UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE MATHÉMATIQUES, SCIENCES DE L'INFORMATION ET DE L'INGÉNIEUR

Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection

no. d'ordre: 1541

THÈSE présentée par :

Giorgos SFIKAS

soutenue le : **7 septembre 2012**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : Traitement du signal et des images

Modèles statistiques non linéaires pour l'analyse de formes. Application à l'imagerie cérébrale.

THÈSE dirigée par :

M. Christian HEINRICH	Professeur des Universités, Université de Strasbourg
M. Christophoros NIKOU	Professeur Assistant, Université de Ioannina

RAPPORTEURS :

M. Jean-Philippe RANJEVA	Professeur des Universités, Aix-Marseille Université
M. Frédéric RICHARD	Professeur des Universités, Aix-Marseille Université

AUTRES MEMBRES DU JURY :

M. Jack FOUCHER	Maître de conférences, praticien hospitalier, HUS
M. Frédéric BLANC	Maître de conférences, praticien hospitalier, HUS
M. Aristidis LIKAS	Professeur Associé, Université de Ioannina
M. Pierre CHARBONNIER	Directeur de recherche, LRPC, Strasbourg

Note

The current text is a version of the complete manuscript after all elements copied or adapted from third-party publications, for which the author did not have a licence for public reproduction, have been removed ("version de diffusion").

Acknowledgements

The following manuscript stands for the representative output of a process that has begun almost four years ago, when I started pursuing my PhD title. While it is the fruit of personal work, ideas –be it that they are original or unoriginal, scientific or unscientific– do not form on one's mind by themselves. I believe, however cliché that may sound for some, that on the one hand ideas, and in general the imaginary, stands in a firm, two-ways, dialectical relation with the objective reality of one's time (a reality that is indeed, in every aspect, changing very rapidly lately all around the world). On the other hand, one's ideas are born, shaped or challenged when in contact –where contact may mean anything from a mundane discussion to a university lecture– with other people, carriers of like or different thoughts and ideas. As such, from another point of view, the present work is the product of many (only I get the credit though !). And it is those people, living factors, that have contributed directly or indirectly, towards its completion, that I feel obliged to thank here.

My thesis advisors, Christian Heinrich and Christophoros Nikou. I thank them both for their guidance and support, all through the time I spent working on my thesis. Christian Heinrich for proposing me an interesting PhD subject, and also importantly, finding me financial support. From the very first moment my thesis started in France, Christian Heinrich has been a steady, patient and endurant source of aid of scientific, as well as more mundane, nature. Christophoros Nikou for his hard-to-find availability and helpfulness in his supervision, despite the long distance that separates Alsace and Epirus.

My heartfelt gratitude to the late Dr. N.P.Galatsanos, who supervised my research mostly prior to the beginning of my PhD. An outstanding scientist, teacher and researcher, he has been a true source of inspiration to all that surrounded him. Especially personally I feel that I owe him very much, within as well as far beyond the context of this thesis.

Jean-Philippe Ranjeva, Frédéric Richard, who have undertaken the task of judging this manuscript ; medical doctors Frédéric Blanc, Nathalie Philippi and Jack Foucher, with whom we have been or are on an ongoing collaboration ; Aristidis Likas for accepting to be member of the jury, despite his busy work schedule in Ioannina.

My collegues and friends in LSIIT, IPB and elsewhere, who I have frequented more or less on a fixed basis. Félix Renard, Slim Karkar, Ahmed Mumtaz, Hanna Eriksson, Erhan Aptoula, Swati Sharma, Shine On, Georges Moustaki, Benjamin Perret, Fatma Abdelmoullah, Akram Belghith, Farouk Nouizi, Abdelqader Lakhsassi, René Albino Perez. All the, admittedly small, greek community of Strasbourg. Especially Evangelos Gintersos, Frans Declerq, Théodoros "Théo" Chiou, Angelina Greka, Anargyros Skaltsas, Billy "Gavrakos" Grec and Paul Kitromèlidès for all the memorables soirées grecques we organised in and out of Café des Anges, and our efforts in spreading 9/8 in Alsace.

In parallel to the long process of pursuing my PhD thesis, I feel I should mention here that I have been part of an equally long and interesting process, that of learning the French language. So thanks are in order here to all the francophones of Strasbourg and elsewhere, who helped in this (ongoing) task, either conciously or not. Learning French has frankly opened up new horizons of thought and experience for me, perhaps as much as the thesis in itself. Perhaps

unexpectedly, it has even resulted in me getting to know my own mother language better.

Finally, last but by no means least, I wish to especially thank my family for their unconditional love and support throughout all these years.

Table of contents

List of abbreviations

AD	Alzheimer's Disease
ADAS	Alzheimer's Disease Assessment Scale
BDE	Boundary Displacement Error
CA	Cornus Ammonis
CDR	Clinical Dementia Rating
CSF	Cerebrospinal Fluid
DBM	Deformation-Based Morphometry
DRS	Dementia Rating Scale
DT-MRI	Diffusion Tensor Magnetic Resonance Imaging
EM	Expectation-Maximization
FMM	Finite Mixture Model
fMRI	Functional Magnetic Resonance Imaging
GCE	Global Consistency Error
GDS	Global Deterioration Stage
GM	Gray Matter
GMM	Gaussian Mixture Model
HRSD	Hamilton Rating Scale for Depression
ISNR	Improvement over Signal-to-Noise Ratio
LDDMM	Large Deformation Diffeomorphic Metric Mapping
m-reps	Medial Representations
MAP	Maximum A Posteriori
MCI	Mild Cognitive Impairment
MCMC	Markov Chain Monte Carlo
MDS	MultiDimensional Scaling
ML	Maximum Likelihood
MM	Majorization-Minimization
MMSE	Mini-Mental State Exam
MRF	Markov Random Field
MRI	Magnetic Resonance Imaging
ncuts	Normalized Cuts
OCSVM	One-Class Support Vector Machine
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PSA	Polarization State Analyzer
pdf	Probability Distribution Function
ROI	Region Of Interest
RI	Rand Index
RVM	Relevance Vector Machine
RBF	Radial Basis Function

SPHARM	Spherical Harmonics
SNR	Signal-to-Noise Ratio
SVDD	Support Vector Data Description
SVGMM	Spatially-Varying Gaussian Mixture Model
St-SVGMM	Student's-t Spatially Varying Gaussian Mixture Model
SVM	Support Vector Machine
TBM	Tensor-Based Morphometry
WM	White Matter
VOI	Variation Of Information
VI	Variational Inference

Abstract

This thesis addresses statistical shape analysis, in the context of medical imaging. In the field of medical imaging, shape analysis is used to describe the morphological variability of various organs and tissues, such as the heart, the brain, or particular structures and tissues in the brain. Pathologies, whose development is characterized by structural change in brain tissue, include Alzheimer's disease, which is the pathology of interest in this thesis, multiple sclerosis and schizophrenia. Structural change in these diseases reflects the evolution of neuronal degeneration and brain atrophy. The qualitative and quantitative characteristics of such structural degeneration can help us assess the development of the pathology, and eventually plan a disease-slowing or symptom prevention treatment.

Our focus in this thesis is on the construction of a generative and discriminative, compact and non-linear model, suitable to the representation of shapes. This model is evaluated in the context of the study of a population of Alzheimer's disease patients and a population of healthy controls. The discriminative component of the model is based on the *Support Vector Data Description* (SVDD) framework. The SVDD is closely related to the *one-class Support Vector Machine* (SVM). Our principal interest here is using the discriminative model to discover morphological differences that are the most characteristic and discriminate best between a given shape class and forms not belonging in that class.

An important aspect of the proposed model is the hypothesis that the data (the shapes) lie on, or sufficiently near, a low-dimensional, non-linear manifold. In this work we originally parametrize shape as a deformation field with respect to a given reference image, a representation which gives a high quality, detailed description, but also a description space of very high dimensionality, in the order of millions of dimensions. After dimensionality reduction, the shape can be represented as a point of a space comprising typically no more than ten dimensions.

The theoretical innovation of our work lies in two principal points : first, we propose a tool to extract discriminative difference in the context of SVDD models ; second, all generated reconstructions are anatomically correct, due to the non-linear and compact character of the model. The application of our model on real medical data shows results coherent with well-known findings in related research. We confirm that particular tissue regions and subfields are damaged the most, like the head and the tail of the hippocampus. We also use the model as a classifier, discriminating healthy from pathological subjects on the basis of structural information.

Résumé

Cette thèse a pour objet l'analyse statistique de formes, dans le contexte de l'imagerie médicale. Dans le champ de l'imagerie médicale, l'analyse de formes est utilisée pour décrire la variabilité morphologique de divers organes et tissus, tels que le cœur, le cerveau, ou des structures et tissus particuliers dans le cerveau. Les pathologies d'intérêt, dont le développement est marqué par des changements structurels dans les tissus cérébraux, incluent la maladie d'Alzheimer, qui est la pathologie d'intérêt de cette thèse, la sclérose en plaques et la schizophrénie. Le changement structurel dans le contexte de ces maladies reflète l'évolution de la dégénération neuronale ou de l'atrophie du cerveau. Les caractéristiques qualitatives et quantitatives de ces dégénérations structurelles peuvent nous aider à évaluer le développement de la pathologie, et de planifier éventuellement un traitement pour ralentir l'évolution de la maladie.

Nous nous focalisons dans cette thèse sur la construction d'un modèle génératif et discriminatif, compact et non-linéaire, adapté à la représentation de formes. Ce modèle est évalué dans le contexte de l'étude d'une population de patients atteints de la maladie d'Alzheimer et d'une population de sujets contrôles sains. La composante discriminative du modèle est fondée sur le cadre *Support Vector Data Description* (SVDD). Le SVDD est étroitement lié au *one-class Support Vector Machine* (SVM). Notre intérêt principal ici est l'utilisation du modèle discriminatif pour découvrir les différences morphologiques les plus discriminatives entre une classe de formes donnée et des formes n'appartenant pas à cette classe.

Un aspect important du modèle proposé est l'hypothèse que les données (les formes) se trouvent sur ou non loin d'une variété non-linéaire de dimension faible. La forme, dans ce travail, est paramétrée originellement comme un champ de déformation par rapport à une image de référence, ce qui donne une description très fine et de haute qualité, mais aussi un espace de description de très haute dimensionalité, de l'ordre de millions de dimensions. Après réduction de dimension, la forme peut se décrire comme un point d'un espace de moins de dix dimensions.

L'innovation théorique apportée par notre modèle réside en deux points principaux : premièrement, nous proposons un outil pour extraire la différence discriminative dans le contexte des modèles SVDD ; deuxièmement, toutes les reconstructions générées sont anatomiquement correctes, ce qui est dû au caractère non-linéaire et compact du modèle. Une application de notre modèle à des données médicales réelles montre des résultats cohérents avec les connaissances médicales. Nous mettons en évidence des points significatifs autour de la tête et de la queue des hippocampes gauche et droit. Nous utilisons également le modèle comme classifieur, permettant de discriminer les sujets sains des sujets pathologiques à partir toujours de données structurelles des régions d'intérêt du cerveau.

Περίληψη

Το αντικείμενο αυτής της διατριβής είναι η στατιστική ανάλυση μορφών, στα πλαίσια της επεξεργασίας ιατρικής εικόνας. Στο πεδίο της επεξεργασίας ιατρικής εικόνας, η ανάλυση μορφών χρησιμοποιείται στο στόχο της περιγραφής της μορφολογικής ποικιλότητας διάφορων οργάνων και ιστών, όπως η καρδιά, ο εγκέφαλος, ή συγκεκριμένες δομές στο εσωτερικό του εγκεφάλου. Συγκεκριμένες ασθένειες, των οποίων η ανάπτυξη χαρακτηρίζεται από σημαντικές δομικές αλλαγές στον εγκεφαλικό ιστό, συμπεριλαμβάνουν την νόσο Αλτσχάιμερ –η οποία αποτελεί και τη μελέτη περίπτωσης στην παρούσα διατριβή– τη σκλήρυνση κατά πλάκας και τη σχιζοφρένεια. Οι δομικές αλλαγές σε αυτές τις ασθένειες αντικατοπτρίζουν την πρόοδο της νευροεκφύλισης και γενικότερα της εγκεφαλικής ατροφίας. Η γνώση των ποιοτικών και ποσοτικών χαρακτηριστικών του δομικού εκφυλισμού μπορούν να βοηθήσουν στην εκτίμηση του τρέχοντος σταδίου και εξέλιξης της ασθένειας, στα πλαίσια ή μη μιας αγωγής επιβράδυνσης της νόσου και των συμπτωμάτων της.

Ο στόχος της παρούσας διατριβής είναι η κατασκευή ενός παραγωγικού και διακριτικού, συμπαγούς και μη γραμμικού μοντέλου, κατάλληλο για την αναπαράσταση ανατομικών μορφών. Το μοντέλο αξιολογείται στο πλαίσιο της μελέτης ενός πληθυσμού ασθενών της νόσου Αλτσχάιμερ και ενός πληθυσμού υγιών εθελοντών. Ο διακριτικός μηχανισμός του μοντέλου βασίζεται στο πλαίσιο Support Vector Data Description (SVDD). Το SVDD είναι εφάμιλλο της μονοταξικής Support Vector Machine (One class Support Vector Machine, OCSVM). Το κύριο ενδιαφέρον μας εδώ είναι η χρήση του διακριτικού μοντέλου στον στόχο της αποκάλυψης εκείνων των μορφολογικών διαφορών, οι οποίες είναι οι πιο χαρακτηριστικές στη διάκριση μεταξύ ανατομικών μορφών-μελών μιας δοσμένης τάξης (υγιείς) και μορφών που δεν ανήκουν στην τάξη (ασθενείς).

Ένα σημαντικό χαρακτηριστικό του μοντέλου είναι η χρήση της υπόθεσης ότι τα δεδομένα (οι ανατομικές μορφές) αποτελούν μέλη ενός μη-γραμμικού πολυπτύγματος (πολλαπλότητας) χαμηλής διάστασης. Σε αυτή την εργασία παραμετροποιούμε τις μορφές σαν ένα διανυσματικό πεδίο παραμόρφωσης σε σχέση με μια εικόνα αναφοράς. Αυτή η αναπαράσταση δίνει μια περιγραφή υψηλής ποιότητας και ανάλυσης, με το τίμημα ενός χώρου μορφών πολύ υψηλής διάστασης, τάξης εκατομμυρίων. Μετά την μείωση της διάστασης, χάρη στο βήμα εκμάθησης του πολυπτύγματος, η μορφή μπορεί πλέον να αναπαρασταθεί σαν ένα σημείο χώρου λιγότερο των, τυπικά, δέκα διαστάσεων.

Σε θεωρητικό και μεθοδολογικό επίπεδο, η καινοτομία της παρούσας εργασίας έγκειται σε δύο βασικά σημεία. Πρώτον, προτείνουμε ένα εργαλείο εξαγωγής διακριτικής διαφοράς στα πλαίσια του μονοταξικού μοντέλου SVDD. Δεύτερον, το μη-γραμμικό και συμπαγές μοντέλο εγγυάται την ορθότητα των παραχθείσων αναπαραστάσεων, στο βαθμό που το βήμα εκμάθησης του πολυπτύγματος καταγράφει με ακρίβεια την δομή των δεδομένων. Η εφαρμογή του μοντέλου σε πραγματικά ιατρικά δεδομένα δίνει αποτελέσματα συμβατά με προηγούμενα σχετικά ευρήματα. Επιβεβαιώνουμε ότι συγκεκριμένες περιοχές παρουσιάζουν την πιο εκτεταμένη ατροφία, όπως η κεφαλή και η ουρά του αριστερού και δεξιού ιππόκαμπου. Χρησιμοποιούμε επίσης το μοντέλο σαν ταξινομητή, στο στόχο της διάκρισης υγιών από ασθενείς ιστούς, στη βάση πάντα των ανατομικών δεδομένων.

Résumé étendu

Introduction

Analyse de formes en imagerie médicale

Le but de l'analyse statistique de formes est la construction de modèles capables de décrire des formes et leur variabilité. Dans le champ de l'imagerie médicale, l'analyse de formes est utilisée pour comprendre la variabilité morphologique de divers organes et tissus, tels que le cœur, le cerveau, ou des structures et tissus particuliers dans le cerveau. L'étude de telles données nous permet de mieux comprendre la fonction du corps et du cerveau, d'étudier l'évolution anatomique, de diagnostiquer certaines pathologies, d'étudier leur développement et de planifier un traitement pour ralentir l'évolution de la maladie. Des différences anatomiques entre groupes peuvent être étudiées, où groupe peut signifier similitude d'âge, de condition clinique ou diagnostique. Un facteur important pour l'analyse de formes a été l'introduction de l'imagerie haute résolution et de l'Imagerie par Résonance Magnétique (IRM) en particulier, ce qui a permis des études *in vivo*. L'étude morphologique dans le contexte de l'imagerie cérébrale a été à son origine réduite à de simples mesures d'aire et volume. Aujourd'hui, nous pouvons quantifier des formes avec une résolution et une précision sans cesse meilleures. Les pathologies d'intérêt, dont le développement est marqué par des changements structurels dans les tissus cérébraux, incluent la maladie d'Alzheimer, qui est la pathologie d'intérêt de cette thèse, la sclérose en plaques et la schizophrénie. Le changement structurel dans le contexte de ces maladies reflète l'évolution de la dégénération neuronale ou de l'atrophie du cerveau. Les caractéristiques qualitatives et quantitatives de ces dégénérations structurelles jouent le rôle de biomarqueurs. Dans la mesure où un modèle peut capter et décrire précisément la variabilité anatomique, l'intérêt de la variabilité en tant que biomarqueur est important. La variabilité anatomique présente un intérêt pour les études intra-population. Elle est tout aussi importante pour l'étude de la différenciation anatomique entre un groupe de patients et un groupe de sujets sains (contrôles).

Problématique et contributions de la thèse

Nous nous focalisons dans cette thèse sur la construction d'un modèle compact, non-linéaire, adapté à la représentation de formes. Ce modèle est évalué dans le contexte de l'étude d'une population de patients atteints de la maladie d'Alzheimer et d'une population de sujets contrôles. Le modèle proposé fonctionne à la fois comme un modèle discriminatif et comme un modèle génératif. La composante discriminative du modèle est fondée sur le cadre *Support Vector Data Description* (SVDD). Dans la bibliographie, le SVDD est utilisé pour l'estimation de support, la classification, la détection de points aberrants et le débruitage. Le SVDD est étroitement lié au *one-class Support Vector Machine* (SVM). Notre intérêt principal ici est l'utilisation du modèle discriminatif pour mettre en évidence les différences morphologiques les plus discriminatives entre une classe de formes donnée et des formes n'appartenant pas à cette classe. Un outil similaire (la direction discriminative) a été proposé [Golland 2000] dans le cadre SVM. Notre approche, basée sur le SVDD et que nous appelons *trajectoire discriminative*, génère un

processus d'instances de formes, à partir d'une forme donnée vers le support de classe estimé. Toutes les instances de ce processus sont des projections sur un support estimé, pour différentes valeurs de paramètres du modèle.

Un aspect important du modèle proposé est l'hypothèse que les données (les formes) se trouvent sur ou non loin d'une variété non-linéaire de dimension faible. Cette hypothèse a été largement utilisée dans la bibliographie depuis l'introduction de méthodes telles que l'Isomap [Tenenbaum et al. 2000]. La pertinence de l'hypothèse a été confirmée depuis lors dans plusieurs domaines applicatifs. En pratique, elle nous permet de construire un modèle compact de représentation de formes. Dans notre travail, les formes sont paramétrées par un champ de déformation par rapport à une image de référence (*template*), ce qui donne une description très fine et de haute qualité, mais aussi un espace de description de très haute dimensionalité, de l'ordre de millions de dimensions. Après réduction de dimension, la forme peut se décrire comme un point d'un espace de moins de dix dimensions.

En ce qui concerne la trajectoire discriminative introduite dans cette thèse, l'hypothèse d'existence de la variété de formes nous sert à contraindre la trajectoire à correspondre à des formes anatomiquement correctes, dans la mesure où la structure intrinsèque des données est captée avec précision. La génération de formes incorrectes constitue un problème connu qui touche des modèles, comme par exemple la direction discriminative originale, qui négligent la non-linéarité de la structure des données.

L'innovation théorique apportée par notre modèle se reflète dans deux nouveautés principales de la trajectoire discriminative : premièrement, cette dernière constitue un outil pour extraire la différence discriminative dans le contexte des modèles SVDD ; deuxièmement, elle ne permet que des reconstructions anatomiquement correctes. D'un autre côté, afin de décrire succinctement les différences de groupe entre sujets sains et sujets non sains, nous visualisons les différences avec une carte de p-valeurs, superposée à une forme de référence. Les p-valeurs sont calculées voxel par voxel, comme une fonction de grandeurs calculées d'après le modèle appris. Cette carte sert à indiquer les zones de différences de formes significatives entre classe et non-classe. Une application de notre modèle à des données médicales réelles montre des résultats cohérents avec les connaissances médicales. Nous utilisons également le modèle comme classifieur, permettant de discriminer les sujets sains des sujets pathologiques à partir toujours de formes cérébrales. Pour toutes les applications, la non-linéarité et la compacité du modèle proposé se révèlent très utiles.

Contexte de la thèse et collaborations

Ce travail a été effectué dans le cadre d'une collaboration entre l'équipe « Modèles, Images et Vision » du Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (Université de Strasbourg, UMR CNRS – UDS 7005) et le *Information Processing and Analysis Group* (département d'informatique, Université de Ioannina, Grèce).

D'un point de vue méthodologique, cette thèse est liée aux travaux de Félix Renard [Renard 2011], avec lequel l'auteur a collaboré pendant son séjour au LSIIT. Elle est aussi liée aux travaux de Torbjørn Vik [Vik 2004] sur des modèles probabilistes, linéaires et non-gaussiens et également à ceux de Matthieu Brucher [Brucher 2008] sur l'analyse de formes à l'aide de modèles non-linéaires.

Pendant la thèse, nous avons collaboré avec Frédéric Blanc et Nathalie Philippi (Service de Neuropsychologie, département de Neurologie, Hôpitaux Universitaires de Strasbourg) qui ont fourni une partie des données anatomiques que nous traitons dans ce travail. Une collaboration avec Jack Foucher et Daniel Roquet (INSERM U666 - Hôpitaux Universitaires de Strasbourg) sur une étude de sujets victimes d'hallucinations est également en cours.

Plan de la thèse

Le mémoire est organisé en trois parties. La première partie introduit des méthodes pour décrire les formes. La deuxième partie présente les composantes discriminative et générative de notre modèle. La troisième partie constitue une application du modèle à notre étude de cas, à savoir des populations d'hippocampes dans le contexte de la maladie d'Alzheimer.

La première et la deuxième partie présentent le modèle proposé dans un contexte général d'analyse de formes plutôt abstrait. Nous présentons d'abord l'état de l'art pour décrire la forme. Nous constatons qu'une représentation de formes n'est pas forcément compacte. Ensuite nous appliquons les techniques d'apprentissage de variétés tel que l'Isomap, afin de construire un descripteur de forme compact.

La deuxième partie présente les différentes composantes du modèle proposé. Nous introduisons d'abord le SVDD, le modèle discriminatif de base dans cette thèse. Ensuite nous utilisons le SVDD pour extraire de l'information à propos de caractéristiques de formes qui discriminent le mieux entre des objets dans la classe et des objets hors de la classe. Nous présentons la trajectoire discriminative en termes de projections SVDD, une opération initialement proposée dans le contexte du débruitage de données. Ensuite nous connectons toutes les composantes présentées, tout en créant une chaîne de traitement cohérente. Le modèle est complété avec la mise en place d'une composante de régression multidimensionnelle, qui permet de passer de l'espace réduit vers l'espace de description original, en générant effectivement de nouvelles formes.

La troisième partie présente une application sur des données médicales réelles. Nous introduisons brièvement des aspects-clefs de l'anatomie du cerveau et de la maladie d'Alzheimer. Nous parlons ici surtout de la corrélation entre l'évolution structurelle de certains tissus de cerveau et l'évolution de la maladie d'Alzheimer. Ensuite nous appliquons notre modèle à des données réelles.

