Multi Atlas Fusion Methods for Medical Image Segmentation

THÈSE N° 5747 (2013)

PRÉSENTÉE LE 7 JUIN 2013
À LA FACULTÉ DES SCIENCES ET TECHNIQUES DE L'INGÉNIEUR
LABORATOIRE DE TRAITEMENT DES SIGNAUX 5
PROGRAMME DOCTORAL EN GÉNIE ÉLECTRIQUE

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Subrahmanyam VENKATA RAVI MOHANA SAI GORTHI

EPFL

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2013

Abstract

Automated segmentation of anatomical structures is still an open challenge in medical imaging. In particular, when the structures to be segmented do not possess distinct image characteristics (like well defined edges, shapes or region properties), it is quite challenging to obtain accurate and robust segmentations. In such cases, inclusion of more prior knowledge into the segmentations process is essential.

Atlas-based methods are widely used for exploiting prior knowledge. Atlas is a reference image, which is similar to the target image to be segmented, and in which structures of interest have been accurately segmented, usually by hand, or by semi-automatic methods. The general procedure for obtaining the segmentations in the target image can be summarized in a two-step procedure: first, a dense deformation field that maps the atlas onto the target image is computed using appropriate registration methods. Second, the deformation field computed above is applied to the already segmented structures of interest in the atlas, and this provides segmentations of the corresponding structures in the target image. It has been shown in many recent works that, instead of using just a single atlas, if the segmentations obtained from multiple atlases are merged together with an appropriate fusion strategy, then the resulting segmentations are more accurate and robust.

In this thesis, we therefore focus on two important aspects of atlas-based methods that have a significant impact on the accuracy of segmentation: (i) developing efficient registration methods for computing the dense deformation field, and (ii) developing efficient fusion methods for merging segmentations obtained from multiple atlases.

There are mainly four methodological contributions from this thesis.

First, we present a new and original variational framework for atlas-based registration. This framework is quite general and encompasses many state-of-the-art registration methods. It integrates both the active contour-based registration forces and the dense deformation fields of the optical flow framework. We also demonstrate the advantages of this framework by comparing it with other classical registration methods.

Second, we propose a Markov Random Field (MRF) based framework that performs simultaneously both fusion and edge-preserving smoothing of labels. We show how the existing fusion methods like majority voting, global weighted voting, local weighted voting and shape based averaging (SBA) can be reformulated to profit from the proposed framework. This framework has resulted in more accurate as well as contiguous regions

by getting rid of the unwanted holes and islands in the output segmentations.

Third, we propose two new fusion methods: "global weighted shape based averaging" and "local weighted shape based averaging". These methods extend the well known SBA method by additionally incorporating the similarity information between each atlas and the target image. The proposed methods have resulted in more accurate segmentations compared to the existing SBA method.

Fourth, we propose a new similarity-weighted pairwise term that models spatially, neighborhood priors for each possible label pair combination. Thanks to our MRF-based framework, it facilitates combining this neighborhood model with any of the above mentioned fusion methods. When compared to the existing fusion methods that do not use such prior information, the accuracy of segmentations has significantly improved with the inclusion of this new additional pairwise term.

Finally, from the applications perspective, we have performed various evaluations in the context of segmentation of lymph nodes in the 3D Head and Neck (H&N) CT images. These experiments have clearly illustrated the advantages of our registration model. The newly proposed SBA-based methods are found to improve the segmentation results significantly compared to the original SBA methods. These experiments have also confirmed the advantages of our MRF-based fusion framework that additionally incorporates edge-preserving smoothness term along with the neighborhood priors information.

To summarize, this thesis focuses on atlas-based segmentation. It presents (i) a variational framework for nonrigid registration, (ii) an MRF-based framework for simultaneously performing both fusion and edge-preserving smoothing, (iii) two new fusion methods, and (iv) a neighborhood model for incorporating more prior knowledge about label pair combinations into the fusion framework.

Keywords: Medical Imaging, Image Segmentation, Atlas-based Segmentation, Image Registration, Non-rigid Registration, Label Fusion, Multi Atlas Fusion, MRF, Lymph Nodes, SBA.

Résumé

La segmentation automatique des structures anatomiques est toujours une question ouverte en imagerie médicale. En particulier, lorsque les structures à segmenter ne possèdent pas de caractéristiques distinctives observables dans l'image (comme des bords bien définis, des formes ou des régions homogènes), il est assez difficile d'obtenir des segmentations précises et robustes. Dans ces cas, l'introduction de connaissances a priori dans le processus de segmentation est essentielle.

Les méthodes basées sur des atlas sont largement utilisées pour l'exploitation de connaissances a priori. L'atlas est une image de référence, qui est semblable à l'image à segmenter, et dans laquelle les structures d'intérêt ont été segmentés manuellement. La procédure générale d'obtention des segmentations d'une image peut être ainsi résumée en deux étapes : D'abord, un champ de déformation dense qui met en correspondance l'atlas avec l'image à segmenter est calculé en utilisant des méthodes de recalage appropriées. Deuxièmement, le champ de déformation ainsi calculé est appliquée aux structures d'intérêt déjà segmentées dans l'atlas, ce qui donne une segmentation des structures correspondantes dans l'image originale. De nombreux travaux récents ont démontré qu'il est possible obtenir des segmentations plus précises et robustes en utilisant plusieurs atlas, et en fusionant correctement leurs segmentations, au lieu d'utiliser simplement un atlas unique.

Dans cette thèse, nous avons donc étudié deux aspects importants des méthodes basées sur des atlas qui ont un impact significatif sur la précision de la segmentation : (i) le développement de méthodes efficaces de recalage pour le calcul du champ de déformation, et (ii) le développement de méthodes efficaces pour la fusion des segmentations obtenues à partir de plusieurs atlas.

Il y a principalement quatre contributions méthodologiques dans cette thèse.

Tout d'abord, nous présentons une nouvelle méthodologie pour le recalage basé sur des atlas. Cette méthodologie est très générale et englobe de nombreuses méthodes actuelles de recalage d'images. Il intègre à la fois les forces de recalage basés sur des contours actives et les champs de déformation denses typiques du flux optique. Nous avons également démontré les avantages de cette méthodologie en le comparant à d'autres méthodes de recalage classiques.

Deuxièmement, nous proposons une framework basée sur les champs aléatoires de Markov (MRF) qui effectue simultanément la fusion et le lissage des labels en préser-

vant les bords de l'image. Nous montrons comment les méthodes de fusion existantes, comme majority voting, global weighted voting, local weighted voting et shape based averaging (SBA) peuvent être reformulées et profiter de la méthodologie proposée. Ce framework a donné lieu à des segmentations plus précises avec des régions contigues, en se débarrassant des trous et îles non désirés.

Troisièmement, nous proposons deux nouvelles méthodes de fusion : "global weighted shape based averaging" et "local weighted shape based averaging". Ces méthodes améliorent la méthode de fusion SBA en tenant compte de la similitude entre chaque atlas et l'image à segmenter. Les méthodes proposées ont donné lieu à des segmentations plus précises par rapport à SBA.

Quatrièmement, nous proposons un nouveau terme qui tient compte de l'information a priori. A chaque pixel, ce terme mesure la similarité entre les voisinages de chaque possible combinaison de paires d'étiquettes. Notre méthode MRF facilite la combinaison de ce modèle avec l'une des méthodes de fusion mentionnées auparavant. En comparaison avec les méthodes de fusion existantes, qui n'utilisent pas ces informations a priori, la précision de la segmentation s'est considérablement améliorée grâce à l'introduction de ce terme supplémentaire.

Enfin, du point de vue des applications, nous avons procédé à diverses évaluations dans le contexte de la segmentation des ganglions lymphatiques dans des images CT du cou et de la tête 3D (H & N). Ces expériences ont clairement montré les avantages de notre modèle de recalage. Les méthodes basées sur SBA que nous avons proposées améliorent la segmentation significativement par rapport aux méthodes originales SBA. Ces expériences ont également confirmé les avantages de notre méthode de fusion MRF, qui incorpore un terme de lissage capable de préserver les bords de l'image et un terme avec l'information a priori.

Pour résumer, le but de cette thèse est la segmentation basée sur des atlas. Elle présente : (i) une méthode variationnelle de recalage non rigide, (ii) un framework basé sur les MRF pour effectuer simultanément la fusion et le lissage en préservant les bords de l'image, (iii) deux nouvelles méthodes de fusion, et (iv) un modèle de voisinage qui incorpore les connaissances a priori sur les combinaisons de paires d'étiquettes dans le cadre de la fusion.

Mots-clés : imagerie médicale, segmentation d'images, segmentation basée sur des atlas, recalage d'image, recalage non-rigide, fusion de labels, multi-fusion d'atlas, MRF, tête et cou, ganglions lymphatiques, la SBA.

Riassunto

La segmentazione automatica di strutture anatomiche rappresenta a tutt'oggi una sfida
aperta nel campo dell'imaging biomedicale. In particolare, quando le strutture da
segmentare non posseggono caratteristiche dell'immagine sufficientemente differenziate
(come ad esempio contorni, forme e regioni ben definite) è piuttosto difficile ottenere
segmentazioni robuste ed accurate. In questi casi, l'inclusione di informazioni aggiuntive
sulle strutture da segmentare è essenziale per ottenere risultati soddisfacenti.

Per poter sfruttare queste informazioni aggiuntive sono largamente utilizzati dei
metodi cosiddetti atlas-based, ovvero basati su atlanti anatomici. In questo contesto
un atlante è una immagine di riferimento, simile all'immagine da segmentare, e nella
quale le strutture di interesse sono state precedentemente segmentate molto accurata-
mente, solitamente a mano o tramite procedure semi-automatiche. Il procedimento per
ottenere le segmentazioni nell'immagine di destinazione a partire dalle informazioni
degli atlanti può essere schematizzato come segue. Per prima cosa viene stimata tramite
appropriati algoritmi di registrazione una trasformazione, o dense deformation field,
che mappi l'atlante nello spazio dell'immagine di destinazione. Successivamente, questa
trasformazione viene applicata alle strutture di interesse precedentemente segmen-
tate nell'atlante, al fine di ottenere le segmentazioni delle corrispondenti strutture
anatomiche anche nell'immagine di destinazione. In molteplici lavori recentemente
pubblicati è stato dimostrato che la qualità delle segmentazioni risulta più accurata
e robusta se, al posto di utilizzare un singolo atlante, vengono combinate assieme le
segmentazioni ottenute utilizzando diversi atlanti tramite apposite strategie di fusione.

In questa tesi l'attenzione è posta su due aspetti molto importanti di questi algoritmi
atlas-based, che possono avere un impatto significativo sulla qualità delle segmentazioni
finali: (i) sviluppo di efficienti algoritmi di registrazione per la stima del dense deforma-
tion field, e (ii) sviluppo di metodi adeguati per una fusione efficace delle segmentazioni
ottenute da atlanti diversi.

In questa tesi proponiamo essenzialmente quattro contributi metodologici.

Il primo è rappresentato da un variational framework innovativo ed originale per la
registrazione atlas-based. Questo framework è piuttosto generale e comprende molti
tra i più avanzati metodi di registrazione attualmente disponibili in letteratura. Inoltre
permette di integrare sia le active-contour-based forces per la registrazione che i dense
deformation field dell'optical flow framework. I vantaggi di questo framework verranno

dimostrati confrontandolo con altri metodi classici di registrazione.

Il secondo contributo proposto in questa tesi é un framework basato su Markov Random Field (MRF) che è in grado di effettuare simultaneamente la fusione e lo smoothing delle etichette delle strutture di interesse preservandone i contorni. Verrà dimostrato come i più diffusi algoritmi di fusione esistenti, ad esempio majority voting, global weighted voting, local weighted voting e shape based averaging (SBA), possano essere riformulati per poter trarre vantaggio dal nostro framework. Questo framework ha permesso di eliminare drasticamente i buchi e le isole indesiderate dalle segmentazioni finali, e quindi di ottenere regioni più contigue ed accurate.

Vengono inoltre introdotti due nuovi metodi di fusione: "global weighted shape based averaging" e "local weighted shape based averaging". Quest'ultimi estendono il ben noto metodo SBA incorporando informazioni di similarità tra ciascun atlante e l'immagine di destinazione. I metodi proposti hanno permesso di ottenere segmentazioni più accurate rispetto alla procedura SBA classica.

Infine, si propone un nuovo termine pairwise similarity-weighted per modellare spazialmente l'informazione sul "vicinato" tra ciascuna possibile coppia di etichette delle strutture di interesse. Grazie al nostro framework basato su MRF, risulta più semplice combinare questo nuovo modello con ciascuno dei metodi di fusione menzionati in precedenza. In un confronto con metodi esistenti di fusione che non fanno uso di questa informazione aggiuntiva, l'accuratezza delle segmentazioni mostra un significativo miglioramento con l'inclusione di questo nuovo termine.

Per quanto riguarda l'aspetto applicativo, numerose valutazioni sono state effettuate nel contesto della segmentazione dei linfonodi da immagini TAC 3D di testa e collo (H&N). Questi esperimenti hanno chiaramente messo in luce i vantaggi del nostro modello di registrazione. Le segmentazioni ottenute con i nuovi metodi presentati in questa tesi evidenziano una qualità significativamente superiore rispetto ai metodi SBA originali. Inoltre, questi esperimenti hanno anche confermato i vantaggi del nostro framework di fusione basato su MRF, il quale incorpora un termine di smoothing che preserva i contorni oltre alla già citata informazione sul vicinato.

In conclusione, questa tesi è focalizzata sulla segmentazione basata su atlanti (atlas-based segmentation). In particolare essa presenta (i) un variational framework per la registrazione non lineare, (ii) un framework basato su MRF per effettuare simultaneamente fusione e smoothing delle strutture di interesse preservandone i contorni, (iii) due nuovi metodi di fusione, e (iv) un "modello di vicinato" per incorporare informazione aggiuntive su ciascuna coppia di strutture nel framework di fusione.

Parole chiave: imaging medicale, segmentazione di immagini, segmentazione basata su atlante, registrazione di immagini, registrazione non-lineare, fusione di etichette, fusione da atlanti multipli, MRF, linfonodi, SBA.

Acknowledgements

It gives me great pleasure to express my sincere thanks and gratitude to all those who have guided, helped, inspired, motivated, co-travelled, and supported me in various ways.

First of all, I would like express my deepest gratitude to my research supervisor: Prof. Jean-Philippe Thiran. It has been a great pleasure and a privilege for me to work under your guidance Sir. I had a very pleasant work-environment and could learn many things throughout all these years of my stay in Switzerland; this became possible mainly because of the freedom and encouragement you have always given. I am indebted to you for the way you mentored, and for everything that you have done for me.

I am thankful to my thesis jury members: Prof. Dimitri Van De Ville, Dr. A. S. Allal, Dr. Meritxell Bach Cuadra, and also to the president of the jury: Prof. Pierre Vandergheynst, for reading this manuscript and giving your valuable comments. My special thanks to you Meri, for the nice interactions we had, for all the collaborations with you in different projects. Many thanks to Dr. Allal and his team of medical experts: Pierre-Alain Tercier, Ulrike Schick and Javier Villafruela Vicario, for providing the database of CT images, for doing those laborious manual delineations of various structures, and for your feedback on the output results of our methods; the evaluations performed in this thesis would not have been possible without those valuable database of atlases.

I am grateful to the Swiss National Science Foundation (SNSF) for generously providing the financial support for my research work and travel.

I am ever indebted to the research guides of my Master's studies at the Indian Institute of Science (IISc), Bangalore: Dr. Atanu Mohanty and Prof. Anindya Chatterjee; without the initial training, teaching and caring that both of you bestowed upon me, I could not have come to an institute like EPFL, and I could not have done my PhD. I would like to express my gratitude to many other teachers also, starting from my primary school education till this PhD level, for educating, training, and inspiring me in the right direction.

I would like to thank my office mate Virginia for providing a cheerful and a greenery workplace :-) I relish all the technical and "not so technical" discussions we had! Special thanks to you for helping me in translating my thesis abstract to French. Thank you so much for your cordiality and care taking nature.

It's now time to mention about my friend - Alessandro, whom *I am not going to thank*.

Acknowledgements

Contents

Contents

List of Figures

List of Tables

1 Introduction

1.1 Atlas-based Segmentation

Accurate segmentation of anatomical structures is a fundamental task in medical imaging. This is required particularly for medical diagnosis, quantitative analysis, treatment planning and outcome evaluation of various pathologies. Manual delineation of the anatomical structures is often laborious as well as time consuming task, and thus, there is a need for robust and accurate automated segmentation methods. While it is relatively easy to automatically segment the structures that have distinct image characteristics, it is hard to directly segment those structures that do not possess distinct boundaries or other similar distinct features (like shape and region-based characteristics) with respect to the surrounding structures; in order to automatically delineate such complex structures, it is essential to incorporate some sort of prior knowledge into the segmentation framework. There are indeed numerous application that require incorporation of such prior knowledge; to name a few, the applications include segmentation of structures like hippocampus, putamen and thalamus in Magnetic Resonance (MR) images [3–5], segmentation of lymph nodes in the head and neck Computed Tomography (CT) images [6–8], and segmentation of heart ventricles in cardiac MR images [9–12].

Atlas-based methods are well known for effectively incorporating prior knowledge in order to segment complex anatomical structures [2, 13–15]. An atlas contains a reference image, which is similar to the target image to be segmented, and in which structures of interest have been accurately segmented, usually by hand, or by semi-automatic methods. The typical atlas-based segmentation procedure can be summarized in two-steps. First, a dense deformation field that maps each pixel in the atlas image onto the target image is computed using appropriate registration methods. Depending on the nature of the images and the target structures to be segmented, the registration methods use various types of forces, like pixel-based forces [2, 8, 16, 17], region-based forces [1, 8], statistical features [18–20], and geometrical features [21, 22]. Second, the deformation field computed in the preceding step is applied to the labels of the already segmented

1
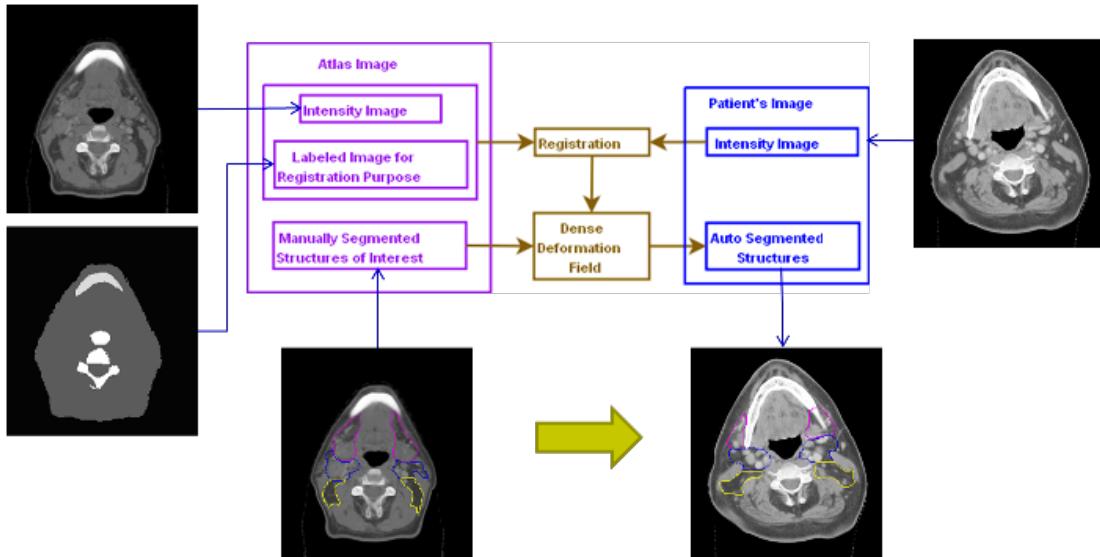
Figure 1.1: Block diagram illustrating the atlas-based segmentation procedure used for segmenting the structures of interest in the target image.

structures of interest in the atlas image, and this provides the segmentations of the corresponding structures in the target image.

Figure 1.1 illustrates the aforementioned segmentation procedure in more detail with the help of a block diagram. We consider here the example of segmentation of lymph nodes in the 3D head and neck CT images. In this figure, the contours of the lymph nodes are shown in one of the axial slices. Please note that, by saying "atlas," we refer combinedly to both the reference CT intensity image, and the labeled-image that specifies the structures to be used for driving the registration process. The labeled-image is however not used (or needed) for certain registration algorithms, and this point will become more clear to the reader in the following chapters. Finally, also note that the "labeled-image" is different from the image containing manual delineations of structures to be segmented in the target image. We repeat that "labeled-image" defines regions of interest to be used for driving the registration process; the dense deformation field computed through this registration is then applied to the manually segmented structures of interest in the atlas for which we want find out automated segmentations in the target image. With these notes, the above figure should be self-explanatory.

A typical underlying assumption in medical image registration is that, there is a one-to-one correspondence between the voxels of the atlas and the target image to be segmented [8]. However, in practice, there can be huge anatomical variations among different subjects, and this can consequently lead to inaccurate output segmentations. Several strategies have been proposed in the literature to overcome this problem. One simple approach is to adaptively select the most similar atlas for each target image to be segmented, from a database of available atlases; the similarity between the two images

for atlas selection is assessed in different ways ranging from visual inspection to, based on certain criteria like smallest resultant "mean square error," or least "deformation field," or maximum "normalized mutual information" between the two images [23–25].

Single-atlas-based approaches, however, require sufficiently large database of atlases for covering all possible anatomical variations. Further, it has been shown in many recent works [23, 26–28] that automated segmentations obtained by merging results from multiple atlases are more reliable and accurate than the results that are based on just one atlas despite using single-best-atlas-selection strategies. This is because, when the results from multiple atlases are combined together, registration errors coming from a particular propagated atlas are less likely to affect the final segmentations; when multiple atlases are combined, as the proportion of errors incurred during propagation of individual atlases are generally independent, they are averaged out [29, 30].

1.2 Multi Atlas Fusion

There are broadly two classes of approaches for fusing the information coming from multiple atlases. The first class of approaches tries to combine the information from several atlases by creating an average atlas or a probabilistic atlas [6, 31–36], and then register that to the target image. The second class of approaches registers each atlas independently to the target image, and then merges the segmentation results obtained from each individual atlas, based on certain optimal criteria [23, 26, 27]. It has been noticed in many recent works that the second class of approaches is more robust to anatomical variations, and it can also profit more from the information coming from multiple atlases than the first class of approaches [13, 23, 37]. In this thesis, we hence focus on this second class of approaches.

We now briefly introduce the state-of-the-art multi-atlas fusion methods, and please refer to the next chapter for a detailed description of some of those methods. The widely used fusion methods include majority voting (MV) [23, 27], global weighted voting (GWV) [27, 38], local weighted voting (LWV) [26, 27, 38], STAPLE [39, 40], and shape based averaging (SBA) [41]. MV method assigns for each voxel, a label that maximum number of atlases agree on. Unlike MV, GWV attaches a weight to each atlas while counting its vote. The weight for each atlas is determined globally, based on its similarity to the image to be segmented; the more the similarity, the higher the weight, and vice versa. LWV is similar to GWV except that, not a single global weight is assigned to the entire atlas; rather, for each voxel, an individual weight is assigned based on the local similarity to the corresponding voxel in the image to be segmented.

Another popular multi-atlas fusion method is "Simultaneous Truth And Performance Level Estimation" (STAPLE) [39]. In this method, the probabilistic estimate of the true segmentation is formed by estimating the optimal combination of segmentations,

weighting each segmentation depending upon estimated performance level. Although STAPLE was originally proposed for combining manual segmentations done by multiple experts, it is also used for combining automated segmentations obtained from multiple atlases. Several modifications and extensions were proposed to the original STAPLE algorithm. For instance, in [40], STAPLE algorithm is extended to multi-label segmentation. A continuous STAPLE for scalar, vector and tensor images is presented in [42]. Recently, [43] proposed a local Maximum A Posteriori (MAP) formulation of STAPLE algorithm that can account for spatially varying performance parameters while combining multiple segmentations.

When compared to the methods described above, SBA method looks at the fusion problem from a different perspective. For each voxel in the output image, SBA assigns a label that results in minimum "Signed Euclidean Distance" (SED) when summed up over all input atlas images; these SEDs are computed with respect to each label. Please refer to [41] for a nice pictorial illustration of this idea.

1.3 Objectives and Contributions

The main goal of this thesis is to develop new methodologies that can potentially improve the accuracy and reliability of atlas-based segmentation. It is evident from the discussions presented in the preceding sections that there are mainly two important components that have a significant impact on the quality of automated segmentations: first, the nonrigid registration approach used for finding out the point-by-point mapping between the atlas and the target images; second, the fusion strategy used for merging segmentation results obtained from multiple atlases.

Hence, the two main objectives of this thesis are:

- To develop a new registration framework that facilitates integrating (or utilizing) various types of registration forces (like pixel-based and region-based forces)

- To propose new models that can facilitate inclusion of more prior knowledge into atlas fusion methods.

Here are the main contributions with regard to the registration:

- We have presented a new variational framework for nonrigid registration that generalizes many of the existing state-of-the-art registration methods. It allows to select structures (or regions) that drive the registration; these structures can be chosen based on their consistency between the two images to be registered, and also by considering their robustness to registration errors. The proposed

framework also facilitates combining different registration forces in an hierarchical fashion. This work has been published in [8].

- The proposed registration framework is evaluated in the context of segmentation of lymph nodes in the 3D head and neck (H&N) CT images. First, we have evaluated the results from our method by comparing them with the results obtained from two other commonly used registration algorithms in medical imaging, viz, "Radial Basis Function" (RBF) algorithm [44] proposed by Rhode et al., and Demons algorithm [16] proposed by Thirion. Finally, as our framework facilitates integrating both the region-based forces coming from the Yezzi's model [1], and the pixel-based forces coming from the Vemuri's model [2], we have demonstrated the advantage of combining these forces in an hierarchical manner by comparing our results with the results obtained from those two methods. Parts of this work have been published in [7, 25].

There are three main contributions with regard to the atlas fusion methods:

- The first contribution is based on the observation that, although the segmentations obtained from individual atlases are contiguous, the merged segmentations can be fragmented, containing undesirable holes and islands [26]. To deal with this problem, the segmentation results are generally post-processed using the approaches like morphological operations and Gaussian smoothing. However, such approaches have many disadvantages like blurring of edges. Moreover, it is not elegant to handle "fusion" and "smoothing" as two different independent problems. To address this issue, we propose and evaluate a new MRF-based framework that performs simultaneously both fusion and edge-preserving smoothing. We also show how some of the existing fusion methods can be reformulated to profit from our framework. Parts of this work have been published in [45, 46].

- The second contribution is based on the observation that the widely used SBA fusion method do not benefit from the information regarding how similar the atlas and target images are. In this thesis, we propose two new fusion methods: "Global Weighted Shape Based Averaging" (GWSBA) and "Local Weighted Shape Based Averaging" (LWSBA). GWSBA and LWSBA methods extend the existing SBA by additionally incorporating the global and local similarity information respectively. We first show how the existing SBA can be reformulated into our MRF-based fusion framework using a logistic function-based transformation. We then propose how to incorporate the global and local similarity information into this model. The results from the proposed methods are also evaluated in the context of H&N lymph nodes segmentation.

- The third contribution is based on the observation that the existing fusion methods currently do not incorporate any prior knowledge regarding the probabilities for

each possible label pair combination at each voxel location. We notice that such information can be easily extracted from the transformed atlas images. In this thesis, we propose a neighborhood prior term that models, at each voxel location, similarity-weighted neighborhood priors for each possible label pair combination. Thanks to the MRF-based framework proposed in this thesis, it facilitates integrating this neighborhood prior model with any of the above mentioned fusion methods. Finally, we have evaluated the effects of adding this neighborhood prior term in the context of H&N lymph nodes segmentation.

Finally, in terms of clinical applications, the main contribution of this thesis is that we have built a complete system for performing automated segmentation of lymph nodes in the 3D H&N CT images. For this purpose, we have first constructed a valuable database of atlases; each atlas in this database contains 3D H&N CT intensity images, manual delineations of structures that are used for driving the region-based registration, and manual delineations of all lymph node structures. In order to be able to use the automated segmentation results in radiotherapy treatment planning, those segmentations have to be exported to the standard DICOM "Radiotherapy Structure Set" (RTSTRUCT) format. However, there are no publicly available tools to perform this task. In order to bridge this gap, we have developed an open-source software tool for exporting the segmentations of the structures to DICOM-RTSTRUCT format. This work has been published in [47]. We have also performed a comprehensive evaluation of various atlas fusion methods for performing H&N lymph nodes sementation. Parts of this work have been published in [28, 46, 48].

1.4   Outline

The structure of this thesis closely follows the sequence of aforementioned contributions.

In chapter 2, we present a review of the related state-of-the-art methods in image registration, multiple atlas fusion, and H&N lymph nodes segmentation. We also summarize different evaluation metrics that are used throughout this thesis.

In chapter 3, we present a new variational framework for atlas registration. We also evaluate the proposed framework with other widely used registration methods.

In chapter 4, we present our MRF-based framework for atlas fusion, and then show how the existing methods can be reformulated to fit into this framework. We also present a comprehensive evaluation of this framework.

In chapter 5, we present a reformulation of SBA, and then present two new SBA-based fusion methods that additionally incorporates similarity information between the atlas and the target images.

In chapter 6, we present a new neighborhood prior model, and integrate it with our MRF-based fusion framework. We then present an evaluation of this model along with the evaluation of the new SBA-based fusion methods that are proposed in the preceding chapter.

In chapter 7, we summarize the contributions reported in this thesis. Scope for the future work is also presented.

Finally, in appendices, we present two of our works that supplement the main objectives of this thesis. In appendix A, we formulate brain segmentation problems as an MRF optimization problem, and evaluate recent energy minimization methods for solving this problem. From the mathematical formulation perspective, it has certain similarities with the MRF-based fusion framework. This work has been published in [49]. In appendix B, we present an evaluation of the existing fusion methods for the segmentation of bilateral parotid glands in the H&N CT images. This work has been published in [28].

2 Related Work

The main purpose of this chapter is to establish the base required for describing our contributions that will be presented in the later chapters. We first present here a review of the relevant state-of-art registration methods. We then present the fusion methods which we will be reformulating later in chapter 4 and chapter 5. Since we will be performing in this thesis various evaluations on the H&N lymph nodes segmentation application, we will also present in this chapter a review of different approaches that have been used previously for this specific application. Finally, we present a description of various quantitative measures that are used in this thesis for evaluating the accuracy of lymph nodes segmentations.

2.1 Image Registration

As described in chapter 1.1, atlas-based segmentation is a widely used approach for automated segmentation of medical images. It requires performing a nonrigid registration between the atlas and the target image. This nonrigid registration essentially gives a point-to-point spatial correspondence between the atlas and the target images.

One of the main advantages of the atlas-based segmentation methods is that the dense deformation field, interpolated on the whole image from the registration of visible image features, allows to easily estimate, in the target image, the position of structures with fuzzy or no visible contours. Moreover, this approach allows to segment at the same time several contours of any types (closed, open, connected or disconnected) provided that they have been defined in the atlas image. The accuracy of the segmentation results depend on the segmentation of the reference image and mainly on the quality of the registration between the atlas and the target image. A wide range of image registration techniques that allow to deform a given atlas to a subject have been developed over the last 20 years [16–18, 20–22, 50–57]. To compute the deformation field, they generally optimize some global similarity measure (such as mutual information, sum of squared

differences or cross-correlation) coming from the atlas intensity image.

The main limitation of the image registration methods commonly used for atlas registration is that they often lead to a compromise between the accuracy of the registration and the smoothness of the deformation. When at some places the registration is not accurate enough, a widely used solution is to globally or locally allow more variability in the registration model in order to obtain more local deformation, but with the risk of creating irregularities in the deformation field [58]. Also, this does not assure that the desired level of precision will be obtained. Another limitation of these methods is that they assume that a point to point correspondence exists between the atlas and the images to be segmented. Such assumption can lead to an inaccurate registration, particularly, in the presence of content-based inconsistencies between the atlas and the target image[1].

On the other hand, active contour (AC)-based methods are well known in computer vision community, for directly segmenting the structures of interest in a given image. The active contour models detect the closest contour(s) from an initial position, based on certain optimization criteria. For example, the active contour model (or snake model) that was first introduced by Kass et. al. in [59] locates the sharp intensity edges by deforming the initial contour towards the edges of the object. Please refer to [60] for a comprehensive review of the developments that took place in AC-based segmentation methods.

We now present a brief review of previous works that attempted to combine registration and active contour (AC) segmentation. Concerning the algorithms deduced from the level set evolution equation, we cite as the first reference, the "morphing AC model" proposed by Bertalmío et al. in [61], despite the fact that this method combines morphing instead of registration to AC segmentation. This model deforms the moving image (i.e., atlas) to the target image as in a registration process but the corresponding geometric transformation is not determined explicitly. The goal is to identify in the target image, a contour that corresponds to the object segmented in the reference image. This is done by solving a system of two partial differential equations (PDEs). The first one is in charge of morphing the two images. The second one is a tracking equation that makes the level set function evolve with the same velocity given by the morphing equation. Vemuri et al. in [2] have proposed an algorithm very close to the "morphing AC model" of Bertalmío [61], but dedicated for atlas registration. The main difference is that the image matching is not anymore tracked by a level set function but by a deformation field and that the morphing and the tracking PDEs are combined in a single PDE. To get the segmentation of a particular object in the target image, the computed deformation is applied to the segmentation of the corresponding object in the moving image following the principles of the atlas-based segmentation method.

---

[1]The term *content-based inconsistencies* in this thesis refers to the presence of certain structures (for example, a tumor), either in the atlas, or in the target image, but not in both.

The first attempt to model the atlas-based segmentation directly from the energy equations is made by Yezzi et al. in [1]. This model first defines two segmentation energies, one in the moving image and another one in the target image, that aim to segment the same object in both the images. Then both energies are coupled by defining the active contour of the target image as being the active contour of the moving image under a particular deformation. The advantage of this method is that it can combine multi-modal information since the segmentation energies are defined independently on both the images.

Inspired by the Yezzi's model [1], several authors have proposed other joint registration and AC segmentation models. [62] have shown that Yezzi's model is very sensitive to the initial active contour position and deformation. To avoid local minima, the active contours of both images have to be initially already well superposed to the objects of interest. Thus they have proposed to define the active contour of both images (not only of the target image as in the Yezzi's model) as being the initial curve under a particular deformation. They have also introduced a pre-registration step to find out the best initial registration parameters that compensate in both images the difference of positions between this initial curve and the objects of interest. [63] have generalized the rigid framework of Yezzi to non-rigid registration. [64] have further included a prior segmentation term in their model that computes the intensity difference between the prior shape and the shape to be segmented by the active contour in the target image. This term in fact corresponds to the matching image term used in the Bertalmío and Vemuri's models with the difference that it is computed only inside the prior shape. [65] have proposed to refine in a second step, the segmentation result obtained by the Yezzi's model with the "morphing AC model" of Bertalmío. Finally, [66] have presented two joint registration and AC segmentation models. The first model is similar to Yezzi's model, but optimizes a nonrigid deformation that is then propagated from the contour to the whole image. Their second model did not use anymore a level set function to represent the contour. It is based on an energy that optimizes the detection of common contours between the atlas and target images. Thus close, open, connected and disconnected contours can be considered for the registration. However, this model is limited to contour-based registration, and cannot exploit region-based features.

Other types of models close to this joint registration and active contour segmentation framework have been proposed. First, there are algorithms derived from joint optical flow registration and active contour segmentation that aim to generate a dense but "discontinuous" deformation field on the whole image for the study of motion in image sequences [67, 68]. Also [69] have extracted the non-rigid deformation between 2D geometric shapes by representing them with a level set function. This last model extracts the deformation only on the active contours.

The main contribution of the new registration framework that we will be proposing in the next chapter is that it combines the advantages of the models described above and

overcomes some of their limitations. First, as in [64], we consider the moving image as a prior image (the atlas). The initial shapes of the active contours are given by the objects of interest that are manually delineated in this prior image. Then, as in [62], we perform an initial global registration step in order to compensate for the initial differences of position between the atlas and the target image. Also, as in [66], our algorithm computes a dense nonrigid deformation field on the whole image. Similar to Vemuri's model [2], our method has been inspired first by the general evolution equation of the level set function. One of the main differences is that in our scheme, we propose to model several connected and/or disconnected active contours with a new label function representation. This label function permits in particular to consider different segmentation/registration forces at different areas of the image. Inspired by [64, 65], we propose an approach to combine the object-based forces coming from the AC segmentation technique with the pixel-based forces used in Vemuri's model [2], AC segmentation models with shape prior [70], or in optical flow registration algorithms [71].

2.2 Multi Atlas Fusion

We have already presented a summary of atlas fusion methods in chapter 1.2. We now present, in more detail, some of those methods that will be reformulated later, in chapter 4 and chapter 5.

2.2.1 Majority Voting

"Majority voting" (MV) is perhaps the most simple fusion method. It assigns for each voxel a label that maximum number of atlases agree on.

