UNIVERSIDAD NACIONAL DE COLOMBIA

Multi-Atlas Structure Segmentation On Medical Images

Segmentación De Estructures En Imágenes Médicas Basada en Multi-Atlas

Henry Mauricio Orbes Arteaga

Universidad Nacional de Colombia
Faculty of Arquitecture and Engineering
Department of Electrical, Electronics and Computer Engineering
Manizales, Colombia
2016

Multi-Atlas Structure Segmentation On Medical Images

Segmentación De Estructures En Imágenes Médicas basada en Multi-Atlas

Henry Mauricio Orbes Arteaga

Thesis presented as partial requeriment to obtain the degree:
Master of Engineering

Advisor:
César Germán Castellanos Domínguez, PhD.
Co-advisor:
David Cárdenas Peña, M.Sc.

Research line:
Industrial Automation
Research Group:
Signal Processing and Recognition Group

Universidad Nacional de Colombia
Faculty of Arquitecture and Engineering
Department of Electrical, Electronics and Computer Engineering
Manizales, Colombia
2016

Dedicated to my family

Acknowledgements

I would like to thank everyone who contributed with this work, special thanks to my advisor, German Castellanos, for his constant support, and David Cardenas, my co-advisor for his motivation and guidance in this research process. Also all the people from the Signal Processing and Recognition Group. Finally, thank to my family for trusting me all the time.

This research was carried out under the projects

- 1110-569-34461 Desarrollo de un sistema efectivo y apropiado de estimación del volumen de tejido activo cerebral para el mejoramiento de los resultados terapéuticos en pacientes con enfermedad de parkinson intervenidos quirúrgicamente, funded by COLCIENCIAS.
- 1110-657-40687 Estimación de los parámetros de neuro modulación con terapia de estimulación cerebral profunda en pacientes con enfermedad de parkinson a partir del volumen de tejido activo planeado, funded by COLCIENCIAS.
- 22506 Desarrollo de un sistema automático de mapeo cerebral para el monitoreo y tratamiento de enfermedades neurodegenerativas funded by UNIVERSIDAD NACIONAL DE COLOMBIA.

Abstract

Structure segmentation is an important task in medical applications as it give a quantitative knowledge of volume, shape or location of the structures in consideration, enabling the understanding of several pathologies. Manually segmentation is considered the current gold standard to obtain accurate segmentation. However, this process is time and resource consuming becoming impractical for large database studies. Recently, Multi-atlas based methods have been used to support automate brain structure segmentations. The advantage of this methods relies on the capacity to provide spatial information while encoding anatomical variability by using a set of prelabeled atlases. However, the accuracy of segmentation depends on the capability of each atlas to propagate the labels to the target image as well as the employed methodology for fusing the label conflicts. In this sense, atlas selection and label fusion become important directions to improve the performance of muti-atlas segmentation. In the present work several approaches to enhance atlas selection and label fusion are proposed. Firstly, a kernel based representation is proposed aiming to map the original space domain in a low dimensional embedding space where latent groups in the data are highlighted, and the intrinsic similarities are better reflected. Secondly, a supervised similarity measure is proposed which take advantage of local similarities and supervised information for linking similarity appearance with the segmentation performance. Finally, label fusion is casting as non-local probabilistic atlas-based segmentation, where the strengths of patch based and atlas-based approaches are combined to improve the segmentation accuracy. The proposed approaches are compared with conventional state of the art techniques in segmentation task. Attained results show the potential of the proposed approach to improve the segmentation outperforming the conventional state of the art methods.

Keywords: Kernel-based representation, Embedding space, Supervised atlas selection, Probabilistic atlas construction, Patch based method segmentation)

Resumen

La segmentación de estructuras es muy importante en aplicaciones médicas ya que proporciona un conocimiento cuantitativo del volumen, forma o posición de las estructuras en consideración, lo cual permite el análisis y entendimiento de diferentes patologías. La técnica estándar de segmentación es la proporcionada manualmente por un experto clínico. Sin embargo este proceso es computacionalmente costoso dificultando el análisis en grandes bases de datos. Recientemente, los métodos basados en Multi-atlas han sido usados para apoyar las tareas de segmentación de imágenes cerebrales. La principal ventaja de estos métodos es debido a que son capases de proporcionar información espacial a su vez que la variabilidad anatómica es capturada mediante el uso de un conjunto de atlases etiquetados. Sin embargo la precisión de la segmentación depende de la capacidad de cada atlas para propagar sus etiquetas a la imagen objetivo, así como el método empleado para combinar las estimaciones de cada atlas. Por esta razón, la selección de atlas y el método de combinación o fusión de etiquetas, son dos importantes direcciones para mejorar el desempeño de los métodos basados en Multi-atlas. En este trabajo se proponen diferentes enfoques con el fin de mejorar la selección de atlases y la combinación de etiquetas. En primer lugar se propone un método de representación de imágenes médicas basado en Kernels la cual permite mapear el espacio original en un espacio embebido de baja dimensión donde se destacan los grupos latentes en los datos y refleje las similitudes intrínsecas. En segundo lugar, se propone una medida similitud supervisada entre imágenes, esta usa medidas de similitudes locales e información supervisada para correlacionar similitudes en apariencia con el desempeño en la segmentación. Por último, el problema de combinación de etiquetas es enmarcado como un método probabilístico no local de segmentación, donde se combinan las fortalezas de los métodos basados en parches y los métodos basados en atlas probabilísticos con el objetivo de mejorar la precisión de la segmentación. Los métodos propuestos son comparados con métodos convencionales del estado del arte en tareas de segmentación. Los resultados obtenidos muestran el potencial de los enfoques propuestos para mejorar la segmentación ya que superan los métodos convencionales.

Palabras Clave: Representación basada en Kernels, Espacio embebido, selección de atlas supervisada, construcción de atlas probabilísticos, métodos de segmentación basados en atlas.

Contents

Abstract 8

Resumen 9

1 Introduction 2
- 1.1 Motivation . . . 2
- 1.2 State of the art . . . 3
- 1.3 Problem Statement . . . 8
- 1.4 Objectives . . . 9
  - 1.4.1 General Objective . . . 9
  - 1.4.2 Specific Objectives . . . 9
- 1.5 Contributions . . . 10

2 A Kernel-based Representation to Support 3D MRI Unsupervised Clustering and Atlas Selection 12
- 2.1 Materials And Methods . . . 13
  - 2.1.1 Image Similarity Measure . . . 13
  - 2.1.2 Spectral clustering . . . 13
  - 2.1.3 Kernel Based Image Representation . . . 14
  - 2.1.4 Bayesian medical image segmentation . . . 14
  - 2.1.5 Learning model parameters . . . 15
- 2.2 Experimental setup . . . 16
  - 2.2.1 Preprocessing . . . 16
  - 2.2.2 Proposed Image Feature Extraction . . . 16
  - 2.2.3 Kernel-based Image Similarity . . . 17
- 2.3 Kernel Based Representation for clustering task . . . 17
  - 2.3.1 Database . . . 17
  - 2.3.2 Parameters Tuning and Results . . . 18
- 2.4 **K**ernel-based **A**tlas **I**mage **S**election from the **E**mbedding **R**epresentation (KAISER) . . . 23
  - 2.4.1 Database . . . 23
  - 2.4.2 Parameters Tuning and Results . . . 24

2.4.3 Tissue Labeling Performance . . . 25
2.5 Discussion . . . 26

3 Supervised Similarity Measure supporting Atlas-Selection 29
3.1 Materials and Methods . . . 30
3.1.1 Image similarity metrics . . . 30
3.1.2 Spatial Enhancement of Image Metrics . . . 31
3.1.3 Supervised Image Measure Learning . . . 32
3.1.4 Label Fusion . . . 33
3.2 Experimental Set-Up . . . 34
3.2.1 Database . . . 34
3.2.2 Image Preprocessing . . . 35
3.2.3 Metric parameter learning . . . 35
3.2.4 Evaluation of Similarity Metrics . . . 36
3.3 Discussion . . . 36

4 Target Specific Patch-atlas Construction 39
4.1 Methods . . . 40
4.1.1 Target specific probabilistic atlas segmentation . . . 40
4.1.2 Probabilistic atlas generation . . . 40
4.1.3 Iterative labeling . . . 42
4.2 Experiments . . . 45
4.2.1 Database Description . . . 45
4.2.2 Experimental results Target-specific atlas construction . . . 45
4.2.3 Fast multi-point and the hierarchical fast multi-point algorithm . . . 47
4.2.4 Influence of deformable registration . . . 49
4.3 Discussion . . . 51

5 Conclusions and future work 53
5.1 Conclusions . . . 53
5.2 Future work . . . 55

6 Publications 56

1 Introduction

1.1 Motivation

Medical images are of vital importance in medical applications. These provide to the clinicians the visualization of patient anatomy in a non-invasive way, supporting the diagnosis and treatment planning of several diseases. There are several techniques to obtain medical images being computed tomography(CT) and Magnetic Resonance (MR) the most common used modalities for 3D structural analysis. Depending on the application, each of them can be used taking into account their advantages and disadvantages. CT provided high resolution images and fast acquisition, however, this technique is based in ionizing radiation that might be harmful to the tissues and cause diseases such as cancer and eye cataract. Therefore, CT images are usually employed in emergency situations and other applications as cancer detection, bone injuries, and intensity modulate radiotherapy planning which became the state of the art method for the treatment of head-neck cancer. On the other hand, MR scans provide images with high resolution without damage the tissues in the acquisition process. However, the acquisition takes more time than CT generating some imaging artifacts. MR images has been widely used in detection and follow-up of the main brain pathologies such as Alzheimer Schizophrenia and Parkinson, enabling to estimate anatomical or functional structure changes [19]. Specifically in the sub-cortical structures as the (Hypothalamus (HYPO), Amygdala (AMYG), Putamen (PUT), Caudate Nucleus (CAUD), Thalamus(THAL), and Pallidum (PAL)). Image segmentation is one of the most important tasks in the above mentioned applications, as it provide a quantitative knowledge of volume, shape or location of the structures of interest which is essential for the analysis and understanding of pathologies, as well as allows find biomedical markers for neurological disorders, helping to diagnose any disease or its planning treatment. The segmentation task basically consists in assign a label to each voxel in the image according to a set of characteristic, in such way that voxels with the same label belong to the same structure. The current gold standard image segmentation is performed manually by expert clinicians which become an impractical task for large studies given that is very time and resource consuming. In addition, the quality in manual labeling depends on the performance of the expert clinician. Therefore, the inter and intra-rater observer bias produced by the clinicians segmentation hinders the statistical analysis of segmentations. For all these reasons, automatic

segmentation methods have emerged as an alternative to manual segmentation, becoming an important research area. The primary objective of this methods is obtaining robust and accurate segmentation improving the reliability and repeatability on the segmentation while the computational and human requirements are reduced. Nonetheless, this can be a difficult task due to the presence of noise and artifacts which affect overall image quality, restricting their performance to specific applications. Particularly, segmentation of subcortical and other structures within the deep brain are more complicated. This due mainly to the following issues: *i)* Surrounding structures can exhibit similar intensities. *ii)* Structures can be composed of different tissues which exhibit inhomogeneities on the intensities. *iii)* There are shape and size variations between the subjects. Hence, prior anatomical information is crucial for guiding the segmentation. In the last years, template-based techniques have been raised in an efficient manner to provide spatial information. These methods using one or a set the labeled training images to model spatial distributions. Even these have become the most successful path for obtaining accurate segmentation. There are still many limitations which bias their performance.

1.2 State of the art

Segmentation of medical images has been performed on several ways by different methods. Segmentation techniques vary according to the type of image information they used as well as the prior knowledge about the problem and internal constraints of the segmented geometry [64]. Therefore, the performance depends on the application and the type of used image[10]. Early segmentation methods were based on the direct modeling of the intensity characteristics present on the images, proving be robust in the segmentation the brain tissues such the white and gray matter. These methods assume that interested structures or organs have distinctive quantifiable features such as the intensity or gradient magnitude [54]. In this sense, voxels can be classified according to an established threshold or edge based decision rule. Examples of this type of methods are the watersheds algorithms where intensity and edge information are combined, or the Region Growing method[1] which introduce a decision rule based on voxel connectivities. More sophisticated methods employ pattern recognition techniques aiming to cluster and classify the voxels within the regions. In this sense segmentation can be carried out in unsupervised or supervised way. Common unsupervised approaches include fuzzy C-means [15],Data analysis techniques[17] and unsupervised neural network [16] these algorithms alternate between segmenting the image and characterizing the properties of each class on-line. On the other hand, supervised methods such as Neuronal networks [2], support vector machine[79] and active appearance models[31] require training samples for its initialization. Therefore, the accuracy on these methods is sensitive to the initial set of training samples and the initial conditions. On the other hand,

Statistical classifiers also are used for segmentation. These approaches are based on the modeling of intensity distributions where the labels are estimated according to probability values derived from the employed model. Generally, finite mixture models are used for intensity modeling, several approaches include Gaussian mixture models, a parzen windows and Markov random fields [88]. Although intensity-based segmentation methods have proven to be successful for specific segmentation tasks such gray and withe matter classification. Segmentation of sub-cortical structures requires from prior spatial information, due to the low inter-structure contrast and lack of clearly defined edges between some brain structures which bias the performance on the intensity-based methods. Hence,spatial functions termed as atlas or templates have been proposed for modeling structure distribution. To this end templates are provided as a set of shape, intensity and/or functional models. In this sense, the objective of atlas-based segmentation is to transfer the prior information provided by an atlas to the unlabeled target image. For the majority of the atlas-based methods an atlas is defined as an image and its corresponding segmentation performed manually. Early methods use a single atlas chosen for being representative for an determined population, then, labels are transferred from the atlas to the unlabeled images. For this purpose, the atlas is registered (deformed) to the target image and the obtained transformation is used for mapping the structure labels from the atlas to the target space, this process is known as label propagation [53]. In this sense, segmentation becomes a registration problem, consequently, atlas based methods have driven their efforts on the development of new registration techniques. Generally, a complete registration process needs a linear transformation for global alignment, and the local deformations to deal with anatomical differences. Although, linear alignment is enough for intra-subject medical applications [60], for the majority of medical applications that requires inter-subject analysis non-linear registrations is indispensable. Several methods in the literature perform non-linear registration using mathematic transformation as as B-spline curves, level set partial differential equations, or transformation based on physical models as elasticity [14, 40, 55, 44, 34] , deformation vectors [27], fluid mechanics [24], thermodynamics [35], optical flow [47]. Other approaches propose new cost functions as second order mutual information or gradient mutual information. A complete review of registration methods is presented on [46] where the importance of registration accuracy on atlas-based segmentation methods is analyzed and conclude that symmetric normalization(SyN) [13] algorithm perform the best. However, single-atlas based methods assume that exact correspondence can be achieved between the atlas and the target image after registration. Therefore, the performance on these methods is limited due to a single atlas is insufficient to capture wide anatomical variations, doing the registration errors more evident especially when the atlas is unrepresentative of the target image. To solve this problem more recent methods combined information of different subjects aiming to construct probabilistic atlas [59]. To this end, all templates in the population are co-register to a common space where probabilistic models for each structure can be inferred. To segment a new image this needs to be registered to the atlas space, then the segmentation can be estimated

by an statistical framework. For instance, [52] incorporate probabilities for each tissues as an energy term in a grap-cut segmentation framework. In [39] the volume is defined as a Markov Random field(MRF) and the parameters are optimized using iterative conditional models, then the segmentation is sequentially actualize until the optimum is found. [26, 45] combine atlas probabilities with other image features such as voxel intensities or spatial coordinates to train a classifier. The main advantage of probabilistic-atlas-based methods is that only a single registration is required to segment a new image. However, find the common space for the atlas can be difficult, especially if high anatomical variability exist on the atlas set[64]. On the other hand, Multi-atlas approaches have proposed other alternative to deal with anatomical variability when multiples atlas are available. On this type of methods, each atlas is registered with the target image and the individual propagations are combined to obtain a final segmentation. In this way, the effects of the registration errors are minimized while the anatomical variability is better modeled. This methods have prove be robust for segmenting different structures such as the hippocampus, head and neck , prostate and abdomen, becoming the baseline for obtaining accurate segmentation [10]. Multi-atlas segmentation was first proposed by [63, 64] and adapted to medical applications in [47, 42]. The most straightforward strategy consist on register all the atlases to the target image and assign to each voxel the label with the most agreement, this scheme is commonly known as majority voting. However, there are a lot of more sophisticated methods on the state of the art which employs techniques of computer vision and machine learning aiming to take full advantage of the set of atlas. Nevertheless, Two issues constantly predominate on multi-atlas segmentation, the first one refers to which and how many atlases are necessary in the scheme. The second one concerns to how combine them in final segmentation to obtain the highest accuracy. Consequently, state of the art methods mainly focusing their contribution on i) Atlas selection, and ii) label fusion strategies.

Atlas selection is crucial in multi-atlas methods because not only can improve the segmentation accuracy but also reduce the computational cost. Given that label propagation requires non-liner registration which is usually expensive, keep the lower number of atlas can significantly improve the computational efficiency. On the other hand, if the most similar atlases to the target image are selected, effects of unrepresentative atlases are avoided improving the final segmentation. In the earliest work[63] just the most similar atlas from the dataset was used for label propagation. However, in [4, 42] has been proved that using more atlases improve the segmentation. Generally, selection is supported by an image similarity metric such a mutual information cross correlation or means squares [6, 86, 3, 4] which are computed on voxel-wise basis over the whole image or in structure basis as in [84] where mutual information is used to measure similarities on determined regions. Successful of similarity measures depends of some factors as was studied by [53, 3], for instance mean squares and cross correlation provides accurate selection if the image intensities have been normalized, meanwhile, mutual information does not require image normalization being more appropri-

ate for multi-modal studies or when the images are obtained from different scanners. Also, meta information such age have displayed reliable selection[4] however this information is not always available. On the other hand, some authors propose to use information from the registration process, such the deformation field (e.g) [29] propose to measure the amount of deformation using the euclidean norm, meanwhile [61] use the the Jacobian determinant or its harmonic energy [78]. However, use deformation information implies that deformable registration has to be performed before selection, therefore, the computational cost is not reduced. In this sense, similarity measures have an advantage as they can be computed in the a common space. However, these are computed on the high dimensional image space, where the intrinsic morphological properties may not be highlighted[20]. Hence, new methodologies have been introduced to project the images in a low dimensional space where intrinsic similarities can be better reflected. Techniques as manifold learning [21] and locality preserving projections [20] or principal component analysis [12] have been used in this regard. On the other hand, selection can be driven as a clustering problem where off-line or on-line clusters can be identified for the most similarity atlas. [57, 48] The main contribution of these approaches is that clusters most representative from the target image can be used for initializing a more refined selection, especially when a large databases is available. Other approaches use information on the manual segmentation to guide the appropriate selection as a learning problem[68]. In this sense, atlas are ranked based on the expected performance for the final segmentation instead of high similarities, providing a more suitable selection given that higher similarity not necessarily represent the most accurate segmentation. Nevertheless, current learning methods are not generalized beyond of the training data [43].

Posterior to atlases selection, label propagations have to be combined to obtain a final segmentation in a label fusion scheme. The simplest strategy known as majority voting has been widely used in several approaches proving be robust [47, 42]. However, on this scheme, all atlases have the same contribution on the final segmentation which can be undesirable due to the presence of outliers. This drawback can be overcome by weighted voting, where the atlases are weighted based on a performance or similarity criterion. In this regard, weights can be computed according to intensity similarity measures between the atlases and the target image as in atlas selection. Normalize mutual information was used in [5], also square differences and normalize cross correlation were evaluate in [6, 23]. Other more sophisticated strategies compute global weights based on shape differences [65] or solving a cost function of minimal reconstruction error [21] where it is assumed that the target image intensities are equal to the lineal combination of the atlas intensities. However, in the aforementioned methods weights are computed globally, so local differences are not captured. As an alternative, local weighting strategies propose to compute spatially varying weights for each atlas. For this purpose, similarity measures can be assessed on a local neighborhood for each voxel as in [6], they evaluate the performance of global and local weighting using Cross correlation, Normalize mutual information, and square differences, and conclude that local

weights perform the best. In [77, 76] optimal weights are computed by taking into account the correlation of errors produced by different atlas. Also, local weighting can be derived from a probabilistic models of the data e.g [67] employs a generative model in a voxel basis which stands label fusion as a classification problem. Meanwhile, in [51] spatial structure and appearance models are combined with a regularization term in a probabilistic framework. In contrast, atlas weighting has also been introduced in an statistical context. The well know approach Simultaneous Truth and Performance Level Estimation (STAPLE) iteratively estimate the performance for each atlas using the expectation maximization[80], this methodology has been widely studied aiming to overcome some limitations in the original method. Some variations incorporate prior probabilities [30], or intensity probability map instead of natural labels [81], as well as, spatially varying performance parameters aiming to account for spatial varying difficulty [8, 28], also, STAPLE was combined with atlas selection and intensity information (SIMPLE) [49]. However, the above mentioned label fusion methods typically assume that the target image is a voxel-wise representation of multiples atlases [11]. Therefore, the labels are estimated by direct correspondence from the intensity and label information after non-linear registration. In consequence, these methods are highly dependent on the accuracy of the employed registration method, which is challenging mainly due to anatomical differences often exist. Recently, non-local patch based strategies have been proposed to alleviate effects of registration errors. In these methods, images are represented by local volumetric patches, then, relationships between these patches are used to seek the label which have a perfect correspondence with each voxel in the target image. In the label fusion step, each target voxel is labeled by using a set of candidate patches which are extracted from a local neighborhood on the training dataset. There are two ways to use the patch dictionary in label fusion. The first correspond to label a target patch by weighting the contributions of all candidate patches. weighting can be estimated by a intensity similarity criterion [33], or based on the reconstruction of the target patch, where the target patch can be seen as an sparse linear combination from the patch dictionary. [87, 71], one advantage of sparse reconstruction is that effects of the misleading patches are minimized. However, reconstruction and similarity based approaches rely on the assumption that patches with similar appearance could have the same anatomical label. Nonetheless, patch similarity and label affinity among voxels may be unrelated especially when only the central voxel in the patch is taken into account for segmentation[70]. On the other hand the patch dictionary and their corresponding labels can be used to learn a classification function [41]. However, all the patches in the dictionary, have the same contribution on the learning procedure [69].