Nous notons que nous avons réservé un appendice, afin de présenter une partie du travail de l'auteur qui s'est effectuée comme suite de travaux de recherche commencés avant le début de la thèse. La thématique de ces travaux n'est pas directement liée à l'axe principal de la thèse, que nous venons de discuter ici. Les travaux présentés dans l'appendice se focalisent sur des modèles bayesiens et sur l'inférence bayesienne appliquée principalement à la segmentation d'images. Nous présentons une application à la segmentation d'images naturelles, la segmentation d'IRM cérébrales en trois types de tissus (matière grise, matière blanche, liquide céphalo-rachidien), et une application à la segmentation et restoration d'images polarimétriques.

Représentations compactes de formes

Le point d'entrée d'un cadre d'analyse statistique de formes est un jeu d'apprentissage fini, représenté de manière à permettre l'élaboration d'un modèle bâti sur ces données. Le modèle peut être considéré comme une structure bâtie sur la base du mode de représentation des données, et la qualité de la représentation est alors étroitement liée à la qualité globale du modèle. Une bonne représentation de formes doit être capable de décrire toute forme possible dans le contexte de l'application d'intérêt. Il est également souhaitable que le modèle ne représente que des instances valides, et utilise en même temps un minimum de paramètres dans ce but. Nous tenons à atteindre ces buts en utilisant des méthodes d'état de l'art de représentation, tout en mettant en œuvre des outils d'apprentissage de variétés afin de construire une représentation compacte et non-linéaire.

Représentation de formes

Dans la bibliographie il existe un grand nombre de méthodes de description de formes, comme par exemple les nuages de points, les transformées en distance (*distance transforms*) [Golland 2000], les m-reps [Pizer et al. 2003], les harmoniques sphériques [Chung et al. 2010], les représentations par amers et les champs de déformations à partir d'une image de référence [Cootes et al. 2008, Rueckert et al. 2003].

Voyons brièvement les plus importantes de ces méthodes.

Un *nuage de points* traite la forme comme un ensemble de voxels. Une telle représentation peut être obtenue facilement à partir d'une image d'entrée (e.g. une IRM tri-dimensionnelle), et il existe des scanners qui captent et fournissent la forme directement en tant que nuage de points. Cette modalité est souvent utilisée comme une étape de base dans le but de construire une représentation plus élaborée.

Les *harmoniques sphériques* (spherical harmonics, SPHARM) sont un modèle populaire, basé sur la représentation de la surface de la forme. La forme est vue comme une fonction qu'il faut exprimer au moyen d'un nombre de termes d'un jeu de fonctions de base. En ce sens, elle est conceptuellement liée à la représentation de fonctions et signaux en série de Fourier. La première étape est la construction d'une carte de correspondance entre la surface de la forme d'entrée et la surface d'une sphère. Ensuite, la fonction de correspondance est analysée en fonction de bases pertinentes. Les paramètres de cette analyse constituent le descripteur de la forme. L'inclusion d'un nombre fini de termes de la série infinie, ou, ce qui est équivalent, l'exclusion de termes de hautes fréquences, a pour conséquence le lissage automatique de la forme.

La représentation par *m-reps* (medial representations) est fondée sur l'idée de l'axe médian de la forme et de la transformée de l'axe médian (medial axis transform, MAT). L'axe médian est défini comme l'ensemble des points de la forme qui ont plus d'un point plus proche à la surface de la forme. La transformée de l'axe médian est définie comme l'axe médian complété par l'information concernant les sphères inscrites centrées sur les points de l'axe médian. Une représentation m-rep est constituée d'une collection des points échantillonnés à partir de l'axe médian, qui sont appelés atomes médians (medial atoms). Chaque atome médian comprend de l'information concernant les frontières les plus proches, permettant ainsi de reconstruire la forme à partir de la grille des atomes.

Un *champ de déformation* peut jouer le rôle de descripteur de forme. Après avoir recalé de manière non rigide la forme d'entrée vers une image cible ou template, le champ de déformation qui en résulte et l'image cible peuvent être utilisés pour reconstruire la forme originale et peuvent ainsi représenter la forme. Si nous voulons construire des représentations pour un jeu de formes, nous prenons soin d'avoir la même image de référence pour toutes les mises en correspondance, afin que toutes les représentations se trouvent dans le même cadre de référence. Une caractéristique souhaitable du recalage est de produire des champs qui soient difféomorphiques. En pratique, ceci signifie que nous ne permettons ni déchirement ni pliage pendant le recalage, un déchirement ou un pliage n'étant pas réaliste ici.

Dans ce travail, nous choisissons d'utiliser des champs de déformation en tant que méthode de représentation de base. Chaque champ de déformation est vu comme un champ de vecteurs en 3D. L'ensemble des composantes de chaque vecteur, pour chaque voxel, constitue le descripteur de la forme donnée. Notons bien aussi que, du point de vue méthodologique, la chaîne de traitement proposée ici resterait intacte si nous utilisions une méthode pour décrire la forme autre que le champ de déformation.

L'emploi d'un champ de déformation, ou d'une autre methode générique de description de forme comme celles que nous venons de discuter brièvement ici, malgré son utilité, fournit une représentation en général non-compacte, et un espace de vecteurs non-dense (ou non-spécifique, d'après la nomenclature de [Davies et al. 2008]). Nous contournons ces lacunes en utilisant

l'apprentissage de variété afin de construire une représentation plus pertinente.

Représentation compacte par apprentissage de variétés

Alors que nous augmentons le nombre de paramètres d'un type de représentation de formes, en général nous pouvons avoir une représentation plus fine et exacte. A titre d'exemple, une grille de 256^3 voxels fournit plus de détails qu'une grille de 128^3 voxels. Pourtant, la variabilité des données restera toujours équivalente au nombre de degrés de liberté, typiquement très faible. Par exemple, un ensemble d'objets géométriques qui diffèrent seulement par leur taille aurait exactement un degré de liberté. Un jeu de données d'IRM cérébrales présenterait le même type de problème. Un modèle générique de description de formes échouerait également à capter la structure intrinsèque d'une image de cerveau, dont un élément serait à titre d'exemple que les voxels extérieurs correspondent aux voxels du crâne, ou que le cerveau est quasi-symétrique et que les deux hémisphères sont joints au moyen du corps calleux. En termes de modèle, une représentation présentant une grande disparité entre le nombre de ses paramètres et le nombre de degrés de liberté est une représentation non-compacte.

Une représentation non-compacte peut se révéler problématique pour diverses raisons. Tout d'abord, un modèle compact serait par définition plus expansif que nécessaire. D'autre part, les modèles de grande dimensionalité sont caractérisés par ce qui est connu dans la littérature comme la *malédiction de la dimension*. La malédiction de la dimension exprime le fait que, afin de construire un modèle statistique sur un espace donné, un nombre de points d'apprentissage en relation exponentielle avec le nombre de dimensions est nécessaire [Bellman 1961]. Sinon, tout modèle construit sur cet espace est susceptible de disposer de propriétés d'induction très faibles.

Par conséquent, il serait souhaitable d'incorporer au modèle un moyen d'obtenir un espace de dimension réduite. Nous utilisons à ce propos l'apprentissage de variété, et plus particulièrement l'Isomap [Tenenbaum et al. 2000]. Dans les méthodes d'apprentissage de variété, le principe de base est que l'espace de formes possibles constitue une variété de dimensionalité faible, qui occupe un volume qui ne constitue qu'une faible fraction de l'espace englobant. Autrement dit, les données sont supposées faire partie (ou être assez proches) d'une variété.

Les détails de la chaîne de traitement qui permet de déterminer la variété à partir des données d'entrée sont spécifiques à chaque méthode d'apprentissage de variété. Dans le cas de l'Isomap, qui est la méthode retenue pour ce travail, l'entrée de l'algorithme est constituée par les distances entre les formes du jeu d'apprentissage. Après une étape convertissant ces distances en distances géodésiques, l'algorithme cherche des points d'un espace de faible dimension correspondant aux formes d'entrée, tels que les distances entre les points soient préservées. En pratique, cette opération est répétée pour plusieurs dimensionalités de l'espace cible, et la dimension est choisie au moyen d'une règle heuristique. Les nouveaux points générés constituent la représentation compacte. Idéalement, chaque point de ce nouvel espace correspond à une observation possible et en ce sens valide, en tant qu'élément de la variété considérée.

Construction du modèle

Nous examinons maintenant les composantes discriminative et générative du modèle, qui sont bâties sur la couche de représentation compacte de formes que nous venons de discuter.

Un modèle discriminatif est essentiellement une machine qui répond à la question : étant donné un ensemble d'objets, réparti en classes, inférer la classe à laquelle appartiendra un objet qui n'appartient pas à l'ensemble original. Dans ce travail, nous nous focalisons sur un tel modèle, le SVDD [Tax & Duin 1999]. Le SVDD calcule la frontière de la classe dite normale,

en formant une hyperboule autour des points normaux d'entrée. Le rayon et la position de cette hyperboule, qui joue le rôle du support de la distribution des formes normales, sont déterminés en tant que compromis entre deux buts antagonistes. Ces buts sont de minimiser le volume du support et d'englober autant de formes normales d'entrée que possible. Dans le contexte de l'analyse de formes, une forme nouvelle serait identifiée comme normale si elle tombait dans le support, et anormale sinon.

Notons ici que la forme d'hyperboule peut se généraliser, en employant une application ϕ de l'espace original vers un espace auxiliaire, dit espace de caractéristiques, où le support prend la forme d'une hyperboule. Comme la correspondance entre l'espace original et l'espace de caractéristiques est en général non-linéaire, la forme de l'antécédent de l'hyperboule dépend finalement de la flexibilité de la fonction de correspondance utilisée.

La classification d'un nouvel arrivant est considérée comme l'utilisation « par défaut » d'un modèle discriminatif. Une utilisation autre que la classification d'un nouvel arrivant est l'estimation du type de différence de formes le plus marquant entre les formes normales et les formes anormales. Dans ce travail, nous sommes plutôt intéressés à cette dernière question. Dans le contexte médical, la question est particulièrement importante, puisque la classe peut faire référence au diagnostic ou à l'état clinique du sujet associé. Chercher la différence discriminative de formes peut, dans ce cas-ci, mettre en évidence des détails importants du développement de la pathologie d'intérêt.

Cette différence entre objets dans la classe et objets hors de la classe, ou différence discriminative, a été abordée dans la bibliographie au moyen d'outils tels que la direction discriminative [Golland 2000, Zhou et al. 2008] et la structure discriminative (*discriminative pattern*) [Koutsouleris et al. 2009], tous fondés sur le Support Vector Machine (SVM) en tant que machine discriminative.

La direction discriminative [Golland 2000] donne la direction d'un point donné vers la frontière du SVM. En faisant évoluer le point le long de cette direction, la forme change graduellement vers un objet de l'autre classe. Pourtant, la direction discriminative ne peut pas garantir la validité des formes ainsi générées. Dans [Zhou et al. 2008], une version de la direction discriminative est proposée, comprenant un terme de régularisation dans la fonction objectif.

L'outil proposé ici dans ce contexte, la trajectoire discriminative, est posé en tant que partie d'un modèle SVDD, au lieu du SVM des autres approches. La trajectoire est définie au moyen de la notion de projection SVDD, qui est le point du support le plus proche de l'objet d'intérêt. Nous écrivons ici $p_R(x)$ pour la projection du point x sur le support SVDD de taille optimisée R. La différence entre le point donné x et le point projeté $p_R(x)$ est ici interprétée comme l'élimination de la différence de forme qui est non-caractéristique de la classe normale. La trajectoire discriminative est définie comme un ensemble de projections SVDD. Chaque instance de l'ensemble correspond à un paramètre spécifique du modèle SVDD. Les instances du processus correspondent à toute la gamme de taille d'hyperboules SVDD, de la plus petite et optimale, à la plus grande, où la projection s'identifie au point initial. Formellement, la trajectoire discriminative $\delta(u)$ serait définie comme

$$\delta(u) = p_{R'}(x), \qquad \text{avec} \quad R' = (1-u)R_{\max} + uR, \tag{0.1}$$

où $u \in [0,1]$, et $R_{\max}$ désigne une taille de l'hyperboule SVDD telle que $p_{R_{\max}}(x) = x$, c'est-à-dire que le rayon est maximum.

D'autre part, tous les points de la trajectoire discriminative sont anatomiquement corrects. Dans la mesure où l'étape d'apprentissage de la variété explique correctement la structure non-linéaire des données, tous les points générés seront valides, et une correction de la trajectoire discriminative n'est pas nécessaire.

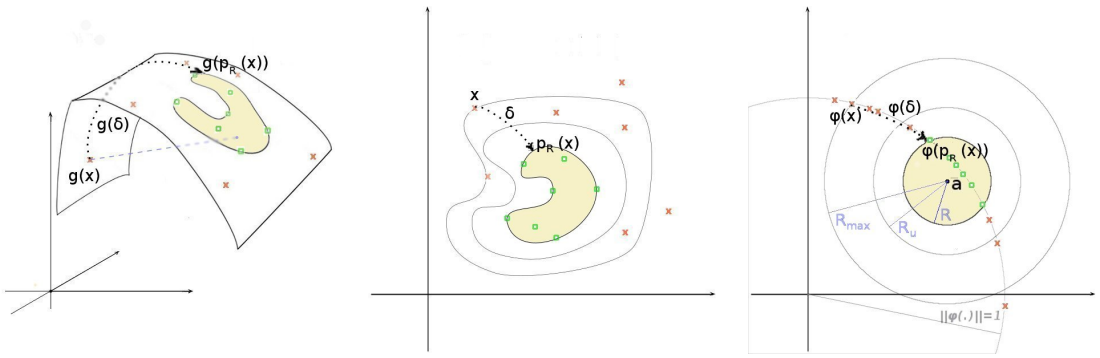
Il reste à produire des formes à partir des points et trajectoires dans l'espace réduit. Nous utilisons dans ce but les couples des formes d'apprentissage et des coordonnées de l'espace réduit,

afin de construire des fonctions de base de régression. Chaque nouvelle forme apparaît comme un nouveau champ de déformation, qui se calcule en tant que somme pondérée de fonctions de base centrées sur les formes du jeu d'apprentissage. Formellement, la forme reconstruite y est liée au point de l'espace réduit x par la relation

$$y = g(x) + \epsilon = W^T \beta(x) + \epsilon,$$

où W est une matrice de poids pour les fonctions de base β. Les poids correspondant aux fonctions de base sont obtenus simplement par une solution d'un problème de moindres carrés.

A titre de résumé, rappelons que les données provenant du jeu d'apprentissage sont représentées de trois façons (voir fig. 0.1) : en tant que points de l'espace des champs de déformation ; en tant que points de l'espace réduit, cet espace étant obtenu au moyen de l'étape d'apprentissage de la variété ; en tant que points de l'espace des caractéristiques, espace où le support SVDD forme une hyperboule autour des données. Nous appelons ces espaces respectivement espace original, espace réduit, et espace de caractéristiques. La chaîne de traitement peut se décrire ainsi en tant qu'interaction entre ces trois espaces : les formes d'apprentissage de l'espace original se transforment en points de l'espace réduit, et leur image dans l'espace des caractéristiques aide à déterminer le modèle discriminatif. De nouveaux points et trajectoires de l'espace réduit se transforment en formes et processus de formes dans l'espace original.



(a) (b) (c)

FIGURE 0.1 – Illustration des notions d'espace ambiant, d'espace réduit et d'espace de caractéristiques. Les cercles verts représentent des données normales, les croix rouges représentent des données anormales. (a) Espace ambiant. Cet espace est très peu densément rempli de données. La structure ayant une allure de feuille –la variété– est un espace de dimensionalité plus faible que celle de l'espace ambiant. Les données font partie de la variété. En utilisant des techniques d'apprentissage de variété, nous pouvons représenter la variété comme un espace réduit (b). Cet espace n'est pas observé, mais il est dense, au sens où ses points correspondent en principe tous à des observations valides. Nous nous servons de l'espace de caractéristiques (c) pour estimer la distribution des points normaux (support). Sur chacun des graphes, $p_R(x)$ correspond à la projection de la donnée x sur le support normal (*projection SVDD*). La *trajectoire discriminative* δ, représentée en pointillés noirs, est le chemin joignant les points x et $p_R(x)$, inclus dans la variété.

Nous testons la chaîne de traitement proposée sur des données de synthèse. Nous appliquons des déformations artificielles sur des formes géométriques et des anatomies médicales pour former les points anormaux des jeux de données de synthèse. Le modèle proposé permet de reconstruire les déformations originales. La trajectoire discriminative produit un processus de formes à partir d'une forme déformée vers la version normale correcte, alors que nous pouvons résumer la différence discriminative à l'aide d'une carte de p-valeurs. Cette carte mesure la différence statistique des points normaux et des points pathologiques voxel par voxel. Les points

significatifs sont marqués d'une p-valeur faible, alors que les points non significatifs sont marqués d'une p-valeur grande.

Les formes de synthèse sont délibérément construites afin de constituer une structure non-linéaire de l'espace des champs de déformations (espace ambiant). Ceci est réalisé en appliquant divers types de déformations pour créer des données pathologiques. Nous montrons que notre modèle peut capter et reproduire tout type de déformation, grâce à sa composante de réduction de dimension. Nous démontrons également qu'au contraire, un modèle qui ne respecterait pas la structure non-linéaire intrinsèque des données échouerait à capter correctement la différence discriminative.

Étude de cas : dégéneration de l'hippocampe dans le contexte de la maladie d'Alzheimer

La dernière partie de la thèse présente une application du modèle proposé à des données médicales réelles. L'étude de cas est la maladie d'Alzheimer. La progression de la maladie d'Alzheimer, comme celle de nombre d'autres pathologies, est fortement liée à l'évolution structurelle de certaines régions d'intérêt dans le cerveau. Ces régions incluent les hippocampes, les amygdales et les ventricules latéraux. La dégéneration de ces organes cause également la dégénération de leur fonction, causant une déficience cognitive liée au progrès de la maladie.

L'anatomie du cerveau, le développement de plaques amyloïdes et d'autres symptômes de la maladie peuvent être considérés comme des indicateurs de la gravité de la maladie. Ces indicateurs sont connus comme étant des biomarqueurs de la progression de la maladie. Idéalement, un biomarqueur doit être corrélé aux différentes étapes de la maladie, et pouvoir aider à prévoir l'évolution clinique du patient. La recherche en prévention et ralentissement de la maladie pourrait donc bénéficier d'une information portée par le biomarqueur.

Dans le contexte de la production d'un biomarqueur structurel, aussi informatif que possible, l'étude de la structure de l'évolution anatomique du cerveau ou d'un tissu particulier du cerveau avec autant de précision que possible reste un défi important. Ceci est l'objet de nombreuses études en anatomie numérique, et c'est en particulier l'objet de la partie applicative du présent travail.

Expériences sur données médicales réelles

Nous appliquons la chaîne de traitement que nous avons mise en place ici sur des données réelles. Nous nous servons de deux jeux de données IRM, comprenant des hippocampes et amygdales qui proviennent d'une part de sujets atteints de la maladie d'Alzheimer et d'autre part de sujets sains.

Le premier jeu de données est constitué par la base de données libre OASIS. Nous répartissons les données disponibles en deux sous-ensembles, qui contiennent respectivement des hippocampes gauches et des hippocampes droits, provenant de patients âgés de plus de 60 ans. Nous avons 198 structures pour chaque sous-ensemble. Nous étiquetons les données suivant l'information clinique disponible en trois classes : sujets sains, sujets malades, et sujets d'état intermédiaire.

Le deuxième jeu de données provient d'une étude menée au sein des Hôpitaux Universitaires de Strasbourg (Centre Hospitalo-Universitaire de Strasbourg, Service de Neurologie, Unité de Neuropsychologie). Nous répartissons les données disponibles en quatre sous-ensembles, qui contiennent respectivement des hippocampes gauches, des hippocampes droits, des amygdales gauches et des amygdales droites. Nous avons 26 structures pour chaque sous-ensemble. Toutes correspondent à des sujets âgés de 69 à 84 ans. Les sujets sont déjà étiquetés et répartis en deux

classes suivant leur état, malade ou sain.

Nous appliquons nos outils d'estimation de la différence discriminative entre classes –trajectoires discriminatives (voir fig. 0.2) et cartes de significativité– au premier jeu de données OASIS. Nous découvrons des points significatifs autour de régions de la tête et de la queue et de grandes parties de la région CA1 des deux hippocampes. La tête et la queue de l'hippocampe sont connus pour être des régions importantes pour l'évolution de la maladie d'Alzheimer [Morra et al. 2010]. L'atrophie de CA1 est également connue pour être associée en général à la diminution de la capacité de récupération de mémoire et de consolidation, et elle est associée à l'apparition de la maladie d'Alzheimer [West et al. 2000]. Ce résultat est donc en adéquation avec les résultats des études antérieures.

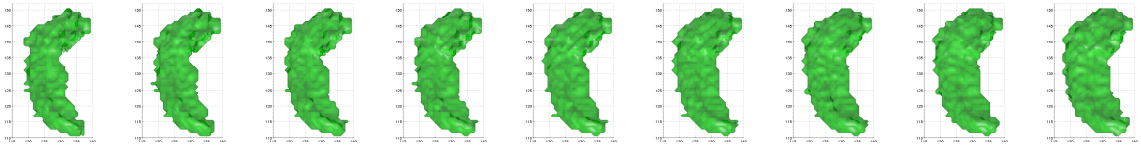


Figure 0.2 – Échantillons le long d'une trajectoire discriminative. De gauche à droite, les formes évoluent graduellement d'une forme extrême de la pathologie vers une anatomie saine et non-atrophiée.

Nous utilisons le deuxième jeu de données pour évaluer la capacité du modèle à la classification en groupe de caractéristiques pathologiques homogènes, à partir de données structurelles. Les résultats montrent une performance numérique proche de celle de l'état de l'art de la bibliographie. Il est montré également que la composante de réduction de dimension du modèle a également un effet bénéfique sur les performances du modèle en tant que classifieur.

Conclusion et perspective

En conclusion, nous avons présenté dans ce travail une modélisation non-linéaire compacte et une chaîne de traitement permettant l'étude de formes. Etant donné un jeu de formes d'apprentissage, nous estimons d'abord la distribution de ces formes dans un espace de description, en utilisant la théorie SVDD. Nous nous servons ensuite du support estimé afin de discriminer les points normaux des points anormaux. Posé dans le contexte de l'imagerie médicale, ceci se transpose en la discrimination entre tissus sains et tissus pathologiques. A partir de ce modèle discriminatif, nous introduisons la trajectoire discriminative, comme outil pour estimer la différence discriminative. La distribution des données sous-jacente, modélisée comme une variété non-linéaire, est partie intégrante de la chaîne de traitement.

Dans la mesure où l'apprentissage de la variété explique correctement la structure de données sous-jacente, le modèle est compact, ce qui nous mène à une description de données plus fidèle à la réalité. D'un côté, les nouvelles formes générées grâce au modèle sont anatomiquement correctes, puisqu'elles sont contraintes à appartenir à la variété. D'un autre côté, tous les éléments de la trajectoire discriminative sont également contraints à faire partie de la variété, ce qui nous sert à fournir une estimation précise de la différence discriminative. Nous illustrons ce point en appliquant le modèle à des jeux de données géométriques et à des jeux de données anatomiques de synthèse. L'application du modèle à des données réelles donne des résultats qui nous mènent à des conclusions cohérentes avec les résultats des recherches antérieures.

Nous confirmons que suite à la manifestation et pendant le déroulement de la maladie d'Alzheimer, certaines régions dans le cerveau, et plus particulièrement des régions des deux hippocampes, sont atteintes plus que les autres régions. Il s'agit en particulier de la tête et de la queue des hippocampes. Ceci a été montré à l'aide de la trajectoire discriminative et des

cartes de significativité. Les trajectoires discriminatives expliquent la différence discriminative en tant que processus entre un seul sujet vis-à-vis d'une classe, alors que la carte des p-valeurs fournit un résumé de la différence en tant que différence de classe vis-à-vis de la non-classe. Bien que la sortie de la fonction SVDD présente un certain degré de corrélation avec des scores cliniques, qui sont largement une fonction du temps, notons bien ici que le paramètre libre de la trajectoire discriminative n'est pas identique à un paramètre de temps. En perspective, un jeu de données longitudinal, couplé à un modèle pouvant générer de telles données, pourrait mener éventuellement à des conclusions plus concrètes en modélisant la forme directement en tant que fonction du temps.

