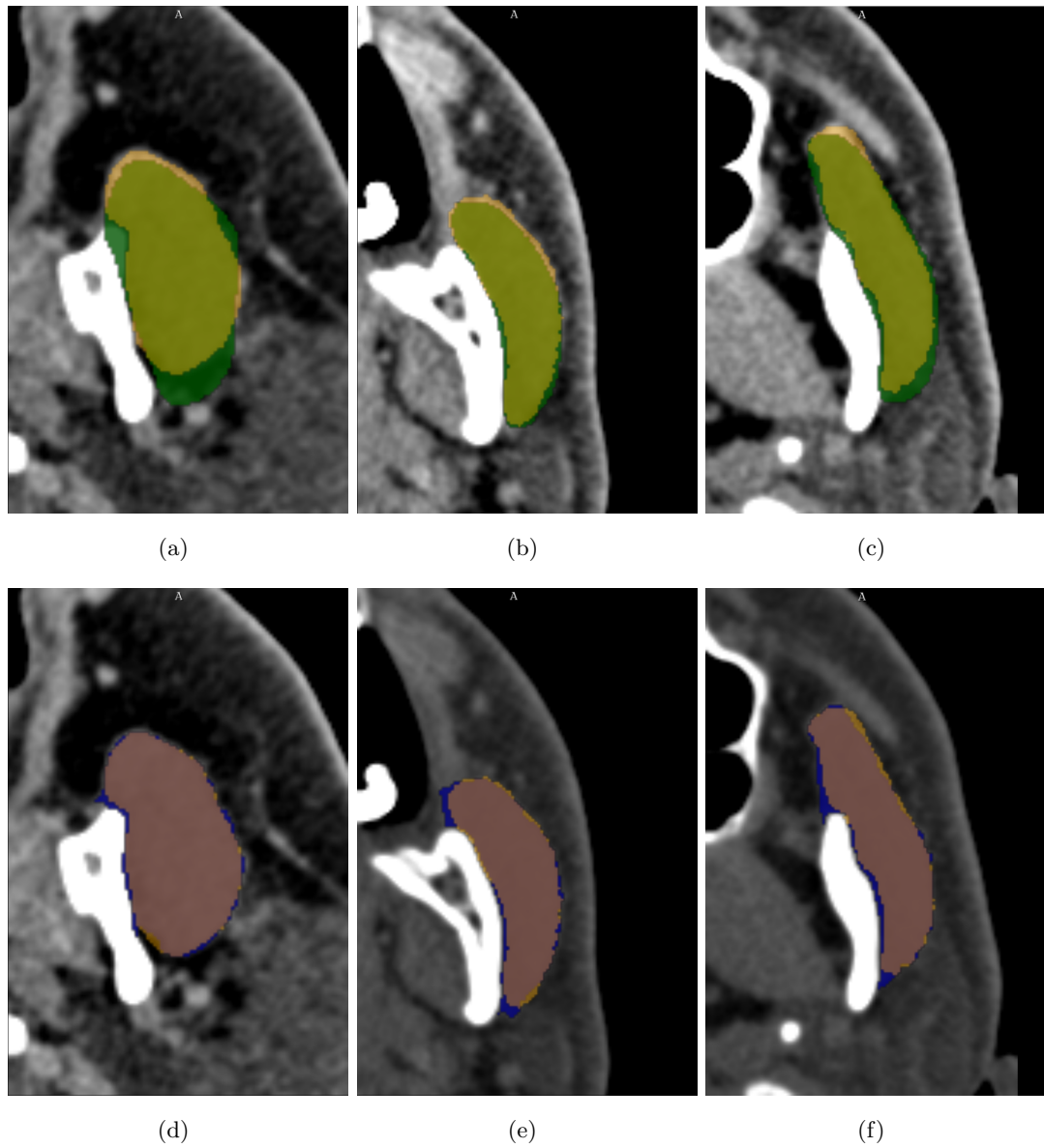


Figure 4.9: Shape prior and the segmentation result in 2D. (a-c) show shape prior (green) overlaid with ground truth (orange), while (d-f) show the segmentation (blue).

compared to Table 3.1, that there is an increase of the segmentation accuracy from a mean value of 0.869 for App. A to 0.884 for App. B1 of the dice coefficient. It can also be seen that App. B1 is more stable than the previous approach (App. A) as the standard deviation of the current approach (App. B1) is lower than that of the previous

4. STATISTICAL MODEL BASED SEGMENTATION

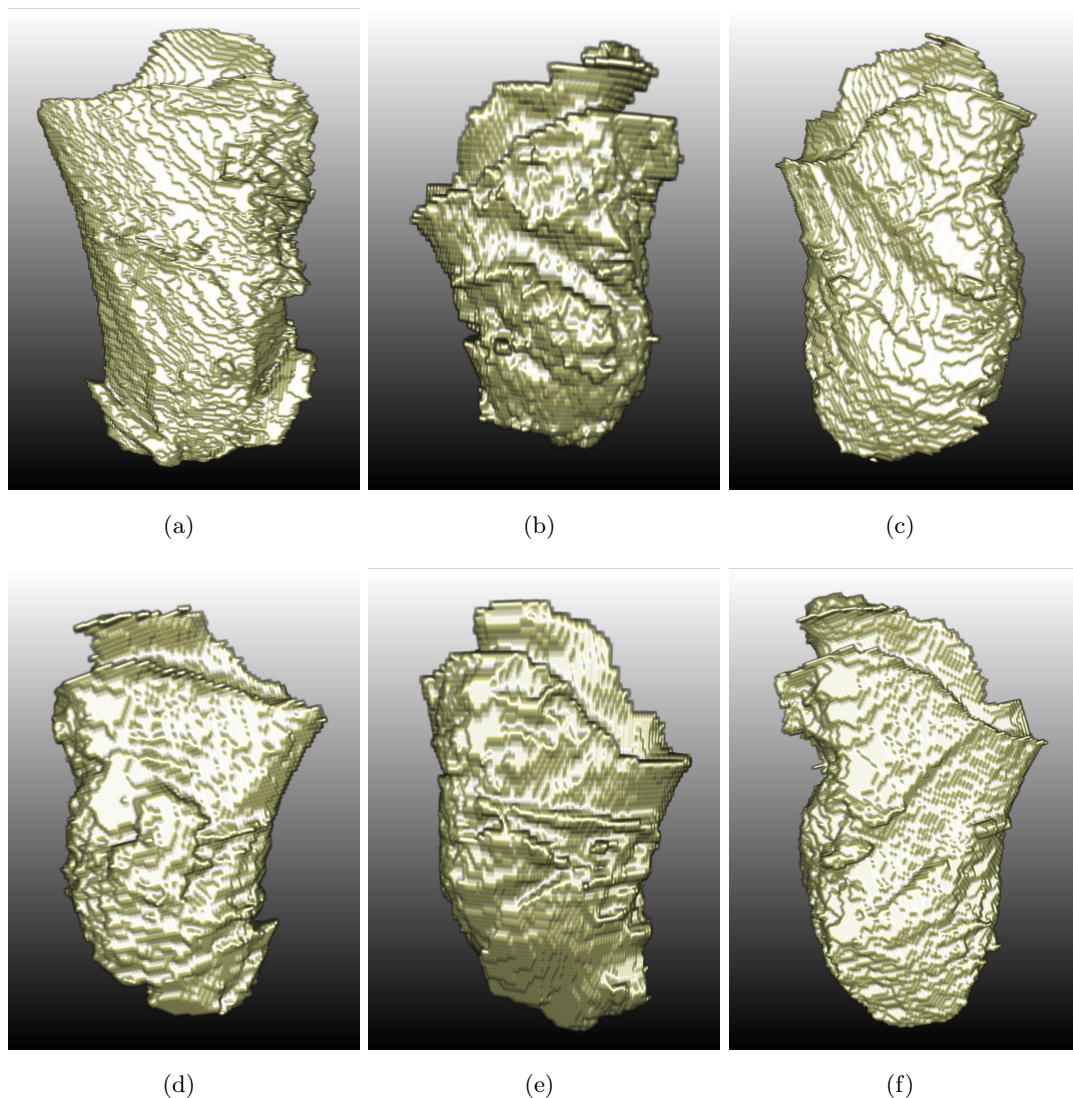


Figure 4.10: Segmentation result of the left (a-c) and the right muscles (d-e) in 3D.

approach.

A statistical comparison of App. B1 with App. F was also conducted and statistical analysis (t-test with significance level 0.05) was performed over the dice coefficient values to determine whether the two methods are statistically significantly different. Statistical analysis showed that App. B1 performed statistically significantly better (p-value = 0.008) for all datasets. App. B1 also performed statistically significantly

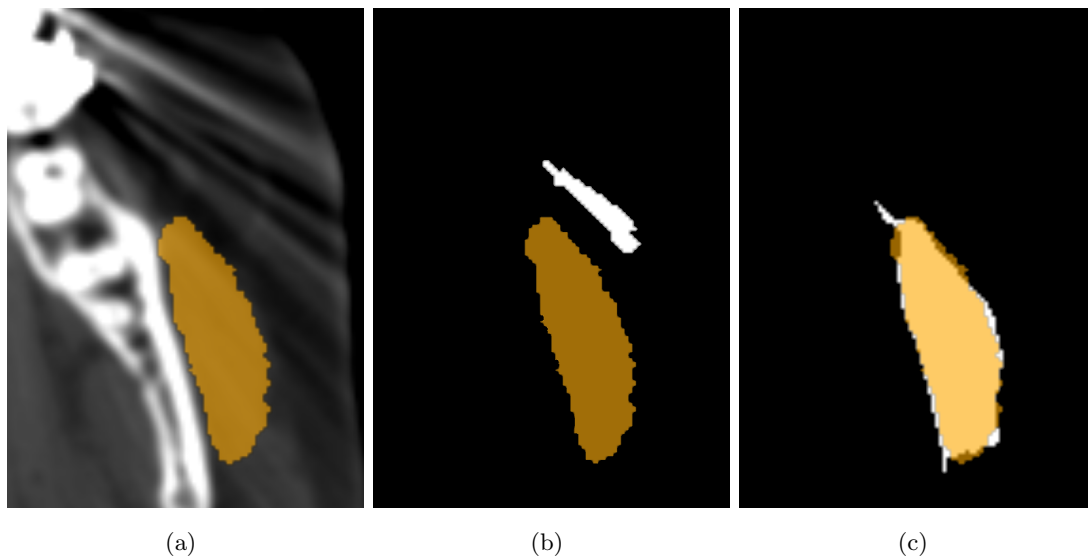


Figure 4.11: High density imaging artifact corrupted slice (a) with ground truth in orange overlaid in all the figures while the segmentation result is in white in (b,c). Segmentation result with the shape prior in boundary term (App. B2) in (b) is worse than the shape prior in regional term (App. B1)(c).

better (p-value = 0.01) than App. B2 for all datasets. This clearly shows the increased robustness gained by including a shape prior in the regional term of the energy function.

The shape prior in the boundary term has the effect of smoothing out regions which results in a portion of the muscle being smoothed out as can be seen in Fig. 4.11(b)). This can be dealt with by incorporating the shape prior in the regional term (App. B1) which allows the flexibility to encode shape knowledge in areas where the intensity alone does not provide sufficient knowledge because of the artifacts. In such regions, the shape prior in the regional term can offset the intensity likelihood and estimate the likelihood of voxels belonging to the muscle based on the spatial location of the voxel with respect to the shape prior as can be seen in Fig. 4.11(c)). This is the main motivation of adding a shape prior in the regional term only. In the previous approach App. A the shape knowledge was added in the regional term in addition to the boundary term to circumvent the problem with respect to the high density image artifacts. With App. B1, it was found that adding the shape knowledge just in the regional term of the energy function is sufficient to obtain a robust segmentation with respect to the high

4. STATISTICAL MODEL BASED SEGMENTATION

density artifacts.

As App. F used a fixed shape template as shape prior in the boundary term, their results were worse than the results obtained by App. B1 that adds the shape prior in the regional term especially for severely corrupted and noisy datasets. The reason being that a fixed shape template cannot adapt to the different anatomical shape variations of the muscle. App. A have also used a fixed shape template added in both the regional and the boundary terms with a 6-neighborhood system. Here, higher segmentation accuracy has been achieved as compared to App. A for the masseter muscle segmentation by employing a statistical model that accommodates for increased shape variability and a 26-neighborhood system. While Malcolm *et al.* [102] can generate non-linear combinations of shapes which encode a higher flexibility than App. B1 based on linear statistical model, it has been shown, that App. B1 can still achieve comparable segmentation results, and a complex nonlinear shape model is thus not required.

Figure 4.9 shows qualitative results of App. B1 in 2D, while the qualitative results in 3D are shown in Fig. 4.10. App. B1 used a 26-neighborhood system which has resulted in better segmentation as compared to a 6-neighborhood system. The experimental results obtained using App. B1 is clinically acceptable as validated by medical experts from the department of CMF surgery of University Hospital Basel.

4.6 Conclusion

A segmentation approach has been presented that combines a 3D statistical shape model with an MRF-based segmentation framework, which is robust with respect to high-density artifacts. The statistical shape model was learned using a PPCA based approach which was instrumental in providing the model with a constrained variability. Using a constrained variability statistical model has proven to be a more sensible alternative for constraining the model. The shape knowledge, provided as a triangulated surface mesh, was transformed into the local shape constraints of the regional term (App. B1) of the MRF-based energy functional, and in a novel manner, the shape was updated over the edge-weighted unsigned distance map in an iterative process. The method's performance was shown by segmenting the masseter muscle from CT datasets which is a challenging task because of the presence of soft-tissues in close proximity as well as imaging artifacts. The results demonstrated that by incorporating prior shape

knowledge, clinically acceptable results can be achieved. Although App. B1 has been applied for the segmentation of the masseter muscle, the method itself is general and can be used for the segmentation of any object of interest from 2D or 3D datasets, whenever a statistical model is available.

4. STATISTICAL MODEL BASED SEGMENTATION

5

Approach C: Graph Cut Segmentation using a Constrained Statistical Model with Non-Linear and Sparse Shape Optimization

5.1 Introduction

The material presented in this chapter is an extended version of the paper that was published in Majeed *et al.* [21].

The previous approach App. B1 incorporated a 3D statistical model as shape knowledge in the MRF framework. Although it provided better segmentation results compared to the technique which incorporated a fixed shape template App. A, it can still be further improved by making the shape update mechanism non-linear. The shape update mechanism in App. B1 used a linear shape update mechanism where all vertices incur the same penalty regardless of their variability.

This approach named App. C here onwards also combines shape knowledge obtained from a variability constrained statistical model into the MRF segmentation framework and outlines a segmentation method with a non-linear shape update mechanism and is based on the earlier approach App. B1 but extends it in several ways. In particular

5. STATISTICAL MODEL BASED SEGMENTATION WITH NON-LINEAR AND SPARSE OPTIMIZATION

a non-linear cost function together with L^1 regularization [137] has been introduced to provide bolder and more accurate shape update than that of App. B1 which uses a linear cost function. The employed statistical model based on PPCA allows computing local information about the remaining variance *i.e.* uncertainty about the correct segmentation boundary. This knowledge about the local segmentation uncertainty is then used to construct a prior with a non-linear shape update mechanism, where a higher cost is incurred in locations with little uncertainty and a low cost for shifting the segmentation boundary in locations with high uncertainty. In other words, a vertex that has lower variability incurs a higher penalty if it moves the same amount of distance than a vertex with higher variability. This helps in finer shape fitting of the statistical model to the current segmentation in areas where the shape variability is high.

Experimental results for segmenting the masseter muscle from CT data are presented showing the advantage of including the knowledge about local segmentation uncertainties into the segmentation framework. The main advantages of App. C is that non-linear cost function and L^1 regularization provides better shape update and guard against the statistical model from degenerating.

This chapter is organized as follows: Section 5.2 lays out the segmentation framework and how shape knowledge is extracted from the variability constrained statistical model. The creation of the non-linear cost function over which the statistical model is optimized is detailed in Sec. 5.3. The complete algorithm is given in Sec. 5.4. Section 5.5 provides the results of applying the proposed method App. C to segment masseter muscle and finally Sec. 5.6 concludes this chapter.

5.2 Segmentation Framework

The segmentation problem is casted as a binary labeling problem in the MRF framework. Let $\mathbf{L} = \{0, 1\}$ be the set of binary labels, “1” for object and “0” for background, \mathbf{P} be the set of voxels of the volume dataset and $\mathbf{z} = \{z_p : p \in \mathbf{P}, z_p \in \mathbf{L}\}$ be the set of labellings which defines the segmentation. The goal of segmentation is to find a labeling z , which is a mapping from $\mathbf{P} \mapsto \mathbf{L}$ by minimizing the energy functional

$$E(\mathbf{z}|\mathbf{I}, \mathbf{x}^*) = \sum_{p \in \mathbf{P}} \left\{ V_p(z_p|\mathbf{I}) + \mu V_p(z_p|\mathbf{x}^*) \right\} + \lambda \sum_{p \in \mathbf{P}} \sum_{q \in N_p} V_{p,q}(z_p, z_q|\mathbf{I}), \quad (5.1)$$

where $\mathbb{N} = \{N_p | \forall p \in \mathbf{P}\}$ is an unordered 26-neighborhood system over \mathbf{P} , \mathbf{I} is the observed intensity data, \mathbf{x}^* is the shape prior, λ is the smoothness parameter and μ is the shape parameter. $V_p(z_p|\mathbf{I})$ and $V_{p,q}(z_p, z_q|\mathbf{I})$ are the data and the smoothness terms respectively, based on the image intensity information. The data term encodes how likely a voxel is to belong to object and background given its intensity while the smoothness term encodes the prior assumption about the target object that it consists of a homogeneous region, therefore, the smoothness term assigns a penalty, whenever adjacent voxels p, q are assigned different labels z_p and z_q . The data and the smoothness terms are based on the traditional GC intensity based energy functional of Boykov & Jolly [78].

$V_p(z_p|\mathbf{x}^*)$ is the shape data term which encodes how likely a particular voxel p is to belong to the object “1” and the background “0”, given the shape prior \mathbf{x}^* obtained from a PPCA based statistical model as explained in Sec. 2.3. Shape knowledge is encoded by creating a probability map both for the object and the background from the unsigned distance map of the shape prior’s contour and it is similar to the one explained in Sec. 3.2.2. Based on the closeness to the shape’s contour; the object probability map is created for the voxels enclosed by the contour while the background probability map is created for the voxels not enclosed by the contour.