Let N be the number of atlases. Let V be the number of voxels in the target image. Let Y_p denote the label assigned to the pth voxel in the output image. Let X^j represent the jth input labeled image (corresponding to the jth atlas) after applying the transformation that maps the jth atlas to the output intensity image. Let X_p^j be the label assigned to the pth voxel of X^j. Let D be the total number unique label values (including the background label) in the input labeled images. Let $L = \{l_1, \cdots, l_D\}$ represent the set of all possible labels to be assigned. The original formulation of majority voting [27] is a maximization problem of the following form.

$$Y_p = \max \left[\sum_{j=1}^{N} \delta \left(X_p^j, l_1 \right), \sum_{j=1}^{N} \delta \left(X_p^j, l_2 \right), \cdots, \sum_{j=1}^{N} \delta \left(X_p^j, l_D \right) \right], \tag{2.1}$$

where δ is a Kronecker delta function. It is easy to see from the above equation that, if "l_d" is the label that maximum number of atlases assign to a voxel p, then the associated term (i.e., $\sum_{j=1}^{N} \delta \left(X_p^j, l_d \right)$) will have the maximum value among all D terms, and thereby,

label l_d will be assigned to Y_p.

2.2.2 Global Weighted Voting

If an atlas is more similar to the target image to be segmented when compared to other atlases, it is more likely that the automated segmentations obtained from that particular atlas are relatively more reliable and accurate. Majority voting method, while merging segmentation results obtained from different atlases, however does not take this information into account; it simply gives equal weight to all decisions coming from various atlases. Unlike majority voting, "Global Weighted Voting" (GWV) attaches a weight to each atlas while counting its vote. The weight for each atlas is determined *globally*; the more the similarity, the higher the weight, and vice versa. The original formulation of GWV [27] is as follows.

$$Y_p = \max \left[\sum_{j=1}^{N} w^j \delta \left(X_p^j, l_1 \right), \sum_{j=1}^{N} w^j \delta \left(X_p^j, l_2 \right), \cdots, \sum_{j=1}^{N} w^j \delta \left(X_p^j, l_D \right) \right], \qquad (2.2)$$

where w^j is the global weight assigned to the decisions made by the jth atlas. As just mentioned, the only difference between the MV formulation in equation (2.1), and the GWV formulation in equation (2.2) is the inclusion of this global weighting term. Thus, the interpretation of GWV maximization formulation can be done in similar lines to that of MV which has been already explained in the preceding subsection.

The global weight for each atlas can be computed using various similarity metrics. For instance, the similarity metric can be based on mean square difference (MSD) of intensities, or normalized cross correlation, or mutual information between the transformed atlas and the output intensity image. The metric needs to be selected appropriately depending on the nature of the atlas and the target images. For example, for the application of lymph nodes segmentation in the H&N CT images that we consider in this thesis, since both the atlas and the target images are expected to have similar intensity distributions, we have chosen MSD as the similarity metric; the lower the total MSD value between the transformed atlas and target intensity images, the higher the similarity, and vice versa. Hence, the weights for each atlas are computed as follows.

$$w^j = \frac{1}{\sum_{p=1}^{V} \left(I_p^j - I_p^T \right)^2 + \epsilon}, \qquad (2.3)$$

where I_p^j is the intensity of the pth voxel in the jth transformed atlas, I_p^T is the intensity of the pth voxel in the target intensity image, and ϵ is a small positive constant used for avoiding singularity that could occur when both the transformed atlas and the target images are exactly identical at all voxels.

2.2.3 Local Weighted Voting

As mentioned in the preceding subsection, GWV assigns a single global weight for each atlas as a measure of the similarity between the atlas and the target images. However, although two images may differ significantly in some particular regions, it is possible that they may be very similar in some other regions, and "Local Weighted Voting" (LWV) is indeed based on this observation. Instead of computing a single similarity measure for the entire image (like in GWV), LWV method computes the similarity measure *locally* for each voxel, within a specified neighborhood. In other words, LWV is similar to GWV except that, not a single global weight is assigned to the entire atlas; rather, for each voxel, an individual weight is assigned based on the *local* similarity. Let w_p^j represent the weight assigned to the jth atlas, at p^{th} voxel. Then the original formulation of LWV [27] can be expressed as follows.

$$Y_p = \max \left[\sum_{j=1}^{N} w_p^j \, \delta \left(X_p^j, l_1 \right), \sum_{j=1}^{N} w_p^j \, \delta \left(X_p^j, l_2 \right), \cdots, \sum_{j=1}^{N} w_p^j \, \delta \left(X_p^j, l_D \right) \right]. \tag{2.4}$$

Notice that the formulation for GWV presented in equation (2.2), and the above formulation for LWV are very similar except that the weight assigned for each atlas is not fixed anymore for the entire image; rather, the weights are now dependent on the voxel location.

The weights for each atlas at each voxel can be computed similar to GWV, using any of the similarity metrics like mean square difference (MSD) of intensities, normalized cross correlation, or mutual information, and the selection of appropriate metric depends on the nature of atlas and target images. We selected the MSD metric for the H&N lymph nodes segmentation application that we present in this thesis. Let $\aleph_p$ be the set of all voxels in the predefined neighborhood range of the pth voxel (with the set $\aleph_p$ including the pth voxel also). Then the MSD metric-based weight assigned to the jth atlas, at p^{th} voxel is given by

$$w_p^j = \frac{1}{\sum\limits_{\forall q \in \aleph_p} \left(I_q^j - I_q^T \right)^2 + \epsilon}. \tag{2.5}$$

2.2.4 Shape Based Averaging

Notice that the fusion methods that have been described so far in the preceding subsections are essentially *voting-based* methods since they make labeling decisions at each voxel by looking at the labels assigned by each atlas at that voxel; they differ from each other only in the way they weigh those votes (or decisions) coming from different atlases. In contrast to the above mentioned methods, "shape based Averaging" (SBA) method [41] looks at the fusion problem from a different perspective. Just to give an

intuitive interpretation, SBA first computes in all the transformed atlas images, how "deep" (inside), or how "faraway" (outside) a given voxel is from the contours of each possible label; this is tracked with the help of "signed Euclidean distance" (SED) measure; it then selects for each voxel a label that has the least resultant SED at that voxel, when summed up over all atlases. Please refer to [41] for a nice pictorial illustration of this idea, and we now proceed to its mathematical formulation.

Let $u_p^j(l)$ represent the SED for label l, at p^{th} voxel, in X^j. In the original formulation of SBA [41], for each voxel p, it assigns independently, that label which results in the minimum value of the following summation.

$$Y_p = \arg\min_l \frac{1}{N} \sum_{j=1}^{N} u_p^j(l). \tag{2.6}$$

2.3   Head and Neck Lymph Nodes Segmentation

"Automated segmentation of lymph nodes in the 3D H&N CT images" is the main application that we have considered throughout this thesis for evaluating our methodologies. It is because this application is both clinically significant as well as technically challenging. In order to make the importance of this application more clear to the reader, this section presents its clinical context, and it is followed by a review of the segmentation approaches that have been used so far in the literature for performing lymph nodes segmentation.

Intensity modulated radiotherapy (IMRT) is a very high precision technique to accurately deliver X-ray radiation treatment for different tumor locations of the patients. However, one of the significant obstacles in the widespread implementation of IMRT, for head and neck (H&N) cancer, concerns the complexity of target definition. In the case of H&N carcinomas radiotherapy, besides the gross tumor volume, the radiation oncologist has to segment the clinical target volume and the complicated planning target volume which contains different lymph node levels. Each lymph node level or group of levels correspond to a potential area of spread for a given tumor sub-location. Since the IMRT approach prerequisites the segmentation of all the volumes to be treated as well as the organs at risk, it is easy to understand that its routine use for H&N tumors is not common at the present time. Besides the precise contouring of primary H&N tumors that is often difficult, the accurate, reproducible and time-efficient contouring of elective nodal risk regions represents an even greater challenge.

The tentative to implement lymph nodes levels segmentation in the clinical environment were initially based on the translation of the surgical lymph nodes levels to CT-based regions which meant meticulous segmentation of each CT regions on each slice of the planning CT scan, a laborious process that was considered as incompatible

(a) Axial slice   (b) Sagittal slice   (c) Coronal slice

(d) Front view   (e) Side view   (f) Back view
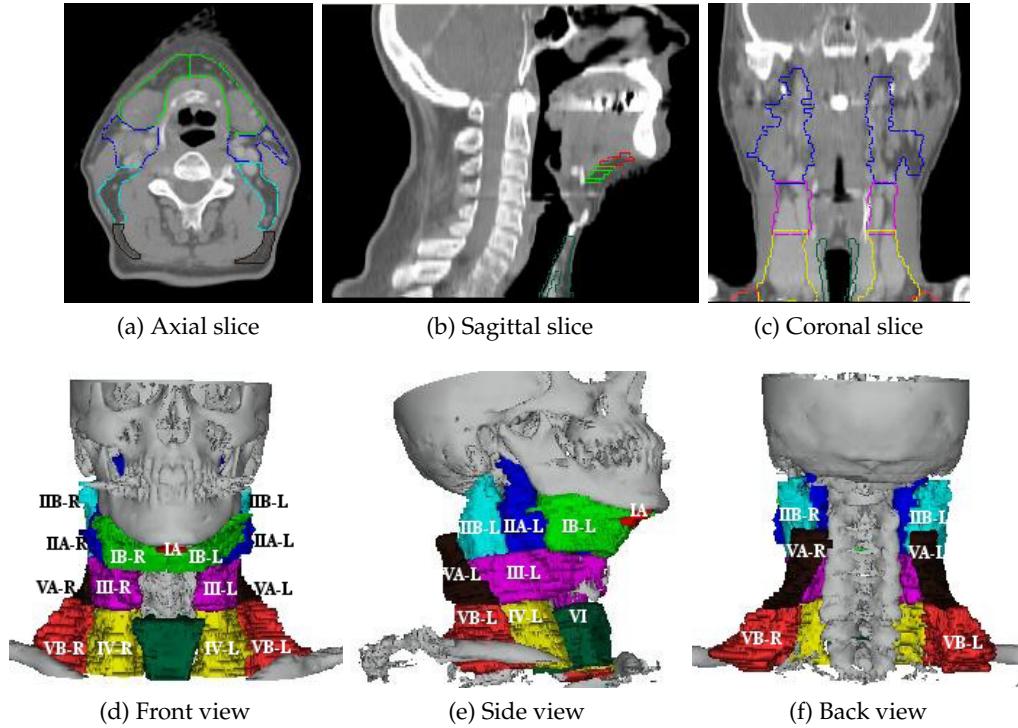
Figure 2.1: Manually delineated lymph nodes in a 3D CT image. The top row shows the contours of the lymph nodes in one of the axial, sagittal, and coronal slices. The bottom row shows the lymph node volumes on a thresholded CT volume that contains only bones.

with a routine clinical practice [72]. Indeed, experienced H&N cancer specialists generally spend several hours to fully contour and refine desired targets for a single H&N IMRT case. In a study reported by Song et al. [73], the average physician working time to design a H&N treatment contours for the target definition was 2.7 hours for IMRT approach compared to 0.3 hours for the conventional three field plan. In summary, the major challenge in the routine clinical implementation of IMRT for H&N region is to delineate the lymph nodes automatically and accurately.

Grégoire et al. [74] presented guidelines for delineating the lymph nodes in the H&N region. Figure 2.1 shows lymph node levels: IA, IB-left, IB-right, IIA-left, IIA-right, IIB-left, IIB-right, III-left, III-right, IV-left, IV-right, VA-left, VA-right, VB-left, VB-right and VI on the CT images. The top row shows the manually delineated contours of the lymph nodes in one of the axial, sagittal and coronal slices. The bottom row shows the lymph node volumes on a thresholded CT volume containing only bones. It is easy to notice from this figure that these lymph nodes do not possess distinct boundaries with respect to the surrounding structures; rather they are defined with respect to other distinct landmark structures in the H&N region and hence posing challenges in the automated segmentation.

16

The lymph node segmentation techniques that have been reported so far can be broadly classified into two categories. The first category of techniques assume that at least a portion of the lymph node to be segmented has a distinct boundary with respect to the surrounding structures. The second category of techniques do not assume the existence of any such distinct boundaries.

In the first category, Rogowska et al. [75] used various basic techniques like threshold selection, sobel/watershed technique and deformable contour algorithm for the segmentation of lymph nodes. Their evaluation was on synthetic images. In [76], Honea et al. semi-automatically segmented the lymph nodes with slice-wise active contours and active surface models. Their evaluation was also on synthetic images. In [77], Yan et al. proposed an improved 2D fast marching method with an intensity weighted speed term. [78] extended it to 3D images using a similar idea with watershed transform. In both [77, 78], boundary leaking is avoided through a hard stopping criterion, by manually bounding a circle around the lymph node; the circle should be very close to the actual boundary of the lymph node. In [79], Dornheim et al. proposed a mass-spring model for 3D lymph node segmentation, and the formulation is based on the characteristic gray value range, directed contour information and shape knowledge

The second category of techniques are atlas-based segmentation methods. The majority of atlas-based methods that have been tried so far were actually developed in the context of brain segmentation. Unfortunately, those methods are of limited use for lymph nodes segmentation due to the presence of high anatomical variability, particularly in the nodal regions. To cope up with this problem, Teng et al. [80] proposed to use the BSpline algorithm of Mattes et al. [81] along with the landmark points. The landmark points are selected on the mandible and hyoid bones, because of their proximity to the lymph nodes, high contrast and consistency among the patients. These landmarks are used to initialize the deformation field at the start of every resolution level of the BSpline algorithm. The authors show that local constraints on the atlas registration lead to a more accurate delineation of the lymph nodes. The main drawback of this method is that it requires a pre-processing step including the segmentation of the bones of interest in the patient image, the extraction of the landmark points on the bones surfaces and the computation of landmarks correspondence between the atlas and the patient image. In [6], Commowick et al. try to better catch the high variability of the H&N region by using a mean atlas built from a database of 45 patients. The atlas is then registered to the patient image by using a block matching algorithm [82, 83]. However, this method shows limitations when the patient anatomy is too different from the mean atlas or when the structures of the patient image are drastically deformed by a lesion.

Finally, we conclude this section by briefly pointing out the contributions and evaluations made in this thesis in the specific context of H&N lymph nodes segmentation application. The registration framework proposed in the next chapter has facilitated selecting structures with distinct image features to drive the registration process. It

has also facilitated using multiple registration forces in an hierarchical manner. In the same chapter, we compare the lymph nodes segmentation results obtained from our framework with the results from two other widely used registration methods (viz, RBF and Demons algorithms). From chapter 4 onwards, instead of using single atlas-based approaches, we switch our focus to multi atlas-based methods since they have resulted in more accurate segmentations. In chapter 4, we evaluate various fusion methods, both with and without the MRF-based smoothness term (that we proposed in the same chapter), for performing lymph nodes segmentation. In chapter 6, we extend our evaluation to two new SBA-based fusion methods (that we proposed in chapter 5). In the same chapter, we also evaluate the effect of our MRF-based neighborhood model in improving the accuracy of lymph nodes segmentation.

2.4 Evaluation Metrics

In this section, we put together the details of various statistical and geometric measures that are used in this thesis for evaluating the accuracy of automated lymph nodes segmentations.

(i) True Positive (TP)

For a lymph node under consideration, an voxel is treated as a "True Positive" (TP) voxel if it is present in the lymph node volumes of both the gold standard and the automated segmentation result.

(ii) True Negative (TN)

For a lymph node under consideration, an voxel is treated as a "True Negative" (TN) voxel if it is not present in the lymph node volumes of both the gold standard and the automated segmentation result.

(iii) False Positive (FP)

For a lymph node under consideration, an voxel is treated as a "False Positive" (FP) voxel if it is present in the automated segmentation result of the lymph node volume, but is not present in the gold standard.

(iv) False Negative (FN)

For a lymph node under consideration, an voxel is treated as a "False Negative" (FN) voxel if it is present in the gold standard of the lymph node volume, but is not present

in the automated segmentation result.

(v) Sensitivity

Sensitivity is a statistical measure of "true positive fraction," and is defined as follows:

$$\text{Sensitivity} = \frac{N_{TP}}{N_{TP} + N_{FN}},$$

where N_X represents the number of voxels that belong to the category specified by the subscript X. Sensitivity is thus a measure of under-segmentation, and is not sensitive to over-segmentation. Notice that higher sensitivity values indicate better segmentation results, and the sensitivity value for an ideal automated segmentation is 1.0.

(vi) Specificity

Specificity is a statistical measure of "true negative fraction," and is defined as follows:

$$\text{Specificity} = \frac{N_{TN}}{N_{TN} + N_{FP}}.$$

Specificity is thus a measure of over-segmentation, and is not sensitive to under-segmentation. Hence, sensitivity and specificity metrics are complementary to each other. Notice that higher specificity values indicate better segmentation results, and the specificity value for an ideal automated segmentation is 1.0.

The specificity metric is generally not considered in our evaluations of lymph node segmentations as these values are almost 1 for all methods for the following reason. The major issue with specificity is their dependence on the relation between the size of image and object under investigation [84]. Since the sizes of many lymph nodes are small compared to the size of the image, N_{TN} value is very large when compared to N_{FP}, and hence, the specificity values are close to the ideal value ($= 1$) for all the lymph nodes.

(vii) Dice Similarity Metric (DSM)

Dice Similarity Metric (DSM) is a statistical measure of spatial overlap between the gold standard and the automated segmentation. DSM is defined as follows [2]:

$$\text{DSM} = \frac{2\,N_{TP}}{2\,N_{TP} + N_{FP} + N_{FN}}.$$

---

[2]Note that the equation that we use here for computing DSM is exactly equivalent to the one used in some other papers like [4, 85], and it only differs in the notation or formulation followed.

The ideal value of DSM is 1. It is easy to notice that, unlike the preceding two statistical measures (i.e., sensitivity and specificity measures), DSM is sensitive to both over-segmentation and under-segmentation. Thus, it is relatively a more reliable measure for quantitative evaluation. DSM is also often referred to as "Dice Similarity Coefficient" (DSC).

(viii) Hausdorff Distance (HD)

Hausdorff distance is a measure of resemblance between two sets of data that are superposed on one another [86]. Let the sets $A = \{a_1, \ldots a_m\}$ and $B = \{b_1, \ldots b_n\}$ respectively represent the points of a lymph node in the gold standard and automated segmentations. The Hausdorff distance $H(A, B)$ is defined as

$$H(A, B) = \max(h(A, B), h(B, A)),$$

where

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|(a - b)\|,$$

and $\|.\|$ is the L_2 norm between the points of A and B. The Hausdorff distance $H(A, B)$ is the maximum of $h(A, B)$ and $h(B, A)$. Thus it measures the degree of mismatch between the gold standard and the automatically segmented lymph node regions, by measuring the distance of the points of gold standard that is farthest from any point of automatically segmented lymph node regions and vice versa. Lesser the Hausdorff distance, better the resemblance between the gold standard and the auto segmentation.

(ix) Mean Contour Distance (MCD)

Mean Contour distance (MCD) computes the mean distance between the boundaries of nonzero regions of a given lymph node in the gold standard and automated segmentation results. Hence, intuitively, the difference between HD and MCD is that while HD computes the *maximum distance* between the two contours, MCD computes the *mean distance* between the two contours. The mean contour distance $M(A, B)$ is therefore defined as

$$M(A, B) = \max(m(A, B), m(B, A)),$$

where

$$m(A, B) = \operatorname*{mean}_{a \in A} \min_{b \in B} \|(a - b)\|.$$

Please note that MCD metric is some times also referred to as "Mean Hausdorff Distance" (MHD)

(x) Number of Connected Regions per Label

The output segmentations of each lymph node (label) should ideally contain a single contiguous region. Hence, in certain experiments, we evaluate the fusion algorithms also based on the number of connected regions it has created per each label (ideal value = 1); we count both islands and holes when computing the number of connected regions.

(xi) Wilcoxon Signed-Rank Test

In order to evaluate the statistical significance of difference between the results obtained from two different methods, we use Wilcoxon signed-rank test [87]. This is a non-parametric test procedure for the analysis of matched-paired data. Note that Wilcoxon signed-rank test can be seen as non-parametric alternative to paired-student test since it does not make any assumptions regarding the distributions of the data population.

2.5 Discussion and Conclusions

In this chapter, we have presented a review of the relevant state-of-the-art registration methods and multi-atlas fusion methods. We have also presented a description of various quantitative metrics that will be used in the later chapters for evaluating the accuracy of automated segmentations. In the next chapter, we present our contributions related to the variational formulation of image registration problem.

3 Active Deformation Fields

In this chapter, we present a new variational framework for atlas registration, and is called "active deformation fields". As the name suggests, it integrates both the *active* contour framework, and the dense *deformation fields* of optical flow framework. We also present a new label function representation that can represent multiple labels using a single level set function, and it enables to perform the registration based on only the selected consistent structures of interest in atlas and target images. We then illustrate how this framework facilitates using multiple registration forces in an hierarchical manner. Finally, we compare the segmentation results obtained from the proposed framework with the results from two other widely used registration methods [1].

3.1   Introduction

First of all, please refer to chapter 2.1 for a review of relevant state-of-the-art registration methods. The main source of inspiration for our joint registration and segmentation algorithm is the partial differential equation (PDE)-based model proposed by Vemuri et al. in [2]. The formulation of Vemuri's model has been intuitively deduced from the following general level set evolution equation introduced by [88]:

$$\frac{\partial \phi_D(x,t)}{\partial t} = \nu(\phi_D(x,t))|\nabla \phi_D(x,t)|, \qquad (3.1)$$

where ν is the velocity of the flow or speed function that contains the local segmentation and contour regularization constraints, and $\phi_D : \Omega \to \mathbb{R}$ (for the image domain Ω) is the signed distance function often used to represent implicitly the active contour by its zero level. The original idea brought by Vemuri's model is to replace in (3.1), ϕ_D by the intensity function of the image to register $\phi_I : \Omega \to \mathbb{R}$ (the moving image). This leads to

---

[1]The work presented in this chapter is a joint work with Valérie Duay; both Valérie and myself contributed equally to this work.

the following equation:

$$\frac{\partial \phi_I(x,t)}{\partial t} = \nu(\phi_I(x,t))|\nabla G_\sigma * \phi_I(x,t)|, \tag{3.2}$$

where G_σ is a Gaussian kernel with standard deviation of σ, and $*$ is the convolution operator. The image ϕ_I is prior smoothed with a Gaussian filter because the gradient computation is very sensitive to noise. The level sets considered in the segmentation process correspond to the contours naturally present in the moving image, i.e., the curves of high image gradient. The speed function used in both Bertalmío's model [61] and Vemuri's model is the intensity difference term $\nu(\phi_I(x,t)) = (\phi_I(x,t) - \phi_T(x))$, where $\phi_T(x)$ is the intensity function of the target image.

Note that equation (3.2) gives only the intensity evolution, but not the geometric transformation between the images. So, in order to explicitly track the deformation field, they have intuitively derived one more equation presented below. The deformation field vector $u : \mathbb{R}^n \to \mathbb{R}^n$ (typically, $n = \{2, 3\}$) is given by:

$$\frac{\partial u(x,t)}{\partial t} = (\phi_I(x,t) - \phi_T(x)) \frac{\nabla G_\sigma * \phi_I(x,t)}{|\nabla G_\sigma * \phi_I(x,t)|_\epsilon}, \tag{3.3}$$

where $|\Gamma|_\epsilon = \sqrt{|\Gamma|^2 + \epsilon^2}$, and ϵ is a small positive constant used for avoiding the singularity.

The intensity function ϕ_I at time t is given by the deformation field $u(x,t)$ and the initial intensity function $\phi_I(x,0)$ such that:

$$\phi_I(x,t) = \phi_I(x + u(x,t), 0), \tag{3.4}$$

which ensures that the evolution of the intensity function exactly corresponds to the current deformation.

The deformation forces of Vemuri's model are pixel-based forces computed on the whole image. This has two main consequences. First, this algorithm presents some limitations when it has to recover global differences because the registration forces that it uses are very local. Secondly, it fails to register regions when texture is different with similar intensity distributions. This situation could, for example, arise in natural images [89] in which two regions (object and background in this example) have similar intensity distributions, but different textures. This can happen in medical images also, particularly, in the segmentation of ultrasound images [90]. Finally, Vemuri's model has no scheme to prevent the registration of inconsistent regions.

3.2 Our Active Contour-based Atlas Registration Framework

3.2.1 Our Model: Active Deformation Fields

The general formulation of our model is based on tracking of the level set function motion based on the conservation of morphological (i.e., shape) description. Since level sets are basically a morphological notion, the conservation of morphological description essentially states that level set function is preserved under small deformation field and a short period of time. Let "ϕ_G" represent a general level set function; it means that wherever ϕ_G is used, it is applicable to different types of level set functions like signed distance function (ϕ_D), intensity function (ϕ_I), or the label function (ϕ_L) that will be introduced in the next section. Now, the above mentioned conservation implies that ϕ_G remains constant for small displacements and a short period of time. This results in the following equation:

$$\phi_G(x,t) = \phi_G(x+du, t+dt) \Rightarrow d\phi_G(x,t) = 0, \tag{3.5}$$

where $d\phi_G$ is the total derivative of ϕ_G. By using the chain rule, this constraint can be rewritten as the evolution equation of a vector flow:

$$\frac{\partial u(x,t)}{\partial t} = -\frac{\phi_{G,t}}{|\nabla \phi_G|} N_{\phi_G}, \tag{3.6}$$

where $\phi_{G,t}$ is given by (3.1) (with a change of subscript), and it represents the variation of the level set function according to the desired forces such as supervised segmentation, shape prior knowledge or contour regularization; N_{ϕ_G} in the above equation represents the normal of level set, and is given by:

$$N_{\phi_G} = \begin{cases} \frac{\nabla \phi_G}{|\nabla \phi_G|}, & \text{if } \nabla \phi_G \neq 0; \\ 0, & \text{otherwise.} \end{cases} \tag{3.7}$$

Note that the widely used morphological descriptor in registration problems is the "luminance function". With the luminance function, the above mentioned conservation is equivalent to optical flow (OF) assumption, which states that the brightness of the moving image stays constant for small displacements, over short periods of time. In that sense, in our discussions, we will be referring to "conservation of morphological description" simply as "OF approach".

By introducing the evolution equation of the level set segmentation model (3.1) in (3.6), we obtain the following formula *merging the active contour segmentation framework*

with the image registration task:

$$\frac{\partial u(x,t)}{\partial t} = -v(\phi_G(x,t)) \, N_{\phi_G}. \tag{3.8}$$

The level set function ϕ_G does not evolve with the usual finite difference scheme. Its position at time t is given by the deformation field $u(x,t)$ and the initial level set function $\phi_G(x,0)$ such that:

$$\phi_G(x,t) := \phi_G(x + u(x,t),0), \tag{3.9}$$

where $\phi_G(x,0)$ is the initial active contour position. This ensures that the evolution of the level set function exactly corresponds to the current deformation. Notice that when intensity function is used as the level set function (i.e., when $\phi_G = \phi_I$), the general evolution equation 3.8 is indeed equivalent to gradient flows, where each voxel warps along its intensity-gradient, and its corresponding deformation field is taken into account by 3.9. Introducing (3.9) in (3.8) yields:

$$\frac{\partial u(x,t)}{\partial t} = -v(\phi_G(x + u(x,t),0)) \, N_{\phi_G}. \tag{3.10}$$

This equation defines a displacement vector at each point of the level set function. The level set function models the contours of the objects selected in the atlas to drive its registration.

As mentioned earlier, different types of level set functions can be used in place of ϕ_G in the above formulation. For instance, signed distance function (ϕ_D) representation can be used; however, the main drawback of ϕ_D representation is that it can model only two regions. Multi-phase representations have been proposed by combining several signed distance function (see for instance [91, 92]), but such solution will drastically increase the computational complexity of this model. On the other hand, while the representations like intensity function (ϕ_I) can model any type of contours, they are limited to only pixel-based registration forces (see section 3.3 for more details on ϕ_D and ϕ_I).

To cope with these limitations, we propose a new label function representation of the level sets that can model any number of regions as well as represent various types of registration forces, using a single function. This new multi-phase active contour representation is described in detail in the following section [2].
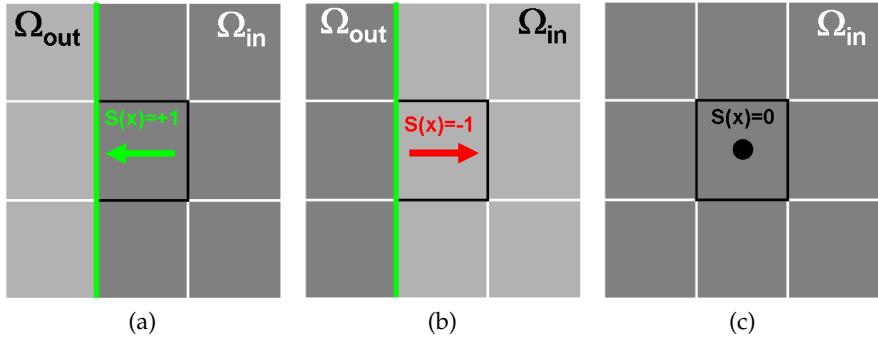
Figure 3.1: Illustration of the proposed sign function $S(\phi_L(x))$: (a) $S(\phi_L(x))$ is taken as 1 when the gradient direction of the label function coincides with the direction of level set evolution. (b) $S(\phi_L(x))$ is taken as -1 when the level set evolution direction has to be opposite to that of the gradient direction of the label function. (c) $S(\phi_L(x))$ is 0 when there is no gradient.

3.2.2 New Multi-Phase Active Contours Representation

For the registration of multiple regions, we propose a new label function ϕ_L to represent the active contours selected in the atlas. In the ϕ_L representation, the contours do not anymore correspond to the zero level set, but to the discontinuities of piecewise constant level set function. This label function permits to define an arbitrary number of regions as follows:

$$\phi_L : x \in \Omega_k \rightarrow \phi_L(x) = k, k \in [1, .., n] \text{ if } x \in \Omega_k, \tag{3.11}$$

where Ω_k is the kth labeled region and n is the number of regions. In this representation, active contours are modeled by the *discontinuities* of ϕ_L.

The main advantage of this label function representation is that it can distinguish any number of regions with a single function. The implementation of this representation is an important challenge; when multiple regions are represented using ϕ_L, it does not contain the polarity information that is essential to compute the widely used region-based forces of the AC segmentation. Note that the term polarity in this context is defined as the information of direction that indicates the inside (Ω_{in}) and the outside (Ω_{out}) for each modeled region. The importance of polarity information will be more clear with the description of region-based forces in section 3.3.1. We now present in detail on how we handle this issue of polarity information.

In order to accurately represent the polarity information for each region, we introduce

---

[2]We note that in the AC segmentation framework, the idea of using labels to perform a multi-phase segmentation has been presented (see for instance [93]). The difference with our work is that this representation has been proposed previously for particular variational energy-models and we present a scheme for any type of PDE-based models.
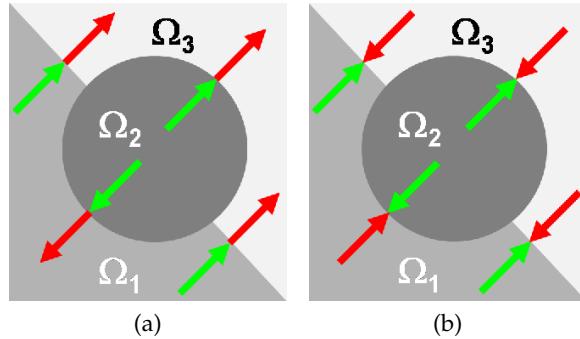
Figure 3.2: Example of a label function with three regions. (a) Orientations of the label function gradients. (b) Orientations of $(S(\phi_L(x)) \triangledown \phi_L(x))$, i.e., with the inclusion of polarity information. Arrows are marked in red when orientations of $(\triangledown \phi_L(x))$ and $(S(\phi_L(x)) \triangledown \phi_L(x))$ are in the opposite directions; arrows are marked in green when both the orientations are same.

in the general formulation of our model (3.10), a sign function: $S(\phi_L(x))$. The purpose of this sign function is to manipulate the overall direction of the evolution such that it is always oriented in the direction of Ω_{in} to Ω_{out}, independent of the orientation of the gradient of the label function $\phi_L(x)$ (which is always from the low intensity label to the high intensity label). In other words, irrespective of the direction of $\triangledown \phi_L(x)$, the direction of $(S(\phi_L(x)) \triangledown \phi_L)$ should always be from inside to outside of that region.

For an easy understanding, we explain the design of the function $S(\phi_L(x))$ using an example shown in figure 3.1. The green line enhances the interface between a high-intensity and a low-intensity label, each representing a different region. Each panel shows the current pixel (enhanced in bold) surrounded by its 8 neighbors. The arrows represent the direction of the gradients of the label function which are always from the low intensity label to the high intensity label. Now, if the neighbors of x have higher intensity label values than $\phi_L(x)$, it means that the desired level set evolution direction coincides with the gradient direction; so, $S(\phi_L(x))$ value is taken to be 1 (figure 3.1a). On the other hand, if the neighbors of x have lower intensity label values than $\phi_L(x)$, it means that the desired level set evolution direction is opposite to that of the gradient direction of the label function; so, $S(\phi_L(x))$ value is taken to be -1 (figure 3.1b). Finally, if all the neighbors of x have same label values as that of x, it means that the gradient value of the label function is zero, and there is no evolution of the level set for that pixel; so, $S(\phi_L(x))$ value is taken to be 0 (figure 3.1c). Thus, with the inclusion of this polarity information and the label function, the general formulation of our registration model (3.10) becomes:

$$\frac{\partial u(x,t)}{\partial t} = -S(\phi_L(x)) \, v(\phi_L(x + u(x,t), 0)) \, N_{\phi_L}. \tag{3.12}$$

To further illustrate about the polarity information, another example of a label function containing three regions (Ω_1, Ω_2, Ω_3) is presented in figure 3.2. The arrows in figure 3.2a show the orientations of the gradient $\bigtriangledown \phi_L$ at the interfaces of the label function. Figure 3.2b shows the orientations of ($S(\phi_L(x)) \bigtriangledown \phi_L(x)$). In both these figures, arrows in the regions where the polarity information and the gradient are in the opposite direction, are marked in red; if both are in the same direction, they are marked in green.

We now present other important implementation details of the label function. As mentioned earlier, ϕ_L value at time t and with the deformation field u is computed using equation (3.9). Note that this computation is done using a nearest neighbor interpolation so that values of the level set function remain fixed ($[1, .., n]$) during the registration process. Another point to be mentioned is, since ϕ_L is not a continuous function across its borders, we convolve it with a Gaussian kernel G_σ prior to the gradient computation. One of the possible alternatives to Gaussian kernel is using a regularized version of the Heaviside function as in [94]; however, [94] does not treat all level sets equally, giving more weight to the zero level set. Thus, image registration with the proposed multi-label level set representation is more robust if all the level sets are treated equally, which is done with the Gaussian smoothing approach. Further, to stabilize the numerical computation when $\bigtriangledown \phi_L$ is close to zero, equation (3.12) is modified as follows:

$$\frac{\partial u(x,t)}{\partial t} = -S(\phi_L(x)) \, v(\phi_L(x + u(x,t), 0)) \, \frac{\bigtriangledown (G_\sigma * \phi_L)}{|\bigtriangledown G_\sigma * \phi_L|_\epsilon}. \tag{3.13}$$

As mentioned earlier, ϵ is a small positive constant used for avoiding the singularity. In all our experiments, we set ϵ to $1e^{-6}$; we notice that as long as ϵ is small enough such as $\epsilon < 1e^{-4}$, there are no problems of sensitivity of the method to this parameter.

3.3 Driving Forces and Regularization Constraints

In this section, we first present various types of registration / segmentation forces that can be used in our registration framework. Secondly, we derive two of the existing atlas registration models (Yezzi's model [1] & Vemuri's model [2]) from our general framework, and compare them conceptually with the model that we have proposed in section 3.2.2. Finally, we present the smoothing and the bijectivity constraints that we include in our AC-based registration framework.