Chapter 1

Introduction

In this introduction, we shall present the work realized in the context of this thesis. We shall discuss the core theoretical notions, application context, and contributions of our work, as well as the related collaborations and projects.

1. General context

1.1 Shape and neuroimaging

In this thesis, we are interested in statistical shape analysis, applied in a medical, neuroimaging context. Statistical shape analysis refers to constructing models capable of describing shape and variability of shape. In medical imaging, shape analysis is used to model the morphological variability of various body organs and tissues, such as the heart, the brain or specific brain tissues. In turn, study of such data allows us to better understand body and brain function, study anatomical evolution, diagnose certain pathologies, identify their time of onset and study their development, realize the planning of therapeutical treatment, and monitor its impact. Group anatomical differences can be studied, where group difference can refer to a partition by age-group, pathological vs control or gender anatomical difference.

Paramount to medical shape analysis has been the introduction of high resolution 3D imaging such as Magnetic Resonance Imaging (MRI), which has allowed *in vivo* study of medical shape. Various MRI modalities exist today. To name a few, these include structural MRI, which is suitable to describe anatomical structure, functional MRI, which associates function and activation to brain structure, and diffusion MRI, which measures water molecule diffusion in biological tissue. For the purpose of tissue shape analysis, structural MRI is the most suitable. Morphological study in neuroimaging was originally based on simple measurements of area and volume. Such measurements provide a crude, elementary measure of morphological variation, and as such can only capture shape complexity in a limited degree.

Pathologies of interest, whose development induces structural change in brain tissue, include multiple sclerosis, schizophrenia and Alzheimer's disease (AD), the latter being the case study of

this thesis. Structural change in these diseases reflects the progress of neuronal degeneration, or brain atrophy. The qualitative and quantitative characteristics of the progress of this structural degeneration play the role of the so-called disease *biomarker*. A biomarker is in general defined as a biological or biologically derived indicator of a process, event, or condition such as disease (or other conditions such as aging, or exposure to a toxic substance) [Medical Dictionary 2012]. To the extent that a model can describe anatomical variability as precisely as possible, its value as a biomarker of disease progression is more important. Anatomical variability is of interest as intra-population variability, as well as the study of anatomical difference between a group of patients and a group of healthy controls.

1.2 Statistical shape analysis

Quantitative shape analysis is generally admitted to date originally back to the work of D'Arcy Thompson [Thompson 1917]. The context of Thompson's works is the study of the relation between form and function. A group of biological forms is put into anatomical correspondence, by defining points of interest in a template shape and their corresponding points in a second shape. These points define a grid of lines (see fig. 1.1) which is accordingly deformed to create a dense correspondence between shape points ("method of coordinates"). This approach presents several practical difficulties, such as defining the interpolation allowing the passage from point correspondence to straight line correspondence, or the definition of the position of the landmark points. These problems were left unanswered originally [Heinrich 2008]. The milestone work of Thompson includes, albeit in a premature form, some aspects and problems of statistical shape analysis recurrent in models and methods of today. In a modern, computational context, description of a set of shapes is performed by defining a correspondence between shape data, either explicitly or implicitly. Numerous shape description methods exist, which in general permit to map each shape-object to a mathematical object, playing the role of a shape descriptor. Landmark descriptors, spherical harmonics and deformation fields are some of the most important shape descriptors (see chapter 2).

A model, in shape analysis or elsewhere, seeks to describe observed data as manifestations of some underlying pattern, as exactly and yet as simply as possible [Davies et al. 2010]. Given a finite training set of shape data, generative or discriminative models can be constructed. The goal of a generative model is to generate shape instances that are in some sense similar to the training shapes. This necessarily presupposes the estimation of the shape distribution in a shape description space. A discriminative model supposes the partition of the data into a finite set of classes (for example, pathological and non-pathological class), and its goal is to decide the class of new, unseen data. Regardless of its exact utility and type, a good statistical shape model should have, according to [Davies et al. 2008, Davies et al. 2010], a set of traits that can be summarized to (1) generalization ability, so that the model can represent any possible shape instance (2) specificity, so that it can only represent valid instances of the modelled class of object (3) compactness, so that the model can be represented with few parameters. We might add here that all of these traits are correlated to one another. For example, a non-compact model will in all likelihood also be non-specific, and vice versa. In this work we are interested in building compact shape representations, and dense -in the sense of highly specific- shape spaces. Manifold learning is a tool that will aid us in this direction (see chapter 3).

A basic step in creating a model is to define a shape description space where data reside. Two typical assumptions can be made at this point [Davies et al. 2010] : (1) data form a linear substructure in the description space that can be captured using a related estimation method such as Principal Component Analysis (PCA) and (2) shape data distribution can be captured using a Gaussian probability density function. The undeniable advantage of these assumptions is exactly their simplicity. However data may not always fit such simplistic hypotheses. The

goal of the current thesis, and main axis of the presented work, is the creation of a statistical shape model that relaxes the aforementioned hypotheses and describes shape in a compact and precise manner.

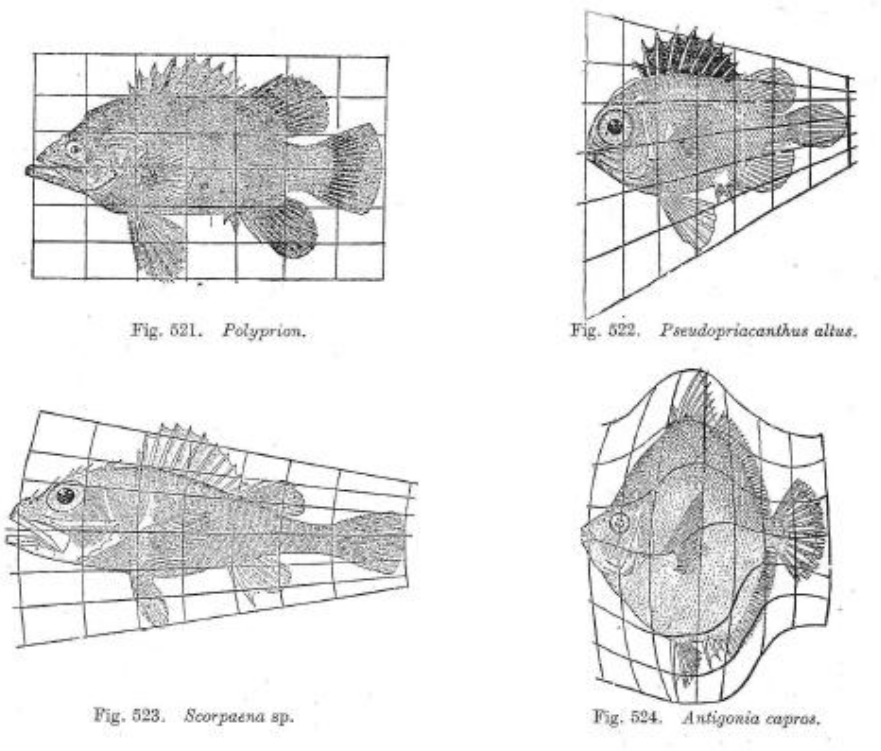


FIGURE 1.1 – Putting shapes into correspondence with the method of coordinates. Figure reproduced from [THOMPSON 1917].

2. Contributions

The material presented in this thesis centers around the construction of a compact, non-linear model suitable to represent shape. This model is evaluated in the context of a population study of Alzheimer's disease patients and controls. The proposed model acts both as a discriminative and a generative model.

The discriminative aspect of the model is based on the Support Vector Data Description (SVDD) model. In the literature, uses of SVDD include support estimation and classification, anomaly detection and denoising. SVDD is very closely related to the one-class Support Vector Machine [SCHÖLKOPF et al. 2001]. It should also be noted that, while SVDD is based on a one-class versus-all philosophy, it can be extended to a multi-class model [LEE & LEE 2007], competing fairly with state-of-the art discriminative models. Our main interest here is to use the discriminative model to uncover shape difference between a given class of shapes versus shapes not in the class, that is the most characteristic of the class difference. An analogous tool, called discriminative direction, has been proposed originally in [GOLLAND 2000] in the context of a SVM model. Our proposed SVDD-based approach, which we call *discriminative trajectory*, generates a process of shape instances starting from a given shape, towards the estimated class support. All process instances share the trait of being projections onto the estimated support, given varying model parameters. Our naming of the proposed scheme as discriminative trajectory, serves to differentiate from related models in the literature and to underpin our formulation of discriminative difference explicitly as a process.

An important aspect of the proposed model is that we use the hypothesis that shape data lie on (or sufficiently near) a non-linear manifold of low dimension. This *manifold hypothesis* has been extensively used in the literature since the introduction of milestone non-linear manifold learning techniques such as isomap [Tenenbaum et al. 2000]. The usefulness of the manifold hypothesis has been confirmed in various domains (see chapter 3). In practice, it allows to construct a compact model of shape representation. With respect to the proposed discriminative trajectory, the manifold hypothesis constraints trajectory shape instances to correspond only to anatomically correct shapes, to the extent that the intrinsic data structure is captured faithfully. Conversely, models that neglect this non-linear structure, such as the discriminative direction [Golland 2000], cannot guarantee anatomical correctness. (see [Zhou et al. 2008] for a related discussion).

Methodologically, the contribution of this thesis is reflected in two aspects of the proposed discriminative trajectory. First, it is a tool for extracting discriminative shape difference in the context of SVDD models. Second, it allows only for anatomically correct reconstructions. To the end of describing group difference between healthy and pathological subjects, we visualize as a p-value map on residuals associated to each shape voxel. This map indicates the locus of significant difference between class and non-class. The application of our model on real anatomical data, shows results consistent with related literature findings. The non-linear, compact character of the model is shown to have overall a beneficial impact in any aspect or type of application.

3. Collaborations

This thesis has been realized in the context of a partnership between two research groups : the "Modèles, Images et Vision" (MIV) group, part of the Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédetection (LSIIT, UMR 7005, Université de Strasbourg - CNRS), and the "Information Processing and Analysis" (IPAN) group, part of the Department of Computer Science (University of Ioannina, Greece).

This thesis is part of a line of research projects of the MIV group of LSIIT, under the general topic of "Traitement d'Images Médicales" (Medical image processing). The thesis is also part of research work conducted by the federation "Imagerie et Robotique Médicale et Chirurgicale" (IRMC). More generally, this research topic is supported by the pôle of competitivity "Innovation thérapeutiques" and is part of pôle image Région Alsace. This thesis was funded by the Région Alsace.

From a methodological point of view, this thesis is closely related to the work of Félix Renard [Renard 2011], with which the author of the thesis has collaborated. It is also related to the work of Torbjørn Vik [Vik 2004] in probabilistic linear, non-Gaussian models and that of Matthieu Brucher [Brucher 2008] in shape analysis with non-linear models.

We have used two main sources of real neuroimaging data. The first source is the free-access OASIS database [Marcus et al. 2007]. The other is a study conducted at the University Hospital of Strasbourg through the service of Neurology and the Neuropsychology unit, [Philippi et al. 2012], courtesy of Frédéric Blanc and Nathalie Philippi of the same unit. A collaboration with the latter group, aiming at exploring correlations of memory and cognitive markers with structural deformation is currently in progress.

A collaboration with Jack Foucher and Daniel Roquet (INSERM U666 - University Hospital of Strasbourg) is also in progress. This project targets at building a suitable discriminative model for hallucinating and non-hallucinating schizophrenic patients, plus controls. The input is functional connectivity data processed with the Independent Component Analysis (ICA) model.

4. Thesis outline

We present the proposed statistical shape analysis model in three parts.

The first and second part presents the proposed model in an abstract, general shape analysis context. Chapter 2 and chapter 3 constitute part I, and discuss shape description. In chapter 2, we discuss state-of-the-art shape methods for shape description. We observe that these do not necessarily lead to a compact representation. In the next chapter (chapter 3), we use the technique of manifold learning to create a compact shape descriptor.

The second part presents the discriminative and generative component of the model. In chapter 4, we introduce elements of SVDD theory. In chapter 5, we use SVDD to extract information about shape characteristics that best discriminate objects of a given class from objects not in the class. This is encoded in the notion of the discriminative trajectory. We present the latter in terms of the SVDD projection, an operation originally proposed for data denoising, and discuss the relation of the proposed discriminative trajectory to the SVM-based discriminative direction. In chapter 6, we "connect" the manifold learning component with the SVDD component of the model, creating a coherent processing pipeline. We express the model in terms of three description spaces – the ambient space, the reduced space, and the feature space – and the relation of each one to the other. Finally we apply our model to geometrical and synthetic anatomical shape, in order to assess the validity of our claims. We show that the proposed discriminative trajectory, as part of the present model, reveals discriminative difference in detail that a non-compact model would not achieve.

The third part of the thesis presents an application of the proposed model to real neuroimaging data. In chapter 7, we discuss the correlation of structural evolution to the progress of Alzheimer's disease, and the related notion of a disease biomarker. In chapter 8, we use our model on a group of patients, intermediate condition subjects and controls, and create discriminative trajectories for the left and right hippocampi. We also illustrate the use of our model as a classifier, deciding over the likelihood of a hippocampus or amygdala originating off a healthy subject or not, based always on structural data.

We have reserved a fourth -appendix- part to present work conducted since the beginning of the thesis, but not directly thematically connected to the main presented subject. The work presented in the appendix is centered around using Bayesian models and inference for image segmentation. We discuss an application in natural image segmentation, brain MR image segmentation in three tissue classes (gray matter, white matter, cerebrospinal fluid) and an application in polarimetric image segmentation and restoration.

Part I

Compact shape representation

Introduction

The entry point of any statistical shape analysis framework is a finite training set of shapes, represented in a way suitable to allow for a subsequent elaboration of a model built on this data. In a way, a model is a superstructure constructed on the basis of the way shape is represented ; hence the quality of the representation is organically related to the quality of the model.

In the first part of this thesis we discuss how to describe shape. First we present a brief state of the art of shape description. It is in the interest of the model as a whole to use a descriptor that can accurately describe any possible shape in the application of interest, as well as to ease the construction of an accurate, flexible generative or discriminative model. Next we discuss how to process such a shape descriptor with manifold learning in order to obtain a compact shape representation. As we shall see, compactness is a trait that can lead to more useful and stronger models of shape.

Chapter 2

Shape description

Encoding the shapes of the input training set in some computationally usable, numeric form, is a basic first step in any statistical shape analysis pipeline. There exist plenty of ways in the related literature to represent shape, like distance transforms [Golland 2000], m-reps [Pizer et al. 2003], spherical harmonics [Chung et al. 2010], landmark-based representations deformation vector fields with respect to a template [Cootes et al. 2008, Rueckert et al. 2003]. In this chapter we shall review some of the most widely used models for shape representation. We bear in mind that we shall be eventually specifically interested in training sets in the context of medical imaging. Most of the models that follow have been widely used in related applications of the same area. We shall imply a three-dimensional shape, unless otherwise stated. In most cases that we examine, passage from the formulation of a 2D planar shape descriptor to its corresponding version in 3D –or vice versa– is straightforward.

1. Models for shape representation

1.1 Point cloud

The most straightforward way to describe a shape, given a three-dimensional voxel intensity map (such as the $T1$-weighted intensity information of an MRI), is the intensity map itself, or a thresholding/segmentation over the input intensities producing a binary map. In this map, true-valued voxels correspond to voxels that are part of the shape, and false-valued voxels correspond to background. This is what is called a *point cloud*. Point clouds are the default output of $3D$ scanners, and can be further processed to build three-dimensional surfaces meshes. Besides medical imaging, point clouds have been employed in a wide spectrum of applications, including computer-aided design, computer generated-imagery (CGI) visualization and animation, archaeology and antiquities preservation [Allen et al. 2004].

1.2 Spherical harmonics (SPHARM)

The spherical harmonics model (SPHARM) [Brechbühler et al. 1995,Schönefeld 2005] is a popular, surface-based representation method. It views shape as a function which is to be expressed in terms of a suitable set of basis functions, as is the case for example for the Fourier or the Legendre series in 1D. The first requirement in a spherical harmonics representation is to build a proper parametrization of the shape to be represented. This means defining a continuous, one-to-one mapping from the surface of the original shape to a two-coordinate space with spherical topology. Mapping our object to a sphere immediately implies that the modelled object is necessarily a single, simply connected object.

The parametrization is formulated as a constrained optimization problem. The variables of the optimization problem are the coordinates of all vertices, i.e. the positions on the unit sphere to which the vertices are mapped. The parametrization is constrained, so that any object surface region maps to a region of proportional area on the sphere. The objective is to miminimize the distortion of the surface net in the mapping. To fulfill this goal perfectly, a target object facet should map to a "spherical square".

The previous step produces a one-to-one mapping where each pair of spherical coordinates (θ, ϕ) is connected to a point $x \in \mathbb{R}^3$ in Cartesian coordinates, which corresponds to a position on the surface of the object to be represented. Coordinates (θ, ϕ) are defined over a range of $(0, \pi)$ and $(0, 2\pi)$ respectively. The decomposition of the function $x(\theta, \phi)$ with spherical harmonics is defined as a series of the form

$$x(\theta, \phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} c_l^m Y_l^m(\theta, \phi), \tag{2.1}$$

where Y_l^m denotes a basis function of degree l and order m [Brechbühler et al. 1995] and c_l^m is a tri-variate complex vector representing the expansion coefficient of the same degree and order. Basis functions Y_l^m form an orthonormal set and are defined as

$$Y_l^m(\theta, \phi) = \sqrt{\frac{2l+1}{4} \frac{(l-m)!}{(l+m)!}} P_l^m(cos\theta) e^{im\phi}, \qquad Y_l^{-m}(\theta, \phi) = (-1)^m Y_l^{m*}(\theta, \phi), \tag{2.2}$$

where

$$P_l^m(x) = (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x), \qquad P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2-1)^l. \tag{2.3}$$

The coefficients c_l^m are formally defined as

$$c_l^m = \int_0^{\pi} \int_0^{2\pi} x(\theta, \phi) Y_l^m(\theta, \phi) d\phi \sin\theta d\theta. \tag{2.4}$$

In practice, an approximation of eq. (2.4) is used to calculate the expansion coefficients [Brechbühler et al. 1995]. The higher the degree of an expansion set and its coefficients, the finer the structural information they describe. Low-degree terms carry coarse structural information. In practical applications, only a limited number of the lowest degree coefficients are used, and any further analysis is performed over them. Using a limited number of coefficients naturally serves as an implicit form of mesh smoothing [Chung et al. 2010].

Let us note that the spherical harmonics have a global support, so the decomposition coefficients encapsulate necessarily global information about the shape. Decomposition using other sets of bases has been considered in the literature, like Fourier descriptors [Staib & Duncan 1992], and more recently wavelets [Nain et al. 2007,Yu et al. 2007]. The latter have a non-global support, and in principle each coefficient can carry local shape information.

1.3 Medial axis representations

1.3.1 Medial axis and medial axis transform

[Blum 1967, Blum & Nagel 1978] introduced the medial axis transform as a shape descriptor, originally for a biological application context. The medial axis transform is based on the notion of a shape medial axis. Formally, the medial axis, or topological skeleton, of a shape is defined as the set of points with more than one closest point to the shape boundary. It turns out that the medial axis forms a tree-like hierarchical structure, intuitively similar to a sort of an object "backbone". The medial axis transform (MAT) is defined as the medial axis plus information about the largest inscribing spheres to any of its points. Most of the algorithms for computing the MAT typically require a binary segmentation of the object as input [Golland et al. 2000, Leymarie & Levine 1992].

1.3.2 Medial representations (m-reps)

M-rep is a representation based on the medial axis and the medial axis transform. M-reps come to amend weaknesses of the medial axis transform representation. Namely, the medial axis is difficult to manipulate in its continuous version, and is not robust to noise, in the sense that small surface perturbations can cause changes in the medial axis [Fletcher et al. 2001]. The m-rep is defined as a structure made up of so-called *medial atoms*. Medial atoms are objects sampled over the shape medial axis, from which the entire medial axis (and the original shape itself) can be reconstructed. Each medial atom thus encapsulates information about the medial axis in its immediate neighborhood. The mesh of the medial atoms forms a structure referred to as a medial surface. With the exception of atoms on the medial surface edge and branch positions, each medial atom defines two equal length vectors, named the "port vector" and "starboard vector". These vectors are incident and normal to the shape surface at their ends.

A medial atom is defined as a four-tuple

$$m = \{x, R, F, \theta\},$$

where $x \in \mathbb{R}$ is a position on the medial axis, R is the radius of the inscribing sphere, θ is one half of the angle between the port and the starboard vector. F is a three-dimensional frame with origin at x and fitted to the geometry of the medial axis. The frame F is a rotation of the standard Euclidean basis; the first basis vector is defined as the bisector of the port and the starboard vector, the second is perpendicular to the first, so that together they span the plane tangent to the medial surface on x. The remaining frame vector is the normal at the same point x.

Simple enough shapes can be described using a single medial atom mesh. If needed, *hinge* atoms can be added to the main medial atom mesh, so as to represent protrusions or intrusions on the original shape (fig. **??**). In this manner, complex objects can be represented as a collection of connected sub-objects.

1.4 Distance transforms

A shape descriptor conceptually related to the medial axis transform, also with reference to the object surface, is the distance transform [Leventon et al. 2000, Golland 2000]. The distance transform is a function of the original object voxels, where each voxel is assigned a scalar equal to the distance of the given voxel to the surface. Hence surface points are by definition assigned a value equal to zero. Points outside the boundary are assigned a negative distance. Like the MAT, distance transforms can be easily computed given a binary map of the target object.

1.5 Landmark-based representation

Landmark-based methods represent shape using a set of discrete points scattered over the target shape in a specific manner. These points can be placed along the shape boundary, such as in a point-distribution model (PDM) [Cootes & Taylor 1995]. Using an interpolant between these points produces the continuous original surface. We can alternatively have a full net of control points distributed across the whole volume of the target shape [Cootes et al. 2008]. Again, exact positions for the points in-between the control points is computed using interpolation.

Proper positions for the landmarks can be determined manually, to identify special locations [Bookstein 1996]. Such locations can be for example the position of the fingertips in a set of human hand shapes, the position of bone connections in a set of body shapes, or the position of the hippocampi or other structures in a set of brain MR images. However, manual positioning of the landmarks can be a very challenging and time-consuming task, especially in 3D. Methods exist therefore for automatic placement of landmarks, exploiting the semantical value of geometrical properties such as curvature [Pennec et al. 2000]. In [Cootes et al. 2008] grids of control points are initialized to a uniform distribution across the shape area, in a multiresolution manner. Finding the correct positions for the control points in this case, when done with respect to a specific template, is essentially equivalent to non-rigid image registration (which we shall see in more detail below). Group-wise landmark position optimization has been considered in [Davies et al. 2001, Davies et al. 2008], determining correspondence between and in the context of a specific group of shapes.

1.6 Deformation vector fields with respect to a template

The representation of this type is based on non-rigid matching of a template shape to an input shape. The resulting registration produces a vector field, that is one vector per original shape voxel. This vector field forms a dense correspondence between the template and the input shape. Numerous models have been proposed which focus on producing a correspondence mapping that is diffeomorphic [Beg et al. 2005, Cootes et al. 2008, Vercauteren et al. 2008, Vercauteren et al. 2009]. A diffeomorphism is a one-to-one mapping f, where both f and its inverse f^{-1} are differentiable. In practice, this means that in matching the template to the input shape no tearing or folding is allowed, which is a desirable trait when it comes to dealing with anatomical data.

Given a training set of a number of shapes, using a single template provides a common coordinate frame for the shape descriptors. The shape that plays the role of the template can be picked out of the training set, or can be a pre-computed standard image. Choosing a single template shape can bias the representation, and methods that create multiple templates have been proposed [Blezek & Miller 2007, Sabuncu et al. 2009, Sfikas et al. 2010A]. However, there is no straightforward way to use a multiple template collection as a common coordinate frame, to the end of representing shape in our case.

Deformation vector fields have been widely used in modeling medical imaging anatomical data (whole 3D brain volume [Ashburner et al. 1998], cortical surfaces [Thompson et al. 2000], hippocampus [Joshi et al. 1997], and cingulate gyrus [Csernansky et al. 2004]).