5.3 Adaptive Shape Optimization

For updating the shape prior with respect to the segmentation, it is proposed to use a shape optimization based on adaptive weights with respect to the remaining variances of the statistical model and sparse shape optimization.

5.3.1 Shape Cost Function

Creating the cost function for shape optimization is the second major step of the algorithm after the segmentation corresponding to the energy function $E(\mathbf{z}|\mathbf{I}, \mathbf{x}^*)$, Eq. 5.1. The surface mesh is very dense consisting of 39156 vertices, thus adjacent vertices are close enough to occupy adjacent voxels. Once a vertex is in a voxel, the cost of the vertex can be directly read out of the voxel, defined by

$$C(v) = \beta C_{obj}(v) + \eta C_{edge}(v) + C_{seg}(v), \quad (5.2)$$

5. STATISTICAL MODEL BASED SEGMENTATION WITH NON-LINEAR AND SPARSE OPTIMIZATION

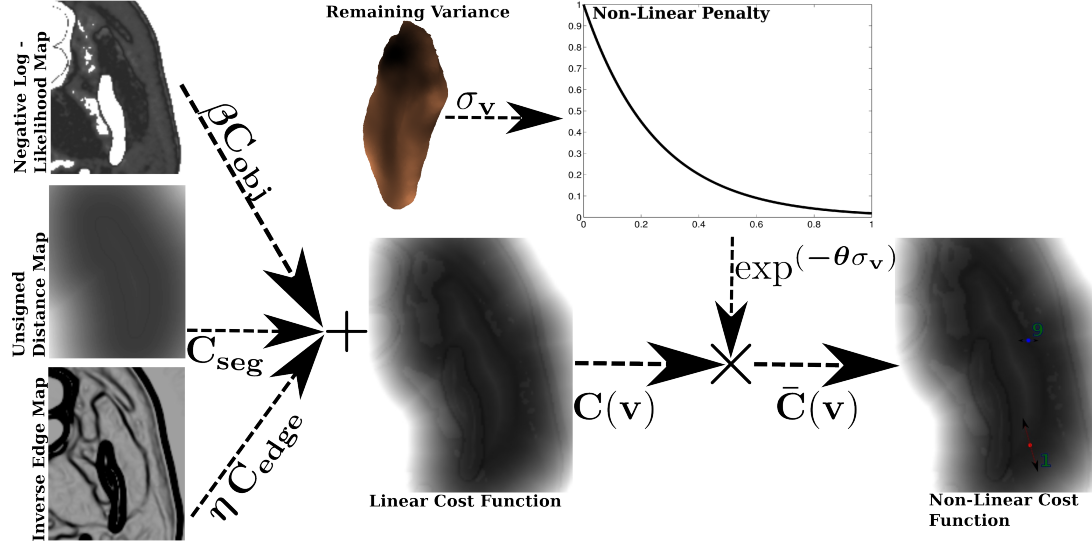


Figure 5.1: Generating non-linear cost function $\bar{C}(v)$.

where η and β are weighting parameters, $v = (v_x, v_y, v_z)$ represents the x , y , z coordinates of a vertex of \mathbf{x} . The object intensity negative log-likelihood map (C_{obj}) is calculated using the parzen window estimation that has already been estimated during the GC segmentation when object and background likelihood are computed. An inverse edge map C_{edge} provides low values where there is an edge. The third term in Eq. 5.2 is the unsigned distance map (C_{seg}) which is calculated from the segmentation boundary. All the maps are then linearly combined (see Eq. 5.2 and Fig. 5.1) to generate the linear cost function $C(v)$ which is then weighted with non-linear variance penalties $\exp(-\theta\sigma_v)$ to generate the non-linear cost function $\bar{C}(v)$, shown in Fig. 5.1.

5.3.2 Shape Optimization

Once the non-linear cost function has been created, the next step is to optimize the statistical model over the generated cost function. It is proposed here (App. C) to use an adaptive cost instead of the linear cost employed in App. B1 in order to make the statistical model more robust against local minima encountered during the shape update. The cost is adapted with respect to the remaining variance of the vertex σ_v . The sum of the cost of all the vertices for a particular setting of shape coefficients α

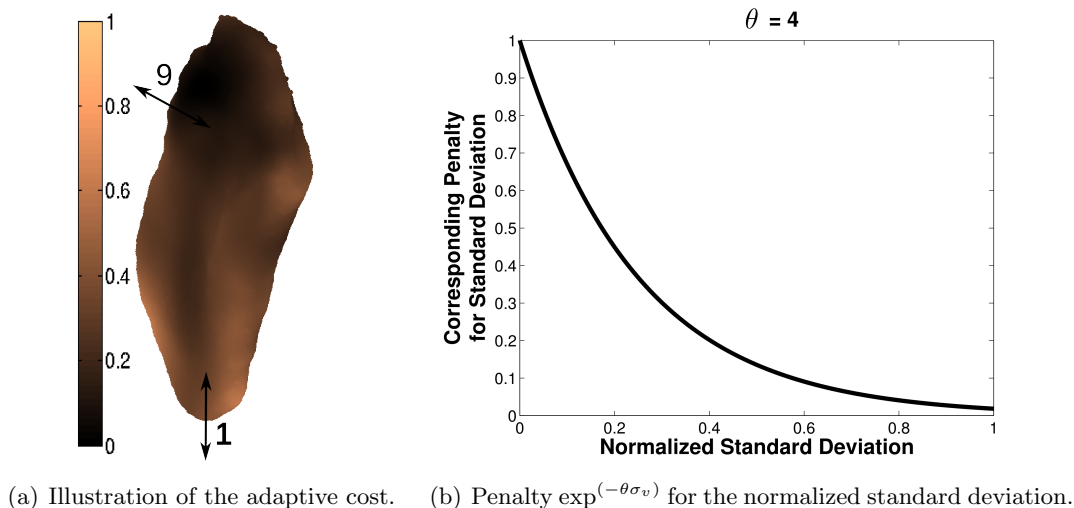


Figure 5.2: Normalized standard deviations and their corresponding penalty.

gives the cost of the shape as follows

$$\bar{C}(\mathbf{x}^*) = C(\tilde{\mathbf{x}} + (\mathbf{U}\tilde{\mathbf{U}})\mathbf{D}_{\mathbf{x}_a|\mathbf{x}_b}\boldsymbol{\alpha}) \exp(-\theta\sigma_v) = \sum_v C(v) \exp(-\theta\sigma_v), \quad (5.3)$$

where σ_v is the remaining variance of vertex v and θ is a weighting parameter.

Here vertices with higher variance incur lower cost in comparison to vertices with lower variance. As a consequence, a vertex with higher remaining variance (color coded in light golden in Fig. 5.2(a)) as given by the statistical model is allowed to move further with less cost (cost 1) while a vertex with lower remaining variance (color coded in black) incurs higher cost (cost 9) when it moves the same distance. With linear cost, vertices irrespective of their remaining variance in the statistical model would incur equal cost when they move equal distances. The adaptive weights with respect to the variances are shown by the graph in Fig. 5.2(b).

It is trivial to compute the variability of each vertex from the variability constrained statistical model $\mathcal{N}(\tilde{\mathbf{x}}, \boldsymbol{\Sigma}_{\mathbf{x}_a|\mathbf{x}_b})$ which has already been calculated as explained in Sec. 4.2.1 and Sec. 4.2.2. In total 10,000 shapes are sampled from the variability constrained statistical model using $\boldsymbol{\alpha} \sim \mathcal{N}(0, I_n)$ within ± 3 standard deviation. This gives a total of 10,000 samples for each vertex that are used to compute a statistical model for each vertex having a normal probability distribution $\mathcal{N}(\bar{\mathbf{v}}_i, \boldsymbol{\Sigma}_{v_i})$, where $\bar{\mathbf{v}}_i$ is the mean and $\boldsymbol{\Sigma}_{v_i}$ is the variance of the i^{th} -vertex. The original variance of the

5. STATISTICAL MODEL BASED SEGMENTATION WITH NON-LINEAR AND SPARSE OPTIMIZATION

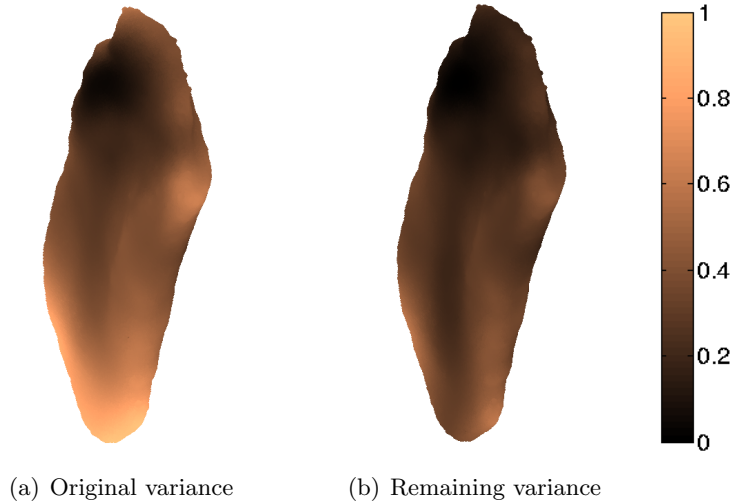


Figure 5.3: Normalized variance of the statistical model.

vertices of the statistical model built from the samples database (see Fig. 5.3(a)) and the remaining variability of each vertex after conditioning it on the partial information (landmarks) (see Fig. 5.3(b)) is color coded in Fig. 5.3.

The statistical model is optimized by minimizing the sum of the cost of vertices over the non-linear cost function $\bar{C}(v)$. The coefficients α corresponding to the main modes of variation are obtained by solving the minimization problem

$$\min_{\alpha} \left\{ \bar{C}(\mathbf{x}^*) + \xi |\alpha|_{L^1} \right\}, \quad (5.4)$$

where ξ is a weighting parameter. Since the adaptive cost is used, the statistical model has more flexibility, therefore, it is required that the model be regularized to constrain the solution space and generate smoother shape priors. Note that we use L^1 regularization [137] to constrain the solution space and generate sparse and more accurate solutions. Once the optimal α are found, the optimized shape is then constructed using Eq. 4.11 and used as a shape prior for the next iteration.

5.4 Algorithm

Figure 5.4 outlines the algorithm. The algorithm starts with the initial shape prior $\tilde{\mathbf{x}}$ obtained from the shape reconstruction from partial information (see Sec. 4.2.1) is used for the first iteration. The shape prior is used to generate the probability maps for the

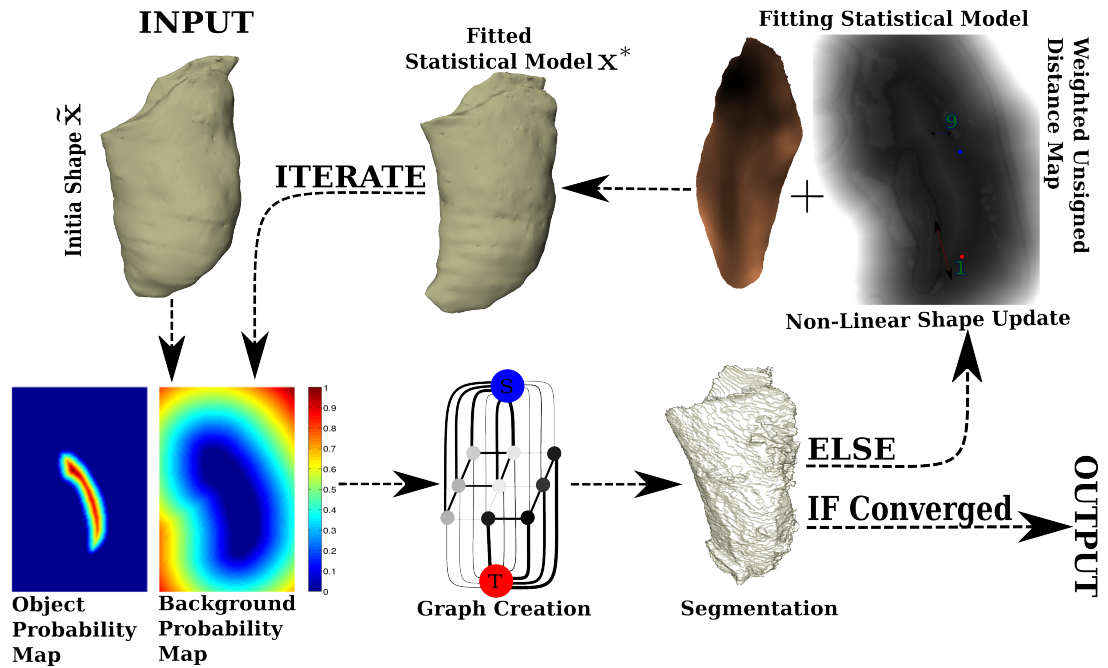


Figure 5.4: Segmentation process.

object and the background which is similar to the one used in Chap. 4 and encodes the shape knowledge. These maps are then used to create a graph corresponding to the energy function $E(\mathbf{z}|\mathbf{I}, \mathbf{x}^*)$ given by Eq. 5.1 and then the GC algorithm of Boykov & Kolmogorov [125] is used to obtain the muscle segmentation. If the segmentation has not converged then a non-linear cost function $\bar{C}(\mathbf{x}^*)$ outlined in Sec. 5.3.1 is created over which the shape prior is updated. Once the statistical model has been fitted to the current segmentation (see Sec. 5.3.2), the fitted statistical model provides a better and more accurate prior shape for the next iteration. This process is repeated until segmentation converges. Segmentation convergence is evaluated by comparing the current segmentation with the previous segmentation and if no voxels have changed label then the convergence has been reached or the algorithm run for a maximum of 15 iterations. The update of the shape prior is required as the initial estimate of the shape is not perfect, therefore, previous segmentation is used to update shape knowledge and get a better fitting of the statistical model to the specific patients muscle anatomy.

5. STATISTICAL MODEL BASED SEGMENTATION WITH NON-LINEAR AND SPARSE OPTIMIZATION

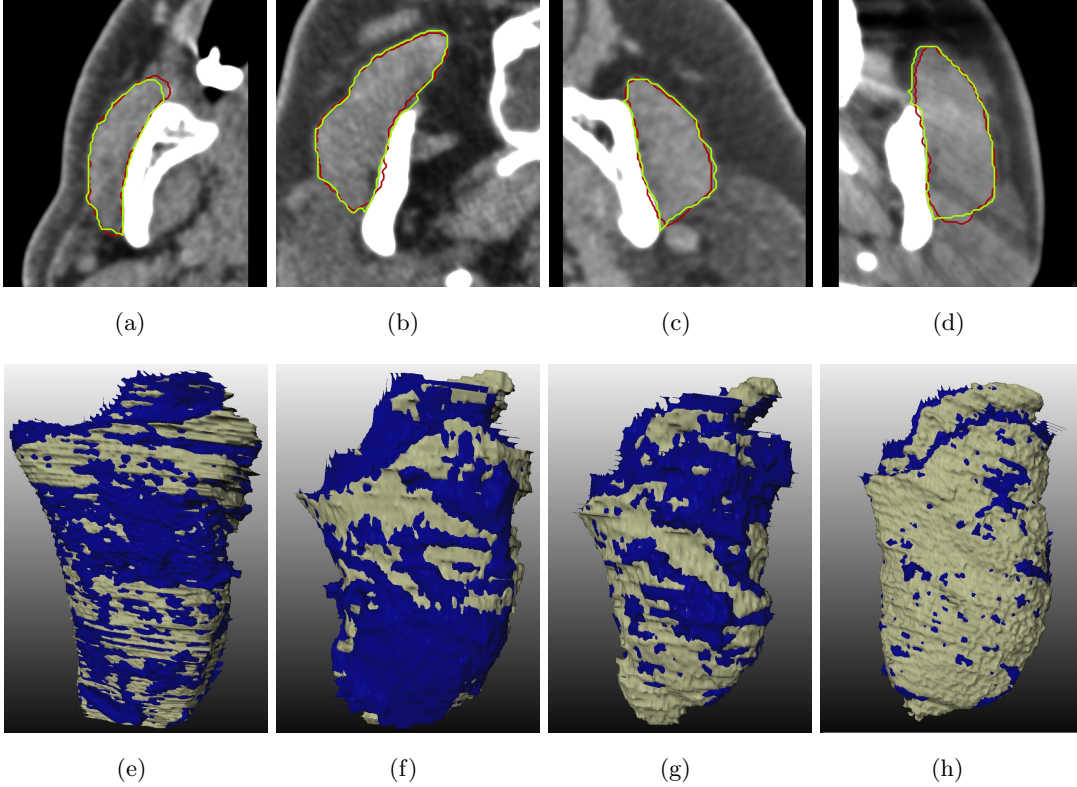
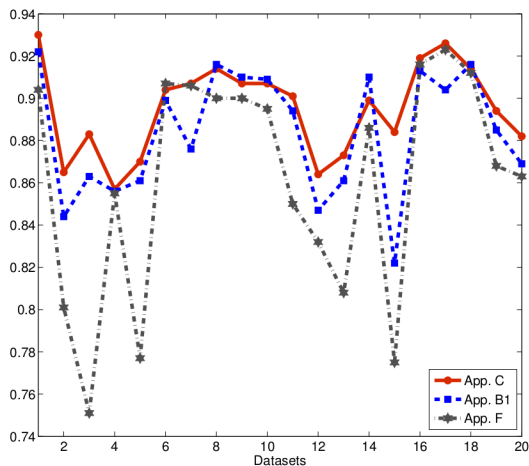


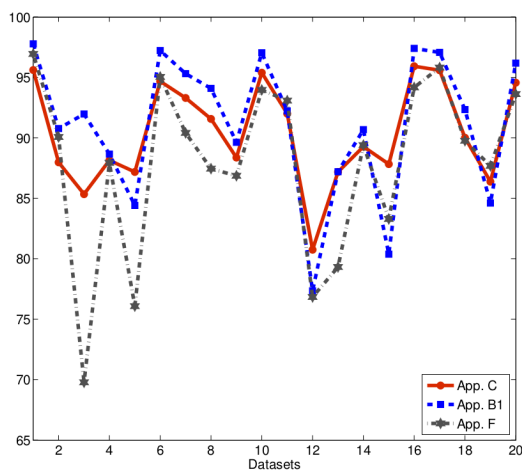
Figure 5.5: (a-d) Qualitative segmentation result in 2D where red is the ground truth and green is the segmentation boundary. (e,h) Qualitative segmentation result in 3D where ground truth is in gray and segmentation is in blue.