1.3 Problem Statement

There are several issues that hinder automatic segmentation, these vary depending on the application an can be more relevant for a determined modality or the kind of structure to be segmented. Common artifact such noise streaks or motion artifacts are originated on the acquisition process. Although, there are techniques to reduce their effects, these artifacts affect the quality of the images, doing the identification of structures more complicated. On the other hand, the major challenge for segmentation approaches is due to some structures present similar intensity values. Particularly, in the subcortical ones there is low contrast between structures and the borders are poorly defined, besides some structures are composed by different tissue types and there are large anatomical variability for some structures such the hippopotamus. In addition, segmentation of parotid glands or brain stem on CT images has similar challenges. Parotid glands exhibit large variations in anatomy, meanwhile the segmentation of brainstem is challenging due to surrounding structures have similar intensity. For all above mentioned reasons, segmentation can not be accurately performed based on intensity information. Hence, spatial information is required for guiding the segmentation. multi-atlas strategies have been raised as an effective way to provide spatial information and modeling the anatomical variability, becoming the state of the art on automatic segmentation. However, the successful of these schemes mainly dependent on atlas selection and label fusion steps which are challenging mainly for the next issues.

- Generally, label fusion and atlas selection involve finding similarities or relationships between images which have a high dimensional space. This fact affects not only the computational efficiency on the algorithm but also assessing similarities in this high dimensional space could yield undesirable results due to intrinsic similarities are not accurately reflected in the high dimensional space [21]. Moreover, similarities are usually assessed voxel-to-voxel, hence relationships that might exist among the measurements on different voxels or regions are obviated. On the other hand, images can be projected onto a low dimensional-manifold space [82, 20]. However, the choice of the manifold function, as well as the manifold parameters is not generalized [36].

- The aim of selection is to find the atlases which provide the best propagate labels. However, when a new target image is evaluated, any information about its segmentation is known. Hence, Comparison between atlases relies on the intensity appearance similarities under the assumption that the most similar in appearance would provide the best segmentation. However, as have been demonstrated in [68], there are not always a correlation between similarity appearance and the performance to segment a target image. On the other hand, incorporate previous knowledge about of inter-atlas segmentation reliability usually lacks of generalization for extending to the target image. Moreover, if the entire image is used to compute similarities or for constructing a

manifold, pairwise affinity could be biased for undesirable regions.

- Label fusion strategies generally depend on accurate registration, as the labels are voxelwise propagated from atlases to the target image, exact correspondence should be achieved after registration. However, registration errors are often present in real world applications limiting the performance of segmentation. Patch based methods minimize registration errors due to do not rely on voxelwise correspondence. In these approaches each voxel is represented by a volumetric patch and a set of non-local patches is searched on the training atlas for label fusion. However, there are some issues that could affect the performance of label fusion strategy. In general, patch-based methods assume that patches with similar appearance have the same label. However, patch similarity and label affinity may be unrelated, for instance, on the boundaries background and structure patches share high appearance similarity, also, if surrounding structures have similar intensity values, wrong labels could produce inaccurate results. In addition, current methods treat each patch independently for label fusion, and labels are not taken into account to measure their contribution. Finally, all regions in the image have the same importance, therefore, the same search and fusion strategy is used for all points on the image. However, labeling difficulty vary spatially and some regions with high variability could need a better estimation.

1.4 Objectives

Based on the state of the art and their corresponding open issues, we develop the current research work around one general objective which is further expanded into three specific objectives, as follow:

1.4.1 General Objective

To develop a Multi-atlas segmentation framework capable of obtaining accurate segmentation for several medical applications by improving selection and fusion stages on a multi-atlas scheme.

1.4.2 Specific Objectives

- To propose a strategy for MRI representation in order to map images from a high-dimensional space in a low dimensional space where the intrinsic morphological sim-

ilarities can be better reflected, and the data distribution is highlighted, enabling to find clusters which are representative of a determinate population.

- To develop a supervised similarity measure that incorporates label information and local measurements aiming to learn the relationship between the local appearance similarities and the segmentation performance. In this sense, the similarity measure is used to select a set of atlas based on their expected performance to label a target image.

- To develop a new strategy for label fusion capable of accounting for registration errors and spatial labeling difficulty, while, intensity and label information are incorporated in the fusion step. To this end, non-local patch searches are used in a probabilistic atlas construction, and the label fusion is carried out by a Bayesian inference.

1.5 Contributions

Due to the capability to model anatomical variability Multi-atlas segmentation has become one of the most successful strategies to obtain accurate segmentation. On this thesis, we focus in improving atlas selection and label fusion which are the main components on multi-atlas schemes. Special attention is put in perform atlas selection in a common space thus, not only the best atlas are selected for label propagation, but the computational cost is reduced significantly. On the other hand, an appropriate strategy for label fusion is mandatory to obtain accurately segmentation. We focus especially on patch based method due to the capability to deal with registration errors. Although, this methods are expensive in some applications, this can be combined with atlas selection to reduce the computational cost. The main contributions of this thesis are as follow.

- A new **Kernel based MRI represetation** is proposed in order to obtain a low-dimensional space where pairwise similarities can be computed. The proposed approach, use an effective way to extract the main shape information based on inter-slice similarities. Moreover, a low dimensional space is constructed enabling to identify demographic clusters such the gender and age. In this sense, our proposal is suitable to support MRI clustering which is helpful for pre-selection of images on large databases allowing to construct patient specific atlases. On the other hand, the proposal provides a powerful tool to compute pairwise image similarities in an embedding low dimensional space where the intrinsic similarities can be better reflected, enabling to support atlas selection in Multi-atlas schemes.

- Spatial similarities are introduced in a **Supervised similarity measure** aiming to

relate the similarity appearance of the images with the performance to segmentation. While the traditional selection of atlas is based on similarity measures, has been demonstrated that similarity appearance is not always correlated with the performance of segmentation. Moreover, as measures are evaluated over the whole images local differences in non-relevant regions can bias the measure. In this work, a supervised kernel matrix is constructed from the labeled training atlases to learn appropriately the combination of local similarities in such way that the resultant measure is related to the segmentation performance.

- Muti-atlas label fusion is addressed as a **probabilistic target patch atlas construction** scheme. Where strengths of both probabilistic atlas and patch based segmentation methods are combined to obtain accurate segmentation. Thus, effects spatial variability information are relaxed by patch local search in target atlas construction, meanwhile, label fusion is carried out by a Bayesian framework which intuitively incorporate intensity and label information in the fusion step. In addition, the proposed approach account for labeling task difficulty, enabling to refine the segmentation for regions where is harder obtain reliable segmentation. To this end, the segmentation is carried out by an iterative framework where segmentation for the confident regions are used to driving the segmentation for those less confident. Also, Hierarchical fast multipoint scheme is proposed to reduce the computational time where a few points are used to obtain the whole image segmentation.

2 A Kernel-based Representation to Support 3D MRI Unsupervised Clustering and Atlas Selection

Multi-atlas methods provide an effective strategy to deal with anatomical variability by using a set of manually labeled dataset. However, to obtain the final segmentation all the atlases are registered to the query image, therefore, the computational burden increase proportionally with the number of atlases on the database becoming impractical for large datasets. Moreover, the existence of unrepresentative atlases on the dataset could bias the segmentation [4]. Therefore, there is a need for selecting a smaller and more representative subset of atlases from a large set. A simple way for selection is based on demographic image information like demographic affinity[4]. In this way, the images could be clustered aiming to build probabilistic atlases to represent a given population. Nevertheless, demographic-based approaches assume that image distributions depend only on few considered grouping categories (e.g age and gender). Moreover, accounting for all possible demographic groups is a not tractable task. On the other hand, Atlas selection is usually based on intensity measures. However, measures are assessed in the image domain where estimator convergence can not be guaranteed on such high-dimensional spaces. As an alternative [18] use the mean shift algorithm to find modes on the dataset. Thus, clustering results are provided by means of multidimensional scaling. Nevertheless, the original image space is not shown to be compact enough to guarantee the mean estimator to converge. Moreover, there is not an interpretation of the resulting modes. This drawback can be overcome by introducing a high dimensionality reduction stage aiming to work in more compact spaces. Thus, [36] compares three manifold learning techniques for computing compact spaces where images are compared. Comparison of a new query against the dataset is estimated as the accumulation of the atlas-image-to-single-template and single-template-to-query image measures. This aspect may lead to a suboptimal selection, given that the employed measure is not guaranteed to match all distance properties. [50] Proses to reduce the number of features by grouping voxels into anatomical regions through the alignment of labeled atlases. Then a classification is performed using classical machine learning techniques. Nevertheless, all those approaches require a previous alignment stage so being constrained to the computational burden issues.

In this chapter, we propose to reduce the original feature space dimension using a kernel based MRI representation. The proposed approach provides an effective way to encode inter-slice similarities so that the main shape information is kept in a lower dimensional space. Additionally, a spectral clustering technique is employed to build a low-dimensional space where regular distances can be more accurately calculated, and the latent data structure is highlighted. The proposed methodology is used in a clustering task aiming to identify the demographic categories age and gender. In addition, pairwise similarities computed on the embedding space are used to select a set of the most representative atlases to segment an unlabeled target image. Finally, Multi-atlas segmentation is stated in a probabilistic framework where a target-specific probabilistic atlas is generated for each unlabeled target images using the selected set of atlases.

2.1 Materials And Methods

2.1.1 Image Similarity Measure

Let $\mathcal{X} = \{\boldsymbol{X}_n \in \Omega : n = 1, \ldots, N\}$, a dataset that holds N MR images. In order to encode the affinity between a couple of images $\{\boldsymbol{X}_n, \boldsymbol{X}_m\}$, we introduce the following kernel function.

$$\zeta(\boldsymbol{X}_n, \boldsymbol{X}_m) = \langle \varphi(\boldsymbol{X}_n), \varphi(\boldsymbol{X}_m) \rangle \tag{2-1}$$

where $\varphi(\cdot)$ maps from the original domain, Ω, into a Reproduced Kernel Hilbert Space $\mathcal{H}$. Notation $\langle \cdot, \cdot \rangle$ stands for the inner product. Generally, it holds that $|\mathcal{H}| \to \infty$, so that $|\Omega| \ll |\mathcal{H}|$ can be assumed. Nevertheless, there is no need for computing $\varphi(\cdot)$ directly. Instead, the well-known *kernel trick* is employed for computing the elements of the matrix $\boldsymbol{Z} \in \mathbb{R}^{N \times N}$ encoding pair-wise image similarities. The matrix $\boldsymbol{Z}$ is estimated from the set $\mathcal{X}$ and holding elements $z_{nm} = \zeta(\boldsymbol{X}_n, \boldsymbol{X}_m)$ with $z_{nm} \in \mathbb{R}^+$.

2.1.2 Spectral clustering

Aiming to build a low dimensional space from the data, the set $\boldsymbol{X}$ is seen as a graph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ composed of a set of N vertices $\mathcal{V} = \{v_n \in \mathbb{N} : n = 1, \ldots, N\}$ and a set of edges $\mathcal{E} = \{e_{nm} \in \mathbb{R}^+ : n, m = 1, \ldots, N\}$ linking them, the goal of spectral clustering is to find K disjoint subsets from $\mathcal{V}$. In this regard, $\boldsymbol{Z}$ can be seen as the weighting edge matrix of the undirected, fully connected graph $\mathcal{G}$, where each entry represents the similarity between

each image pair, $e_{nm} = z_{nm} = \zeta(\boldsymbol{X}_n, \boldsymbol{X}_m)$. Since the number of connected vertices in the graph $\mathcal{G}$ corresponds to the eigenvalue multiplicity of the normalized Laplacian matrix of $\boldsymbol{Z}$, such a matrix is defined as [56]:

$$\boldsymbol{F} = \boldsymbol{D}^{-1/2}\boldsymbol{Z}\boldsymbol{D}^{-1/2}$$

where $\boldsymbol{D} \in \mathbb{R}^{N\times N}$ is a diagonal matrix, known as the degree matrix with elements $d_{mm} = \sum_{n=1}^{N} z_{mn}$. Hence, a spectral decomposition of $\boldsymbol{F}$ is required. Specifically, if the K largest eigenvectors of $\boldsymbol{F}$ are stacked as columns on a matrix $\boldsymbol{Y} \in \mathbb{R}^{N\times k}$ and each row is scaled to have unit length, thus mapping the original data points into a unit hypersphere [38], data modes can be inferred by simple clustering techniques while enhancing latent distributions [74]. Aiming to minimize redundant information we use the principal components analysis(PCA) that decomposes the laplacian matrix as: $\boldsymbol{F}=\boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{\top}$, where $\boldsymbol{\Lambda}\in\mathbb{R}^{N\times N}$ is a diagonal matrix containing the ranked in descending-order eigenvalues of $\boldsymbol{F}$ and matrix $\boldsymbol{V}\in\mathbb{R}^{N\times N}$ holds its column eigenvectors. In this sense, to represent $\boldsymbol{X}_n$ we use a vector $\boldsymbol{u}_n\in\mathbb{R}^p$ obtained as: $\boldsymbol{u}_n=\sum_{m\in N} z_{nm}\tilde{\boldsymbol{v}}_m$, where $\tilde{\boldsymbol{v}}_m$ is the m-th column of $\boldsymbol{V}$ truncated to the p most relevant components in terms of its corresponding eigenvalue.

2.1.3 Kernel Based Image Representation

In a 3D MR image, the original feature space corresponds to $\Omega=\mathbb{R}^{W\times H\times L}$. Therefore each image can be arranged as an ordered slice set $\boldsymbol{X} = \{\boldsymbol{S}_k : k \in [1, L]\}$, being $\boldsymbol{S}_k = \{\boldsymbol{x}_r : r = (i, j, k)\}$, with $\boldsymbol{S}_k\in\mathbb{R}^{W\times H}$ the k-th $(W \times H)$-sized matrix slice.

Assuming smooth variations between adjacent slices on $\boldsymbol{X}$, inter-slice relationship is encoded by the kernel function: $\gamma(\boldsymbol{S}_k, \boldsymbol{S}_{k'}) = \langle\varphi(\boldsymbol{S}_k), \varphi(\boldsymbol{S}_{k'})\rangle$. Thus, the output symmetric matrix $\boldsymbol{G}\in\mathbb{R}^{L\times L}$ with real-valued elements, $g_{kk'}=\gamma(\boldsymbol{S}_k, \boldsymbol{S}_{k'})$, becomes the inter-slice kernel (ISK) of $\boldsymbol{X}$. Moreover, provided $g_{kk'} = g_{k'k}$ and $g_{kk} = 1$, each image $\boldsymbol{X}_n$ can be represented by the upper triangular ISK elements as a vector $\boldsymbol{y}_n\in\mathbb{R}^{L(L-1)/2}$. Therefore, the image similarity measure in Eq. (2-1) is represented as: $\zeta(\boldsymbol{X}_n, \boldsymbol{X}_m) = f(\boldsymbol{y}_n, \boldsymbol{y}_m)$, where $f(\cdot, \cdot)$ is the introduced kernel function.

2.1.4 Bayesian medical image segmentation

As proposed in [7, 39], automatic labeling of medical images can be stated within a Bayesian classification framework as estimation of a label set, $\boldsymbol{L}=\{l_r\in[1, C]:r\in\Omega\}$, from a given set of measurements (or *image*), $\boldsymbol{X}=\{\boldsymbol{x}_r\in\mathbb{R}^d\}$, where a single label, l_r, is assigned to the r-th

spatial element (or *spel*), depending on the d-dimensional measurement vector $\boldsymbol{x}_r$, where $C\in\mathbb{N}$ is the total number of considered labels or classes, and Ω is the spel domain.

In the Bayesian framework, provided a given query image, the probability of having a label set, $P(\boldsymbol{L}|\boldsymbol{X})$, the probability of occurring the image given the labeling, $P(\boldsymbol{X}|\boldsymbol{L})$, and the label prior probability, $P(\boldsymbol{L})$, are together related as:

$$P(\boldsymbol{L}|\boldsymbol{X}) \propto P(\boldsymbol{X}|\boldsymbol{L})P(\boldsymbol{L}) \tag{2-2}$$

In most of medical imaging applications, measurement vector distributions of different tissues are overlapped. Consequently, the probabilities $P(\boldsymbol{X}|\boldsymbol{L})$ and $P(\boldsymbol{L})$ (termed *atlas functions*), can vary spatially, so that a set of Bayesian classifiers is applied along Ω, where each one deals with independent small regions. Hence, the probability of obtaining a label set given a query image is represented as $P(\boldsymbol{L}|\boldsymbol{X})=\prod_{r\in\Omega} P(l_r=c|\boldsymbol{x}_r)$, where $P(l_r=c|\boldsymbol{x}_r)=P(\boldsymbol{x}_r|l_r=c)P(l_r)/P(\boldsymbol{x}_r)$ is the probability of obtaining the label $c\in[1,C]$ at spel r and given the measurement $\boldsymbol{x}_r$, $P(\boldsymbol{x}_r|l_r=c)$ is the conditional probability of observing the measurement $\boldsymbol{x}_r$ at r given the label $l_r=c$, $P(l_r)$ is the label prior probability, and $P(\boldsymbol{x}_r)$ is the evidence probability.

2.1.5 Learning model parameters

Since the N-image set is mostly heterogeneous, some objects can bias the segmentation obtained over a given query image. Hence, we propose to enhance the above explained approach by computing the atlas functions for a given subject from a subpopulation of the entire dataset, which in turn depends on a new introduced pair-wise image similarity. Thus, the segmentation task relies on the estimation of the atlas functions from a given sub-set of M pre-labeled atlas images, $\mathcal{X}^*=\{\boldsymbol{X}_m,\boldsymbol{L}_m:1,\cdots,M\}$, whit $\mathcal{X}^*\subset\mathcal{X}$ where the r-th spel of the m-th image is assigned to the measurement vector $\boldsymbol{x}_r^m$ and the class l_r^m. Hence, provided $\mathcal{X}$, the prior $P(l_r=c)$ and evidence $P(\boldsymbol{x}_r)$, both the atlas functions can be computed as:

$$P(l_r=c)=\boldsymbol{\mathcal{E}}\{\delta(l_r^m-c):\forall m\in M\} \tag{2-3a}$$

$$P(\boldsymbol{x}_r)=\sum\nolimits_{c\in C} P(\boldsymbol{x}_r|l_r=c)P(l_r=c) \tag{2-3b}$$

where $\delta(\cdot)$ is the delta function and notation $\boldsymbol{\mathcal{E}}\{\cdot\}$ stands for the expectation operator.

In each class, $\boldsymbol{x}_r$ is assumed to be normally distributed [39], $\boldsymbol{x}\sim\mathcal{N}(\boldsymbol{\mu}_r^c,\boldsymbol{\Sigma}_r^c)$, where mean $\boldsymbol{\mu}_r^c\in\mathbb{R}^d$ and covariance $\boldsymbol{\Sigma}_r^c\in\mathbb{R}^{d\times d}$ class parameters are estimated as:

$$\boldsymbol{\mu}_r^c=\boldsymbol{\mathcal{E}}\{\boldsymbol{x}_r^m|l_r^m=c:\forall m\in M\} \tag{2-4a}$$

$$\boldsymbol{\Sigma}_r^c=\boldsymbol{\mathcal{E}}\{(\boldsymbol{x}_r^m-\boldsymbol{\mu}_r^c)(\boldsymbol{x}_r^m-\boldsymbol{\mu}_r^c)^\top|l_r^m=c:\forall m\in M\}, \tag{2-4b}$$

2.2 Experimental setup

The proposal is used in two specific tasks, the first one correspond to clustering tasks aiming to find demographic groups on the population. On the other hand, similarities computed in the embedding space are use to select the most similar atlases in multi-atlas segmentation scheme. In this sense the proposal is outline on three main stages necessary for both tasks *i*) the image preprocessing stage is aimed to reference all images in the datasets to the same intensity space and spatial framework, *ii*) the image representation stage is performed to code the high-dimensional information of each image on a lower feature space, and *iii*) the image embedding stage is employed to build up a new low dimensional space provided with a better interpretability and where inferences can be carried out easily.

2.2.1 Preprocessing

Two preprocessing steps are performed for all images on the dataset. Firstly, each image is registered to the MNI305 template by an affine transform so that the whole dataset is referenced to Talairach space. [37], Due to the registering each volume is resampled to $197 \times 233 \times 189$ size. Lastly, an intensity normalization procedure is performed by scaling each voxel intensity, so that the mean intensity of the white matter is fixed to be 110. Both preprocessing steps, normalization and registrering, are performed with the Freesurfer image analysis suite, which is documented and freely available for download online [1].