2. Feature vector space as a model basis

Building a feature vector given a specific shape model can be done by simply concatenating model parameters into a single vector. Each variate contains a different model parameter, with care taken for the concatenation to be always in the same order when we are dealing with

modeling a set of multiple shapes.

Let $\boldsymbol{y} \in \mathbb{R}^D$ be the produced feature vector, assuming D parameters, and in the general case where all parameters are real scalars. This representation immediately implies a space of dimensionality D where feature vectors reside. Further statistical analysis can then be performed in terms of this space.

3. Conclusion

In this thesis we shall use deformation vector fields with respect to a template to describe shape. We use Diffeomorphic demons [Vercauteren et al. 2008] to register shapes one to another and produce the required deformation fields. [1] The template is chosen as the shape with the minimum sum of distances to the rest of the shapes, where shape distance is defined as in [Gerber et al. 2009] (see also chapter 3 for a related discussion of shape distance).

Using a deformation field for shape description –and any of the methods described in this section, up to a certain degree– in spite of its usefulness, produces a representation that is in general non-compact, and a feature vector space that is non-dense (or non-specific, using the terminology of [Davies et al. 2008]). In this sense, our current choice of shape representation is a component that could be replaced with any of the other alternative presented in this chapter, leaving the rationale of the proposed pipeline and model intact. In the next chapter we shall discuss how to produce a compact representation of shape, and its relation with its non-compact peer.

1. We have used and adapted code from http ://www.insight-journal.org/browse/publication/644 and http ://darwin.bio.uci.edu/ cestark/roial/Distribution/source/DemonsRegistration.cxx .

Chapter 3

Compact representation with manifold learning

In the previous chapter, we discussed shape representation, and we reviewed a number of related parametric models. However, regardless of the specific model of parametric representation, the following contradiction typically arises : on the one hand, the more parameters the model uses, in general the higher the fidelity to the actual shape ; for example, a 256^3 voxel grid would allow us much more detail for a three-dimensional shape than a 128^3 voxel grid. On the other hand, data variability is represented by a very low number of degrees of freedom. For example, a set of geometrical objects that would differ only in their size would have exactly 1 (one) degree of freedom, corresponding to scale. As another example, let us consider a dataset of structural brain MR images. A generic shape description model would fail to capture *a priori* the intrinsic structure of any brain MR image, such as the fact that the external voxels correspond to skull voxels, or that the brain is quasi-symmetric in structure and the two hemispheres are joined by the corpus callosum, etc. In model terms, a representation with a large disparity between the number of possible model parameters and the number of actual degrees of freedom is not *compact*.

A non-compact representation can be problematic in a number of ways. Aside from the obvious fact that a non-compact model is by definition more "expensive" than it should, high-dimensional models are characterized by what is known as the *curse of dimensionality* in the literature [Bellman 1961]. The curse of dimensionality, in brief, expresses that in order to build useful models over a given space, a number of training data that is in exponential relation to the number of the space dimensions is required. If not, then any model, be it discriminative or generative, built on the high-dimensional space in question runs a serious risk of having very poor inductive properties.

In anatomical population analysis, training sets are at the very best in the order of 1000 subjects [Jack Jr. et al. 2008], so a means to cope with the high dimensionality of the data is required. Taking advantage of *a priori* information on a given pathology can and has been used. Examples of such information include extracting and working with particular tissues and brain regions of interest that are known to be the most affected, instead of working with the

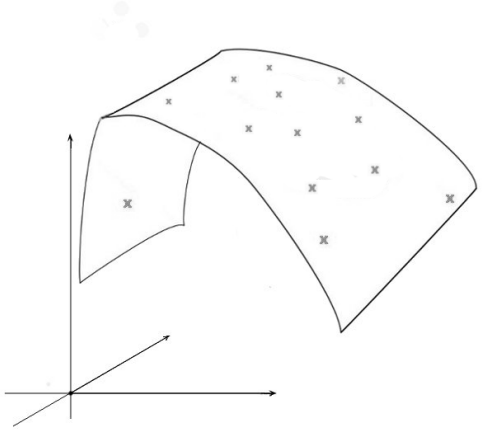


FIGURE 3.1 – Illustration of the manifold data structure hypothesis. The data are formally vectors in a high-dimensional space –the "ambient space"–, yet they can only cover a very small part of it –the "embedding". Here, for the purpose of visualization the ambient space is 3-dimensional and the embedding is 2-dimensional. In practice, the differences of dimensionality between the ambient space and the embedding can be several orders of magnitude greater.

whole brain. Feature selection methods [FAN et al. 2007, MORRA et al. 2010] and recently foremost *dimensionality reduction* or *manifold learning* methods [GERBER et al. 2010, PLESS & SOUVENIR 2009, TENENBAUM et al. 2000, WOLZ et al. 2010] are used to the end of remedying the undesirable effects associated with working with high-dimensional, sparsely populated spaces. In this thesis we shall work with the latter solution, that of manifold learning.

In the following sections of this chapter, we first see the basic hypotheses underlying manifold learning. We discuss the important issue of defining a suitable distance metric, that will capture manifold topology. We choose the distance metric we shall use in this work, and see how to estimate the manifold with Isomap.

1. Manifold hypothesis

In manifold learning methods, we postulate that the space of possible shapes forms a low-dimensional continuous manifold, occupying only a small volume fraction of the containing space. In other words, data are assumed to lie on or close to a manifold. We shall refer to this assumption as the "manifold hypothesis" (see fig. 3.1). This idea has been introduced originally for natural images [DONOHO & GRIMES 2006, LU et al. 1998]. The parametrization of the low-dimensional embedding, or *intrinsic* parametrization, is often found to have a very intuitive character [LU et al. 1998, PLESS & SOUVENIR 2009, TENENBAUM et al. 2000, THAYANANTHAN et al. 2006, VIK et al. 2007] such as lighting, pose or scale of the imaged object. Principal Component Analysis (PCA) [BISHOP 2006] in the context of image analysis (e.g. [VIK et al. 2007], or the Eigenfaces model [TURK & PENTLAND 1991] among others) can be considered as a model adhering to the image manifold hypothesis, in its most simple form. The resulting eigenanalysis bases of PCA form a subspace, that is clearly a smooth manifold, constrained to be flat everywhere. In real problems however, the flatness / linearity constraint can easily prove to be an inadequate assumption : in [DONOHO & GRIMES 2006], a set of simple artificial examples is presented, where the articulated image produces curved manifolds even in the case of articulations as simple as a rigid translation ; in [LU et al. 1998], sets of face images are studied, and found to form manifolds which have a curvature too high to be considered linear. The

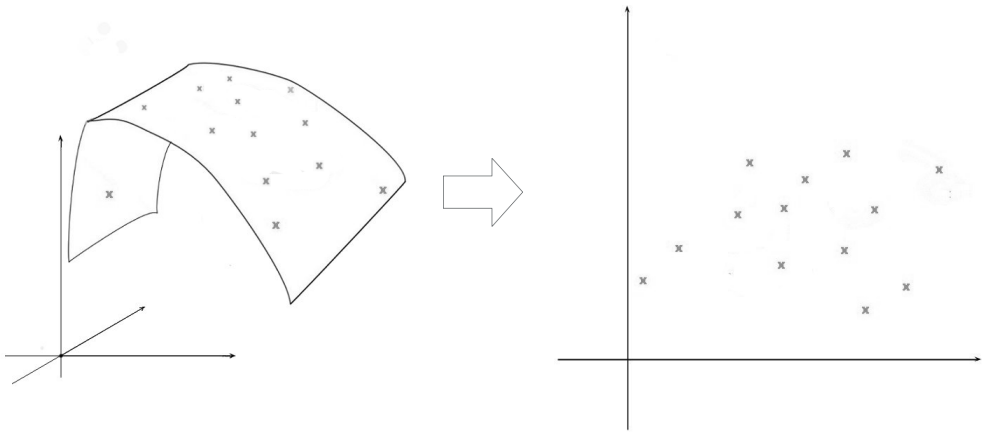


FIGURE 3.2 – Illustration of manifold learning. Geodesic distances between point pairs in the ambient space (left) are preserved as Euclidean distances between images of the same points in the reduced space (right).

application of the manifold hypothesis and learning covers a wide field spectrum, which includes tracking [NASCIMENTO & SILVA 2010], content-based image retrieval [SPARKS & MADABHUSHI 2011], image registration [WACHINGER & NAVAB 2010], pose estimation [THAYANANTHAN et al. 2006, PTUCHA et al. 2011]. In the context of medical imaging, low-dimensional representations of medical data have been used including both linear [VIK et al. 2007, ZHOU et al. 2011] and nonlinear models [CAO et al. 2011, FLETCHER 2004, GERBER et al. 2010, HAMM et al. 2010, WACHINGER et al. 2010, WACHINGER & NAVAB 2010, WOLZ et al. 2010, WOLZ et al. 2010, ZHANG et al. 2006]. The reader may also refer to the related section of chapter 7 for a summary of applications of the manifold hypothesis in medical imaging.

2. Manifold learning

We assume an input set Y of N shapes, under a common representation framework. In this work, we have used deformation fields to describe shape for our experiments (see chapter 2). All members of $Y = \{y_n\}_{n=1}^{N}$ reside in a space of dimensionality D, which we shall call the ambient space $\mathcal{Y}$. Formally, the *manifold hypothesis* consists in the assertion that the same elements are at the same time constrained to lie on a continuous manifold $\mathcal{X} \subset \mathbb{R}^d$, $d \ll D$, embedded in the containing, ambient space $\mathcal{Y}$. We consider all points / shapes $y \in \mathcal{Y}$ to be related to $x \in \mathcal{X}$ as

$$y = g(x) + noise, \tag{3.1}$$

where g defines a mapping $g : \mathcal{X} \to \mathcal{Y}$ and x corresponds to the *intrinsic* parametrization or coordinates of datum y over the embedding $\mathcal{X}$. The noise parameter expresses a (small) margin of model misrepresentation. We are not interested in a more precise form for the mapping g for the moment, as our goal here is to obtain the compact representation for our shape set, that is pass from y_n to x_n. We discuss again the mapping g in chapter 6.

We can learn the values $x_n \in X$ for our observation set Y using manifold learning techniques [PLESS & SOUVENIR 2009, VAN DER MAATEN et al. 2009]. Numerous algorithms have been proposed to learn a nonlinear manifold such as locally linear embedding [ROWEIS & SAUL 2000], Hessian eigenmaps [DONOHO & GRIMES 2003], semidefinite embedding [WEINBERGER & SAUL 2006], Laplacian eigenmaps [BELKIN & NIYOGI 2003] and Isomap [TENENBAUM et al. 2000]. Isomap remains one of the most widely used manifold learning methods, and it is the algorithm we use in this thesis.

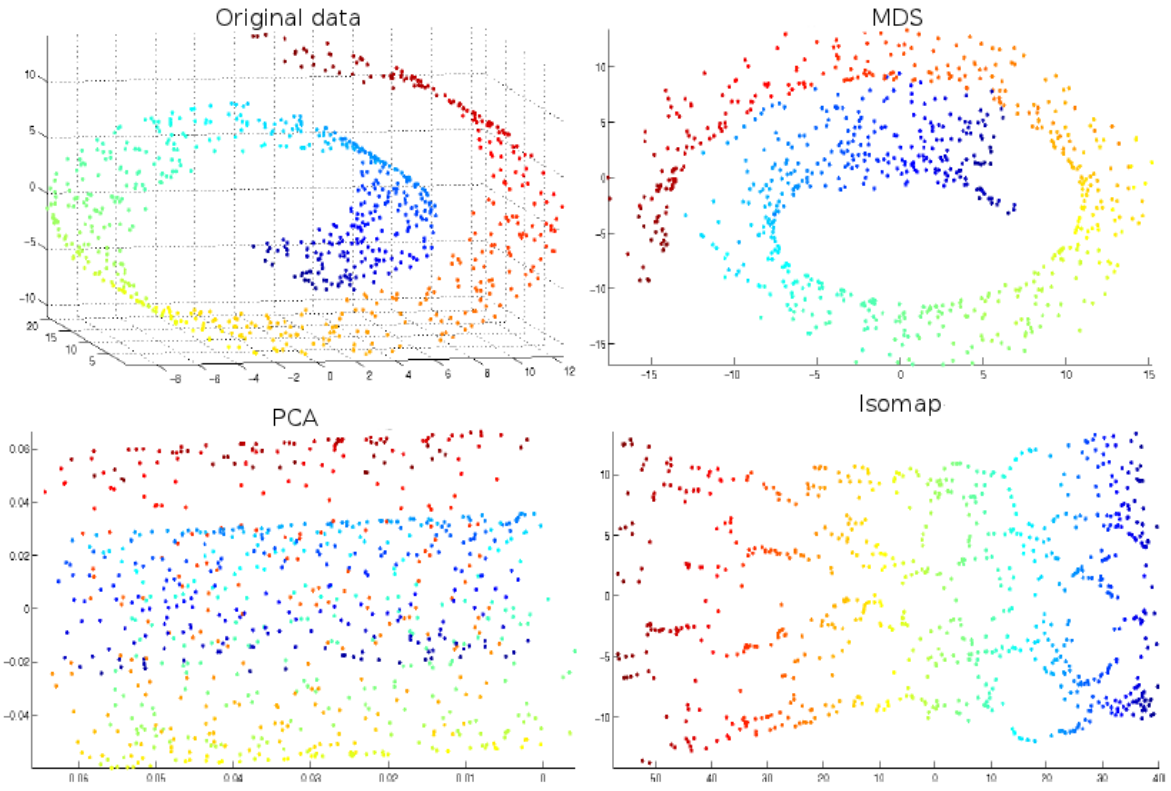


FIGURE 3.3 – The swiss roll dataset. Results of various manifold learning algorithms. Only Isomap captures the data structure correctly. Image produced using software in http ://www.math.ucla.edu/ wittman/edu

2.1 Learning the manifold with Isomap

Isomap works by computing an isometric mapping of a dataset. In other words, the Euclidean distances between all points of the output set $X = \{x_n\}_{n=1}^N$ are constrained to be equal (or reasonably close) to geodesic distances taken as input, for all corresponding interpoint distances in Y (see fig. 3.2). Geodesic distance is defined as the length of the shortest path that links two points y_a, y_b and all points of this path are part of the embedded manifold. The input in Isomap can be regarded as a $N \times N$ matrix D with $D_{ij} = d(y_i, y_j)$, where $d(\cdot, \cdot)$ is the input distance operator. Based on this matrix, Isomap constructs k-nearest neighbor neighborhoods and computes shortest paths over the new induced graph, where each point is a vertex forming edges with its neighbors (Floyd-Dijkstra algorithm). This produces the output geodesic distance matrix D_G. Isomap then relies on the Multi-dimensional scaling (MDS) algorithm [COX et al. 2000]. MDS determines the new points $\{x_n\}_{n=1}^N$, constraining them to miminize the cost

$$\underset{x_1, x_2, \cdots, x_N}{\operatorname{argmin}} \frac{\sum_{n=1}^N \sum_{m=1, m\neq n}^N (\|x_n - x_m\|_2 - D_{G^{nm}})^2}{\sum_{n=1}^N \sum_{m=1, m\neq n}^N D_{G^{nm}}^2}. \tag{3.2}$$

Note that in classical MDS, the original distances D are used in place of D^G. See fig. 3.3 for an illustration of using Isomap, MDS and the linear PCA on a test dataset.

2.2 Distance metrics

The hypothesized distance metric between data is a defining aspect of the induced embedding. The distance metric can be used to induce topological properties for a given space [Munkres 2000]. Otherwise, moving over the manifold has the meaning of infinitesimal deformations that gradually transform shape y_a to shape y_b. Hence, the manifold geodesic distances are expected to mesure intuitive distance between shapes. Different metrics are expected to capture different notions of shape difference, and prioritize different aspects of shape as more important and more responsible for shape difference.

In this work we have used a *deformation field induced distance* [Gerber et al. 2009, Aljabar et al. 2010, Vercauteren et al. 2009] based on the estimation of a diffeomorphic mapping between the two shapes or images. Diffeomorphisms, i.e. transforms that are everywhere smooth and invertible, are widely used in computational anatomy [Beg et al. 2005, Trouvé 1998, Durrleman et al. 2008] to assess shape difference. In the widely used LDDMM framework [Trouvé 1998], a flow of diffeomorphic transformations is constructed on coordinates of the shapes to be matched. Here we follow the related framework and discussion as is posed in [Gerber et al. 2009, Gerber et al. 2010], and consider the diffeomorphism in terms of a time-varying vector field. We denote the diffeomorphic coordinate transformation as $\phi(x, 1)$, where $\phi(x, t) = x + \int_0^t u(\phi(x, \tau), \tau) d\tau)$ and $u(x, t)$ is a smooth, time-varying vector field. A Riemannian metric can be defined on the space of diffeomorphic transformations [Gerber et al. 2010] as

$$d(y_i, y_j) = min_u \int_0^1 \|u(x, \tau)\|_L \, d\tau$$

$$s.t. \int_\Omega \|y_i(x + \phi(x, 1)) - y_j(x))\|_2^2 \, dx = 0$$

The metric prioritizes the possible mappings, and with an appropriate choice of the differential operator L, can ensure smoothness. For two shapes that are relatively similar to one another, we can approximate the coordinate transform by a single vector field, because the velocities of the geodesics are smooth in time [Younes et al. 2009]. Formally we have

$$\phi(x, 1) \approx u(x, 0) = v(x), \text{and } d(y_i, y_j) \approx min_v \, \|v(x)\|_L \tag{3.3}$$

subject always to the condition that the diffeomorphism matches well enough the two shapes of the pair to be matched. Assuming that the distance between the given shapes is small is legitimate, since the manifold learning layer is based on distances between nearby samples (larger distances are eventually disregarded anyway, due to the isomap Floyd-Dijkstra step).

We again follow [Gerber et al. 2010, Aljabar et al. 2010] and define the differential operator L as $\alpha\nabla + (1 - \alpha)Id$. Finally the formula for our distance in eq. (3.3) becomes

$$d(y_i, y_j) = \int_\Omega \alpha \, \|\nabla v(x)\|^2 + (1 - \alpha) \, \|v(x)\|^2 \, dx. \tag{3.4}$$

where we set $\alpha = 0.9$ [Gerber et al. 2010]. [1]. In practice, a symmetrized version of the resulting distance matrix is fed to Isomap. Note that it is an open problem to show that the resulting distance metric satisfies the triangle inequality and hence is a metric or a semimetric [Gerber et al. 2010].

There are several other options for distance metrics in use in the literature in similar contexts. A related, recently proposed framework is that of *currents* [Durrleman et al. 2008, Vaillant & Glaunes 2005]. Under this framework, shapes are represented as mathematical objects

1. We have used Diffeomorphic demons [Vercauteren et al. 2008] to register shapes and produce the required deformation fields. We also used the same method in section 2 to create shape descriptions.

that encode normal vectors of surfaces, which are in turn elements of a vector space equipped with a computable distance. In this manner, a matching is performed not only in terms of extrinsic volume coordinates, but also based on intrinsic values as are the surface normal vectors, resulting theoretically to a more elaborate result.

Possibly the most inexpensive alternative would be a *voxel-to-voxel distance*. As reported in the literature [Gerber et al. 2009, Souvenir & Pless 2007, Twining & Marsland 2008], a direct application of a measure such as a L_1 or a L_2 distance is in general not the most suitable to capture shape differences. However, measured over pairs of (usually affinely or rigidly) registered images, it has given useful results in a number of related medical imaging applications [Blezek & Miller 2007, Wolz et al. 2010, Wolz et al. 2010]. Other distance metrics used for comparing shape include normalized mutual information [Studholme et al. 1999, Wolz et al. 2010], normalized cross-correlation [Aljabar et al. 2010] and the Hausdorff distance [Souvenir & Pless 2007, Thayananthan et al. 2006].

3. Conclusion

Under the *manifold hypothesis*, the duality *ambient* space - *reduced* space (or ambient space - embedding) is introduced. On the one hand, we have the ambient space which is directly observable (usually) but high-dimensional and sparsely populated, while on the other hand the reduced space is latent but low-dimensional and densely populated, in the sense that ideally *all* of its points correspond to a possible discrete observation. The spirit that characterizes all cases, regardless of application, is that provided the data do indeed lie in a low-dimensional embedding, one can produce a model that represents the data more faithfully and compactly than models that neglect this intrinsic structure. Classification, clustering, anomaly detection, population comparison can thence be performed over the dimension-reduced data instead of the original space input. To the extent that the dimension-reduced space is free of the *curse of dimensionality*, the above operations will be performed in a more robust and reliable fashion than without dimensionality reduction.

One of the main focuses of this thesis is to build a discriminative model capable of tracking the exact structural difference that is responsible of the model class difference. We discuss this point in the next part of this thesis, where we also show that compactness proves to be a crucial trait for the achievement of our claims, through the proposed model.

Part II

Discriminative and generative shape models

Introduction

In this part we discuss the discriminative and generative components of the model, and we examine how to put them together as a working processing pipeline. Our model is built on the compact representation layer, which was presented in the previous part. The discriminative model is essentially a SVDD, one-class machine, built over the reduced dimension embedding previously discussed. Aside from providing a decision function which tells us which part of the shape space can be considered "normal" or "abnormal", the SVDD model can also allow us to estimate the so-called discriminative difference or discriminative pattern between class members and class non-members. In the final section of this part, we discuss how we can generate new shapes with regression, and review all the presented model components as a whole. We use datasets of geometric and synthetic anatomical shapes, where we train our model and show that its compact, non-linear qualities –stemming from the fact that it deals inherently with shape as a low-dimensional, non-linear manifold structure– render it more powerful. We illustrate this point by showing that the proposed discriminative trajectory, as part of the present model, reveals discriminative difference in detail that would not be achieved by a non-compact model.

Chapter 4

Discriminative shape models using SVDD

The basic premise of any discriminative model is as follows : given a number of objects, partitioned into a number of classes, infer class membership for a new unseen object. In this thesis we focus on one such discriminative model, the support vector data description (SVDD) model [Sakla et al. 2011, Sjöstrand et al. 2007, Tax & Duin 1999, Tax & Duin 2004]. In the shape analysis context, the role of a discriminative model –such as SVDD– is, simply put, to tell if a target shape can be categorized as part of a learned class of shapes, or not. In forthcoming sections we shall examine applications of SVDD other than classification.

Even if SVDD is based on a one-class versus-all philosophy, it can in principle be extended to a multi-class model [Lee & Lee 2007, Yang et al. 2011]. In [Lee & Lee 2007] for example, a different support for each class is trained. Then, a classifier is built using Bayes optimal decision theory with the class-conditional densities being approximated using the SVDD boundaries. The resulting model competes fairly with other classification techniques, with similar or better results.

In what follows we shall present elements of the SVDD framework, and then briefly review some models closely related to SVDD from a theoretical standpoint.

1. Support vector data description theory elements

Support vector data description (SVDD) is a one-class support vector learning method [Tax & Duin 2004] used for novelty / outlier detection. The main idea in SVDD is to approximate the support of data *a priori* tagged as *normal* using a hypersphere. Thence the support is used typically for classification of new unseen data as normal or abnormal, depending if they happen to lie respectively inside or outside the SVDD support. In the context discussed here for example, the healthy normal control group plays the role of the SVDD normal class.

Consider a hypersphere with center a and radius R, and the normal, training data set X consisting of objects $\{x_n\}_{n=1}^{M} \in \mathcal{C}$ with $\mathcal{C} \subset \mathbb{R}^d$. Consider also measures ξ_n for each datum ;

these measures form a penalty term for data lying outside the hypersphere boundary. In SVDD learning, we wish to minimize the SVDD hypersphere radius R while keeping the ξ_n penalties in check. Formally this is expressed as solving

$$\arg\min_{R,a,\xi_n} R^2 + C\sum_{n=1}^{M}\xi_n,$$

$$s.t. \quad \|x_n - a\|^2 \leqslant R^2 + \xi_n, \qquad \xi_n \geqslant 0, \qquad n = 1, \cdots, M, \tag{4.1}$$

where C is a parameter controlling the trade-off between the relative importance of minimizing the hypersphere radius R, versus minimizing the penalties ξ_n. Optimally $\xi_n = \|x_n - a\|^2 - R^2$ for data outside the hypersphere of radius R and center a, and $\xi_n = 0$ for the rest. Larger values for C result in greater hypersphere radius R, and larger estimated support area. In practice, parameters C and slack variables ξ_n allow for a number of the training set to be finally labelled as outliers, in the interest of computing a more compact support.