5.5 Experimental Setting and Results

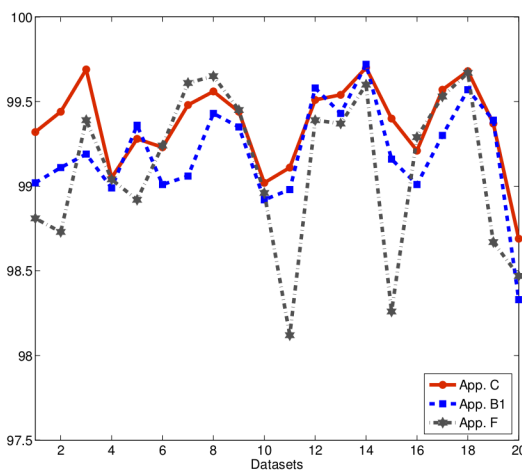
The proposed segmentation method (App. C) was tested on 20 CT datasets with the ground truth provided by a medical expert, using a Leave-One-Out cross-validation. The dataset dimensions were $79-156 \times 148-214 \times 125-384$ voxels and spacing $0.3-0.5 \times 0.3-0.5 \times 0.3-1$ mm³. All datasets possessed moderate high-density artifacts while 10 of them had severe high-density artifacts. The parameter $\sigma = 10$, mentioned in Sec. 4.2.1, controls the remaining variance of the statistical model, λ was chosen as 0.016 and $\mu = 0.0037$. The parameters were optimized on three different datasets and used throughout the entire segmentation experiments. The parameters $\beta = 0.01$, $\eta = 0.07$, $\theta = 4$ were used to generate the non-linear cost function while $\xi = 600,000$ was used for sparse shape optimization. The dice coefficient, sensitivity and specificity of the segmentation



(a) Dice Coefficient



(b) Sensitivity



(c) Specificity

Figure 5.6: Quantitative segmentation results.

were calculated as similarity measures to determine the accuracy of App. C.

Shape convergence was achieved within 5 – 11 iterations. The algorithm is computationally quite fast; it takes on average 4.1 ± 1.5 minutes. Computational time of 4 minutes is not real-time but, on the other hand, it takes over an hour and a medical expert to segment the muscle manually. It should be noted that although the employed mesh is very dense, the algorithm itself is independent of the density of the mesh. The cost function is evaluated where the vertices end up and that gives the total cost of the

5. STATISTICAL MODEL BASED SEGMENTATION WITH NON-LINEAR AND SPARSE OPTIMIZATION

Data	Similarity Measure								
	Dice Coefficient			Sensitivity%			Specificity%		
	App. C	App. B1	App. F	App. C	App. B1	App. F	App. C	App. B1	App. F
1	0.930	0.922	0.904	95.61	97.77	96.94	99.32	99.02	98.81
2	0.865	0.844	0.801	87.98	90.77	90.07	99.44	99.11	98.73
3	0.883	0.863	0.751	85.34	91.97	69.78	99.69	99.19	99.39
4	0.857	0.856	0.855	88.14	88.66	87.93	99.05	98.99	99.04
5	0.870	0.861	0.777	87.18	84.41	76.08	99.28	99.36	98.92
6	0.904	0.899	0.907	94.68	97.22	95.04	99.23	99.01	99.24
7	0.907	0.876	0.906	93.30	95.31	90.41	99.48	99.06	99.61
8	0.914	0.916	0.900	91.57	94.09	87.45	99.56	99.43	99.65
9	0.907	0.910	0.900	88.37	89.63	86.86	99.44	99.35	99.45
10	0.907	0.909	0.895	95.39	97.04	93.96	99.02	98.92	98.96
11	0.901	0.894	0.850	92.00	92.23	93.08	99.11	98.98	98.12
12	0.864	0.847	0.832	80.75	77.34	76.84	99.51	99.58	99.39
13	0.873	0.861	0.808	87.18	87.20	79.32	99.54	99.43	99.37
14	0.899	0.910	0.886	89.27	90.68	89.40	99.70	99.72	99.60
15	0.884	0.822	0.775	87.83	80.38	83.29	99.40	99.16	98.26
16	0.919	0.913	0.916	95.92	97.40	94.17	99.21	99.01	99.29
17	0.926	0.904	0.923	95.59	97.07	95.81	99.57	99.30	99.53
18	0.914	0.916	0.912	90.01	92.34	89.76	99.68	99.57	99.67
19	0.894	0.885	0.868	86.38	84.61	87.68	99.37	99.39	98.67
20	0.882	0.869	0.863	94.56	96.18	93.62	98.69	98.33	98.47

Table 5.1: Quantitative comparison of App. C with App. B1 and App. F.

	DC (Mean \pm Std)	DC (Median)	DC (Smallest - Largest)
App. C	0.895 \pm 0.022	0.900	(0.857 - 0.930)
App. B1	0.884 \pm 0.029	0.890	(0.822 - 0.922)
App. F	0.861 \pm 0.054	0.877	(0.751 - 0.923)

Table 5.2: The table list the mean & the standard deviation, the median as well as the smallest and the largest dice coefficient of App. C, App. B1 and App. F

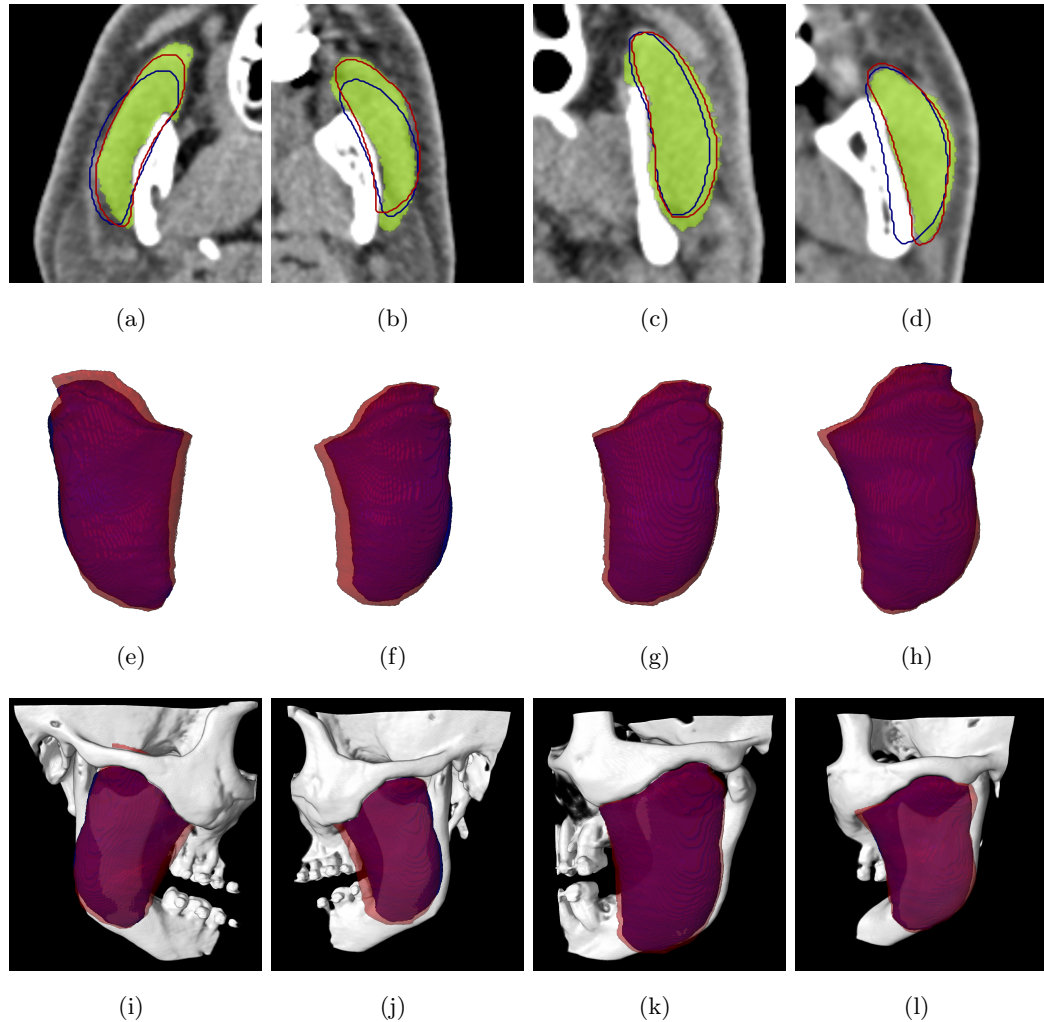


Figure 5.7: The shape update in different datasets. The top row (a...d) shows the initial shape contour in blue while the final shape contour in red. The ground truth is shown in green in the top row. The middle row (e...h) depicts the initial 3D mesh of the masseter muscle in blue while the final transparent masseter muscle mesh is shown in red. The bottom row (i...l) shows the same mesh overlaid in the dataset.

shape. The algorithm will work equally well with a less dense mesh.

Figure 5.5 shows the qualitative results in 2D (a-d) and in 3D (e-h). The experimental results obtained using App. C are clinically acceptable as validated by medical experts from the department of CMF surgery of University Hospital Basel.

The graphs in Fig. 5.6 show the results of App. B1 using a linear cost function (blue

5. STATISTICAL MODEL BASED SEGMENTATION WITH NON-LINEAR AND SPARSE OPTIMIZATION

curve) and App. C with non-linear cost function (red curve). The gray curve represents the results of using App. F. The dice coefficients, sensitivity and specificity values of the different datasets are listed in Table 5.1

A statistical comparison was conducted between the approaches App. C, App. B1 and App. F. Statistical analysis (t-test with significance level 0.05) was performed over the dice coefficient values to determine whether App. C is statistically significantly different than the other two methods. We found that App. C is statistically significantly better than both App. F with a p-value = 0.001 and App. B1 with a p-value = 0.005. The improvement over the results App. B1 is mainly due to the use of the non-linear cost function using L^1 regularization. The dice coefficient (see Fig. 5.6(a)) and specificity (see Fig. 5.6(c)) for App. C is better for all datasets except for 6. Table 5.2 lists the mean, median, standard deviation and the smallest and the largest dice coefficient values for the methods.

The segmentation presented in the previous approach App. B1 were already close to the human/expert segmentation, therefore, drastic improvements are difficult to achieve. Nevertheless, it has been demonstrated that the proposed technique is still statistically significantly better. It has also been shown that this approach (App. C) allows to further improve the segmentation accuracy and that a statistical model can not only be used to restrict the shape variability during segmentation but also how to make use of the remaining shape variability in the statistical models to even further improve the segmentation. Figure 5.7 shows the initial and the final shape of the statistical mesh model both in 2D (a...d) and in 3D (e...l).

5.6 Conclusion

An improved segmentation approach (App. C) has been proposed that combines a constrained statistical model with an MRF-based segmentation approach. As compared to the state-of-the-art methods, a non-linear cost function was used to fit the statistical. This new cost function has shown to be superior as it generates more consistent shape updates. The method's performance has been evaluated on 20 masseter CT dataset and quantitatively compared to state-of-the-art segmentation approaches. Although the method has been shown and evaluated on the masseter muscle it is of general use and can be applied whenever a statistical model is available.

6

Facial Soft Tissue Segmentation

6.1 Introduction

An approach to segment the facial muscles and other soft-tissues (facial fat and facial skin) is presented in this chapter based on the algorithm presented in Chap. 3. Furthermore, a facial muscle template is used to separate the facial muscles. The aim of this chapter is to develop a semi-automatic facial soft-tissue extraction method from CT datasets. The motivation behind separating the facial soft-tissues has been presented in Chap. 1. A step-by-step approach has been presented to extract the facial soft-tissues that has been in part motivated by the works of Tasdizen *et al.* [32]. They segmented different tissues in the human head from multispectral MRI, where the separability of different soft-and hard-tissues is easier. One soft-tissue layer was peeled off at a time such as to obtain a brain and non-brain separation, followed by bone removal from the non-brain region and finally the extraction of blood vessels, sinuses, eyes, bone marrow, muscle and fat. Skin was not processed. However, in this work only either a CT or MRI dataset is used to perform the separation of the soft-tissues in a step-by-step approach.

6.2 Segmentation of Facial Soft-Tissues

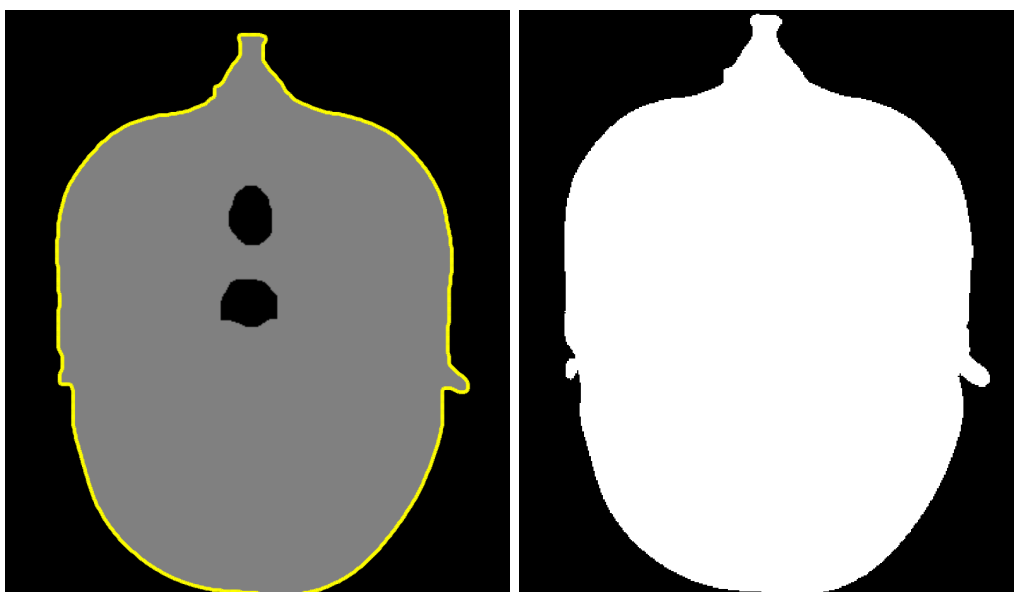
The proposed method starts by the separation of the background and the head region. The skin is then peeled off from the head region which is then followed by the separation of bone. The next step is then to segment the brain from this head region, followed by the removal of the sinuses and air regions from the dataset. All extracted regions form

6. FACIAL SOFT TISSUE SEGMENTATION



(a) Original dataset.

(b) Threshold, morphological closing and region growing operations.



(c) Head region contour selection.

(d) Segmented head region.

Figure 6.1: Head region and background separation.

new datasets that are used as hard constraints for the segmentation of fat. Finally, the relevant facial muscles are segmented, where, again, all previous segmentations are

used as hard constraints. The details of each step are now presented in the following sub-sections.

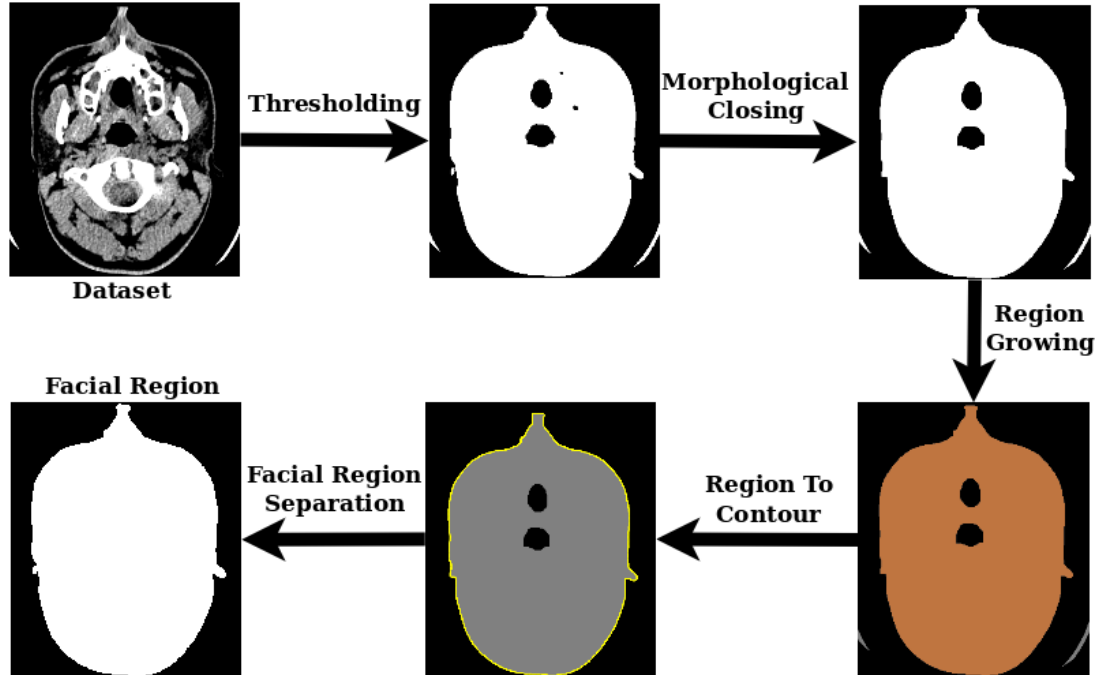
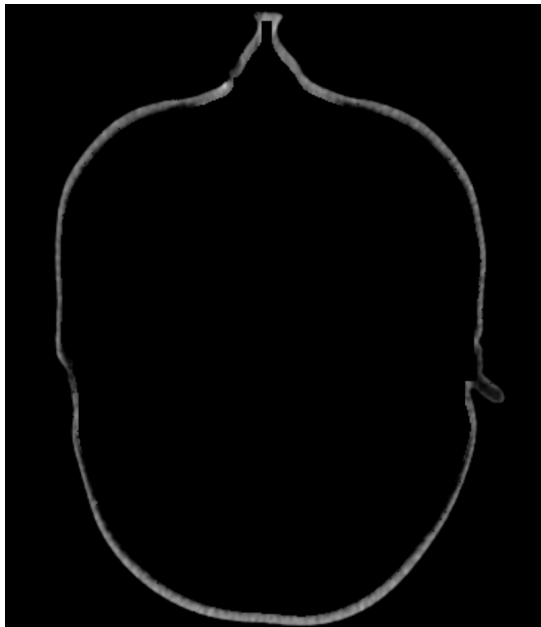


Figure 6.2: Workflow for head region and background separation.