3.3.1 Registration/Segmentation Forces

One of the main contributions of our registration framework is that it facilitates using various types of AC-based as well as OF-based forces. Thus, this framework can be easily adapted to various specific applications. Figure 3.3 classifies the AC-based and OF-based forces according to their effect in a contour matching process. The most used
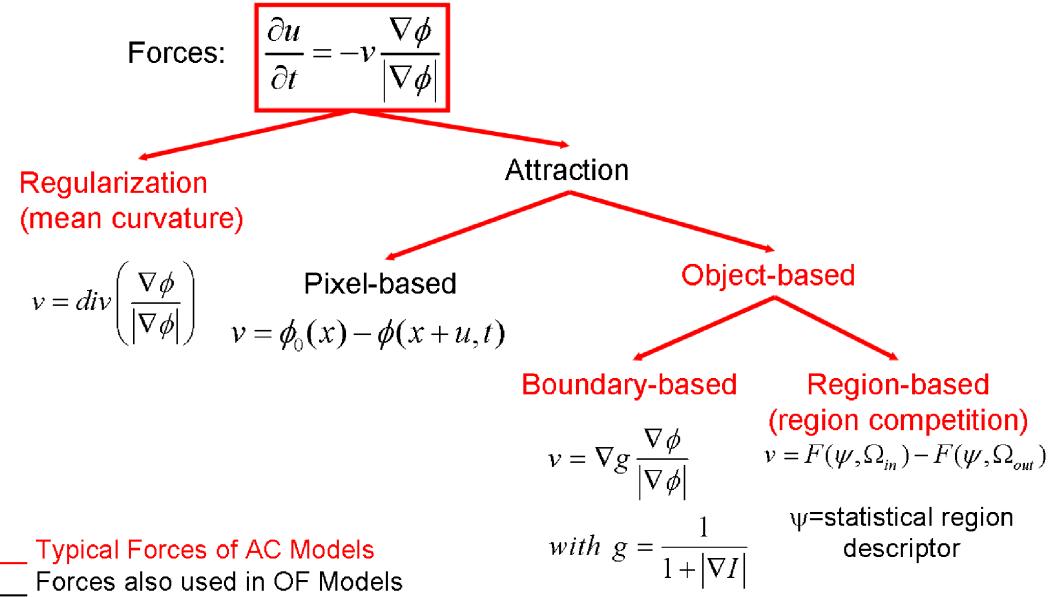
Figure 3.3: Classification of the AC-based and OF-based forces according to their effect in a contour matching process.

regularization force of the AC framework is the mean curvature force. This force smoothes the level sets by minimizing their length. They can be applied to any type of contour representation. The *pixel-based forces* are based on the smallest image feature, the pixel value. They allow the local registration of the whole moving image domain or selected regions. Pixel-based forces are the typical segmentation forces of the OF model. In AC model, these forces are rather used to include intensity or shape prior knowledge in a segmentation process. These forces can match any type of contours (closed or open) and can also be used with any type of representation. However, they are very sensitive to image noise and are limited to recover small deformations. The *object-based forces* can register image regions. Finally, *region-based forces* are very efficient forces of the AC framework because they are less sensitive to noise than the *boundary-based forces*. They can also perform supervised segmentation, i.e., they can use prior knowledge extracted from a reference image. In summary, the current framework allows to choose any of these forces depending on the specific application.

Region-based Registration Forces

For many of the applications presented in this thesis, we use a region-based force inspired by the unsupervised region-based segmentation model proposed by Chan and Vese [94]. The following speed function is used for this purpose:

$$ v = \int_{\Omega_{in}} |I(x) - \mu_{in}^{prior}|^2 dx - \int_{\Omega_{out}} |I(x) - \mu_{out}^{prior}|^2 dx, \tag{3.14} $$

where Ω_{in} is the image area inside the contour and Ω_{out} is the image area outside the contour, μ^{prior} is the prior mean of a given region extracted from a reference image (the atlas) and I is the intensity function of the image to be segmented. This force assumes that corresponding regions between the reference and the target images have similar means[3].

Note that even for multi-phase segmentation with more than 2 regions, the above energy equation remains same except that the parameters of the label function are interpreted in a slightly different manner; those parameters are not defined anymore with respect to the entire label function. Instead, they are defined for each distinct region of the label function (i.e., based on the value of $\phi_L(x)$).

With the label function representation and the above mean-based force, the equation of the derived AC-based atlas registration model is:

$$\frac{\partial u(x,t)}{\partial t} = -S(\phi_L(x)) \left(\int_{\Omega_{in}} |I(x) - \mu_{in}^{prior}|^2 dx - \int_{\Omega_{out}} |I(x) - \mu_{out}^{prior}|^2 dx \right) \frac{\bigtriangledown (G_\sigma * \phi_L)}{|\bigtriangledown G_\sigma * \phi_L|_\epsilon}.$$
(3.15)

The parameters and the numerical approach used for solving the above equation are same as those that are already presented for pixel-based registration (equation 3.13) except with the following additional important details: Since μ^{prior} does not evolve during the registration process, it is computed *only* once, on the atlas image, at the beginning of the registration. The sign function $S(\phi_L(x))$ is computed over 8-pixel neighborhood for 2D images, and over 27-voxel neighborhood for 3D images. The implementation details for $\phi_L(x)$ are already presented in section 3.2.2.

We would like to mention here regarding the sensitivity of registration towards manual delineations of labels. Notice that the above region-based registration approach requires delineation of selected structures only in the atlas, but not in the target image. It is assumed that the structures of interest are accurately delineated in the atlas. In case of any errors in those manual delineations, deformation field estimated from the region-based registration can get affected accordingly. However, thanks to the pixel-based forces, if they are used in the next level (whenever applicable), since they do not depend on manual delineations, they can indeed correct the previous errors as long as the effect of errors caused in the preceding step are within a retrievable local neighborhood range. The bijectivity constraint can also automatically compensate for small errors in manual segmentations.

Finally, note that although we presented here only the classical mean-based region-forces that use sum of squared intensity differences as the similarity measure for the

---

[3]Possible intensity differences between both images can be reduced in a pre-process step by histogram matching.

registration, this framework is not at all limited to only these forces; it can be easily adapted to various other types of metrics, as well as registration forces. For instance, it is indeed showed in [95] how marginal entropies and joint entropy can be used for driving the region-based registration. Thus, this framework is suitable for even multi-modal registration. It also allows to easily incorporate other measures (like mutual information and its variants) as well, and it could be probably extended to other statistical registration models like [96]. Regarding the incorporation of other types of registration forces, it is showed in [19] how shape-based forces can also be integrated into our framework.

3.3.2 Derived Atlas-based Registration Models

Different models can be derived by simply changing the type of active contour representation ϕ_G, and/or the speed term ν that is used to generate the segmentation / registration forces in equation (3.10). Note that the possible selections of ν are determined/limited by the selection of the level set function ϕ_G. We present here 3 important special cases of this generalized framework.

(i) If ϕ_D is used as the level set function in equation (3.10), it results in *Yezzi's model* [1]. With the ϕ_D representation, only the speed terms related to the object-based forces and the regularization forces can be used, but not the pixel-based forces. As mentioned earlier, ϕ_D representation can be used only for the two-phase segmentation.

(ii) If ϕ_I is used as the level set function, it results in the *Vemuri's model* [2]. With the ϕ_I representation, only the speed terms related to the pixel-based forces can be used; neither the object-based, nor the regularization forces can be used.

(iii) If the new ϕ_L representation proposed in section 3.2.2 is used, it results in the segmentation model presented in equation (3.13). With this new ϕ_L representation, speed terms related to any of the of the three forces can be used. Further, ϕ_L representation can perform multi-phase segmentation.

The above discussion has been summarized in table 3.1. Thus, it is clear that the proposed framework generalizes two completely diverse state-of-the-art atlas-based methods, and with our proposed ϕ_L representation, multi-phase segmentation can be performed besides benefitting from any of the registration forces.

3.3.3 Regularization Constraints

In order to constrain the output deformation field to be uniform, we use two types of regularization constraints: (i) smoothing constraints, (ii) bijectivity constraint; these are described in more detail below. Incorporation of these constraints is required because of the ill posed nature of the nonrigid transformation.

Table 3.1: Comparison of our registration model that uses ϕ_L representation, with Yezzi's model [1] and Vemuri's model [2].

	Yezzi's Model	Vemuri's Model	Our Model with ϕ_L
Level set function	ϕ_D	ϕ_I	ϕ_L
Regularization forces	✓	×	✓
Pixel-based forces	×	✓	✓
Object-based forces	✓	×	✓
Multi-phase Segmentation	×	✓	✓

Smoothing Constraints

Linear diffusion smoothing is the commonly used regularization constraint in non-parametric registration algorithms (see [2, 16]). It is applied at the end of each iteration on the current global deformation field. Its PDE corresponds to the well known heat equation:

$$\frac{\partial v(x,t)}{\partial t} = \Delta v(x,t),$$
$$v(x, t = 0) = u_*(x), \qquad (3.16)$$

where u_* is the solution of equation (3.10) at the point x, and Δ is the Laplacian operator. This technique permits not only to smooth the deformation field, but also to diffuse the contour deformation in a narrow-band around it. The fastest way to perform this diffusion is by filtering. The filter corresponding to the heat equation is the Gaussian filter.

As mentioned in section 3.3.1, mean curvature is commonly used in active contour segmentation models for smoothing. Opposed to the linear diffusion, mean curvature approach acts directly on the evolution equation. Its goal is to constrain the evolving contours of the moving image to stay smoothed during their evolution. Hence, in this thesis, we apply mean curvature smoothing while using region-based forces, and apply linear diffusion smoothing (through Gaussian filtering) while using pixel-based forces.

Bijectivity Constraint

In recent years, there are significant developments in the estimation of smooth and invertible transformations with the main goal being the preservation of the topology of objects. In the framework of registration containing large deformations, diffeomorphisms are powerful and mathematically elegant transformations that ensure a one-to-one smooth and continuous mapping with nonsingular Jacobian determinant; they guarantee both invertibility and preservation of topology. Earlier diffeomorphic registration algorithms

are based on differential equations-based modeling of "viscous fluid" [97,98]. Recent diffeomorphic algorithms [17,99–101] resulted in more accurate estimations, but still, most of these algorithms have a high computational cost.

In a small deformation framework, many registration approaches still use non-diffeomorphic transformations; these algorithms, unlike diffeomorphic methods, are based on simply adding non-parametric displacement fields to an identity transform. However, such approaches do not necessarily preserve the topology, and do not guarantee the bijectivity constraint *unless* some explicit schemes are implemented like in [16,102]. Note that our deformation field computation is similar to the "additive demons iterations scheme" of Thirion's approach [16] in terms of the way the deformation field is updated, and the type of iterative scheme. [16] computes the forward and the backward transformations independently, following the approach suggested by [103], and modifies these transformations at each iteration to maintain their compatibility.

The backward transformation required in Thirion's approach can be computed in different ways. The simplest method could be to use a second label function defined on the target image, and perform registration with the first image. However, creating a label function on the target image could be cumbersome when segmentation of the structures that are used in the labeled image cannot be obtained using simple segmentation procedures like thresholding. Hence, in order to avoid the segmentation of labeled structures in the target image, we use another method inspired by the Thirions's algorithm [16]. In this approach, the inverse transform is computed by diffusing the object of the target image through the contours of the source label image. This leads to the following equation:

$$\frac{\partial u(x,t)}{\partial t} = \nu(\phi_G(x + u(x,t),0))\, N_{\phi_G}. \tag{3.17}$$

Notice that the above equation is identical to equation (3.10) except that the label function is not deformed and the driving forces have an opposite sign because they do not have to attract the contour to the target image; rather, it has to diffuse the target object inside the contour. While the above approach, like in [16], is good enough for small deformation framework, it may not be an accurate method in case of large deformations. Recently, a diffeomorphic model for the demons algorithms has been proposed in [17]. We could also probably adapt a similar approach to our framework in order to deal with large deformations.

3.4 Hierarchical Approach to Registration Forces

As described in section 3.3, our proposed framework facilitates to use multiple registration forces coming from both OF and AC frameworks. Further, the multi-phase AC representation proposed in section 3.2.2 enables to select specific regions to be used for

each of those registration forces. To benefit fully from these features, we propose to use the registration / segmentation forces in an hierarchical manner, based on their relative global (or local) characteristics. For instance, we know that the pixel-based forces are more local forces compared to the region-based forces. Hence, logically, performing the registration first with the relatively global region-based forces, and then with the pixel-based forces could avoid converging to a local minimum. In addition, we point out that there are other possible types of hierarchical approaches that can be easily incorporated into our framework, like, hierarchy of structures to be used [104].

We now illustrate, through a 2D synthetic example and a real clinical application, the advantages of using multiple registration forces in an hierarchical manner. In particular, in both these examples, we compare three registration models: In the first case, we use our framework with only the pixel-based forces for performing the registration; thus, this is equivalent to using the Vemuri's model [2]. In the second case, we use our framework with only the region-based forces for performing the registration; thus, this is equivalent to using Yezzi's model [1]. In the final case, we consider both the region-based and pixel-based forces in an hierarchical manner. Notice that this hierarchical approach became possible because of the proposed framework. As mentioned earlier, since region-based forces are more global than the pixel-based forces, during the hierarchical registration, we first use the region-based forces, and then followed by the pixel-based forces.

3.4.1   2D Synthetic Example

Figure 3.4 shows the synthetic images used in this illustration. Figures 3.4a and 3.4b show the target and atlas images respectively. Labeled image of the atlas is shown in figure 3.4c, and is used during the region-based registration. This labeled image, also called as label function, defines four connected regions. Notice that three of these regions (the background, the left and right regions) are consistent between the atlas and the target image, i.e., they have similar mean intensities and texture; on the other hand, the last one (the central region) has a similar mean in both images but the orientation of the texture pattern is different. The target contours are overlayed in red onto the atlas and target images to visualize the initial differences. The arrows on figure 3.4b indicate that in the right region, non-corresponding atlas and target texture patterns having similar intensity distribution are initially superposed.

Figure 3.5a shows the results obtained from Vemuri's model ([2]). It can be noted that it cannot register correctly in the right and central regions due to the following reasons: (i) The algorithm fails to register global differences because the registration forces that it uses (pixel-based forces) are very local. (ii) Since it has no scheme to prevent the registration of inconsistent regions, it obviously tries to register even the inconsistent texture patterns in the central region.
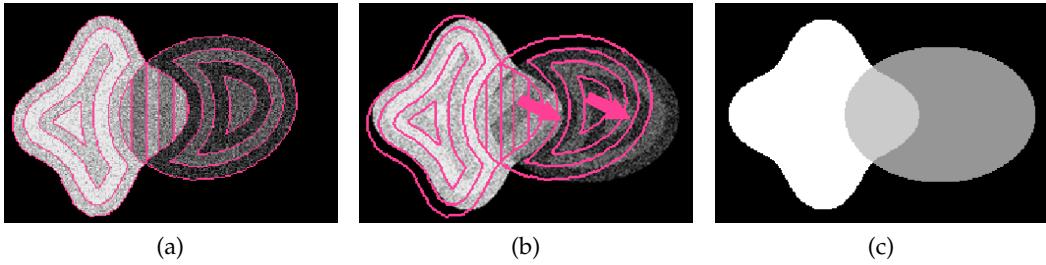
Figure 3.4: A 2D synthetic example for illustrating the advantages of hierarchical registration approach. Input data: (a) Target image. (b) Atlas image. (c) Labeled image corresponding to the atlas. Contours of the target image are superposed over the intensity images in order to highlight the initial differences.
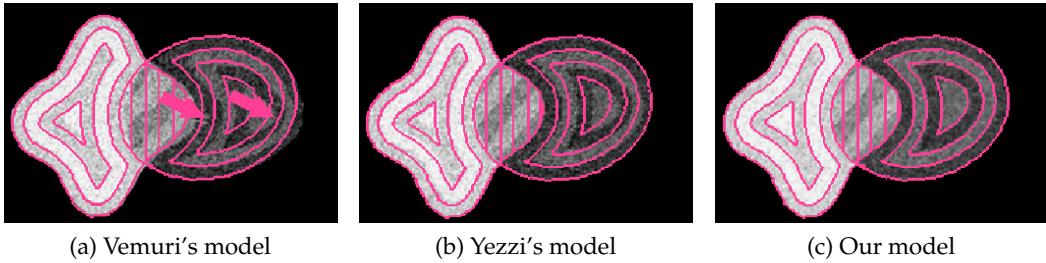
Figure 3.5: Registration results from different methods applied on the synthetic data in figure 3.4.

Figure 3.5b shows the results obtained from Yezzi's model ([1]) that uses mean-based forces to register the four regions marked by the labeled image. For this example, while the Yezzi's model has globally registered the regions better than the Vemuri's model, one can see the limitations of using region-based forces alone. Notice from the results that, since the deformation is based on only the contours of interest, it has more registration errors in the texture when we are far away from the registered contours. Moreover, we can model only closed contours with the Yezzi's model.

We finally perform the registration with our framework, using the hierarchical approach, and the results are shown in figure 3.5c. In this hierarchical approach, region-based registration is performed followed by the pixel-based registration. Since the texture of the central region is different between the images to be registered, this region is not considered during the pixel-based registration, and thanks to the label function to determine where to compute the pixel-based forces; we just let the texture pattern in the central region to follow the deformation of the region-based contours. We can see from the results on these simple synthetic 2D images that our combined region-based and pixel-based hierarchical model has the potential to give better results than the methods that use only pixel-based forces or, only region-based forces.

36

3.4.2 Segmentation of Head and Neck Lymph Nodes

As detailed in chapter 2.3, automated segmentation of lymph nodes in 3D CT images is a crucial and challenging step for intensity modulated radiotherapy (IMRT) treatment of H&N cancer. Figure 3.6 shows, for one of the patients' image, the manually delineated ground truth segmentations of lymph node levels: IB-Left, IB-Right, IIA-Left, IIA-Right, III-Left, III-Right, IV-Left, IV-Right, VA-Left, VA-Right, VB-Left, VB-Right, and VI. The ground truth segmentations are performed by a medical expert, under the supervision of a radiation oncologist.

The data set used for this application contains the H&N CT images of 13 patients, acquired at Divisions of Radiotherapy, Geneva University Hospital (HUG), during the routine clinical practice. The resolution of CT images in X and Y directions is varying in the range of $0.45\,mm$ to $0.94\,mm$; the typical resolution in the Z direction is $3\,mm$. The size of each axial slice is 512×512 pixels. The Field Of View (FOV) was not same for all the images, and hence, the images are cropped during the preprocessing to contain the same FOV. The number of slices for the data set after cropping is varying in the range of $61 - 73$.

The structures that have been selected for driving the region-based registration are: (i) external contours of the H&N, (ii) bones, and (iii) trachea, thus forming four distinct labels (including the background). Figure 3.7 shows these labels for one of the patients. Notice that all the above selected structures have unique intensity characteristics (i.e., mean values of intensities inside and outside the selected structures), and they also influence the location of the lymph nodes to be segmented. During the region-based registration, curvature-based regularization is also used on the resulting contour. Gaussian smoothing with a sigma of $1.5mm$ is used in the region-based registration whereas a sigma of $3.5mm$ is used in the pixel-based registration.

There can be huge anatomical variations in the H&N region among different patients. In addition to the registration method used, similarity / closeness of the atlas to the patient's image to be segmented also plays an important role in the segmentation accuracy. If the atlas is more similar to the patient to be segmented, better will be the segmentation results. Hence, a single arbitrarily selected image is not used as atlas for all the images in the data set; rather, for each image to be segmented, an atlas is adaptively selected from the rest of the images in the data set, based on the Mean Square Error (MSE) metric. For this purpose, each image to be segmented is first affinely registered to the other images in the data set, using leave-one-out strategy; then MSE value is computed over the entire image region for each pair of affinely registered images; the image that has given the least MSE is chosen as the atlas for that patient's image to be segmented.

Notice that use of statistical similarity measures compared to MSE metric could

probably reduce the sensitivity to atlas selection. However, the type of impact of the reduced-sensitivity on the accuracy of segmentation can largely depend on the nature of structures to be segmented. For instance, when the structures to be segmented are thin or soft tissue structures with slowly varying intensities, statistical similarity measures could unfortunately be very insensitive in accurately registering those regions (i.e., negative impact of decreased-sensitivity); thus, for such structures, they are not preferable over MSE metric. On the contrary, for example, if structures are not thin and containing some artifacts, statistical measures could probably result in more accurate as well as robust registration results than MSE metric. Since lymph nodes are soft tissue structures, we preferred MSE for this application.

Results

Figure 3.6 shows one of the patients' image to be segmented. Ground truth segmentations of lymph nodes for that patient are superposed over the same image. Figure 3.7 shows the atlas image that was selected for the above patient's image, based on MSE criteria. Labeled image used for driving the region-based (first level) registration is superposed over the same image. In order to qualitatively illustrate the advantage of the proposed hierarchical approach, segmentation results from the three methods in one of the axial slices, along with its ground truth, are presented in figure 3.8; the names of the lymph nodes are labeled in the sub-figure that shows ground truth segmentations. We can see the advantage of the hierarchical approach, particulary for lymph nodes IB and VA. Notice that, in this slice, these are the lymph nodes that are close to the structures considered during region-based registration, and thus, profiting better from both Yezzi's model and Vemuri's model. Hence, with reference to the ground truth, we can see that these results are better compared to using only one of the two methods.

The segmentation results are quantitatively compared using two statistical metrics: (i) sensitivity, (ii) Dice Similarity Measure (DSM), and a geometrical metric: (i) Mean Hausdorff Distance (HD). Please refer to the chapter 2.4 for a detailed description pf these metrics. Figure 3.9 shows the box plots of Sensitivity, DSM and mean HD when different types of registration forces are used. The corresponding mean and standard deviation values are summarized in table 3.2. It can be noted from these results that for all the three measures, our hierarchical registration approach combining region-based and pixel-based forces gave better results than the Vemuri's model (using only pixel-based forces) and the Yezzi's model (using only region-based forces).

We also performed tests to evaluate the statistical significance of differences between the methods. We used Wilcoxon signed-rank test (described in chapter 2.4) for this purpose. We performed this test on DSM statistics, for each permutation of pair of methods chosen from the three methods, with the alternative hypothesis being: "segmentation results from the first method are statistically better (greater) than the second method."

(a) Axial slice   (b) Sagittal slice   (c) Coronal slice
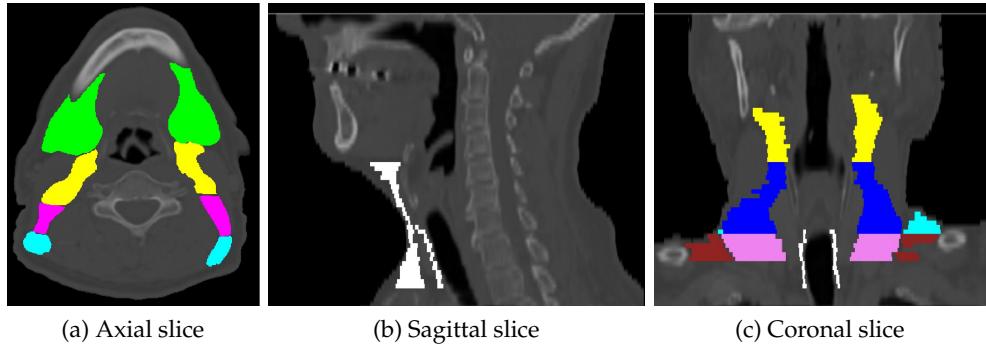
Figure 3.6: One of the patients' image to be segmented in the H&N data set. Ground truth segmentations of the lymph nodes are superposed over the image.

(a) Axial slice   (b) Sagittal slice   (c) Coronal slice
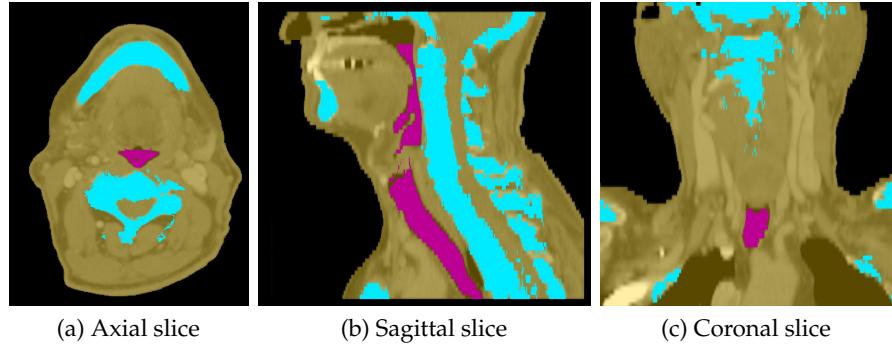
Figure 3.7: Labeled image superposed over the atlas image in the axial, sagittal and coronal views. The labeled structures in the image are (1) External-contour, (2) Bones, and (3) Trachea. This image is used for driving the region-based registration.

Based on these results, the following conclusions are drawn for this application at 0.05 significance level: (i) Segmentation results from Vemuri's model are better than Yezzi's model (with a p-value of 0.0004). (ii) Segmentation results from our hierarchical model are better than Vemuri's model (with a p-value of 0.009) as well as Yezzi's model (with a p-value of 0.0001).

3.5 Comparison of Registration Methods

In this section, we present a comparison of lymph nodes segmentation results obtained from our registration framework with two other commonly used methods in medical imaging, viz, radial basis function (RBF) algorithm [44], and Demons algorithm [16].

(a) Ground truth   (b) Vemuri's model   (c) Yezzi's model   (d) Hierarchical model
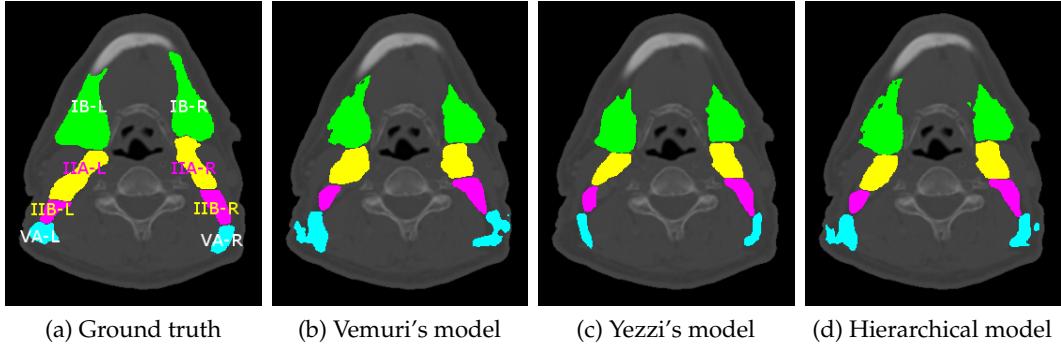
Figure 3.8: Screen shots in one of the axial slices, of the ground truth and automated lymph nodes segmentations obtained from Vemuri's model, Yezzi's model, and our hierarchical model.
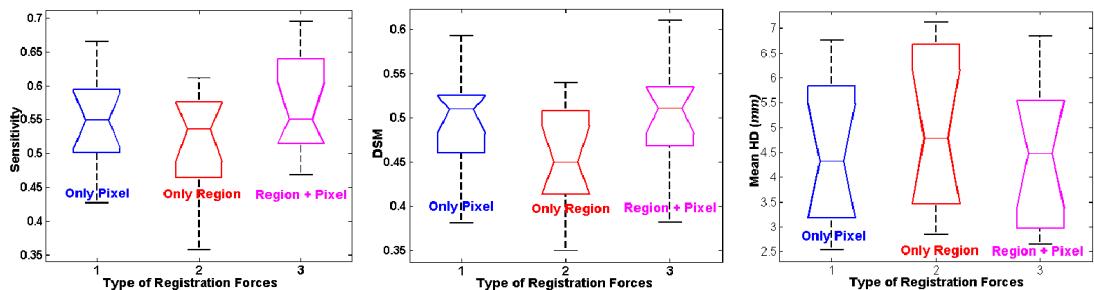
Figure 3.9: Box plots comparing average values sensitivity, DSM and mean HD for lymph nodes segmentation, using (a) only pixel-based forces (Vemuri's model) (b) only region-based forces (Yezzi's model) and (c) our hierarchical model.

Radial Basis Function Algorithm (RBF)

This is a mutual information-based technique proposed by Rhode et al. [44]. The deformation that registers the intensity atlas onto the patient image is modeled with a linear combination of radial basis functions with finite support.

Demons Algorithm

It is an intensity-based algorithm developed by Thirion [16]. Demons algorithm is close to our method for the following reasons. First both are non-parametric. Then, both are specially designed to match contours and also both the algorithms use linear diffusion to extend the deformation computed on the contours to the whole image. Finally both algorithms rely their registration force on polarity. However in the Demons algorithm, the polarity depends on intensity differences and in our algorithm it depends on the inside and outside of the objects to be registered.

The data set used for this evaluation contains the H&N CT scans of 10 patients, ac-

Table 3.2: Mean and standard deviations of Sensitivity, DSM, and mean HD measures for lymph nodes segmentation obtained from (a) Vemuri's model [2], (b) Yezzi's model [1], and (c) our hierarchical model.

Method	Type of forces	Sensitivity	DSM	Mean HD (*mm*)
Vemuri's Model	Only Pixel	0.550 ± 0.066	0.494 ± 0.056	4.507 ± 1.412
Yezzi's Model	Only Region	0.518 ± 0.072	0.451 ± 0.062	5.043 ± 1.669
Our Model	Region + Pixel	**0.574 ± 0.070**	**0.503 ± 0.057**	**4.449 ± 1.418**

quired during routine clinical practice, at Divisions of Radiotherapy, Geneva University Hospital (HUG). The lymph nodes are manually delineated by the radiation oncologist for all the 10 patients, and those are considered as the gold standard for evaluating the automated segmentations. Sagittal views of four of those patients are shown in figure 3.10 for illustrating the anatomical variability present the current data set. The resolutions, size, and the preprocessing procedure for this data set are very similar to the data set used previously, in section 3.4.2. The CT image shown in figure 3.11 is arbitrarily chosen as the atlas image for the data set, and automated segmentations are performed on the remaining 9 CT images [4].

It is common with most of the nonrigid image registration methods to perform an affine or rigid registration prior to the final nonrigid registration in order to recover large deformations. However, our method is found to be robust enough to accurately perform the registration without requiring any such initial registration and is indeed, because of the labeling approach. So we do not use any initial registration and this is one of the advantages of our registration model, whereas we perform an initial affine registration for the other two methods that we use in this evaluation.

In this evaluation, for our method, we just use region-based forces combined with curvature-based regularization. With the inclusion of more registration forces (for e.g., using additionally, pixel-based forces in an hierarchical manner, as illustrated in the evaluation presented in the preceding section) the results from our method could be improved even further. The structures selected for driving this region-based registration are described in figure 3.12. Concerning the parameters setting, we use 4 levels of resolutions for the multi-resolution approach and a $\sigma = 1$ for the Gaussian filtering performing the linear diffusion.

Results

We first present the qualitative results obtained from our registration method for one of the patients' images in the data set. Figure 3.11 and figure 3.12 show respectively the

---

[4]Please note that we conducted this comparison well before performing another different evaluation that we have already presented in section 3.4.2. It is because of this reason that the size of this data set and the atlas selection strategy are slightly different for these two sets of experiments.
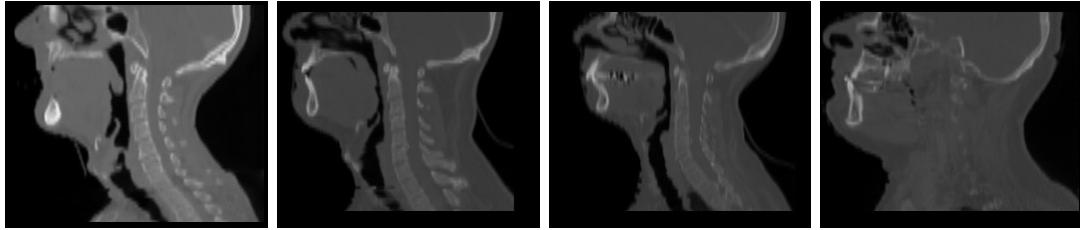
Figure 3.10: Sagittal views of four patients in the current data set. These images illustrate the anatomical variability present in the current data set.

(a) Axial slice   (b) Sagittal slice   (c) Coronal slice
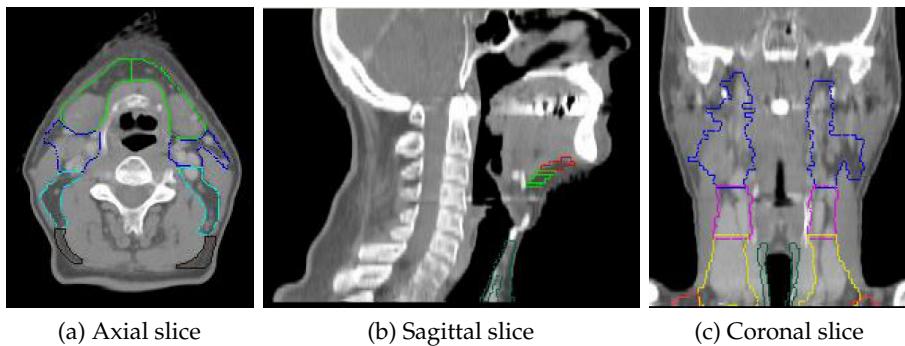
Figure 3.11: Atlas image superposed with manually delineated lymph nodes.

intensity image and the labeled image of the atlas. Manually delineated lymph nodes of the atlas image are shown on the intensity image itself in figure. 3.11. Figure 3.13 shows the CT scans of a target image to be segmented.

The contours of the deformed structures of the labeled image are superposed over the patient's image, and are shown in figure 3.14. It can be seen that the algorithm has registered the selected structures well, except over the lower portion of few axial slices. Although the active contour method is actually capable of registering the contours perfectly, the deviation is due to the trade off between the regularization term that smoothens the deformation field and the region-based term. The auto segmented lymph node regions results are superposed over the patient's image and are shown in figure 3.13.

Finally, figure 3.15 shows 3D segmentations of lymph node volumes on their respective thresholded images; the top row of this figure shows manually delineated lymph node volumes on a thresholded atlas image, and the bottom row shows the lymph node volumes on the thresholded target image obtained from our automated atlas-based segmentation method. These results are visually inspected by an expert oncologist and found to be qualitatively good for most of the lymph nodes.

The quantitative evaluation is performed over all the 10 CT images in the data

(a) Axial slice   (b) Sagittal slice   (c) Coronal slice
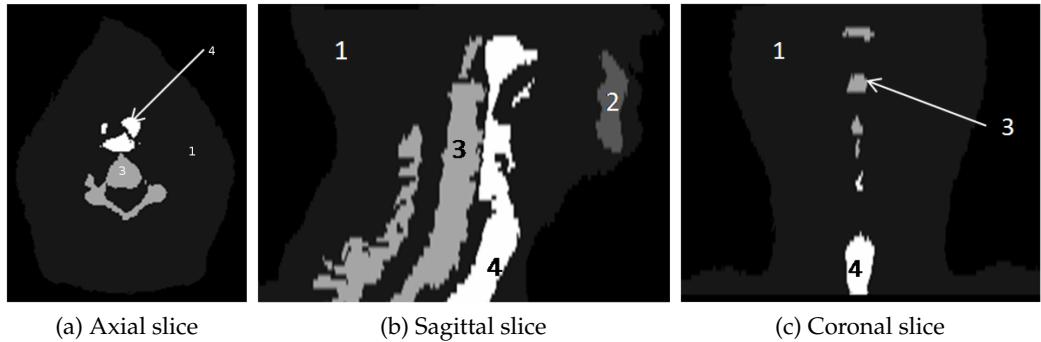
Figure 3.12: Labeled image used for region-based registration. The labeled structures in the image are (1) External-contour, (2) Mandible, (3) Vertebra and (4)Trachea.


(a) Axial slice   (b) Sagittal slice   (c) Coronal slice
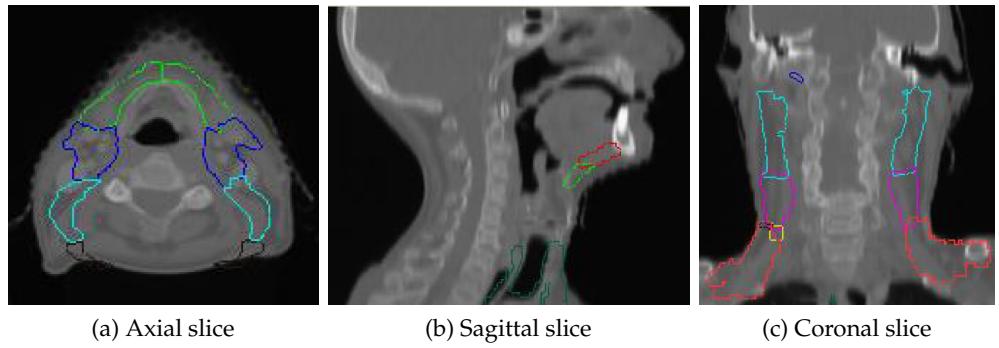
Figure 3.13: One of the patients' image to be segmented. Automated segmentations of the lymph nodes are superposed over the patient's image.

set. The statistics for the left right side pairs of the same lymph node are combined while presenting the results because they are mostly symmetrical. For instance the statistics for lymph nodes IB-left (IB-L) and IB-right (IB-R) are represented as IB itself. The evaluation is performed using three statistical metrics: sensitivity, specificity, Dice Similarity Coefficient (DSC), and a geometrical metric: Hausdorff distance. The details of these metrics are already presented in chapter 2.4.