The optimization problem of eq. (4.1) can be reformulated in dual form [Park et al. 2007, Tax & Duin 2004] as

$$\arg\max_{\alpha} \sum_{n=1}^{M}\alpha_n x_n^T x_n - \sum_{n=1}^{M}\sum_{m=1}^{M}\alpha_n\alpha_m x_n^T x_m,$$

$$s.t. \quad \sum_{n=1}^{M}\alpha_n = 1, \qquad \alpha_n \in [0, C], \qquad n = 1, \cdots, M, \tag{4.2}$$

where α_n are M Lagrange multipliers, $\alpha = [\alpha_1\alpha_2\cdots\alpha_M]$. Eq. (4.2) is a quadratic optimization problem with linear constraints. It can be solved using interior point methods [Wright 2005]. Since its objective function is strictly concave, it has a unique solution.

The basic assumption of SVDD in the form introduced here, that is that the support to be approximated has the form of a hypersphere, is quite limiting. The hypersphere form constraint can be relaxed by means of mapping the original data to an auxiliary space $\mathcal{F}$, called feature space. This mapping is typically nonlinear. The solution is obtained using the so-called *kernel trick* or *kernel substitution* [Bishop 2006, Shawe-Taylor & Cristianini 2004, Lemm et al. 2011]. The kernel trick, most notably also used in SVM, is based on the observation that the input data appear in the problem formulation exclusively through inner products. These inner products can be replaced by a corresponding *kernel function* between the same input data. Kernel functions are functions of the form $k : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, satisfying the condition that for any choice of points X in $\mathbb{R}^d$ the matrix K formed as $K_{nm} = k(x_n, x_m)$, with $x_n, x_m \in X$, must be positive semi-definite. Hence in this notation, we have $k(x_n, x_m) = \phi(x_n)^T\phi(x_m)$. The replacement of inner products by valid kernel functions -valid in the sense that they satisfy the semi-definite positivity condition- implies that the transformation $x \mapsto \phi(x)$ [Bishop 2006] does not need to be defined explicitly in practice. Equation (4.2) presents the trivial case, termed linear kernel in the literature, where $\phi(x) = x$ and $k(x_n, x_m) = x_n^T x_m$. Using kernel choices other than the linear kernel leads practically to a SVDD support form that is more flexible than the standard hypersphere. In fact, the support does not in this case form a hypersphere in the original space but it does form a hypersphere in feature space $\mathcal{F}$, where the function ϕ maps the input x to. A hypersphere in the feature space produces an arbitrary form in the original space $\mathfrak{X}$, this form depending on the choice of $k(.,.)$, and consequently on $\phi(.)$.

In this work, we shall assume a Gaussian kernel function, $k(x_n, x_m) = \exp(-d_{\mathcal{C}}^2/\sigma^2)$, with $d_{\mathcal{C}}$ the distance metric in the original space $\mathcal{C}$, typically the Euclidean $d_{\mathcal{C}} = \|x_n - x_m\|$. The notation $k(.,.)$ will refer to the Gaussian kernel unless otherwise stated. The Gaussian kernel is used

extensively already in contexts other than SVDD, most notably in support vector machines. For kernel bandwidth $\sigma \to +\infty$, the SVDD solution tends to the solution under the linear kernel [Tax & Duin 1999], so that the Gaussian kernel can also be viewed as a convenient generalization of the trivial linear kernel. Mapping using a nonlinear kernel comes in handy versus a linear kernel when the number of observations is much larger than the number of data dimensions [Hsu et al. 2010], which is the case with the reduced-dimension data we shall use. The dual problem of SVDD is reformulated [Park et al. 2007] as

$$\arg\min_{\alpha} \sum_{n=1}^{M}\sum_{m=1}^{M} \alpha_n \alpha_m k(x_n, x_m),$$

$$s.t. \quad \sum_{n=1}^{M} \alpha_n = 1, \qquad \alpha_n \in [0, C], \qquad n = 1, \cdots, M, \tag{4.3}$$

where $k(x_n, x_n) = 1$ for the RBF (here Gaussian) kernel. After optimizing over α, the hypersphere radius that abides to the constraint expressed in eq. (4.1) is given by

$$R^2 = 1 - 2\sum_{n=1}^{M} \alpha_n k(x_n, x_s) + \sum_{n=1}^{M}\sum_{m=1}^{M} \alpha_n \alpha_m k(x_n, x_m), \tag{4.4}$$

where we used $d_{\mathcal{C}}(x, x) = 0$, which must hold for any true metric $d_{\mathcal{C}}$. Also, x_s stands for any support vector lying on the hypersphere surface. The SVDD decision function [Sakla et al. 2011, Tax & Duin 1999] for an unseen point x is

$$f_{SVDD}(x) = R^2 - \|\phi(x) - \sum_{n=1}^{M} \alpha_n \phi(x_n)\|^2 =$$

$$\sum_{n=1}^{M} \alpha_n k(x_n, x) - \frac{1 + \sum_{n=1}^{M}\sum_{m=1}^{M} \alpha_n \alpha_m k(x_n, x_m) - R^2}{2}. \tag{4.5}$$

Non-negative values of f_{SVDD} correspond to points estimated to be part of the support, or normal. Negative values are considered lying outside the support, or outliers. See fig. 4.1 for an illustration of support estimation with SVDD.

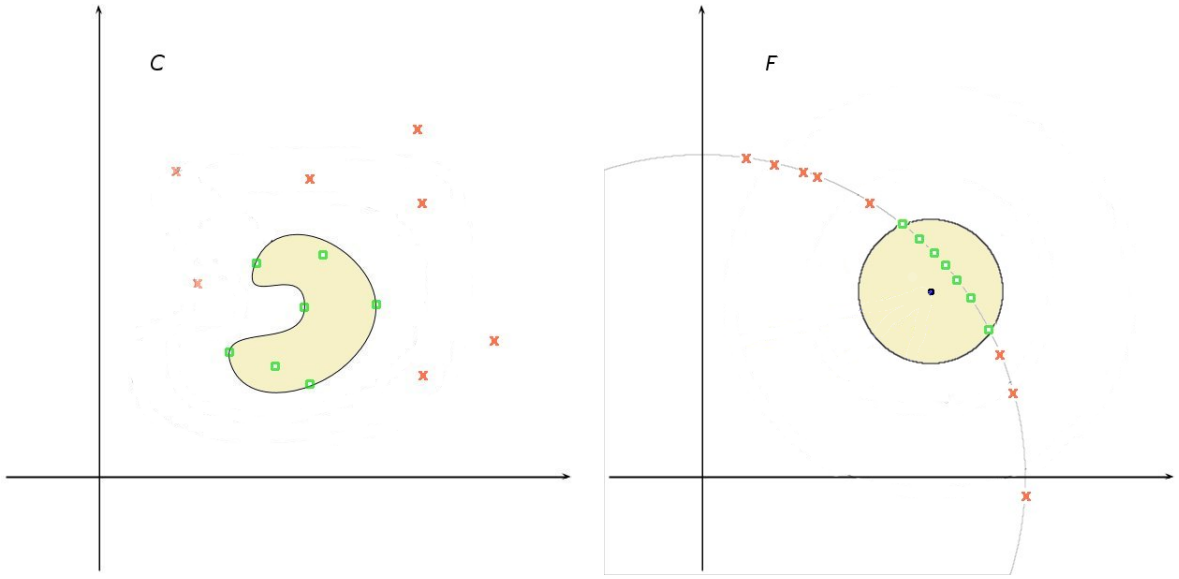
2. Related models

2.1 Support vector machines

Closely related to the SVDD, and historically much prior to the SVDD, is the celebrated Support Vector Machine (SVM) [Cortes & Vapnik 1995]. The SVM assumes two classes. These classes are divided by a hyperplane in the feature space. The dividing hyperplane in the SVM is essentially the counterpart of the hypersphere as used in the SVDD.

Let X be the training set consisting of objects $\{x_n\}_{n=1}^{M} \in \mathcal{C}$ with $\mathcal{C} \subset \mathbb{R}^d$. *A priori* we also know the set of labels $\{y_n\}_{n=1}^{M} \in \{-1, +1\}$, with $y_n = +1$ if datum x_n belongs to class A, and $y_n = -1$ if datum x_n belongs to class B. The training of SVM seeks to find a hyperplane that separates the data, under two constraints. The first constraint seeks to reward separating hyperplanes with a large margin –where margin is the distance from the closest datum to the separating hyperplane. The second constaint seeks to minimize the number of wrongly classified training data. Formally we have

$$\arg\min_{w, \xi_n} \frac{1}{2} \|w\| + C\sum_{n=1}^{M} \xi_n,$$



(a) (b)

FIGURE 4.1 – Support estimation with SVDD. Green circles are normal data, red crosses are abnormal data. The normal data support corresponds to the beige-color filled area. (a) The original (shape) space. (b) The feature space. The support boundary forms a hypersphere in the feature space. This translates to a nonlinear structure in the original space.

$$s.t. \quad y_n(w^T x_n - b) \geqslant 1 - \xi_n, \qquad n = 1, \cdots, M, \tag{4.6}$$

where w is a vector controlling the margin of the hyperplane, $1/\ \|w\|$ is the margin of the hyperplane, b is a scalar bias term, and $[\xi_1 \xi_2 \cdots \xi_M]$ are slack variables. These slack variables allow for the corresponding data to be classified on the wrong side of the hyperplane, if need be. C is a user-defined parameter controlling the tradeoff between the two optimization goals, that is maximization of the margin and minimization of the number of wrongly classified data.

The problem as stated in eq. (4.6) can be reformulated in its dual as follows [CORTES & VAPNIK 1995, BISHOP 2006]

$$\arg\max_{\alpha,\xi_n} \sum_{n=1}^{M} \alpha_n - \frac{1}{2} \sum_{n=1}^{M} \sum_{m=1}^{M} \alpha_n \alpha_m y_n y_m x_n^T x_m,$$

$$s.t. \quad \xi_n \geqslant 0 \qquad n = 1, \cdots, M. \tag{4.7}$$

Note the similarities of the above with the SVDD problem formulation (section 1. of this chapter). The problem formulation allows use of the kernel trick here as well. The SVM decision function is given by

$$f_{SVM}(x) = \sum_{n=1}^{M} \alpha_n y_n k(x_n, x) + \frac{\sum_{n \in S} y_n - \sum_{n=1}^{M} \sum_{m=1}^{M} \alpha_m y_m k(x_n, x_m)}{|S|}, \tag{4.8}$$

where S is the set of support vectors ($\alpha_n = 0$). The interested reader is referred to [CORTES & VAPNIK 1995] or [BISHOP 2006, SHAWE-TAYLOR & CRISTIANINI 2004] for more details on support vector machine theory and applications.

2.2 One-class support vector machines

At about the same time with SVDD, the one-class support vector machine (OCSVM) was introduced [Schölkopf et al. 2001]. Intuitively, the object separating normal examples from abnormal in OCSVM is a hyperplane, like in SVM, which is set to separate input data points from the origin. The solution of SVDD and that of OCSVM are shown to be equivalent, when the chosen kernel is a radial basis function [Sjöstrand et al. 2007], such as the Gaussian kernel.

3. SVDD in shape analysis

Let us put the SVDD discussion in the more special context of analyzing shapes. First, we need to define the original space $\mathcal{C}$ as a space representing shapes. We have already reviewed a number of ways to describe shape in section 2.

Assuming that small distances correspond to similar shapes, then the SVDD support, as it defines a continuous area characterized by small distance to all or part of the training set, defines equivalently the set of shapes considered as close enough to the input to be considered of a similar enough shape. For example, a shape with a deformity that does not appear in any of the shapes of the training set would lie outside the SVDD support, and thus be labelled abnormal. Likewise, in a neurodegenerative disease setting, a tissue with a deformity (or deformities) not apparent in the control group should be tagged as pathological. This is the case for the hippocampus dataset that we shall examine in chapter 8. The idea of using shape as a disease biomarker [Tarawneh & Holtzman 2010], that is to classify subjects according to their progress with respect to a given neurodegenerative pathology, has been recently extensively used and is an active field of research (see for example [Cuingnet et al. 2011]).

4. Conclusion

In this chapter we have seen the formal definition of the SVDD, the SVM and the one-class SVM. The SVDD is a one-class classifier. Our focus in this work is not classification in itself, as much as using the induced discriminative model to extract information about what kind of shape difference is responsible for the class difference. We review this in the following chapter.

Chapter 5

Discriminative trajectory

In this chapter, we shall discuss the following question : given a set of shape data, what is the type of shape difference that discriminates best between objects that belong to a given class and objects that do not ? The answer implies the existence of a discriminative model trained on the data, which however does not answer the question by itself. We shall present a novel solution to the problem, that is based on SVDD and which we call *discriminative trajectory*. Related work, where the concept was originally posed, is based on the 2-class SVM. Golland et al. [Golland 2000, Golland 2001, Golland et al. 2005] proposed the SVM-based *discriminative direction*, and other related proposals have followed [Koutsouleris et al. 2009, Zhou et al. 2008].

In a medical context, the question is of special importance, as object class may refer to the object (organ, tissue) and associated subject being diagnosed with a certain pathology or not. Thus, looking for the discriminative shape difference in this case can unveil details of pathology development in terms of structural change.

1. Estimating discriminative difference with SVDD

In this section, we introduce a method to estimate discriminative difference, suitable for SVDD models. This estimate comes in the form of a *discriminative trajectory*, which is defined explicitly as a process of shapes, given an initial shape x. Each moment of the process is in turn defined as a projection on a specific data support trained using SVDD. We must hence first discuss the definition of this special projection, to which we shall refer as SVDD projection.

1.1 SVDD projection

In [Park et al. 2007], the SVDD theory is used in a denoising application. Given a noisy pattern x and an SVDD ball trained on a set of normal, noiseless data, the closest point of the SVDD support to the noisy pattern will be a denoised version of x. We shall call this point the

SVDD projection of x and represent it by $p_R(x)$, where R is the optimal hypersphere radius (eq. (4.4)). Formally, given the space $\mathcal{C}$ where data reside, the SVDD projection can be thought of as a mapping $p_R : \mathcal{C} \to \mathcal{C}$. Based on the idea of the SVDD projection for denoising, we use the same theoretical tool in the context of shape analysis. Principally, the difference with denoising is that the noise that is to be eliminated is identified here as shape features that are uncharacteristic of normal data.

For the sake of illustration let us consider a set of geometric shapes defined as normal, for example the set of all geometrical solids having exactly five faces which would include any pentahedron including a normal pyramid (see fig. 5.1). Consider a pyramid with a sawed-off top. Given a pertinent topology and metric for $\mathcal{C}$, this object would be considered abnormal, since it has six faces. The projected image of this new object, so to speak, should be a pyramid, since as a shape, compared with the other shapes in the normal set, the pyramid is perceptually closest to the sawed-off top pyramid in comparison to other "normal" structures, such as a five-face prism.

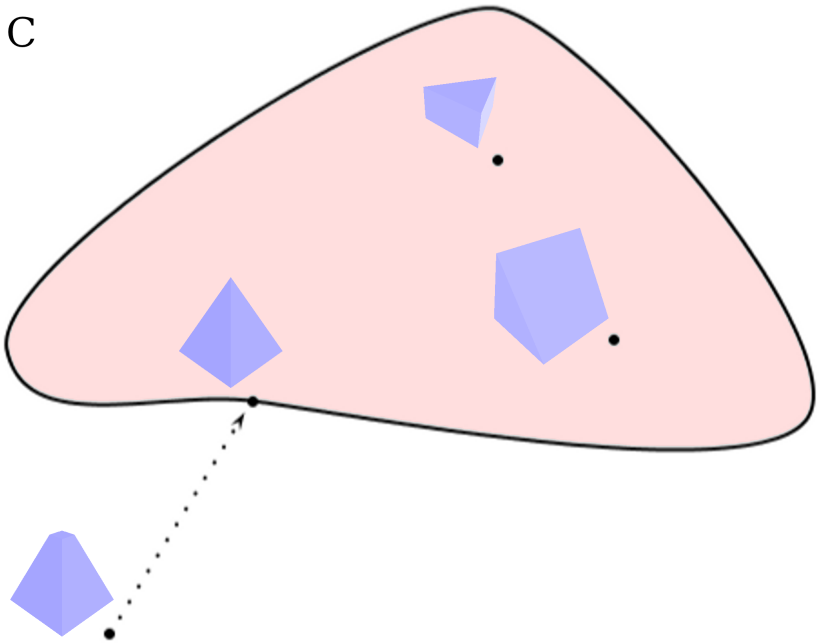


FIGURE 5.1 – Illustration of the projection on the normal data support. The shape description space $\mathcal{C}$ is partitioned in normal and abnormal data, lying correspondingly inside and outside the shaded area. The set of normal shapes is comprised of five-face geometrical solids. This partition consequently labels a pyramidal shape as normal. A truncated pyramid with a sawed-off top is tagged as abnormal, since it comprises six faces. The dotted line joining the truncated pyramid to the pyramid illustrates the notion of projecting an abnormal datum to its corresponding normal datum image.

In the context of a SVDD-inferred support, this operation is captured by the SVDD projection, that is the projection of the abnormal datum on the normal support. If we should see the problem in terms of denoising as in [PARK et al. 2007], the noise would here be represented in the sawed-off top of the new pyramid - provided that the distance metric between geometric solids has been appropriately defined.

The defining aspect about the SVDD projection is that it does not only serve as a means to map noisy patterns onto normal patterns, but also that patterns are mapped specifically to the closest normal datum. In the shape analysis context, this means that through the SVDD projection shapes undergo *just enough* deformation required to become normal. More particu-

larly in the neurodegenerative disease setting, the SVDD projection would result in the image of the healthy tissue that is at the same time the closest to the input pathological tissue.

The SVDD projection of any given input $x \in \mathbb{R}^d$ can be computed –supposing that d is not too large– simply with a grid search at the neighborhood of the input training data. Out of all points fullfilling $f_{SVDD}(.) \geq 0$ (eq. (4.5)), the point closest to the input x is the SVDD projection. In [Park et al. 2007], a non-brute force method to compute the SVDD projection is proposed. The projection is computed first on the feature space in terms of its kernel product. Then, its preimage on the original space is estimated. This latter method is independent of d, however the end-result depends on the estimate quality issued on the preimage step. In our experiments we used the grid search method when d was sufficiently small, falling back to the method of [Park et al. 2007] when grid search proved prohibitively expensive.

1.2 Discriminative trajectory

Previously we have discussed the basic concepts of SVDD theory and introduced the idea of the SVDD projection, as a projection onto the support induced by the training, "normal" data. Let us consider the set of continuous paths Δ between an abnormal point x and its SVDD projection $p_R(x)$ given the hypersphere (R, a). Both endpoints x and $p_R(x)$ are lying in space $\mathcal{C}$. Any element $\gamma \in \Delta$ can be written as a parametric continuous function $\gamma : u \in \mathbb{R} \mapsto \gamma(u) \in \mathcal{C}$, where $\gamma(0) = x$ and $\gamma(1) = p_R(x)$.

In the context of shape analysis, where $\mathcal{C}$ is a space where elements represent shapes and distance corresponds to shape dissimilarity, continuous movement in $\mathcal{C}$ represents gradual shape change. For the sake of illustration, assume $\mathcal{C}$ to be the voxel space, i.e. each vector of $\mathcal{C}$ is an ordered grouping of all volume voxel intensities. Movement from $\gamma(u)$ to $\gamma(u + du)$ then represents shape change in the sense of infinitesimal change in the value of some or all voxel values. Another choice is to assume $\mathcal{C}$ to be the space of all $3D$ deformation vector fields, with respect to a given volume-template. A movement along γ would then represent a change in the sense of an introduced quantum of deformation, or in other words that an infinitesimal amount of contraction or dilation is introduced in a part or all areas of the shape. The same observations would be also valid as if we were discussing a continuous path not in Δ. For elements in $\gamma \in \Delta$ in particular, movement from $\gamma(u)$ to $\gamma(u + du)$ can be interpreted as infinitesimal shape movement towards the normal element $p_R(x)$ if $du > 0$, or infinitesimal shape movement towards the abnormal element x if $du < 0$.

We are interested in a specific path $\delta \in \Delta$, defined as a series of SVDD projections for all valid values of the hypersphere radius. We saw already that by definition a SVDD projection regardless of parameters lies on the $\|\phi(x)\| = 1$ hypersphere, and has a valid preimage. As such, δ is well-defined on both spaces $\mathcal{C}$ and $\mathcal{F}$. Formally we define the path $\delta(u)$ as

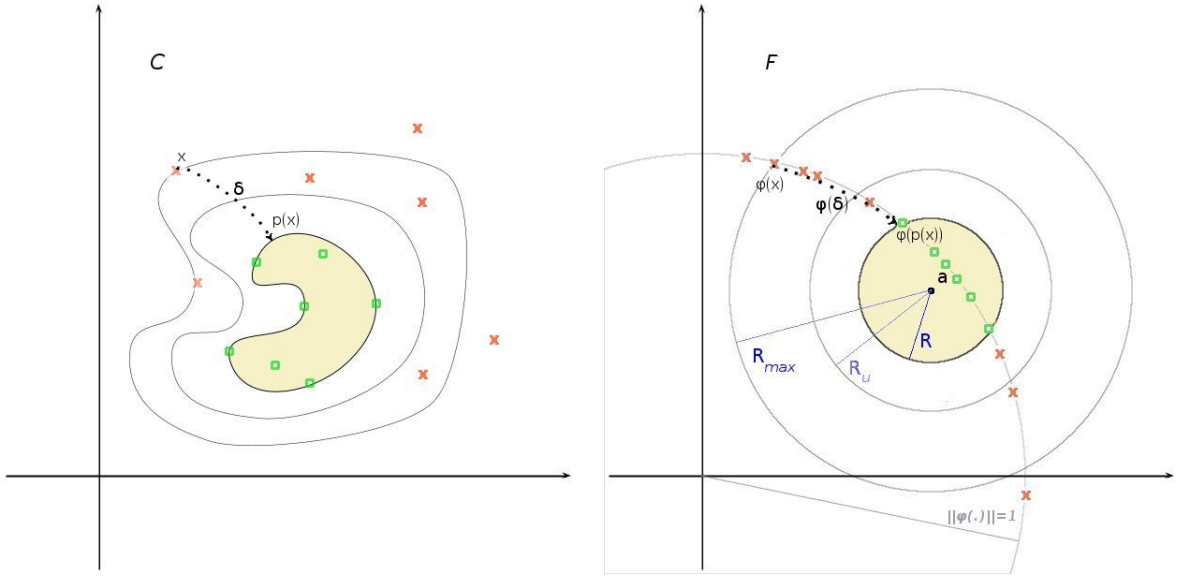
$$\delta(u) = p_{R'}(x), \qquad \text{where} \quad R' = (1 - u)R_{\max} + uR. \tag{5.1}$$

Each path point has the attractive property, that since it is a projection on the SVDD support of normal elements, it represents as such the abnormal input x minus "just enough" deformation to be considered normal (recall the previous discussions in chapter 4 and subsection 1.1 of the current chapter) under a varying radius threshold governing the support size. For this reason we shall refer to δ in this work as a *discriminative trajectory*. Visually, the discriminative trajectory thus defined corresponds to projecting on a series of dilations of the SVDD hypersphere, between the optimal radius R (optimal in the sense of eq. (4.1), and solved as in eq. (4.4)) and the maximal radius R_{max} (see fig. 6.1). Endpoint $\delta(0)$ is defined as the input x itself, so the maximal radius R_{max} must be equal to the distance of x from the hypersphere center, in the feature space.

Expanding $R_{max} = \|\phi(x) - \sum_{n=1}^{M} \alpha_n \phi(x_n)\|$, we have

$$R_{max}^2 = 1 - 2\sum_{n=1}^{M} \alpha_n k(x_n, x) + \sum_{n=1}^{M}\sum_{m=1}^{M} \alpha_n \alpha_m k(x_n, x_m).$$

See fig. 5.2 for a visual illustration of the discriminative trajectory in the original shape space and its image in the feature space.