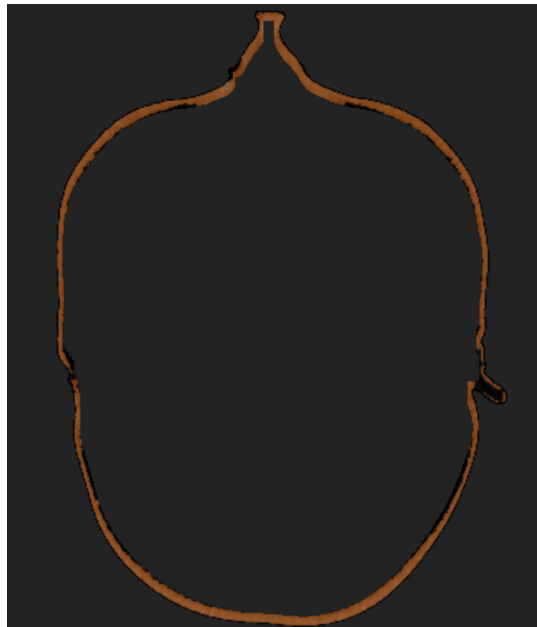
6.2.1 Step 1 - Head Separation

The first step in the extraction of soft-tissues is to separate the head region that includes skull, brain, soft-tissues and everything inside the skin (see Fig. 1.2) from the background, depicted in Fig. 6.1. To separate the head from the background first, a threshold is applied on the dataset followed by morphological closing to remove the small sinuses and the holes created due to high density artifacts (see Fig. 6.1(b)). Region growing is used to select only the head region and discard the table on which the patient is lying (see Fig. 6.1(b)). The head region's contour (in yellow in Fig. 6.1(c)) in a sort of a floodfill method is used to select the entire head region and to separate it from the background. The final segmented head region is shown in Fig. 6.1(d). Figure 6.2 shows the workflow of separating the background from the head region.

6. FACIAL SOFT TISSUE SEGMENTATION



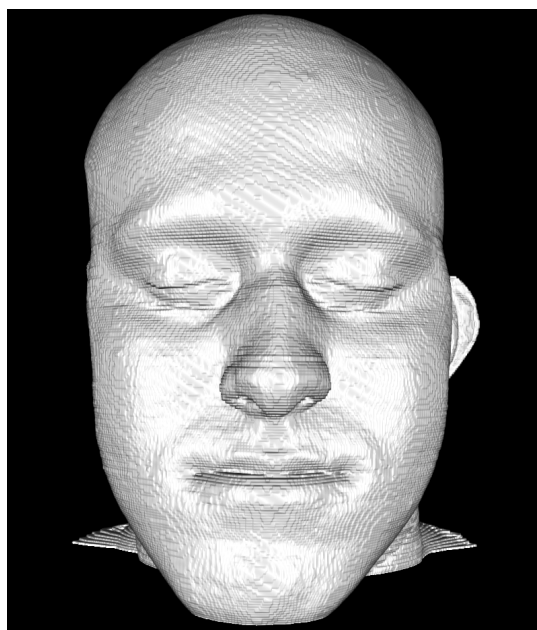
(a) Band around skin.



(b) Region growing.



(c) Skin layer segmentation in green in 2D view.



(d) Skin layer segmentation in 3D view.

Figure 6.3: Skin layer segmentation shown in green.

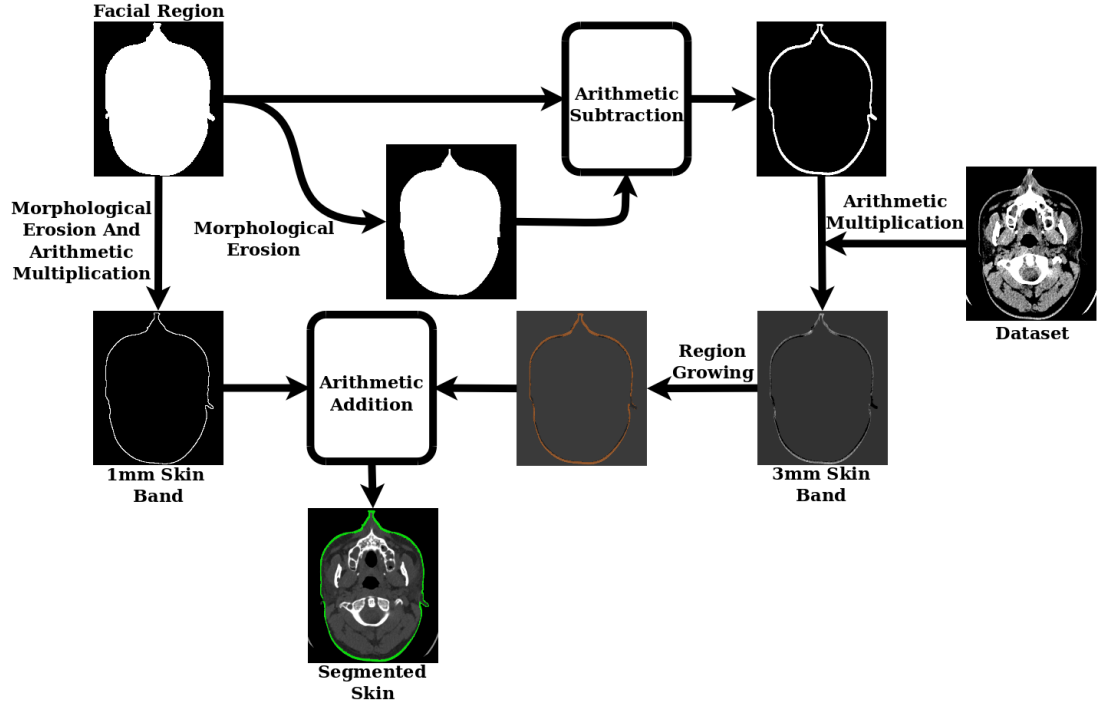


Figure 6.4: Skin layer segmentation workflow.

6.2.2 Step 2 - Skin Layer Segmentation

Skin is segmented using morphological, arithmetic and region growing operators. The head region separated in Step 1 is used to create a band of approximately 3 – 4 mm around the skin using morphological and arithmetic operators. This is done in order to keep the original image intensities inside the band while transforming all intensities outside the band to the imaging dataset minimum intensity as shown in Fig. 6.3(a). Then region growing is applied with a threshold of $[(-30) \cdots (+135) \text{ HU}]$ inside the interval to segment the skin, see Fig. 6.3(b). The region growing result is not perfect due to partial volume effect and high density artifacts which create holes in the skin segmentation. To bridge the holes in the skin, a thin layer of approximately 1 mm thickness is added to the outcome of region growing to obtain a hole-free skin segmentation as shown in 2D and 3D in Fig. 6.3(c) as well as Fig. 6.3(d) respectively. The complete workflow for the segmentation of skin is depicted in Fig. 6.4.

6. FACIAL SOFT TISSUE SEGMENTATION

6.2.3 Step 3 - Bone Segmentation

Bone is segmented using simple thresholding $[(+170) \cdots (+4000) \text{ HU}]$ as shown in the 2D and 3D views of Fig. 6.5. The same process was used by Kale *et al.* [30] and Sener [14]. The selected range of the intensity values allows to separate the spongy bones and thin bones, even their boundaries affected by the partial volume effect. The HU values of different soft- and hard-tissues in the CT dataset are listed in Table 2.1.

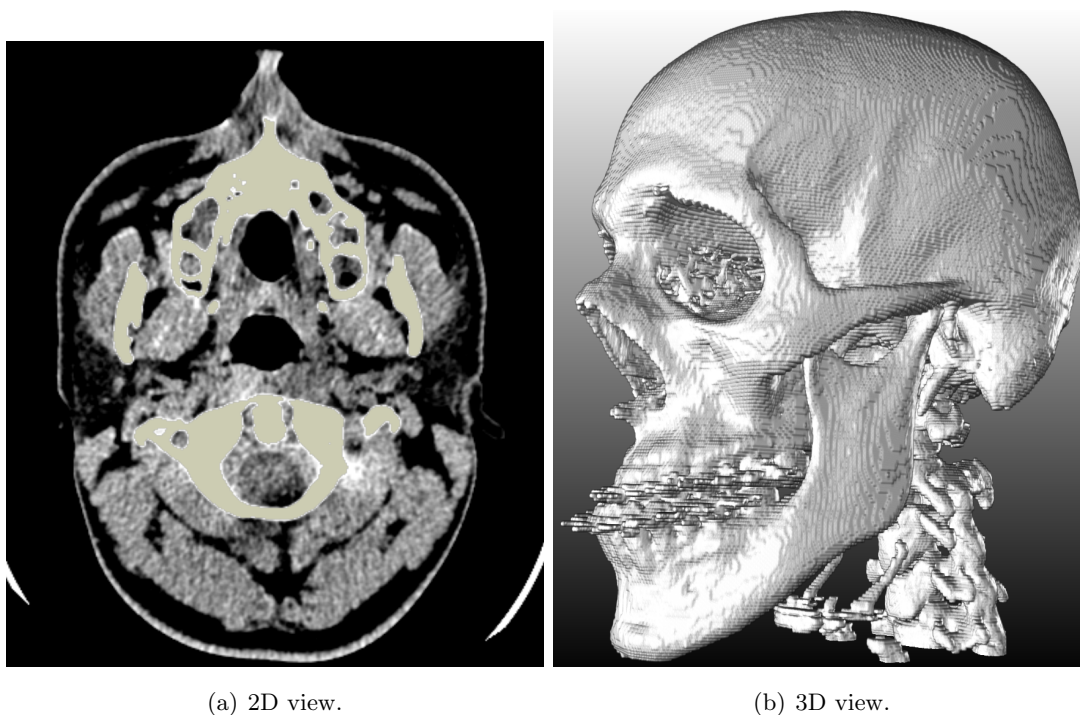


Figure 6.5: Bone segmentation in cream color.

6.2.4 Step 4 - Brain Segmentation

In order to segment the facial soft-tissues, the brain has to be extracted first as it possesses an overlapping intensity profile with one of the facial muscles, and therefore, it is imperative that the brain be separated from the head region (see Fig. 6.6). This procedure was inspired by the work of Shattuck *et al.* [31], who applied it for the segmentation of the brain and its subparts. As a first step, they separated the brain from the rest of the head region and then proceeded to segment the inner structures

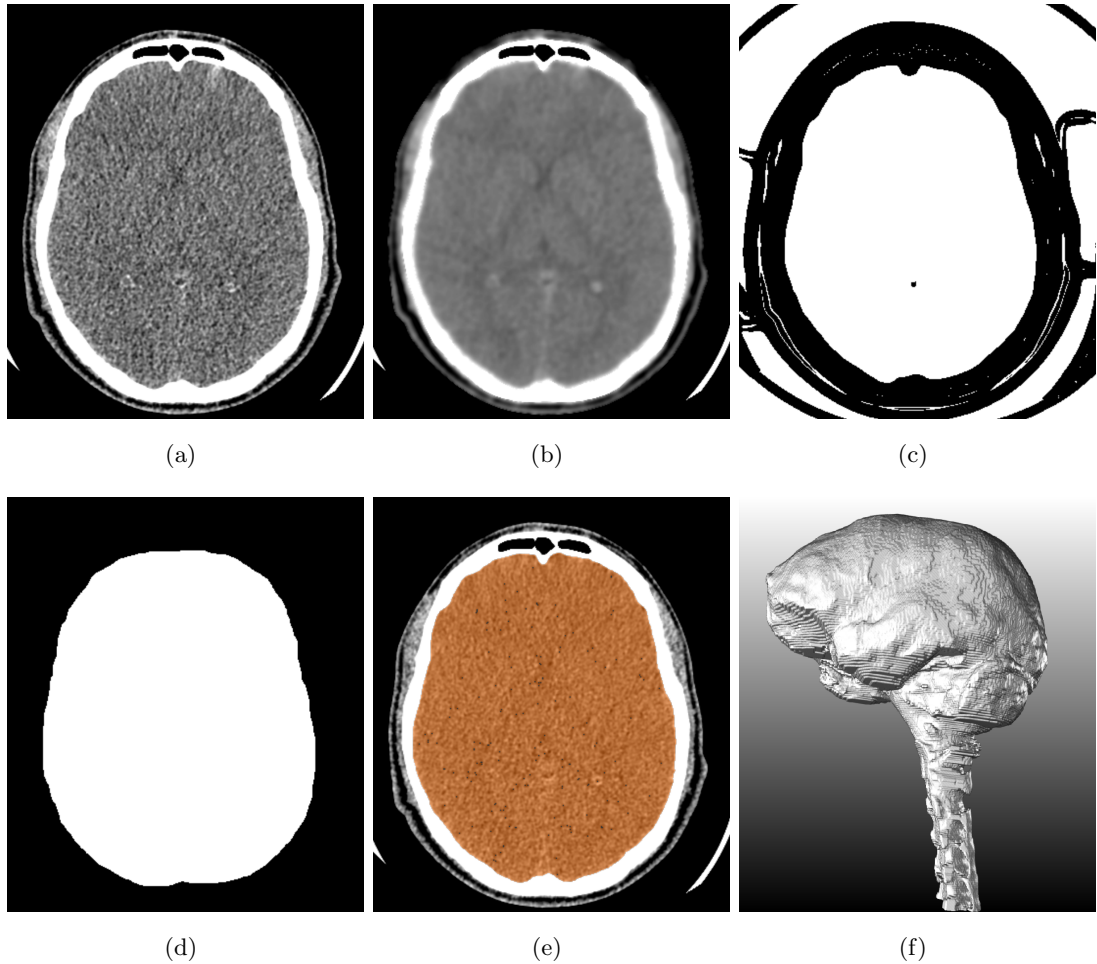


Figure 6.6: Brain segmentation. (a) Original dataset, (b) anisotropic diffusion filtering, (c) edge detection and gradient thresholding, (d) morphological and arithmetic operation, (e) brain segmentation in 2D and (f) brain segmentation in 3D.

of the brain. Here, the interest is in segmenting the structures in the head region. Our initial steps to segment the brain are the same as those described by Shattuck *et al.* [31]. They used a combination of anisotropic diffusion filtering shown in Fig. 6.6(b), edge detection, see Fig. 6.6(c), morphological and arithmetic operators depicted in Fig. 6.6(d). Finally, brain segmentation was performed using region growing in 2D, Fig. 6.6(e) and in 3D, Fig. 6.6(f). The region growing, morphological and arithmetic operations were added on top of those of Shattuck *et al.* [31] and are used to stop the region growing leaking into other parts of the face. This process of brain segmentation

6. FACIAL SOFT TISSUE SEGMENTATION

is also sometimes referred to as *Skull Stripping*. The workflow is shown in Fig. 6.7.

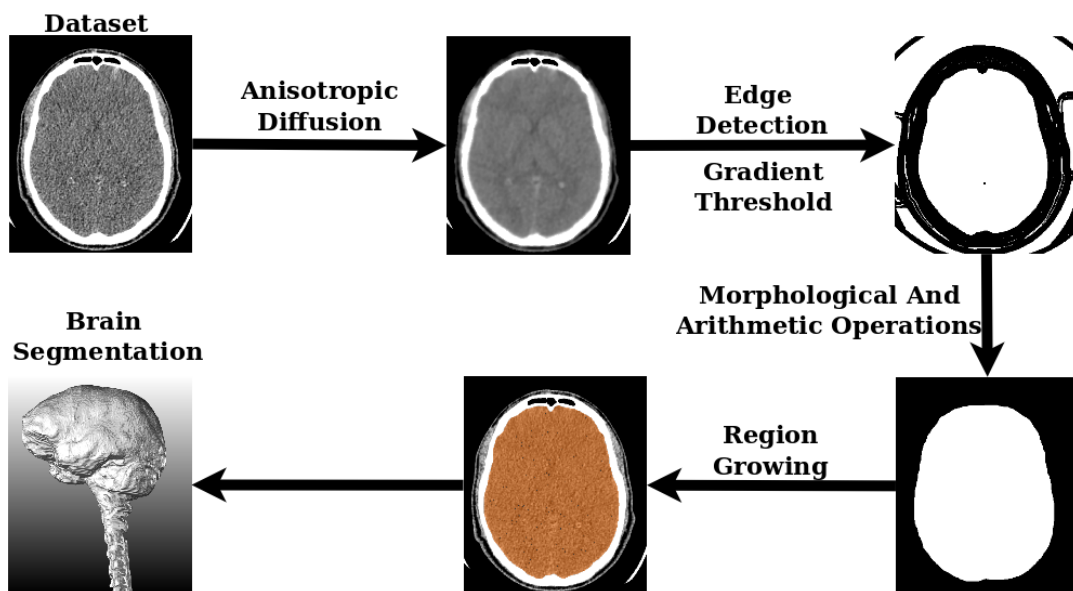


Figure 6.7: Brain segmentation workflow.

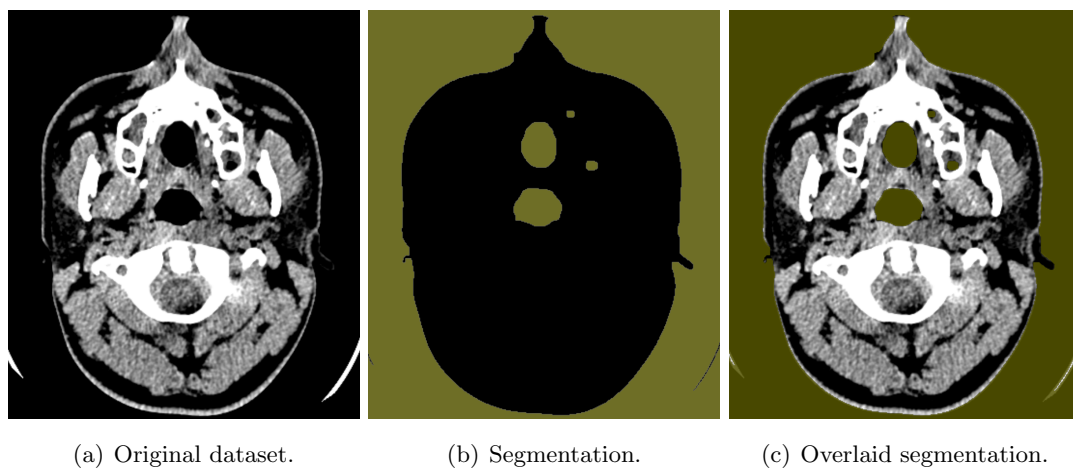


Figure 6.8: Air segmentation.