2.2.2 Proposed Image Feature Extraction

We note the ISK of a given image as $\boldsymbol{y}_n \in \mathbb{R}^{L(L-1)/2}$. Hence, a new representation space of order 10^4 is achieved, instead of the original image domain of order 10^6. Specifically, to compute $\gamma_{kk'}$ of each image, we use the Gaussian kernel

$$\gamma\left(\boldsymbol{S}_k, \boldsymbol{S}_{k'}\right) = \exp\left(\frac{-||\mathbf{S}_k - \mathbf{S}_{k'}||_F^2}{(2\sigma_\kappa^2)}\right) \tag{2-5}$$

where $\sigma_\kappa \in \mathbb{R}^+$ is a scale parameter and notation $\|\cdot\|_F$ stands for the Frobenius norm. Hence, two important issues have to be highlighted. Firstly ISK can be Computed Along Three different axes(namely Axial, Sagittal and Coronal): Therefore, all axes are considered for subsequent analysis. Secondly, σ_κ parameter has to be tuned. In the concrete case taken into account that $\lim_{\sigma_k \to 0} \mathrm{Var}(\boldsymbol{G}(\sigma_k)) = 0$, $\lim_{\sigma_k \to \infty} \mathrm{Var}(\boldsymbol{G}(\sigma_k)) = 0$, and an appropriate σ_κ value

[1] `http://surfer.nmr.mgh.harvard.edu/`

spans widely the values of $\boldsymbol{G}$. The optimization of the scale parameter is performed by maximizing the following element-wise matrix variance:

$$\sigma_{\kappa}^{*} = \arg\max_{\sigma} \left\{ \mathrm{Var}\{\boldsymbol{\gamma}_{kk'}(\sigma) : \forall k \neq k'\} \right\}, \tag{2-6}$$

2.2.3 Kernel-based Image Similarity

Starting from the aforementioned representation approach, an embedding low dimensional space is build by using the aforementioned explained spectral decomposition. Visualization clustering and regression are performed over a new more compact space. For this specific task, the common Gaussian kernel is employed:

$$\zeta(\boldsymbol{X}_n, \boldsymbol{X}_m) = \exp\left(\frac{-d(\boldsymbol{X}_n, \boldsymbol{X}_m)^2}{2\sigma_{\zeta}^2}\right) \tag{2-7}$$

Where $\sigma_{\zeta} \in \mathbb{R}^{+}$ is the scale parameter and $d(\boldsymbol{X}_n, \boldsymbol{X}_m) \in \mathbb{R}^{+}$ is the distance function between the n-th and m-th images.

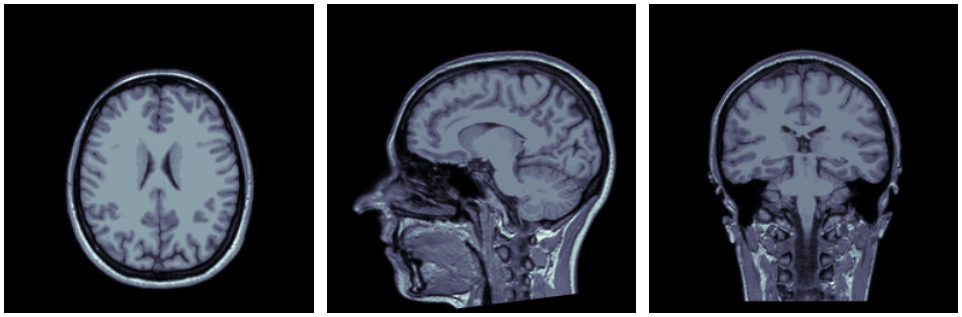
2.3 Kernel Based Representation for clustering task

2.3.1 Database

The `IXI` dataset is a brain imaging study, holding MR images from 575 normal subjects which age between 20 and 80 years. Subjects are provided with T1, T2, PD, DTI and angiogram volumes. The image sequences were acquired with three different scanners (Philips 1.5T, Philips 3T and a GE 3T), anonymised and converted to NIFTI format. Additionally, basic demographic information for each subject is included (age, gender, ethnicity, among others). The whole dataset is freely available online [2].

Since the target of the current paper is related to atlas construction, only the T1 sequences of $N = 322$ subjects (acquired with the GE 3T scanner) were taken into account. T1 sequences are composed of $256 \times 256 \times 150$-sized volumes with a voxel size of $0.9375 \times 0.9375 \times 1.2$mm. Thus, the considered subset is composed of 141 male, 175 female and 6 unknown subjects. Figure 2-1 shows an example of the MR image for a given subject along three different views.

[2] `http://www.brain-development.org/`

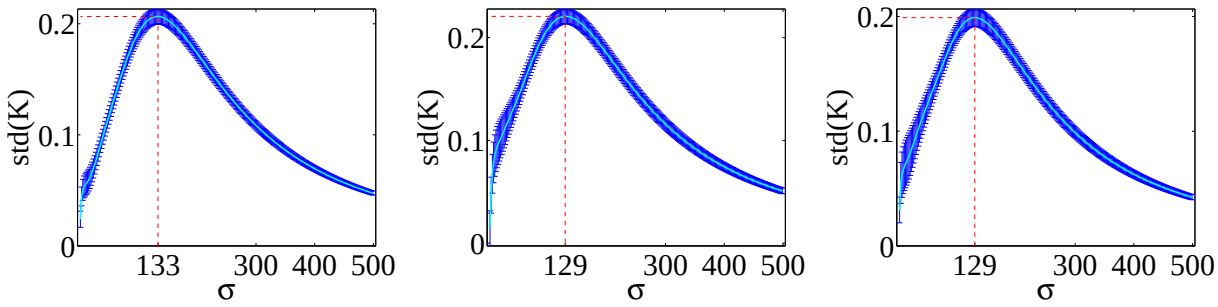


(a) Axial (b) Sagittal (c) Coronal

Figure **2-1**: Volume sample from the IXI database. Subject 002

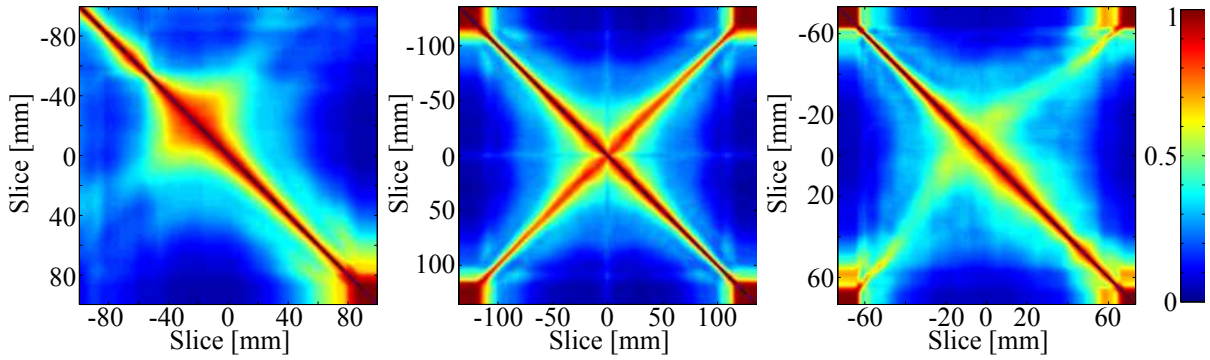
2.3.2 Parameters Tuning and Results

Proposed approach is compared against the conventional MR image similarity voxe-wise sum of square differences on the task of clustering the demographic categories age and gender. Obtained results are shown in terms of cluster distributions and centroid distances. In this sense two image representation techniques are employed for comparison. The first one is a baseline where each voxel on the image is used as a feature. While the second one corresponds to the proposed approach where each image is represented by an ISK matrix, noted as $\boldsymbol{G}$. Given that, ISK analysis can be achieve on three different axes (namely Axial, Sagittal and Coronal). Therefore, all axes are considered for analyses. Figure **2-2** shows Obtained tuning curves for the scale parameter σ_κ in the three considered axes. Mean and standard deviation are computed for a randomly selected subset of 30 MR images. Figure **2-3** shows an example of the proposed image representation for a given MR image along the axial, sagittal and coronal axes.



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

Figure **2-2**: Sigma tunning for Inter Slice Kernels along the three possible axes. Mean and standard deviation are plotted.



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

Figure **2-3**: Inter Slice Kernel similarity for a given image. All axes are measured in mm.

From both aforementioned representation approaches, and embedding low dimensional spaces is build by using the Gaussian kernel as in the equation 2-7 which encode pairwise image correspondences. In this sense, a distance function in the equation is used for each representation. For the baseline image representation approach, that distance correspond to the voxel-wise image euclidean norm, defined as:

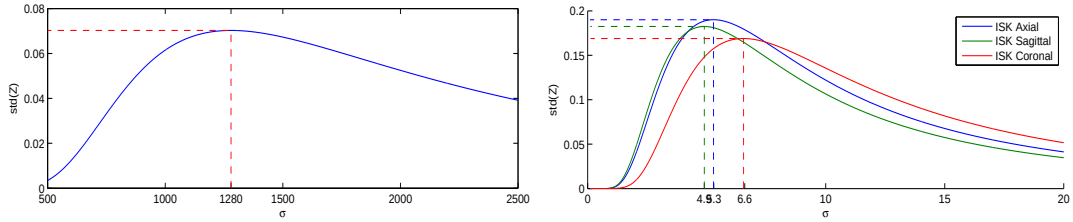
$$d_{VW}(\boldsymbol{X}_n, \boldsymbol{X}_m) = ||vec(\boldsymbol{X}_n) - vec(\boldsymbol{X}_m)||_2$$

while for the proposed representation, the distance is computed as the ISK matrix Frobenius norm as follows:

$$\begin{aligned} d_{ISK}(\boldsymbol{X}_n, \boldsymbol{X}_m) &= ||G_n - G_m||_F \\ &= ||\boldsymbol{y}_n - \boldsymbol{y}_m||_F \end{aligned}$$

Figure **2-4** shows obtained tuning curves for all considered image representation using the criterion proposed on Equation (2-6).

Resulting kernels for considered representations are shown in Figure **2-5**, where subjects are sorted by gender and age values. Although all dataset information is encoded on matrices and some small subsets can be identified, it is still hard to group subjects on categories as gender and age. Therefore, a PCA-based projection space is computed from the laplacian of above matrices. The four largest decomposition eigenvectors are shown in Figures **2-6** and **2-7**. Obtained projection allows to come up the following statements: i) The first and second decomposition eigenvectors build a space able where the age is "unfolded/unwraped". ii) The fourth decomposition eigenvector decodes the gender category more accurately than remaining components.



(a) Voxel-Wise (b) Inter Slice Kernel

Figure **2-4**: σ_ζ tuning curve for considered image representations: (a) Voxel-Wise approach and (b) Proposed ISK along the three possible axes.

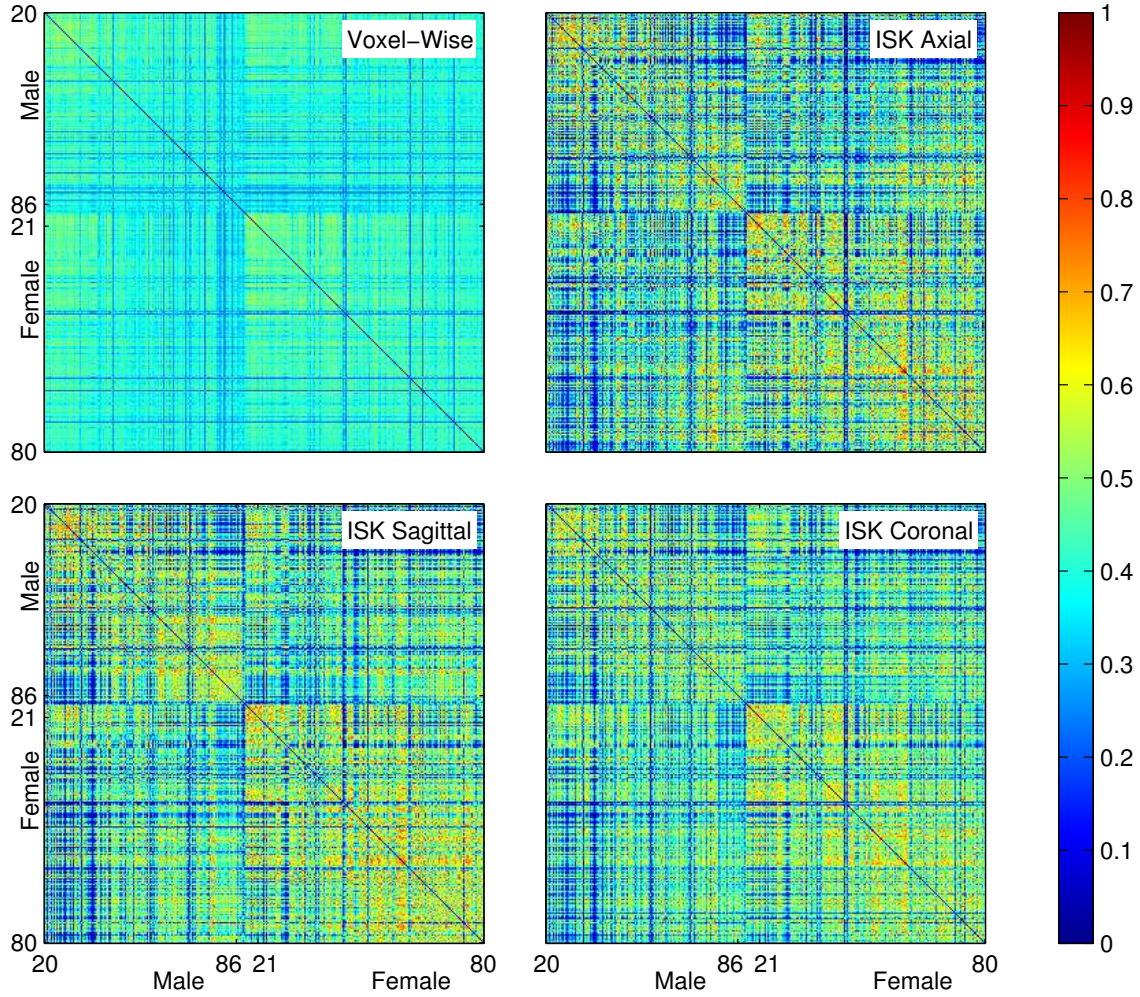


Figure **2-5**: IXI dataset kernel matrices for both considered image representations. Each row and column on the matrices corresponds to a given image. Images are ordered by gender (firstly) and age (secondly). The color range is normalized between the interval $[0, 1]$

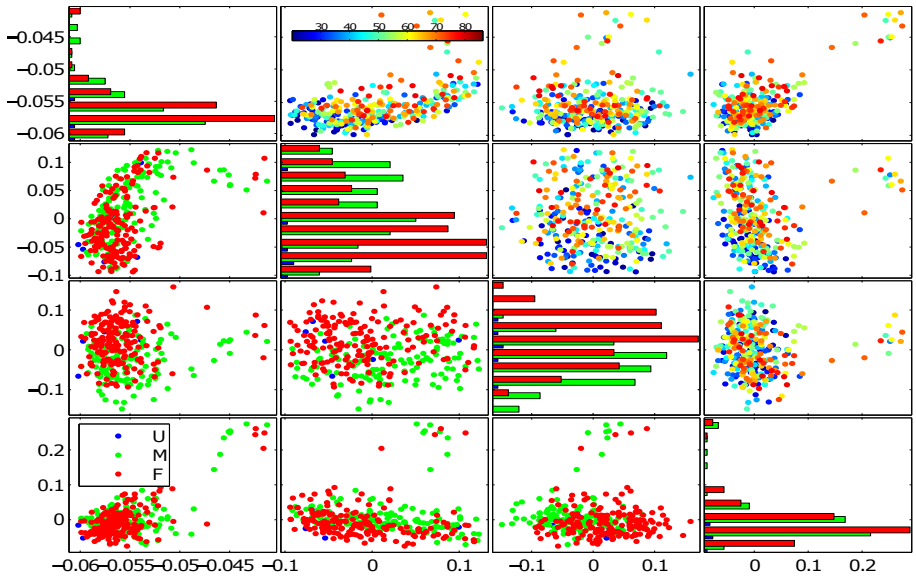


Figure **2-6**: Scatter matrix for the four largest decomposition eigenvectors using the voxel-wise approach. Upper scatters are grouped by age. Lower scatters are grouped by gender.

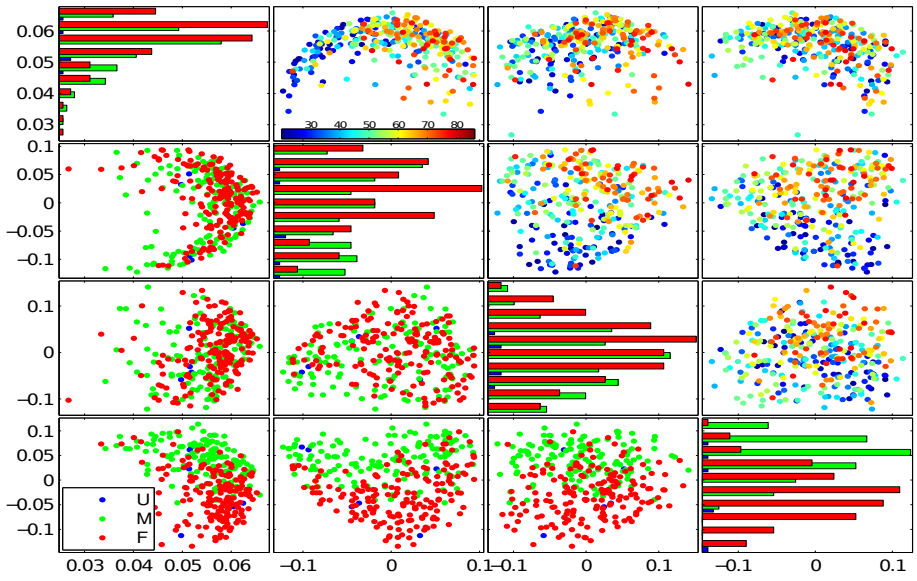
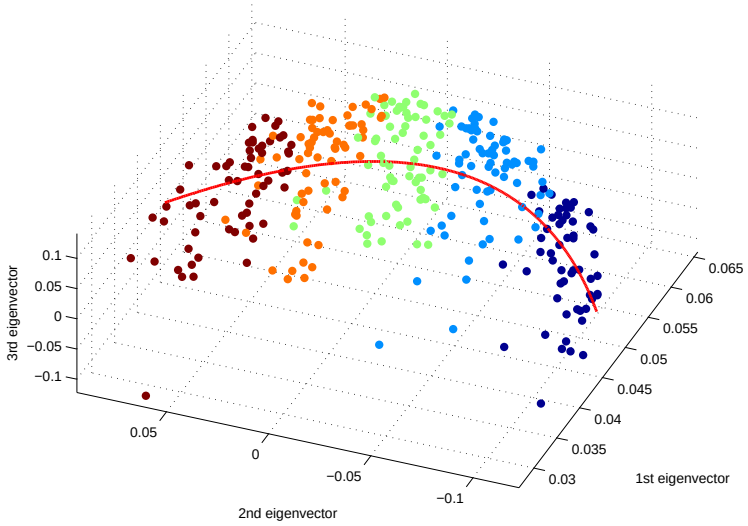


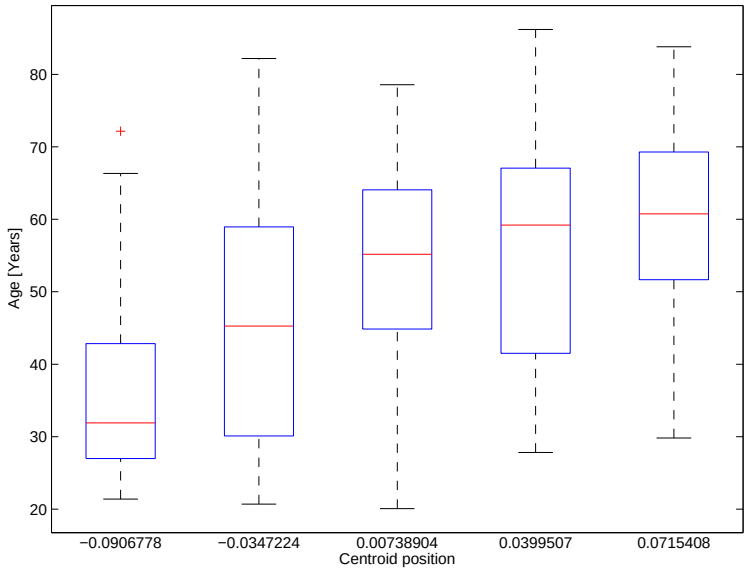
Figure **2-7**: Scatter matrix for the four largest decomposition eigenvectors using the ISK along the axial axis. Upper scatters are grouped by age. Lower scatters are grouped by gender.

Aiming to prove the first of above statement, two subsequent analyses are performed on the obtained representation space. Firstly, a quadratic regression is computed, so that the trend along the second component is emphasized. Secondly, since the new representation

enhances each cluster properties, making the natural groups easily detectable, a simple k-means clustering algorithm is used to find, with less difficulties, the natural groups on the new space. As expected, resulting regression (colored in red line) and clustering results (Figure 2.8(a)) show a trend along the second axis. Besides, the cluster age distribution is provided in Figure 2.8(b) showing age clustering, proving that the subject age tends to increase along the computed trend curve.



(a) Quadratic regression and clustering



(b) Age distribution for each cluster

Figure **2-8**: Trend estimation along the second decomposition axis results using the ISK axial image representation.

Lastly, for proving the second statement, a cluster measure is employed to quantify the

separability of male and female clusters. The measure is computed from the inter-cluster over the intra-cluster variance, so that the larger the value, the farther the cluster distributions. The test is carried out along each considered decomposition eigenvectors and over the four dimensional representation space. Obtained separability measures (see Table **2-1**) show a larger separation through the fourth component for most of the ISK representations.

Approach	**1st**	**2nd**	**3rd**	**4th**	**All**
Baseline	2.33	5.23	8.33	2.62	12.22
ISK Axial	2.74	2.25	2.06	14.66	15.98
ISK Sagittal	3.03	4.61	8.48	11.20	18.74
ISK Coronal	2.57	3.05	12.79	2.72	14.50

Table **2-1**: Cluster separability measure for the first four components of considered representation approaches.

2.4 Kernel-based Atlas Image Selection from the Embedding Representation (KAISER)

The proposed methodology is tested on a manually labeled database to support brain tissue segmentation. To this end, A kernel-based atlas Image selection computed in the embedding representation space (termed KAISER) is used to construct target-specific probabilistic functions which guide the brain tissue segmentation. The proposal is compared against the probabilistic segmentation using the whole-population atlas and demographic population atlas.