(a) (b)

FIGURE 5.2 – Green circles are normal data, red crosses are abnormal data. The normal data support corresponds to the beige-color filled area. (a) The original (shape) space. (b) The feature space. In both panels, $p(x)$ corresponds to the projection of a datum x onto the normal support (*SVDD projection*). The *discriminative trajectory* δ is drawn as a black dotted path.

2. Related work in support vector machines

2.1 Discriminative direction

The concept of a tool to estimate the kind of shape deformation that differentiates one class of objects from non-members has been originally introduced in [GOLLAND 2000], in the form of the *discriminative direction*. The underlying discriminative model is the support vector machine, in the case of the discriminative direction. We recall that SVMs are 2-class discriminative models (see chapter 4). In SVM, the two classes are separated by a hyperplane in the feature space. The SVM decision function is of the form $f_{SVM}(x) = w^T\phi(x) + b$ (see eq. (4.8)). This implies that the classifier function depends only on the projection of vector $\phi(x)$ on w, ignoring totally the component that is perpendicular to $\phi(x)$. In other words, in order to create a displacement of $\phi(x)$ that would correspond to the difference between the two classes, one should change the input vector projection onto the w component while keeping the orthogonal component intact. Thus essentially the direction along which we need to move the datum image $\phi(x)$ is w.

At this point arises the following problem : moving the datum image x along w would be straightforward if $\phi(.)$ would be the identity mapping. If not, we need to compute the preimage of $\phi(x) + dz$, where dz is a displacement in the feature space $\mathcal{F}$. However, it is not *a*

priori guaranteed that $\phi(x) + dz$ has a valid preimage. In the typical scenario where $\mathcal{F}$ is the feature space for the RBF kernel case, points of $\mathcal{F}$ with a valid preimage necessarily reside on hypersphere $\phi(x)^T\phi(x) = 1$. In the discussion on the discriminative trajectory, the problem was solved by defining each moment of the trajectory as an SVDD projection, which has a preimage by construction.

The solution that is proposed in [Golland 2000, Golland et al. 2005] is to search for the direction around the datum x in the original space that minimizes the divergence of its image $\phi(x)$ from the direction of the projection vector w. This is called the discriminative direction, and represents the direction that affects the output of the classifier while introducting as little spurious difference as possible.

Formally, moving from x to a new point $x + dx$ in the shape space $\mathcal{C}$, the image vector in the feature space $\mathcal{F}$ changes by

$$dz = \phi(x + dx) - \phi(x), \tag{5.2}$$

which creates a deviation e from w, equal to

$$e = dz - \frac{dz^T w}{\|w\|}\frac{w}{\|w\|} = dz - \frac{dz^T w}{w^T w}w. \tag{5.3}$$

Solving for the discriminative direction dx is achieved by a constrained optimization problem, set to minimize the error norm $\|e\|$:

$$\arg\min_{dx} \|e\|^2 = dz^T dz - \frac{(dz^T w)^2}{w^T w}, \tag{5.4}$$

$$s.t.\ \|dx\|^2 = \epsilon. \tag{5.5}$$

For the Gaussian RBF kernel case, the above yields a closed form solution, which is the following [Golland et al. 2005]

$$\sum_{n=1}^{M} \alpha_n y_n K(x, x_n)(x - x_n), \tag{5.6}$$

where we recall that α_n are given by the solution of the SVM objective function (see chapter 4), and y_n correspond to $+1$ or -1 depending on which class training datum x_n belongs.

2.2 Regularized discriminative direction

In [Zhou et al. 2008] it is recognized that moving a shape along the discriminative direction may prove to be problematic if the low-dimensional manifold structure of the shape distribution in the used shape space is not taken under consideration. We also discuss this important point in chapter 3 and chapter 6. Simply put, [Zhou et al. 2008] point out that the discriminative direction as defined in [Golland 2000, Golland et al. 2005] may produce shape samples that come from areas that are off the hypothesized shape manifold. This means that in practice generated shapes will include characteristics that are more or less invalid, or in other words, shapes that do not look like the original shapes. Likewise, the discriminative deformation will then contain spurious shape difference. To deal with this, [Zhou et al. 2008] propose a modification of the discriminative direction optimization objective function (eq. (5.4)). They add a regularization term of the form

$$\beta(x + dx - \mu)^T \Sigma^{-1} (x + dx - \mu)^T \tag{5.7}$$

where β is the scalar non-negative weight of the regularization term, and μ and Σ are the parameters of the hypothesized probability distribution of shapes, that should reflect the low-dimensional manifold shape structure. Mean μ is set to x, and Σ is set to the covariance of the

training data in $\mathcal{C}$, assuming μ as their mean [Zhou et al. 2009]. Solving the modified problem results in an iterative scheme.

Note that in the model that is proposed in this thesis, we do address the same issue, albeit in a different, implicit manner. Instead of working to modify the discriminative trajectory so that the trajectory in $\mathcal{C}$ is rectified to account for the manifold hypothesis, we modify in a sense the space $\mathcal{C}$ itself, using dimensionality reduction. We have seen dimensionality reduction in chapter 3 and we discuss the connection with the discriminative aspect of the model in chapter 6. Also note that, in comparison with the scheme of [Zhou et al. 2008], we hypothesize the existence of a non-linear manifold using state-of-the-art manifold learning methods [Tenenbaum et al. 2000], consequently we do not need to alter the model objective function in any way.

2.3 Discriminative pattern

A map showing the areas and the type of discriminative deformation over a template tissue has been recently proposed [Costafreda et al. 2011, Koutsouleris et al. 2009], again under the SVM framework. The proposed discriminative pattern is computed by first selecting pairs of support vectors that are separated by the smallest distance accross opposite sides of the separating margin. For each of these pairs, their difference is computed. This difference plays a role similar to that of the discriminative direction of [Golland 2000], albeit defined here differently. The mean of these differences, evaluated for support vector pairs, forms the discriminative pattern.

The main advantage of this discriminative pattern is that it allows for the creation of a simple and functional map of discriminative difference. The discriminative pattern is essentially a function of a discriminative direction evaluated at a selection of points, adapted so as to visualize information carried by the discriminative direction easily. The discriminative pattern is a means to obtain discriminative difference as a single, global, group vs group trait. This advantage of the discriminative pattern is also its disadvantage, as the resulting map is made up of multiple difference components. Due to their summing up to a single pattern, they are rendered impossible to discern one from the other.

3. Conclusion

In this chapter we have posed the question, what is the type of deformation that makes the difference between class member and non-member, and have discussed possible answers to this question.

We have proposed a novel method to estimate discriminative deformation, that fits the SVDD model. We have named the proposed scheme discriminative trajectory, which serves not only to differentiate from the SVM-based schemes discussed in this chapter, but also underpins our viewing of discriminative difference explicitly as a process. The endpoints of this process are the original shape itself and the projection of the input on the normal data support. In principle, the difference associated with one endpoint of the process can be very different from the one associated on the other trajectory endpoint, as well as from points in-between.

In the next chapter, aside from completing the presentation of our model, we shall see that in synergy with the manifold hypothesis model component, the proposed discriminative trajectory can uncover discriminative difference that models that do not take the manifold structure of shape data into consideration, cannot.

Chapter 6

Proposed model

In this chapter, we review the components presented previously, and present them as a connected whole. We discuss the proposed model in terms of the interaction between three data spaces –ambient space, reduced space, and feature space– and define how to regress on embedding coordinates and produce new shapes. Next, we present two artificial datasets which we use to illustrate and evaluate the processing pipeline, before continuing with an application on real anatomical data in the next part of this thesis. Model function is illustrated and evaluated through discriminative trajectories and p-value significance maps which summarize discriminative difference.

1. Combining SVDD and the manifold hypothesis

Let us recall the definition of the *membra disjecta* of our model, and review them in their unison. In chapter 4 we have introduced the duality *original space* $\mathcal{C}$ - *feature space* $\mathcal{F}$. In chapter 3 we have introduced another space duality, that between the reduced space $\mathcal{X}$ and the ambient space $\mathcal{Y}$. The reduced space $\mathcal{X}$ is the output of manifold learning given a set of N observations in $\mathcal{Y}$. In contrast to $\mathcal{Y}$, $\mathcal{X}$ is –in principle– a compact model of shapes, where each point of $\mathcal{X}$ corresponds to a shape similar to the training data shapes. Hence, making the link between the SVDD theory part and the manifold hypothesis part that we so far discussed, we identify the original space of SVDD with the reduced space, that is we set $\mathcal{C} \equiv \mathcal{X}$. In all, our data are then represented in three ways :

1. the representation in the ambient space $\mathcal{Y}$ which is the only one directly observed. In the experiment part of this work (section 2. chapter 8) we shall assume $\mathcal{Y}$ to be the space of non-rigid deformation vector fields with respect to a coordinate frame, given by a predefined template image.
2. the intrinsic parametrization in the reduced space, or embedding $\mathcal{X}$.
3. the representation in the feature space $\mathcal{F}$. This latter comes into play for SVDD support estimation.

A subset of M normal observations out of the total N training data is used to induce the normal data support defined in $\mathcal{X}$ (identified here with $\mathcal{C}$) (see fig. 6.1 for a graphical overview of these spaces).

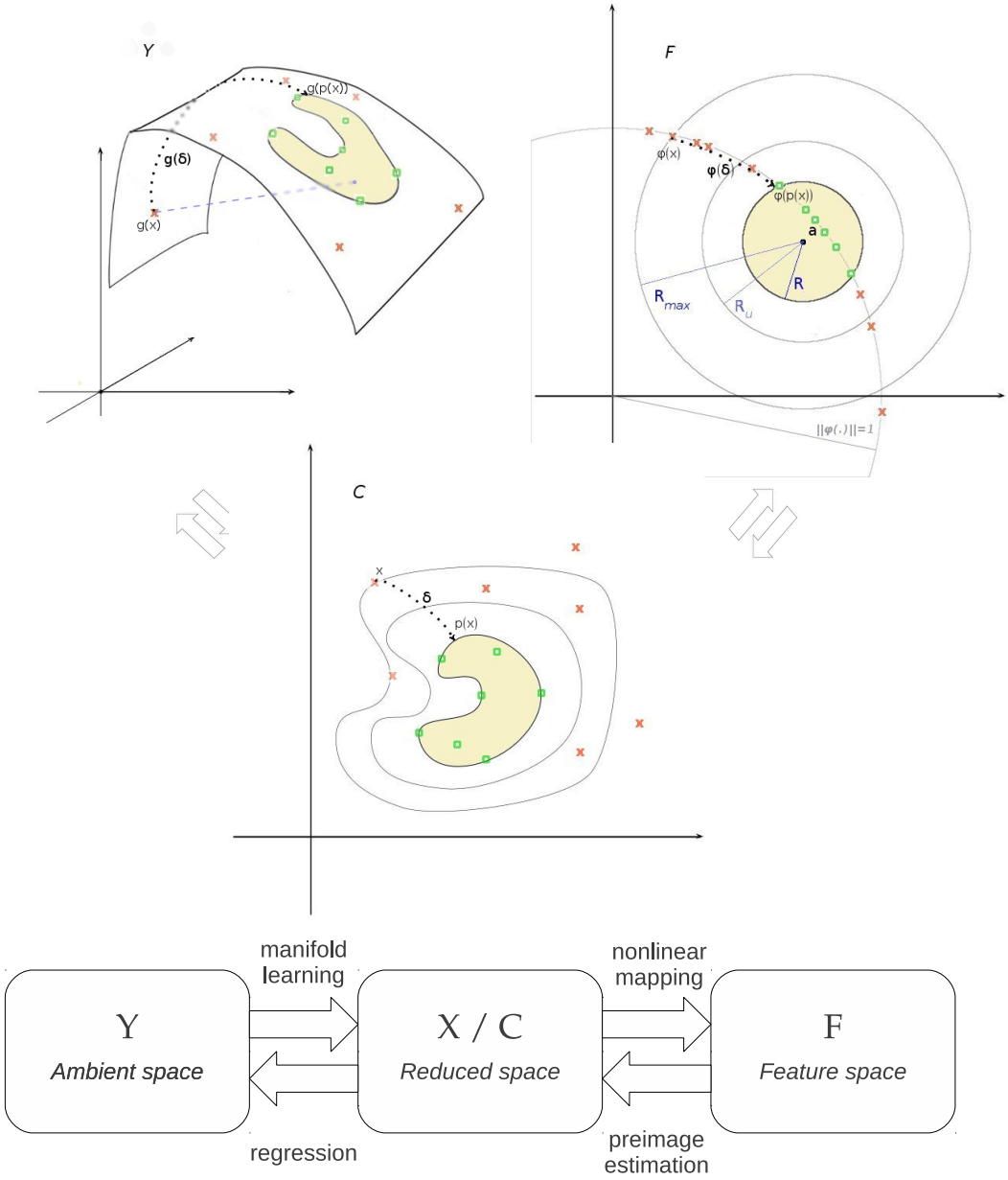


FIGURE 6.1 – The processing pipeline of the proposed model, in terms of the different shape representation spaces used (ambient space $\mathcal{Y}$, reduced space $\mathcal{C}$, feature space $\mathcal{F}$). Green circles are normal data, red crosses are abnormal data. The ambient space is observed, but typically is very sparsely populated by data. The "sheet" structure –a *manifold*– is a space of dimensionality lower than the one of the ambient space, where data are assumed to lie upon. Using manifold learning we can represent it as a reduced space. This space is not observed, but in principle points of this space correspond to valid, possible observations. The feature space is useful when estimating the normal data distribution (support). In all panels, $p_R(x)$ corresponds to the projection of a datum x onto the normal support (*SVDD projection*). The *discriminative trajectory* δ, drawn as the black dotted path, is a path joining x and $p_R(x)$, constrained to pass through the manifold. In the ambient space graph, the blue path joins x and the support, but passes through non-manifold space.

The notions of SVDD support estimation (see chapter 4), SVDD projection and discriminative trajectory (see chapter 5) all apply to elements of the embedding $\mathcal{X}$. Points in $\mathcal{X}$ are latent variables however ; in order to produce a practically usable form for the SVDD projection and trajectory, we have to set our model in "generative mode" using a regression relation of the form of eq. (6.1). We define this relation more precisely in the following subsection.

1.1 Generating new shape with multivariate regression

Generating new shape consists in employing the couples $\{\{x_1, y_1\}, \{x_2, y_2\}, \ldots, \{x_N, y_N\}\}$ to form basis functions and interpolate over new $\mathcal{X}$ values. Let us recall that a couple $\{x_n, y_n\}$ consists of the non-compact shape descriptor y_n –which consists of parameters of the deformation field descriptor (see chapter 2)– and its compact image x_n on the reduced space, computed using Isomap (see chapter 3). We assume that

$$y = g(x) + \epsilon = W^T \beta(x) + \epsilon, \tag{6.1}$$

fixing the more general eq. (3.1). The above equation forms an ordinary least squares (OLS) multi-output regression [Bishop 2006], with $y \in \mathbb{R}^D, x \in \mathbb{R}^d$ as already discussed, W being a $P \times D$ weight matrix and $\beta(x)$ being a vector containing P coefficients corresponding to the basis functions (rows of W). Noise ϵ accounts for small deviations of y from the manifold surface, with typically $\epsilon_n \sim \mathcal{N}(0, \sigma_n^2)$.

This yields a straightforward solution which is

$$W = (\Phi\Phi^T)^{-1}\Phi\mathfrak{Y}, \tag{6.2}$$

where $\mathfrak{Y}$ is an $N \times D$ matrix with y_n as its n^{th} row, and Φ^T the $N \times P$ *design* matrix with $\beta(x^n)$ as its n^{th} row. We choose h-scale Gaussian basis functions, plus a bias basis :

$$\beta_n(x) = \begin{cases} \exp\{-\frac{1}{2h^2}(x - x_n)^2\} & \text{if} \quad 1 \leq n \leq N, \\ 1 & \text{if} \quad n = N + 1, \end{cases}$$

and therefore, $P = N + 1$. In practice bandwidth h is adjusted by cross-validation. Evidently, the more data N at our disposal, the better the whole model will be. This is true in terms of the regression model, the quality of which increases with N, as well as for other aspects such as the manifold learning or the support estimation step.

The resulting non-linear structure conveys as such an advantage over a linear, non-compact solution (see fig. 6.1(a)). Traversing a path as simple as a straight line in $\mathcal{X}$ corresponds to a curve in the ambient space $\mathcal{Y}$, meaning dynamically changing types and directions of deformation over the studied anatomy. More importantly, elements of the resulting path, hence also of the discriminative trajectory, are constrained to always "look like" elements of the training set.

2. Illustration with synthetic datasets

2.1 Data sets considered

Set name	Type	Anatomy	Number of subjects			
			controls	patients	int. group	total
Synth-I	synth.	spheres	20	30	0	50
Synth-II	synth.	l. ventricles	2	16	18	36

TABLE 6.1 – Overview of synthetic datasets used in the current work.

We have tested the proposed processing pipeline on a number of both real and synthetic data. In this section, we shall present the synthetic data we have elaborated, then build our discriminative and generative model over them. An overview of some characteristics of the synthetic datasets is given in table 6.1.

We created two synthetic data sets, *Synth-I* and *Synth-II*, on which we test our model, for proof of concept. Set *Synth-I* is composed of deformations on a baseline sphere structure. The normal class is comprised of 20 spheres with small, random deformations spread over their surface. The abnormal class is comprised of spheres with 3 distinct types of deformation, with 10 data for each type. In total, the abnormal class comprises 30 deformed spheres. All deformations over the baseline form were produced using the trigonometric kernel model of [COOTES et al. 2008]. To each of the 3 deformation types corresponds a distinct point over the sphere surface, acting as a deformation control point. These three points are chosen to be equidistant and lying on a common sphere perimeter. The first type of deformation corresponds to pushing the control point towards the center of the sphere, which gives a contraction over the area around the control point. The other two control points are pulled outwards, giving a swelling over each of the corresponding areas. In figure 6.2 we show a number samples taken from set *Synth-I*.

Set *Synth-II* is composed of deformations over a baseline of a real left lateral ventricle, taken from the OASIS dataset (see chapter 8 for a detailed description of the OASIS dataset). The normal class here is composed of the baseline ventricle, plus a version of the same ventricle with a slight contraction on the middle of its surface (see fig. 6.3, left column). Over these two forms, considered as normal, the same series of deformations is then applied to create the abnormal class. The series of deformations is as follows : first we deform the baseline form, by elongating its lower part (see fig. 6.3, middle column). Each of the produced forms is gradually more elongated than its corresponding previous one. In this manner we create 5 volumes for each of the two baseline volumes (total 10 up to this point). The most elongated forms now constitute a new baseline. We deform the higher part of the new baseline, by elongating its higher part this time (see fig. 6.3, right column). Again each of the produced forms is gradually more elongated than its corresponding previous one. In this manner we create 3 extra volumes for each of the two baseline volumes (that is an extra 6 volumes, in total 16 in the abnormal set). We have also added a set of points that are tagged intermediate. In real data applications, these would correspond to unlabelled data, or data of unsure/intermediate diagnosis when class refers to pathological state. These play no direct role in training the discriminative model or creating discriminative trajectories, but are used in the manifold learning step, like the rest of the data. In *Synth-II*, we have created a total of 18 shapes tagged as intermediate. The process we used to create them mirrors the one for the aforementioned shapes of the normal and abnormal sets, differing only in the level of contraction in the middle of their surface.

2.2 Tests on synthetic datasets

Running Isomap on the *Synth-I* data produces the $2D$ embedding of fig. 6.4. The normal

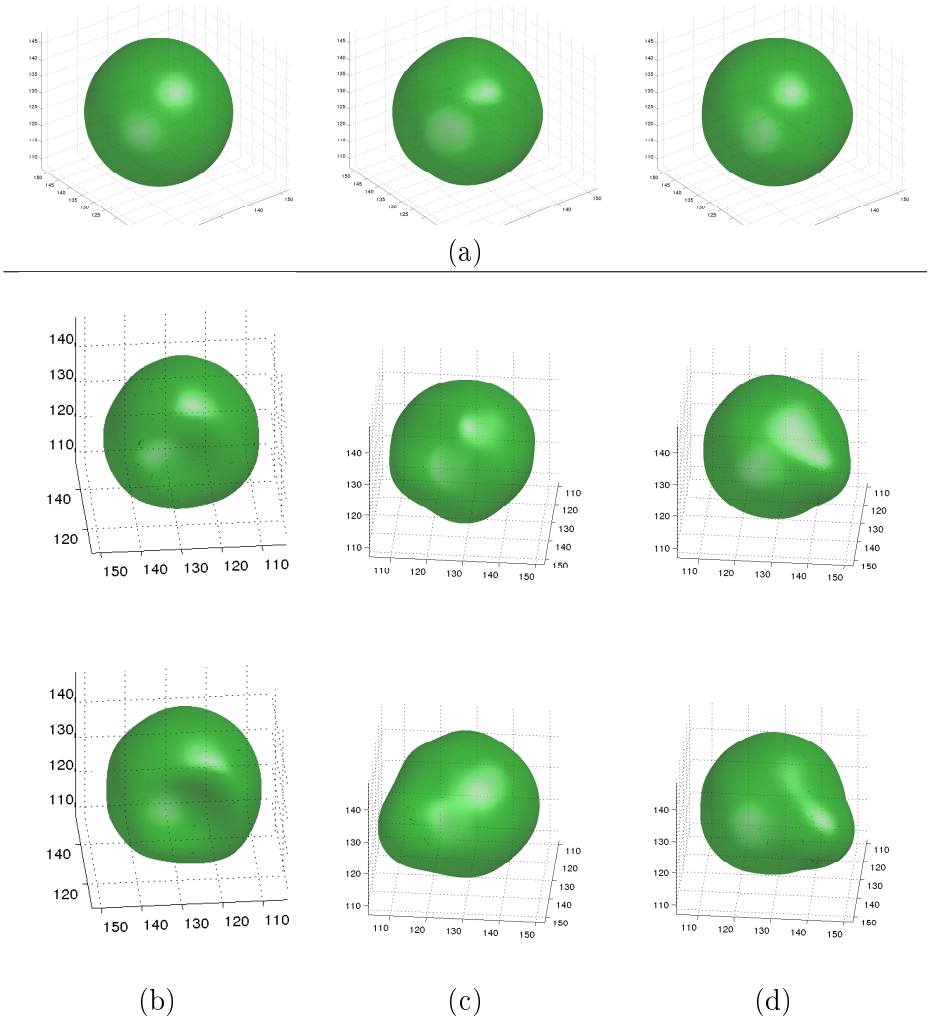
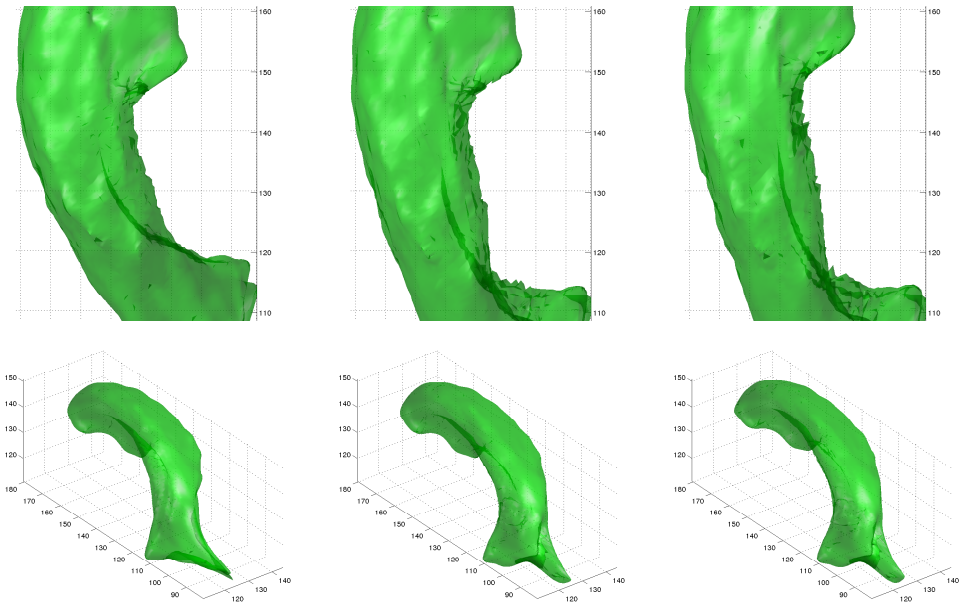
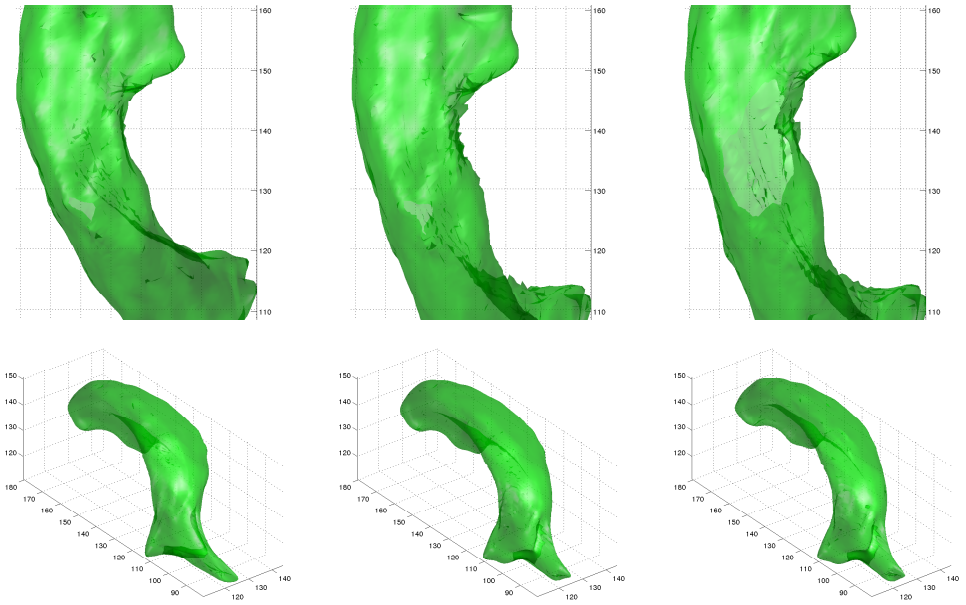


FIGURE 6.2 – Samples taken from the synthetic dataset *Synth-I*. (a) Random sample shapes from the normal group. (b)-(d) Sample shapes from the abnormal group. Same column corresponds to the same type of deformation with each row corresponding to a different magnitude. Columns (b)-(d) correspond to different viewpoints.