6.2.5 Step 5 - Air Segmentation

The air is segmented using simple thresholding $[(-1500) \dots (-300) \text{ HU}]$, see Fig. 6.8. Here, basically sinuses and empty areas within the head are segmented from the dataset,

Fig. 6.8(b), while the rest of the non-head region is redundant as it has already been segmented in Sec. 6.2.1. The same process has been used by Kale *et al.* [30] and Sener [14].

6.2.6 Step 6 - Fat Segmentation

Once the head separation, skin, bone, brain and air segmentation have been done, they are used to create a background datasets for fat segmentation by simple addition as shown in Fig. 6.9(g) and Fig. 6.9(h). The background dataset is then used as a hard constraint in the GC segmentation. The standard GC segmentation approach of Boykov & Jolly [78] was used for the segmentation of facial fat. The energy function is the same as the one used by Boykov & Jolly [78] and presented in Sec. 2.2.1 is given by

$$E(\mathbf{z}|\mathbf{I}) = \sum_{p \in \mathbf{P}} V_p(z_p|\mathbf{I}) + \lambda \sum_{p \in \mathbf{P}} \sum_{q \in N_p} V_{p,q}(z_p, z_q|\mathbf{I}). \quad (6.1)$$

where λ is the smoothness parameter. The facial fat was segmented using the GC algorithm of Delong & Boykov [124]. The code is available free of charge for academic and research purposes and can be downloaded from the website of Sameh Khamis¹. The colored region in Fig. 6.9(h) is used as a background hard constraint, instead of obtaining user defined hard constraints. The fat segmentation results are shown in Fig. 6.10.

6.2.7 Step 7 - Facial Muscle Segmentation

Once the fat regions of a specific skull have been segmented (see Fig. 6.11(a)), they are also added to the fat background dataset (see Fig. 6.11(b)) and used as a hard background constraint during the segmentation of the facial muscles. It is evident that the facial muscles that are essential for pre-operative surgery simulation lie just beneath the facial skin, and therefore, a band around the facial skin is created using morphological erosion and subtraction from the head region, see Fig. 6.11(c). This band is also added to the facial muscle background dataset. A further constraining of the facial muscle region is obtained using the voxelization of the facial muscle template. Since the template is warped to fit the facial muscles, the voxelization of the facial muscle template is morphologically dilated to create another band (see Fig. 6.11(d)).

¹Website: <http://www.samehkhamis.com/>

6. FACIAL SOFT TISSUE SEGMENTATION

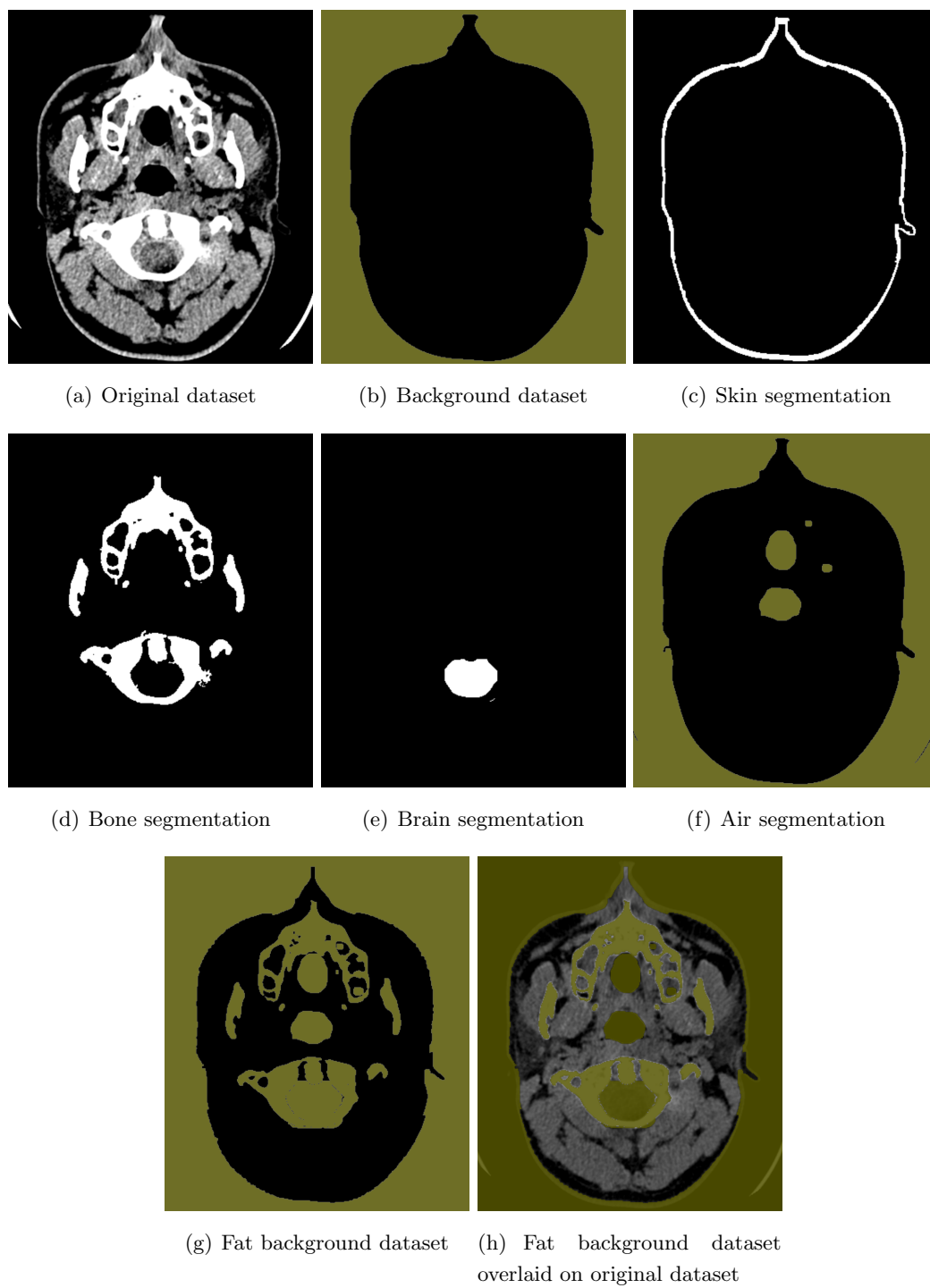


Figure 6.9: Fat background dataset.

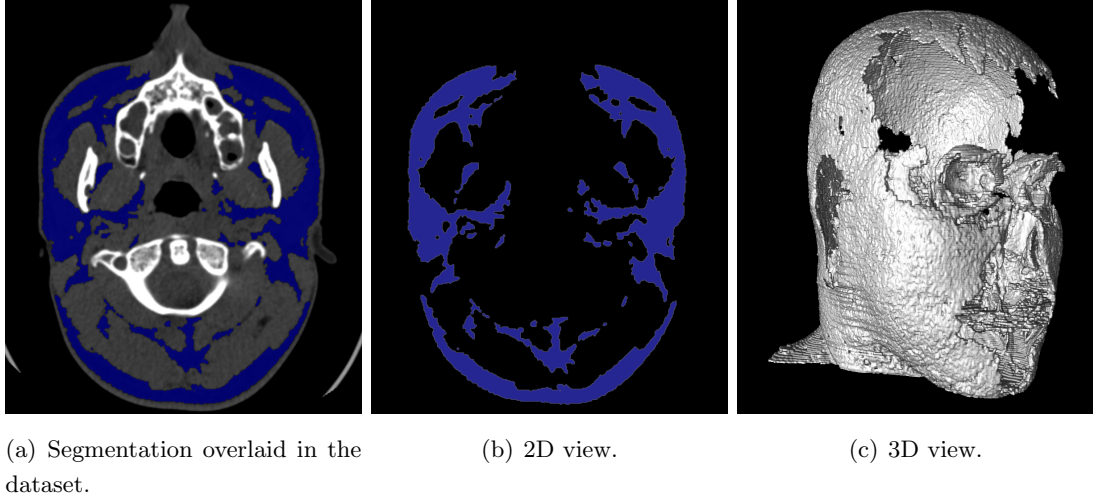


Figure 6.10: Fat segmentation in blue.

The warped template only roughly fits the facial muscles and is not accurately aligned with them, therefore, the morphological dilation of the template is used to create a band which provides information about all the target facial muscles that lay within the band and thus allows to eliminate the other facial soft-tissues and facial muscles that are not the target of the segmentation.

The morphological dilation is kept constant throughout the entire segmentations process and is large enough to encompass all the target facial muscles. This band is also added to the facial fat background dataset. The final facial muscle background dataset is shown in Fig. 6.11(e). The entire shaded region (greenish in Fig. 6.11(e) and Fig. 6.11(f)) show the hard constraints for the background region. For the purpose of pre-operative surgery simulation, it is sufficient to segment all the muscles as one object. The facial muscles were segmented using the GC algorithm of Delong and Boykov [124].

Segmentation Model

The additional shape prior information is incorporated in the form of the muscle template. The facial muscle template mesh (see Fig. 6.12(a)) is aligned with the skull mesh model (see Fig. 6.12(b)) using landmark locations. These landmark locations are the muscle attachments at the facial bones and skull as mentioned in Chap. 4. Once the

6. FACIAL SOFT TISSUE SEGMENTATION

template is aligned to the skull mesh, it is deformed simultaneously together with the skull mesh using the method of Dedner *et al.* [133]. The same deformation fields as those used to adapt the skull have also been used for the vertices of the facial template mesh. This results in the warped facial muscle template mesh $\hat{\mathbf{x}}$ (see Fig. 6.12(c)) which is then used as additional shape information. The energy function is similar to the one used in Eq. 4.13 and is rewritten here for convenience, replacing the optimal masseter muscle shape \mathbf{x}^* with $\hat{\mathbf{x}}$ which represents the facial muscle mesh template as shown in

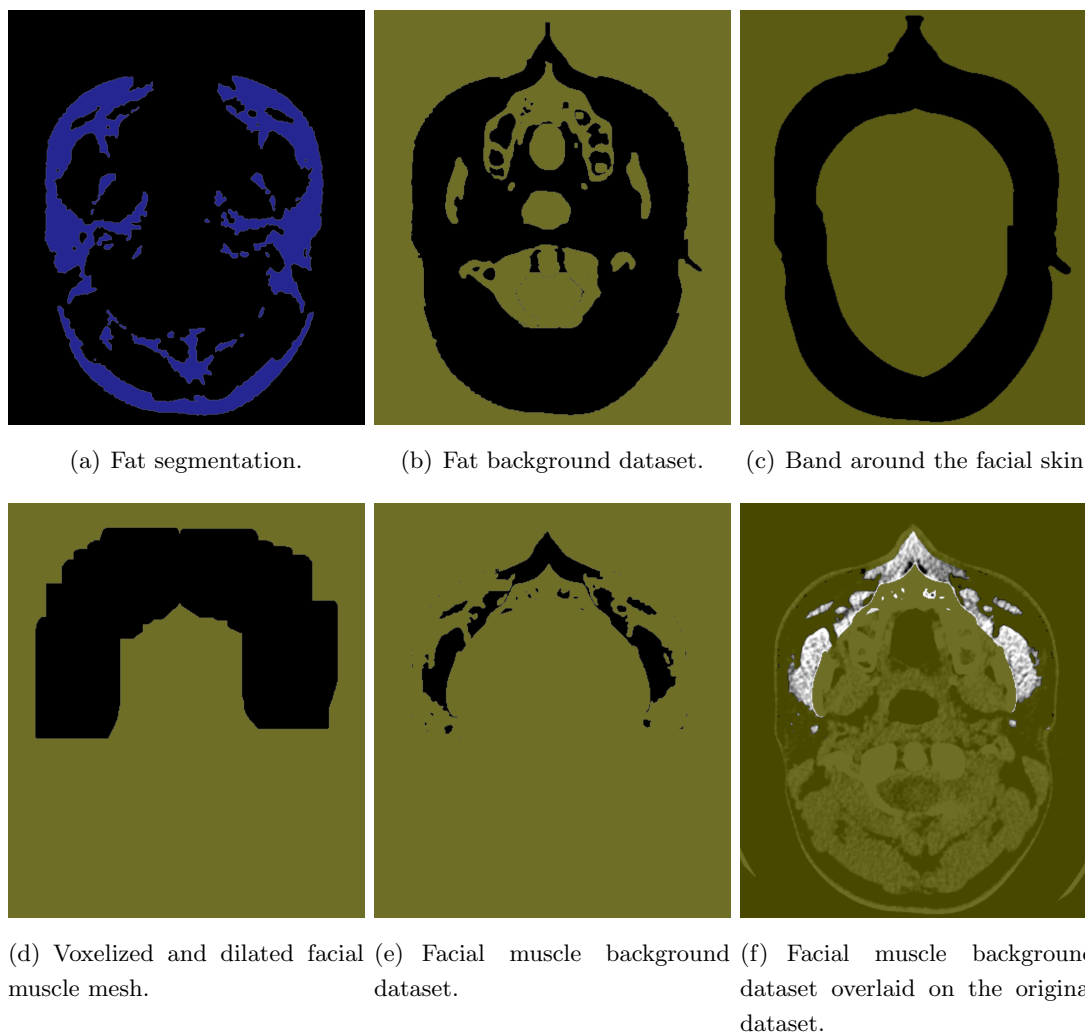


Figure 6.11: Facial muscle background dataset.

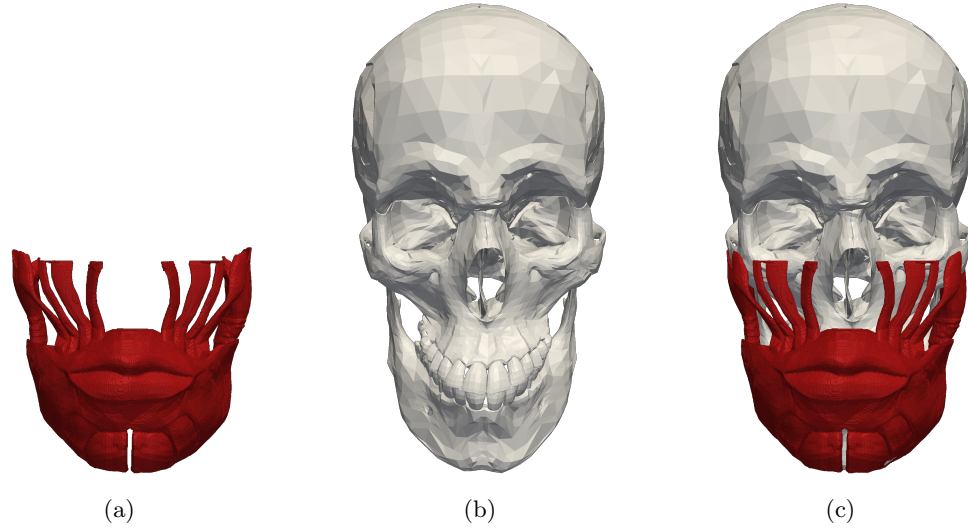


Figure 6.12: (a) Facial muscle template mesh, (b) skull mesh, (c) warped template mesh.

Fig. 6.12(a)

$$E(\mathbf{z}|\mathbf{I}, \hat{\mathbf{x}}) = \sum_{p \in \mathbf{P}} \left\{ V_p(z_p|\mathbf{I}) + \mu V_p(z_p|\hat{\mathbf{x}}) \right\} + \lambda \sum_{p \in \mathbf{P}} \sum_{q \in \mathbf{N}_p} V_{p,q}(z_p, z_q|\mathbf{I}), \quad (6.2)$$

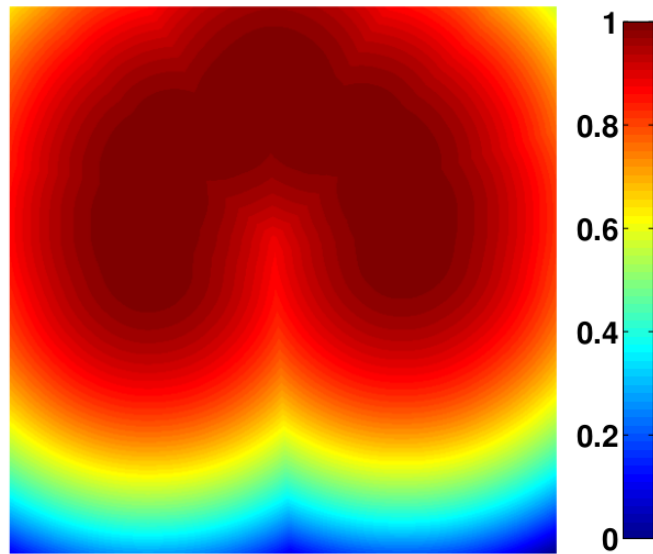


Figure 6.13: Probability map for voxels belonging to the object.