Figure 3.16 and figure 3.17 respectively show the box plots of sensitivity and specificity for each lymph node, computed from the three methods: RBF algorithm, Demons algorithm and our proposed algorithm. The mean and standard deviations for sensitivity are presented in table 3.3. From the comparison of sensitivity results, it can be noted that our model has performed better than RBF and Demons algorithms. In case of specificity measure, the results for all the lymph nodes, from all the methods are around the ideal value (= 1). Specificity is not a good measure for the current evaluation of lymph node segmentation, because of its dependency on the relative size of the lymph node with respect to the image size, and this is already explained in more detail in chapter 2.4.

| (a) Axial slice | (b) Sagittal slice | (c) Coronal slice |
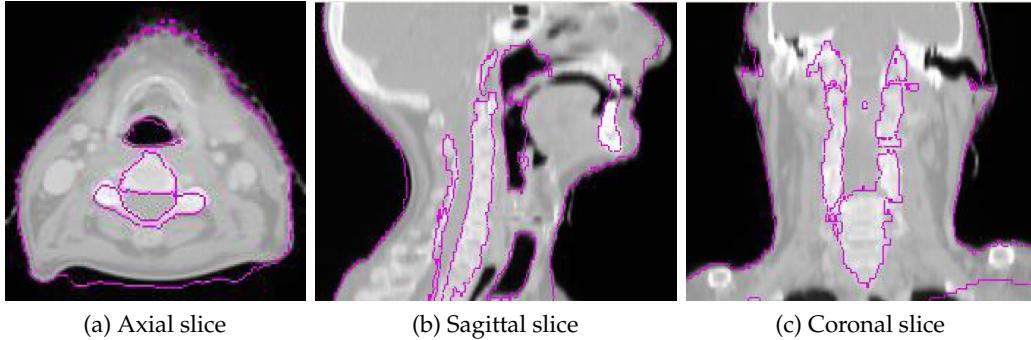
Figure 3.14: The contours of structures in the deformed label are superposed over the patient's image, for qualitative evaluation of the registration.

Table 3.3: Mean and standard deviations for sensitivity and DSC measures, on a data set of 10 patients.

Lymph Node	Sensitivity			DSC		
	RBF	**Demons**	**Our Method**	**RBF**	**Demons**	**Our Method**
IA	0.12 ± 0.23	0.25 ± 0.20	0.60 ± 0.25	0.14 ± 0.25	0.22 ± 0.19	0.39 ± 0.14
IB	0.47 ± 0.22	0.54 ± 0.33	0.78 ± 0.12	0.47 ± 0.22	0.50 ± 0.30	0.61 ± 0.06
IIA	0.37 ± 0.27	0.54 ± 0.26	0.76 ± 0.09	0.31 ± 0.18	0.52 ± 0.23	0.58 ± 0.08
IIB	0.24 ± 0.15	0.48 ± 0.26	0.64 ± 0.14	0.27 ± 0.17	0.48 ± 0.24	0.53 ± 0.12
III	0.31 ± 0.26	0.38 ± 0.24	0.53 ± 0.17	0.29 ± 0.20	0.43 ± 0.27	0.48 ± 0.18
IV	0.24 ± 0.22	0.41 ± 0.27	0.64 ± 0.14	0.20 ± 0.22	0.38 ± 0.25	0.41 ± 0.15
VA	0.04 ± 0.04	0.19 ± 0.16	0.32 ± 0.23	0.05 ± 0.05	0.20 ± 0.17	0.27 ± 0.18
VB	0.28 ± 0.16	0.43 ± 0.28	0.51 ± 0.24	0.19 ± 0.16	0.33 ± 0.22	0.31 ± 0.21
VI	0.15 ± 0.08	0.24 ± 0.14	0.63 ± 0.10	0.19 ± 0.11	0.28 ± 0.13	0.53 ± 0.08
Mean	0.25 ± 0.18	0.38 ± 0.24	0.60 ± 0.16	0.23 ± 0.17	0.37 ± 0.22	0.46 ± 0.13

Figure 3.18 shows the box plots of DSC for each lymph node, computed from all the three methods. The mean and standard deviations for DSC are presented in table 3.3. The DSC values are good considering the complexity of lymph nodes segmentation, but there is still scope for improvement. Since there are no previously reported statistics of DSC for automated segmentation of the H&N lymph nodes, a comparison with the existing literature is not possible. For DSC measure shown in figure 3.18, our method is superior to the RBF method. The reason for lower values of DSC for RBF method is because it is a parametric method and hence cannot cope up with large deformations between the atlas and the patient's image. The mean values of DSC from our method are very close to the Demons method. The variance in DSC is relatively small for our method compared to the Demons method. However, as mentioned earlier, there is a scope for further improving the segmentation results obtained from our method by additionally including more registration forces, as illustrated in section 3.4.2.

(a) Atlas - front view (b) Atlas - side view (c) Atlas - back view

(d) Patient - front view (e) Patient - side view (f) Patient - back view
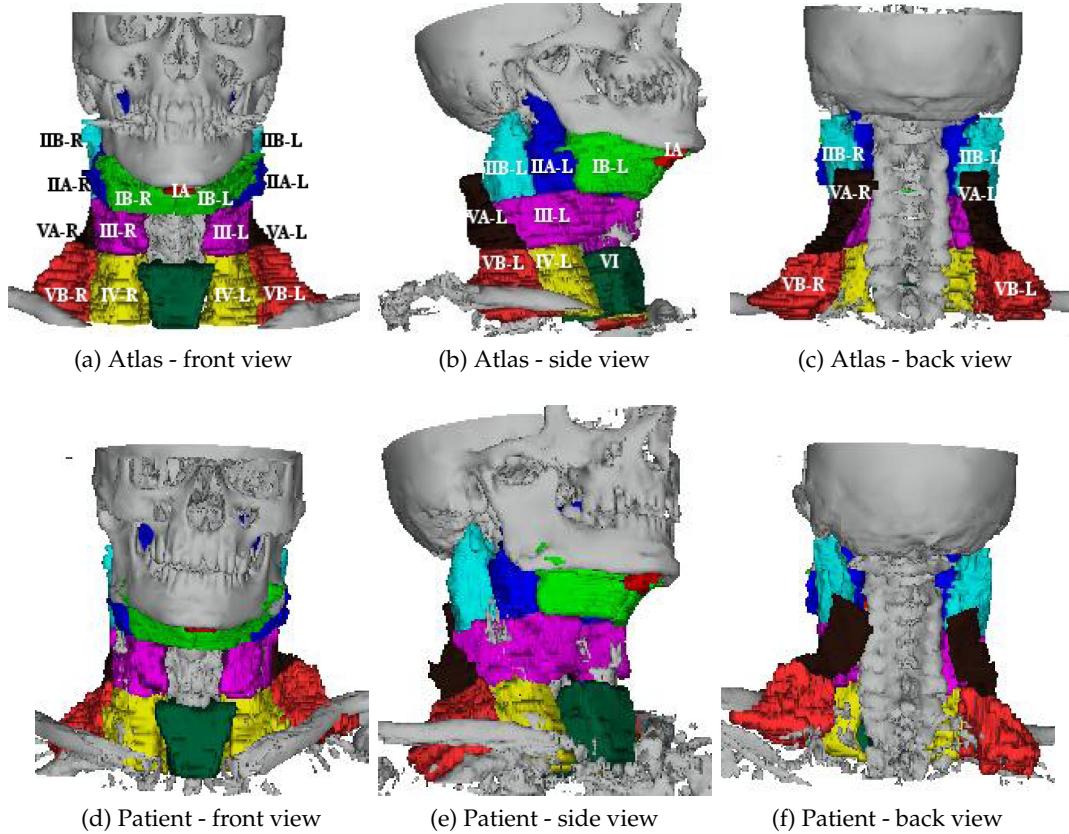
Figure 3.15: Manually delineated lymph node volumes on the thresholded atlas image are presented in the top row. Automated segmentations of the lymph node volumes on the thresholded patient's image are presented in the bottom row.

Figure 3.19 shows the box plots of Hausdorff distance for each lymph node, from all the three methods and the associated mean and standard deviation are presented in table 3.4. As mentioned in chapter 2.4, lower values of Hausdorff distance imply a better resemblance between the gold standard and the automated segmentation. Thus, it can be noted from these box plots that the results are poor for the RBF method. The mean values of Hausdorff distance for our method are close to the Demons. For some lymph nodes, the mean values are slightly better for Demons method and is in the other way some other lymph node regions. However, Our method has the advantage of relatively small value of variance than the Demons method.
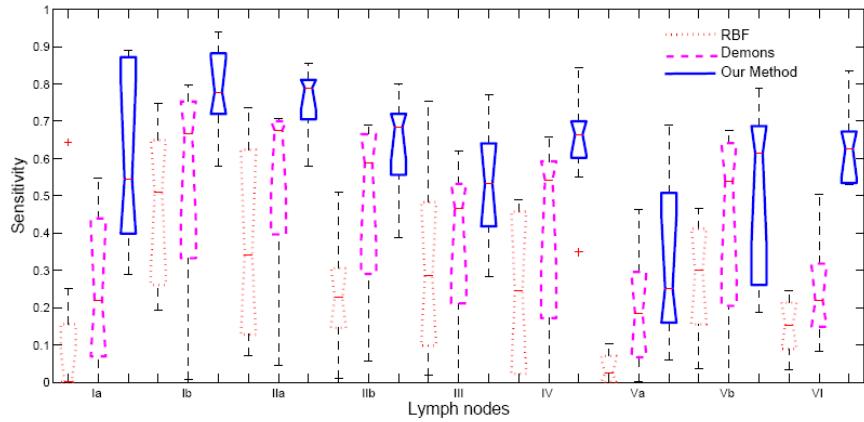
Figure 3.16: Box plots comparing the sensitivity measure of lymph node regions segmentation, using (1)RBF algorithm, (2)Demons algorithm and (3)Our proposed algorithm.
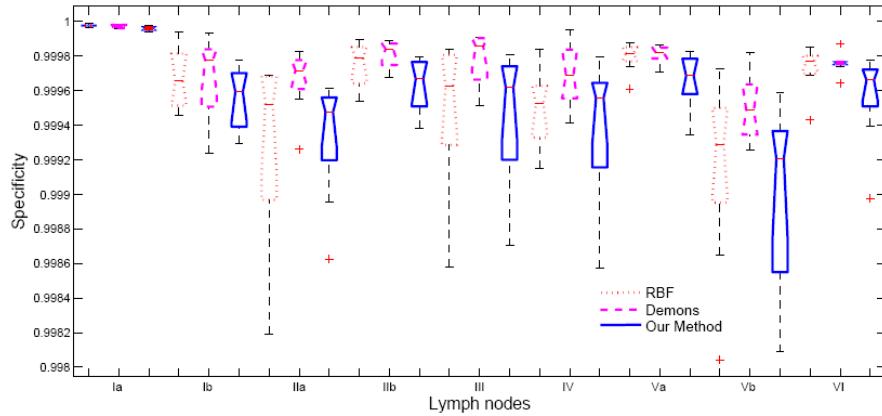
Figure 3.17: Box plots comparing the specificity measure of lymph node segmentations, using (1)RBF algorithm, (2)Demons algorithm and (3)Our proposed algorithm.
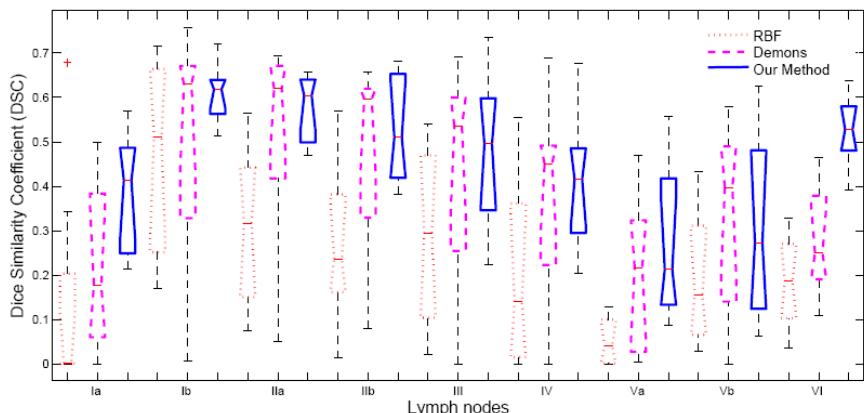
Figure 3.18: Box plots comparing the DSC measure of lymph node regions segmentation, using (1)RBF algorithm, (2)Demons algorithm and (3)Our proposed algorithm.
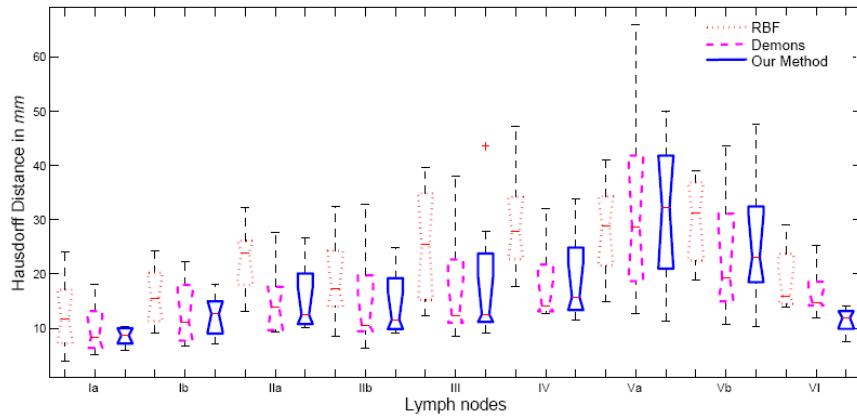
Figure 3.19: Box plots comparing the Hausdorff distance measure for lymph node regions segmentation, using (1)RBF algorithm, (2)Demons algorithm and (3)Our proposed algorithm.

Table 3.4: Mean and standard deviations for Hausdorff distance measure, on a data set of 10 patients.

Lymph Node	Hausdorff Distance (mm)		
	RBF	**Demons**	**Our Method**
IA	12.51 ± 6.76	9.86 ± 4.53	8.49 ± 1.58
IB	15.93 ± 5.69	12.82 ± 6.17	12.31 ± 3.80
IIA	22.66 ± 6.12	14.93 ± 6.21	15.41 ± 6.74
IIB	19.00 ± 7.67	14.84 ± 9.08	14.41 ± 6.11
III	25.36 ± 10.64	17.34 ± 10.93	18.45 ± 11.85
IV	29.31 ± 9.43	17.84 ± 7.31	19.16 ± 9.10
VA	28.19 ± 8.74	32.14 ± 17.95	31.43 ± 13.49
VB	29.89 ± 7.85	23.12 ± 12.02	25.76 ± 12.31
VI	18.92 ± 6.13	16.51 ± 4.44	11.43 ± 2.27
Mean	22.42 ± 7.67	17.71 ± 8.74	17.43 ± 7.47

3.6   Discussion and Conclusions

The main contribution of this chapter is a new variational framework that encompasses many existing registration methods. We have presented how active contour (AC) based registration model of [1], and pixel-based registration model of [2] can be derived from the proposed framework. We have also proposed a new label function for performing multi-phase AC-based registration. The advantage of the proposed label function is that it can distinguish any number of regions with a single function. Our proposed framework allows to select the structures to be used for driving the registration. This way, the atlas can be registered on target images using the objects that are consistent between both the images.

We have proposed to use multiple registration forces in an hierarchical manner, based on their relative global characteristics so that convergence to a local minimum can be avoided besides benefitting from multiple registration forces. We have demonstrated the advantage of such hierarchical approach through a 2D synthetic example and a clinical application. In the clinical application, we have compared the 3D head and neck (H&N) lymph nodes segmentations obtained from the pixel-based Vemuri's model [2] and the region-based Yezzi's model [1] with the results obtained from our registration framework that facilitated combining both region-based and pixel-based forces in an hierarchical manner. The improvements in segmentation accuracy from our registration model are found to be statistically significant when compared to both Vemuri's model and Yezzi's model.

We have compared the 3D H&N lymph nodes segmentation results obtained from our method with RBF and Demons registration algorithms. Demons method and our registration method have clearly outperformed the RBF method. This is probably because, while RBF is a parametric method, the rest of the two methods are non-parametric registration methods. The segmentation results from our method are very close to the results Demons methods, and sometimes, marginally better. However, please note that, in this evaluation, we have used only region-based forces for our registration method. Hence, unlike Demons method, there is still a definite scope for further improving the results from our method as it facilitates including many other registration forces. Moreover, in case of any inconsistencies between the atlas and target images, unlike Demons method, our registration framework also facilitates to select specific structures to drive the registration process. Thus, these evaluations have clearly demonstrated the potential of the proposed registration framework.

In all the experiments presented so far, automated segmentations of each target image are obtained based on just a single atlas which is chosen either arbitrarily, or based on certain optimization criteria. However, as described in chapter 2.2, results obtained by merging segmentations from multiple atlases are more robust and accurate than the single atlas-based results. The rest of the thesis is hence focussed on multi

atlas fusion methodologies. Finally, please note that the atlas fusion methodologies presented in this thesis are independent of the registration methods used for computing the deformation field; hence, although we used in this thesis our "active deformation fields" for performing registration, one can choose their own preferred registration strategies, and still can use any of these fusion methods.

In the next chapter, we present a new Markov Random Field (MRF)-based fusion framework for merging segmentations obtained from multiple atlases. We also present a detailed evaluation of various fusion methods for performing H&N lymph nodes segmentation.

4 | MRF-based Framework for Multi-Atlas Fusion

In this chapter, we propose a general Markov Random Field (MRF) based framework that can perform simultaneously both edge-preserving smoothing and fusion of labels obtained from multiple atlases. More specifically, we formulate the label fusion problem with MRF-based neighborhood model, as an energy minimization problem containing a unary data term and a pairwise smoothness term. We present how the existing methods like majority voting, global weighted voting and local weighted voting can be reframed to profit from the proposed framework. This framework resulted in generating more accurate segmentations as well as more contiguous segmentations by getting rid of unwanted holes and islands. Finally, we present a comprehensive evaluation of various fusion methods, both 'with' and 'without' using MRF-based smoothness term, in the specific context of segmentation of lymph nodes in 3D head and neck CT images.

4.1   Introduction

Multi Atlas fusion methods are well known for providing robust and accurate segmentations when compared to single atlas-based strategies. We have already presented a brief summary of atlas fusion methods in chapter 1.2, and described the formulation of the existing fusion methods in chapter 2.2.

One of the main problems with many of the existing fusion strategies is, although the segmentations obtained from each individual atlas are contiguous, the merged segmentations can be fragmented containing undesirable discontinuities like holes and islands [26]. Some of the atlas fusion approaches like SBA provide more contiguous segmentations compared to other widely used methods; however, those approaches cannot be generalized or extended to other existing fusion methods.

To deal with the above mentioned problem of discontinuous segmentations, the segmentation results are sometimes post-processed. For instance, in [26], the segmentation results for each structure are independently considered as binary masks, and

are post-processed by first smoothing them with a Gaussian kernel, and then thresholded at 0.5; after that, they further perform connected component analysis and retain only the largest component. In [37], the segmentation results are post-processed by morphological closing, and then followed by the extraction of the largest component.

However, such post-processing approaches have many disadvantages. First, they do not preserve sharp edges. Second, depending on the order in which those labels are smoothed, such simple Gaussian smoothing of labels results in conflicting regions across the boundaries of adjacent structures. To avoid this conflict between regions, one could probably introduce approaches like a more complex iterative coupled Gaussian smoothing of multiple labels, but that still will not solve the first problem of preserving the edges. Finally and most importantly, it is not elegant to handle "fusion" and "smoothing" as two different, independent problems.

To address the above mentioned issues, in this chapter, we propose an MRF-based framework that performs simultaneously both fusion and edge-preserving smoothing of multiple labels. The rest of the chapter proceeds as follows. In the next section, we present a new MRF-based framework and reformulate some of the existing fusion methods to fit into our framework. In section 4.3, we present an evaluation of various fusion methods for H&N lymph nodes segmentation. Finally, conclusions are presented in section 4.4.

4.2 MRF-based Fusion Framework

In this section, we first present the MRF-based framework that we propose for performing simultaneously both fusion and smoothing of multiple structures whose segmentations are obtained from multiple individual atlases. We then show how some of the existing fusion methods like majority voting, global weighted voting and local weighted voting can be reformulated to profit from the proposed framework.

Let V be the number of voxels in the target image. Let Y_p denote the label assigned to the pth voxel in the output image. Let Y be the set containing labels assigned to each voxel in the output image, i.e., $Y = \{Y_1, \cdots Y_V\}$. Then, we will be formulating the atlas fusion as a general energy minimization problem of the following form.

$$Y^* = \arg \min_Y \left\{ E_{\text{data}}(Y) + \lambda \, E_{\text{smooth}}(Y) \right\}, \tag{4.1}$$

where the first term is a data term (unary term), and it should be defined in such a way that it reaches to a minimum value when the chosen fusion criteria has been met; the second term is a smoothness term (pairwise term), and in the current context, it should penalize for irregular distribution of labels while allowing for the edge-discontinuities. λ is a weighting parameter between the data term and smoothness term. Energy equation
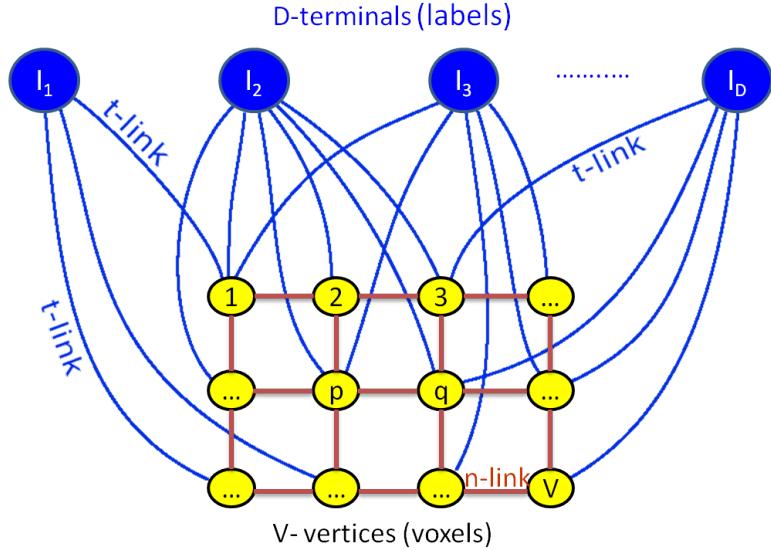
Figure 4.1: Illustration of graph-construction procedure for solving the MRF-based fusion problem using graph-cuts method. This picture is inspired from a presentation of Prof. Bokov on graph-cuts.

of the form (4.1) is ubiquitous in many computer vision problems, and there exists various efficient MRF optimization methods for solving them [49, 105]. In this thesis, we use the graph cuts expansion method [106–108] as it guarantees convergence to a global optimum for the current model. One could also explore other alternative promising algorithms like loopy belief propagation [109, 110], tree reweighted message passing [111, 112], and graph cuts based on linear programming [113].

We now mention the graph construction procedure in more detail with the help of the illustration presented in figure 4.1. Notice from this figure that, each voxel in the target image is represented by a vertex, and thus there are V vertices in the graph representing all the voxels of the image. Then, each possible label to be assigned (denoted as $l_1, l_2, \cdots, l_D$ in the figure) is represented by an additional node, and these nodes are typically called as "terminals." Each of these D terminals are connected to all V vertices, and those connecting edges are called "t-links." (In the figure, we have drawn only few t-links as this figure is just for illustrating the procedure.) The weight of the t-link connecting a particular terminal to a vertex is equal to the resulting data cost for that voxel when it is assigned the label represented by that terminal. The edges that connect a vertex (i.e., voxel) to another (neighboring) vertex are called "n-links." The weight of an n-link is equal to the resulting pairwise smoothness cost for those two vertices, and this is dependent on the respective labels assigned to those two vertices. After the construction of the graph in this manner, the initial MRF-based fusion problem is eventually transformed into computing a minimum cost cut on the graph.

53

Note that although energy formulations similar to the above one were used previously [114] in performing segmentations using individual atlases, and in other computer vision problems like denoising, inpainting and stereo vision [105], it is for the first time such MRF-based energy minimization formulation is used for *merging segmentations* obtained from multiple atlases. We have noticed that, in computer vision, [115] has recently proposed a framework, that has certain similarities with our current framework. [115] uses Markovian Bayesian framework for combining several quickly estimated segmentation maps. Our framework is different from [115] in the sense that, unlike [115], we formulate the fusion problem as an energy minimization problem; further, we propose here a generalized framework in a completely different context, and we also show how some of the existing fusion methods can be reformulated to profit from our framework.

We now reformulate some of the existing fusion methods to fit into the data term of the above mentioned MRF-based energy minimization formulation.

4.2.1 Majority Voting

Let N be the number of atlases. Let X^j represent the jth input labeled image (corresponding to the jth atlas) after applying the transformation that maps the jth atlas to the output intensity image. Let X_p^j be the label assigned to the pth voxel of X^j. Let D be the total number unique label values (including the background label) in the input labeled images. Let $L = \{l_1, \cdots, l_D\}$ represent the set of all possible labels to be assigned. Let δ be a Kronecker delta function.

Majority Voting (MV) assigns for each voxel a label that maximum number of atlases agree on. The original formulation of MV is already described in detail in chapter 2.2.1, and it is a maximization problem of the following form.

$$Y_p = \max \left[\sum_{j=1}^{N} \delta \left(X_p^j, l_1 \right), \sum_{j=1}^{N} \delta \left(X_p^j, l_2 \right), \cdots, \sum_{j=1}^{N} \delta \left(X_p^j, l_D \right) \right]. \tag{4.2}$$

We now reformulate this as a minimization problem. The above maximization formulation of MV is exactly equivalent to minimizing the following energy function.

$$E_{\text{data}}(Y) = \frac{1}{N} \sum_{p=1}^{V} \sum_{j=1}^{N} \left(1 - \delta \left(X_p^j, Y_p \right) \right). \tag{4.3}$$

Notice that, at a pth voxel, if the label assigned to output image (Y_p) is same as the label assigned by a jth atlas (X_p^j), the corresponding term in the above equation (i.e., $1 - \delta \left(X_p^j, Y_p \right)$) becomes 0, and its value is 1 otherwise; hence the summation of this

term computed over all atlases at given voxel ($\sum_{j=1}^{N} 1 - \delta\left(X_p^j, Y_p\right)$) is minimum when the output label assigned is same as the label assigned by majority of atlases. When this data term alone is used in equation (4.1) (i.e., when λ is set to zero), it is evident from the above equation that the labeling decision to be made at each voxel is independent of the labels assigned to rest of the voxels. In other words, the outside summation over each voxel ($\sum_{p=1}^{V}$) present in equation (4.3) is separable to independent voxel-wise optimizations similar to equation (4.2). Thus, when λ is set to zero in equation (4.1), this minimization-formulation is exactly equivalent to the original maximization problem. The reason to additionally include a factor of $(1/N)$ in the above data term is to normalize its magnitude when smoothness term is also used along with this data term. When no smoothness term is used, this normalization term obviously does not affect the output labeling decisions.

4.2.2 Global Weighted Voting

Global Weighted Voting (GWV) is similar to MV except that it attaches a global weight to each atlas while counting its vote. The original formulation of GWV is already described in detail in chapter 2.2.2, and it is a maximization problem of the following form.

$$Y_p = \max\left[\sum_{j=1}^{N} w^j \delta\left(X_p^j, l_1\right), \sum_{j=1}^{N} w^j \delta\left(X_p^j, l_2\right), \cdots, \sum_{j=1}^{N} w^j \delta\left(X_p^j, l_D\right)\right], \qquad (4.4)$$

where w^j is the global weight assigned to the decisions made by the jth atlas.

We now reformulate the maximum-energy formulation of GWV into an equivalent energy minimization problem, and is as follows.

$$E_{\text{data}}(Y) = \frac{1}{N} \sum_{p=1}^{V} \sum_{j=1}^{N} \hat{w}^j \left(1 - \delta\left(X_p^j, Y_p\right)\right), \qquad (4.5)$$

where $\hat{w}^j$ is the normalized global weight assigned to the decisions made by the jth atlas. Notice that in the reformulated equation, unlike in equation (4.4), we normalize the global weights (w^j) to the range 0-1. When the data term in equation (4.5) alone is optimized, this normalization is not required, and even if the normalization has been done, it has no impact on the output labeling. Thus, equation (4.5) then results in exactly the same output labeling as that of equation (4.4) irrespective of the normalization. However, when the above data term is combined with a smoothness term as shown in equation (4.1), such normalization is useful in order to make the relative magnitudes of the data term and the smoothness term independent of the number of atlases.

Finally, w^j values defined in equation (2.3) are normalized as follows.

$$\hat{w}^j = \frac{w^j}{\max\left(w^1, \cdots, w^N\right)}.$$

(4.6)

4.2.3 Local Weighted Voting

While GWV assigns a single global weight for each atlas, Local Weighted Voting (LWV) assigns an individual local weight for each voxel in each atlas. The original formulation of GWV is already described in detail in chapter 2.2.3, and it is a maximization problem of the following form.

$$Y_p = \max\left[\sum_{j=1}^{N} w_p^j\, \delta\left(X_p^j, l_1\right), \sum_{j=1}^{N} w_p^j\, \delta\left(X_p^j, l_2\right), \cdots, \sum_{j=1}^{N} w_p^j\, \delta\left(X_p^j, l_D\right)\right],$$

(4.7)

where w_p^j is the weight assigned to the jth atlas, at p^{th} voxel.

Notice that the formulation for GWV presented in equation (4.4), and the above formulation for LWV are very similar except that the weight assigned for each atlas is not fixed anymore for the entire image; rather, the weights are now dependent on the voxel location. Hence, the corresponding minimum energy formulation for LWV is as follows.

$$E_{\text{data}}(Y) = \frac{1}{N} \sum_{p=1}^{V} \sum_{j=1}^{N} \hat{w}_p^j \left(1 - \delta\left(X_p^j, Y_p\right)\right),$$

(4.8)

where $\hat{w}_p^j$ is the normalized local weight assigned to the jth atlas, at p^{th} voxel. The effects of this normalization are same as those that we described for GWV in the previous subsection.

Finally, w_p^j values defined in equation (2.5) are normalized as follows.

$$\hat{w}_p^j = \frac{w_p^j}{\max\left(w_p^1, \cdots, w_p^N\right)}.$$

(4.9)

4.2.4 Smoothness Term

Regarding the smoothness term, we use here the widely used edge-preserving Potts model [105]. However, one could even use other models that are specific to a given application, that incorporate prior knowledge about the spatial distribution of the

labels [1].

Let $\aleph_p$ be the set of all voxels in the predefined neighborhood of the pth voxel. Let w_{pq} represent spatially varying per-pairing weights. Then the Potts model-based smoothness term is given by

$$E_{\text{smooth}}(Y) = \sum_{p=1}^{V} \sum_{\substack{\forall q \in \aleph_p \\ q \neq p}} w_{pq} \left(1 - \delta(Y_p, Y_q)\right).$$
$\qquad(4.10)$

For the H&N lymph nodes segmentation application that we consider in this chapter, we simply set w_{pq} to $1/$(cardinality of the set $\aleph_p$). However, as shown in the above equation, this framework clearly facilitates to further incorporate, through the term w_{pq}, the prior knowledge of spatially varying probability distributions of neighboring label pairs.

4.3 Evaluation

In this section, we present a comprehensive evaluation of the state-of-the-art fusion methods and the proposed MRF-based framework in the context of segmentation of lymph nodes in a data set of 3D Head and Neck (H&N) CT images. In this evaluation, we consider 10 lymph node structures for automated segmentation: (i) IB-Left, (ii) IB-Right, (iii) IIA-Left, (iv) IIA-Right, (v) IIB-Left, (vi) IIB-Right, (vii) III-Left, (viii) III-Right, (ix) IV-Left, (x) IV-Right. These lymph nodes for one of the patients are shown in figure 4.2. Both the left and right structures that share the same lymph node number are shown in the same color in these figures. (For example, both IB-Left and IB-Right are displayed in magenta color; this is just a convention that we followed for the sake convenience in visualization). Lymph nodes IB, IIA, IIB, III and IV are respectively shown in magenta, cyan, red, green and blue colors in this figure.

The current data set contains 12 atlas images and 8 patients' images to be segmented, acquired at Divisions of Radiotherapy, Geneva University Hospital (HUG), during the routine clinical practice. These images have a resolution of $1mm \times 1mm \times 1mm$ in X, Y and Z directions respectively. The Field of View (FOV) was not same for all the images, and hence, the images are cropped during the preprocessing to contain the same FOV. An expert oncologist has manually delineated lymph nodes on all the images, and those manual delineations are considered as the ground truth segmentations; the manual delineations have been done according to the consensus guidelines published in [74].

Fusion methods evaluated in this chapter are (i) MV, (ii) GWV, (iii) LWV, (iv) SBA, and (v) STAPLE. All the methods (except SBA and STAPLE) are evaluated for both 'with'

---

[1]We will be exploring more in this direction (i.e., towards developing advanced neighborhood prior models) in the coming chapters.

(a) Axial slice   (b) Coronal slice   (c) Lymph node volumes
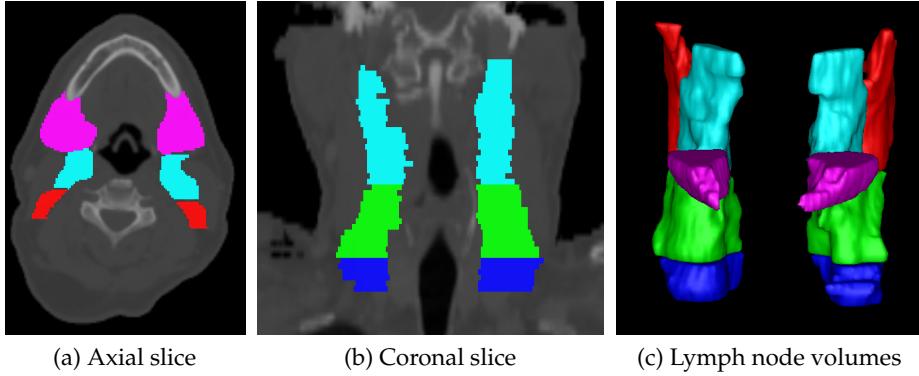
Figure 4.2: Manual delineations of H&N lymph nodes for one of the images. Lymph nodes IB, IIA, IIB, III and IV are shown in magenta, cyan, red, green and blue colors respectively.

and 'without' the MRF-based smoothness term (i.e., respectively with $\lambda \neq 0$ and $\lambda = 0$ in equation (4.1)). Those methods with the smoothness term are denoted either with a suffix "+MRF", or an apostrophe (') added to the name of the method. Although we did not yet explore reformulating STAPLE algorithm to fit into our label fusion framework, we still include this method in the current evaluation; this is because STAPLE is one of the widely used methods, and thus, it is interesting to notice its relative performance in the specific context of H&N lymph nodes segmentation; we use here the multi-label implementation of the STAPLE proposed in [40]. Since the reformulation approach required for SBA is different from the voting-based methods, we will be dealing with it separately in the next chapter.

For all the experiments presented in this chapter, λ is empirically set to 0.5. This value is selected based on visual inspection of results, and is not optimized further. In the context of the current evaluation, such empirical λ selection is sufficient enough since the goal here is to compare the relative performances of different methods, and to demonstrate the advantages of the MRF-based framework. Instead, if the purpose were to obtain more accurate segmentations from individual methods, λ value may be further fine-tuned iteratively, based on the prior knowledge about the output structures (for instance, based on the prior knowledge about the "number of connected regions" in the structures of interest); λ value could also be set based on the quantitative results from a training data set. In the smoothness term, $\aleph_p$ is set to the standard 3D grid of 6 neighbors. Finally, similarity weights for LWV (i.e., $\hat{w}_p^j$) are computed over $9 \times 9 \times 9$ neighborhood.

We also evaluate the performance of the fusion methods for varying number of atlases. For this purpose, the atlases for each patients' image to be segmented are ordered based on the overall DSM between the ground truth and the automated segmentation results obtained from individual atlases. As the ground truth segmentations for target

images may not be available in advance in routine clinical practices, in such cases, the atlases can instead be ordered based on the computation of an appropriate metric (like sum of "mean square difference" (MSD) of intensities computed over all voxels, or deformation field vector) between each transformed atlas and the target image.

Regarding the registration, all the 12 atlases are registered to each patient to be segmented. An initial affine registration is performed followed by a two-level hierarchical nonrigid registration. The affine registration is performed based on MSD of intensities metric. In the first level, region-based nonrigid registration is performed with its parameters and procedure very similar to the application already presented in chapter 3.4.2. The structures that have been selected for driving the region-based registration are (i) external contours of the H&N, (ii) bones, and (iii) trachea. Notice that all the above selected structures have unique intensity characteristics (i.e., mean values of intensities inside and outside the selected structures), and they also influence the location of the lymph nodes to be segmented. This region-based registration is followed by a final pixel-based registration. The parameters for pixel-based registration are also same as those used in the application presented in chapter 3.4.2.

Finally, before presenting the evaluation results, we want to make few comments about the computational time. We implemented all the registration and fusion algorithms in the ITK [2] framework, which is a widely used open-source tool in medical imaging. We have not yet optimized our code for making it computationally efficient. All the experiments are run on the same machine (2.40 GHz, Intel Core 2 Duo processor, 3 GB RAM). With our current implementation, the two-level nonrigid registration for one pair of images is taking around 50 minutes. When compared to the time taken for registration, the fusion procedure is fast. For example, in our current evaluation, the computational timings for merging the segmentation results obtained from 12 atlases are approximately varying from a minimum of 30 seconds to a maximum of 5 minutes depending on the fusion method. Among all the fusion methods, MV is the fastest one whereas LWV needs relatively more time as it requires the computation of local similarity measure at each voxel in the image. When compared to MV, SBA also takes more time as it requires the computation of Euclidean distances at each voxel, to each possible label. Please note that there is further scope for the optimization of our current implementation of these algorithms, and thus, these timings are only indicative.