(a)



(b)

FIGURE 6.3 – Samples taken from the *Synth-II* dataset. Images in (b) are versions of the images in (a), with slightly contracted bodies. Left column : baseline images. Middle column : bottom part elongated. Right column : bottom and top part elongated (top/bottom specifications refer to the 2D view). Only the images on the left column are labelled normal. The difference in structure is best visible in the "C"-shaped interior outline, on the 2D view.

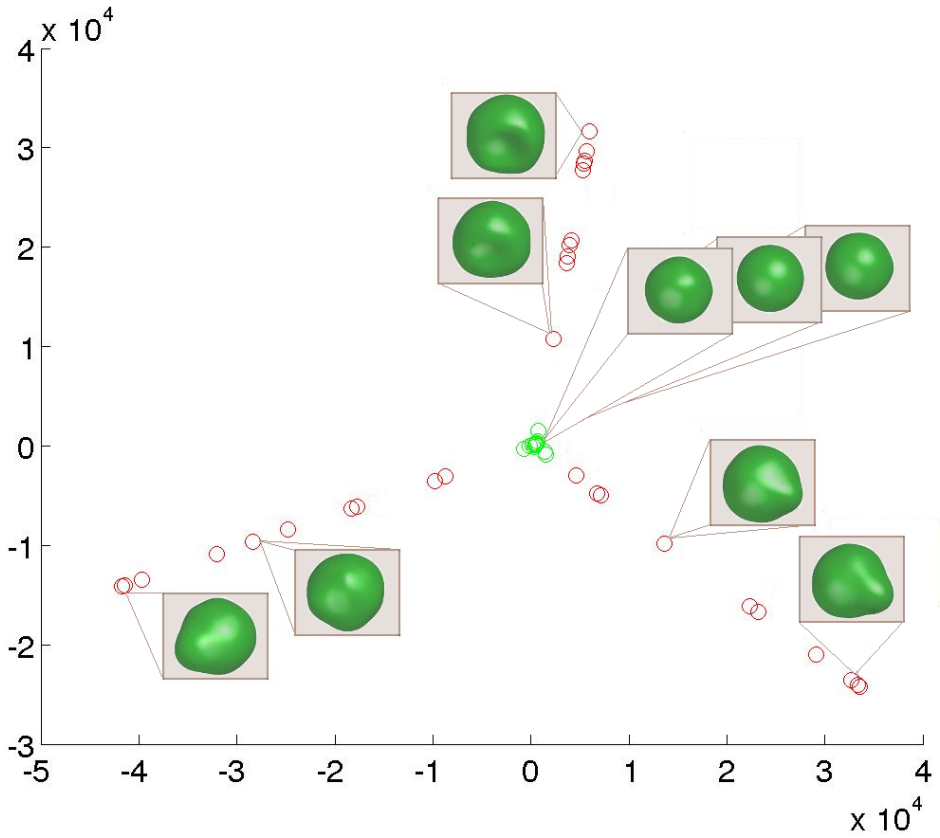


FIGURE 6.4 – Isomap $\mathbb{R}^2$ embedding for synthetic dataset *Synth-I*. Green points are part of the normal set. Red points are part of the pathological set, corresponding to an imaginary degenerative pathology over the sphere "tissue". Normal data are almost perfect spheres, while pathological data are spheres plus an amount of contraction or swelling over a given area. Three types of pathological data were produced, portrayed on the isomap representation as each of the rays protruding from the center.

group data lie on the center, tightly packed and close to one another. This is due to the small shape difference between them, as they are all quasi-spheres. The pathological group is divided into the 3 well-demarcated rays, protruding from the normal center. Correctly, these correspond to the three types of deformation previously discussed, each corresponding to a deformation around a given control point. The further a point from the center, the larger its magnitude of deformation with respect to the baseline perfect sphere.

In order to assess shape difference between class and non-class, we have computed a map of *p-values*. These values show areas of significant difference between two given groups of interest [Rosner 2000]. Let us remind the reader that in our analysis, the elements of the ambient space $\mathcal{Y}$ are non-rigid deformation vector fields with respect to a coordinate frame, given by a template image. We have also seen already that normal subjects are characterized and differentiated from pathological subjects in terms of their SVDD projection. Normal subjects should be very close to the SVDD border or inside the SVDD ball, while pathological subjects should be relatively further away. This stems simply from the fact that normal subjects define the position and form of the SVDD ball. The difference of each subject to its reconstructed SVDD projection is the value of interest for our statistical tests and the related p-value maps.

Formally, we compute for each anatomy $n \in [1, N]$ and each shape voxel v,

$$\epsilon_n^v = \|y_n^v - g^v \circ p_R(x_n)\|,$$

where $g \circ p_R(x_n)$ represents the reconstruction of the SVDD projection of anatomy n, and $g^v \circ p_R(x_n)$ represents the vector value of $g \circ p_R(x_n)$ at voxel v. For normal anatomies $g \circ p_R(x_n)$ should be close to y_n for all voxels, for the reasons we mentioned above. A one-sample t-test is then run over the norm ϵ_n^v corresponding to each voxel, that is we perform one test per voxel. The "pathological" spheres play the role of the sample here. Hence, the test statistic for each voxel is computed as $t = \frac{\bar{x}-\mu_n}{\sigma_s \sqrt{n_s}}$ [Rosner 2000, section 7.3], where $\bar{x}, \sigma_s, n_s$ are the sample mean, standard deviation and cardinal of the sample (abnormal shapes) and μ_n is the normal population mean. The test statistic obtained per voxel leads to a p-value per voxel. This is computed as the area under the tail of a Student's-t distribution of mean $\bar{x} - \mu_n$, standard deviation σ_s and $n_s - 1$ degrees of freedom [Rosner 2000, section 7.3] (see fig. 6.5 for an illustration). Note that using a two-sample test, with the set of normal shapes playing the role of the second sample, would be inappropriate in our case. This is because values of ϵ_n for normal shapes should be zero or close to zero at all voxels, and any small variance would primarily stem on reconstruction error, which is not what we want to measure with this test.

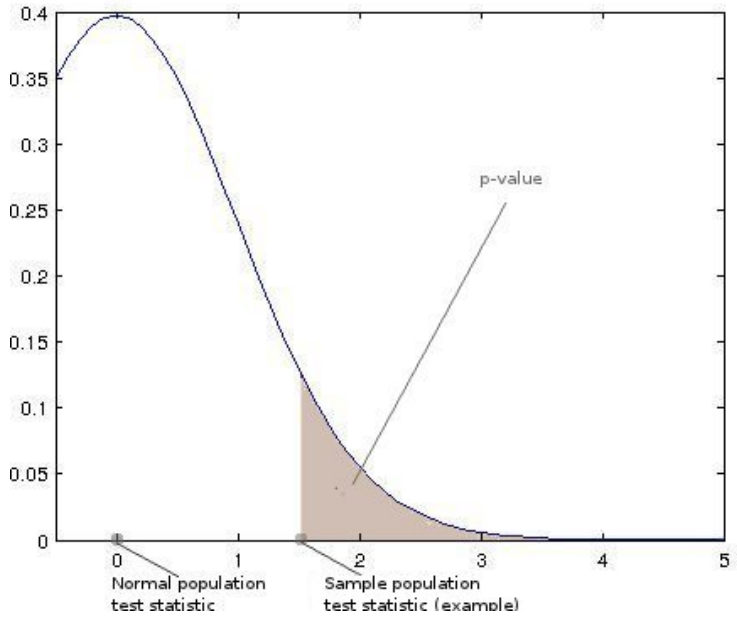


FIGURE 6.5 – The p-value, computed as the dark tail area demarcated by the sample test statistic (see text). The distribution mean represents statistics characterizing the normal set. In our case the test statistic is a value computed per shape voxel (ϵ_n^v). Put simply, this value corresponds to how much a given shape differs from its projection on the normal support, on the given shape voxel. Normal training data are evidently expected to have an identity projection, hence a statistic value equal or close to zero. If the abnormal training data give a consistently high ϵ_n^v value, this will translate to a statistically significant difference denoted by a far-off t-statistic and a low p-value for the specific point (voxel) of the shape.

In fig. 6.6 we see false discovery ratio (FDR) corrected [Benjamini & Hochberg 1995, Storey 2003] p-value maps overlaid on the template sphere. We have used FDR correction as it can give a more useful significance threshold when it comes to multiple tests, as is the case here, where we have one test per voxel.

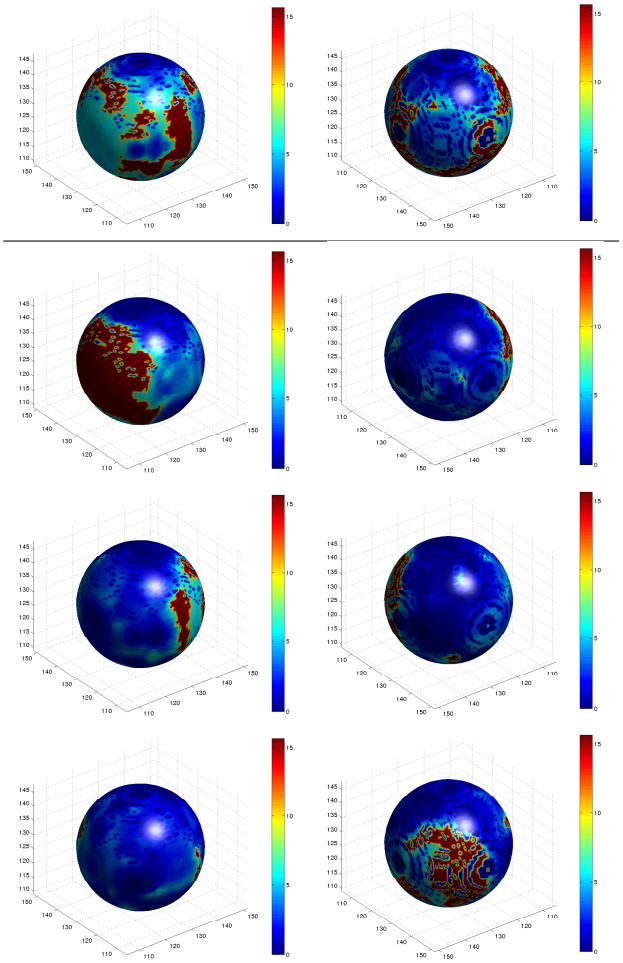


Figure 6.6 – Top row : p-value significance maps for the *Synth-I* dataset. Left and right columns correspond to front and back views of the sphere. Bottom rows : significance maps computed considering the patients set partitioned according to the three synthetized deformation types considered separately (see text). Each of these three rows corresponds to a different type of deformation. Negative log values are shown in all cases. Red corresponds to highly significant zones, blue corresponds to insignificant zones.

The p-values are lower where deformation is deemed *significant* between the normal and the pathological group. These low p-value voxels are correctly identified with the areas around the original deformation. On the downside, while the significance map of fig. 6.6 does identify the correct areas, it cannot give us any information about the diverse types of deformation included in the dataset. In fig. 6.7 the discriminative trajectories over the embedding are shown, corresponding to each of the deformation types of the pathology class. We set as their initial point the farthest sample for each of the three rays.

In fig. 6.8 we can see the Isomap embedding for dataset *Synth-II*. Here moving along the y-axis coincides with the degree of contraction in the middle of the ventricle body. The two normal subjects are at the rightmost part of the embedding plot. The points corresponding to the original OASIS ventricle are above, while the horizontal paths below correspond to the contracted versions. Starting from the two normal points are two point paths corresponding

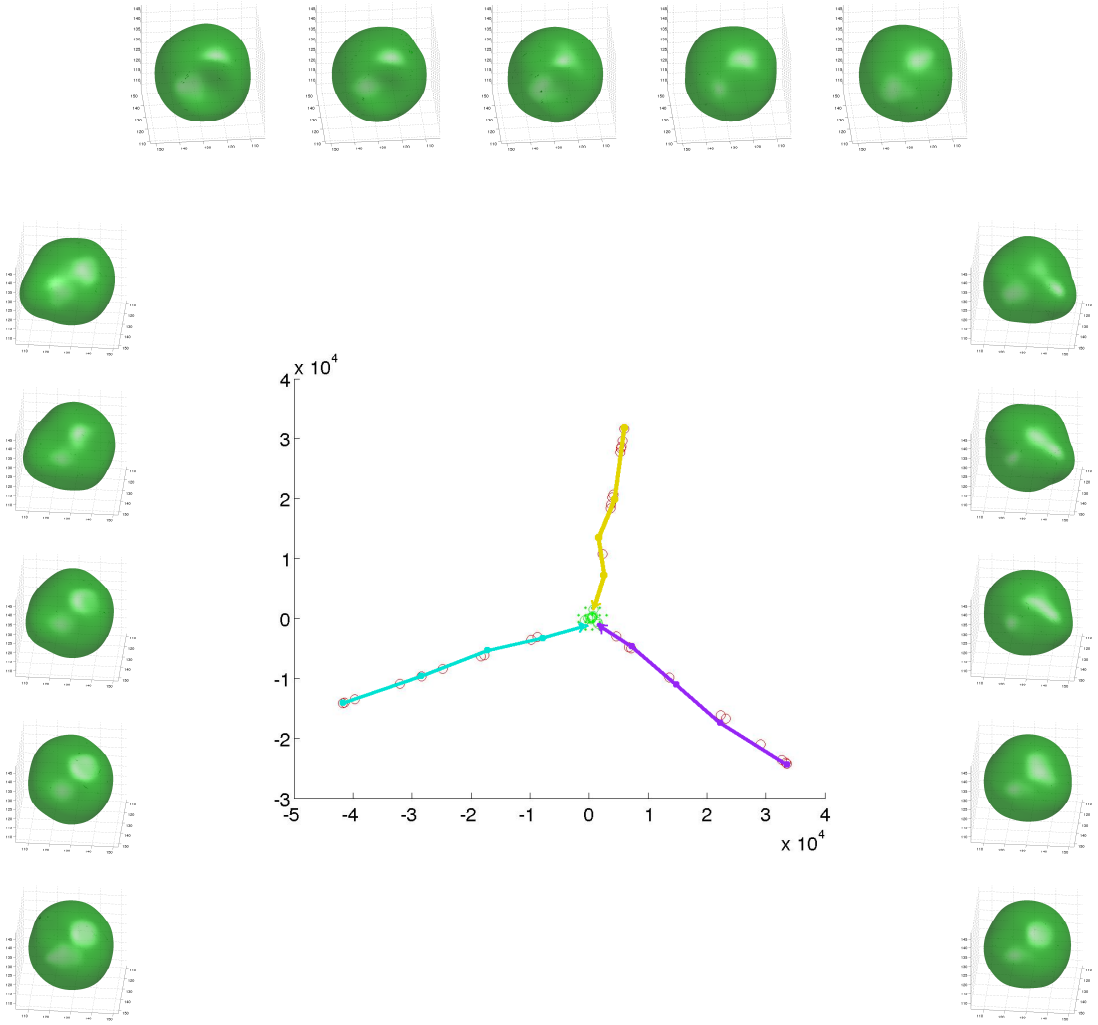


FIGURE 6.7 – Reconstructions sampled over discriminative trajectories, computed for each of the pathology deformation types. The orange-colored trajectory corresponds to the reconstructions shown on the top ; the cyan-colored trajectory corresponds to the reconstructions shown on the left ; the magenta-colored trajectory corresponds to the reconstructions shown on the right.

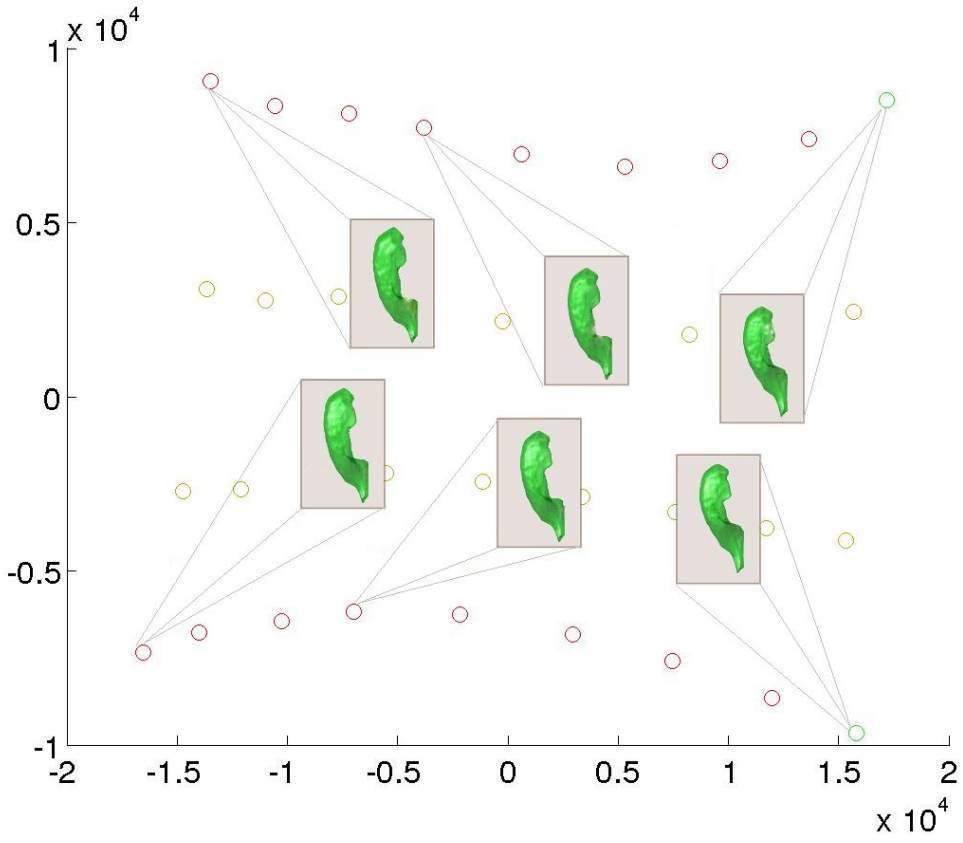


FIGURE 6.8 – Isomap $\mathbb{R}^2$ embedding for synthetic dataset *Synth-II*. Green points are part of the "normal" set ; red points are part of the "pathological" set. Orange points are part of the "intermediate" set.

first to a gradual swelling of the ventricle bottom, then to a swelling of its top. Between these two point paths, on the top and the bottom, are two intermediate point paths, that correspond to intermediate set shapes, which as we discussed earlier, have an intermediate degree of contraction of the ventricle body.

In fig. 6.9 the discriminative trajectories over the embedding are shown, corresponding to each of the two baseline ventricles. We set as their initial point the farthest sample for each of

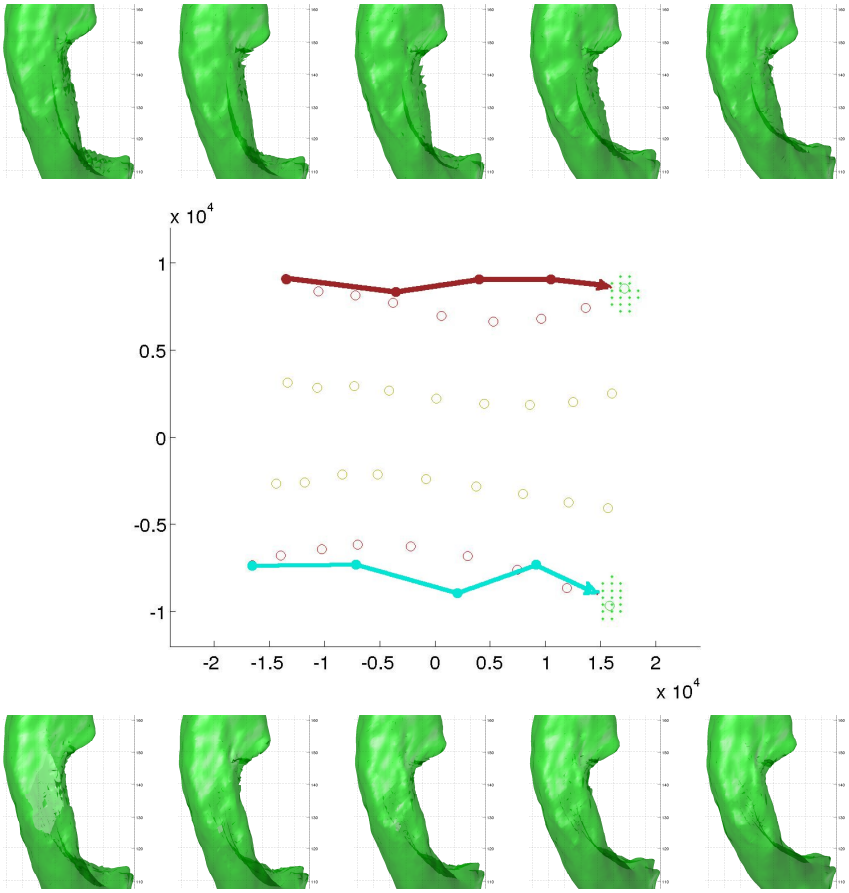


FIGURE 6.9 – Reconstructions sampled over discriminative trajectories, computed for each of the "pathology" deformation types. Note that the series of deformation types, on the top and the bottom of the "C"-like curve, is correctly reconstructed (See text for details). The brown-colored trajectory corresponds to the reconstructions shown on the top, the cyan-colored trajectory corresponds to the reconstructions shown on the bottom.

the two baseline normal ventricles. Note that movement on the x-axis, almost coinciding here with movement along the discriminative trajectories, does not correspond to a single type of deformation, such as a single type of swelling or a contraction on a number of given points. Rather, movement on the x-axis corresponds to a juxtaposition of *two* different types of deformation. This is not surprising, as it stems from the non-linearity of our model. In practice, this means that we can reconstruct correctly the process that leads from the baseline shapes on the rightmost edge of the embedding plot, to the very deformed shapes of the leftmost edge.