2.4.1 Database

The `OASIS` dataset is a brain imaging study, holding an MR image collection from 416 subjects, aged 18 to 96 years old, including diagnosed very mild dementia (70), mild dementia (28), moderate dementia (2), and healthy (316) subjects. For each of them, three or four T1-weighted MR scans obtained within a single imaging session are included (see a sample subject in Fig. **2-9**), from which a motion-corrected co-registered average image is obtained. Additionally, each subject is provided with ground-truth segmented gray matter (GM), white matter (WM) and cerebro-spinal fluid (CSF) structures. A fourth label (BG) is included in the present study, aiming to model the background of images, as the regions with no provided label.

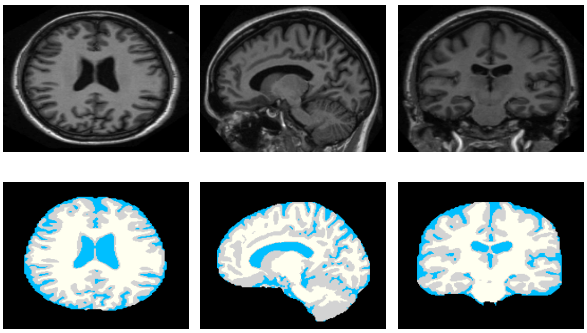


Figure **2-9**: Sample subject of the OASIS dataset publicly available in `http://www.oasis-brains.org`. Top row: Motion corrected MR image. Bottom row: Provided structure segmentation. Left to right: Axial, Sagittal and Coronal axes.

2.4.2 Parameters Tuning and Results

Aiming to perform atlas selection, The proposed feature extraction and embedding space are computed. According to the obtained results on clustering task, ISK representation is computed along the sagittal axis given that provide more separability on the clusters. Figure **2-10** shows obtained tuning curves for the scale parameter σ_κ

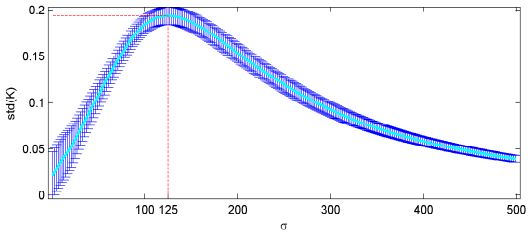
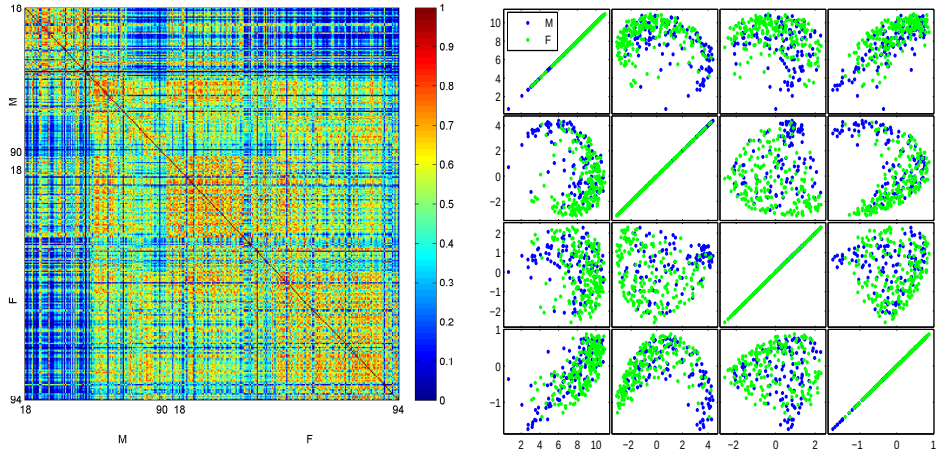


Figure **2-10**: Scale parameter vs. kernel element-wise variance for the sagittal ISK. Mean and standard deviation values are plotted.

Starting from the sagittal ISK representation, the Gaussian kernel in the equation 2-7 is used to encode image similarities in the low dimensional space. A scale of $\sigma_\zeta = 9$ is obtained using the above proposed tuning procedure. The resulting kernel is depicted in Fig. 2.11(a), where each row and column represents a given image on the dataset. Images are ordered by gender (firstly) and age (secondly).

Therefore, the PCA-based projection space is computed from te laplacian of the kernel $\boldsymbol{Z}$. As seen in Fig. 2.11(b) showing the four largest decomposition eigenvectors, there exists an inherent structure, hidden on the image distribution, that is hard to identify in the original space domain (Ω), but easily identifiable in the proposed projected space. As a

result, patient-dependent atlases lead to more accurate segmentation results than the whole-population atlas.



(a) Kernel matrix using the ISK features (b) Data projection along the first four eigenvectors

Figure **2-11**: Kernel matrix (top) and eigen-projection (bottom) representations for the OASIS dataset.

Finally, the well-known L-curve criterion is used to tune the value p to truncate the number of components employed on the projected representation $\boldsymbol{v} \in \mathbb{R}^p$. For such criterion, the minimum distance to origin from a curve, composed by the normalized eigenvalues and the percentage of components, has to be found. For the OASIS dataset, such distance is found at the ninth component. Hence, subsequent analyses are performed with $p = 9$.

2.4.3 Tissue Labeling Performance

Taking into account the image similarity measure described in Section 2.1.1, we propose four strategies for choosing the image subset used for computation of the atlas functions: *i)* using the whole dataset on the Bayesian Classification Framework (FULL) [7], as a state-of-the-art comparative baseline; *ii)* extracting M_k samples using the proposed **K**ernel-based **A**tlas **I**mage **S**election from the **E**mbedding **R**epresentation, termed KAISER; *iii)* randomly choosing a subset of M_r images (RAND), as the comparative approach used in [4], with $M_r = 20$ and 10 folds. *iv)* Selecting demographically affine images to the query image (DEMG), using gender and age categories, as in [72].

For measuring segmentation quality of above described approaches, we assess agreement between resulting segmentations and the labels provided for each dataset image. Namely, we employ the Dice index: $d = 2|\mathcal{A} \cap \mathcal{B}|/(|\mathcal{A}| + |\mathcal{B}|)$, where $\mathcal{A}$ and $\mathcal{B}$ are respectively the

provided and estimated regions of the compared tissue, and $|\cdot|$ stands for the number of spels on a given region. Additionally, to prove the repeatability of the approaches, a leave-one-out (LOO) cross-validation scheme is employed. Specifically for KAISER, the number of neighbors M_k has to be tuned; this procedure is done by intensive searching of the largest average-accuracy-over-standard-deviation ratio in the LOO scheme. Search is performed over the interval $M_k = [5, 21]$, as shown in Fig. **2-12**, where the optimum value is found at $M_k = 13$ neighbors.

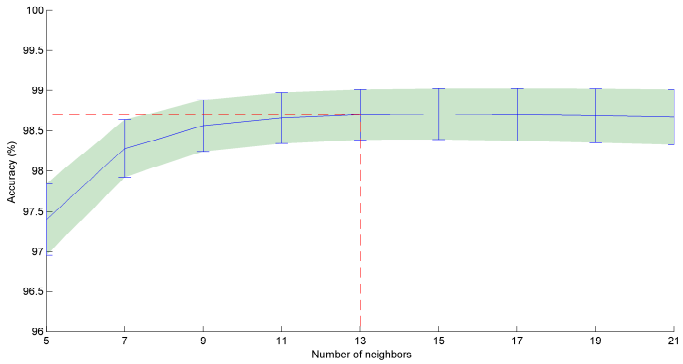


Figure **2-12**: Overall accuracy segmentation versus the number of selected closest neighbors for the proposed KAISER. Errorbar is displayed for a LOO validation scheme.

Once tuned the number of neighbors, the Dice index and overall accuracy are computed for the compared approaches in the LOO validation scheme. Mean and standard deviation results, provided in Table **2-2**, show that proposed KAISER outperforms FULL on all items.

Table **2-2**: Segmentation results for considered selection approaches

	BG	*WM*	*GM*	*CSF*	*Acc*
FULL [7]	100	81.5 ± 8.8	85.8 ± 5.4	84.5 ± 7.6	96.1 ± 1.4
KAISER	100	92.1 ± 1.9	95.1 ± 1.4	96.2 ± 1.0	98.7 ± 0.3
RAND [4]	100	74.6 ± 6.7	81.7 ± 5.1	79.2 ± 7.0	94.8 ± 1.3
DEMG [72]	100	77.2 ± 5.9	83.6 ± 5.7	82.2 ± 6.7	95.4 ± 1.4

2.5 Discussion

The proposed kernel-based image representation is specifically devoted to perform 3D MR image unsupervised clustering. Besides, a low-dimensional, compact space is built by means of spectral clustering, so that the demographic categories are easier distinguished. Moreover, due to the intrinsic similarities are better reflected in this low-dimensional space, the proposed methodology is suitable for atlas selection enabling to construct target specific probabilistic atlases enhancing brain tissue segmentation.

As seen in Figure **2-2**, the estimated parameters for the Gaussian kernel are close to each other. Therefore, since the inter slice difference is taken from an image for the three axes, the difference dynamic range and the parameter search space are the same for all axes. Moreover, given that the confidence interval, produced by the 30-trials-computed standard deviation, is narrow, the proposed parameter tuning proved to be a stable criterion. In this sense, obtained ISK using estimated parameters maximally enhances the inter-slice relationship, as can be seen in Figure **2-3**.

In Figure **2-4**, obtained parameters for ISK tuning (see Figure 2.4(b)) lie around the same value. Such a fact may be explained by a couple reasons: i) the feature values range from 0 to 1 for all axes, so the search space is the same. ii) Although the kernel shape is different for the three axes (Figure **2-3**), the latent phenomenon is the same for all axes. Therefore, if above considerations are meet, tuned parameters for different views tend to converge to the same value.

Regarding the age as a demographic category, by visual inspection of the first and second components depicted in Figure **2-7**, it can be seen that the proposed methodology is able to unfold the age better than any other component pair, even on the baseline decomposition results (see Figure **2-6**). Moreover, a quadratic dependance between second and first eigenvectors can be inferred. Additionally, a larger dispersion is shown on older subjects than on younger ones. This finding can be due to a larger head shape dispersion on older humans, which is according to anatomical head knowledge. It is known that brain anatomy is steady on middle age humans, while change (gray matter volume diminish) faster on older humans.

Aiming to prove above statements, a quadratic regression is performed (see Figure **2-8**), which proves to fit adequately the relation between the first two ISK-Axial eigenvectors. Moreover, since the average age on each cluster tends to increase as the centroid position increases, it can be said that subject age is directly described by the relationship between the first two eigenvectors.

Regarding the gender, on a component-wise analysis, the fourth ISK-Axial and third voxel-wise eigenvectors seem more suitable to distinguish gender than remaining components in Figures **2-6** and **2-7**. Therefore, a two-sample hypothesis test is employed to quantify the separability between gender clusters for the first four eigenvectors of all considered representations. A component-wise and a multivariate test are performed. Results on Table **2-1** show that the largest separability is found when using the first four eigenvectors and proved that our proposal is more suitable to distinguish gender than the voxel-wise baseline.

Taking into account the aforementioned results, the proposed kernel-based representation methodology is proved to find the natural inherent distributions of MR images, namely, age and gender categories. In this sense, our proposal is suitable to support MR image clustering

and similarity measurement tasks required on template-based image segmentation.

Overall, obtained segmentation results (see Table **2-2**) show that the KAISER proposal achieved the largest average and the least deviation values for all Dice indexes and accuracy. Hence, our proposal exhibits the best classification performance and the most repeatability. The above leads to an improvement, regarding the atlas selection, on multi-atlas-based image labeling approaches.

3 Supervised Similarity Measure supporting Atlas-Selection

In the previous chapter, we introduce a similarity measure computed in a low-dimensional space which demonstrated be accurate for highlight the data distribution on the dataset, enabling to perform appropriate atlas selection. In the aforementioned method, only the intensity information was employed to inferred image relationships. Now in this chapter, we incorporate label information aiming to learn the appropriate selection based on the expected segmentation. Current methodologies use similarity measures such mutual information(MI), cross-correlation, or squared differences. More sophisticated methods use a distance on a manifold. However, these methods are limited mainly for two reasons: Firstly, as the measures are evaluated over the entire image, it is biased towards large regions as the background instead of the relevant structures. Secondly, these rely on pairwise similarities for both measure appearance or learn the manifold image, however, image similarities are not necessarily related with the performance on segmentation of target image. In an attempt to cope with the first restriction. Structure-wise atlas selection is suggested in [85] to segment the brain MRIs based on the highest local mutual information. Also, an adaptive method for a local combination is proposed so that a subset of templates and their weighting are estimated independently at image localities [62]. The main drawback of these approaches lies in the requirement of a deformable registration stage measuring the image similarity for all atlases. That procedure is computationally much more expensive than linearly mapping all the images into a common reference space. On the other hand, a recent approach [68] propose to learn relations between pairwise appearance and the labeling performance. Bearing in mind all described above constraints, we propose a new spatially weighting procedure of the well-known image metrics for supporting the atlas selection within a multi-atlas-based segmentation scheme. Specifically, we study the mean squares (MS), histogram correlation coefficient (HCC), normalized mutual information (NMI), and neighborhood cross-correlation (NCC). Our approach computes independently the measurements at regular image partitions; then all partition similarity values are linearly combined to get a single similarity outcome. The combination parameters are properly tuned to match the optimal similarity based on the label affinity between the prelabeled images on the training dataset.

3.1 Materials and Methods

Let $\mathcal{X}=\{\boldsymbol{X}^n, \boldsymbol{L}^n : n=1,\ldots,N\}$ be a labeled MRI dataset holding N pairs of segmented images, where $\boldsymbol{X}^n=\{x_r^n \in \mathbb{R} : r\in\Omega\}$ is the n-th MR image, the value r indexes the spatial elements, and the matrix $\boldsymbol{L}^n=\{l_r^n \in [1,C] : r\in\Omega\}$ is the provided image segmentation into $C \in \mathbb{N}$ classes, which for 3D volumes holds dimension $\Omega=\mathbb{R}^{T_a\times T_s\times T_c}$, with $\{T_a, T_s, T_c\}$ as the Axial, Sagittal, and Coronal real-valued sizes, respectively.

3.1.1 Image similarity metrics

The similarity between a given image pair, $\{\boldsymbol{X}^n, \boldsymbol{X}^m\}$, can be assessed by using one of the following widely employed metrics:

Mean Squares (MS): This metric that is based on the average square difference along the space is embedded into a Gaussian kernel function, yielding the following bounded similarity measure:

$$s\{\boldsymbol{X}^n, \boldsymbol{X}^m\} = \exp\left\{-\frac{1}{2\sigma^2}\boldsymbol{\mathcal{E}}\{(x_r^n - x_r^m)^2 : \forall r \in \Omega\}\right\} \in [0,1]; \tag{3-1}$$

where $\sigma\in\mathbb{R}^+$ is the kernel bandwidth. Notation $\boldsymbol{\mathcal{E}}\{\cdot\}$ stands for the expectation operator.

Histogram Correlation Coefficient (HCC): This metric calculates similarity between image histograms as follows:

$$s\{\boldsymbol{X}^n, \boldsymbol{X}^m\} = \frac{\boldsymbol{\mathcal{E}}\{h(x_r^n, x_s^m)(x_r^n x_s^m - \bar{x}^n\bar{x}^m) : \forall r,s \in \Omega\}}{\boldsymbol{\mathcal{E}}\{h(x_r^n)(x_r^n - \bar{x}^n)^2\}\boldsymbol{\mathcal{E}}\{h(x_r^m)(x_r^m - \bar{x}^m)^2\}} \in [0,1] \tag{3-2}$$

where $h(x^n, x^m)\in\mathbb{R}^+$ is the joint histogram between both input images, and $\bar{x}^\upsilon\in\mathbb{R}$, with $\upsilon\in\{m,n\}$, is the average intensity of the respective input image $\boldsymbol{X}^\upsilon$.

Normalized Mutual Information (NMI): This similarity value measures the normalized mutual information of a couple of images as:

$$s\{\boldsymbol{X}^n, \boldsymbol{X}^m\} = \frac{\mathbb{H}\{\boldsymbol{X}^n\} + \mathbb{H}\{\boldsymbol{X}^m\}}{\mathbb{H}\{\boldsymbol{X}^n, \boldsymbol{X}^m\}} - 1 \in [0,1] \tag{3-3}$$

where notation $\mathbb{H}\{\boldsymbol{X}^n, \boldsymbol{X}^m\}$ stands for the joint entropy between $\boldsymbol{X}^n$ and $\boldsymbol{X}^m$.

Neighborhood Cross-Correlation (NCC): This metric is widely used within the Advanced Normalization Tools (ANTs) framework, and computes the normalized cross correlation of voxel neighborhoods between two images:

$$s\{\boldsymbol{X}^n, \boldsymbol{X}^n\} = \frac{\boldsymbol{\mathcal{E}}\{(\boldsymbol{x}_s^n - \bar{\boldsymbol{x}}_s^n)^2 (\boldsymbol{x}_s^m - \bar{\boldsymbol{x}}_s^m)^2 : \forall s \in \Omega\}}{\boldsymbol{\mathcal{E}}\{(\boldsymbol{x}_s^n - \bar{\boldsymbol{x}}_s^n)^2\}\boldsymbol{\mathcal{E}}\{(\boldsymbol{x}_s^m - \bar{\boldsymbol{x}}_s^m)^2\}} \in [0, 1] \tag{3-4}$$

where $\boldsymbol{x}_s^n \in \mathbb{R}^{q\times q\times q}$ is the set of intensity levels in a q-sized neighborhood around the s-th voxel of the image $\boldsymbol{X}^n$, and $\bar{\boldsymbol{x}}_s^n = \boldsymbol{\mathcal{E}}\{\boldsymbol{x}_s^n : \forall s \in \Omega\}$.

It is worth noting that MS and NMI measures are re-written from their original definition so that all of the above similarity metrics share the same interpretation. Namely, $s=0$ implies the complete mismatch between images, while $s=1$ – an absolute match achieved only if $\boldsymbol{X}_n = \boldsymbol{X}_m$.

3.1.2 Spatial Enhancement of Image Metrics

Since all studied metrics are computed over the whole image, they do not account for local content similarities. Therefore, these measures are biased towards the large similar regions, as the background, masking the relationship between common image structures. Besides, those similarities lack in robustness against artifacts. For instance, the intensity inhomogeneity (being a low-frequency artifact) changes the image intensity distribution along the space. The most common approach to overcome this issue is to compute the metrics at all local regions, which should be further combined adequately into a single metric value. To this end, the image $\boldsymbol{X}$ is split into P different regular blocks, $\{\Omega_p : p \in P\}$. Hence, each image is seen as a set of non-overlapped blocks $\boldsymbol{X} = \{\boldsymbol{\Xi}_p \in \mathbb{R}^{\rho_a \times \rho_s \times \rho_c}\}$, with $P = \prod_v P_v$, $\rho_v = T_v / P_v$, and P_v the number of partitions along the axis v ($v \in \{a, s, c\}$). Consequently, the following P-dimensional vector of metrics holding each the block-wise similarity value is obtained:

$$\boldsymbol{s}\{\boldsymbol{X}^n, \boldsymbol{X}^m\} = \{s_p^{n,m} = s\{\boldsymbol{\Xi}_p^n, \boldsymbol{\Xi}_p^m\}; \forall p \in [1, P]\}$$

With the aim of building a new bounded scalar similarity metric, ζ, we make use of a linear combination of the elemental block-wise measures, namely,

$$\zeta^{n,m} = \boldsymbol{w}^\top \boldsymbol{s}\{\boldsymbol{X}^n, \boldsymbol{X}^m\},\ \zeta^{n,m} \in \mathbb{R}[0, 1]$$

where $\boldsymbol{w} = \{w_p\} \in \mathbb{R}^P$ is the combination vector that is subject to $\sum_{p\in P} w_p = 1$. Since each weight has to account mostly for the influence of the corresponding region on the resulting metric $\zeta^{n,m}$, we assume the vector $\boldsymbol{w}$ to be dependent on the partition size. In the case of the equally-sized blocks, the combination vector is computed as follows [**?**]:

$$w_p = \frac{\rho_a \rho_s \rho_c}{T_a T_s T_c} = \frac{1}{P}, \tag{3-5}$$

As a result, the combination vector estimated in 3-5 leads to the plain averaged block distance, that is, $\zeta^{n,m}{=}\boldsymbol{\mathcal{E}}\{s\{\boldsymbol{\Xi}_p^n, \boldsymbol{\Xi}_p^m{:}\forall p{\in}[1,P]\}\}$. In practice, each block holds a different amount of information depending on its content or its relevance to the task at hand. Regarding this, the contribution of each block can be achieved as its average intensity variance:

$$w_p = \frac{1}{\omega}\boldsymbol{\mathcal{E}}\{\mathrm{var}\left\{\boldsymbol{\Xi}_p^n\right\}; \forall n \in [1,N]\} \tag{3-6}$$

where $\omega{\in}\mathbb{R}^+$ is the normalization factor.