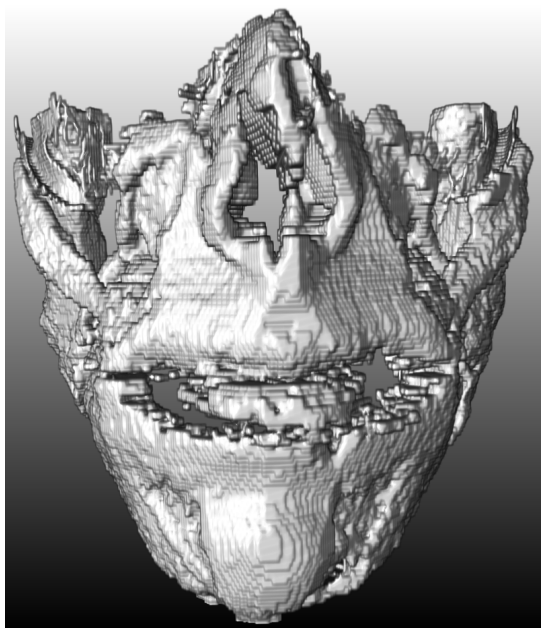
6. FACIAL SOFT TISSUE SEGMENTATION



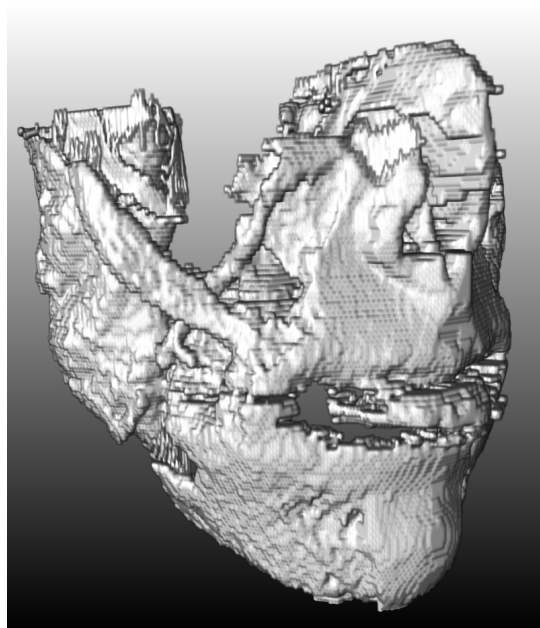
(a) 2D view.



(b) Segmentation overlaid on dataset.



(c) 3D front view.



(d) 3D side view.

Figure 6.14: Facial muscle segmentation in red.

where λ is the smoothness parameter and μ is the shape prior parameter. Furthermore, \mathbf{z} is a binary variable which defines the segmentation, \mathbf{I} is the observed intensity data, $V_p(z_p|\mathbf{I})$ and $V_{p,q}(z_p, z_q|\mathbf{I})$ are the regional and the boundary terms, respectively. The shape term encodes only how likely a particular voxel p belongs to the object “1” given the shape prior $\hat{\mathbf{x}}$ and is shown in Fig. 6.13. The voxels enclosed by the mesh possess higher probability to belong to the object. The probability of belonging to the object decreases as the distance of the voxel to the mesh increases.

Unlike the previous segmentation approaches, where the shape prior was well aligned with the target dataset, the facial muscle template mesh is not as accurately aligned. The muscle template mesh only provides knowledge that the muscle is somewhere in the vicinity of the mesh. As a result, the background probability dataset can not be reliably used. Shape knowledge is deployed based on the fact that the muscle is somewhere in the vicinity of the mesh and as one moves further away from the mesh, the less likely a voxel is to belong to the object. The segmented facial muscles and soft-tissues are shown in Fig. 6.14 and Fig. 6.15 respectively. Here, they are also shown in comparison to each other and with respect to the skull. The skull was generated using simple thresholding and has been smoothed for visualization purposes. Finally, all the soft-tissues were visualized together (see Fig. 6.16(a)) and jointly with skull (see Fig. 6.16(b)) in 3D view and in 2D view overlaid in an axial slice in Fig. 6.16(c).

Segmentation Issues

Although the soft-tissues have been segmented, there is still more to achieve to improve the segmentation accuracy. It is imperative that the individual facial muscles be identified. One way to identify them is to use a labeled template and segment the facial muscles one after the other. If the alignment of the facial muscle template mesh can be improved, it will also aid in improving the segmentation accuracy of the facial muscles.

6.3 Experimental Setting and Results

The proposed approach was tested on six datasets, five of which were CT datasets with dimensions $512 \times 512 \times 191-299$, spacing $0.32-0.49 \times 0.32-0.49 \times 0.5-1 \text{ mm}^3$ and one MRI dataset with dimensions $400 \times 400 \times 80$ voxels and spacing $0.45 \times 0.45 \times 1.5 \text{ mm}^3$. The CT datasets were noisy with almost all possessing mild to severe imaging artifacts.

6. FACIAL SOFT TISSUE SEGMENTATION

The CT datasets originate from the patients of the Orthognathic surgery department of Sao Carlos Hospital, Brazil. There was no ground truth segmentation available from the medical experts, and therefore, the validation strategy chosen was to evaluate how close the facial simulation prediction results using the facial soft-tissue segmentations, are to those of the CMF post operative datasets. This still has to be tested on the simulation results to prove that with the obtained segmentation results, simulation results are getting better. The parameters were manually optimized on one datasets and then the same parameters were used on the rest of the datasets. The parameter values used were $\lambda = 0.4$ and $\mu = 0.6$.

6.4 Conclusion

This chapter presented a method to semi-automatically segment the facial soft-tissues using some simple low-level image processing operations. Segmentation of facial soft-tissues is important as the existing facial surgery simulation and prediction methods regard the space between the bone and skin as homogeneous, while this space consists of a number of different soft-tissues like muscle, vessels, nerves, glandular structures and fat. The anatomy, orientation and organization of each of these structures differ from patient to patient having a different coefficient of elasticity. The simplified model lacks the accuracy needed in a clinical setting, rendering the simulation and prediction software to mere visualization or animation tasks. It is, therefore, necessary to devise segmentation algorithms to extract the important facial soft-tissues so that patient-specific simulation and prediction results can be achieved. This will also increase the accuracy of the simulation and prediction results. In this thesis, the two most important soft-tissues, namely muscles and fat, have been segmented as they fill up, apart from the skin, the majority of the space.

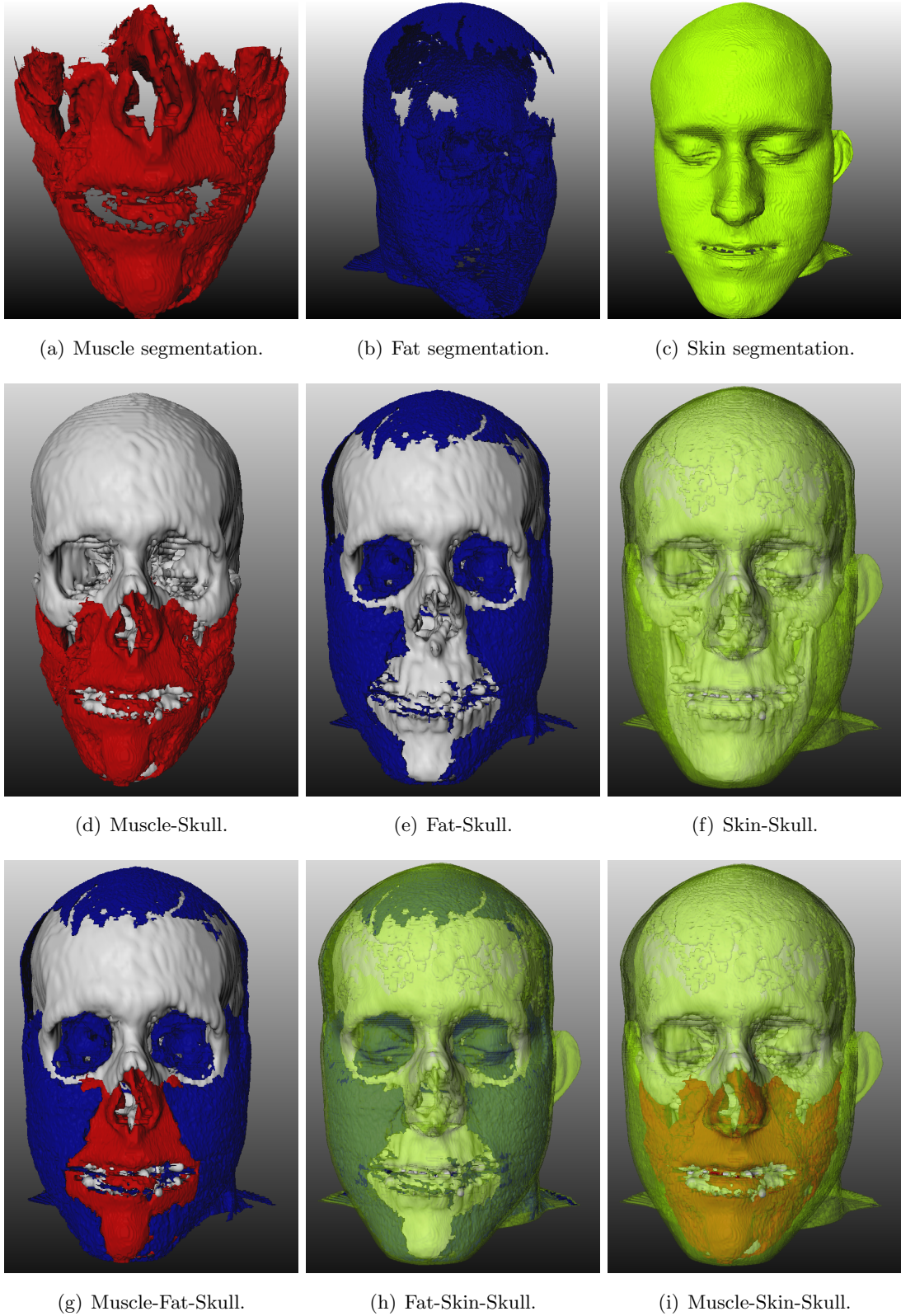
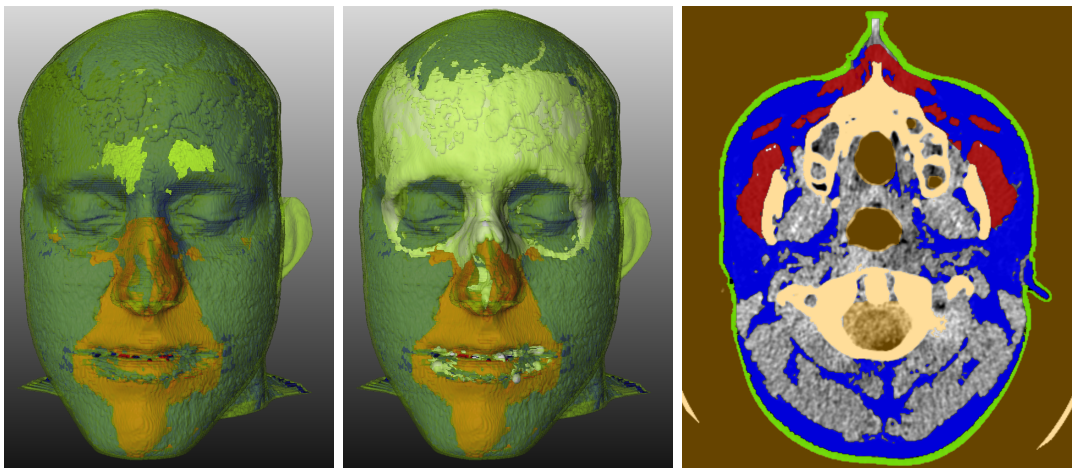


Figure 6.15: Facial soft-tissues segmentation in comparison to each other as well as the skull.

6. FACIAL SOFT TISSUE SEGMENTATION



(a) Soft tissues segmentation in 3D. (b) Soft tissues segmentation along with skull in 3D. (c) Soft tissues segmentation in 2D.

Figure 6.16: Facial soft-tissues segmentation. Facial muscle in red, fat in blue, skin in green, and brownish orange is the background dataset.

7

Conclusion

7.1 Discussion

This thesis proposes incremental approaches in the MRF framework for segmentation of facial soft-tissues. The approaches integrated additional shape knowledge which is necessary to make the segmentation robust with respect to weak edges, noise, imaging artifacts and multiple adjacent objects having similar intensity profiles as the target object. Different ways of integrating the shape prior modeled by a fixed shape template or a statistical model based on PPCA have been proposed. In the first approach App. A, the fixed shape template was incorporated into both the data and the smoothness terms, while in the other approaches (App. B1 and App. C) the statistical model was encoded only in the data term of the MRF energy function. Shape prior augmented MRF energy functions were then minimized using GC to obtain global solutions.

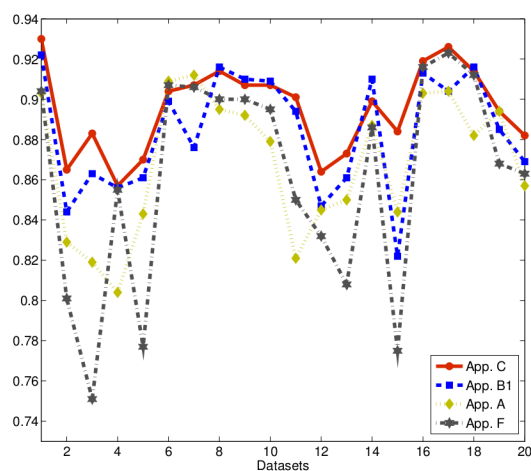
	DC (Mean \pm Std)	DC (Median)	DC (Smallest - Largest)
App. A	0.869 ± 0.034	0.881	(0.804 - 0.912)
App. B1	0.884 ± 0.029	0.890	(0.822 - 0.922)
App. C	0.895 ± 0.022	0.900	(0.857 - 0.930)
App. F	0.862 ± 0.053	0.877	(0.752 - 0.923)

Table 7.1: Comparison of the proposed methods App. A, App. B1 and App. C with App. F.

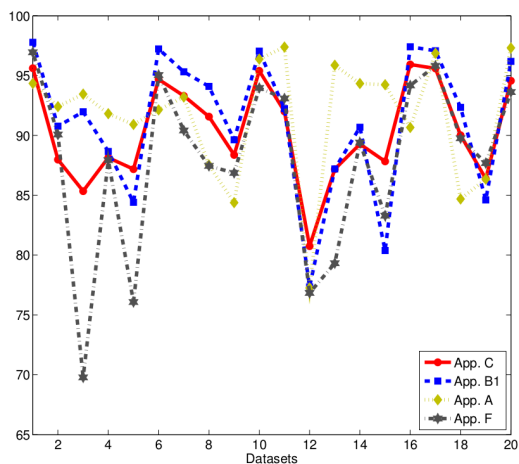
The proposed methods are general enough and can be used to segment any kind of anatomical or non-anatomical object from medical or non-medical image datasets for

7. CONCLUSION

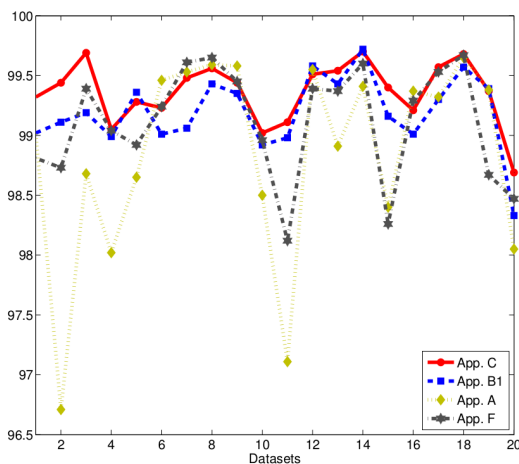
example they can be easily and readily adapted to CT and MRI datasets without any major changes. The proposed methods (App. A, App. B1 and App. C) have been tested on the challenging example of segmenting the masseter muscle. The results showed that by incorporating prior shape knowledge, leaking can be effectively constrained and clinically acceptable results can be achieved.



(a) Dice Coefficient



(b) Sensitivity



(c) Specificity

Figure 7.1: Combined quantitative segmentation results.

Table 7.1 shows the mean and standard deviation, the median, as well as the smallest and the largest value of the dice coefficient of the proposed methods App. A, App. B1,

and App. C in comparison to App. F. The approach App. A uses a fixed shape template (*i.e.* the mean shape) as prior shape knowledge. In general, good segmentation results could be obtained which were clinically acceptable, but subtle anatomical shape variations made it difficult to optimally align the fixed shape template within the target dataset. Furthermore, the use of a fixed shape template was also not robust enough to capture all the shape variations found within the population. This can be seen from the low average dice coefficient of 0.869 in Table 7.1.

The method App. B1 based on statistical shape model as prior shape knowledge has been proposed in Chap. 4, which replaces the fixed shape template of App. A, to better capture the anatomical shape variability found within the population. As can be seen from Table 7.1, the increase in average segmentation accuracy from 0.869 to 0.884 was mainly due to the use of the PPCA based statistical shape model. In an iterative scheme, this resulted in a better fitting of the model to the specific patient’s muscle anatomy in comparison to the mean shape. This led to a decrease in the standard deviation to 0.029 compared to 0.034 of App. A indicating the increase in the stability and the robustness of the method. The increase in the median, smallest and the largest dice coefficient values also lent more weight that the method is superior compared to the method that used just the mean shape. The shape refinement mechanism used was linear, where all vertices incurred the same penalty.

To make the shape update and fitting more robust, a non-linear shape update mechanism was proposed in App. C. The proposed non-linear shape update mechanism further increased the accuracy of the statistical shape model fitting. The method resulted in achieving the highest average dice coefficient (0.895) of all the proposed methods. The standard deviation of App. C is also the lowest among all approaches, which shows that the method is the most stable and robust. Similarly, the non-linear shape update method App. C has the highest median as well as the smallest and largest DC values of all the methods.

The proposed methods App. A, App. B1 and App. C were also compared to App. F, which also used a fixed shape template in the boundary term of the energy function. All the proposed methods were found to be performing better than App. F, which led to the conclusion that adding prior shape knowledge in the data term of the energy function provides better segmentation results. The graphs in Fig. 7.1 show the com-

7. CONCLUSION

bined quantitative results of all the proposed approaches as well as a comparison of them with App. F.

A statistical analysis of the proposed approaches amongst each other, as well as with App. F was also conducted. A t-test with significance level 0.05 was performed over the dice coefficient values (see Table 7.2). It showed that App. C outperformed all the other approaches (first row of Table 7.2), as the method used a 3D statistical model with a non-linear and sparse shape update mechanism. Another important feature of our method is that it incorporates the shape knowledge in the data term of the energy function, which helps in achieving robust segmentation results even in the presence of high-density imaging artifacts.