Results

We start with presenting qualitative results for one of the patients' images; these results are obtained by merging segmentations from 12 atlases using different fusion methods. Figure 4.3 shows, in one of the axial slices, the ground truth segmentations, and the segmentations obtained from various fusion methods. Figure 4.4 shows segmentations in

---

[2]http://itk.org/

(a) Ground Truth   (b) MV   (c) MV+MRF   (d) GWV   (e) GWV+MRF

(f) LWV   (g) LWV+MRF   (h) SBA   (i) STAPLE
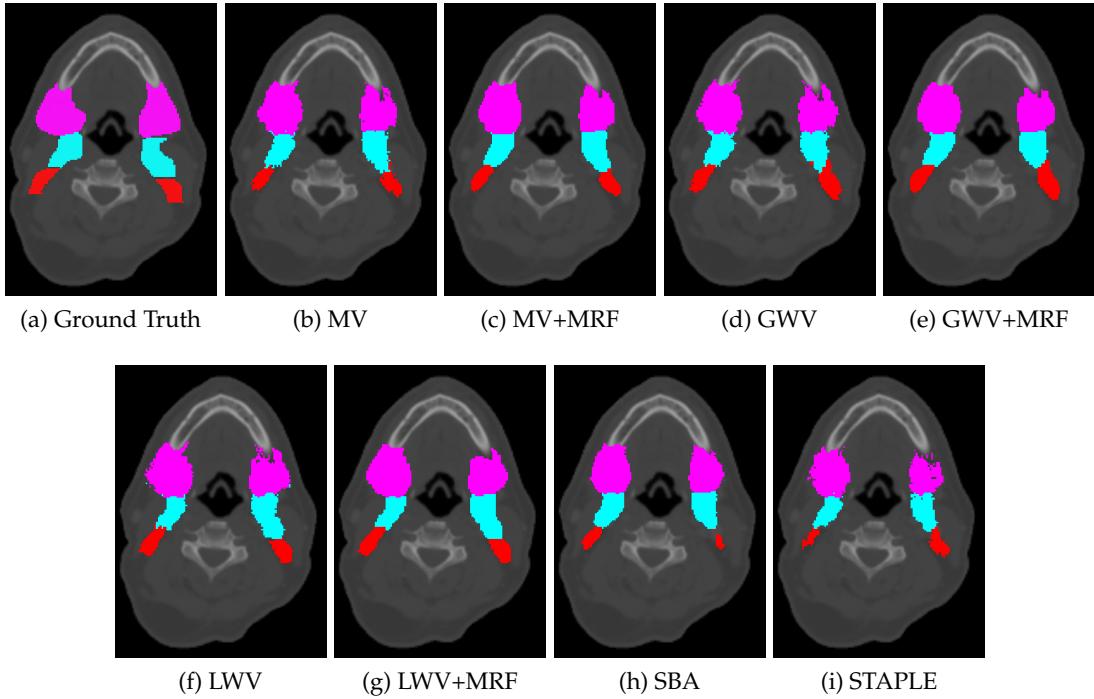
Figure 4.3: Screen shots of H&N lymph nodes segmentations in the axial orientation, from different fusion methods, using 12 atlases.

one of the coronal slices. Similarly, figure 4.5 shows the lymph node volumes constructed from the ground truth segmentations, and from the segmentations obtained from various fusion methods. It can be noted from these figures that, the inclusion of the proposed MRF-based smoothness term, in general, provided more accurate segmentations, and also resulted in getting rid of unwanted holes and islands in the output segmentations when compared to their counterparts that do not use any such MRF-based term.

The quantitative evaluation is performed over the entire data set, using two metrics; (i) Dice Similarity Metric (DSM) and (ii) number of connected regions per label. As we already detailed in chapter 2.4, the output segmentations of each lymph node (label) should ideally contain a single contiguous region. Hence, we evaluate the fusion algorithms also based on the number of connected regions it has created per each label (ideal value = 1).

Figure 4.6 and figure 4.7 present box plots of DSM for each lymph node, and also box plots of average DSM computed across all lymph nodes. The following observations can be made from these box plots. For LWV, GWV and MV based fusion schemes, methods with the proposed MRF-based edge preserving smoothness term generally provided better results than their counterparts that do not incorporate any such smoothness term. If we look at a particular method both 'with' and 'without' the inclusion of MRF-based term as a single class of methods, then, among all the classes of methods, LWV provided
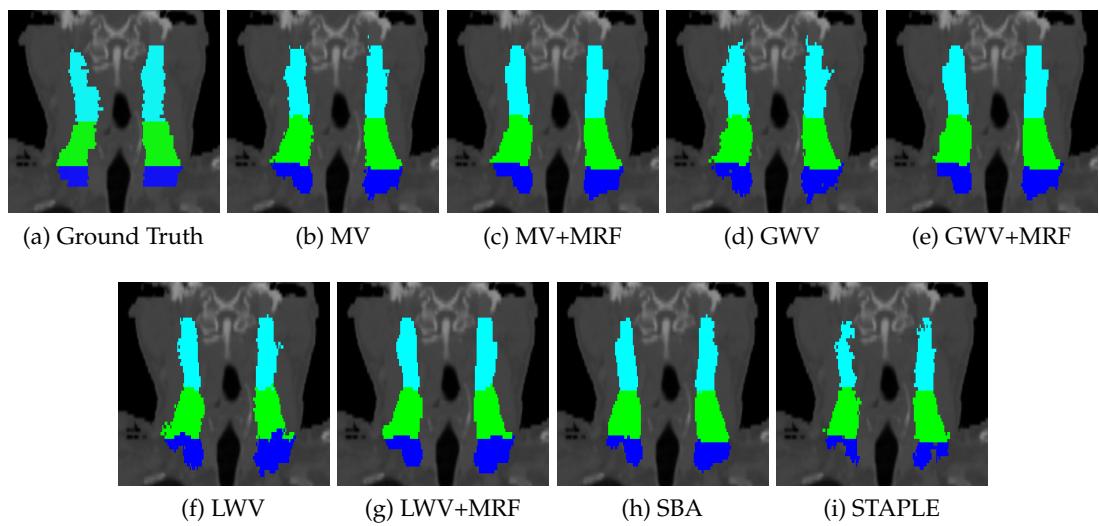
Figure 4.4: Screen shots of H&N lymph nodes segmentations in the coronal orientation, from different fusion methods, using 12 atlases.
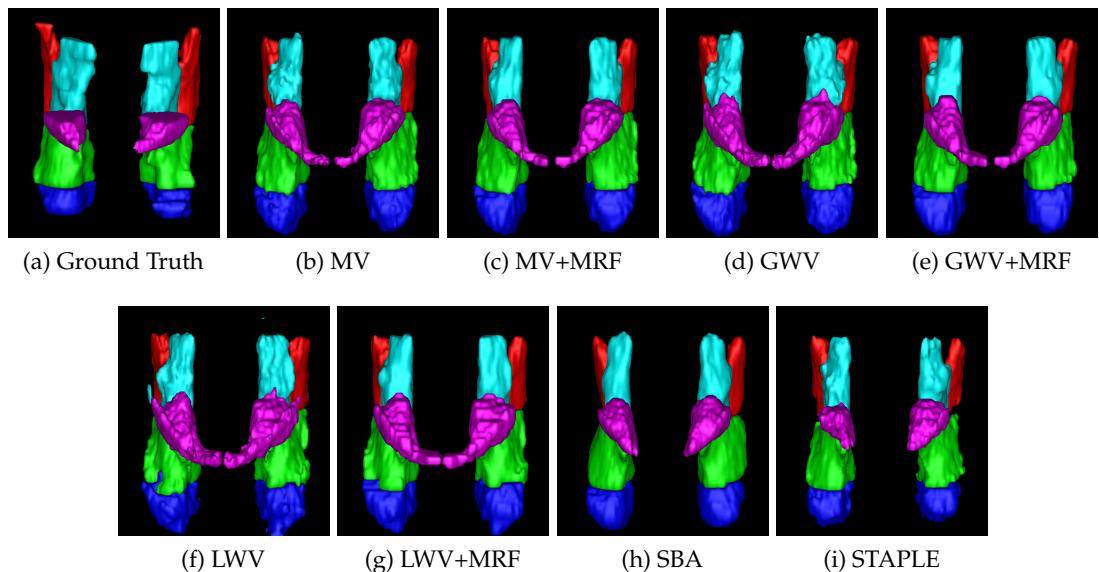
Figure 4.5: Lymph node volumes for one of the patients' images, obtained from different fusion methods, using 12 atlases.

the best overall segmentation results, and it is followed by GWV, MV, SBA and STAPLE respectively.

Figure 4.8 presents box plots of number of connected regions per each label. These statistics are computed over all the lymph nodes, and across all the patients. As mentioned earlier, each lymph node should ideally contain a single contiguous region. The following points can be noted from these box plots. All the MRF-based methods have successfully resulted in almost a single contiguous region per each label. SBA method also resulted in almost a single contiguous region per each structure. Among the rest of the methods, LWV resulted in the highest number of disconnected regions per label, and it is followed by GWV, MV and STAPLE methods respectively.

We further evaluated the statistical significance of the improvements in the segmentation results with the inclusion of MRF-based smoothness term. Wilcoxon signed-rank test (described in chapter 2.4) is used for this purpose. We performed this test on DSM statistics, for each pair of methods (i.e., without and with the inclusion of MRF-based smoothness term), with the alternative hypothesis being: "Segmentation results with the inclusion of MRF-based smoothness term are statistically better (greater) than the original methods that do not use this term." From these experiments, it is found (at 0.05 significance level) that for MV, GWV and LWV methods, the improvements in the segmentation results due to the inclusion of MRF-based smoothness term are statistically significant compared to their counterparts that do not use such MRF-based smoothness term. Thus, these statistical tests also supported the observations from the previous box plots of DSM that, MRF-based smoothness term could improve the segmentation results for MV, GWV and LWV methods.

We now present the evaluation of effects of varying number of atlases on various fusion methods. Figure 4.9 shows, in one of the axial slices of a patient's image, segmentation results for varying number of atlases obtained from LWV+MRF method. It can be noted from this figure that the segmentation results in this slice are quite stable for $N \geq 7$. Instead of presenting qualitative results for each method, we directly present the quantitative results for all the methods as that would make the task of comparison across multiple fusion methods easier.

Figure 4.10 shows box plots of average DSM values obtained from different fusion methods, for varying number of atlases; these average DSM values are computed across all lymph nodes; box plots of methods 'with' and 'without' MRF-based term are shown side-by-side in the same sub-figure in different colors; all sub-figures are plotted with the same range along the Y-axis so that results across multiple fusion methods can be visually compared. Figure 4.11 shows mean values of DSM for different atlas fusion methods, and for varying number of atlases. The following observations can be made from these figures. Local weighting based methods (i.e., LWV+MRF and LWV) performed consistently better than the rest of the methods. With the increasing number
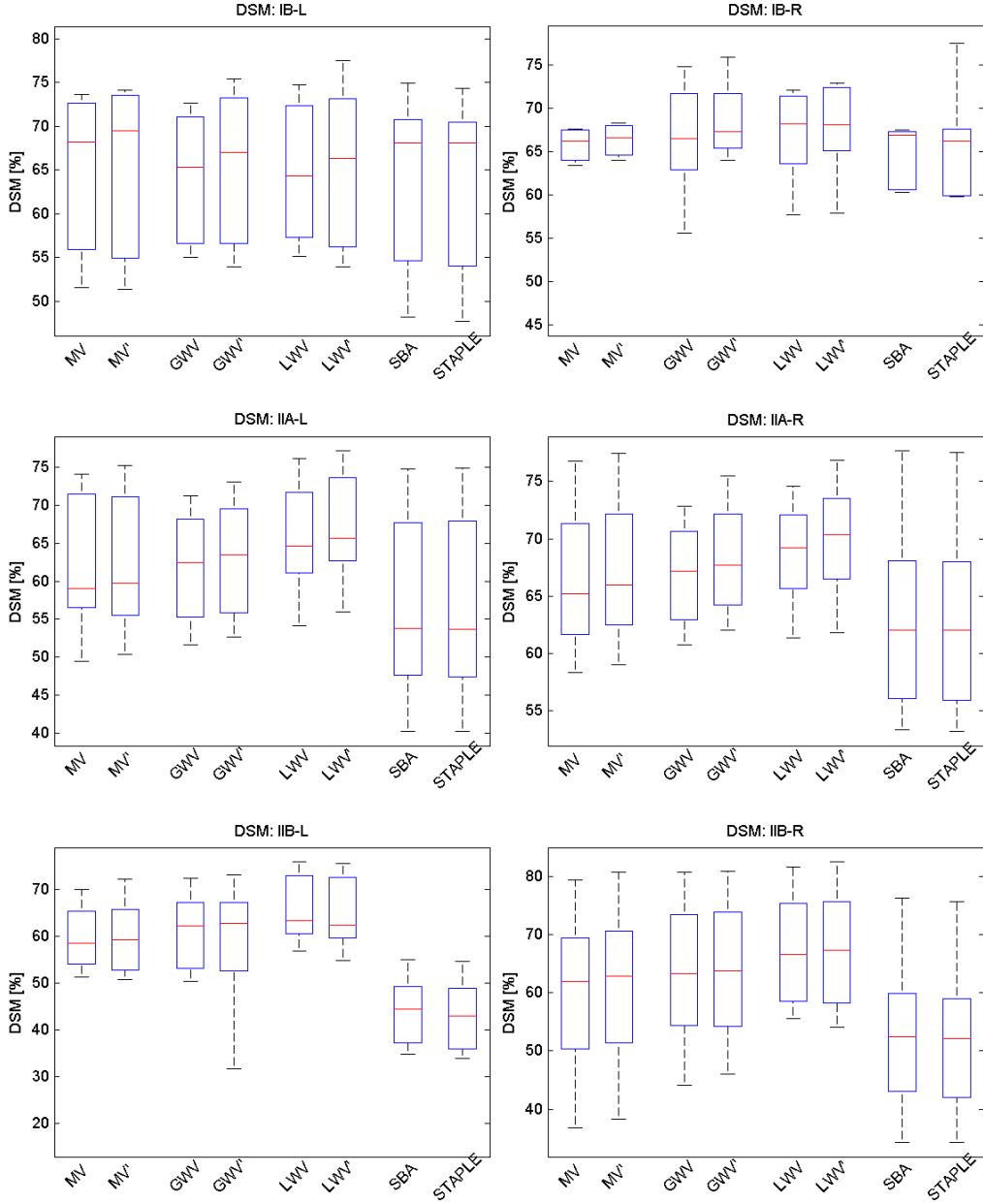
Figure 4.6: Box plots of Dice Similarity Metric (DSM) for the H&N lymph nodes: IB-Left, IB-Right, IIA-Left, IIA-Right, IIB-Left, IIB-Right. These statistics are obtained by the fusion of segmentation results from 12 atlases using different methods. In the above figures, methods using MRF-based smoothness term are denoted with an apostrophe at the end of the method name.
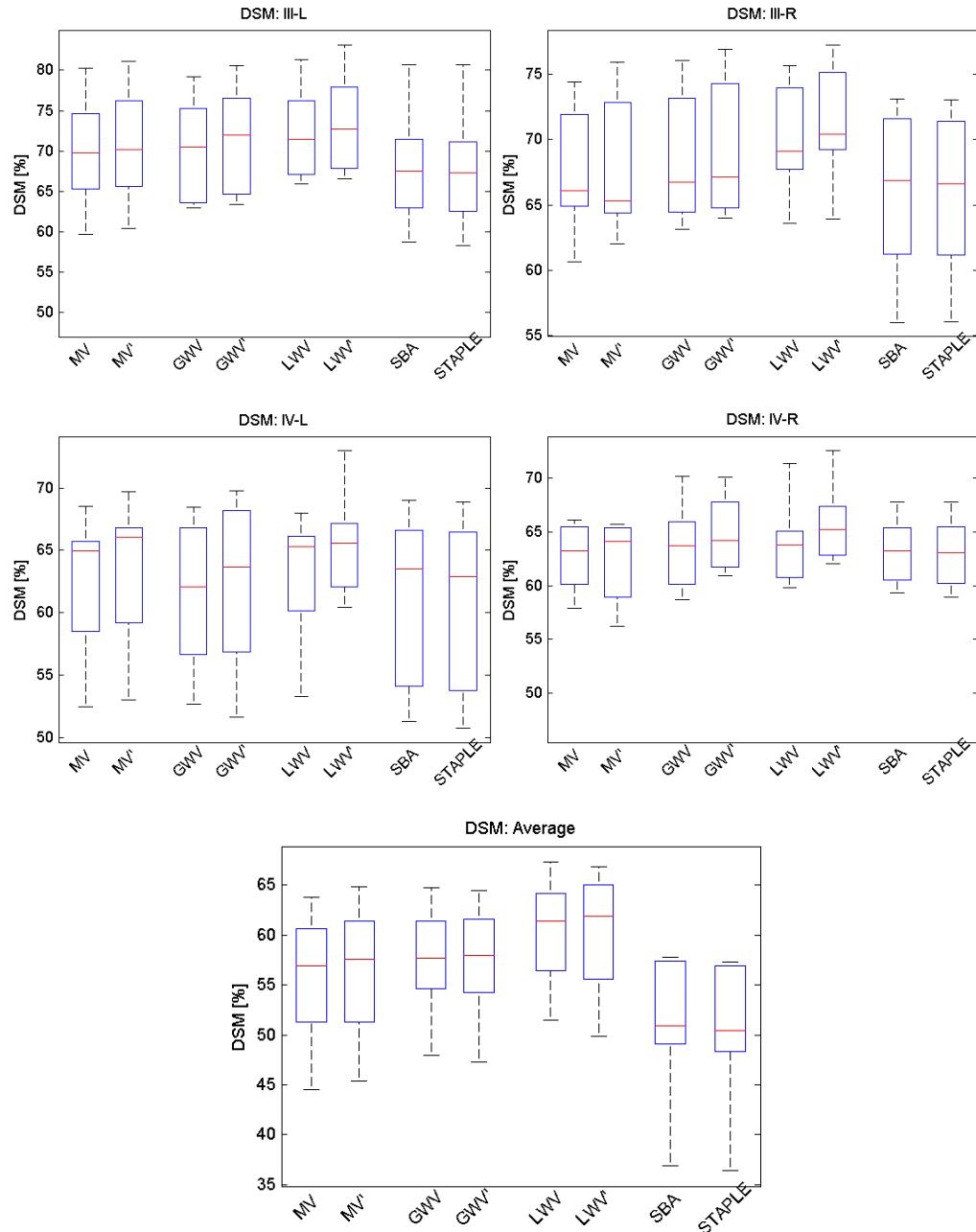
Figure 4.7: Box plots of Dice Similarity Metric (DSM) for the H&N lymph nodes: III-Left, III-Right, IV-Left, IV-Right and average DSM value computed over all structures. These statistics are obtained by the fusion of segmentation results from 12 atlases using different methods. In the above figures, methods using MRF-based smoothness term are denoted with an apostrophe at the end of the method name.
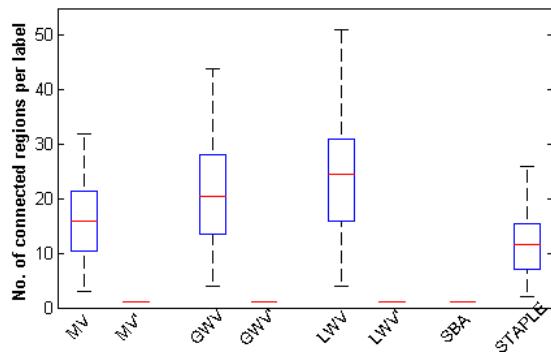
Figure 4.8: Box plots of number of connected regions per lymph node. These statistics are computed over all the 10 lymph nodes, across all the patients, and while merging results from 12 atlases. In the above figure, methods using MRF-based smoothness term are denoted with an apostrophe at the end of the method name.

| (a) Gold stand. | (b) N = 2 | (c) N = 3 | (d) N = 4 | (e) N = 5 | (f) N = 6 |

| (g) N = 7 | (h) N = 8 | (i) N = 9 | (j) N = 10 | (k) N = 11 | (l) N = 12 |
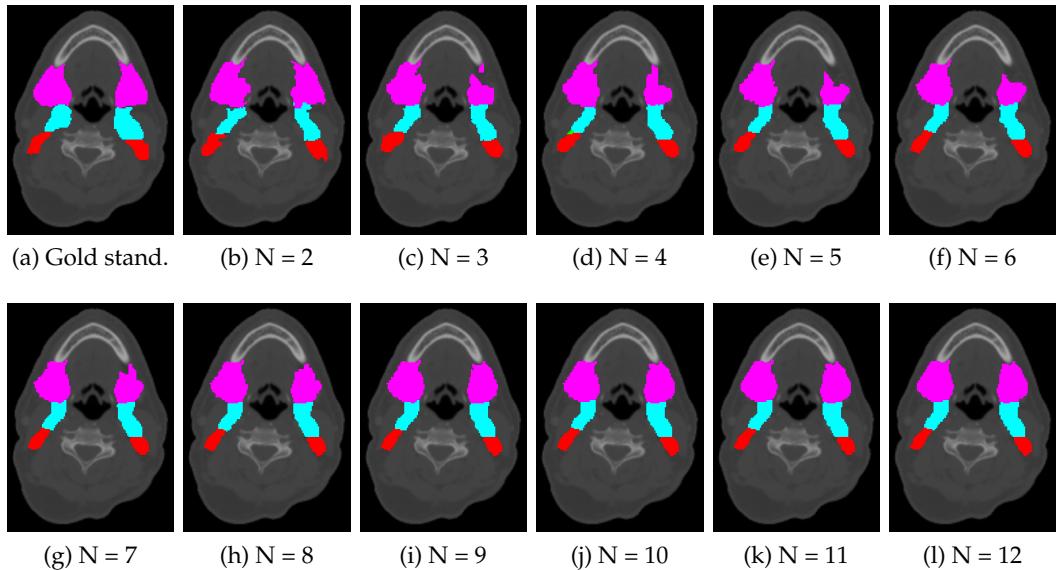
Figure 4.9: Screen shots of H&N lymph nodes segmentations in the axial orientation, from LWV+MRF method, with varying number of atlases.
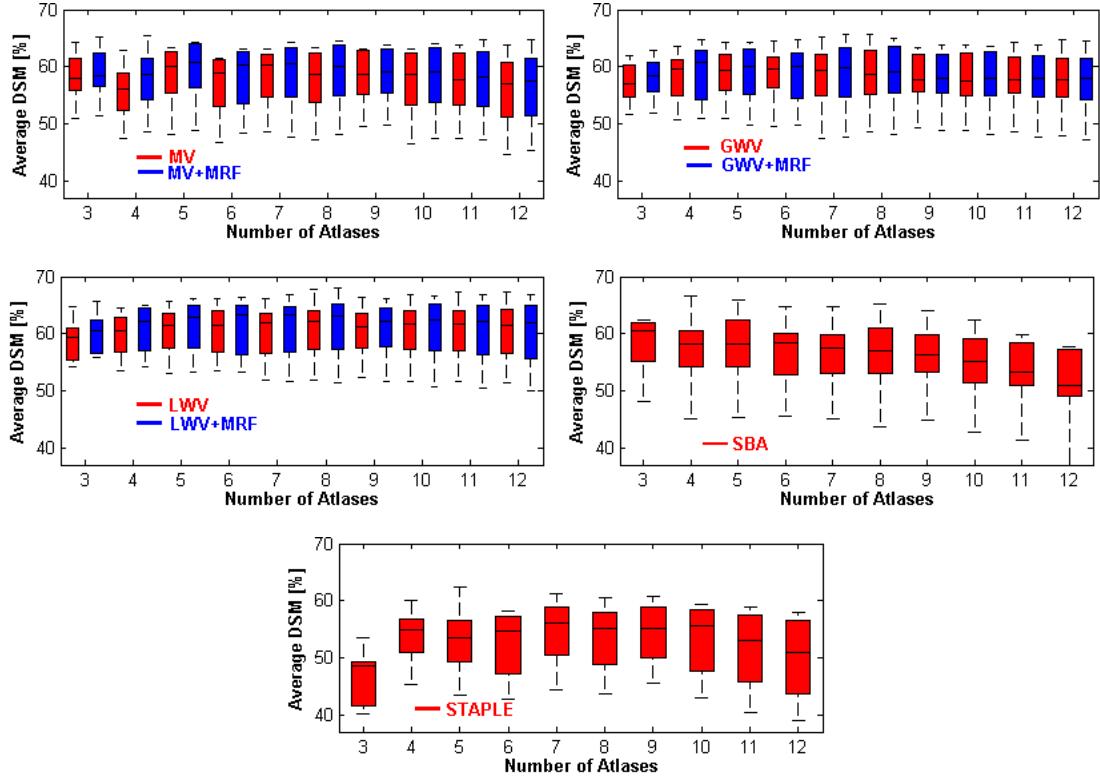
Figure 4.10: Box plots of the average dice similarity metric (DSM) obtained from different fusion methods, for varying number of atlases. Results from methods 'with' and 'without' MRF-based term are shown in the same sub-figure. All sub-figures are plotted with the same range along the Y-axis so that the results across different methods can also be compared visually.

of atlases, the accuracy of segmentation results from those methods that do not take into account any similarity information between the atlas and the target images seem to either fluctuate in a relatively unpredictable manner, or declining. Among all methods, for $N > 4$, the best results are obtained from LWV+MRF; on the other hand, for $N > 4$, the results from STAPLE are the poorest among all the methods, and is followed by SBA.

Finally, figure 4.12 shows mean values of "number of connected regions per label" obtained from different fusion methods, and for varying number of atlases. As mentioned earlier, ideally, a single contiguous region is expected in the output segmentations of each lymph node. In this figure, we did not show the plots for methods that include MRF-based smoothness term since they almost resulted in a single contiguous region per each label irrespective of the number of atlases. Notice from this figure that, the relative trends of the fusion methods in terms of "number of connected regions per label" remained consistent with varying number of atlases. LWV resulted in the highest number of disconnected regions, and is followed by GWV, MV, STAPLE and SBA respectively. As we can see, SBA resulted in very few isolated regions per label even
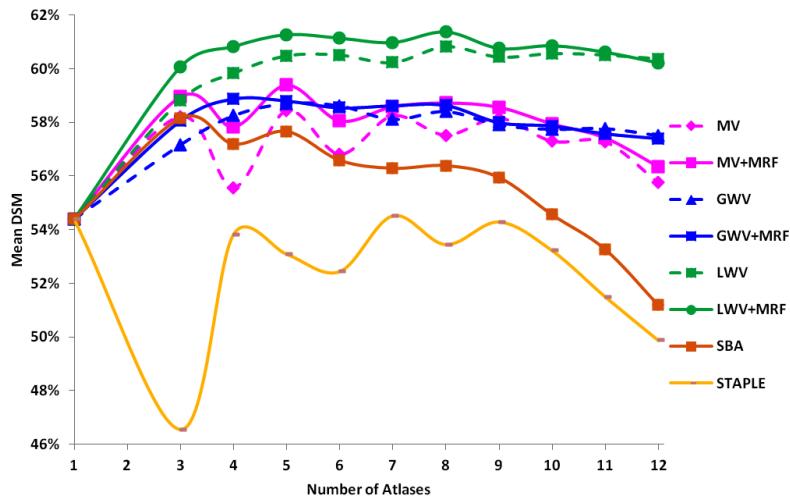
Figure 4.11: Mean values of dice similarity metric (DSM) for different atlas fusion methods, for varying number of atlases.

without the inclusion of any smoothness term.
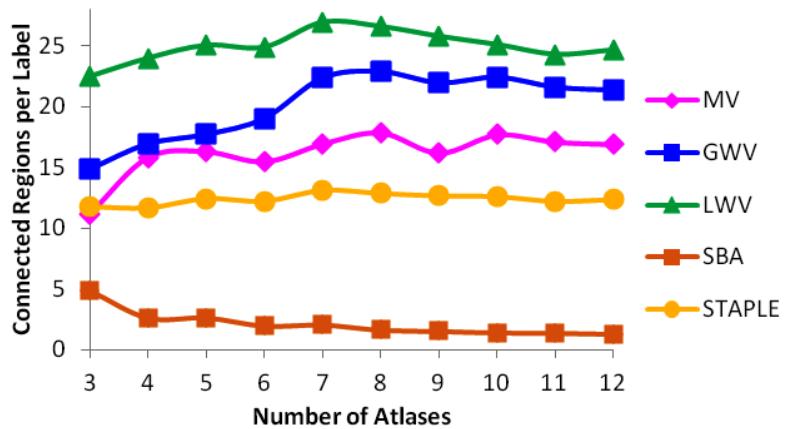
Figure 4.12: Mean values of "number of connected regions per label" for different atlas fusion methods, for varying number of atlases. The plots corresponding to MRF-based methods are not included in this figure since all those methods resulted in a mean value of one, irrespective of the number of atlases.

4.4 Discussion and Conclusions

From the methodological perspective, the main contribution of this chapter is a new MRF-based framework that can simultaneously perform both fusion and edge-preserving smoothing of multiple labels. From the clinical application perspective, the main contribution of this chapter is a comprehensive evaluation of various fusion methods for H&N lymph nodes segmentation in 3D CT images.

We have noticed that the MRF-based minimum energy formulation of fusion methods presented in this chapter has some parallels with the probabilistic framework presented in [38]. One of the main differences compared to the probabilistic framework of [38] is that, unlike our MRF-based framework, it does not include any edge-preserving smoothness term. It would be interesting to further derive an equivalent probabilistic framework for our MRF-based minimum energy formulation.

The evaluations are performed on different fusion methods, both 'with' and 'without' the inclusion of the proposed MRF-based term. In terms of the overlap measure (DSM), among all the fusion methods, local weighted voting (LWV) combined with the edge-preserving term (denoted as LWV+MRF) provided the best segmentation results. It is then followed by LWV, global weighted voting with MRF (i.e., GWV+MRF), GWV, majority voting with MRF (i.e., MV+MRF), MV, shape based averaging (SBA), and STAPLE respectively.

We have also studied the effects of varying the number of atlases on different fusion methods. It is found that, with the increasing number of atlases, for the fusion methods that do not take into account any similarity information between the atlas and the target images (i.e., MV, SBA and STAPLE), the accuracy of segmentation results seems to be fluctuating in a relatively unpredictable manner. On the other hand, for the fusion methods that consider such similarity information (i.e., GWV and LWV), the improvements in the segmentation results with the increasing number of atlases are found to be either increasing or at least stable.

In addition to the overlap measure, the segmentation results obtained from various fusion methods are also evaluated based on the criteria: "how contiguous are the output segmentations without the presence of any unwanted holes and islands?" In this perspective, SBA produced contiguous regions even without the incorporation of any smoothness term. Notice that, unlike voting-based methods, SBA method is based on the distances from each voxel to the contours of each label, and thereby *implicitly* including the neighborhood information of voxels. The evaluation confirmed that such implicit neighborhood information is generally sufficient enough to impose the regularity constraint on labels. On the other hand, for the voting-based method, the MRF-based smoothness term is found to be quite useful in obtaining contiguous as well as more accurate segmentations.

In the next chapter, we reformulate the SBA method using a logistic function-based transformation in order to fit it into our MRF-based fusion framework. We then propose two new fusion methods that extend the SBA method by additionally incorporating the similarity information between the atlas and target images.

5 Weighted Shape Based Averaging Methods

In this chapter, we propose two new fusion methods: "Global Weighted Shape Based Averaging" (GWSBA) and "Local Weighted Shape Based Averaging" (LWSBA). These methods extend the well known Shape Based Averaging (SBA) by additionally incorporating the similarity information between the atlas and the target image to be segmented. In order to fit these methods into our MRF-based fusion framework, we first reformulate the existing SBA method into an equivalent minimization problem, using a logistic function-based transformation. We then propose two new SBA-based methods that additionally incorporate global and local similarity information.

5.1   Introduction

The details of the existing relevant fusion methods are already presented in chapter 2.2. Since the main focus of this chapter is on the "Shape Based Averaging" (SBA) method, we briefly restate here the details of the SBA method. SBA, when compared to the rest of the methods, looks at the fusion problem from a different perspective. Just to give an intuitive interpretation, SBA first computes in all the transformed atlases, how "deep" (inside), or how "faraway" (outside) a given pixel is from the contours of each possible label; this is tracked with the help of "signed Euclidean distance" (SED) measure; it then selects for each pixel a label that has the least resultant SED at that pixel when summed up over all atlases.

We now present the main observations that led us to propose these new methods. First, it is shown in recent works [116, 117] that for certain applications, SBA performs better than some of the widely used methods like "Majority Voting" (MV) and STAPLE; hence, it is worth investigating the possibilities to further improve SBA. Second, notice the parallels between the developments that took place in voting-based fusion methods (i.e., starting from MV, to "Global Weighted Voting" (GWV), and to "Local Weighted Voting"(LWV)) versus SBA: SBA has limitations similar to MV in the sense that, both

methods do not benefit from the information regarding how similar the atlas and the target image are; however, unlike GWV and LWV, no such extensions to SBA are proposed in the literature for incorporating the similarity information. The methods proposed in this chapter are indeed inspired from both these observations.

Before presenting the new fusion methods, as a first step, we reformulate the SBA method in order to fit it into our MRF-based fusion framework.

5.2   Reformulation of SBA

In the preceding chapter, we have shown how MV, GWV and LWV can be reformulated to fit into the data term of the MRF-based framework proposed in equation (4.1). Since the original formulation of SBA [41] can result in negative values as well, it cannot be directly used in the above framework. In order to deal with this problem, we have actually proposed in [46] an approach based on shifting and thresholding the signed distance values. However, we noticed that such approach is not very elegant since it requires additionally, careful selection of the threshold value, and it is also not an exact equivalent to the original SBA formulation. We now present a new logistic function-based approach that results in not only an exact equivalent to original SBA formulation, but also transforms the distances to nonnegative values.