3.1.3 Supervised Image Measure Learning

Basically, we are looking for a measure supporting the Atlas selection within the multi-atlas-based segmentation task. Therefore, we propose to use the provided set of segmented images of the dataset $\mathcal{X}$ to learn the corresponding combination weights for improving the segmentation accuracy. For this purpose, the following similarity matrix $\boldsymbol{Z_w}{=}\{z_{\boldsymbol{w}}^{n,m}{\in}\mathbb{R}^+{:}$ $m,n{=}1\ldots\boldsymbol{N}\}{\in}\mathbb{R}^{N\times N}$ holding all pair-wise metric values is built as a function of the estimated combination weights:

$$\boldsymbol{Z} = \sum_{p=1}^{P} w_p \boldsymbol{S}_p \tag{3-7}$$

$$z_{\boldsymbol{w}}^{n,m} = \zeta\{\boldsymbol{X}^n, \boldsymbol{X}^m\} = \sum_{p=1}^{P} w_p s_p^{n,m}; \forall n,m \in [1,N], \tag{3-8}$$

where $\boldsymbol{S}_p{=}\{s_p^{n,m}{:}m,n{=}1\ldots\boldsymbol{N}\}$ is the similarity matrix attained at the p-th block. Since all considered metrics are equivalent to bounded similarity measures, each of the $\boldsymbol{S}_p$ matrices becomes a positive definite symmetric (PDS) kernel matrix, as well as their linear combination $\boldsymbol{Z_w}$.

Here, we make use of the PDS kernel matrix $\boldsymbol{K}{=}\{k^{n,m}{:}m,n{=}1,\ldots,N\}{\in}\mathbb{R}^{N\mathcal{X}_N}$ to learn the similarity metric ζ through the centered kernel alignment (CKA) measure. Namely, we search for optimal weight vector w_p^* maximizing the correlation between $\boldsymbol{Z_w}$ and the objective kernel matrix $\boldsymbol{K}$ as follows:

$$\boldsymbol{w}^* = \max_{\boldsymbol{w}} \frac{\langle \boldsymbol{Z}'_{\boldsymbol{w}}, \boldsymbol{K}'\rangle_F}{\|\boldsymbol{Z}'_{\boldsymbol{w}}\|_F \|\boldsymbol{K}'\|_F}, \tag{3-9}$$

where notations $\langle\cdot,\cdot\rangle_F$ and $|\cdot|$ stand for the inner product and the Frobenius norm, respectively. $\boldsymbol{Z}'_{\boldsymbol{w}}$ and $\boldsymbol{K}'$ are the centered kernel versions of $\boldsymbol{Z_w}$ and $\boldsymbol{K}$. Here, we compute the

centered version of $\boldsymbol{\Gamma}$ as $\boldsymbol{\Gamma}'=\boldsymbol{H}\boldsymbol{\Gamma}\boldsymbol{H}$, where $\boldsymbol{H}=\left[\boldsymbol{I}-\mathbf{1}\mathbf{1}^\top/N\right]$, and $\mathbf{1}\in\mathbb{R}^{N\mathcal{X}1}$ is the all-ones vector. Generally, the solution for $\boldsymbol{w}$ in the optimization problem of 3-9 is as follows [32]:

Consequently, the solution within the optimization problem in 3-9 (known as the centered kernel alignment –CKA) for calculating $\boldsymbol{w}$ is given by [32]:

$$\begin{aligned}\boldsymbol{w} &= \frac{\boldsymbol{A}^{-1}\boldsymbol{b}}{\|\boldsymbol{A}^{-1}\boldsymbol{b}\|} \\ \boldsymbol{A} &= \{a_{pq}=\langle \boldsymbol{S}'_p, \boldsymbol{S}'_q\rangle_F; \forall p,q\in[1,P]\}\in\mathbb{R}^{P\times P} \\ \boldsymbol{b} &= \{b_p=\langle \boldsymbol{S}'_p, \boldsymbol{K}'\rangle_F\}\in\mathbb{R}^{P\times 1}\end{aligned} \tag{3-10}$$

In order to achieve an image similarity function more related to the segmentation, the knowledge about the expected labels can be included within the optimization framework of 3-9. In particular, we learn the kernel parameters in a supervised scheme by making each one equals to the Dice Index similarity measuring the accuracy of segmentation between the atlases and defined as:

$$k^{n,m}=\frac{2\langle \boldsymbol{L}^n, \boldsymbol{L}^m\rangle}{(\|\boldsymbol{L}^n\|_1+\|\boldsymbol{L}^m\|_1)} \tag{3-11}$$

where notation $\|\cdot\|_1$ stands for the 1-norm.

3.1.4 Label Fusion

The attained measure can be used to select a subset of the most similar atlases to a target image so that the segmentation is improved. For accomplishing the segmentation, a weighted majority voting scheme is employed since it is the most widely used label fusion algorithm for multi-atlas-based segmentation approaches. Such a system provides a single label for each spell determined by collecting weighted votes from all the contributions over the selected templates and assigning the label with the highest vote to each voxel.

So, let $\mathscr{X}_T=\{\boldsymbol{X}^t, \boldsymbol{L}^t : t=1,\dots T\}$ a set of $\boldsymbol{T}$ selected atlases and $\hat{\boldsymbol{L}}^t$ be the estimated segmentation after deformable registration to the target image. Provided the similarity measure s^t between the target image and the selected atlases $\boldsymbol{X}^t$, matrix $\hat{\boldsymbol{L}}^t$ assigns the label $\hat{\boldsymbol{L}}(r)$ to each spel r, the final segmentation for the spels on the target image $\boldsymbol{O}(r)$ is given by:

$$\boldsymbol{O}(r)=\max_{\forall c}\{\boldsymbol{O}_1(r),\dots,\boldsymbol{O}_C(r),\dots,\boldsymbol{O}_C(r)\}$$

where $\boldsymbol{O}_c(r)=\sum_{t=1,T}\delta_c^n(r)$ and $\delta_c^n(r)=\begin{cases} s^t, & c=\hat{\boldsymbol{L}}^t(r) \\ 0, & \text{otherwise} \end{cases}$

3.2 Experimental Set-Up

In order to evaluate all studied metrics within the multi-atlas-based segmentation task, they are performed to select the most similar labeled images for estimating the segmentation of a query image. To this end, the majority voting scheme is considered for labeling each voxel since the segmentation quality is mostly dependant on the selection strategy. Additionally, the well-known Dice Index similarity is measured to evaluate the segmentation performance.

3.2.1 Database

Here, the dataset tested is the one used in the MICCAI 2012 *Multi-Atlas Labeling and Statistical Fusion* challenge[1] that is a subset of the Open Access Series of Imaging Studies (OASIS) database. This data collection holds T1-weighted structural MRI scans from 35 subjects (13 males and 22 females), aging from 18 to 90 years. Each 256×256×287-sized MRI volume has a voxel size of 1×1×1 *mm*. All images were expertly labeled with 26 structures. Due to our research interest lies in Parkinson surgery, only the following structures are considered: hypothalamus (HYPO), amygdala (AMYG), putamen (PUT), caudate nucleus (CAUD), thalamus (THAL) and pallidum (PAL)**3-1** shows a sample of an image subject as well as its provided segmentation.

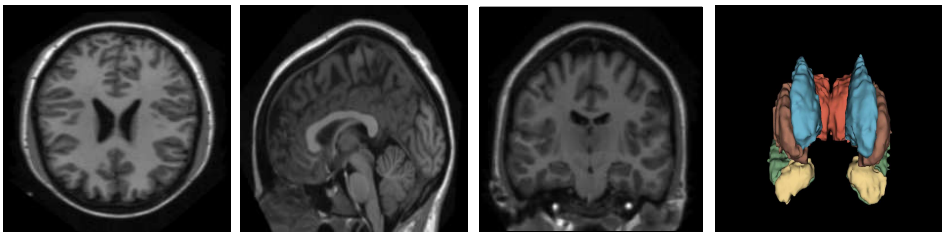


Figure **3-1**: Left to Right: Axial, Sagittal, and Coronal views as well as the ground-truth segmented structures.

[1] https://masi.vuse.vanderbilt.edu/workshop2012

3.2.2 Image Preprocessing

For the sake of comparison within a single common space, all images are spatially normalized into the Talairach space. Thus, each image is rigidly aligned to the ICBM atlas (MNI305-template) allowing to extract the morphological feature set accurately from each considered image. To this end, the Advanced Normalization Tool (ANTS) is employed using a quaternion based mapping and the MI metric as parameters.

In order to perform the label propagation, every pre-labelled image is also spatially mapped into the query image spatial coordinates (*target space*) with a non-linear transformation so that query and atlas images match the best. Further, the registration procedure is performed using the ANTS tool having the following default parameters: elastic deformation as the mapping function (`Elast`), MI as the similarity metric, and 32-bins histograms for estimating the probability density functions. Lastly, to get a finer alignment, the registration is performed at three sequential resolution levels: *i)* the coarsest alignment with a resolution of $1/8 \times$ *Original space*, and 100 iterations, *ii)* the middle resolution $1/4\times$*Original space* and 50 iterations, and *iii)* the finest deformation with a resolution of $1/2\times$*Original space* parameter and 25 iterations, the Gaussian regularization method is employed (σ=3).

3.2.3 Metric parameter learning

For including spatial information, each similarity metric is assessed separately at different locations of the image. To this end, the MRI volumes are regularly partitioned into P=27 blocks (3 partitions along each dimension). Then, the outcome metric is computed for all image pairs as a weighted linear combination of all local measures. Here, three different parameter tuning approaches are considered. The first one assumes that the contribution of each block to the similarity metric is equal for all of them. Therefore, the similarity metric corresponds to the averaging of the local measures. For the second approach, the weights are computed as the average intensity variance in each block as in 3-6. In this way, the shape differences on the brain structure intensity changes are considered as more relevant to the resulting metric than homogeneous regions. For the third approach, the weights are computed based on the contribution to the kernel centered alignment with respect to the objective kernel matrix $\boldsymbol{K}$ as introduced in 3-10.

3-3 shows the 3D scatter plotting all resulting weights, where the coordinates of each element are the spatial location of the image partition while the color and size are directly proportional to the value. As shown in 3.3(a), the central image region has the highest variance. Anatomically, this partition corresponds to basal ganglia location having tissue structures with high variant shape and intensity. However, the scatter plot also shows a substantial

amount of variance on the corners. This dispersion should be related to the partial presence of the scanned head and background, but not necessarily to the intensity dispersion. According with Figures 3.3(b) to 3.3(e), the weight distribution for all metrics exhibit the similar behavior. The corresponding weight to the central region is higher than other ones; i.e. the similarity metric assessed in this region is more correlated to the supervised kernel matrix than in boundary regions.

3.2.4 Evaluation of Similarity Metrics

We consider the leave-one-out validation scheme to evaluate the performance of the resulting metrics. In this case, all metrics are used to carry out an atlas selection task for the atlas-voting segmentation approach in the target image space. Finally, the metric performance is assessed with the Dice Index similarity.

3-2 shows obtained results of multi-atlas segmentation obtained by each tested metric and using all templates selected within the common space. As seen, the achieved accuracy gets close to the global computation when all weights become equal. Although the accuracy obtained by the MS metric should tend to the one of equally-weighted case, in practice both outcomes differ because of computational accuracy issues, but their resulting selection curves are statistically similar. Moreover, the noise artifact produces a high-variability of intensity, leading to unsuitable weights and a biased metric. About the weights computation, the similarity metric slightly improves the accuracy in comparison to the global and equally-weighted metrics if the variance intensity is taken into account. On the other hand, our proposed approach achieves the highest accuracy with an optimal atlas selection for all tested metrics. The approach outperforms not only all other benchmark combination methods, but also the accuracy obtained with the whole data-set.

3.3 Discussion

We propose a new spatially weighting criterion to improve similarity metrics aiming to measure image correspondences to support atlas selection on a multi-atlas segmentation scheme. Our proposed measure outperforms the widely used equal and variance-based weighting on the tested similarity metrics.

In accordance with the obtained results, it is clear that the spatially weighted metrics outperform the global ones. However, the computation of weights becomes an important task. The equal weighting provides spatial information in terms of partition size, but the con-

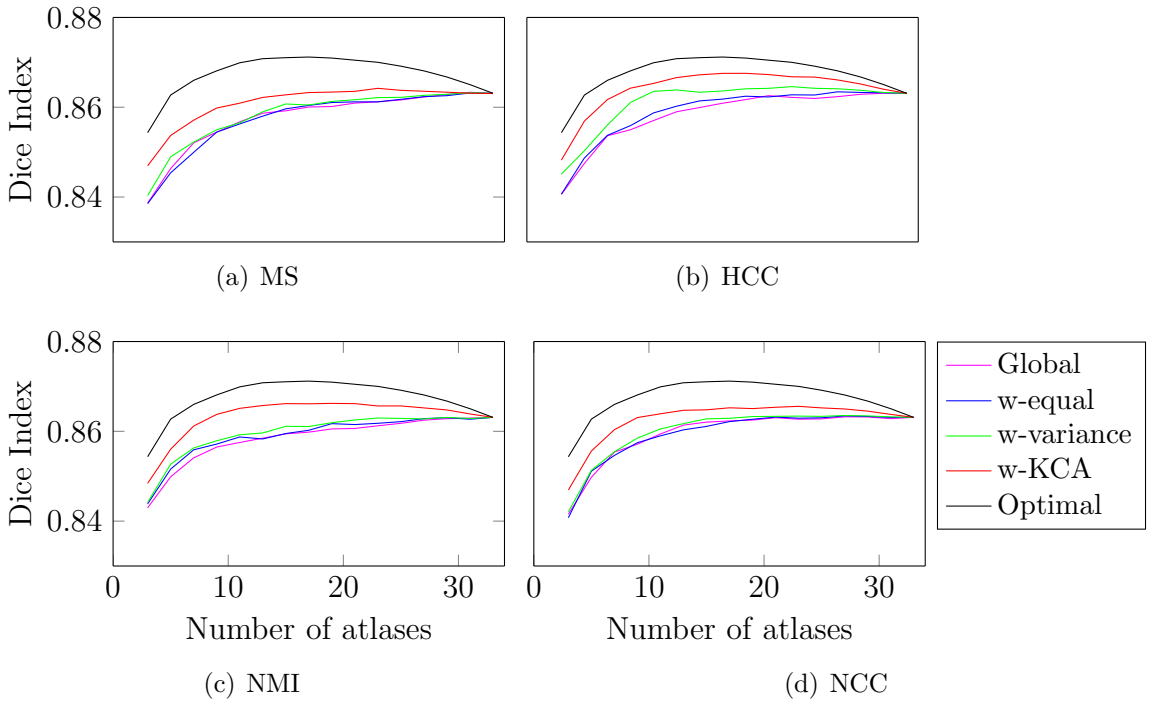
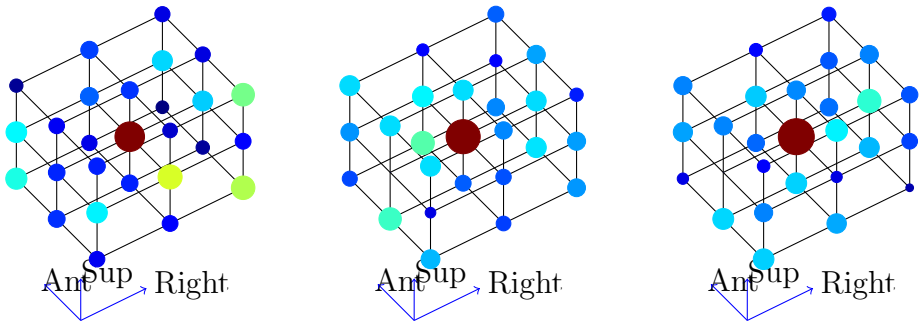


Figure **3-2**: Average Dice Index similarities versus the number of selected atlases for all considered metric tuning approaches.

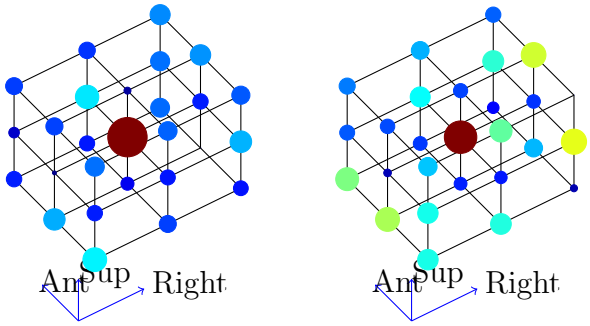
tent inside each locality is not taken into account. Hence, it performs the worst among all weighting approaches.

With the purpose of capturing shape differences on brain structures, the variance criterion weighs each block according to the intensity variance on it. However, the background on the images biases the weights computation towards image edges, which contain just partially the scanned head. Moreover, the inherent noise reduces the performance of a variability-based criterion.

Meanwhile, we propose to use a supervised kernel matrix aiming to learn the combination weights for the similarity metric. In this sense, the weights are computed according to the contribution to the kernel centered alignment with respect to the supervised kernel matrix. An advantage of our proposal is that the closed form to the kernel alignment solution provides an easy implementation for weights computation, while the construction of supervised kernel matrix is carried out only once off-line. As a result, we assess a suitable image similarity metric as an atlas selection criterion. Therefore, the similarity-based atlas ranking correlates correctly with the segmentation accuracy for subcortical structures. Actually, the learned metric can provide a subset of atlases achieving a higher accuracy for the label propagation than the whole dataset provides.



(a) Variance (b) MS (c) HCC



(d) NMI (e) NCC

Figure **3-3**: Resulting weight distribution for the variance criterion and all considered metrics. Markers are located at the center of each partition. Color and size are directly proportional to weight parameter value.

4 Target Specific Patch-atlas Construction

In Chapter 2 we stated Multi-atlas segmentation as a probabilistic atlas-based segmentation, to this end, the proposed similarity measure is used to select a set of most similar atlases aiming to construct a target specific probabilistic atlas which is specialized to segment an unlabeled image. Similarity construction of target specific atlas has been proposed for segmentation [58, 73, 25], specifically for theses methods, spatial prior information is stated as spatial probability maps of structure existence based on the training set, and then these priors are used to initialize the Expectation Maximization algorithm(EM). An advantage of Atlas-based EM methods is that they provide an intuitive way to incorporate spatial label and intensity information for segmentation. However, this method requires accurate registration between the training images and the target, which represents a high computational cost, nevertheless, registration errors can induce to an inadequate priors estimation. Recently, patch-based approaches have been proposed for dealing with registration errors; these methods perform label fusion based on local patch search, where labels are computed based only on local MRI intensity similarities. The key advantage of patch-based approaches is that it is not restricted by the quality of registration, providing better performance than conventional multi-atlas methods. Nevertheless, these methods label each point independently assuming the same labeling difficulty, however, some regions could be harder to segment than others, where the label affinity between the patches could not be related to patch similarity as is the case of the boundaries and low contrast regions. In this chapter, we propose to integrate a patch-based local search aiming to build target specific probabilistic atlases which are more representative to segment an unlabeled target image. Additionally, we account for spatial labeling difficulty performing the segmentation in an iterative scheme where the most confident points are firstly labeled and use this information to drive the segmentation of the less confident points. To this end, the search for the most similarity points is refined aggregating the label information in the previous iterations as well as global and local information to measure patch similarities. Finally, we propose to speed up the method using a Hierarchical scheme using only a few point to label the confident regions and gradually increment the number as the segmentation is less reliable.

4.1 Methods

4.1.1 Target specific probabilistic atlas segmentation

Given a target image $\boldsymbol{X}^q$, and a set of atlas $\mathscr{X}=\{\boldsymbol{X}^n, \boldsymbol{L}^n : n=1,\ldots,N\}$ with N image-segmentation pairs where $\boldsymbol{X}^n=\{x_r^n \in \mathbb{R} : r\in\Omega\}$ is the n-th MR image, the value r indexes all spatial elements, and the matrix $\boldsymbol{L}^n=\{l_r^n \in [1,C] : r\in\Omega\}$ is the provided image segmentation into $C\in\mathbb{N}$ classes. Conventional probabilistic atlas methods use the set $\mathscr{X}$ to construct a specific probabilistic atlas for a target image (usually a subset of the most similarities atlases is selected for atlas construction). Then, the estimation of the label map $\boldsymbol{L}^q$ is stated in a Bayesian framework where the probability of having the label map, $P(\boldsymbol{L}|\boldsymbol{X})$, the probability of occurring the image given the labeling, $P(\boldsymbol{X}|\boldsymbol{L})$ and the label prior, $P(\boldsymbol{L})$ are related as:

$$P(\boldsymbol{L}|\boldsymbol{X}) \propto P(\boldsymbol{X}|\boldsymbol{L})P(\boldsymbol{L}) \tag{4-1}$$

Where the probabilities $P(\boldsymbol{X}|\boldsymbol{L})$, $P(\boldsymbol{L})$ (termed atlas functions) vary along Ω. Then the probability of obtaining the a label map given a target image is represented as $P(\boldsymbol{L}|\boldsymbol{X}) = \prod_{r\in\Omega} P(l_r = C|x_r)$, and the probability of obtaining the label $c\in[1,C]$ at the position r given the voxel x_r is given by

$$P(l_r = c|x_r) = \frac{P(x_r|l_r = c)P(l_r)}{P(x_r)} \tag{4-2}$$

Where $P(x_r|l_r = c)$ is the conditional probability of observing voxel x_r at r given the label $l_r = c$, $P(l_r)$ is the label prior probability, and $P(x_r)$ is the evidence probability which are given by the generated probabilistic atlas. At that point, the segmentation process relies on the estimation of the atlas functions. For this purpose a subset of the most similar images and their respective segmentations are registered to the target image. Then, the prior probability $P(l_r)$ is estimated by computing the probability of occurrence of each class for all $r \in \Omega$ along of all atlases. And $P(x_r|l_r = c)$ is the intensity distribution of each class c which is approximated by a Gaussian distribution, $P(x_r|\mu_c, \sigma_c^2)$ where μ_c, σ_c^2 are estimated using EM algorithm.

4.1.2 Probabilistic atlas generation

As aforementioned, the construction of the atlas functions plays a significant role in the segmentation. However, this process can be affected by registration errors. Even though the most similar atlases are selected for atlas construction, local anatomical differences are

hard to cover by the registration algorithms. Recently, patch fusion methods were proposed to overcome limitations produced by registration errors, on these schemes each voxel is represented by a patch and a set of the neighboring patches are used to estimate a label. We use the patch fusion scheme aiming to construct a probabilistic local model, in this way the resultant probabilistic atlas could be more representative of the target image.