The second row of Table 7.2 shows the comparison of the method App. B1 with the other methods. Statistically it was found to be significantly better than both the methods App. F and App. A. An improvement in the mean dice coefficient was also shown by App. B1 over App. A. The third row of the table shows the comparison of App. A and App. F. Statistically, App. A performed significantly better than App. F only for the datasets with severe image artifacts. All other statistical comparisons were carried out irrespective of the severity of the image artifacts.

	App. C	App. B1	App. A	App. F
App. C	-	p-value<0.05	p-value<0.05	p-value<0.05
App. B1	-	-	p-value<0.05	p-value<0.05
App. A	-	-	-	p-value<0.05
App. F	-	-	-	-

Table 7.2: Statistical comparison of the proposed methods (App. A, App. B1 and App. C) with the approach App. F.

7.2 Future Work

Our statistical model of the masseter muscle was created irrespective of the gender, age and race of the patients. Females tend to have shorter or slimmer muscles as compared to males. Similarly, children have smaller muscles while those of elderly people will show atrophy. Patients with African origin tend to have stronger, thicker or larger muscles

compared to the patients with Caucasian or Asian origin. Thus, building different statistical models based on the gender, age or region will allow to better capture the shape variations found within a particular group.

Separating different facial muscles from each other is still an elusive goal. This thesis mainly concentrates on segmenting all facial muscles as one object as it is sufficient for the targeted CMF surgery simulations. To distinguish different muscles or groups of muscles, Sener [14] and Kale *et al.* [30] have used regions of interest, which is a tedious and time consuming approach. A more feasible and automatic approach could be to use a labeled shape template and segment different muscles one after the other. An unsigned distance map for each muscle will encode shape and size knowledge for that muscle. This requires an accurate automatic alignment of the shape template to the target muscles, which is not a trivial task.

Higher order clique potentials of the MRF based energy function could also be used to achieve better and smoother results [138, 139, 140]. To guarantee that the global minimum of the energy function can be achieved, the higher order clique potential (third or fourth order clique potentials) functions should be submodular as explained by Kolmogorov & Zabih [52].

The proposed method seems to have some potential for future research. It is worth mentioning that the method in Chap. 6 can be used in future to create a pool of segmentations of the facial muscles from which a PPCA based statistical model of the facial muscles can be created. This statistical model can then be used in a similar manner to segment the masseter muscle outlined in this thesis. A statistical model of all the facial muscles together is more suitable than a statistical model of individual muscles. A larger and fatter face will have bigger and thicker facial muscles where as a shorter and slimmer face will have shorter and thinner muscles. This dependence of the facial muscles on the face size can readily be captured by a statistical model. However, to capture the variability of such a complex structure will require a much larger database of the segmented facial muscles than the one used for masseter muscle.

7. CONCLUSION

Bibliography

- [1] S. J. D. PRINCE. *Computer Vision: Models, Learning, and Inference*. Cambridge University Press, 1 edition, June 18 2012.
- [2] L. G. FARKAS. *Anthropometry of the Head and Face*. Raven Press, 1994.
- [3] E. GLADILIN. *Biomechanical Modeling of Soft Tissue and Facial Expressions for Cranio-Facial Surgery Planning*. PhD thesis, Fachbereich Mathematik und Informatik der Freien Universität Berlin, 2003.
- [4] E. GLADILIN, S. ZACHOW, P. DEUFLHARD, AND H.-C. HEGE. **Anatomy and Physics Based Facial Animation For Cranio-Facial Surgery Simulations**. *Medical and Biological Engineering and Computing*, **42(2)**(2):167–170, 2004.
- [5] G. BARBARINO, M. JABAREEN, J. TRZEWIK, AND E. MAZZA. **Physically Based Finite Element Model of the Face**. In *International Symposium on Biomedical Simulation (ISBMS)*, pages 1–10. Springer-Verlag, 2008.
- [6] R. OLSZEWSKI, M. B. VILLAMIL, D. G. TREVISAN, L. P. NEDEL, C. M. D. S. FREITAS, H. REYCHLER, AND B. M. MACQ. **Towards an Integrated System for Planning and Assisting Maxillofacial Orthognathic Surgery**. *Computer Methods and Programs in Biomedicine*, **91(1)**:13–21, 2008.
- [7] J. M. PLOOIJ, T. J. MAAL, P. HAERS, W. A. BORSTLAP, A. M. KUIJPERS-JAGTMAN, AND S. J. BERG. **Digital Three-Dimensional Image Fusion Processes for Planning and Evaluating Orthodontics and Orthognathic Surgery. A Systematic Review**. *International Journal of Oral and Maxillofacial Surgery*, **40(4)**:341–352, 2011.
- [8] E. MAZZA AND G. G. BARBARINO. **3D Mechanical Modeling of Facial Soft Tissue for Surgery Simulation**. *3D Imaging Technologies for Facial Plastic Surgery*, **19(4)**:623–637, 2011.
- [9] R. OLSZEWSKI, Y. LIU, T. DUPREZ, T. M. XU, AND H. REYCHLER. **Three-Dimensional Appearance of the Lips Muscles with Three-Dimensional Isotropic MRI: In Vivo Study**. *International Journal of Computer Assisted Radiology and Surgery (CARS)*, **4(4)**:349–352, 2009.

BIBLIOGRAPHY

- [10] W. MOLLEMANS, F. SCHUTYSER, N. NADJMI, AND P. SUETENS. **Very Fast Soft Tissue Predictions With Mass Tensor Model for Maxillofacial Surgery Planning Systems.** *International Journal of Computer Assisted Radiology and Surgery (CARS)*, **1281**:491–496, 2005.
- [11] W. MOLLEMANS, F. SCHUTYSER, N. NADJMI, F. MAES, AND P. SUETENS. **Predicting Soft Tissue Deformations for a Maxillofacial Surgery Planning System: From Computational Strategies to a Complete Clinical Validation.** *Medical Image Analysis*, **11**(3):282–301, 2007.
- [12] H. KIM, P. JÜRGENS, S. WEBER, L.-P. NOLTE, AND M. REYES. **A new Soft-Tissue Simulation Strategy for Cranio-Maxillofacial Surgery using Facial Muscle Template Model.** *Biomechanical Modelling of Soft Tissue Motion*, **103**(2):284–291, 2010.
- [13] H. KIM, P. JÜRGENS, AND M. REYES. **Soft-Tissue Simulation for Cranio-Maxillofacial Surgery: Clinical Needs and Technical Aspects.** *Patient-Specific Modeling in Tomorrow's Medicine*, **9**:413–440, 2012.
- [14] E. SENER. *Automatic Bayesian Segmentation Of Human Facial Tissue Using 3D MR-CT Fusion By Incorporating Models Of Measurement Blurring, Noise And Partial Volume.* PhD thesis, The Graduate School Of Natural And Applied Sciences Of Middle East Technical University, 2012.
- [15] E. SIFAKIS, A. SELLE, A. ROBINSON-MOSHER, AND R. FEDKIW. **Simulating Speech with a Physics-Based Facial Muscle Model.** In M.-P. CANI AND J. O'BRIEN, editors, *Eurographics Symposium on Computer Animation*, pages 261–270, Vienna, Austria, 2006. Eurographics Association.
- [16] Y. ZHANG, E. C. PRAKASH, AND E. SUNG. **A New Physical Model with Multilayer Architecture for Facial Expression Animation Using Dynamic Adaptive Mesh.** *IEEE Transactions on Visualization and Computer Graphics*, **10**(3):339–352, 2004.
- [17] S. D. GREEF AND G. WILLEMS. **Three-Dimensional Cranio-Facial Reconstruction in Forensic Identification: Latest Progress and New Tendencies in the 21st Century.** *Journal of Forensic Sciences*, **50**(1):12–17, 2005.
- [18] K. ZEPÄ, I. URTANE, Z. KRISJANE, AND G. KRUMINA. **Three-Dimensional Assessment of Musculoskeletal Features in Class II and Class III Patients.** *Baltic Dental and Maxillofacial Journal*, **11**(1):15–20, 2009.
- [19] T. MAJEED, K. FUNDANA, M. LÜTHI, J. BEINEMANN, AND P. CATTIN. **A Shape Prior-Based MRF Model for 3D Masseter Muscle Segmentation.** In *SPIE Medical Imaging: Image Processing*, **8314**, 2012.
- [20] T. MAJEED, K. FUNDANA, M. LÜTHI, S. KIRIYANTHAN, J. BEINEMANN, AND P. C. CATTIN. **Using a Flexibility Constrained 3D Statistical Shape Model for Robust MRF-Based Segmentation.** In *Mathematical Methods in Biomedical Image Analysis (MMBIA)*, pages 57–64, 2012.

-
- [21] T. MAJEED, K. FUNDANA, S. KIRIYANTHAN, J. BEINEMANN, AND P. CATTIN. **Graph Cut Segmentation using a Constrained Statistical Model with Non-Linear and Sparse Shape Optimization.** In *Medical Image Computing and Computer-Assisted Intervention (MICCAI) Workshop on Medical Computer Vision (MCV)*, pages 48–58, 2012.
- [22] H. P. NG, S. H. ONG, S. HUANG, J. LIU, K. W. C. FOONG, P. S. GOH, AND W. L. NOWINSKI. **Salient Features useful for the Accurate Segmentation of Masticatory Muscles from Minimum Slices Subsets of Magnetic Resonance Images.** *Machine Vision and Applications*, **21**(4):449–467, September 2008.
- [23] H. P. NG, Q. M. HU, S. H. ONG, K. W. C. FOONG, P. S. GOH, J. LIU, AND W. L. NOWINSKI. **Segmentation of the Temporalis Muscle from MR Data.** *International Journal of Computer Assisted Radiology and Surgery (CARS)*, **2**(1):19–30, January 2007.
- [24] H. P. NG, S. H. ONG, K. W. C. FOONG, P. S. GOH, AND W. L. NOWINSKI. **Automatic Segmentation of Muscles of Mastication from Magnetic Resonance Images Using Prior Knowledge.** In *International Conference on Pattern Recognition (ICPR)*, **Volume 03**, pages 968–971, September 2006.
- [25] H. P. NG, S. H. ONG, Q. HU, K. W. C. FOONG, P. S. GOH, AND W. L. NOWINSKI. **Muscles of Mastication Model-Based MR Image Segmentation.** *International Journal of Computer Assisted Radiology and Surgery (CARS)*, **1**(3):137–148, October 2006.
- [26] M. FARRUGIA, G. BYDDER, J. FRANCIS, AND M. ROBSON. **Magnetic Resonance Imaging of Facial Muscles.** *Clinical Radiology*, **62**(11):1078–1086, November 2007.
- [27] C. E. HUISINGA-FISCHER, J. M. VAANDRAGER, F. W. ZONNEVELD, AND B. PRAHL-ANDERSEN. **Precision and Accuracy of CT-Based Measurements of Masticatory Muscles in Patients with Hemifacial Microsomia.** *Dentomaxillofacial Radiology*, **33**(1):12–16, 2004.
- [28] M. MEEHAN, M. TESCHNER, AND S. GIROD. **Three-Dimensional Simulation and Prediction of Craniofacial Surgery.** *Orthodontics and Craniofacial Research*, **6**(1):102–107, 2003.
- [29] M. DE MAIO AND B. RZANY. *Botulinum Toxin in Aesthetic Medicine*. Springer, 2 edition, 2009.
- [30] E. H. KALE, E. U. MUMCUOGLU, AND S. HAMCAN. **Automatic Segmentation of Human Facial Tissue by MRI-CT Fusion: A Feasibility Study.** *Computer Methods and Programs in Biomedicine*, **108**(3):1106–1120, 2012.
- [31] D. W. SHATTUCK, S. R. SANDOR-LEAHY, K. A. SCHAPER, D. A. ROTTENBERG, AND R. M. LEAHY. **Magnetic Resonance Image Tissue Classification Using a Partial Volume Model.** *NeuroImage*, **13**(5):856–876, May 2001.
- [32] T. TASDIZEN, D. WEINSTEIN, AND J. N. LEE. **Automatic Tissue Classification for the Human Head from Multispectral MRI.** Technical report, Scientific Computing and Imaging Institute, University of Utah, 2004.
- [33] Y. REZAEITABAR AND I. ULUSOY. **Automatic 3D Segmentation of Individual Facial Muscles Using Unlabeled Prior Information.** *International Journal of Computer Assisted Radiology and Surgery (CARS)*, **7**(1):35–41, 2011.

BIBLIOGRAPHY

- [34] H. P. NG, K. W. C. FOONG, S. H. ONG, J. LIU, P. S. GOH, AND W. L. NOWINSKI. **Shape Determinative Slice Localization for Patient-Specific Masseter Modeling using Shape-Based Interpolation.** *International Journal of Computer Assisted Radiology and Surgery (CARS)*, **2**(1):398–400, 2007.
- [35] H. P. NG, J. LIU, S. HUANG, S. H. ONG, K. W. C. FOONG, P. S. GOH, AND W. L. NOWINSKI. **An Improved Shape Determinative Slice Determination Method for Patient-Specific Modeling of Facial Anatomical Structure.** *International Journal of Computer Assisted Radiology and Surgery (CARS)*, **3**(3):221–230, 2008.
- [36] H. P. NG, S. H. ONG, J. LIU, S. HUANG, K. W. C. FOONG, P. S. GOH, AND W. L. NOWINSKI. **3D Segmentation and Quantification of a Masticatory Muscle from MR Data Using Patient-Specific Models and Matching Distributions.** *Journal of Digital Imaging*, **22**(5):449–462, October 2009.
- [37] N. R. PAL AND S. K. PAL. **A Review on Image Segmentation Techniques.** *Pattern Recognition*, **26**(9):1277–1294, 1993.
- [38] D. L. PHAM, C. XU, AND J. L. PRINCE. **Current Methods In Medical Image Segmentation.** *Annual Review of Biomedical Engineering*, **2**(2000):315–337, 2000.
- [39] R. C. GONZALEZ AND R. E. WOODS. *Digital Image Processing.* Prentice Hall, 2002.
- [40] D. RIGOTTI, M. INGLESE, AND O. GONEN. **Whole-Brain N-Acetylaspartate as a Surrogate Marker of Neuronal Damage in Diffuse Neurologic Disorders .** *American Journal of Neuroradiology*, **28**(10):1843–1849, 2007.
- [41] D. J. WITHEY AND Z. J. KOLES. **A Review of Medical Image Segmentation Methods and Available Software.** *International Journal of Bioelectromagnetism*, **10**(3):125–148, 2008.
- [42] N. OTSU. **A Threshold Selection Method from Gray-Level Histograms.** *IEEE Transactions on Systems, Man and Cybernetics*, **9**(1):62–66, 1979.
- [43] S. A. HOJJATOLESLAMI AND J. KITTLER. **Region Growing: A New Approach.** *IEEE Transactions on Image Processing*, **7**(7):1079–1084, 1998.
- [44] S. A. HOJJATOLESLAMI AND F. KRUGGEL. **Segmentation of Large Brain Lesions.** *IEEE Transactions on Medical Imaging (T-MI)*, **20**(7):666–669, 2001.
- [45] L. CHEN. **Lambda-Connected Segmentation and the Optimal Algorithm for Split-and-Merge Segmentation.** *Chinese Journal of Computers*, **14**:321–331, 1991.
- [46] S. BEUCHER. **The Watershed Transformation Applied To Image Segmentation.** *Scanning Microscopy International*, **6**:299–314, 1991.
- [47] H. K. HAHN AND H.-O. PEITGEN. **The Skull Stripping Problem in MRI Solved by a Single 3D Watershed Transform.** In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 134–143, 2000.

-
- [48] J. F. BARRETT AND N. KEAT. **Artifacts in CT: Recognition and Avoidance.** *Radio Graphics*, **24**:1679–1691, 2004.
- [49] S. Z. LI. *Markov Random Field Modeling in Image Analysis*. Springer-Verlag, 3 edition, 2009.
- [50] A. BESBES, N. KOMODAKIS, G. LANGS, AND N. PARAGIOS. **Shape priors and discrete MRFs for knowledge-based segmentation.** In *Computer Vision and Pattern Recognition (CVPR)*, pages 1295–1302, 2009.
- [51] D. CREMERS, M. ROUSSON, AND R. DERICHE. **A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape.** *International Journal of Computer Vision (IJCV)*, **72**(2):195–215, 2007.
- [52] V. KOLMOGOROV AND R. ZABIH. **What Energy Functions can be Minimized via Graph Cuts?** In *European Conference on Computer Vision (ECCV)*, pages 65–81, London, UK, 2002. Springer-Verlag.
- [53] L. D. COHEN. **On Active Contour Models and Balloons.** *Computer Vision, Graphics and Image Processing: Image Understanding*, **53**(2):211–218, 1991.
- [54] M. KASS, A. WITKIN, AND D. TERZOPOULOS. **Snakes: Active Contour Models.** *International Journal of Computer Vision (IJCV)*, **1**(4):321–331, 1988.
- [55] W. NEUENSCHWANDER, P. FUA, G. SZÉKELY, AND O. KUEBLER. **Making Snakes Converge from Minimal Initialization.** In *International Conference on Pattern Recognition (ICPR)*, pages 613–615, 1994.
- [56] C. Y. XU AND J. L. PRINCE. **Snakes, Shapes, and Gradient Vector Flow.** *IEEE Transactions on Image Processing*, **7**(3):359–369, March 1998.
- [57] T. F. CHAN AND L. A. VESE. **Active Contours without Edges.** *IEEE Transactions on Image Processing*, **10**:266–277, 2001.
- [58] R. RONFARD. **Region-Based Strategies for Active Contour Models.** *International Journal of Computer Vision (IJCV)*, **13**(2):229–251, October 1994.
- [59] V. CASELLES, R. KIMMEL, AND G. SAPIRO. **Geodesic Active Contours.** *International Journal of Computer Vision (IJCV)*, **22**(1):61–79, February 1997.
- [60] A. TSAI, J. A. YEZZI, W. WELLS, C. TEMPANY, D. TUCKER, A. FAN, W. GRIMSON, AND A. WILLSKY. **A Shape-Based Approach to the Segmentation of Medical Imagery Using Level Sets.** *IEEE Transactions on Medical Imaging (T-MI)*, **22**(2):137–154, 2003.
- [61] M. E. LEVENTON, W. E. L. GRIMSON, AND O. FAUGERAS. **Statistical Shape Influence in Geodesic Active Contours.** In *Computer Vision and Pattern Recognition (CVPR)*, pages 316–323, 2000.
- [62] R. MALLADI, J. A. SETHIAN, AND B. C. VEMURI. **Shape Modeling with Front Propagation: A Level Set Approach.** *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, **17**(2):158–175, February 1995.