Let N be the number of atlases. Let V be the number of voxels in the target image. Let Y_p denote the label assigned to the pth voxel in the output image. Let Y be the set containing labels assigned to each voxel in the output image, i.e., $Y = \{Y_1, \cdots Y_V\}$. Let X^j represent the jth input labeled image. Let $u_p^j(l)$ represent the SED at p^{th} voxel, in X^j, for label l. In the original formulation of SBA [41], for each voxel p, it assigns independently, that label which results in the minimum value of the following summation.

$$Y_p = \arg\min_l \frac{1}{N} \sum_{j=1}^{N} u_p^j(l). \tag{5.1}$$

The above equation can also be written as a minimization problem of the following data term.

$$E_{\text{data}}(Y) = \frac{1}{N} \sum_{p=1}^{V} \sum_{j=1}^{N} u_p^j(l). \tag{5.2}$$

However, as mentioned earlier, u_p^j can take any real value, and the goal here is to apply a transformation that maps these values to nonnegative values while retaining all the desired properties. Our following approach is primarily inspired by the work of [118]

where they use Logarithm Odds maps, in a different context, for shape representation [1]. The standard logistic function $P(.)$ that maps a variable $u_p^j \in \mathbb{R}$ to $t_p^j \in (0,1)$ is given by:

$$t_p^j(l) = P(u_p^j(l)) \triangleq \frac{1}{1 + \exp(-u_p^j(l))}. \tag{5.3}$$

Notice that the above logistic function maps negative SEDs to the range $(0, 0.5)$, zero SED (i.e., voxels on the contours of the same label) to 0.5, and positive SEDs to $(0.5, 1)$. It is easy to derive the following two properties of the logistic function-based transformation which we use in our reformulation. First, addition of SEDs is equivalent to the following operation in the logistic transformation space:

$$\sum_{j=1}^{N} u_p^j(l) \leftrightarrow \frac{\prod_{j=1}^{N} t_p^j(l)}{\prod_{j=1}^{N} t_p^j(l) + \prod_{j=1}^{N} \left(1 - t_p^j(l)\right)}. \tag{5.4}$$

Second, multiplication by a scalar (α) in the SED space is equivalent to the following operation in the logistic transformation space, which we will denote as $\alpha \otimes t_p^j$ for the notational convenience:

$$\alpha \, u_p^j(l) \leftrightarrow \alpha \otimes t_p^j(l) \triangleq \frac{\left(t_p^j(l)\right)^{\alpha}}{\left(t_p^j(l)\right)^{\alpha} + \left(1 - t_p^j(l)\right)^{\alpha}}. \tag{5.5}$$

Finally, using the property in equation (5.4), the original SBA minimization problem in equation (2.6) can be equivalently reformulated as the following data term of equation (4.1).

$$E_{\text{data}}(Y) = \frac{1}{N} \sum_{p=1}^{V} \left(\frac{\prod_{j=1}^{N} t_p^j(l)}{\prod_{j=1}^{N} t_p^j(l) + \prod_{j=1}^{N} \left(1 - t_p^j(l)\right)} \right). \tag{5.6}$$

5.3 New Fusion Methods

5.3.1 Global Weighted Shape Based Averaging

As mentioned in the introduction of this chapter, SBA, similar to MV, does not take into account any similarity information of the atlases to the target image. Our goal here is,

---

[1]We would like to thank Dr. Bennett Landman (Assistant Professor in the Electrical Engineering Department of Vanderbilt University) for mentioning us about this reference paper.

similar to GWV, to scale the contributions coming from various atlases in accordance to their global similarity to the target image. As we introduce the global similarity-based weighting into SBA, we call this new method as "Global Weighted Shape Based Averging" (GWSBA). This results in the following minimum energy formulation.

$$E_{\text{data}}(Y) = \frac{1}{N} \sum_{p=1}^{V} \sum_{j=1}^{N} \hat{w}^j \, u_p^j(l), \tag{5.7}$$

where $\hat{w}^j$ is the normalized global weight assigned to the jth atlas. This equation can be reformulated using the properties mentioned in equation (5.4) and equation (5.5), and it consequently results in the following minimization equation.

$$E_{\text{data}}(Y) = \frac{1}{N} \sum_{p=1}^{V} \left(\frac{\prod\limits_{j=1}^{N} \hat{w}^j \otimes t_p^j(l)}{\prod\limits_{j=1}^{N} \hat{w}^j \otimes t_p^j(l) + \prod\limits_{j=1}^{N} \hat{w}^j \otimes (1 - t_p^j(l))} \right). \tag{5.8}$$

The global weights assigned to each atlas (i.e., w^j) are computed similar to GWV (as described in chapter 2.2.2). Thus, when mean square difference (MSD) of intensities between the transformed atlas and the target image is used as the similarity metric, w^j is given by

$$w^j = \frac{1}{\sum_{p=1}^{V} \left(I_p^j - I_p^T \right)^2 + \epsilon}. \tag{5.9}$$

Finally, the normalization of global weights assigned to each atlas is also performed similar to GWV (as in equation 4.6), and is given by

$$\hat{w}^j = \frac{w^j}{\max \left(w^1, \cdots, w^N \right)}. \tag{5.10}$$

5.3.2 Local Weighted Shape Based Averaging

Notice that GWSBA assigns a single *global* weight for each atlas as a measure of the similarity. However, although two images may differ significantly in some particular regions, it is possible that they may be very similar in some other regions, and "Local Weighted Shape Based Averaging" (LWSBA) is indeed based on this observation. Instead of assigning a single similarity measure for the entire image (like in GWSBA), LWSBA computes similarity measure *locally* for each voxel, within a specified neighborhood. So,

the minimum energy formulation is given by

$$E_{\text{data}}(Y) = \frac{1}{N} \sum_{p=1}^{V} \sum_{j=1}^{N} \hat{w}_p^j \, u_p^j(l), \tag{5.11}$$

where $\hat{w}_p^j$ is the normalized global weight assigned to the jth atlas at the pth voxel. Please note that, except in the computation of weight, LWSBA is similar to GWSBA, and thus, the final formulation of LWSBA is as follows:

$$E_{\text{data}}(Y) = \frac{1}{N} \sum_{p=1}^{V} \left(\frac{\prod_{j=1}^{N} \hat{w}_p^j \otimes t_p^j(l)}{\prod_{j=1}^{N} \hat{w}_p^j \otimes t_p^j(l) + \prod_{j=1}^{N} \hat{w}_p^j \otimes (1 - t_p^j(l))} \right). \tag{5.12}$$

The local weights assigned for each voxel in each atlas (i.e., w_p^j) are computed similar to LWV (as described in chapter 2.2.3). Thus, when mean square difference (MSD) of intensities between the transformed atlas and the target image is used as the similarity metric, w_p^j is given by

$$w_p^j = \frac{1}{\sum_{\forall q \in \aleph_p} \left(I_q^j - I_q^T \right)^2 + \epsilon}. \tag{5.13}$$

Finally, the normalization of local weights assigned is also performed similar to LWV (as in equation 4.9), and is given by

$$\hat{w}_p^j = \frac{w_p^j}{\max \left(w_p^1, \cdots, w_p^N \right)}. \tag{5.14}$$

5.4 Discussion and Conclusions

In this chapter, first, we have reformulated the existing shape based averaging (SBA) method using a logistic function-based transformation in order to fit it into our MRF-based fusion framework. Then, we have proposed two new label fusion methods (named - GWSBA and LWSBA) that extend the existing shape based averaging by additionally including similarity information between the atlas and the target images.

We have noticed that in a very recent work [119], a geodesic extension of SBA is proposed in the context of DTI tractography that includes a similarity measure into SBA. However, the goals and formulations of their work are different from our current work. Further, unlike the formulation proposed by [119], our formulation also facilitates inclusion of more prior knowledge into these fusion methods, and this is presented in detail in the next chapter.

The new fusion methods proposed in this chapter will be evaluated in the next chapter; the results obtained from these methods will be compared with the results from the existing SBA method in the context of head & neck lymph nodes segmentation.

6 Neighborhood Prior Model for Atlas Fusion

In this chapter, we propose a new neighborhood prior model that incorporates spatially-varying similarity-weighted neighborhood prior information for each possible label pair combination. This neighborhood prior model is integrated with our MRF-based fusion framework; hence, this model can be combined with any of the existing fusion methods that are reformulated in the preceding chapters to fit into our MRF-based fusion framework, and also with the newly proposed SBA-based fusion methods (i.e., with GWSBA and LWSBA methods). Finally, we present an evaluation of SBA, GWSBA and LWSBA methods, both 'with' and 'without' using the neighborhood priors, in the context of head and neck (H&N) lymph nodes segmentation.

6.1   Introduction

The very goal of atlas-based segmentation is to exploit the prior knowledge coming from a similar reference image (i.e., the atlas) regarding the structures of interest to be segmented. It is a proven fact in practice that the segmentations obtained by merging results from multiple atlases are more robust and accurate than the single atlas-based segmentation results. However, notice that certain prior information (or desirable properties) that the structures of interest possess in the segmentations obtained from individual atlases could be lost during the process of merging segmentations from multiple atlases. For example, we have identified and addressed one such problem in chapter 4, that, although the segmentations obtained from each individual atlas are contiguous, the merged segmentations could be fragmented with unwanted holes and islands. Thus, we realize that in addition to incorporating maximum possible prior knowledge into the atlas-based segmentation procedure, it is also equally important, if not more, to retain and incorporate more prior knowledge into the fusion procedure as well.

We notice that, from all the transformed atlases and their respective output seg-

mentations, we can indeed extract certain prior information regarding the probable neighborhood labels at each voxel location, and for each possible label pair. However, to the best of our knowledge, none of the existing fusion methods benefit from this potentially useful information regarding the probable label-pair combinations at each voxel.

The objective of this chapter is therefore to first propose a model that can incorporate such neighborhood labels information, then to integrate that model with the existing fusion methods, and finally, to evaluate it in the context of H&N lymph nodes segmentation application.

6.2 New Neighborhood Prior Model

Let $\aleph_p$ be the set of all voxels in the predefined neighborhood of the pth voxel. Let $|\aleph_p|$ represent the cardinality of $\aleph_p$. Let $\delta(.)$ be the standard Kronecker delta function. Let X_p^j represent the label assigned to the pth voxel in the jth transformed atlas. Similarly, let $\hat{w}_p^j$ represent the normalized local weight assigned to the pth voxel in the jth transformed atlas. Let Y be the set containing labels assigned to each voxel in the output image, i.e., $Y = \{Y_1, \cdots Y_V\}$. The E_{prior} term that we propose in this thesis is as follows.

$$E_{\text{prior}}(Y) = \sum_{p=1}^{V} \sum_{\substack{\forall q \in \aleph_p \\ q \neq p}} \frac{1}{|\aleph_p|} \zeta(p, q, Y_p, Y_q), \tag{6.1}$$

where

$$\zeta(p, q, Y_p, Y_q) = \begin{cases} \infty, & \text{if } \sum\limits_{j=1}^{N} \delta(X_p^j, Y_p) = 0, \\ 1 - \left(\dfrac{\sum\limits_{j=1}^{N} \hat{w}_p^j \, \delta(X_p^j, Y_p) \, \delta(X_q^j, Y_q)}{\sum\limits_{j=1}^{N} \hat{w}_p^j \, \delta(X_p^j, Y_p)} \right), & \text{otherwise.} \end{cases} \tag{6.2}$$

We now present an intuitive interpretation of the above formulation.

- The first case of the neighborhood prior function $\zeta(.)$ in the above equation implies that, if a label Y_p never occured at the pth voxel in any of the N transformed atlases (i.e., for a given p value, when $\delta(X_p^j, Y_p) = 0$ for all values of j), then we set the value of $\zeta(.)$ to ∞. In other words, we are first imposing the constraint that, in order to assign a label under consideration to the pth output voxel, that particular label should have occurred at least in one of the N transformed atlases at that voxel location.

- Regarding the second case of equation (6.2), notice that, if we ignore the similarity weights (i.e., $\hat{w}_p^j$) for a moment, the summation in the denominator of this equation computes the number of times the label Y_p occurred at the pth voxel, among all the N transformed atlases; in the same way, the summation in the numerator computes the number of times Y_q occurred at X_q^j when Y_p occurred at X_p^j; $\hat{w}_p^j$ is introduced to simply scale the contributions coming from individual atlases based on their respective local similarities to the target image. Finally, notice that the value of the ratio of the summations in this equation is always in the range $[0, 1]$, and it's value is maximum when the output label pair (Y_p, Y_q) to be assigned coincides with the label pair that occurred the most among the transformed atlases, at (p, q) voxels; hence, this ratio is subtracted from 1 in order to transform it into an equivalent minimization problem so that this term can be later integrated with our MRF-based fusion framework.

The normalized similarity weights (i.e., $\hat{w}_p^j$) in equation (6.2) are computed similar to the weights in "Local Weighted voting" (LWV) method, as we described in chapter 4.2.3, and we repeat those equations here for convenience.

$$\hat{w}_p^j = \frac{w_p^j}{\max\left(w_p^1, \cdots, w_p^N\right)},\tag{6.3}$$

where

$$w_p^j = \frac{1}{\sum\limits_{\forall q \in \aleph_p} \left(I_q^j - I_q^T\right)^2 + \epsilon}.\tag{6.4}$$

As mentioned earlier, our objective is to integrate this neighborhood prior model with our MRF-based fusion framework (proposed in chapter 4) so that, the existing fusion methods could benefit from this additional prior information as well. For this purpose, we propose to modify the general energy minimization problem presented in equation (4.1) as follows.

$$Y^* = \arg\min_Y \left\{E_{\text{data}}(Y) + \lambda\, E_{\text{smooth}}(Y) + \beta\, E_{\text{prior}}(Y)\right\},\tag{6.5}$$

where $E_{\text{prior}}(Y)$ is given by equation (6.1) and β is a nonnegative weighting parameter for the neighborhood prior term.

Please note that, $E_{\text{data}}(Y)$ is a unary term whereas $E_{\text{smooth}}(Y)$ and $E_{\text{prior}}(Y)$ are pairwise terms. Hence, while solving this energy minimization problem using graph cuts [49, 105, 106], $E_{\text{data}}(Y)$ contributes to the weights of the data-links of the graph whereas $(\lambda\, E_{\text{smooth}}(Y) + \beta\, E_{\text{prior}}(Y))$ contributes to the neighborhood-links of the graph.
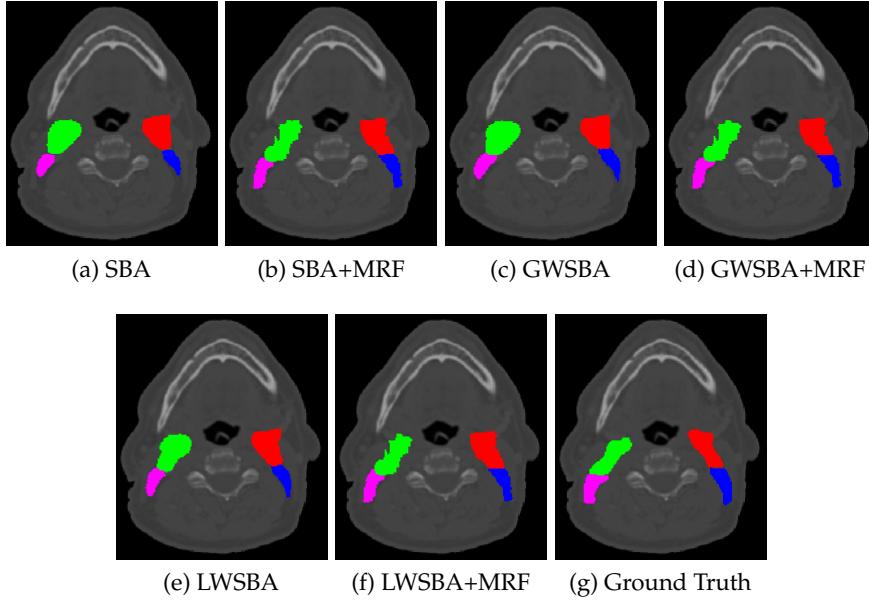
Figure 6.1: H&N lymph nodes segmentations obtained from different fusion methods and the corresponding ground truth are shown in an axial slice, for one of the target images.

Finally, in order to have a relatively easy interpretation of the weighting parameters to be set, we rewrite the equation (6.5) as follows.

$$Y^* = \arg \min_{Y} \left\{ E_{\text{data}}(Y) + \beta \left(E_{\text{prior}}(Y) + \xi \, E_{\text{smooth}}(Y) \right) \right\}, \tag{6.6}$$

where $\xi = \frac{\lambda}{\beta}$. The above equation can be interpreted as follows. $E_{\text{data}}(Y)$ is a unary term representing any of the fusion methods. $\left(E_{\text{prior}}(Y) + \xi \, E_{\text{smooth}}(Y) \right)$ is a pairwise term that is newly introduced in this thesis. β represents the weight assigned to this pairwise term as a whole. ξ can be interpreted as a parameter that determines the relative weight given to the neighborhood prior term (i.e., $E_{\text{prior}}(Y)$) when compared to the edge-preserving smoothness term (i.e., $E_{\text{smooth}}(Y)$). For example, setting ξ value to 0.4 can be interpreted as, if the weight given to the neighborhood prior term is treated as 100%, the relative weight given to the smoothness term is 40%. We hereafter refer to and use equation (6.6) instead of equation (6.5) because of the ease of interpretation it offers regarding the weighting parameters.

6.3   Evaluation

The purpose of the evaluations presented in this section is twofold. First, we evaluate the two new SBA-based methods (viz, GWSBA and LWSBA) that we have proposed in the preceding chapter by comparing the segmentation results obtained from them with

(a) SBA   (b) SBA+MRF   (c) GWSBA   (d) GWSBA+MRF

(e) LWSBA   (f) LWSBA+MRF   (g) Ground Truth
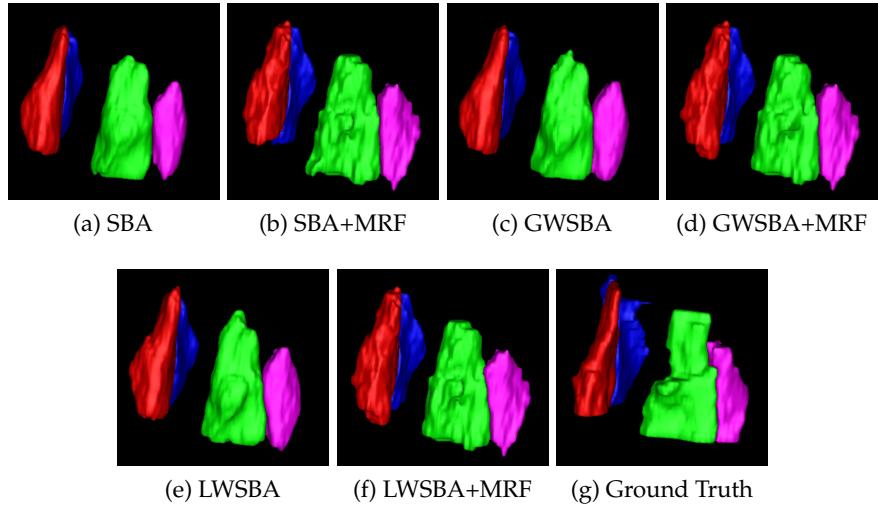
Figure 6.2: H&N lymph node volumes obtained from different fusion methods and the corresponding ground truth volumes are presented for one of the target images.

Table 6.1: Mean and standard deviations of Dice Similarity Metric (DSM) values (in %) for each lymph node. Average values of DSM computed across all the 4 lymph nodes, obtained from each method are also presented.

Method	DSM values in % (mean±std)				
	IIA-Left	IIA-Right	IIB-Left	IIB-Right	Average
SBA	56.19±11.97	62.74±7.76	40.58±11.42	52.39±12.34	55.06±9.26
SBA'	63.13±11.95	68.85±3.65	60.39±16.21	64.42±14.34	66.09±6.52
GWSBA	59.66±9.56	65.60±6.09	46.48±10.28	57.90±12.15	59.43±7.97
GWSBA'	64.81±10.31	68.71±3.61	60.12±18.47	64.23±14.65	66.50±6.33
LWSBA	63.27±9.01	67.10±6.11	50.96±9.36	60.22±11.60	62.18±7.30
LWSBA'	63.98±11.30	68.60±3.59	60.04±18.59	64.21±14.52	66.24±6.53

the results from the existing SBA method. Second, we evaluate the effects of using the pairwise terms proposed in this thesis with SBA, GWSBA and LWSBA methods.

The evaluations are performed in the context of segmentation of lymph nodes in the 3D Head and Neck (H&N) CT images. We consider 4 important lymph nodes for automated segmentation: (i) IIA-Left, (ii) IIA-Right, (iii) IIB-Left and (iv) IIB-Right. The current dataset contains 12 atlas images and 8 patients' images to be segmented, with a typical resolution is 1.5mm×1.5mm×3mm in X, Y and Z directions respectively. An expert oncologist has manually delineated lymph nodes on all the images, and those manual delineations are considered as the ground truth segmentations.

Regarding the registration, all the 12 atlases are registered to each patient to be segmented. An initial affine registration is performed followed by a two-level hierarchical
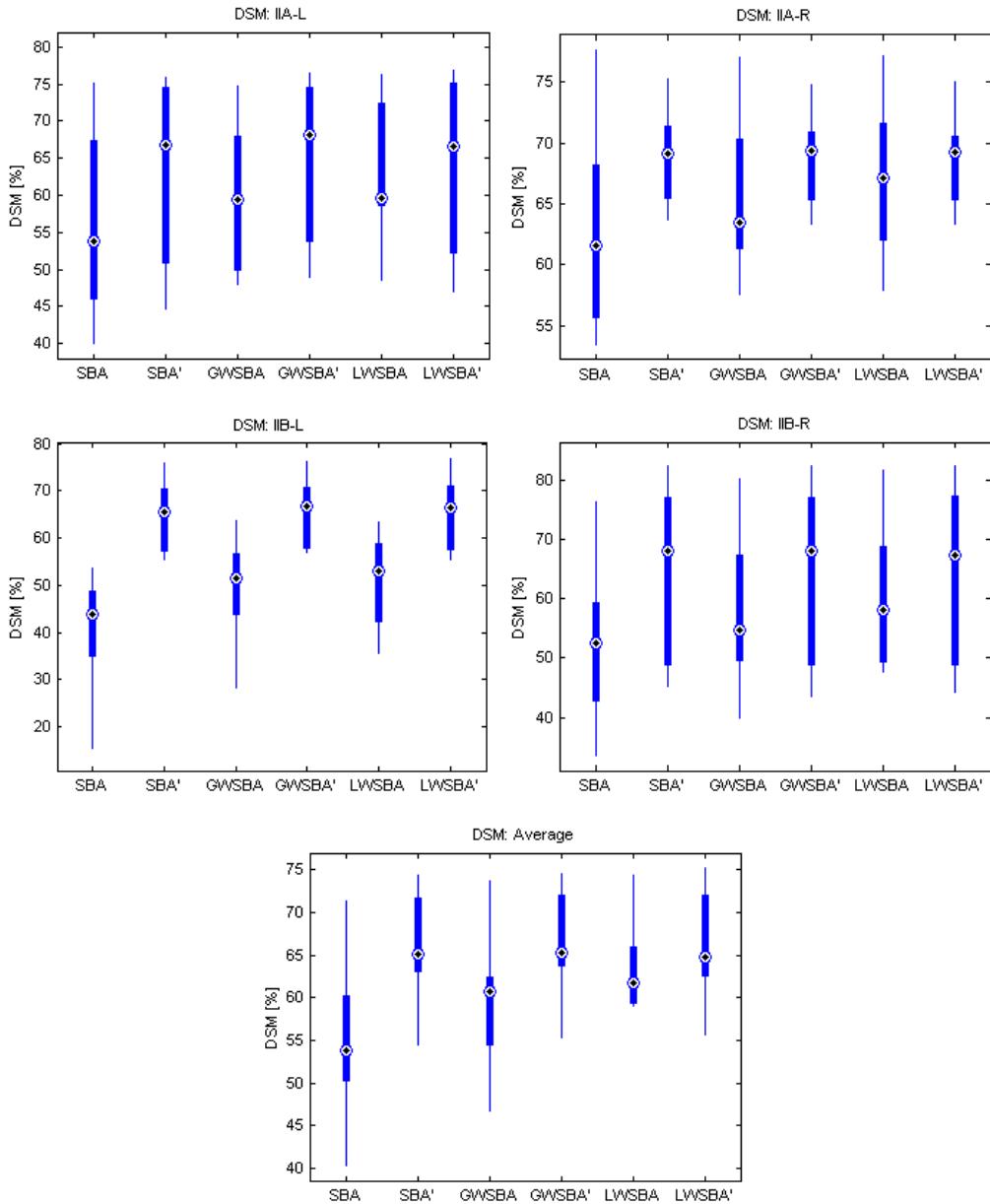
Figure 6.3: Box plots of Dice Similarity Metric (DSM) for the H&N lymph nodes: IIA-Left, IIA-Right, IIB-Left, IIB-Right are presented in the first and second rows. Box plots of average DSM computed across all the four lymph nodes are presented in the last row. Methods using additional neighborhood prior term and smoothness term are suffixed with an apostrophe.

nonrigid registration. In the first level, region-based registration forces are used, and it is then followed by a pixel-based registration. The registration procedure and parameters values for registration are exactly same as those presented in chapter 4.3.

SBA, GWSBA and LWSBA methods are evaluated both 'with' and 'without' the smoothness and neighborhood prior terms (i.e., respectively with $\beta \neq 0$ and $\beta = 0$ in equation (6.6)); methods with E_{prior} and E_{smooth} terms are denoted either with a suffix "+MRF", or an apostrophe (') added to their names. When using the MRF-based pairwise terms, β, ζ and $\aleph_p$ values are empirically set to 3.0, 0.4 and to the standard 3D grid of 6 neighbors respectively. As mentioned before, setting ζ value to 0.4 means: to give for the E_{prior} term, 40% of the relative weight assigned to the E_{prior} term. These values are selected based on visual inspection of the results, and are not optimized further; such empirical selection is however sufficient enough for the current context. Finally, the normalized similarities between each transformed atlas and the target image are computed based on the mean square difference (MSD) of intensities using equation (6.3) and equation (6.4).

Figure 6.1 shows, for one of the target images, in one of the axial slices, the results obtained by merging segmentations from all 12 atlases using different fusion methods, and the ground truth segmentations. Similarly, figure 6.2 shows, for the same target image, the lymph node volumes constructed from different fusion methods, and from the ground truth segmentations. The quantitative evaluation is performed over the entire data set, using "Dice Similarity Metric" (DSM) described in chapter 2.4. Figure 6.3 presents box plots of DSM for each lymph node structure, and also the box plots of average DSM computed across all the four lymph nodes. Finally, table 6.1 presents the corresponding mean and standard deviation values of DSM (in %) for each lymph node.

It can be noted from the above results that the proposed GWSBA and LWSBA methods have resulted in more accurate segmentations than the existing SBA. We can also notice that the inclusion of MRF-based pairwise terms have significantly improved the results when compared to their counterparts that do not use such information. Another interesting observation from these results is that, while there is a significant difference among SBA, GWSBA and LWSBA in terms of segmentation accuracy, all these methods resulted in similar accuracy with the inclusion of the proposed pairwise terms.

We further evaluated the statistical significance of improvements using Wilcoxon signed-rank tests, performed on the segmentation results of 4 lymph nodes, in 8 images, and from 6 methods. It is found that the improvements in DSM from SBA to SBA+MRF are statistically significant with a p-value of 1.6e-5; the improvements from GWSBA to GWSBA+MRF are significant with a p-value of 2.8e-4; the improvements from LWSBA to LWSBA+MRF are significant with a p-value of 0.0047; the improvements from SBA to GWSBA are significant with a p-value of 2.3e-09; finally, the improvements from SBA to LWSBA are also statistically significant with a p-value of 3.2e-08.

6.4   Discussion and Conclusions

In this chapter, we have proposed a new neighborhood prior term that models spatially, neighborhood information for each possible label pair combinations. The proposed model also takes into account the local similarities between each atlas and target images. This model is integrated with our MRF-based fusion framework so that it can be used together with any of the fusion methods like MV, GWV, LWV, SBA, GWSBA and LWSBA.

The evaluations in this chapter are performed in the context head and neck lymph nodes segmentation, with two specific objectives in mind; first, for the two new SBA-based methods proposed in the preceding chapter (viz, GWSBA and LWSBA), we wanted to evaluate the improvements in accuracy when compared to the original SBA method; second, we wanted to simultaneously investigate the effects of incorporating the neighborhood prior term proposed in this chapter with SBA, GWSBA and LWSBA methods.

With regard to the first objective, it is found that the two newly proposed SBA-based fusion methods have significantly improved the accuracy of the lymph nodes segmentations when compared to the original SBA method. For instance, for the SBA method, the mean value of the average DSM computed across all the 4 lymph nodes is around 55% whereas for the GWSBA and LWSBA methods, the mean DSM values are around 59% and 62% respectively. The relative performance of these three methods across all the four lymph nodes is consistent; among the three methods, LWSBA provided the best segmentation results, and it is followed by GWSBA and SBA respectively.

With regard to the second objective, it is found that the inclusion of the proposed pairwise term has significantly improved the accuracy of the lymph nodes segmentations when compared to their respective counterparts that do not use any such prior information. For instance, for the SBA method, with the inclusion of this additional term, the average DSM value has increased from around 55% to 66%; similarly, for the GWSBA method, the average DSM has increased from 59% to 67%, and for LWSBA method, DSM has increased from 62% to 66%. Hence, for all the three methods, across all the four lymph nodes, the inclusion of the proposed pairwise term has always resulted in improving the segmentation accuracy. Another interesting observation is that, although there is a significant difference in segmentation accuracy among SBA, GWSBA and LWSBA, with the inclusion of the proposed pairwise terms, all these three methods resulted in similar accuracies.

Finally, in the future work, we would like to extend this evaluation to MV, GWV and LWV methods combined with the proposed neighborhood prior term. We would like to also study the effects of varying the weighting parameters on the final output segmentations obtained from each fusion method.

7 Conclusions and Future Work

Conclusions

The primary objective of this thesis is to develop new methodologies that are aimed towards improving the accuracy of the atlas-based segmentations, and to evaluate them in clinically important challenging applications like segmentation of lymph nodes in the head and neck CT images.

To fulfil the above objective, we focussed on two crucial aspects of atlas-based segmentation that can have a significant impact on the final output segmentations; the first aspect that we considered is the registration method to be used for computing the point-by-point mapping between the atlas (i.e., the reference image) and the target image to be segmented; the second aspect is the fusion methodologies used for merging the segmentation results obtained from multiple individual atlases. While chapter 3 presents our contributions regarding the registration methods, chapter 4, chapter 5 and chapter 6 present our contributions regarding fusion methods.

In chapter 3, we presented a new variational framework that generalizes and combines some of the existing state-of-the-art registration methods. In particular, it combines the registration forces coming from the region-based active contour (AC) framework and the pixel-based optical flow framework. A new label function that can distinguish any number of regions with just a single function is also presented. Another advantage of our framework is that it facilitates selecting specific structures and/or registration forces for driving the registration process. The first set of evaluations confirmed the advantage of combining multiple registration forces, which became possible because of the proposed framework. In the second set of experiments, we compared the segmentation results obtained from our registration method with two other state-of-the-art methods, and these evaluations clearly demonstrated the potential of our framework.

In chapter 4, we addressed one of the main issues that is encountered by most of the

existing fusion methods: although the segmentations obtained from each individual atlas are contiguous, the merged segmentations could be fragmented with unwanted holes and islands. In order to deal with this problem, we proposed a general MRF-based framework that performs simultaneously, edge-preserving smoothing and fusion of multiple labels. We then reformulated many existing fusion methods so that they can benefit from this framework. It is found that the proposed MRF-based approach succeeded in generating contiguous as well as more accurate automated segmentations.

In chapter 5, we further reformulated one more widely used fusion method, called "shape based averaging" (SBA) method. The reformulation is done using a logistic function-based transformation. We then proposed two new SBA-based fusion methods (called GWSBA and LWSBA) that additionally incorporate the similarity information between each atlas and the target images into the existing SBA method. These two new methods are also fitted into our MRF-based framework. The evaluation of these methods is later performed in the next chapter; GWSBA and LWSBA methods are found to improve the accuracy significantly when compared to the original SBA method.

In chapter 6, we proposed a new prior term that spatially models the neighborhood prior information for each possible label pair combination. This prior term also takes into account the local similarities between each atlas and the target image. This spatially-varying similarity-weighted neighborhood prior term is then integrated with our previously proposed MRF-based fusion framework. We investigated the impact of this neighborhood prior model on SBA, GWSBA, and LWSBA methods; it is found that the neighborhood prior term has improved the segmentation results significantly when compared to their respective counterpart fusion methods that do not include any such information.

Finally, we would like to summarize here the results obtained in the specific context of head and neck (H&N) lymph nodes segmentation application. We performed detailed evaluations for analyzing the impact of simultaneously performing both fusion and edge-preserving smoothing using our MRF-based fusion framework. In this evaluation, we considered majority voting (MV), global weighted voting (GWV), local weighted voting (LWV), SBA and STAPLE methods. Among all these methods, the best segmentation results are obtained from the LWV method that was additionally incorporating the proposed MRF-based smoothness term. We later compared the lymph nodes segmentations obtained from the proposed GWSBA and LWSBA methods with the existing SBA method. Among these three methods, LWSBA provided the best segmentation results, and is followed by GWSBA and SBA respectively. Incorporation of the proposed neighborhood prior term has further improved the accuracy of these three methods significantly.

Future Work

We now present some possible directions of the future research that we have identified. We group them here into two categories: one from the general view point, and the other in the specific context of the H&N lymph nodes segmentation application.

Possible Future Work in General:

1. The approach that we currently follow for the deformation field computation is similar to the "additive demons iterations scheme" of Thirion's approach [16]. They are similar in terms of the way the deformation field is updated, and the type of iterative scheme. While such approach is good enough for small deformation framework, it may not be an accurate method in case of large deformations. In the framework of registration containing large deformations, diffeomorphisms are powerful and mathematically elegant transformations that ensure a one-to-one smooth and continuous mapping with nonsingular Jacobian determinant; they guarantee both invertibility and preservation of topology. A diffeomorphic model has been proposed in recent years for the Demons algorithms [17]. We could also probably adapt a similar approach to our framework to deal with large deformations.

2. In this thesis, many fusion methods are formulated directly as energy minimization problems. Such a direct formulation instead of beginning with a probabilistic formulation is indeed not an uncommon practice in many computer vision problems [105]. It will be however interesting, at least from the theoretical perspective, to also formulate them as equivalent maximum a posteriori estimations of MRF, and analyze their properties.

3. The MRF-based fusion framework proposed in this thesis is currently developed for a regular 3D grid. In the future work, we would like to make an extension of our current fusion framework to a more general graph structure so that it could then cater to even more diverse applications.

4. In the current work, while computing the neighborhood prior terms and smoothness terms, we considered the voxels present on the standard 3D grid of 6 neighbors. It will be interesting to explore other possibilities as well, for instance, considering the voxels in the normal directions since it is in this direction the changes in labeling are more likely to occur.

5. The weighting parameters for the smoothness term and the neighborhood prior term in chapter 6 are empirically set, based on the visual inspection of the results, and they are not optimized anymore. It will be handy to establish some automated criteria for fine tuning these parameters. We would like to also study the effects of varying these weights on different fusion methods.

6. We would like to also extend our fusion framework to other important applications like, automated parcellation of human cerebral cortex [120]. This requires, as a first step, extending our MRF-based fusion framework from a regular 3D grid to even more general graph structure.

7. There are indeed many tasks that can be parallelized both in the registration algorithm and in the fusion algorithms. In order to be able to use these algorithms in real time even when using multiple registration forces and high resolution images, GPU-based implementation of these algorithms will be quite useful.

Possible Future Work in H&N Lymph Nodes Segmentation:

1. In our current data set, manual delineations made by a single expert at one time are treated as the gold standard. Thus, these manual segmentations do not take into account either inter-rater or intra-rater variability. The ground truth segmentations will be more reliable and accurate if the manual delineations made by multiple experts, and by the same expert at different times are merged together and those resultant segmentations are considered as the ground truth. This could even improve the accuracy of the automated segmentations.

2. The current data set contains 12 atlases. The bigger the data set, the larger will be the variability in the anatomies that it can cover. Especially for applications like H&N lymph nodes segmentation, one can expect large anatomical variabilities among different patients. Hence, having a bigger data set could result in generating more reliable and accurate segmentations.

3. We evaluated the automated segmentations using various statistical and geometrical metrics. However, one important metric that is still missing in our evaluations is the average time taken by a doctor to make the final corrections on the results obtained from each fusion method. This is an important metric that has to be included in the future evaluations since the ultimate purpose of using any automated method is to reduce the time that a doctor has to spend in making the corrections.

4. In this thesis, we evaluated the effects of incorporating the proposed neighborhood term with SBA, GWSBA and LWSBA methods. In the future work, we want to extend this evaluation to MV, GWV and LWV methods also.

5. Since the very definition of lymph node structures is based on certain visible and well defined landmarks in the CT images, in practice, it could be a good idea to introduce one more layer of registration which is primarily driven by a set of landmarks marked by an expert oncologist on every atlas and the target image. While considering these landmarks, only those landmarks that can be quickly identified, and are consistently present in all images should be chosen for this purpose.

6. In the region-based registration process, we used four important structures that have distinct boundaries and can impact the position of the lymph nodes; they are - external contour, mandible, vertebrae and trachea. Identifying and including more such relevant structures could probably improve the accuracy further.

7. Finally, in the current registration for lymph nodes segmentation, we used region-based and pixel-based forces. We could further explore incorporating more registration forces like statistical features [18–20] and geometrical features [21, 22].

A Comparison of Minimization Methods for Brain Segmentation

This chapter presents 3D brain tissue classification schemes using three recent promising energy minimization methods for Markov random fields: graph cuts, loopy belief propagation and tree-reweighted message passing. The classification is performed using the well known finite Gaussian mixture Markov Random Field model. Results from the above methods are compared with widely used iterative conditional modes algorithm. The evaluation is performed on a data set containing simulated T1-weighted MR brain volumes with varying noise and intensity non-uniformities. The comparisons are performed in terms of energies as well as based on ground truth segmentations, using various quantitative metrics.

A.1   Introduction

Recent developments in energy minimization methods for Markov random fields (MRF) have resulted in faster and efficient global (or strong local) minimization algorithms [105]. Graph cuts (GC) [106–108], loopy belief propagation (LBP) [109,110] and tree reweighted message passing (TRW) [111,112] are among such most notable algorithms. The applications of these algorithms are spread over a wide variety of early vision problems. Szeliski et al. [105] compared these algorithms in the domains of stereo matching, image stitching, 2D binary image segmentation, denoising and inpainting, for different types of smoothness-based priors. They demonstrated the potential of these global minimization algorithms over the older yet widely used iterated conditional modes (ICM) [121] algorithm. While many of these global minimization algorithms have received the much deserved attention in the domains like stereo matching, they are less explored in the context of medical imaging. For instance, to the best of our knowledge, TRW algorithm, which is found to give consistently strong results in various early vision problems [105], has never been evaluated in the context of medical image segmentation.

The main contribution of this chapter is the convergence study of these optimization

methods for tissue classification of Magnetic Resonance (MR) brain imaging. Brain tissue classification plays an important role in many applications. For instance, this is essential for the quantitative study and analysis of several brain disorders like alzheimer's disease, as well as in understanding the development process of the brain. Further, brain tissue classification is also used as a preprocessing step for many applications like voxel-based morphometry. Among numerous approaches, MRF models are widely used for performing automated 3D brain tissue classification [122]. This work evaluates the recent MRF energy minimization methods on the widely used finite Gaussian mixture MRF (FGMMRF) model [123].