Following the traditional patch fusion schemes, for each spatial position r we extract a set of patches $\mathscr{P}_r = \{\boldsymbol{\beta}_y^n \subset \boldsymbol{X}^n, \boldsymbol{\gamma}_y^n \subset \boldsymbol{L^n}{:}y{\in}\eta(r)\}$, from the registered training dataset, where $\eta(r)$ is a neighborhood around r, and denote the target patch as $\boldsymbol{\beta}_r^q \subset \boldsymbol{X}^q$. In this sense, each patch is an arrange $\boldsymbol{\beta}_y^n = \{x_s^n : \|y-s\| < \xi\}$, and $\boldsymbol{\gamma}_y^n = \{l_s^n : \|y-s\| < \xi\}$, being ξ the patch radius. Let $\mathscr{D}_r = \{\boldsymbol{\beta}_y^t, \boldsymbol{\gamma}_y^t : t = 1 \cdots T\}$ is a patch dictionary that holds the set of most similar patches to the target patch, chosen according to the well known structural similarity measurement ss [33].

$$ss(\beta^q, \beta^t) = \frac{2\mu_q\mu_t}{\mu_q^2 + \mu_t^2} \times \frac{2\sigma_q\sigma_t}{\sigma_q^2 + \sigma_t^2} \tag{4-3}$$

Where (μ_q, σ_q) and (μ_t, σ_t) are the mean and standard deviation of β^q and β^t respectively. Then, a threshold ρ_1 is use to select the most similar patches i.e $\forall_{\beta^t \in \mathscr{D}_r}\ ss(\beta^q, \beta^t) > \rho_1$.

Aiming to construct probabilistic patch function, the eq. (4-1) is rewritten as:

$$P(\boldsymbol{\gamma}|\boldsymbol{\beta}) \propto P(\boldsymbol{\beta}|\boldsymbol{\gamma})P(\boldsymbol{\gamma}) \tag{4-4}$$

Then the objective is to construct the atlas functions $P(\boldsymbol{\beta}|\boldsymbol{\gamma}), P(\boldsymbol{\gamma})$ from the set of patches $\mathscr{D}_r$, for this purpose for all spatial position s in $\boldsymbol{\beta}^t$ the prior $P(l_s = c)$ and the evidence $P(x_s)$ can be computed as:

$$P(l_s = c) = E\{\delta(l_s^t = c) : \forall t \in T\} \tag{4-5}$$

$$P(x_s) = E\{\sum_{c=C} P(x_s|l_s = c)P(l_s = c)\} \tag{4-6}$$

Where $\delta(\cdot)$ is the delta function and notation $E\{\cdot\}$ stands for the expectation operator. Regarding to the intensity distribution $P(x_s|l_s = c)$ which is assumed to be Gaussian ($i.e\ P(x_s|\mu_c, \sigma_c)$), Traditionally, Expectation maximization is used to estimated the parameters μ_c, σ_c. However, even this approach has been successfully used for segmentation of brain tissues [25, 7, 73], the convergence of the algorithm in other structures as the sub-cortical ones could be no garantized, this due to low contrast on the boundaries and intensity inhomogeneities. We account this issue by using the selected patches to construct spatial appearance models. In this way, assuming x_s is normally distributed in each class, $x_s \sim \mathcal{N}(\mu_s^c, \sigma_s^{2c})$. The

mean $u_s^c \in \mathbb{R}^d$ and variance $\sigma_s^{2c} \in \mathbb{R}^d$ are estimated as:

$$\mu_s^c = E\{x_s^t | l_s^t = c : \forall t \in T\} \tag{4-7}$$
$$\sigma_s^{2c} = E\{(x_s^t - \mu_s^c)^2\} \tag{4-8}$$

Note that, μ_s^c and σ_s^c can be computed iteratively, as the x_s^t and l_s^t are aggregated for each point. To this end, we use the online computation of mean and variance proposed by [22]. Then, the segmentation for target voxel can be obtained using the maximum a posteriori (MAP) criterion.

$$l_s^q = \arg\max_c P(l_s^q = c | x_s) \tag{4-9}$$

with

$$P(l_s^q = c | x_s) = \frac{P(x_s^q | l_s = c) P(l_s = c)}{P(x_s^q)} \tag{4-10}$$

Multipoint Overlapping

Multipoint estimation was first analyzed by [66], instead to compute the central voxel of the patch, the whole patch can be segmented using the same weighting scheme. Then the overlaps estimations are fused by majority vote, they conclude that multipoint outperforms the point-wise estimations. Moreover, they evaluate the performance of multipoint strategy using only a subset of points for segmenting the whole image (fast multipoint), although fast multipoint obtain less accurate results, the time for computation was considerably reduced. Similarly in our method, given that the patch extraction is performed for each r along Ω, there are multiples estimations for each spatial position r, which are given by the neighboring points. Therefore, we aggregate the overlapping voxels on the estimation of the atlas function, note that it is possible due to the iterative methodology used to compute the parameters of the appearance models, in the same way, it is easy to compute the priors $P(l_s)$ by aggregating voxels to the models.

4.1.3 Iterative labeling

One of the drawbacks in the majority of patch-based methods is that each voxel is labeled independently assuming the same labeling difficulty. However, some regions as the boundaries or low contrast regions are more difficult to segment than those inside of the structures. Some authors have tackled this issue, for example, [9] performs statistical fusion where spatial varying models of rather performance are constructed by estimating the consensus level of each voxel. On the other hand, [69] label the most confident points and use the partial

segmentation to guide the segmentation of the less confident points iteratively. For this purpose, the obtained labels in each iteration are used to refine the patch selection for the less confident points. Similarly to [69], we use a confidence criterion to estimate the regions where atlas functions are more representative from the target image, being capable of generating confident labels. On the other hand, we assume that atlas functions in the less confident regions need to be recomputed. In this way, confident atlas functions are used to computed a partial segmentation which is used to drive the labeling for the less confident regions in an iterative scheme. In the same way to [69], the computed labels on the previous iteration are used improve the search of the most similar patches. Moreover, we account for the global and local information to build the new refined patch dictionary. Finally aiming to reduce the computational time, we use a fast multipoint estimation, where a subset of points Ω^* is used to label the high confident regions. Then, the subset Ω^* is augmented at each iteration.

Confidence measure

As was discussing in Section 4.1.2 for each target voxel x_r^q we obtain the vector

$[P(l_r^q = 1|x_r^q) \cdots P(l_r^q = C|x_r^q)]$ which represent the probability of a target voxel to belong to each class. Note, that each element in the vector can be seen as the grade of confidence to be labeled with the class c. Given that, the voxel is labeled according to MAP criterion, in each iteration only the points where the maximum likelihood is superior to a threshold λ are labeled, λ is decreased in each iteration, this process is carry out as follow.

$$l_r^{q(t)} = \begin{cases} c & \max\limits_c(P(l_r^q = c|x_r^q)) > \lambda \\ \perp & \max\limits_c(P(l_r^q = c|x_r^q)) < \lambda \end{cases} \tag{4-11}$$

Where, $l_r^{q(t)}$ is the label at r in the iteration t, and $\perp$ denotes the unlabeled point.

Patch dictionary refinement

Selection of an appropriate patch dictionary is of vital importance for obtaining accurate segmentation in patch-based methods. A straightforward way to select similar patches is using intensity similarities [33]. However, selection based only on intensity similarities could bias the performance. Another issue that affects the appropriate dictionary construction is the use of fixed sizes for the patch and search neighborhoods. As was demonstrated in [83], patches with a fixed size prevent capturing the global and local information. Also, a fixed size for the search neighborhood imposes a restriction on the tolerance to registration errors. Even though these parameters can be determined from a given dataset, some regions

can present more variability, therefore, need a larger search neighborhood size to cover the registration errors. Aiming, to overcome these issues, we propose a hierarchical search where the global information is firstly captured using larger sizes for the patches and search windows, then the sizes are reduced gradually to capture the local information. Additionally, we refined the patch dictionary by a label similarity measure employing the partial labels obtained in the previous iterations using the approach proposed in [69].

The process is as follow, suppose that we want to build a patch dictionary for a target patch centered on a given position r. Then at the first level, we fix the patch size with radius ξ_1 and establish the search neighborhood $\eta(r)$ with radius $\xi_1 + 2$ then the set of most similar patches are extracted. For the next levels, let r' the position for one of the selected patches on the previous level, the search neighborhood is established as $\eta(r')$ where $\xi_i = \xi_{(i-1)} - 1$, this process continues until $\xi_i = 1$. At the end of the hierarchical patch selection, we obtain a patch dictionary $\mathscr{D}_r = \{\boldsymbol{\beta}^t, \boldsymbol{\gamma}^t : t = 1 \cdots T\}$, then as was mentioned in section 4.1.3 we use the partial labeling obtained for the confident points for guide the patch selection. Let $L^q(t-1)$ the label map from the previous iteration, then we extract the partial labels for the target patch at the iteration $(t-1)$ denoted as $\widetilde{\gamma^q}$. Thus, the refined patch dictionary denoted as $\widetilde{\mathscr{D}}_r$ is build using only the patches with high label similarity to the partial target label patch, as defined in the following:

$$\widetilde{\mathscr{D}}_r = \{\beta^t, \gamma^t | sim(\gamma^t, \widetilde{\gamma^q}) > \rho_2\} \tag{4-12}$$

where $0 \leq \rho_2 \geq 1$ is a label similarity threshold and $sim(\gamma_t, \widetilde{\gamma^q})$ is a label similarity measure between the label patch γ_t and the partial label target patch $\widetilde{\gamma}^q$ which is de defined as:

$$sim(\gamma_t, \widetilde{\gamma^q}) = \frac{|id(\gamma_t) \cap id(\widetilde{\gamma}^q)|}{|id(\widetilde{\gamma}^q)|} \tag{4-13}$$

Where $id(\widetilde{\gamma}^q)$ is the function which denote the the set of point in $\widetilde{\gamma}^q$ which are labeled in the previous iteration, and $|\cdot|$ denotes the cardinality of the set.

Hierarchical Fast Multi-point estimation

In [66] it was demonstrated that there is not necessary to use all points over the image when the whole target patch is segmented. Instead, they use a fast-multipoint estimation where a subset of points is selected to reduce the overlapping estimates, with the constraint that there is, at least, one label estimate for each voxel. However using all point achieve higher accuracy that fast multipoint estimation. In light of this, we propose to start the segmentation with a set of points over a 3D grid where the distance between the grid points is equal to the half of the employed patch size. Then in each iteration, the distance between the grid points is reduced until this is equal to one. Given that, in each iteration only it

is necessary to label the unconfident points computed in the previous iteration, only those over the grid are used for estimations.

4.2 Experiments

We evaluate the performance of our method in the segmentation of subcortical structures using the SATA Miccai Challenge Dataset. Our proposal is compared against the conventional segmentation based on target specific probabilistic atlas and EM algorithm (**Target-atlas EM**, and the conventional patch-based methods: Non-Local Weighted Voting(**NLWV**) and the Sparse Patch-Based Labeling (**SPBL**). Also, we evaluate the proposed segmentation method based on target-specific patch atlas construction without using the iterative scheme (Target specific Patch Atlas Segmentation **TSPAS**) and our complete iterative algorithm (Hierarchical-Fast Target Specific Patch Atlas Segmentation **HF_TSPAS**), Qualitative evaluation is carried out using the Dice ration index.

4.2.1 Database Description

The SATA MICCAI online challenge has been proposed as a framework for comparison of several approaches which requires a set of training data. Although it has a particular interest in Multi-atlas methods, another type of approaches as Active shape and appearance models have been considered. The SATA Dataset consisted of 35 training samples with manual labels with 14 structures which correspond to both hemispheres of accumbens area, amygdala, caudate, hippocampus, pallidum, putamen, and thalamus, also, all pair deformable registration for all images are provided.

4.2.2 Experimental results Target-specific atlas construction

Aiming to evaluate the performance of target patch atlas construction, we firstly evaluate our posed method without the iterative scheme(**TSPAS**) using all points over the image. Additionally, a leave-one-out strategy is used to compare the performance of the methods above, given that one of our goals is to reduce the computational cost all methods are evaluated using only affine registration as preprocessing. To this end, for each leave-one-out case, each one of the remaining images is affine-registered to the target space using the ANTs toolbox with default parameters and mutual information as a cost function.
For a fair comparison, the size of the patch and search neighborhood are tuned independently for each patch method. Also, due to our proposed estimation take into account patch

overlapping for the atlas functions estimation. We use multi-point estimation for Non-local weighting voting **NLWV** and Sparse based labeling **SPBL** methods. Moreover, multi-point estimation has demonstrated better results than point-wise segmentation (labeling only the central point) [66, 75]. In order to obtain the optimal parameters set, we randomly select 10 subjects from the training set. Additionally, we employ a grid search with patch radius range from $\xi = [2, 3]$ (5×5×5 to 7×7×7) and neighborhood search radius from 2 to 5 (5×5×5 to 11×11×11). Also, we found that $\rho_1 = 0.9$ was the best value for the patch similarity threshold, and the l_1-norm strength in the SPBM is set to $\alpha = 0.01$. Figure **4-1** shows the Dice ratios for all possible configurations, where the highest values for our proposed and SPBL are achieved with patch radius of 3 and search neighborhood with radius 4 while for NLWV are 2 and 4 respectively. Note that our proposed outperform to NLWV for all possible configurations. Meanwhile, the average dice ratio for our proposal using patch size $\xi = 2$ is close to the achieved by SPBL, however, there is an improvement when the patch size is $\xi = 3$.

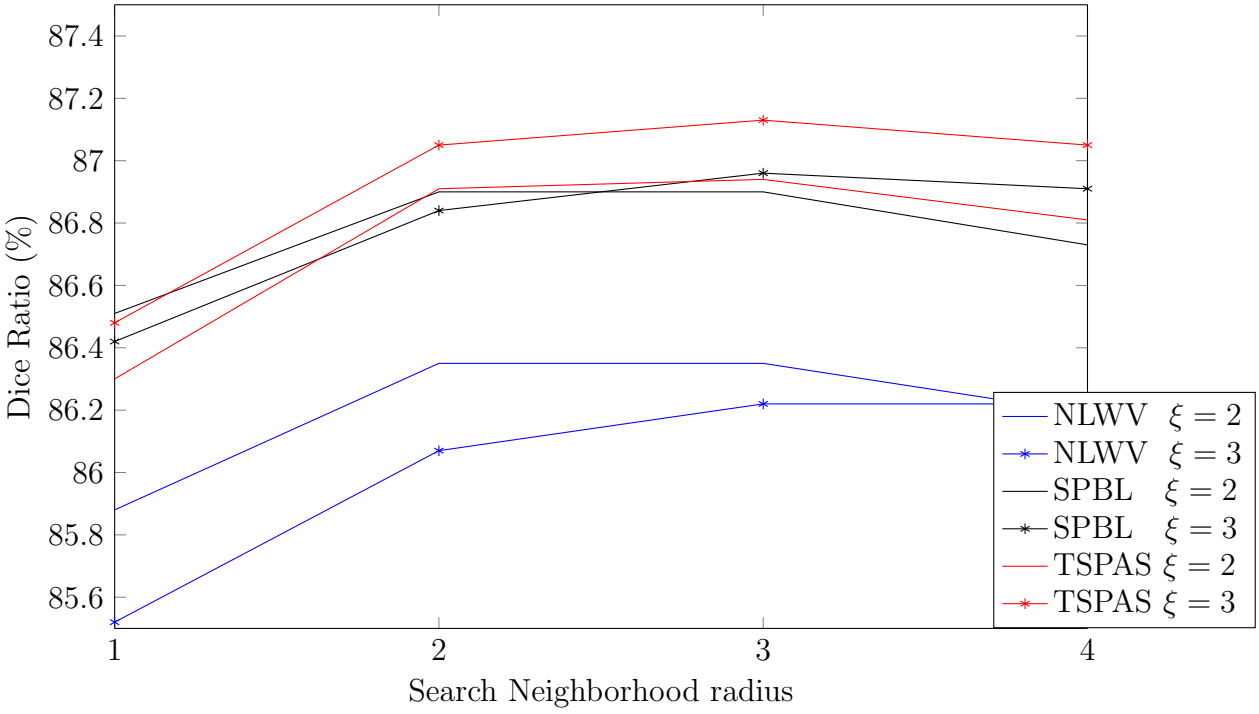


Figure **4-1**: Hirarchical patch selection

Aiming to give an insight of the advantage of incorporate patch search in atlas construction, all methods are evaluated over the whole training data using the best parameter found above. Table **4-1** shows the average Dice ratio for each structure(left and right combined) by each method. It is clear that our proposed method achieves the best labeling accuracy in each

structure, where the overall segmentation is improved in 1% and 0.42% respect to NLWV and SPBL respectively. It is worth noting that our proposed method outperforms in 26.84% to **Target-Atlas EM** method. The worse performance for this method can be due to this requires a good initialization for the EM algorithm which is difficult to achieve if the atlases are only affine-registered to the target. However, our proposed overcome this limitation by including patch search on the target atlas construction.

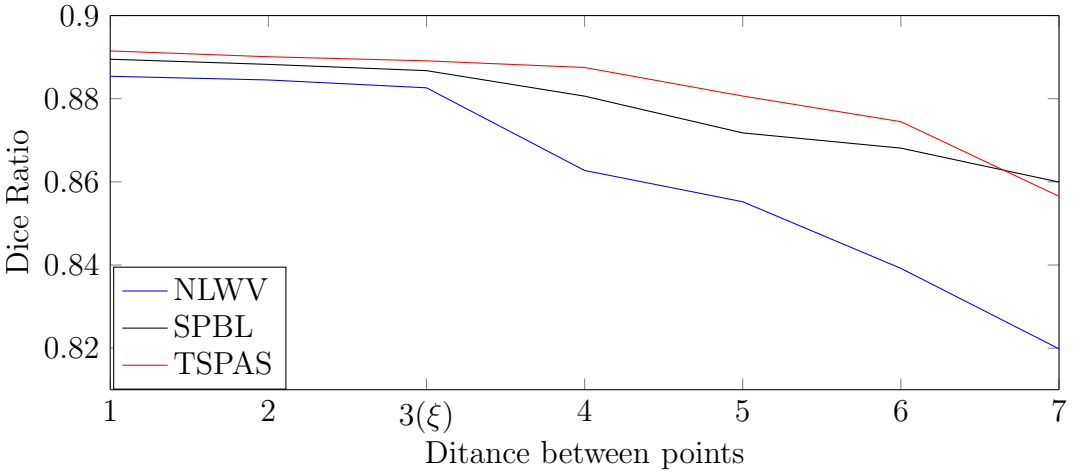
	Target-Atlas EM	**NLWV**	**SPBL**	**TSPAS**
Accumbens	38.70 (38.70)	79.39 (4.02)	80.32 (3.71)	**81.10 (3.74)**
Amygdala	51.91 (51.91)	80.99 (2.64)	81.67 (2.59)	**82.32 (2.89)**
Caudate	63.07 (63.07)	90.24 (2.30)	90.90 (2.32)	**90.97 (2.33)**
Hippocampus	59.61 (59.61)	86.54 (1.95)	87.20 (1.63)	**87.59 (1.66)**
Pallidum	63.31 (63.31)	87.58 (2.70)	87.99 (2.41)	**88.47 (2.42)**
Putamen	70.87 (70.87)	91.47 (2.37)	91.89 (2.24)	**92.30 (2.13)**
Thalamus	80.23 (80.23)	92.30 (1.13)	92.66 (1.00)	**92.83 (1.02)**
Overall	61.10 (61.10)	86.93 (2.45)	87.52 (2.27)	**87.94 (2.31)**

Table **4-1**: The mean and the standard deviation of the Dice ratio (%) for all structures by **Target-Atlas EM**, **NLWV** [33] **SPBL** [87] and our proposal **TSPAS**

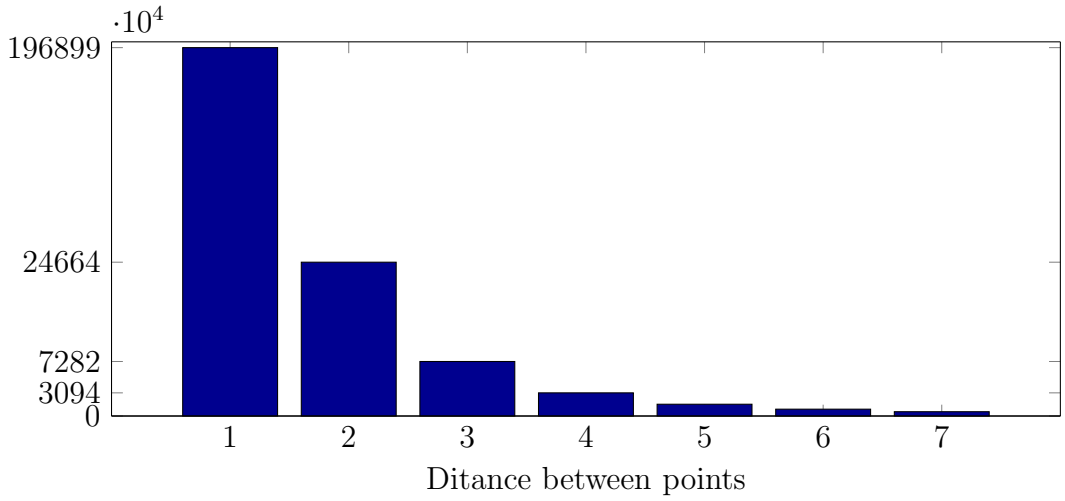
4.2.3 Fast multi-point and the hierarchical fast multi-point algorithm

Given that, our proposal is based on the fast-multipoint estimation we analyze the performance of all patch-based methods under this scheme. To this end, using the same sizes for patch and neighborhood search found for multipoint estimation, we select a subset of points over the image in such a way that the distance between each point is equal to the ratio of the patch ξ. This configuration was selected due to it represent a trade-off between accuracy and computational cost. To demonstrate this point we evaluate the performance of fast multipoint scheme changing the distance between the points. Figure **4-2** (Top) shows the dice ratio using a patch size of 7×7×7 ($\xi = 3$) and a range of distances from 1 to 7 points in all patch methods. As we can see on the top figure, there is a slight decrease for distances between 1 to 3, whereas it is more significant for superior values. Note that taking a distance of one point means use all points on the image, meanwhile with a distance of 7 points ($2\xi + 1$) there are not overlapping on the estimation. On the other hand, if the distance is greater than 7 ($> 2\xi + 1$), some voxels are not estimated. Also, according to Figure **4-2**(Bottom) it is clear that there is an important reduction in the number of points as the distance increases, specifically, a distance of ξ points leads to a speed-up of 27.