BIBLIOGRAPHY

- [63] D. CREMERS. **Nonlinear Dynamical Shape Priors for Level Set Segmentation.** *Journal of Scientific Computing*, **35**(2-3):132–143, 2007.
- [64] X. BRESSON, P. VANDERGHEYNST, AND J. P. THIRAN. **A Variational Model for Object Segmentation Using Boundary Information and Shape Prior Driven by the Mumford-Shah Functional.** *International Journal of Computer Vision (IJCV)*, **68**(2):145–162, June 2006.
- [65] D. CREMERS, S. J. OSHER, AND S. SOATTO. **Kernel Density Estimation and Intrinsic Alignment for Shape Priors in Level Set Segmentation.** *International Journal of Computer Vision (IJCV)*, **69**(3):335–351, September 2006.
- [66] L. DORNHEIM, K. D. TOENNIES, AND J. DORNHEIM. **Stable Dynamic 3D Shape Models.** In *International Conference on Image Processing (ICIP)*, **3**, pages 1276–1279, 2005.
- [67] L. DORNHEIM, J. DORNHEIM, AND K. D. TOENNIES. **Automatic Generation of Dynamic 3D Models for Medical Segmentation Tasks.** In *SPIE Medical Imaging*, 2006.
- [68] Y. BOYKOV AND G. FUNKA-LEA. **Graph Cuts and Efficient N-D Image Segmentation.** *International Journal of Computer Vision (IJCV)*, **70**(2):109–131, November 2006.
- [69] L. HE, Z. G. PENG, B. EVERDING, X. WANG, C. Y. HAN, K. L. WEISS, AND W. G. WEE. **A Comparative Study of Deformable Contour Methods on Medical Image Segmentation.** *Image and Vision Computing*, **26**(2):141–163, February 2008.
- [70] A. K. JAIN, Y. ZHONG, AND M.-P. DUBUISSON-JOLLY. **Deformable Template Models: A Review.** *Signal Processing*, **71**:109–129, 1998.
- [71] T. MCINERNEY AND D. TERZOPOULOS. **Deformable Models in Medical Image Analysis: A Survey.** *Medical Image Analysis*, **1**(2):91–108, 1996.
- [72] T. F. CHAN, S. ESEDOGLU, AND M. NIKOLOVA. **Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models.** *SIAM Journal of Applied Mathematics*, **66**(5):1632–1648, 2006.
- [73] X. BRESSON, S. ESEDOGLU, P. VANDERGHEYNST, J. P. THIRAN, AND S. J. OSHER. **Fast Global Minimization of the Active Contour/Snake Model.** *Journal of Mathematical Imaging and Vision (JMIV)*, **28**(2):151–167, June 2007.
- [74] E. S. BROWN, T. F. CHAN, AND X. BRESSON. **Completely Convex Formulation of the Chan-Vese Image Segmentation Model.** *International Journal of Computer Vision*, **98**(1):103–121, 2012.
- [75] A. CHAMBOLLE AND T. POCK. **A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging.** *Journal of Mathematical Imaging and Vision*, **40**(1):120–145, 2011.
- [76] T. POCK, A. CHAMBOLLE, D. CREMERS, AND H. BISCHOF. **A Convex Relaxation Approach for Computing Minimal Partitions.** In *Computer Vision and Pattern Recognition (CVPR)*, pages 810–817, 2009.

-
- [77] D. M. GREIG, B. T. PORTEOUS, AND A. H. SEHEULT. **Exact Maximum A Posteriori Estimation for Binary Images.** *Journal of the Royal Statistical Society*, **51**(2):271–279, 1989.
- [78] Y. Y. BOYKOV AND M.-P. JOLLY. **Interactive Graph Cuts for Optimal Boundary and Region Segmentation of Objects in N-D Images.** In *International Conference on Computer Vision (ICCV)*, **1**, pages 105 – 112, 2001.
- [79] J. WU AND A. C. S. CHUNG. **Cross Entropy: A New Solver for Markov Random Field Modeling and Applications to Medical Image Segmentation.** In J. S. DUNCAN AND G. GERIG, editors, *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, **3749** of *Lecture Notes in Computer Science*, pages 229–237, Heidelberg, September 2005. Springer-Verlag.
- [80] J. WU AND A. C. S. CHUNG. **Markov Random Field Energy Minimization via Iterated Cross Entropy with Partition Strategy.** In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, **1**, pages 457–460, 2007.
- [81] A. X. FALCO, J. K. UDUPA, S. SAMARASEKERA, S. SHARMA, B. E. HIRSCH, AND R. DE A. LOTUFO. **User-Steered Image Segmentation Paradigms: Live Wire and Live Lane.** *Graphical Models and Image Processing*, **60**(4):233–260, July 1998.
- [82] E. N. MORTENSEN AND W. A. BARRETT. **Interactive Segmentation with Intelligent Scissors.** *Graphical Models and Image Processing*, **60**(5):349–384, September 1998.
- [83] I. J. COX, S. B. RAO, AND Y. ZHONG. **'Ratio Regions': A Technique for Image Segmentation.** In *International Conference on Pattern Recognition (ICPR)*, **2**, pages 557–564, 1996.
- [84] I. H. JERMYN AND H. ISHIKAWA. **Globally Optimal Regions and Boundaries.** In *International Conference on Computer Vision (ICCV)*, page 904, 1999.
- [85] J. DORNHEIM, D. LEHMANN, L. DORNHEIM, B. PREIM, AND G. STRAUSS. **Reconstruction of Blood Vessels from Neck CT Datasets using Stable 3D Mass-Spring Models.** In *Eurographics Workshop on Visual Computing for Biomedicine*, pages 77–82, 2008.
- [86] J. DORNHEIM, H. SEIM, B. PREIM, I. HERTEL, AND G. STRAUSS. **Segmentation of Neck Lymph Nodes in CT Datasets with Stable 3D Mass-Spring Models.** In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 904–911, 2006.
- [87] J. PEARL. **Reverend Bayes on Inference Engines: A Distributed Hierarchical Approach.** *Association for the Advancement of Artificial Intelligence (AAAI)*, pages 133–136, 1982.
- [88] J. S. YEDIDIA, W. T. FREEMAN, AND Y. WEISS. **Understanding Belief Propagation and its Generalizations.** Technical report, Mitsubishi Electric Research Laboratories, nov 2001.
- [89] Y. Y. BOYKOV AND M.-P. JOLLY. **Interactive Organ Segmentation Using Graph Cuts.** In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 276–286. Springer-Verlag, 2000.

BIBLIOGRAPHY

- [90] L. R. FORD AND D. R. FULKERSON. *Flows in Networks*. Princeton University Press, 1960.
- [91] G. G. SLABAUGH AND G. UNAL. **Graph Cuts Segmentation Using an Elliptical Shape Prior**. In *International Conference on Image Processing (ICIP)*, pages 1222–1225, 2005.
- [92] D. FREEDMAN AND T. ZHANG. **Interactive Graph Cut Based Segmentation with Shape Priors**. In *Computer Vision and Pattern Recognition (CVPR)*, pages 755–762, 2005.
- [93] D. R. CHITTAJALLU, S. K. SHAH, AND I. A. KAKADIARIS. **A Shape-Driven MRF Model for the Segmentation of Organs in Medical Images**. In *Computer Vision and Pattern Recognition (CVPR)*, pages 3233–3240, San Francisco, CA, USA, June 13–June 18 2010.
- [94] M. FREIMAN, A. KRONMAN, S. J. ESSES, L. JOSKOWICZ, AND J. SOSNA. **Non-parametric Iterative Model Constraint Graph Min-Cut for Automatic Kidney Segmentation**. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, **13** of *MICCAI'10*, pages 73–80, Berlin, Heidelberg, 2010. Springer-Verlag.
- [95] O. VEKSLER. **Star Shape Prior for Graph-Cut Image Segmentation**. In *European Conference on Computer Vision (ECCV)*, **5304**, pages 454–467, 2008.
- [96] Q. SONG, Y. LIU, Y. LIU, P. K. SAHA, M. SONKA, AND X. WU. **Graph Search with Appearance and Shape Information for 3-D Prostate and Bladder Segmentation**. *Medical Image Computing And Computer-Assisted Intervention (MICCAI)*, **13**:172–180, 2010.
- [97] A. M. ALI, A. A. FARAG, AND A. S. EL-BAZ. **Graph Cuts Framework for Kidney Segmentation with Prior Shape Constraints**. In *Medical Image Computing And Computer-Assisted Intervention (MICCAI)*, **10**, pages 384–392, 2007.
- [98] M. FREIMAN, N. BROIDE, M. NATANZON, E. NAMMER, O. SHILON, L. WEIZMAN, L. JOSKOWICZ, AND J. SOSNA. **Vessels-Cut: A Graph Based Approach to Patient-Specific Carotid Arteries Modeling**. In *Modelling The Physiological Human*, 2009.
- [99] P. DAS, O. VEKSLER, V. ZAVADSKY, AND Y. BOYKOV. **Semi-Automatic Segmentation with Compact Shape Prior**. *Image and Vision Computing*, **27**(1):206–219, January 2009.
- [100] N. EL-ZEHIRY AND A. ELMAGHRABY. **Graph Cut Based Deformable Model with Statistical Shape Priors**. In *International Conference on Pattern Recognition (ICPR)*, pages 1–4, 2008.
- [101] N. VU AND B. S. MANJUNATH. **Shape Prior Segmentation of Multiple Objects with Graph Cuts**. In *Computer Vision and Pattern Recognition (CVPR)*, 2008.
- [102] J. MALCOLM, Y. RATHI, AND A. TANNENBAUM. **Graph Cut Segmentation with Nonlinear Shape Priors**. In *International Conference on Image Processing (ICIP)*, **4**, pages 365–368, 2007.
- [103] J. ZHU-JACQUOT. **Graph Cuts Segmentation with Geometric Shape Priors for Medical Images**. In *IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pages 109–112. IEEE Computer Society, 2008.

-
- [104] A. WIMMER, G. SOZA, AND J. HORNEGGER. **A Generic Probabilistic Active Shape Model for Organ Segmentation**. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, **5762**, pages 26–33, 2009.
- [105] J. T. BUSHBERG, J. A. SEIBERT, E. M. LEIDHOLDT, AND J. M. BOONE. *The Essential Physics of Medical Imaging*. Lippincott Williams & Wilkins, 2 edition, December 2001.
- [106] P. FUA AND C. BRECHBÜHLER. **Imposing Hard Constraints on Soft Snakes**. In *European Conference on Computer Vision (ECCV)*, pages 495–506, 1996.
- [107] S. Z. LI. **Markov Random Field Models in Computer Vision**. *European Conference on Computer Vision (ECCV)*, **801**:361–370, 1994.
- [108] S. GEMAN AND D. GEMAN. **Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images**. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, **6**(6):721–741, November 1984.
- [109] N. VILLAIN, Y. GOUSSARD, J. IDIER, AND M. ALLAIN. **Three-Dimensional Edge-Preserving Image Enhancement for Computed Tomography**. *IEEE Transactions on Medical Imaging (T-MI)*, **22**(10):1275–1287, 2003.
- [110] J. WOO, E. Z. MURANO, M. C. STONE, AND J. L. PRINCE. **Reconstruction of High-Resolution Tongue Volumes From MRI**. *IEEE Transactions on Biomedical Engineering*, **59**(12):3511–3524, 2012.
- [111] Y. CHEN AND A. A. AMINI. **A MAP Framework for Tag Line Detection in SPAMM Data Using Markov Random Fields on the B-Spline Solid**. *IEEE Transactions on Medical Imaging (T-MI)*, **21**(9):1110–1122, 2002.
- [112] M. SVENSN, F. KRUGGEL, AND D. Y. VON CRAMON. **Probabilistic Modeling Of Single-Trial FMRI Data**. *IEEE Transactions on Medical Imaging (T-MI)*, **19**(1):25–35, 2000.
- [113] A. C. S. CHUNG, J. A. NOBLE, AND P. SUMMERS. **Vascular Segmentation of Phase Contrast Magnetic Resonance Angiograms Based On Statistical Mixture Modeling And Local Phase Coherence**. *IEEE Transactions on Medical Imaging (T-MI)*, **23**(12):1490–1507, 2004.
- [114] T. CHEN AND D. METAXAS. **A Hybrid Framework for 3D Medical Image Segmentation**. *Medical Image Analysis*, **9**(6):547–565, December 2005.
- [115] D. J. C. MACKAY. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003.
- [116] J. M. HAMMERSLEY AND P. CLIFFORD. **Markov Fields on Finite Graphs and Lattices**. 1971.
- [117] C. M. BISHOP. *Pattern Recognition and Machine Learning*. Springer, 2 edition, 2007.
- [118] S. KIRKPATRICK, C. D. GELATT, JR., AND M. P. VECCHI. **Optimization by Simulated Annealing**. *Science*, **220**(4598):671–680, 1983.

BIBLIOGRAPHY

- [119] J. BESAG. **On the Statistical Analysis of Dirty Pictures.** *Journal of the Royal Statistical Society*, **48**(3):259–302, 1986.
- [120] C. PETERSON AND B. SÖDERBERG. **A New Method for Mapping Optimization Problems onto Neural Networks.** *International Journal of Neural Systems*, **1**(3):3–22, 1989.
- [121] A. ROSENFELD, R. A. HUMMEL, AND S. W. ZUCKER. **Scene Labeling by Relaxation Operations.** *IEEE Transactions on Systems, Man and Cybernetics*, **6**(6):420–433, 1976.
- [122] A. BLAKE AND A. ZISSERMAN. *Visual Reconstruction*. MIT Press, 1987.
- [123] Y. BOYKOV, O. VEKSLER, AND R. ZABIH. **Fast Approximate Energy Minimization via Graph Cuts.** *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, **23**(11):1222–1239, November 2001.
- [124] A. DELONG AND Y. BOYKOV. **A Scalable Graph-Cut Algorithm for N-D Grids.** In *Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2008.
- [125] Y. BOYKOV AND V. KOLMOGOROV. **An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision.** *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, **26**(9):1124–1137, September 2004.
- [126] R. SEDGEWICK. *Algorithms in C++ Part 5: Graph Algorithms*. Addison-Wesley Professional, 2001.
- [127] T. H. CORMEN, C. E. LEISERSON, R. L. RIVEST, AND C. STEIN. *Introduction to Algorithms*. MIT Press, 2009.
- [128] T. F. COOTES AND C. J. TAYLOR. **Statistical Models of Appearance for Medical Image Analysis and Computer Vision.** In *SPIE Medical Imaging*, pages 236–248, 2001.
- [129] T. F. COOTES, C. J. TAYLOR, D. H. COOPER, AND J. GRAHAM. **Active Shape Models; Their Training and Application.** *Computer Vision and Image Understanding*, **61**(1):38–59, January 1995.
- [130] VOLKER BLANZ AND THOMAS VETTER. **A Morphable Model for the Synthesis of 3D Faces.** In *Conference on Computer Graphics and Interactive Techniques*, pages 187–194, New York, NY, USA, 1999. ACM Press/Addison-Wesley Publishing Co.
- [131] M. TURK AND A. PENTLAND. **Eigenfaces for Recognition.** *Cognitive Neuroscience*, **3**(1):71–86, 1991.
- [132] M. LÜTHI, T. ALBRECHT, AND T. VETTER. **Probabilistic Modeling and Visualization of the Flexibility in Morphable Models.** In *Mathematics of Surfaces*, pages 251–264, 2009.
- [133] A. DEDNER, M. LÜTHI, T. ALBRECHT, AND T. VETTER. **Curvature Guided Level Set Registration using Adaptive Finite Elements.** In *Pattern Recognition*, pages 527–536, 2007.

BIBLIOGRAPHY

- [134] M. E. TIPPING AND C. M. BISHOP. **Probabilistic Principal Component Analysis**. *Journal of the Royal Statistical Society*, **61**(3):611–622, 1999.
- [135] P. E. DANIELSSON. **Euclidean Distance Mapping**. *Computer Graphics and Image Processing*, **14**:227–248, 1980.
- [136] V. BLANZ AND T. VETTER. **Reconstructing the Complete 3D Shape of Faces from Partial Information**. In *Informationstechnik und Technische Informatik*, **44**, pages 295–302, 2002.
- [137] S. ZHANG, Y. ZHAN, M. DEWAN, J. HUANG, D. N. METAXAS, AND X. S. ZHOU. **Sparse Shape Composition: A New Framework for Shape Prior Modeling**. In *Computer Vision and Pattern Recognition (CVPR)*, pages 1025–1032, 2011.
- [138] H. ISHIKAWA. **Higher-Order Clique Reduction in Binary Graph Cut**. In *Computer Vision and Pattern Recognition (CVPR)*, pages 2993–3000, 2009.
- [139] A. M. ALI, A. A. FARAG, AND G. L. GIMELFARB. **Optimizing Binary MRFs with Higher Order Cliques**. *European Conference on Computer Vision (ECCV)*, **5304**:98–111, 2008.
- [140] P. KOHLI, M. P. KUMAR, AND P. H. S. TORR. **P3 & Beyond: Solving Energies with Higher Order Cliques**. In *Computer Vision and Pattern Recognition (CVPR)*, 2007.

Curriculum Vitae

PERSONAL DATA

Name	Tahir Majeed
Date of Birth	17.06.1981
Citizenship	Pakistan

EDUCATION

2009 – 2014	PhD in Biomedical Engineering Medical Image Analysis Center University of Basel, Switzerland Facial Soft Tissue Segmentation Supervisor: Prof. Dr. Philippe C. Cattin
2005 – 2009	M.Sc. in Computational Visualistic Fakultät für Informatik (FIN) Otto-von-Guericke-Universität Magdeburg, Germany
2001 – 2004	B.S. in Computer Science Department of Computer Science University of Karachi, Pakistan

REFERENCES

Prof. Dr. Philippe C. Cattin
Medical Image Analysis Center
Universität Basel
+41 (0)61 265 96 55
philippe.cattin@unibas.ch

Dr. Ketut Fundana
Medical Image Analysis Center (MIAC)
+41 (0)61 265 96 58
ketut.fundana@unibas.ch