We note that except in very few works [124], none of the global (or strong local) optimization methods are evaluated for brain tissue classification. [124] compared brain tissue classification results obtained from ICM with graph cuts. However, in their comparison, while graph cuts-based method was using tissue-priors information, this information is not used with ICM-based method; further, the 3D segmentation in [124] is done slice by slice on a 2D grid, but not on the original 3D grid. Thus, no comparisons of MRF energy minimization methods for brain tissue classification are available in the literature that are performed under identical parameters.

The following algorithms are evaluated in this chapter: (i) two most popular versions of GC, known as *expansion-move* and *swap-move* algorithms [106–108], (ii) an LBP implementation derived by Kolmogorov from TRW-S [112], called *BP-S* algorithm, (iii) an improved version of the original TRW algorithm [111], called the *sequential TRW (TRW-S)* [112], and (iv) *ICM* algorithm [121].

Regarding the implementation of these algorithms, thanks to the MRF library [105] [1], it has served as a basis for our current implementation. The above mentioned MRF library can, however, handle neighborhood priors only on a 2D grid which is not suffice for 3-D medical imaging applications. We have now enhanced it to handle 3D grid and also integrated it with ITK [2], which is a widely used open-source tool in medical imaging.

The rest of the chapter is organized as follows. In section A.2, we briefly present the FGMMRF model and the optimization methods. Evaluation results are presented in section A.3. Discussion and conclusions are presented in section A.4.

---

[1]http://vision.middlebury.edu/MRF/
[2]http://itk.org/

A.2 Methodology

A.2.1 Energy Model

Let ν represent the set of all voxels in a given image, and X_p be the label (tissue-class) assigned to the pth voxel. Let N be the number of voxels in the set ν. Let X be the set containing the labels assigned to ν, i.e., $X = \{X_1, \cdots, X_N\}$. Then, the brain tissue classification is often formulated as an energy minimization problem of the form:

$$E(X) = \sum_{\forall p \in \nu} \psi_p(X_p) + \beta \sum_{\forall p \in \nu} \sum_{\forall q \in \aleph_p} \psi_{pq}(X_p, X_q),$$

where the first term is a data term (unary function) representing the intensity modeling of the tissue classes, and the second term is a smoothness term (pairwise function) representing the neighborhood-prior modeling. β is a weighting parameter between the data term and smoothness term.

a) Data term: Let I_p be the observed intensity of the pth voxel, and $I = \{I_1, \cdots, I_N\}$. Let $L = \{l_1, \cdots l_D\}$ be the set of labels to be assigned. Let $\theta_i = \{\mu_i, \sigma_i\}$ are mean and variance associated with label l_i, and $\theta = \{\theta_1, \cdots, \theta_D\}$. For the brain tissue classification, intensities of the tissue classes are generally modeled with Gaussian distribution [123]. Hence, the likelihood that a given labeling X and θ have produced I is given by:

$$P(I|X, \theta) = \prod_{\forall p \in \nu} \left(\frac{1}{\sqrt{2\pi}\,\sigma_{X_p}} \exp\left(\frac{-\left(I_p - \mu_{X_p}\right)^2}{2\,\sigma_{X_p}^2} \right) \right).$$

By taking a negative logarithm of the above equation, the problem of maximizing the likelihood can be transformed into an equivalent energy minimization problem with the following unary function:

$$\psi_p(X_p) = \frac{(I_p - \mu_{X_p})^2}{2\,\sigma_{X_p}^2} + \log(\sqrt{2\pi}\,\sigma_{X_p}).$$

b) Smoothness term: The smoothness term for FGMMRF model is given by the discontinuity preserving Pott's model:

$$\psi_{pq}(X_p, X_q) = \begin{cases} 0, & \text{if } X_p = X_q; \\ 1, & \text{otherwise.} \end{cases}$$

Note that the above smoothness term does not penalize when neighborhood voxels are assigned the same label. The finite Gaussian mixture MRF (FGMMRF) model studied here [123] aims to classify the image voxels into one of the three brain tissue types: cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM). The background

(BKG) is represented by an additional label. Thus, $D = 4$ for this model, and the set of labels is given by $L = \{\text{BKG, CSF, GM, WM}\}$.

A.2.2   Optimization Methods

We briefly mention here the general characteristics of the optimization methods that we evaluate on the energy model described in the preceding subsection. For a more detailed discussion, the readers are referred to [105]. The classical ICM algorithm uses a greedy approach, and converges faster than the rest of the algorithms evaluated in this chapter. However, ICM provides local minima, and thus it is very sensitive to initialization of the labeling. Graph cuts, if applicable, generally gives very accurate results. They are however applicable to limited pairwise potentials. For example, for obtaining a guaranteed global optimum with graph cuts, the pairwise potential term should be a metric in case of GC-Expansion and semimetric for GC-Swap. On the other hand, LBP is applicable to any energy function, but it does not always converge. Regarding TRW-S, it is noted in [105] that it can be a good competitor to graph cuts in certain applications. TRW-S is applicable to any energy function. However, TRW-S is computationally slower than graph cuts, and convergence of energy is not guaranteed.

A.3   Results

The evaluation is performed on the simulated brain volumes from the *Brain Web* database of the McConnell Brain Imaging Centre [125]. We use a data set of 12 brain volumes that contain images with varying noise levels (0%, 1%, 3%, 5%, 7% and 9%), and two levels of intensity non-uniformities (INU 20% and 40%).

Among different parameters to be set, the most sensitive parameter in obtaining accurate tissue classification is the *mean intensity* of each tissue class (μ_l). Hence, we have iteratively updated the mean intensity values of each label using the expectation maximization approach in [122], by computing the a posteriori probabilities of the labels, and is given by:

$$\mu_l^{(k)} = \frac{\sum\limits_{\forall p \in v} (I_p)\, P^{(k)}(X_p = l | \, I_p, \theta_l^{(k-1)})}{\sum\limits_{\forall p \in v} P^{(k)}(X_p = l | \, I_p, \theta_l^{(k-1)})}.$$

The σ value for each tissue class is fixed to the typical value 5 [122]. β is fixed empirically to 1.0 similar to [105, 122], based on the visual inspection of the segmentation results.

In order to compare the energy convergence results from all the methods under identical conditions, final μ values obtained by iteratively updating them in one of

the methods (GC-Expansion) are used for the remaining methods. For instance, final μ values with GC-Expansion, obtained for CSF, GM and WM, for the image with 5% noise and 20% INU are, 46.3, 92.9 and 126.4 respectively, and the same values are used with other MRF optimization methods. However, we noticed that, even allowing each method to independently update the μ values has resulted in similar mean values with small variations in their first decimal.

All the experiments are run on the same machine (2.26 GHz Intel Xeon Processor, 12 GB RAM). The results from all 12 brain volumes have shown consistently similar behavior. We show here energy results for just one image with 5% noise and 20% INU, and visual results for one of the axial slices extracted from the above brain volume. However, the quantitative results presented here are computed over the entire data set of 12 volumes.

TRW-S has the ability to compute a lower bound on the energy of the optimal solution. Similar to [105], we take advantage of this lower bound; instead of comparing the absolute energies, we normalize the energies by dividing them with the best lower bound computed by TRW-S, and then compare the energies.

Figure A.1 shows the energy convergence results for the brain volume with 5% noise and 20% INU. The energy differences among these methods, including ICM, are found to be marginal. In terms of time taken for convergence, ICM is obviously faster compared to global optimization methods since, its convergence is based on local optimization criteria. Among the global optimization methods, expansion-move version of graph cuts is the fastest one. The energy convergence results for the remaining brain volumes in the data set are also similar to the above mentioned results.

Figure A.2 shows an axial slice extracted from the image brain volume with 5% noise and 20% INU, and its ground truth segmentation. Figure A.3 shows automated tissue classification results obtained from all methods. Quantitative evaluation is performed using the commonly used metrics: (i) sensitivity (a measure of true-positive fraction), (ii) specificity (a measure of true-negative fraction), (iii) dice similarity metric (DSM) (a measure of overlap between ground truth and automated segmentation), and (iv) % of error in volume; the results are shown in table A.1, table A.2, table A.3, and table A.4. The differences among the methods for all these metrics are found to be marginal. The results from ICM in terms of these metrics are, surprisingly, quite close to the best ones.

A.4   Discussion and Conclusions

In this chapter, we have presented tissue classification of MR brain volumes with different MRF optimization algorithms, and compared them using identical parameters. The evaluated algorithms are: expansion-move and swap-move versions of the graph
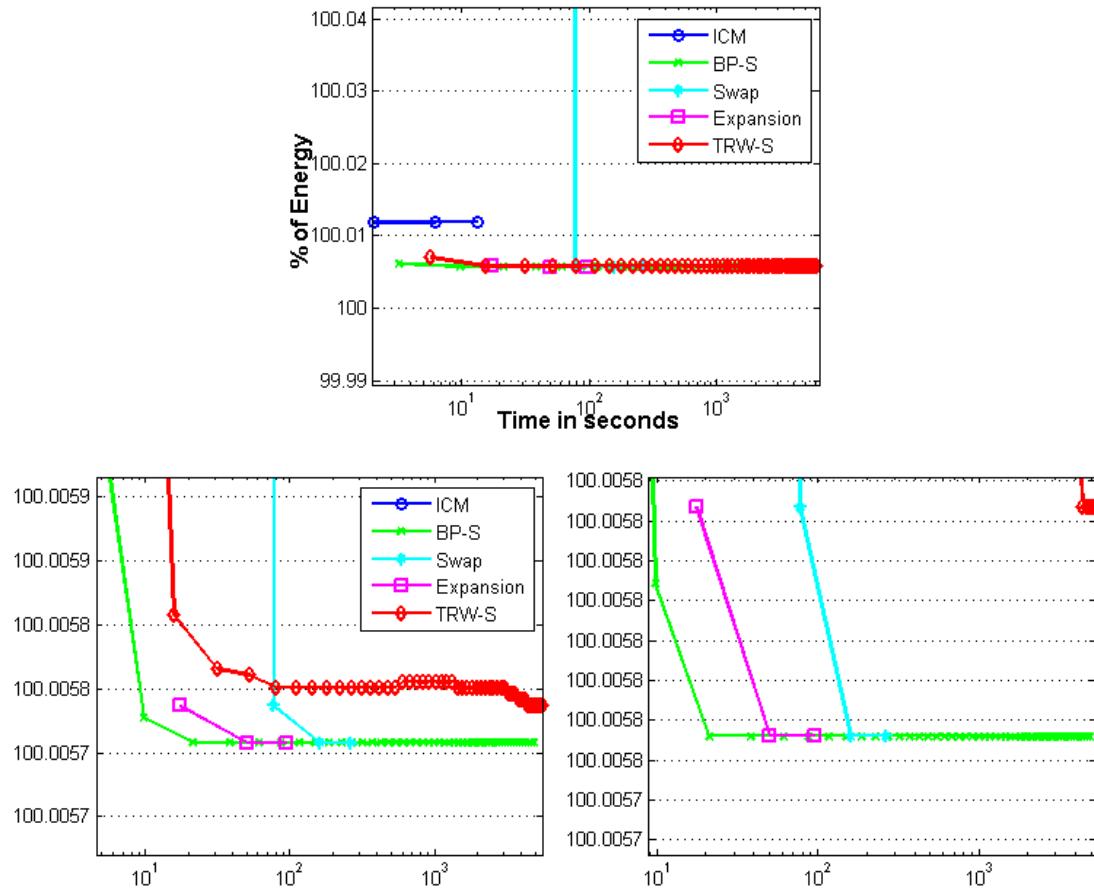
Figure A.1: Comparison of energy results for the MRF minimization methods. These are the results obtained from FGMMRF model, for the simulated brain volume with 5% noise and 20% INU. The second and third plots are the zoomed versions of the first plot.
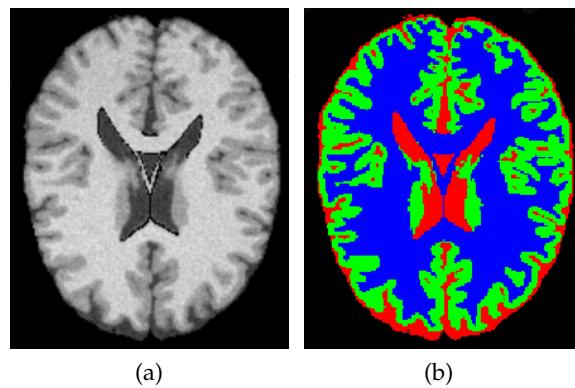
Figure A.2: (a) One of the axial slices extracted from the brain volume with 5% noise and 20% INU. (b) Ground truth segmentation of tissues for the same slice; CSF, GM and WM tissue classes are represented with labels of red, green and blue colors respectively.

(a) Ground truth   (b) ICM   (c) BP-S

(d) GC-Swap   (e) GC-Expansion   (f) TRW-S
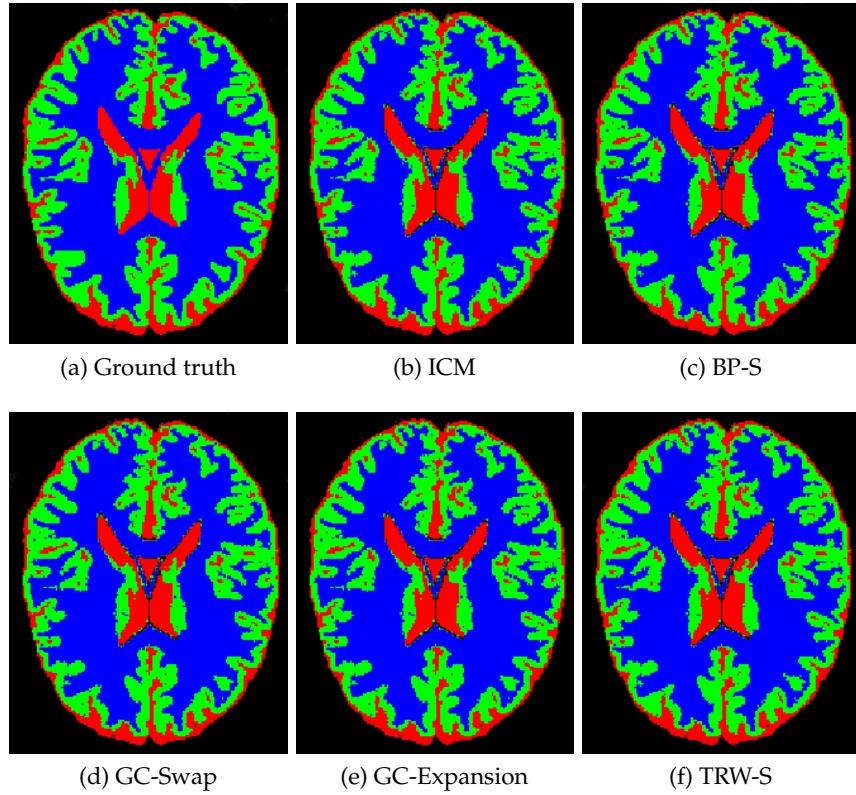
Figure A.3: Qualitative comparison of segmentation results in one of the slices in the axial direction, extracted from the brain volume with 5% noise and 20% INU. Results from all the methods are very close, and are quite similar to the ground truth segmentation.

Table A.1: Mean and standard deviations of sensitivity of CSF, GM and WM tissues, obtained from different MRF optimization methods.

Method	Sensitivity(%)		
	CSF	**GM**	**WM**
ICM	93.42 ± 4.41	92.23 ± 3.41	94.55 ± 2.81
BP-S	93.43 ± 4.39	92.21 ± 3.39	94.62 ± 2.77
GC-Swap	93.23 ± 4.45	91.74 ± 3.59	95.14 ± 2.46
GC-Exp	93.26 ± 4.43	92.01 ± 3.46	94.84 ± 2.61
TRW-S	93.43 ± 4.39	92.21 ± 3.39	94.62 ± 2.76

Table A.2: Mean and standard deviations of specificity of CSF, GM and WM tissues, obtained from different MRF optimization methods.

Method	Specificity(%)		
	CSF	GM	WM
ICM	99.28 ± 0.85	99.21 ± 0.47	99.21 ± 0.39
BP-S	99.26 ± 0.88	99.22 ± 0.46	99.20 ± 0.38
GC-Swap	99.16 ± 1.11	99.25 ± 0.45	99.10 ± 0.42
GC-Exp	99.27 ± 0.89	99.23 ± 0.45	99.16 ± 0.40
TRW-S	99.26 ± 0.88	99.22 ± 0.46	99.20 ± 0.38

Table A.3: Mean and standard deviations of dice similarity metric of CSF, GM and WM tissues, obtained from different MRF optimization methods.

Method	Dice Similarity Metric(%)		
	CSF	GM	WM
ICM	90.89 ± 8.30	93.32 ± 3.34	93.60 ± 3.07
BP-S	90.79 ± 8.48	93.33 ± 3.32	93.62 ± 3.04
GC-Swap	90.11 ± 9.75	93.17 ± 3.39	93.48 ± 3.04
GC-Exp	90.81 ± 8.48	93.27 ± 3.33	93.55 ± 3.01
TRW-S	90.79 ± 8.48	93.33 ± 3.32	93.62 ± 3.04

Table A.4: Mean and standard deviations of % error in volume of CSF, GM and WM tissues, obtained from different MRF optimization methods. These values are computed over a data set of 12 brain volumes with varying noise (0-9%) and INU (0-20%).

Method	%Error in Volume		
	CSF	GM	WM
ICM	6.49 ± 11.43	-2.35 ± 0.75	2.05 ± 1.67
BP-S	6.81 ± 12.05	-2.42 ± 0.74	2.16 ± 1.64
GC-Swap	8.46 ± 15.98	-3.09 ± 0.86	3.60 ± 2.00
GC-Exp	6.39 ± 12.16	-2.70 ± 0.82	2.77 ± 1.76
TRW-S	6.81 ± 12.06	-2.42 ± 0.74	2.17 ± 1.64

cuts method, loopy belief propagation, sequential tree reweighted message passing, and the older iterated conditional modes algorithm. The evaluation is performed on the widely used FGMMRF model, using a data set of 12 brain volumes with varying noise and INU.

In order to draw definite conclusions about the accuracy of the widely used FGMMRF model (or any other model in general), it is essential to make sure that the solution has converged to a global optimum. Otherwise, it will be ambiguous whether the resulting errors are due to the model itself, or, whether the errors are related to the convergence of optimization method used for solving the model. The current study is important in this perspective also.

It is well known that since ICM converges to a local optimum, its results depend very much on the initialization of the labels. In this work, for ICM, each voxel is initialized with that tissue-class label with whose μ value the voxel's intensity difference is minimum. From the results, we notice that although ICM is a local optimization method, the above mentioned initialization resulted in an accurate classification. However, if a data set to be segmented is quite different from the above tested volumes from Brainweb, it may be safe to use GC-Expansion rather than ICM so that convergence to a global minimum is guaranteed. Note that among the global optimization methods, GC-Expansion converged quickly. In the future work, we want to study in more detail, the effects of different initializations on the convergence of ICM. We would like to extend this evaluation to real brain volumes. Further, we would like to perform similar investigation on other tissue classification models like partial volume models [122, 126]. It would be also interesting to evaluate other fast optimization methods like [113].

B Evaluation of Fusion Methods for Parotid Glands Segmentation

This chapter presents the segmentation of bilateral parotid glands in the Head and Neck (H&N) CT images using an active contour-based atlas registration. We compare segmentation results from three atlas selection strategies: (i) selection of "single-most-similar" atlas for each image to be segmented, (ii) fusion of segmentation results from multiple atlases using STAPLE, and (iii) fusion of segmentation results using majority voting. Among these three approaches, fusion using majority voting provided the best results. Finally, we present a detailed evaluation on a dataset of eight images (provided as a part of H&N auto segmentation challenge conducted in conjunction with MICCAI-2010 conference) using majority voting strategy.

B.1 Introduction

Automated segmentation of structures in the Head and Neck (H&N) CT images is a challenging as well as important task for radiation treatment of H&N cancer [7]. Among various structures in the H&N region, parotid glands are one of the important organs at risk that need to be accurately segmented in treatment planning. Automated segmentation of parotid glands is challenging because of their low contrast and lack of distinctly visible boundaries with the surrounding structures. Thus, for an accurate segmentation of these structures, inclusion of prior knowledge is essential. Atlas-based segmentation methods are widely used for exploiting prior anatomical knowledge. There are two factors that mainly affect the accuracy of atlas-based segmentation methods: the type of registration algorithm used for mapping the atlas to the image to be segmented, and the closeness/similarity of the atlas to the image to be segmented.

In this chapter, we perform segmentation of bilateral parotid glands in the H&N CT images, using an Active Contour-Based Atlas Registration (ACBAR) framework. This framework has been already proven to be successful in the segmentation of other structures in the H&N CT images, like lymph nodes [7], mandible and brainstem [25].

As we mentioned, atlas selection is another important factor that affect the accuracy of segmentation. In this chapter, we mainly focus on the atlas selection strategies.

B.2 ACBAR

In the section, we present a brief description of Active Contour-Based Atlas Registration (ACBAR) framework. Please refer to [7, 25] for more details. ACBAR framework combines the forces coming from both optical flow framework (like pixel-based forces), and active contour framework (like region-based and boundary-based forces). The formulation of ACBAR has been intuitively derived from level set equation [88]. The generalized evolution equation of the registration model is given by:

$$\frac{\partial u(x,t)}{\partial t} = -S(x)\, v(\phi_L(x + u(x,t), 0))\, \frac{\nabla \phi_L}{|\nabla \phi_L|} \tag{B.1}$$

where $u(x,t)$ is the deformation field vector at time t and spatial coordinates x, v is a speed function that contains local segmentation and contour regularization constraints, $S(x)$ is a sign function that provides polarity information, and ϕ_L is a labeled level set representation of the contour proposed in [7].

In the current application, we perform a two-level hierarchical registration. In the first level, we perform a region-based registration driven by two selected structures of interest: external contour of the image and the mandible. The reason for using mandible is because it is adjacent to the actual structure to be segmented (parotid glands), and thus can influence in accurately segmenting the parotid glands on the target image. The region-based energy term used is inspired by region-based segmentation model proposed by Chan and Vese [94], and it is same the one described in [25]. In the second level, the registration is driven by pixel-based forces, and the corresponding speed term is same as the one used by [2]. The details of the registration parameters, preprocessing and postprocessing used here for the segmentation of parotid glands are presented in section B.4.

B.3 Atlas Selection Strategies

Atlas selection strategy is a key issue for achieving accurate results in atlas-based segmentation [23, 30]. In majority of works, the much deserved attention has not been given for atlas selection, and a single segmented image is almost randomly selected as an atlas for all the images to be segmented. Such random selection can lead to significant undesirable bias. One of the simple and effective approaches to overcome this problem is, for each image to be segmented, to adaptively select the most similar atlas from a given database of atlases [23]. Alternatively, instead of using a single atlas, multiple atlases can be used. For instance, in [23], it is shown for the segmentation of brain in the

confocal microscopy images, that, using multiple atlases can give more accurate results than with a single-best-atlas.

Segmentation results from multiple atlases can be fused in different ways [23, 30, 39]. In this work, we consider two popular fusion approaches: "Majority voting" [23] and "STAPLE" [39]. In majority voting, each voxel is assigned with the label that most segmentations agree [30]. Another popular algorithm is "Simultaneous Truth And Performance Level Estimation" (STAPLE) [39]. Intuitively, majority voting gives equal weight to segmentations from all atlases, while in STAPLE, the probabilistic estimate of the true segmentation is formed by estimating an optimal combination of the segmentations, weighting each segmentation depending upon the estimated performance level. In this work, we assess three atlas selection strategies: (i) adaptively selecting the "most similar atlas" for each image to be segmented, (ii) fusion of segmentation results from multiple atlases using STAPLE, and (iii) fusion of segmentation results from multiple atlases using majority voting.

B.4   Results

The data set is provided by Princess Margaret Hospital, Toronto, as a part of H&N auto-segmentation challenge [127]. It consists of 18 CT images. At the time of evaluation, expert segmentation for only 10 images (training data) are available to the participants, whereas for the remaining 8 images (testing data), expert segmentations are hidden from the participants and are available only with the organizers of the challenge; automated segmentations for the testing data are submitted by the participants to one of the organizers and then, the organizer has generated the evaluation results for the testing data. Out of 10 images in the training data, for 3 images (numbered: 01, 04 and 10), there are considerable artifacts in the mandible region; so we have constructed an atlas database with only the remaining 7 images. We note that we could actually remove those 3 images from the training data as well. However, we deliberately did not do that because, in practice, we can only select the images to be used in the atlas database, but not the images to be segmented. On the other hand, we could have introduced a preprocessing step before segmenting them, in order to remove the artifacts and thereby obtain more accurate evaluation.

In order to speed up the registration, all images are cropped in the Z-direction; the images are cropped such a way that they include all the structures of interest (parotid glands) as well as 3 additional axial slices both at the top and bottom of the images.

First of all, an affine registration is performed between the images to be segmented versus all the images in the atlas database. We use the AffineTransform available in ITK [1] along with Mean Square Error (MSE) as similarity metric. Then, for each image

---

[1]http://www.itk.org/
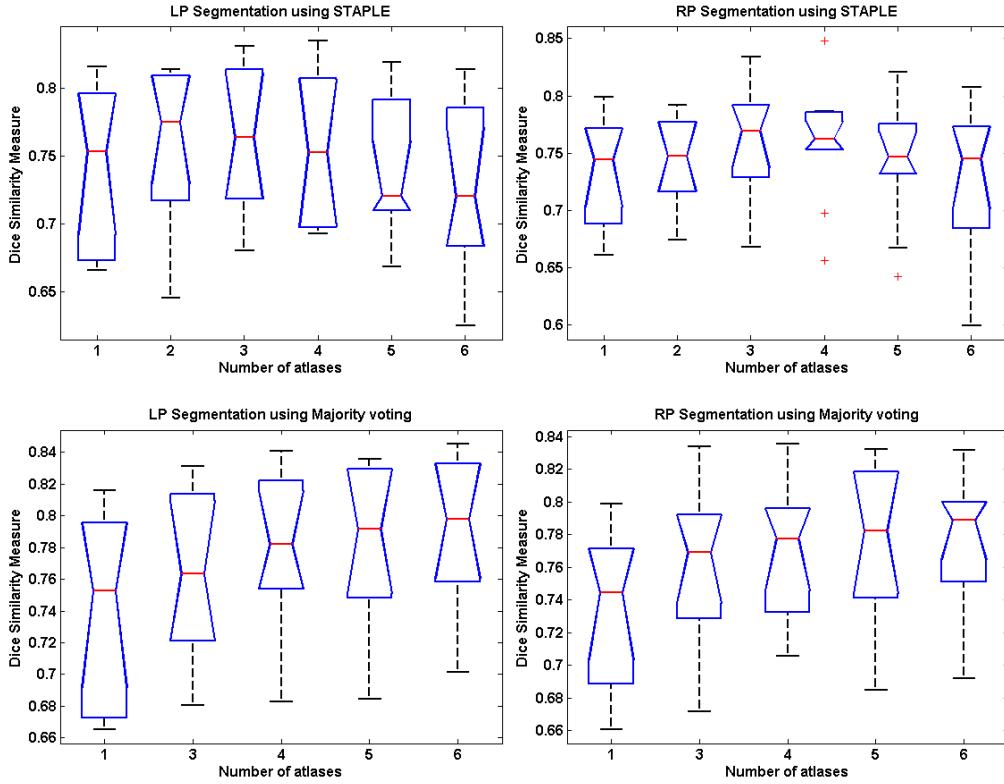
Figure B.1: Box plots of Dice Similarity Measure (DSM) for the segmentation of Left Parotid (LP) and Right Parotid (RP) glands, for varying number of atlases. Box plot corresponding to a single atlas represents the results for single-best-atlas selection.

to be segmented, atlases are ordered based on the resulting MSE at the end of affine registration. This is based on our hypothesis that smaller MSE indicates better similarity of the atlas to the image, and thus, is ordered first. This ordering is used later while merging segmentations from multiple atlases. Then, as described in section B.2, a 2-level active contour-based registration is performed; in the first level, region-based forces and curvature forces are used, and in the second level, pixel-based forces are used. The parameters for ACBAR are same as those used in [25]. For the segmentation of parotid glands on testing data, merged segmentation results from multiple atlases are post-processed by connected-component-thresholding to contain only the principal component, followed by Gaussian smoothing.

B.4.1 Comparison of Atlas Selection Strategies

Three types of atlas-selection strategies are compared on the training data of 10 images. While selecting atlases for each image, obviously, the same image in the atlas database is left out and is not considered as an atlas-candidate. Figure B.1 shows box plots of Dice Similarity Measure (DSM) for the segmentation parotid glands, with varying number of
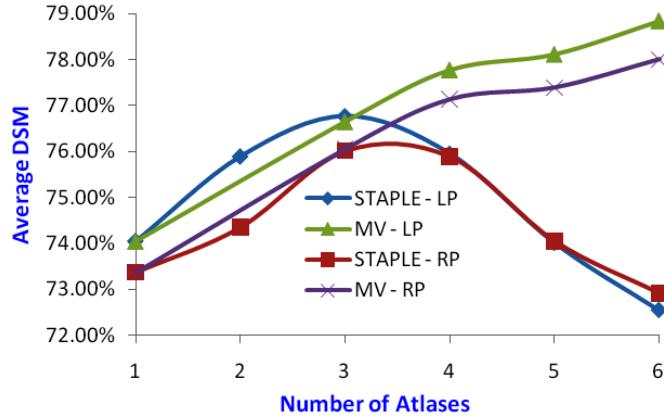
Figure B.2: Graph showing the average values of DSM for the segmentation of Left Parotid (LP) and Right Parotid (RP) glands, with varying number of atlases. Note that since atlases for each image are ordered in the decreasing order of similarity, the values shown for single atlas (i.e., values corresponding to x-label: "1") represent the average DSM values for "single-best-atlas" selection strategy.

Table B.1: Comparison of the three approaches for parotid glands segmentation.

Atlas Selection Strategy	Optimal no. of Atlases	Corresponding DSM (mean $\pm$ SD)	
		Left Parotid	Right Parotid
Adaptive Single Best Atlas	-	74.05% $\pm$ 6.02%	73.37% $\pm$ 5.06%
STAPLE	3	76.77% $\pm$ 5.13%	76.02% $\pm$ 5.05%
Majority Voting	6	**78.85% $\pm$ 5.01%**	**78.01% $\pm$ 4.18%**

atlases. Top row and bottom row respectively show the statistics for atlas fusion using STAPLE and Majority Voting (MV) algorithms. First and second columns respectively show the statistics for left and right parotid glands. Please note that for each image to be segmented, atlases are ordered in the decreasing order of similarity; in other words, the first atlas for each image represents the "most similar atlas" for that image. Thus, in the box plots of both STAPLE and majority voting approaches, the boxes corresponding to single atlas (i.e., the first box in each figure with x-label: "1") actually represent the statistics for "single-most-similar" atlas selection. Because of this reason, we have not separately shown the box plots for single best atlas selection strategy. Figure B.2 shows average values of DSM for parotid glands, over the complete training data, with varying number of atlases. Table B.1 summarizes the comparison of the three atlas selection approaches. Note that number of atlases for which the average value of DSM is maximum, is considered as the "optimal number of atlases" for that strategy.

Multi-atlas-based segmentation results clearly outperformed the single-best-atlas selection, with an optimal selection of "number of atlases". Between the multi-atlas selection strategies, STAPLE gave its best results with 3 atlases whereas results from

Table B.2: Hausdorff distance(HD) statistics for left parotid segmentation.

Dataset No.	Mean HD	Median HD	No. of slices (HD > 3 mm)
11	9.65	9.61	34 (34)
12	7.40	7.44	28 (28)
13	10.34	9.79	26 (26)
14	11.65	9.69	24 (24)
15	6.42	5.12	23 (23)
16	6.69	4.98	29 (28)
17	14.26	14.71	33 (33)
18	8.76	8.22	24 (24)
Mean±SD	**(9.40±2.68)**	**(8.70±3.11)**	-

MV kept improving till 6 atlases. There is approximately 2.7% improvement in DSM from single-best-atlas to the best results of STAPLE, and there is further improvement of 2.1% from the best results of STAPLE to that of MV. Based on these results, for the final evaluation on the testing data, we have chosen majority voting strategy with 6 atlases.

B.4.2 Evaluation on Testing Data

As mentioned in the preceding subsection, the final evaluation on the testing data of 8 images is performed using majority voting with 6 atlases. The quantitative evaluation is performed using various Hausdorff Distance (HD)-based and overlap (OV)-based metrics. Some of these metrics are evaluated both slice-wise and volumetric-wise.

Figure B.3 shows ground truth and automated segmentation of parotid glands for one of the images. Table B.2 and table B.3 respectively present the Hausdorff distance and overlap statistics for left parotid gland segmentation. Similarly, table B.4 and table B.5 respectively present the Hausdorff distance and overlap statistics for right parotid gland segmentation. As expected, the quantitative evaluation results for left and right parotid glands have similar behavior, because of the similarities of these two structures. Average total volume overlap is around 75.5% whereas the average mean-HD is around 9 *mm*.

B.5 Conclusions

In this chapter, we have presented the segmentation of parotid glands in the H&N CT images using active contour-based atlas registration framework. We have evaluated three atlas selection strategies: single-best-atlas, STAPLE and majority voting; among the three approaches, majority voting gave the best results. We then performed the final evaluation on the testing data using majority voting.

Although the current automated segmentations still need manual corrections before
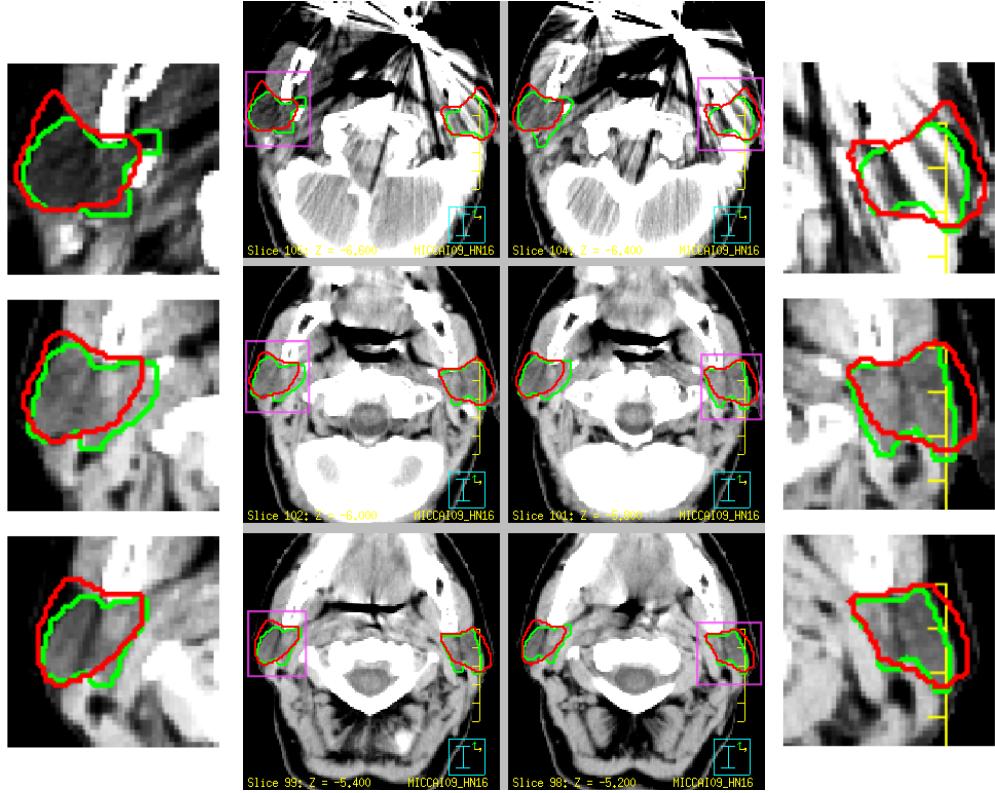
Figure B.3: Middle two columns show the segmentation of parotid glands for one of the images in the testing data. First and last columns show a zoom of the selected regions in the middle image. Ground truth and automated segmentations are respectively shown in green and red colors.

Table B.3: Overlap(OV) statistics for left parotid segmentation on testing data.

Dataset No.	Average slice OV	Median slice OV	Total volume OV
11	69.4 %	72.3 %	77.3 %
12	72.1 %	78.8 %	77.0 %
13	69.0 %	75.7 %	74.2 %
14	68.6 %	78.5 %	74.6 %
15	81.9 %	87.6 %	85.9 %
16	73.0 %	79.0 %	75.9 %
17	60.7 %	63.0 %	63.5 %
18	69.4 %	73.8 %	75.4 %
Mean±SD	**(70.51±5.89)%**	**(76.09±7.02)%**	**(75.48±6.10)%**

using them in treatment planning, these results are quite promising considering the low contrast of the parotid glands and the presence of artifacts. These two issues can be clearly noticed from figure B.3.

Table B.4: Hausdorff distance(HD) statistics for right parotid segmentation.