Aiming to evaluate the performance of our complete hierarchical algorithm (**HF_TSPAS**), we use the same ten SATA subjects for parameter tuning. For the initial patch radius we



(a)



(b)

Figure **4-2**: The top Figure shows the evolution of the Dice ratio as the distance between the points increase. Bottom Figure shows the total number of points required to segment an image.

use $\xi_1 = 9$ and the threshold for label similarity measure was $\rho_2 = 0.9$. The results of the average dice ratios for all structures are reported in Table **4-2** for the baseline patch methods and our proposals. It is worth noting, that our proposed method using the fast multi-point estimation(TSPAS) outperforms both NLWV and SPBL by 1.69% and 0.57% respectively. Regarding our complete hierarchical method HF_TSPAS, we can see that it increases the difference respect to patch based methods obtaining improvements of 2.12% and 1% for NLWV and SPBL. Most importantly, our proposed method outperform all methods on multipoint estimation. Thus, it outperforms by $27\%, 1.25\%, 0.66\%$ to Target-atlas EM, NLWV

and SPBL respectively, also an improvement of 0.24% is obtained with respect to TSPAS. However, the hierarchical algorithm requires considerably less time for its computation.

	Fast multi-point Estimation			
	NLWV	**SPBL**	**TSPAS**	**HF_TSPAS**
Accumbes	78.72 (4.43)	79.92 (3.44)	80.68 (3.73)	**81.44 (3.65)**
Amygdala	80.02 (3.11)	81.02 (2.67)	82.00 (2.88)	**82.77 (2.65)**
Caudate	90.20 (2.56)	90.60 (2.52)	90.84 (2.43)	**91.08 (2.36)**
Hippocampus	84.72 (4.41)	86.81 (1.58)	87.44 (1.69)	**87.85 (1.68)**
Pallidum	86.01 (3.79)	87.71 (2.57)	88.34 (2.58)	**88.72 (2.21)**
Putamen	90.60 (2.87)	91.66 (2.37)	92.17 (2.19)	**92.52 (1.87)**
Thalamus	92.16 (1.21)	92.51 (1.02)	92.76 (1.05)	**92.85 (0.98)**
Overall	86.06 (3.19)	87.18 (2.31)	87.75 (2.36)	**88.18 (2.20)**

Table **4-2**: The mean and the standard deviation of the Dice ratio (%) for all structures by **NLWV** [33] **SPBL** [87] and our complete Hierarchical method(**HF_TSPAS**) under fast-point scheme.

In order to give an insight about the contribution of the iterative methodology, we evaluate the evolution of our complete hierarchical method with iterations, on the other hand, given that, our proposal adds Hierarchical search to the label similarity proposed by [69]. We also tested the hierarchical algorithm using only the label similarity measure for patch dictionary refinement. Figure **4-3** show the evolution of the dice ration with iterations, as we can see, the segmentation accuracy increases most significantly during the first three iterations, where it stabilizes. Therefore, it is not necessary to carry out through all iterations which lead to reducing the computational cost. Also, it is worth nothing that using the hierarchical search leads to improve the segmentation accuracy outperforming the patch dictionary with label similarity alone.

4.2.4 Influence of deformable registration

Given that, one of our goals is to reduce the computational time, our proposed method was first applied using linear registration. However, in [66] was demonstrated the advantage of combine patch methods with non-linear registration. In light of this, we apply our proposed method after performing deformable registration, to this end, we use the pair-wise registration provided by MICCAI Sata challenge. Due to deformable registration reduce the spatial anatomical variability, patch and neighborhood search size were tuned again using a grid search with radius range of 1 to 3 ($3\times3\times3$ to $7\times7\times7$) for both patch and search neighborhood. We found that the best configuration for all patch-based methods is a patch size radius

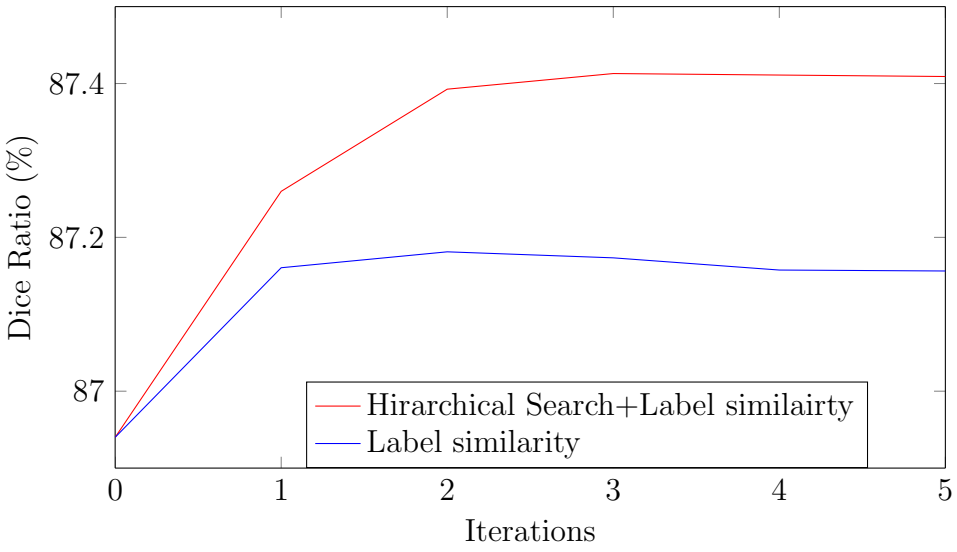


Figure **4-3**: Evolution of Dice ratio with iterations of our Hierarchical method

$\xi = 2$ and neighborhood search radius 1. Note that there is a considerable reduction in the search neighborhood radius, this can due to the high-quality of the registrations. Similarly, to experiments using affine registration we first evaluate the performance using all points on the image for all patch-based methods and our proposal (TSPAS). Also, we evaluate the segmentation performance for the Target-atlas EM method, results are shown in Table **4-3**. Again our method TSPAS perform the best segmentation, however, the improvements respect to patch-based methods are lower compared with results using affine registration 1% against 0.45% for NLWV and 0.42% against 0.13% for SPBL . Also, it is worth noting that performance for the Target-atlas EM method using deformable registrations improves by 24.96% respect to the obtained results using affine registrations, this fact can be due to deformable registration provided a better initialization of the EM algorithm. However, our proposal outperforms Target-Atlas EM by 3.1% demonstrating that include patch search leads a better target specific atlas construction. Also, we evaluate the performance of our complete hierarchical algorithm and compare the results with all patch fusion methods in a fast multipoint scheme. The results are shown in the Table **4-4**, where our complete hierarchical algorithm achieves the best performance. Moreover, our proposed method without hierarchical methodology (TSPAS) outperform both patch-based methods. However, it is worth noting that our complete hierarchical method (HF_TSPAS) improves in 0.13% to our proposal without hierarchical methodology (TSPAS), and 0.06% to our proposal using all point. Which are lower compared with those obtained using affine registration (0.43% and 0.24%). Also, we can see that the performance of the complete hierarchical method using affine registration is comparable to the obtained using deformable registration. However, the computational cost to perform accurate non-linear registration is significantly higher compared with the required for affine registration.

	Target-Atlas EM	**NLWV**	**SPBL**	**TSPAS**
Accumbes	78.16 (78.16)	80.84 (3.95)	81.21 (3.86)	**81.50 (3.93)**
Amygdala	80.29 (80.29)	82.86 (2.42)	83.04 (2.27)	**83.17 (2.51)**
Caudate	82.31 (82.31)	90.67 (2.72)	**91.28 (2.53)**	91.21 (2.61)
Hippocampus	83.75 (83.75)	86.95 (1.84)	87.58 (1.53)	**87.73 (1.51)**
Pallidum	88.22 (88.22)	88.49 (2.26)	88.71 (2.24)	**88.95 (2.21)**
Putamen	92.11 (92.11)	92.32 (2.07)	92.50 (2.01)	**92.58 (2.00)**
Thalamus	91.11 (91.11)	92.67 (1.03)	92.76 (1.04)	**92.80 (1.03)**
Overall	85.14 (85.14)	87.83 (2.33)	88.15 (2.21)	**88.28 (2.26)**

Table **4-3**: The mean and the standard deviation of the Dice ratio (%) for all structures by **Target-Atlas EM**, **NLWV** [33] **SPBL** [87] and our proposal **TSPAS** using deformable registration.

	Fast multi-point Estimation			
	NLWV	**SPBL**	**TSPAS**	**HF_TSPAS**
Accumbes	80.61 (3.95)	81.10 (3.79)	**81.46 (3.87)**	81.24 (3.78)
Amygdala	82.53 (2.46)	82.86 (2.24)	83.05 (2.41)	**83.16 (2.31)**
Caudate	90.53 (2.66)	91.17 (2.49)	91.15 (2.63)	**91.34 (2.63)**
Hippocampus	86.38 (2.42)	87.40 (1.53)	87.66 (1.53)	**88.00 (1.43)**
Pallidum	88.14 (2.31)	88.57 (2.26)	88.88 (2.19)	**89.05 (2.08)**
Putamen	92.06 (2.13)	92.41 (2.03)	92.53 (2.00)	**92.72 (1.89)**
Thalamus	92.52 (1.02)	92.68 (1.02)	92.76 (1.03)	**92.87 (1.01)**
Overall	87.54 (2.42)	88.03 (2.19)	88.21 (2.24)	**88.34 (2.16)**

Table **4-4**: The mean and the standard deviation of the Dice ratio (%) for all structures by **NLWV** [33] **SPBL** [87], **TSPAS** and our complete Hierarchical method **HF_TSPAS** under fast-point scheme using deformable registration.

4.3 Discussion

The proposed method in this chapter addresses the Multi-atlas segmentation by introducing non-local patch search in the construction of target specific atlases. Thus, segmentation could be stated in a Bayesian framework where local bayesian classifiers are built for each point on the image. In this sense, the strength of both probabilistic atlas and patch based segmentation approaches are combined, enabling to deal with anatomical variability whereas a parametric model leads the segmentation. In addition, the proposed method is enhanced to account for spatial labeling difficulty, allowing to refine the segmentation for those regions where the segmentation could be harder to estimate. Finally, the computational burden is reduced by employing a Hierarchical fast multipoint strategy where a few points over the

image with a small overlapping are used to segment the regions which are reliable for labeling. Meanwhile, the overlapping is augmented for regions which require better estimation for its segmentation.

In order to give an insight of the contributions of our proposal, we assessed the segmentation performance when local patch search are introduced on atlas construction alone (TSPAS), and then the complete hierarchical segmentation method was applied for segmentation(HF_TSPAS). Obtained results using affine registrations shows that TSPAS achieve a considerable improvement respect to conventional Target-atlas based segmentation methods, demonstrating the advantage to include local patch search in atlas construction. Moreover, our proposal outperforms conventional patch-based methods in both multipoint and fast multipoint scheme, allowing to infer that bayesian classification provides and effective way obtain accurate segmentation compared with traditional weighted fusion steps.

Regarding to results using the provided deformable registration, we can note a considerable improvement on the conventional Target-atlas based segmentation method, this can due to deformable registration reduces effects of variability on the training set providing a better initialization for the EM algorithm. On the other hand, even thought the proposed method outperform to patch-based methods. The performance difference is lower compared with the obtained results using affine registration. Especially, with the sparse based labeling method, where TSPAS is slightly superior, this fact can be due to a lower amount of patches are required for label fusion when a proper registration is provided. Thus sparse patches labeling is appropriate to eliminate the misleading patches or similar neighboring patches in the same image wich could bias the segmentation. However, obtaining tuned size for the neighborhood search suggests there is high quality on the registration which represents a high computational cost.

The obtained results after evaluating our complete hierarchical method(HF_TSPAS) using the affine registration shows that our proposal outperforms both Target atlas-based method and the patch-based methods using fast multipoint and multipoint estimation. Most important, HF_TSPAS outperforms to TSPAS in a multipoint scheme. However, the HF_TSPAS reduces the computational time considerably. On the other hand, comparing the results using deformable registration, it can be noted that there slightly improvement between HF_TSPAS and TSPAS, it can be due to hierarchical patch search have a poor effect when there is accurate deformable registration. However, it is worth nothing that HF_TSPAS using affine registration achieve comparable performance to that using deformable registration, nevertheless, achieve accurate registration represent increases the computational burden.

5 Conclusions and future work

5.1 Conclusions

This work addresses structure segmentation of medical images using a set of labeled images termed atlas. Specifically, this research focuses on multi-atlas segmentation where a subset of most representative atlases are combined based on the relationships between the atlases and the target of interest to obtain the target segmentation. Under this scheme, several methodologies to enhanced the appropriate selection and combination of atlases are presented. Overall attained results demonstrated that the proposed approaches improve the segmentation accuracy outperforming conventional multi-atlas methods. Following, the main concluding remarks regarding each proposed approach are described.

- A kernel based MRI representation that allows highlight the intrinsic relationships between images is presented to support MRI clustering task and atlas selection. Our approach is able to reduce the original image space encoding the inter-slice similarities so that the main shape information is kept in a lower dimensional space. In addition, a kernel-based embedding space is computed to enhance the measurement of image correspondences. Thus, distances calculated in this compact space are able to reflect the intrinsic similarities between images, enabling to highlight the latent data structure of the data. For the sake of comparison, our proposed approach is tested as MRI representation strategy to support clustering task aiming to identify demographic categories age and gender. Attained results demonstrated that proposed method is able to identify latent groups on the data related to the gender and age, achieving better clustering performance that state-of-the-art methodologies. Also, our proposal is suitable for clustering and measurements tasks required on multi-atlas based segmentation methods, enabling to construct population-probabilistic or target specific atlas by selecting the appropriate subset of images to segment the underlying target image. To demonstrate this point, multi-atlas segmentation was stated as probabilistic atlas segmentations scheme where similarities computed in the embedding space were used to select the appropriate set of images to construct target specific probabilistic atlases. Attained results, shown that selection in embedding space, outperform the conventional demographic atlas selection, as well as the whole population atlas on

brain tissue segmentation.

- A supervised similarity measure that learn the relationship between local appearance correspondences and the segmentation performance is presented to enhance multia-atlas segmentation results. Our approach computes independently local similarities in regular partitions then similarity values are linear weighted combined to get a single similarity outcome. Thus, a supervised similarity measure which encodes the label affinity between the images in the training dataset is used to learn the most appropriate combination weights, so the resulting similarity measure between a pair of images is highly related to the similarity of their provided labelings. To this end, the centered kernel alignment(CKA) between the supervised kernel and the image kernel is maximized about the weight values. In this sense, the proposed similarity measure account for capturing local relationships, meanwhile provide appropriate atlas selection based on expected performance to segment a target image. The introduced similarity measure was compared with the global measures assessment as well as two state of the art combination approaches. Attained results allow highlighting the contributions of our approaches, since the local measures outperform the global ones, while, supervised weighted methodology outperform those unsupervised.

- A new method for label fusion that integrates patch based and probabilistic atlas-based segmentation methods was presented. Our approach integrates patch based local search to enhance target specific atlas construction, in such a way that the probabilistic atlas is more representative to label the underlying target image. In this sense, the strengths of both approaches are combined allowing to reduce the effects the registration while intensity and label information are incorporated in the label fusion step by a bayesian framework. Furthermore, we have presented an iterative framework which gradually estimate labels on the images according to labeling difficulty. Thus, the estimations for the regions which present a reliable segmentation are use to refine estimations for those where is less reliable. For refining of estimations, the proposed method focuses on the enhancing of the local patch search. Thus, the refine search is achieved by incorporating label information from the segmentation on the previous iterations as well as incorporating local and global information, besides, the constraint to the registration errors is relaxed by changing the size of the of the neighborhood search. Finally, our proposal is embedded in a Hierarchical fast multipoint scheme, where overlapping between patches is reduced allowing to use fewer points to label the whole image which leads to reducing the computational time significantly. Attained results in the segmentation of subcortical structures show that our proposed approach outperforms both probabilistic atlas and patch-based segmentation methods.

5.2 Future work

There is, many issues that can be addressed to improve the above proposed approaches:

- Specifically, for the computing embedding representation space, Since obtained decomposition eigenvectors showed non-linear relationships, other non-linear embedding techniques e.g Laplacian eigenmaps and local linear embedding can be used to highlight the essential structure. Supervised decomposition techniques will be provet to find representation to distinguish other categories such as ethnicity or pathology subclasses. Also, new feature extraction methodologies based on mixing inter slide kernel(ISK) along of three main axes can be explored aiming to encode all dynamics into shorter and more compact versions of Images.

- New similarity metric combination approaches can be considered for the introduce supervised framework, as they can take advantage of different relationships between image partitions. For instance, a gain factor for each weight can be used to span the combination weights distribution. Also, new experiments for evaluating the metric for different demographic categories as age, gender, and neurodegenerative disease, as well as, presence of tumors will be performed. This is because every morphological change will be reflected into the measure performance.

- Several directions can be taken into account to improve the proposed target specific patch atlas segmentation method. spatial priors which related information of the cliques could improve the appearance model construction. Also,Markov model estimation could be introduced to account for spatial relationships. On the other hand, patch search can be improved by adaptatively changing the patch and neighborhood radius using a variability criterion.

6 Publications

This thesis has lead to six works presented in international conferences and one participation in a segmentation challenge. Furthermore, a manuscript have been sent to peer-reviewed Medical Image Analysis journal .

- **Orbes-Arteaga**, M., Cardenas-Pena, Castellanos-Dominguez, G. Head and Neck Auto Segmentation Challenge based on Non-Local Generative Models. Head and Neck Auto Segmentation MICCAI challenge workshop(2015).

- **Orbes-Arteaga**, M., Cardenas-Pena, D., Alvarez, M. A., Orozco, A. A., Castellanos-Dominguez, G. Kernel Centered Alignment Supervised Metric for Multi-Atlas Segmentation. ICIAP 2015. ***Best Young Paper Awards Finalist***

- **Orbes-Arteaga**, M., Cardenas-Pena, D., Alvarez, M. A., Orozco, A. A., Castellanos-Dominguez, G. Magnetic Resonance Image Selection for Multi-Atlas Segmentation Using Mixture Models. CIARP 2015.

- **Orbes-Arteaga**, M., Cardenas-Pena, D., Alvarez, M. A., Orozco, A. A., Castellanos-Dominguez, . Spatial-Dependent Similarity Metric Supporting Multi-atlas MRI Segmentation. IbPRIA 2015.

- Cardenas-Pena, D., **Orbes-Arteaga**, M., Castellanos-Dominguez, G. (2015). Supervised brain tissue segmentation using a spatially enhanced similarity metric. IWINAC 2015.

- Cardenas-Pena, D., **Orbes-Arteaga**, M., Castellanos-Dominguez, G. (2014, August). Kernel-based Atlas Image Selection for brain tissue segmentation. In Engineering in Medicine and Biology Society (EMBC) International Conference of the IEEE.

- Cardenas-Pena, D., **Orbes-Arteaga**, M., Castro-Ospina, A., Alvarez-Meza, A., Castellanos-Dominguez, G. (2014, August). A kernel-based representation to support 3d mri unsupervised clustering. In 2014 22nd International Conference on Pattern Recognition (ICPR) IEEE.