Dataset No.	Mean HD	Median HD	No. of slices (HD > 3 mm)
11	6.81	5.69	32 (32)
12	12.30	9.62	30 (30)
13	9.69	7.01	26 (26)
14	12.23	9.81	24 (24)
15	8.13	5.27	26 (24)
16	7.03	6.02	31 (30)
17	10.17	10.28	27 (27)
18	6.62	5.90	25 (25)
Mean±SD	**(9.12±2.33)**	**(7.45±2.10)**	-

Table B.5: Overlap(OV) statistics for right parotid segmentation on testing data.

Dataset No.	Average slice OV	Median slice OV	Total volume OV
11	77.8 %	83.1 %	81.7 %
12	66.6 %	76.2 %	71.7 %
13	72.4 %	75.8 %	77.9 %
14	62.1 %	70.6 %	68.1 %
15	77.7 %	85.1 %	84.4 %
16	69.5 %	75.4 %	73.8 %
17	64.7 %	72.2 %	67.5 %
18	73.8 %	82.3 %	80.2 %
Mean±SD	**(70.58±5.86)%**	**(77.59±5.30)%**	**(75.66±6.34)%**

The main contribution of this work is the evaluation of three atlas selection strategies. We note that while ordering the atlases for fusion, we have arbitrarily selected MSE as a measure of similarity. In future work, we would like to evaluate and compare the effect of using other similarity measures, like normalized correlation coefficient and mutual information.

Bibliography

[1] A. Yezzi, L. Zollei, and T. Kapur, "A variational framework for joint segmentation and registration," *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (CVPR-MMBIA)*, pp. 44–49, 2001.

[2] B. C. Vemuri, J. Ye, Y. Chen, and C. M. Leonard, "Image registration via level-set motion: Applications to atlas-based segmentation," *Medical Image Analysis*, vol. 7, no. 1, pp. 1–20, 2003.

[3] D. W. Shattuck, M. Mirza, V. Adisetiyo, C. Hojatkashani, G. Salamon, K. L. Narr, R. A. Poldrack, R. M. Bilder, and A. W. Toga, "Construction of a 3D probabilistic atlas of human cortical structures," *NeuroImage*, vol. 39, no. 3, pp. 1064 – 1080, 2008.

[4] A. Klein, J. Andersson, B. A. Ardekani, J. Ashburner, B. Avants, M.-C. Chiang, G. E. Christensen, D. L. Collins, J. Gee, P. Hellier, J. H. Song, M. Jenkinson, C. Lepage, D. Rueckert, P. Thompson, T. Vercauteren, R. P. Woods, J. J. Mann, and R. V. Parsey, "Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration," *NeuroImage*, vol. 46, no. 3, pp. 786 – 802, 2009.

[5] J. M. Lotjonen, R. Wolz, J. R. Koikkalainen, L. Thurfjell, G. Waldemar, H. Soininen, and D. Rueckert, "Fast and robust multi-atlas segmentation of brain magnetic resonance images," *NeuroImage*, vol. 49, no. 3, pp. 2352 – 2365, 2010.

[6] O. Commowick, V. Grégoire, and G. Malandain, "Atlas-based delineation of lymph node levels in head and neck computed tomography images," *Radiotherapy Oncology*, vol. 87, no. 2, pp. 281–289, 2008.

[7] S. Gorthi, V. Duay, N. Houhou, M. Bach Cuadra, U. Schick, M. Becker, A. S. Allal, and J.-P. Thiran, "Segmentation of head and neck lymph node regions for radiotherapy planning using active contour-based atlas registration," *IEEE Journal on Selected Topics in Signal Processing*, vol. 3, no. 1, pp. 135–147, 2009.

[8] S. Gorthi, V. Duay, X. Bresson, M. Bach Cuadra, S. Castro, F. Javier, C. Pollo, A. S. Allal, and J.-P. Thiran, "Active deformation fields: Dense deformation field

estimation for atlas-based segmentation using the active contour framework," *Medical Image Analysis*, vol. 15, no. 6, pp. 787–800, 2011.

[9] M. Lorenzo-Valdes, G. Sanchez-Ortiz, R. Mohiaddin, and D. Rueckert, "Atlas-based segmentation and tracking of 3D cardiac MR images using non-rigid registration," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2002*, vol. 2488, 2002, pp. 642–650.

[10] M. Kaus, J. von Berg, W. Niessen, and V. Pekar, "Automated segmentation of the left ventricle in cardiac MRI," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003*, vol. 2878, 2003, pp. 432–439.

[11] M. Lorenzo-Valdés, G. I. Sanchez-Ortiz, A. G. Elkington, R. H. Mohiaddin, and D. Rueckert, "Segmentation of 4D cardiac MR images using a probabilistic atlas and the EM algorithm," *Medical Image Analysis*, vol. 8, no. 3, pp. 255 – 265, 2004.

[12] A. Bistoquet, J. Oshinski, and O. Skrinjar, "Left ventricular deformation recovery from cine MRI using an incompressible model," *IEEE Transactions on Medical Imaging*, vol. 26, no. 9, pp. 1136 –1153, 2007.

[13] T. Rohlfing, R. Brandt, R. Menzel, D. B. Russakoff, and C. R. Maurer, Jr., "Quo vadis, atlas-based segmentation?" in *The Handbook of Medical Image Analysis – Volume III*.   Kluwer Academic, 2005, ch. 11, pp. 435–486.

[14] M. Bach Cuadra, C. Pollo, A. Bardera, O. Cuisenaire, J. Villemure, and J.-P. Thiran, "Atlas-based segmentation of pathological MR brain images using a model of lesion growth," *IEEE Transactions on Medical Imaging*, vol. 23, no. 10, pp. 1301–1314, 2004.

[15] O. T. Carmichael, H. A. Aizenstein, S. W. Davis, J. T. Becker, P. M. Thompson, C. C. Meltzer, and Y. Liu, "Atlas-based hippocampus segmentation in alzheimer's disease and mild cognitive impairment," *NeuroImage*, vol. 27, no. 4, pp. 979 – 990, 2005.

[16] J.-P. Thirion, "Image matching as a diffusion process: an analogy with maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243–260, 1998.

[17] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage*, vol. 45, no. 1, pp. S61–S72, 2009.

[18] D. Rueckert, A. Frangi, and J. Schnabel, "Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration," *IEEE Transactions on Medical Imaging*, vol. 22, no. 8, pp. 1014 –1025, 2003.

[19] N. Houhou, A. Lemkaddem, V. Duay, A. Alla, and J. Thiran, "Shape prior based on statistical map for active contour segmentation," in *15th IEEE International Conference on Image Processing*, 2008, pp. 2284–2287.

[20] J. Xie, Y. Jiang, and H. tat Tsui, "Segmentation of kidney from ultrasound images based on texture and shape priors," *IEEE Transactions on Medical Imaging*, vol. 24, no. 1, pp. 45 –57, 2005.

[21] P. Van den Elsen, J. Maintz, E.-J. Pol, and M. Viergever, "Automatic registration of CT and MR brain images using correlation of geometrical features," *IEEE Transactions on Medical Imaging*, vol. 14, no. 2, pp. 384 –396, 1995.

[22] J. Maurer, C.R., G. Aboutanos, B. Dawant, R. Maciunas, and J. Fitzpatrick, "Registration of 3-D images using weighted geometrical features," *IEEE Transactions on Medical Imaging*, vol. 15, no. 6, pp. 836 –849, 1996.

[23] T. Rohlfing, R. Brandt, R. Menzel, and C. R. Maurer Jr., "Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains," *NeuroImage*, vol. 21, no. 4, pp. 1428–1442, 2004.

[24] O. Commowick and G. Malandain, "Efficient selection of the most similar image in a database for critical structures segmentation." in *Medical image computing and computer-assisted intervention MICCAI*, vol. 10, no. 2, 2007, pp. 203–210.

[25] S. Gorthi, V. Duay, M. Bach Cuadra, P.-A. Tercier, A. S. Allal, and J.-P. Thiran, "Active contour-based segmentation of head and neck with adaptive atlas selection," in *MICCAI Workshop on 3D Segmentation Challenge for Clinical Applications*, 2009.

[26] I. Išgum, M. Staring, A. Rutten, M. Prokop, M. Viergever, and B. van Ginneken, "Multi-atlas-based segmentation with local decision fusion - application to cardiac and aortic segmentation in CT scans," *IEEE Transactions on Medical Imaging*, vol. 28, no. 7, pp. 1000–1010, 2009.

[27] X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz-de Solorzano, "Combination strategies in multi-atlas image segmentation: Application to brain mr data," *IEEE Transactions on Medical Imaging*, vol. 28, no. 8, pp. 1266–1277, 2009.

[28] S. Gorthi, M. Bach Cuadra, U. Schick, P.-A. Tercier, A. S. Allal, and J.-P. Thiran, "Multi-atlas based segmentation of head and neck CT images using active contour framework," in *MICCAI workshop on 3D Segmentation Challenge for Clinical Applications*, 2010.

[29] R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, "Automatic anatomical brain MRI segmentation combining label propagation and decision fusion," *NeuroImage*, vol. 33, no. 1, pp. 115 – 126, 2006.

[30] P. Aljabar, R. Heckemann, A. Hammers, J. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy," *NeuroImage*, vol. 46, no. 3, pp. 726 – 738, 2009.

[31] T. Rohlfing, R. Brandt, C. R. Maurer, Jr., and R. Menzel, "Bee brains, b-splines and computational democracy: Generating an average shape atlas," in *IEEE workshop on mathematical methods in biomedical image analysis*, 2001, pp. 187–194.

[32] H. Park, P. H. Bland, and C. R. Meyer, "Construction of an abdominal probabilistic atlas and its application in segmentation," *IEEE Transactions on Medical Imaging*, vol. 22, no. 4, pp. 483–492, 2003.

[33] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, pp. S151–S160, 2004.

[34] M. De Craene, A. Du Bois D'Aische, B. Macq, and S. K. Warfield, "Multi-subject registration for unbiased statistical atlas construction," in *Lecture Notes in Computer Science*, vol. 3216, 2004, pp. 655–662.

[35] P. Lorenzen, B. Davis, and S. Joshi, "Unbiased atlas formation via large deformations metric mapping." *Medical Image Computing and Computer-Assisted Intervention MICCAI*, vol. 8, no. 2, pp. 411–418, 2005.

[36] G. Grabner, A. L. Janke, M. M. Budge, D. Smith, J. Pruessner, and D. L. Collins, "Symmetric atlasing and model based segmentation: an application to the hippocampus in older adults." *Medical image computing and computer-assisted intervention MICCAI*, vol. 9, no. 2, pp. 58–66, 2006.

[37] L. Ramus and G. Malandain, "Multi-atlas based segmentation: Application to the head and neck region for radiotherapy planning," in *MICCAI workshop on 3D Segmentation Challenge for Clinical Applications*, 2010.

[38] M. Sabuncu, B. Yeo, K. Van Leemput, B. Fischl, and P. Golland, "A generative model for image segmentation based on label fusion," *IEEE Transactions on Medical Imaging*, vol. 29, no. 10, pp. 1714–1729, 2010.

[39] S. K. Warfield, K. H. Zou, and W. M. Wells, "Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation," *IEEE Transactions on Medical Imaging*, vol. 23, no. 7, pp. 903–921, 2004.

[40] T. Rohlfing, D. B. Russakoff, and C. R. Maurer, Jr., "Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation," *IEEE Transactions on Medical Imaging*, vol. 23, no. 8, pp. 983–994, 2004.

[41] T. Rohlfing and C. Maurer, "Shape-based averaging," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 153 –161, 2007.

[42] O. Commowick and S. Warfield, "A continuous staple for scalar, vector, and tensor images: An application to dti analysis," *IEEE Transactions on Medical Imaging*, vol. 28, no. 6, pp. 838–846, 2009.

[43] O. Commowick, A. Akhondi-Asl, and S. K. Warfield, "Estimating a reference standard segmentation with spatially varying performance parameters: Local MAP STAPLE," *IEEE Transactions on Medical Imaging*, no. 99, 2012, in Press.

[44] G. Rohde, A. Aldroubi, and B. Dawant, "The adaptive bases algorithm for intensity-based nonrigid image registration," *IEEE Transactions on Medical Imaging*, vol. 22, no. 11, pp. 1470–1479, Nov. 2003.

[45] S. Gorthi, M. Bach Cuadra, U. Schick, P.-A. Tercier, A. S. Allal, and J.-P. Thiran, "Fusion of multi-atlas segmentations with spatial distribution modeling," in *MICCAI Workshop on Multi-Atlas Labeling and Statistical Fusion*, 2011.

[46] S. Gorthi, M. Bach-Cuadra, U. Schick, P.-A. Tercier, A. S. Allal, and J.-P. Thiran, "Evaluation of atlas fusion strategies for segmentation of head and neck lymph nodes for radiotherapy planning," in *Proceedings of the IEEE International Symposium on Biomedical Imaging*, 2012.

[47] S. Gorthi, M. Bach Cuadra, and J.-P. Thiran, "Exporting contours to DICOM-RT Structure Set," *Insight Journal*, 2009. [Online]. Available: http://hdl.handle.net/1926/1521

[48] S. Gorthi, M. Bach Cuadra, J. Villafruela Vicario, P.-A. Tercier, J.-P. Thiran, and A. S. Allal, "Multi-atlas fusion methods for segmentation of head and neck lymph nodes for radiotherapy planning," in *Proceedings of the 2nd ESTRO Forum*, 2013.

[49] S. Gorthi, J.-P. Thiran, and M. Bach Cuadra, "Comparison of energy minimization methods for 3-D brain tissue classification," in *Proceedings of IEEE International Conference on Image Processing*, 2011, pp. 57–60.

[50] M. Bach Cuadra, O. Cuisenaire, R. Meuli, and J. Thiran, "Automatic segmentation of internal structures of the brain in MR images using a tandem of affine and non rigid registration of an anatomical brain atlas," in *IEEE International Conference on Image Processing*, 2001, pp. 1083–1086.

[51] R. Bajcsy and S. Kovacic, "Multiresolution elastic matching," *Computer Vision, Graphics, and Image Processing*, vol. 46, no. 1, pp. 1–21, April 1989.

[52] M. Bro-Nielsen and C. Gramkow, "Fast fluid registration of medical images," in *International Conference on Visualization in Biomedical Computing*. London, UK: Springer-Verlag, 1996, pp. 267–276.

[53] G. E. Christensen, R. D. Rabbitt, and M. I. Miller, "3D brain mapping using a deformable neuroanatomy," *Physics in Medicine and Biology*, vol. 39, no. 3, pp. 609–618, 1994.

[54] L. D. Collins, C. J. Holmes, T. M. Peters, and A. C. Evans, "Automatic 3-d model-based neuroanatomical segmentation," *Human Brain Mapping*, vol. 3, no. 3, pp. 190–208, 1995.

Bibliography

[55] J. C. Gee, M. Reivich, and R. Bajcsy, "Elastically deforming 3d atlas to match anatomical brain images," *Journal of computer assisted tomography*, vol. 17, no. 2, pp. 225–236, 1993.

[56] D. V. Iosifescu, M. E. Shenton, S. K. Warfield, R. Kikinis, J. Dengler, F. A. Jolesz, and R. W. McCarley, "An automated registration algorithm for measuring MRI subcortical brain structures," *NeuroImage*, vol. 6, no. 1, pp. 13–25, 1997.

[57] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast mr images." *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, August 1999.

[58] V. Duay, P.-F. D'Haese, R. Li, and B. Dawant, "Non-rigid registration algorithm with spatially varying stiffness properties," in *IEEE International Symposium on Biomedical Imaging: Nano to Macro*, 15-18 2004, pp. 408 – 411.

[59] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. V1, no. 4, pp. 321–331, January 1988.

[60] X. Bresson, "Image segmentation with variational active contours," Ph.D. dissertation, EPFL, Lausanne, 2005.

[61] M. Bertalmío, G. Sapiro, and G. Randall, "Morphing active contours," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 7, pp. 733–737, 2000.

[62] M. Moelich and T. Chan, "Joint segmentation and registration using logic models," UCLA CAM Report, Tech. Rep., 2003.

[63] G. Unal and G. Slabaugh, "Coupled pdes for non-rigid registration and segmentation," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, 2005, pp. 168–175.

[64] J. H. An, Y. Chen, F. Huang, D. Wilson, and E. Geiser, "A variational PDE based level set method for a simultaneous segmentation and non-rigid registration." in *Medical Image Computing and Computer-Assisted Intervention*, vol. 8, 2005, pp. 286–293.

[65] Y.-N. Young and D. Levy, "Registration-based morphing of active contours for segmentation of CT scans," *Mathematical Biosciences and Engineering*, vol. 2, no. 1, pp. 79–96, 2005.

[66] M. Droske, W. Ring, and M. Rumpf, "Mumford-shah based registration: A comparison of a level set and a phase field approach," *Computing and Visualization in Science*, vol. 12, no. 3, pp. 101–114, 2009.

[67] N. Paragios, M. Rousson, and V. Ramesh, "Matching distance functions: A shape-to-area variational approach for global-to-local registration," in *European Conference on Computer Vision (ECCV)*, 2002, pp. 775–789.

[68] T. Amiaz and N. Kiryati, "Piecewise-smooth dense optical flow via level sets," *International Journal of Computer Vision*, vol. 68, no. 2, pp. 111–124, 2006.

[69] N. Paragios, M. Rousson, and V. Ramesh, "Non-rigid registration using distance functions," *Computer Vision and Image Understanding*, vol. 89, no. 2-3, pp. 142–165, 2003.

[70] Y. Chen, H. D. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. S. Gopinath, R. W. Briggs, and E. A. Geiser, "Using prior shapes in geometric active contours in a variational framework," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 315–328, 2002.

[71] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, 1994.

[72] O. B. Wijers, P. C. Levendag, T. Tan, E. B. Van Dieren, J. Van Sornsen De Koste, H. Van Der Est, S. Senan, and P. J. C. M. Nowak, "A simplified CT-based definition of the lymph node levels in the node negative neck," *Radiotherapy and Oncology*, vol. 52, no. 1, pp. 35–42, 1999.

[73] S. Song, W. A. Tome, M. P. Mehta, and P. M. Harari, "Emphasizing conformal avoidance versus target definition for IMRT planning in H&N cancer," *International Journal of Radiation Oncology.Biology.Physics*, vol. 57, no. 2, Supplement 1, pp. S299–S300, Oct. 2003.

[74] V. Grégoire, P. Levendag, K. K. Ang, J. Bernier, M. Braaksma, V. Budach, C. Chao, E. Coche, J. S. Cooper, G. Cosnard, A. Eisbruch, S. El-Sayed, B. Emami, C. Grau, M. Hamoir, N. Lee, P. Maingon, K. Muller, and H. Reychler, "CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC, RTOG consensus guidelines," *Radiotherapy and Oncology*, vol. 69, no. 3, pp. 227–236, 2003.

[75] J. Rogowska, K. Batchelder, G. Gazelle, E. Halpern, W. Connor, and G. Wolf, "Evaluation of selected two-dimensional segmentation techniques for computed tomography quantification of lymph nodes," *Investigative Radiology*, vol. 13, 1996.

[76] D. M. Honea and W. E. Snyder, "Three-dimensional active surface approach to lymph node segmentation," K. M. Hanson, Ed., vol. 3661, no. 1.   SPIE, 1999, pp. 1003–1011.

[77] J. Yan, T.-g. Zhuang, B. Zhao, and L. H. Schwartz, "Lymph node segmentation from CT images using fast marching method," *Computerized Medical Imaging and Graphics*, vol. 28, no. 1-2, pp. 33–38, Jan. 2004.

Bibliography

[78] J. Yan, B. Zhao, L. Wang, A. Zelenetz, and L. H. Schwartz, "Marker-controlled watershed for lymphoma segmentation in sequential CT images," *Medical Physics*, vol. 33, no. 7, pp. 2452–2460, 2006.

[79] 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: Segmentation of neck lymph nodes," *Academic Radiology*, vol. 14, no. 11, pp. 1389–1399, Nov. 2007.

[80] C.-C. Teng, L. G. Shapiro, and I. Kalet, "Head and neck lymph node region delineation using a hybrid image registration method," in *2006 3rd IEEE International Symposium on Biomedical Imaging: From Nano to Macro - Proceedings*, vol. 2006, 2006, pp. 462–465.

[81] D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellen, and W. Eubank, "PET-CT image registration in the chest using free-form deformations," *IEEE Transactions on Medical Imaging*, vol. 22, no. 1, pp. 120–128, 2003.

[82] S. Ourselin, A. Roche, S. Prima, and N. Ayache, "Block matching: A general framework to improve robustness of rigid registration of medical images," *Proc. of Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 557–566, 2001.

[83] O. Commowick and G. Malandain, "Evaluation of atlas construction strategies in the context of radiotherapy planning," in *Proceedings of the SA2PM Workshop (From Statistical Atlases to Personalized Models)*, Copenhagen, October 2006, held in conjunction with MICCAI 2006.

[84] A. Popovic, M. Engelhardt, and K. Radermacher, "Segmentation of skull-infiltrated tumors using itk: Methods and validation," *Insight Journal*, 2005.

[85] K. H. Zou, S. K. Warfield, A. Bharatha, C. M. C. Tempany, M. R. Kaus, S. J. Haker, W. M. Wells III, F. A. Jolesz, and R. Kikinis, "Statistical validation of image segmentation quality based on a spatial overlap index," *Academic Radiology*, vol. 11, no. 2, pp. 178–189, 2004.

[86] D. Huttenlocher, D. Klanderman, and A. Rucklige, "Comparing images using the Hausdorff distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 9, pp. 850–863, September 1993.

[87] F. Wilcoxon, "Individual comparisons by ranking methods," *Biometrics Bulletin*, vol. 1, no. 6, pp. 80–83, 1945.

[88] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed - algorithms based on hamilton-jacobi formulations," *Journal of Computational Physics*, vol. 79, no. 1, pp. 12–49, 1988.

[89] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," *International Journal of Computer Vision*, vol. 46, no. 3, pp. 223–247, 2002.

[90] E. Hui, S. Mohamed, M. Salama, and A. Fenster, "Prostate TRUS image region-based feature extraction and evaluation," in *Image Analysis and Recognition*. Springer Berlin / Heidelberg, 2009, vol. 5627, pp. 759–771.

[91] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the mumford and shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.

[92] H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, "A variational level set approach to multiphase motion," *Journal of Computational Physics*, vol. 127, no. 1, pp. 179–195, 1996.

[93] J. Lie, M. Lysaker, and X. . Tai, "A variant of the level set method and applications to image segmentation," *Mathematics of Computation*, vol. 75, no. 255, pp. 1155–1174, 2006.

[94] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[95] V. Duay, S. Luti, G. Menegaz, and J. Thiran, "Active contours and information theory for supervised segmentation on scalar images," in *15th European Signal Processing Conference*, 2007, pp. 1769–1773.

[96] M. Toews and W. Wells, "Bayesian registration via local image regions: Information, selection and marginalization," in *Information Processing in Medical Imaging*, 2009, pp. 435–446.

[97] M. Miller, S. Joshi, and G. Christensen, "Large deformation fluid diffeomorphisms for landmark and image matching," *Brain Warping*, pp. 115–131, 1999.

[98] G. Christensen, R. Rabbitt, and M. Miller, "Deformable templates using large deformation kinematics," *IEEE Transactions on Image Processing*, vol. 5, no. 10, pp. 1435–1447, 2002.

[99] S. Marsland and C. Twining, "Constructing diffeomorphic representations for the groupwise analysis of nonrigid registrations of medical images," *IEEE Transactions on Medical Imaging*, vol. 23, no. 8, pp. 1006–1020, 2004.

[100] M. Beg, M. Miller, A. Trouvé, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *International Journal of Computer Vision*, vol. 61, no. 2, pp. 139–157, 2005.

[101] B. Avants, C. Epstein, M. Grossman, and J. Gee, "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain," *Medical image analysis*, vol. 12, no. 1, pp. 26–41, 2008.

[102] G. Christensen and H. Johnson, "Consistent image registration," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 568–582, 2002.

[103] D. J. Burr, "A dynamic model for image registration," *Computer Graphics and Image Processing*, vol. 15, pp. 102–112, 1981.

[104] N. Houhou, V. Duay, A. S. Allal, and J.-P. Thiran, "Medical images registration with a hierarchical atlas," in *13th European Signal Processing Conference*, 2005.

[105] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, "A comparative study of energy minimization methods for markov random fields with smoothness-based priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 6, pp. 1068–1080, 2008.

[106] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.

[107] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 2, pp. 147–159, 2004.

[108] 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*, vol. 26, no. 9, pp. 1124–1137, 2004.

[109] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient belief propagation for early vision," *International Journal of Computer Vision*, vol. 70, no. 1, pp. 41–54, 2006.

[110] M. F. Tappen and W. T. Freeman, "Comparison of graph cuts with belief propagation for stereo, using identical MRF parameters," in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2, 2003, pp. 900–907.

[111] M. J. Wainwright, T. S. Jaakkola, and A. S. Willsky, "Map estimation via agreement on trees: Message-passing and linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 11, pp. 3697–3717, 2005.

[112] V. Kolmogorov, "Convergent tree-reweighted message passing for energy minimization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 10, pp. 1568–1583, 2006.

[113] N. Komodakis and G. Tziritas, "Approximate labeling via graph cuts based on linear programming," *IEEE transactions on pattern analysis and machine intelligence*, pp. 1436–1453, 2007.

[114] F. van der Lijn, T. den Heijer, M. Breteler, and W. Niessen, "Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts," *NeuroImage*, vol. 43, no. 4, pp. 708–720, 12 2008.

[115] M. Mignotte, "A label field fusion bayesian model and its penalized maximum rand estimator for image segmentation," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1610 –1624, 2010.

[116] N. Robitaille and S. Duchesne, "Hippocampal occurrence aging models obtained through label fusion," in *MICCAI Worskshop on Multi-Atlas Labeling and Statistical Fusion*, 2011.

[117] X. Zhuang, K. Leung, K. Rhode, R. Razavi, D. Hawkes, and S. Ourselin, "Whole heart segmentation of cardiac MRI using multiple path propagation strategy," in *Medical image computing and computer-assisted intervention MICCAI*, 2010, pp. 435–443.

[118] K. Pohl, J. Fisher, M. Shenton, R. McCarley, W. Grimson, R. Kikinis, and W. Wells, "Logarithm odds maps for shape representation," in *Medical Image Computing and Computer-Assisted Intervention*, 2006, pp. 955–963.

[119] M. Cardoso, G. Winston, M. Modat, S. Keihaninejad, J. Duncan, and S. Ourselin, "Geodesic shape-based averaging," in *Medical Image Computing and Computer-Assisted Intervention*, 2012, pp. 26–33.

[120] D. Zosso and J.-P. Thiran, "Geodesic Active Fields," Ph.D. dissertation, 2011.

[121] S. Birchfield and C. Tomasi, "On the statistical analysis of dirty pictures (with discussion)," *J. Royal Statistical Soc., Series B*, vol. 20, no. 4, pp. 259–302, 1986.

[122] M. Bach Cuadra, L. Cammoun, T. Butz, O. Cuisenaire, and J.-P. Thiran, "Comparison and validation of tissue modelization and statistical classification methods in T1-weighted MR brain images," *IEEE Transactions on Medical Imaging*, vol. 24, no. 12, pp. 1548– 1565, 2005.

[123] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden markov random field model and the expectation-maximization algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, no. 1, pp. 45–57, 2001.

[124] Z. Song, N. Tustison, B. Avants, and J. Gee, "Adaptive graph cuts with tissue priors for brain MRI segmentation," in *IEEE International Symposium on Biomedical Imaging: Nano to Macro, 2006.*, 2006, pp. 762–765.

[125] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sied, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE Transactions on Medical Imaging*, vol. 17, no. 3, pp. 463–468, 1998.

Bibliography

[126] 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*, vol. 13, no. 5, pp. 856–876, 2001.

[127] "Head and Neck auto-segmentation challenge 2010." [Online]. Available: http://www.grand-challenge.org/index.php/All_Challenges

Curriculum Vitae

Subrahmanyam Gorthi

Professional Address	Private Address
ELD239, EPFL-STI-IEL-LTS5, Signal Processing Laboratory (LTS5), Institute of Electrical Engineering (IEL), Swiss Federal Institute of Technology Lausanne (EPFL), Switzerland. **Tel:** +41-216934682 Email: subrahmanyam.gorthi@epfl.ch	Chemin de Vermont 18, 1006, Lausanne, Switzerland. **Mobile:** +41-767154915 subrahmanyam.gorthi@gmail.com

Websites: http://lts5srv2.epfl.ch/~gorthi/index.html http://people.epfl.ch/subrahmanyam.gorthi

Research Interests

I have been working in the areas of **medical image registration and segmentation**.

More specifically, the main areas of my work are:
- Atlas-based image segmentation [J3]
- Variational formulation of image registration [J1]
- MRF-based fusion framework for atlas-fusion [C2, C5]
- Model-based image segmentation [B1]

Other areas of my work include:
- Discrete optimization methods (like Graph Cuts) [C6]
- GPU implementation of convex segmentation methods [T1]
- Software development for exporting radiotherapy structures to DICOM format [J2]

My past research areas include mathematical modeling and analysis of MEMS actuators [J4].

Education and Employments

Oct 2008 – till date	**PhD in Medical Imaging**
	Signal Processing Laboratory (LTS5),
	Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
	<u>Supervisor</u>: Prof. Jean-Philippe Thiran

Oct 2008 – till date **PhD in Medical Imaging**
Signal Processing Laboratory (LTS5),
Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
<u>Supervisor</u>: Prof. Jean-Philippe Thiran

Sept 2007 – Oct 2008 **Research Assistant**
Signal Processing Laboratory (LTS5),
Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland
<u>Supervisor</u>: Prof. Jean-Philippe Thiran

July 2006 – Aug 2007 **Software Engineer**
SoftJin Technologies Private Limited,
Bangalore, India

Jan 2006 – June 2006 **Research Assistant**
Supercomputer Education and Research Centre (SERC), Indian Institute
Of Science (IISc), Bangalore, India
<u>Supervisors</u>: Prof. Atanu Mohanty and Prof. Anindya Chatterjee

2003 – 2005 **M.Sc. (Engg.) in Computational Science**
Supercomputer Education and Research Centre (SERC), Indian Institute
Of Science (IISc), Bangalore, India
<u>Supervisors</u>: Prof. Atanu Mohanty and Prof. Anindya Chatterjee

1998 -2002 **B.Tech. in Electrical and Electronics**
V.R. Siddhartha Engineering College, Nagarjuna University, Andhra
Pradesh, India

Publications

Please see the separate list at the end.

Academic Activities

- Reviewer for IEEE Transactions on Image Processing, Medical Image Analysis Journals, and for some of the Conferences in medical imaging.

- I have been a teaching assistant (TA) for various courses at EPFL, including "Image Processing," "Image Analysis and Pattern Recognition" and "Advanced Image Processing and Analysis."

- I assisted in writing a 3 year research proposal (2008-2011) to Swiss National Science Foundation (SNSF) that got accepted and was successfully completed.

122

Programming Skills

- Programming Languages: C, C++, Java, Python

- Technical Software: MATLAB

- Medical Imaging Specific: ITK

- Operating Systems: Linux, Windows

Projects Assisted

Master Projects:
- Shima Sepehri, EPFL Switzerland; "Convex formulation of region-based image registration," Sept 2010-Jan 2011.

- Arnaud Le Carvennec, ENS Cachan-antenne de Bretagne, Cachan cedex, France; "GPU implementation of pixel-based and region-based image registration," May 2011-Sept 2011.

Semester Projects:
- Fredederic Salvador, EPFL Switzerland; "Development of a software tool to register computed tomography and ultrasound images," Apr 2009 – Jun 2009**.**

- Arnaud Le Carvennec, ENS Cachan-antenne de Bretagne, Cachan cedex, France; "GPU accelerated convex multi-phase segmentation," Jul 2010-Sept 2010.

Personal Details

Date of Birth: June 7, 1981.

Citizenship: Indian.

Marital Status: Single.

List of Publications

Journal Articles

[J1] **S. Gorthi**, V. Duay, X. Bresson, M. Bach Cuadra and S. Castro, C. Pollo, A. S. Allal, and J.-Ph. Thiran. Active Deformation Fields: Dense Deformation Field Estimation for Atlas-based Segmentation using the Active Contour Framework, in Medical Image Analysis, vol. 15, num. 6, p. 787-800, 2011.

[J2] **S. Gorthi**, M. Bach Cuadra and J.-Ph. Thiran. Exporting Contours to DICOM-RT Structure Set, in Insight Journal, num. 1, 2009. (ITK Open Access/Open Review/Open Source Journal)

[J3] **S. Gorthi**, V. Duay, N. Houhou, M. Bach Cuadra and U. Schick, M. Becker, A. S. Allal, and J.-Ph. Thiran. Segmentation of head and neck lymph node regions for radiotherapy planning, using active contour based atlas registration, in IEEE Journal of selected topics in signal processing, vol. 3, num. 1, p. 135-147, 2009.

[J4] **S. Gorthi**, A. Mohanty and A. Chatterjee. Cantilever beam electrostatic MEMS actuators beyond pull-in, in Journal of Micromechanics and MicroEngineering, vol. 16, num. 9, p. 1800-1810, 2006.

Conference Papers

[C1] **S. Gorthi**, M. Bach Cuadra, J. Villafruela Vicario, P.A. Tercier, J.-Ph. Thiran, and A.S. Allal. Multi-atlas fusion methods for segmentation of head and neck lymph nodes for radiotherapy planning. 2nd ESTRO Forum 2013, Geneva, Switzerland, 2013.

[C2] **S. Gorthi**, M. Bach Cuadra, U. Schick, P.-A. Tercier and A. S. Allal, and J.-Ph. Thiran. Evaluation of Atlas Fusion Strategies for Segmentation of Head and Neck Lymph Nodes for Radiotherapy Planning. IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, Spain, 2012.

[C3] O. Esteban-Sanz, **S. Gorthi**, G. Wollny, A. Daducci and M. J. Ledesma-Carbayo et al. Brain tissue segmentation on diffusion weighted magnetic resonance data. IEEE International Symposium on Biomedical Imaging (ISBI) 2012, Barcelona, Spain, 2012.

[C4] T. Kober, A. Roche, O. Esteban-Sanz, **S. Gorthi** and D. Ribes et al. Atlas-free brain tissue segmentation using a single T1-weighted MRI acquisition. International Society for Magnetic Resonance in Medicine (ISMRM), Melbourne, Australia, 2012.

[C5] **S. Gorthi**, M. Bach Cuadra, U. Schick, P.-A. Tercier and A. S. Allal, and J.-Ph. Thiran. Fusion of Multi-Atlas Segmentations with Spatial Distribution Modeling. MICCAI Worskshop on Multi-Atlas Labeling and Statistical Fusion, Toronto, Canada, 2011.

[C6] **S. Gorthi**, J.-Ph. Thiran and M. Bach Cuadra. Comparison of Energy Minimization Methods for 3-D Brain Tissue Classification. IEEE International Conference on Image Processing, Brussels, Belgium, IEEE International Conference on Image Processing ICIP, 2011.

[C7] **S. Gorthi**, M. Bach Cuadra, U. Schick, P.-A. Tercier and A. S. Allal, and J.-Ph. Thiran. Multi-Atlas based Segmentation of Head and Neck CT Images using Active Contour Framework. MICCAI workshop on 3D Segmentation Challenge for Clinical Applications, Beijing, China, 2010.

[C8] **S. Gorthi**, V. Duay, M. Bach Cuadra, P.-A. Tercier and A. S. Allal, and J.-Ph. Thiran. Active Contour-Based Segmentation of Head and Neck with Adaptive Atlas Selection. MICCAI workshop on 3D Segmentation Challenge for Clinical Applications, London, 2009.

[C9] M. Bach Cuadra, **S. Gorthi**, F. I. Karahanoglu, F. Salvador and A. Pica, H.P. Do, A. Balmer, F. Munier, and J.-Ph. Thiran. Model-based Segmentation and Image Fusion of 3D Computed Tomography and 3D Ultrasound of the Eye for Radiotherapy Planning. VIPIMAGE, Porto, Portugal, 2009.

[C10] V. Duay, N. Houhou, **S. Gorthi**, A. S. Allal and J.-Ph. Thiran. Hierarchical Image Registration with an Active Contour-Based Atlas Registration Model. 16th European Signal Processing Conference, Lausanne, 2008.

[C11] **S. Gorthi**, A. Mohanty and A. Chatterjee. MEMS Parallel Plate Actuators: Pull-in, Pull-out and Other Transitions. International Conference on MEMS and Semiconductor Nanotechnology, IIT Kharagpur, India, 2005.

Book Chapters

[B1] M. Bach Cuadra, **S. Gorthi**, F. I. Karahanoglu, B. Paquier and A. Pica et al. Model-based Segmentation and Fusion of 3D Computed Tomography and 3D Ultrasound of the Eye for Radiotherapy Planning, in Computational Vision and Medical Image Processing, Computational Methods in Applied Sciences 19, 2011, Springer 2011.

Technical Reports

[T1] **S. Gorthi**, A. Le Carvennec, H. Copponnex, X. Bresson and J.-Ph. Thiran. GPU-accelerated Convex Multi-phase Image Segmentation, 2011.