Bibliography

[1] Adams, Rolf ; Bischof, Leanne: Seeded region growing. En: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16 (1994), Nr. 6, p. 641–647. – ISBN 0162–8828

[2] Alirezaie, J ; Jernigan, M E. ; Nahmias, C: Neural network based segmentation of magnetic resonance images of the brain. En: *1995 IEEE Nuclear Science Symposium and Medical Imaging Conference Record* 3 (1995), p. 1397–1401

[3] Aljabar, Paul ; Heckemann, R ; Hammers, Alexander ; Hajnal, Joseph V. ; Rueckert, Daniel: Classifier selection strategies for label fusion using large atlas databases. En: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*. Springer, 2007, p. 523–531

[4] Aljabar, Paul ; Heckemann, Rolf A. ; Hammers, Alexander ; Hajnal, Joseph V. ; Rueckert, Daniel: Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. En: *Neuroimage* 46 (2009), Nr. 3, p. 726–738

[5] Artaechevarria, Xabier ; Muñoz-Barrutia, Arrate ; Ortiz-de Solorzano, Carlos: Efficient classifier generation and weighted voting for atlas-based segmentation: Two small steps faster and closer to the combination oracle. En: *Medical Imaging* International Society for Optics and Photonics, 2008, p. 69141W–69141W

[6] Artaechevarria, Xabier ; Munoz-Barrutia, Arrate ; Ortiz-de Solórzano, Carlos: Combination strategies in multi-atlas image segmentation: Application to brain MR data. En: *Medical Imaging, IEEE Transactions on* 28 (2009), Nr. 8, p. 1266–1277

[7] Ashburner, John ; Friston, Karl J.: Unified segmentation. En: *Neuroimage* 26 (2005), Nr. 3, p. 839–851

[8] Asman, Andrew J. ; Landman, Bennett [u. a.]: Formulating spatially varying performance in the statistical fusion framework. En: *Medical Imaging, IEEE Transactions on* 31 (2012), Nr. 6, p. 1326–1336

[9] Asman, Andrew J. ; Landman, Bennett A.: Robust statistical label fusion through

consensus level, labeler accuracy, and truth estimation (COLLATE). En: *Medical Imaging, IEEE Transactions on* 30 (2011), Nr. 10, p. 1779–1794

[10] Asman, Andrew J. ; Landman, Bennett a.: Non-local statistical label fusion for multi-atlas segmentation. En: *Medical image analysis* 17 (2013), Februar, Nr. 2, p. 194–208. – ISSN 1361–8423

[11] Asman, Andrew J. ; Landman, Bennett A.: Non-local statistical label fusion for multi-atlas segmentation. En: *Medical image analysis* 17 (2013), Nr. 2, p. 194–208

[12] Asman, Andrew J. ; Landman, Bennett A.: Hierarchical performance estimation in the statistical label fusion framework. En: *Medical image analysis* 18 (2014), Nr. 7, p. 1070–1081

[13] Avants, Brian B. ; Epstein, Charles L. ; Grossman, Murray ; Gee, James C.: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. En: *Medical image analysis* 12 (2008), Nr. 1, p. 26–41

[14] Bajcsy, Ruzena ; Lieberson, Robert ; Reivich, Martin: A computerized system for the elastic matching of deformed radiographic images to idealized atlas images. En: *Journal of computer assisted tomography* 7 (1983), Nr. 4, p. 618–625

[15] Balafar, M.A. ; ramli, A.R. ; saripan, M.I. ; mahmud, R. ; mashohor, S.: Medical image segmentation using fuzzy C-mean (FCM) and dominant grey levels of image. En: *IET Conference Proceedings* (2008), January, p. 314–317(3)

[16] Bezdek, J. C. ; Hall, L. O. ; Clarke, L. P.: Review of MR image segmentation techniques using pattern recognition. En: *Medical Physics* 20 (1993), Nr. 4, p. 1033–1048

[17] Biswal, M. ; Shah, N. ; Newark, S. ; Trivedi, M. ; Nayak, S. ; Dave, H.: Automatic segmentation of pancreatic images using ISODATA alorithm. En: *Journal of Clinical Oncology, 2006* 24 (2006), June, p. 14149

[18] Blezek, Daniel J. ; Miller, James V.: Atlas stratification. En: *Medical image analysis* 11 (2007), Nr. 5, p. 443–457

[19] Cabezas, Mariano ; Oliver, Arnau ; Lladó, Xavier ; Freixenet, Jordi ; Cuadra, Meritxell B.: A review of atlas-based segmentation for magnetic resonance brain images. En: *Computer methods and programs in biomedicine* 104 (2011), Dezember, Nr. 3, p. e158–77. – ISSN 1872–7565

[20] Cao, Yihui ; Yuan, Yuan ; Li, Xuelong ; Turkbey, Baris ; Choyke, Peter L. ; Yan, Pingkun: Segmenting images by combining selected atlases on manifold. En: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*. Springer, 2011, p. 272–279

[21] Cao, Yihui ; Yuan, Yuan ; Li, Xuelong ; Yan, Pingkun: Putting images on a manifold for atlas-based image segmentation. En: *Image Processing (ICIP), 2011 18th IEEE International Conference on* IEEE, 2011, p. 289–292

[22] Chan, Tony F. ; Golub, Gene H. ; LeVeque, Randall J.: Algorithms for computing the sample variance: Analysis and recommendations. En: *The American Statistician* 37 (1983), Nr. 3, p. 242–247

[23] Chen, A ; Niermann, KJ ; Deeley, MA ; Dawant, BM: Evaluation of multiple-atlas-based strategies for segmentation of the thyroid gland in head and neck ct images for imrt. En: *Physics in medicine and biology* 57 (2012), Nr. 1, p. 93

[24] Christensen, Gary E. ; Rabbitt, Richard D. ; Miller, Michael [u. a.]: Deformable templates using large deformation kinematics. En: *Image Processing, IEEE Transactions on* 5 (1996), Nr. 10, p. 1435–1447

[25] Chu, Chengwen ; Oda, Masahiro ; Kitasaka, Takayuki ; Misawa, Kazunari ; Fujiwara, Michitaka ; Hayashi, Yuichiro ; Nimura, Yukitaka ; Rueckert, Daniel ; Mori, Kensaku: Multi-organ segmentation based on spatially-divided probabilistic atlas from 3D abdominal CT images. En: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013* (2013), p. 165–172

[26] Ciofolo, Cybèle ; Barillot, Christian: Atlas-based segmentation of 3D cerebral structures with competitive level sets and fuzzy control. En: *Medical Image Analysis* 13 (2009), Nr. 3, p. 456–470

[27] Collins, D L. ; Holmes, Colin J. ; Peters, Terrence M. ; Evans, Alan C.: Automatic 3-D model-based neuroanatomical segmentation. En: *Human brain mapping* 3 (1995), Nr. 3, p. 190–208

[28] Commowick, Olivier ; Akhondi-Asl, Alireza ; Warfield, Simon K.: Estimating a reference standard segmentation with spatially varying performance parameters: Local MAP STAPLE. En: *Medical Imaging, IEEE Transactions on* 31 (2012), Nr. 8, p. 1593–1606

[29] Commowick, Olivier ; Malandain, Grégoire: Efficient selection of the most similar image in a database for critical structures segmentation. En: *Medical Image Computing*

and Computer-Assisted Intervention–MICCAI 2007. Springer, 2007, p. 203–210

[30] Commowick, Olivier ; Warfield, Simon K.: Incorporating priors on expert performance parameters for segmentation validation and label fusion: a maximum a posteriori STAPLE. En: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*. Springer, 2010, p. 25–32

[31] Cootes, T.F. ; Edwards, G.J. ; Taylor, C.J.: Active Appearance Models. En: *Proc. European Conference on Computer Vision (ICCV)* 2 (1998), p. 484–498. – ISBN 3540646132

[32] Cortes, Corinna ; Mohri, Mehryar ; Rostamizadeh, Afshin: Algorithms for learning kernels based on centered alignment. En: *The Journal of Machine Learning Research* 13 (2012), Nr. 1, p. 795–828

[33] Coupé, Pierrick ; Manjón, José V ; Fonov, Vladimir ; Pruessner, Jens ; Robles, Montserrat ; Collins, D L.: Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. En: *NeuroImage* 54 (2011), Nr. 2, p. 940–954

[34] Davatzikos, Christos: Spatial transformation and registration of brain images using elastically deformable models. En: *Computer Vision and Image Understanding* 66 (1997), Nr. 2, p. 207–222

[35] Dawant, Benoit M. ; Thirion, Jean-Philippe ; Maes, Frederik ; Vandermeulen, Dirk ; Demaerel, Philippe: Automatic 3D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations. En: *Medical Imaging'98* International Society for Optics and Photonics, 1998, p. 545–554

[36] Duc, Albert K H. ; Modat, Marc ; Leung, Kelvin K. ; Cardoso, M J. ; Barnes, Josephine ; Kadir, Timor ; Ourselin, Sébastien ; Initiative, ADN [u. a.]: Using manifold learning for atlas selection in multi-atlas segmentation. En: *PloS one* 8 (2013), Nr. 8, p. e70059

[37] Evans, A.C. ; Collins, D.L. ; Mills, S. R. ; Brown, E. D. ; Kelly, R. L. ; Peters, T.M.: 3D statistical neuroanatomical models from 305 MRI volumes. En: *Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record.*, 1993. – ISBN 0780314875, p. 1813–1817

[38] Filippone, Maurizio ; Camastra, Francesco ; Masulli, Francesco ; Rovetta, Stefano: A survey of kernel and spectral methods for clustering. En: *Pattern recognition* 41 (2008), Nr. 1, p. 176–190

[39] Fischl, Bruce ; Salat, David H. ; Busa, Evelina ; Albert, Marilyn ; Dieterich, Megan ; Haselgrove, Christian ; Van Der Kouwe, Andre ; Killiany, Ron ; Kennedy, David ; Klaveness, Shuna [u. a.]: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. En: *Neuron* 33 (2002), Nr. 3, p. 341–355

[40] Gee, Jim C. ; Reivich, Martin ; Bajcsy, Ruzena: Elastically deforming 3D atlas to match anatomical brain images. En: *Journal of computer assisted tomography* 17 (1993), Nr. 2, p. 225–236

[41] Hao, Yongfu ; Wang, Tianyao ; Zhang, Xinqing ; Duan, Yunyun ; Yu, Chunshui ; Jiang, Tianzi ; Fan, Yong: Local label learning (LLL) for subcortical structure segmentation: Application to hippocampus segmentation. En: *Human brain mapping* 35 (2014), Nr. 6, p. 2674–2697

[42] Heckemann, Rolf A. ; Hajnal, Joseph V. ; Aljabar, Paul ; Rueckert, Daniel ; Hammers, Alexander: Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. En: *NeuroImage* 33 (2006), Nr. 1, p. 115–126

[43] Iglesias, Juan E. ; Sabuncu, Mert R.: Multi-Atlas Segmentation of Biomedical Images: A Survey. En: *arXiv preprint arXiv:1412.3421* (2014)

[44] Iosifescu, Dan V. ; Shenton, Martha E. ; Warfield, Simon K. ; Kikinis, Ron ; Dengler, Joachim ; Jolesz, Ferenc A. ; McCarley, Robert W.: An automated registration algorithm for measuring MRI subcortical brain structures. En: *Neuroimage* 6 (1997), Nr. 1, p. 13–25

[45] Kamber, Micheline ; Shinghal, Rajjan ; Collins, D L. ; Francis, Gordon S. ; Evans, Alan C.: Model-based 3-D segmentation of multiple sclerosis lesions in magnetic resonance brain images. En: *Medical Imaging, IEEE Transactions on* 14 (1995), Nr. 3, p. 442–453

[46] Klein, Arno ; Andersson, Jesper ; Ardekani, Babak A. ; Ashburner, John ; Avants, Brian ; Chiang, Ming-Chang ; Christensen, Gary E. ; Collins, D L. ; Gee, James ; Hellier, Pierre [u. a.]: Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. En: *Neuroimage* 46 (2009), Nr. 3, p. 786–802

[47] Klein, Arno ; Hirsch, Joy: Mindboggle: a scatterbrained approach to automate brain labeling. En: *NeuroImage* 24 (2005), Nr. 2, p. 261–280

[48] Langerak, Thomas R. ; Berendsen, Floris F. ; Van der Heide, Uulke A. ; Kotte,

Alexis N. ; Pluim, Josien P.: Multiatlas-based segmentation with preregistration atlas selection. En: *Medical physics* 40 (2013), Nr. 9, p. 091701

[49] Langerak, Thomas R. ; Van Der Heide, Uulke ; Kotte, Alexis N. ; Viergever, Max ; Van Vulpen, Marco ; Pluim, Josien P. [u. a.]: Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE). En: *Medical Imaging, IEEE Transactions on* 29 (2010), Nr. 12, p. 2000–2008

[50] Lao, Zhiqiang ; Shen, Dinggang ; Xue, Zhong ; Karacali, Bilge ; Resnick, Susan M. ; Davatzikos, Christos: Morphological classification of brains via high-dimensional shape transformations and machine learning methods. En: *Neuroimage* 21 (2004), Nr. 1, p. 46–57

[51] Van der Lijn, Fedde ; De Bruijne, Marleen ; Klein, Stefan ; Heijer, Tom D. ; Hoogendam, Yoo Y. ; Van der Lugt, Aad ; Breteler, Monique ; Niessen, Wiro J.: Automated brain structure segmentation based on atlas registration and appearance models. En: *Medical Imaging, IEEE Transactions on* 31 (2012), Nr. 2, p. 276–286

[52] van der Lijn, Fedde ; den Heijer, Tom ; Breteler, Monique M. ; Niessen, Wiro J.: Hippocampus segmentation in MR images using atlas registration, voxel classification, and graph cuts. En: *Neuroimage* 43 (2008), Nr. 4, p. 708–720

[53] Lötjönen, Jyrki M. ; Wolz, Robin ; Koikkalainen, Juha R. ; Thurfjell, Lennart ; Waldemar, Gunhild ; Soininen, Hilkka ; Rueckert, Daniel: Fast and robust multi-atlas segmentation of brain magnetic resonance images. En: *NeuroImage* 49 (2010), Februar, Nr. 3, p. 2352–65. – ISSN 1095–9572

[54] Ma, Zhen ; Tavares, Joao Manuel R S. ; Jorge, Renato N. ; Mascarenhas, T: *A review of algorithms for medical image segmentation and their applications to the female pelvic cavity.* 2010. – 235–46 p.. – ISBN 1025584090

[55] Miller, Michael I. ; Christensen, Gary E. ; Amit, Yali ; Grenander, Ulf: Mathematical textbook of deformable neuroanatomies. En: *Proceedings of the National Academy of Sciences* 90 (1993), Nr. 24, p. 11944–11948

[56] Ng, Andrew Y. ; Jordan, Michael I. ; Weiss, Yair [u. a.]: On spectral clustering: Analysis and an algorithm. En: *Advances in neural information processing systems* 2 (2002), p. 849–856

[57] Nouranian, Saman ; Mahdavi, S S. ; Spadinger, Ingrid ; Morris, William J. ;

Salcudean, Septimu E. ; Abolmaesumi, Purang: A Multi-Atlas-Based Segmentation Framework for Prostate Brachytherapy. En: *Medical Imaging, IEEE Transactions on* 34 (2015), Nr. 4, p. 950–961

[58] Oda, Masahiro ; Nakaoka, Teruhisa ; Kitasaka, Takayuki ; Furukawa, Kazuhiro ; Misawa, Kazunari ; Fujiwara, Michitaka ; Mori, Kensaku: Organ segmentation from 3D abdominal CT images based on atlas selection and graph cut. En: *Abdominal Imaging. Computational and Clinical Applications*. Springer, 2011, p. 181–188

[59] Park, Hyunjin ; Bland, Peyton H. ; Meyer, Charles R.: Construction of an abdominal probabilistic atlas and its application in segmentation. En: *Medical Imaging, IEEE Transactions on* 22 (2003), Nr. 4, p. 483–492

[60] Pham, Dzung L. ; Xu, Chenyang ; Prince, Jerry L.: Current methods in medical image segmentation 1. En: *Annual review of biomedical engineering* 2 (2000), Nr. 1, p. 315–337

[61] Ramus, Liliane ; Commowick, Olivier ; Malandain, Grégoire: Construction of patient specific atlases from locally most similar anatomical pieces. En: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*. Springer, 2010, p. 155–162

[62] van Rikxoort, Eva M. ; Isgum, Ivana ; Arzhaeva, Yulia ; Staring, Marius ; Klein, Stefan ; Viergever, Max A. ; Pluim, Josien P. ; van Ginneken, Bram: Adaptive local multi-atlas segmentation: application to the heart and the caudate nucleus. En: *Medical image analysis* 14 (2010), Nr. 1, p. 39–49

[63] Rohlfing, Torsten ; Brandt, Robert ; Menzel, Randolf ; Maurer, Calvin R.: Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. En: *NeuroImage* 21 (2004), Nr. 4, p. 1428–1442

[64] Rohlfing, Torsten ; Brandt, Robert ; Menzel, Randolf ; Russakoff, Daniel B. ; Maurer, Calvin R.: Quo Vadis , Atlas-Based Segmentation ? En: *The Handbook of Medical Image Analysis: Registration Models*. 2005. – ISBN 978–0–306–48607–4, p. 435–486

[65] Rohlfing, Torsten ; Maurer Jr, Calvin R.: Shape-based averaging. En: *Image Processing, IEEE Transactions on* 16 (2007), Nr. 1, p. 153–161

[66] Rousseau, Frederic ; Habas, Piotr A. ; Studholme, Colin: A supervised patch-based approach for human brain labeling. En: *Medical Imaging, IEEE Transactions on* 30

(2011), Nr. 10, p. 1852–1862

[67] Sabuncu, Mert R. ; Yeo, BT T. ; Van Leemput, Koen ; Fischl, Bruce ; Golland, Polina: Supervised nonparametric image parcellation. En: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*. Springer, 2009, p. 1075–1083

[68] Sanroma, Gerard ; Wu, Guorong ; Gao, Yaozong ; Shen, Dinggang: Learning to Rank Atlases for Multiple-Atlas Segmentation. En: *Medical Imaging, IEEE Transactions on* 33 (2014), Nr. 10, p. 1939–1953

[69] Sanroma, Gerard ; Wu, Guorong ; Gao, Yaozong ; Thung, Kim-Han ; Guo, Yanrong ; Shen, Dinggang: A transversal approach for patch-based label fusion via matrix completion. En: *Medical image analysis* 24 (2015), Nr. 1, p. 135–148

[70] Sanroma, Gerard ; Wu, Guorong ; Thung, Kim ; Guo, Yanrong ; Shen, Dinggang: Novel Multi-Atlas Segmentation by Matrix Completion. En: *Machine Learning in Medical Imaging*. Springer, 2014, p. 207–214

[71] Tong, Tong ; Wolz, Robin ; Hajnal, Joseph V. ; Rueckert, Daniel: Segmentation of brain MR images via sparse patch representation. En: *MICCAI Workshop on Sparsity Techniques in Medical Imaging (STMI)*, 2012

[72] Valdés-Hernández, Pedro a. ; von Ellenrieder, Nicolás ; Ojeda-Gonzalez, Alejandro ; Kochen, Silvia ; Alemán-Gómez, Yasser ; Muravchik, Carlos ; Valdés-Sosa, Pedro a.: Approximate average head models for EEG source imaging. En: *Journal of neuroscience methods* 185 (2009), Dezember, Nr. 1, p. 125–32. – ISSN 1872–678X

[73] Van Leemput, Koen ; Maes, Frederik ; Vandermeulen, Dirk ; Suetens, Paul: Automated model-based tissue classification of MR images of the brain. En: *Medical Imaging, IEEE Transactions on* 18 (1999), Nr. 10, p. 897–908

[74] Von Luxburg, Ulrike: A tutorial on spectral clustering. En: *Statistics and computing* 17 (2007), Nr. 4, p. 395–416

[75] Wachinger, Christian ; Brennan, Matthew ; Sharp, Greg C. ; Golland, Polina: On the importance of location and features for the patch-based segmentation of parotid glands. En: *MICCAI Workshop on Image-Guided Adaptive Radiation Therapy*, 2014

[76] Wang, Hongzhi ; Suh, Jung W. ; Das, Sunil R. ; Pluta, John B. ; Craige, Caryne ; Yushkevich, Paul [u. a.]: Multi-atlas segmentation with joint label fusion. En: *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 35 (2013), Nr. 3, p. 611–623

[77] Wang, Hongzhi ; Suh, Jung W. ; Pluta, John ; Altinay, Murat ; Yushkevich, Paul: Optimal weights for multi-atlas label fusion. En: *Information Processing in Medical Imaging* Springer, 2011, p. 73–84

[78] Wang, Jiahui ; Vachet, Clement ; Rumple, Ashley ; Gouttard, Sylvain ; Ouziel, Clémentine ; Perrot, Emilie ; Du, Guangwei ; Huang, Xuemei ; Gerig, Guido ; Styner, Martin: Multi-atlas segmentation of subcortical brain structures via the AutoSeg software pipeline. En: *Frontiers in neuroinformatics* 8 (2014)

[79] Wang, Song ; Zhu, Weiyu ; Liang, Zhi-pei: Shape deformation: SVM regression and application to medical image segmentation. En: *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001* 2 (2001), p. 209–216. ISBN 0–7695–1143–0

[80] Warfield, Simon K. ; Zou, Kelly H. ; Wells, William M.: Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. En: *Medical Imaging, IEEE Transactions on* 23 (2004), Nr. 7, p. 903–921

[81] Weisenfeld, Neil ; Warfield, Simon K. [u. a.]: SoftSTAPLE: Truth and performance-level estimation from probabilistic segmentations. En: *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on* IEEE, 2011, p. 441–446

[82] Wolz, Robin ; Aljabar, Paul ; Hajnal, Joseph V. ; Hammers, Alexander ; Rueckert, Daniel ; Initiative, Alzheimer's Disease N. [u. a.]: LEAP: learning embeddings for atlas propagation. En: *NeuroImage* 49 (2010), Nr. 2, p. 1316–1325

[83] Wu, Guorong ; Kim, Minjeong ; Sanroma, Gerard ; Wang, Qian ; Munsell, Brent C. ; Shen, Dinggang ; Initiative, Alzheimer's Disease N. [u. a.]: Hierarchical multi-atlas label fusion with multi-scale feature representation and label-specific patch partition. En: *NeuroImage* 106 (2015), p. 34–46

[84] Wu, Minjie ; Rosano, Caterina ; Lopez-Garcia, Pilar ; Carter, Cameron S. ; Aizenstein, Howard J.: Optimum template selection for atlas-based segmentation. En: *Neuroimage* 34 (2007), Nr. 4, p. 1612–1618

[85] Wu, Minjie ; Rosano, Caterina ; Lopez-Garcia, Pilar ; Carter, Cameron S. ; Aizenstein, Howard J.: Optimum template selection for atlas-based segmentation. En: *Neuroimage* 34 (2007), Nr. 4, p. 1612–1618

[86] Xie, Qiuliang ; Ruan, Dan: Low-complexity atlas-based prostate segmentation by combining global, regional, and local metrics. En: *Medical physics* 41 (2014), Nr. 4, p.

041909

[87] Zhang, Daoqiang ; Guo, Qimiao ; Wu, Guorong ; Shen, Dinggang: Sparse patch-based label fusion for multi-atlas segmentation. En: *Multimodal Brain Image Analysis*. Springer, 2012, p. 94–102

[88] Zhang, Yongyue ; Brady, Michael ; Smith, Stephen: Segmentation of Brain MR Images Through a Hidden Markov Random Field Model and the Expectation-Maximization Algorithm. (2001), Nr. January, p. 45–57