In fig. 6.10 we see samples of the discriminative trajectory in a model where we do not use dimension reduction (cf. fig. 6.1(a)). The samples generated do not resemble to the actual training set, in the sense that their top and their bottom are *simultaneously* deformed. In the original training set for example, fig. 6.3, the majority of the pathological ventricles have their lower part elongated but their higher part remains normal. It is worth noting that a linear dimensionality reduction model would fail to capture the data structure as well, since as we

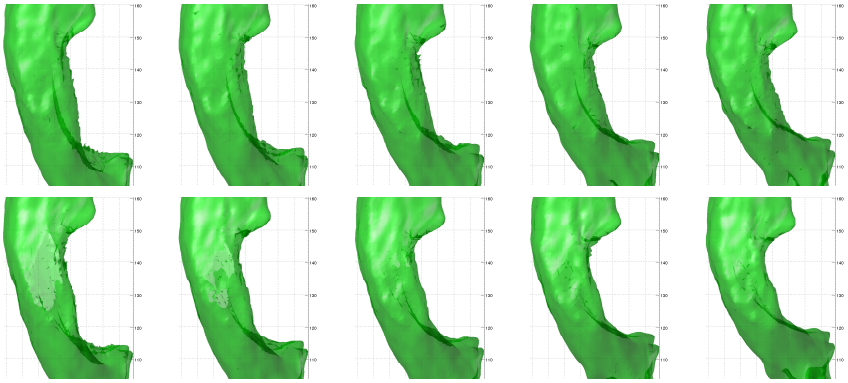


FIGURE 6.10 – Reconstructions for *Synth-II* made without the dimensionality reduction step. (see ambient space panel of fig. 6.1 and the "shortcut" path that does not follow the embedding). See text for details.

have seen the structure of *Synth-II* is by construction non-linear.

In fig. 6.11 we see a significance p-value map overlaid on a template baseline ventricle. We have used the same procedure to produce this map as the one used for the *Synth-I* set. Note that principally the lower ventricle area is found significant, which is natural, yet this discards information on the deformation of the higher ventricle that characterizes part (40%) of the pathological data. While the p-value map gives a possibly more convenient, single global map of class versus non-class discriminative difference, the discriminative trajectory (fig. 6.9) can reconstruct correctly all relevant deformations about the pathological class. In this sense, the significance map and the discriminative trajectory reconstructions are two tools complementary to one another. Note also, that the discriminative trajectories comprise and reconstruct *only* relevant deformation (which is represented here by movement along the embedding x-axis), barring out irrelevant deformation (which is represented here by movement along the embedding y-axis).

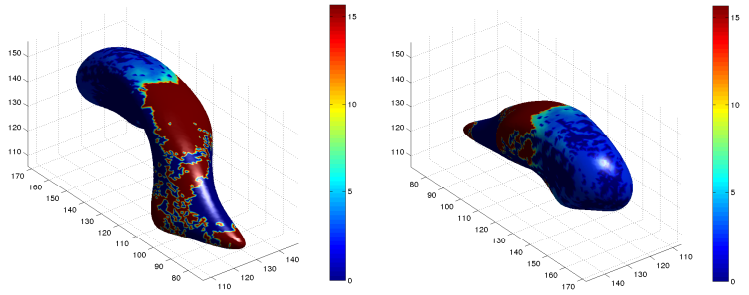


FIGURE 6.11 – p-value significance maps for the *Synth-II* dataset. Left and right image correspond to different views of the same plot. Negative log values are shown in all cases. Red corresponds to highly significant zones, blue corresponds to insignificant zones.

3. Conclusion

With this chapter we have completed this part of the thesis, and the theoretical presentation of the proposed compact, non-linear model. To the extent that the manifold learning step captures the underlying shape data structure correctly, the resulting model is compact, and can lead to a more accurate model. We have seen with an application on synthetic data how this higher model accuracy is expressed : newly generated shapes are on the one hand anatomically

correct, since they are constrained to lie on the hypothesized manifold ; on the other hand, discriminative trajectory moments are inherently also constrained to lie over the hypothesized manifold, leading to an accurate discriminative difference estimate.

Part III

Case study : hippocampus neurodegeneration in Alzheimer's disease

Introduction

The third part of the thesis presents an application of the proposed model to real neuroimaging data. The case study of choice is the progression of Alzheimer's disease, which is strongly correlated to structural evolution of certain regions of interest in the brain. Such regions include the hippocampus, amygdala and lateral ventricles. Our focus is mainly on the relation of the hippocampus to AD. We use our model on a group of patients, intermediate condition subjects, and controls, and create discriminative trajectories for the left and right hippocampi. Finally we illustrate the use of our model as a classifier, deciding over the likelihood of a hippocampus or amygdala originating off a healthy subject or not.

Chapter 7

Brain anatomical degeneration in Alzheimer's Disease

This chapter is a brief introduction to the structure of the brain, and to how this is affected during the course of Alzheimer's disease. First, we discuss some basic points about the structure of the human brain, and more especially about the hippocampus and the amygdala. In the next section, we discuss Alzheimer's disease and the effect its progression has over the brain. Structural and neuronal damage can play the role of a disease progression marker, or *biomarker*. Clinical symptoms of the disease can be measured using clinical assessment scores, that are typically included in patient population studies. Finally, we review state-of-the-art computational models for neuroimaging studies in AD.

1. Elements of brain anatomy

The brain is the center of the nervous system in humans. The cerebrum is the most prominent structure, approximately split in two quasi-symmetric hemispheres. The two cerebral hemispheres form its largest part and are situated above most other brain structures. The outer part of the cerebrum is formed as a highly convoluted layer, known as the cerebral cortex. Different parts of the cerebral cortex are involved in different cognitive and behavioral functions. Anatomically each hemisphere is divided in four lobes, the frontal lobe, parietal lobe, occipital lobe, and temporal lobe. This demarcation has to do mostly with spatial position rather than functional or tissue coherence. Histologically, a common division of the brain is in gray matter, white matter and cerebrospinal fluid. The gray matter is formed by neurons and their unmyelinated fibers, while the white matter contains mostly myelinated axons. Neurons are the basic building blocks of the nervous system. Axons or nerve fibers on the other hand serve to connect different organism areas and pass electro-chemical signals. Gray matter includes brain areas responsible for

muscle control, sensory perception such as seeing and hearing, memory, emotions, and speech. It is distributed at the surface of the cerebral cortex and of the cerebellum (cerebellar cortex), as well as in the depth of the cerebrum (thalamus ; hypothalamus ; sub-thalamus, basal ganglia - putamen, globus pallidus, nucleus accumbens ; septal nuclei), deep cerebellar nuclei (dentate nucleus, globose nucleus, emboliform nucleus, fastigial nucleus), brainstem (substantia nigra, red nucleus, olivary nuclei, cranial nerve nuclei) and spinal white matter (anterior horn, lateral horn, posterior horn). White matter is located in the inner layer of the cortex, the optic nerves, the central and lower areas of the brain (the brainstem) and surrounding the gray matter in the spinal cord. The cerebrospinal fluid is a transparent fluid that surrounds the brain and the spinal cord. Cerebrospinal fluid (CSF) is produced by an organ, called the choroid plexus, in the ventricles and by the surface of the roof of the third and fourth ventricles. The major functions of CSF are to protect the brain from injury and to maintain chemical stability. An important part of the brain is the limbic system, which is a collection of brain structures that form the internal border of the cerebral cortex. The limbic system is a collection of anatomical units with common functions [Duvernoy 1988], including the fornix, cingulate gyrus, amygdala, the hippocampus, the parahippocampal gyrus and parts of the thalamus. It supports functions such as long term memory, emotion, behavior, and olfaction.

A variety of pathologies may during their course cause the degeneration of important structures and brain areas. Alzheimer's, Parkinson's, frontotemporal disease or schizophrenia are some examples. The hippocampus and the amygdala are two cognitively important structures, that are known to be associated to the advent and development of neurodegenerative disease. Note that here exists one hippocampus and one amygdala per cerebral hemisphere. We examine these structures in more detail in the subsections that follow.

1.1 The hippocampus

The hippocampus is located in the medial temporal lobe, underneath the cortical surface. The hippocampus lies adjacent to the temporal horn of the lateral ventricle, and it forms an elongated "C"-like structure. Its general appearance does indeed resemble a sea horse in shape, as its name suggests. It can be divided into three segments : (1) A body, or middle segment, which is sagittally oriented. (2) A head, or anterior segment, which is transversely oriented and dilated and which shows elevations, the digitationes hippocampi. (3) A tail, or posterior segment, which is also oriented transversely and which gradually narrows until disappearing beneath the splenium [Duvernoy 1988].

The hippocampus is part of the limbic system [Duvernoy 1988]. Histologically-wise the hippocampus is subdivided transversely into several subfields, including the dentate gyrus, subiculum, and cornu ammonis subfields (CA1- CA4), which have unique connections to other subcortical and cortical regions in the brain [Greene et al. 2012]. These subfields are independent of subregional definitions of the hippocampus along its longitudinal axis (head, body, and tail). While all subfields are present within the head, body, and tail of the hippocampus, a majority of the dentate gyrus was found to lie within the hippocampal body, while a majority of CA1–CA3 were detected in the hippocampal head in [Malykhin et al. 2010].

It is generally admitted that the hippocampus has a critical role in learning and memory [Kandel et al. 1995, Nikita 2009, Duvernoy 1988]. Information arising from large isocortical zones converges to the entorhinal cortex area and then to the hippocampus. The entorhinal cortex is an area located at the anterior end of the temporal lobe, also important in learning and memory function. The entorhinal area, despite its small size, is the principal input to the hippocampus [Duvernoy 1988]. Newly acquired items cross the hippocampal filter before being fixed in the isocortex. It is possible to distinguish the memory of new or recent items, which depends on the hippocampus (short-term memory), from that of old ones (long-term memory),

which depends on the isocortex.

Patients with ambilateral hippocampus damage exhibit a total inability of conscious learning, while short-term memory is spared [Kandel et al. 1995, Eichenbaum 1999]. The most well-known suggestion about the exact role of the hippocampus in memory is the hypothesis that the hippocampus constitutes some form of a "cognitive map" [Nikita 2009]. Recent findings confirm the multifacet implication of the hippocampus in memory and strengthen the validity of the conclusion that the hippocampus is the mechanism of production and recall of mnemonic representations and stored data [Nikita 2009]. Hippocampus degeneration, as neuronal loss and atrophy, is well-known to be associated with the normal aging process [Driscoll et al. 2003], reflecting the deterioration of cognitive functions. From the aspect of behavioral pathology, the human hippocampus (and especially its anterior parts) constitutes the most frequent area of epileptic seizure manifestation [Nikita 2009]. In tests on rodents notably, the rodent hippocampus showed a relatively high degree of sensitivity in experimental *in vivo* epilepsy models. With some types of epilepsy there is associated a form of structural degeneration of the hippocampus, temporal lobe sclerosis [Nikita 2009]. Other pathologies related with hippocampal degeneration include hippocampal sclerosis, Parkinson's disease, schizophrenia and Alzheimer's disease. The hippocampus notably exhibits serious atrophy in the course of Alzheimer's disease [Cuingnet et al. 2011, Frisoni et al. 2010, Greene et al. 2012]. Neurofibrillary tangle development commences early on the hippocampus during AD [Greene et al. 2012]. There is evidence that the hippocampal head and body undergo increased atrophy in mild cognitive impairment (MCI) [Greene et al. 2012]. At the mild dementia stage of AD, hippocampal volume is already reduced by $15 - 30\%$ relative to controls, and in the amnestic variant of MCI the volume is reduced by $10 - 15\%$ [Frisoni et al. 2010].

The interested reader is referred to [Duvernoy 1988] for a much more detailed description of the hippocampus.

1.2 The amygdala

Anatomically the amygdala is a heterogenous group of nuclei situated in the dorsomedial portion of the temporal lobe. The amygdala has widespread connections with cortical and subcortical areas, including the aforementioned hippocampus. Closely connected with the much more voluminous hippocampus, structurally, spatially and functionally, the amygdala has been studied together with the hippocampus as a single hippocampus-amygdala complex [Golland et al. 2005, Shenton et al. 2002]. The amygdala is an important modulator between external sensory environmental influences and internal appropiate autonomic responses, and seems to be involved in cognitive functions including memory processing [Vereecken et al. 1994].

During the process of AD the amygdala shows abundant senile plaques and neurofibrillary tangles [Vereecken et al. 1994], and its degeneration is as such correlated to the disease development. Other pathologies related to amygdala degeneration include schizophrenia [Golland 2000] and autism [Chung et al. 2010].

2. Alzheimer's disease

2.1 Disease general characteristics

Alzheimer's disease is a hitherto incurable and terminal disease of unknown etiology. AD is the most common form of dementia on people over the age of 60, and today the disease affects more than 30 million people worldwide [Goate & Holtzman 2010, Jellinger 2006]. The central quest of research on AD is to identify the steps in its pathogenesis that, if inhibited, would slow or prevent the disease. An important step for AD pathogenesis shown as necessary accord-

ing to hitherto study is amyloid-β protein (Aβ) accumulation [Selkoe 1997]. Aβ accumulation in the cerebral cortex is an early and invariant event in the development of AD pathology, preceding other brain lesions and clinical symptoms by many years or decades. All AD patients develop neuritic plaques in brain areas subserving memory and cognition. These plaques consist of extracellular masses of $A\beta$ filaments (long protein chains). Aβ also accumulates in many nonfilamentous extracellular deposits. Typical forms of accumulated Aβ proteins are isoforms Aβ_{40} and Aβ_{42}. Virtually all patients also have many neurofibrillary tangles. Neurofibrillary tangles are intraneuronal bundles of paired helical filaments composed of tau protein.

Risk of developping AD increases dramatically with old age, age being the evident risk factor of AD. Also, among the groups with heightened AD risk are people with Down's syndrome and patients with certain familial genes. Such genes include β amyloid precursor protein (βAPP), apolipoprotein E4 (apoE4), presenilin (PS) 1 and 2. These genes have been shown to enhance Aβ production or deposition (or both) even in simple *in vitro* systems [Selkoe 1997]. On the other hand, the presence of apolipoprotein E2 (apoE2) is known to reduce AD risk [Hua et al. 2008B]. The gradual cerebral buildup of Aβ causes local neurons and their processes to be injured, leading to profound metabolic changes. The clinically important consequence of the inflicted changes is eventually synaptic loss and multiple neurotransmitter deficits [Selkoe 1997].

The loss and degeneration of neurons and their connections is commonly known under the general term *brain atrophy*. Structurally, brain atrophy is reflected in volume loss of various brain structures and areas such as the hippocampus, the amygdala and the entorhinal cortex (see also subsection 2.2). As a consequence, these organs degenerate also in function, causing cognitive impairment proportional to the progress of the disease.

2.1.1 Mild cognitive impairment and AD

Mild cognitive impairment (MCI) is a transitional stage between the cognition of normal aging and mild dementia [Petersen et al. 2001]. MCI patients convert to AD at a rate of 10–25% annually, compared with normal subjects who convert at a rate of 1-2% annually [Petersen et al. 2001]. While MCI can be thought of as a preclinical stage of AD [Petersen et al. 2009], it is in reality a heterogeneous group, with some of the subjects converting to AD faster than others, and others never developping AD, or developping non-AD forms of dementia [Filipovych & Davatzikos 2011]. Numerous AD computational studies treat their data input as a labelled 3-class set (e.g. [Filipovych & Davatzikos 2011]), namely partitioned in controls, MCI subjects, and definite AD subjects. Others use less classes (controls and AD) or more (for example separating MCI into MCI converters and MCI non-converters, [Cuingnet et al. 2011]).

In the next section we shall see how the structural, histological, functional and cognitive symptoms can be utilized so as to assess disease development, and discuss the related notion of biomarkers.

2.2 Biomarkers and Alzheimer's disease

We have seen that the evolution of Alzheimer's disease is strongly correlated with anatomical changes in brain structures. Brain anatomy, plaque development and other symptoms of the disease can be seen as a source of measures and indicators of disease severity. These measures are known collectively as *biomarkers* of disease progression. Ideally, a biomarker should have well-known dynamics during the different stages of the disease and predict clinical evolution [Chen et al. 2011, Frisoni et al. 2010, Mueller et al. 2006, Rothenberg et al. 2011, Tarawneh & Holtzman 2010]. Disease-slowing and symptom prevention treatment research on AD could

greatly benefit from an efficient biomarker, on the basis of which future treatment proposals and potential therapeutic agents could be approved or disapproved. Biomarkers could also help differentiate individuals with preclinical and probable AD from healthy subjects or from patients of other dementing disorders [Goate & Holtzman 2010, Jack Jr. et al. 2003, Mueller et al. 2006, Tarawneh & Holtzman 2010]. CSF $A\beta_{42}$, amyloid imaging, CSF tau and structural MRI are some of the most important biomarkers.

2.2.1 Blood and CSF biomarkers

Several promising markers in the blood and the CSF of individuals with AD have been identified [Tarawneh & Holtzman 2010]. Related studies focus on markers that directly reflect pathological symptoms such as amyloid plaques and neurofibrillary tangles. Other potential candidates include proteins involved in cytoskeletal maintenance, cellular trafficking, cellular stress response, redox homeostasis, transcription, and DNA repair [Tarawneh & Holtzman 2010]. Levels of $A\beta_{42}$ and $A\beta_{42}$ in plasma and cerebrospinal fluid have been among the first and perhaps most widely investigated biomarkers. Several studies have suggested that baseline plasma levels of $A\beta_{42}$ are higher in patients with AD, and that the plasma $A\beta_{42}/A\beta_{40}$ ratio predicts a high risk of progression to dementia in cognitively normal individuals [Tarawneh & Holtzman 2010].

As cerebrospinal fluid is a compartment which is in direct contact with brain parenchyma, CSF biomarkers have shown more promise as reliable and early indicators of disease. The CSF biomarkers with the highest diagnostic accuracy are $A\beta_{42}$, tau, and p-tau.

2.2.2 Structural degeneration as a biomarker

Radiological biomarkers, obtained *in vivo* using neuroimaging, could play the role of the disease progression assessment tool. Alongside cerebrospinal fluid (CSF) biomarkers, brain computational imaging and MRI in particular is the most promising biomarker option in AD [Schmand et al. 2011, Tarawneh & Holtzman 2010], and constitutes an active area of research.

Among the brain areas and tissues of interest which undergo significant structural degeneration are the hippocampus, the CSF-containing lateral ventricles, the amygdala and the entorhinal cortex. Structural data of affected areas can so be a useful biomarker [Cuingnet et al. 2011]. Such data may be a simple function of structural change, as simple as a measure of hippocampal volume, or cortical thickness [Cuingnet et al. 2011], or a much more precise and elaborate input in the form of a shape model. Neurodegeneration is detectable on MRI up to 5 years before clinical expression of the disease [Thompson et al. 2004]. Neurofibrillary tangle development commences early on the hippocampus during AD [Greene et al. 2012]. Among all brain neurodegenerative deformation, hippocampal atrophy is one of the best established and validated [Frisoni et al. 2010]. Combined with the fact that the boundaries of the hippocampus are relatively easier to detect than other AD-related areas [Frisoni et al. 2010], the hippocampus constitutes a very useful region of interest.

On the hippocampus, particular subfields that undergo comparably the most significant levels of degeneration during AD progression have been identified [Costafreda et al. 2011, Csernansky et al. 2005, Greene et al. 2012, Mueller et al. 2011, West et al. 2000]. These are the foremost subfields CA1 and subiculum, which suffer neuronal loss and tangle accumulation most notably on early disease stages. The head and tail of the hippocampus have also been known as hotspots, that show extensive volume reduction [Morra et al. 2010]. A combination of hippocampal head volume with neuropsychological test scores and CSF measures gave as much as 100% prediction accuracy of conversion to AD in a recent study [Greene et al. 2012].

The same study [Greene et al. 2012] raises the issue of the differential effects of AD on the head, body, and tail of the hippocampus as measured. To the end of producing a structural biomarker as informative as possible, studying the pattern of structural evolution of the brain tissue in question in as much detail and precision possible, remains a continuous challenge. This is the general direction of numerous computational studies, including the work presented in the current thesis.

3. Clinical assessment scores

Cognitive impairment is measured with clinical score trials. These are typically based on data gathered through a collection of questions and simple cognitive tasks addressed to the patient, and to collateral sources. The result is recorded in the form of a score scale. Mini-mental state exam (MMSE) [Cockrell & Folstein 2002] and CDR (Clinical dementia rating) [Marcus et al. 2007, Morris 1997, Petersen et al. 2009] are two of the most widely-used tests.

The MMSE evaluates five cognitive domains : orientation, registration, attention and calculation, recall, and language. The maximum MMSE score is 30, corresponding to no cognitive impairment, with lower scores signifying damaged cognition functions. Scores of 24 or lower are generally consistent with dementia [Cockrell & Folstein 2002].

The CDR rates subjects for impairment in each of the six domains : memory, orientation, judgment and problem solving, function in community affairs, home and hobbies, and personal care. CDR equal to 0 indicates no dementia and CDR 0.5, 1, 2, and 3 represent very mild, mild, moderate, and severe dementia, respectively [Morris 1997, Morris et al. 2001].

Other related rating scales include Alzheimer's disease assessment scale (ADAS) [Stonnington et al. 2010], dementia rating scale (DRS) [Stonnington et al. 2010] and the global deterioration stage (GDS) [Petersen et al. 2001].

4. State of the art computational models and approaches for the study of AD

There exists a large body of work in neuroimaging AD-related studies, as well as a variety in the brain imaging methods employed. Modalities such as structural T1-weighted MRI [Cuingnet et al. 2011, Klöppel et al. 2008], positron emission tomography (PET) [Chen et al. 2011, Nordberg 1999], diffusion tensor MRI (DT-MRI) [Bozzali et al. 2002], and morphometry types such as deformation and tensor based morphometry (DBM, TBM) [Koikkalainen et al. 2011, Hua et al. 2008B, Hua et al. 2008A, Hua et al. 2009], are included in the arsenal of methodological tools employed in statistical analysis of AD brain populations. A typical framework in neuroimaging studies, shared among others by the celebrated voxel-based morphometry (VBM), involves comparing control versus pathology groups using an independent multivariate voxel-to-voxel analysis [Ashburner & Friston 2000, Thompson et al. 2004, Hua et al. 2008B, Hua et al. 2008A, Hua et al. 2009]. The intensity of each voxel in the tissue under study is analyzed in comparison with homologous voxels on other anatomies. Voxel intensity may represent a range of local qualities, depending on the modality and morphometry type used. Such voxel-wise, mass-univariate methods cannot capture multivariate dependency in the data, which in turn limits the diagnostic value of such models [Davatzikos 2004]. To overcome this limitation, methods inspired by machine learning and pattern recognition theory are employed [Lemm et al. 2011], which view the image or volume in study as a point in a high-dimensional space and hence can capture multivariate dependency inherently.

In this context, various classical machine learning tools have been used with success. Support

vector machines (SVM) are such an example, used widely as a discriminative model [Bishop 2006, Shawe-Taylor & Cristianini 2004, Lemm et al. 2011]. We have already briefly presented SVM along with the closely related SVDD model in chapter 4 of this thesis. Let us name here a few applications of SVM in MRI neuropathology population analysis, for the sake of illustration : in [Klöppel et al. 2008], SVMs over whole brain MRI anatomical data are used to discriminate Alzheimer's patients from controls ; in [Fan et al. 2007], after feature extraction, feature selection and ROI grouping over structural MRI data, schizophrenia patients are classified to pathological or sane ; in [Golland 2000, Golland et al. 2005], hippocampus-amygdala anatomies segmented from a population of schizophrenia patients and controls are used to train a SVM, with the anatomical information represented as the shape distance transform. In the same work the notion of the *discriminative direction* is introduced, corresponding to the type of anatomical change that has to be applied on a given anatomy, so that it would look like an object of the opposite class. In [Zhou et al. 2011], a hierarchical map of brain ROI correlations is computed given the raw T1-weighted MRI intensity map. After a feature selection step and a supervised, linear dimension reduction step, aiming at preserving discriminative class information, the processed features are fed onto a SVM. In [Mourão-Miranda et al. 2011], a one-class SVM is used to model an fMRI study of depression patients as an outlier detection problem. Significant correlation between model distance from the healthy controls support and a clinical depression measure (HRSD scale) is identified.

Compact representations, and notably manifold learning methods have been very much the center of attention in medical imaging recently. We have presented manifold learning in chapter 3 of this thesis. Especially the nonlinear manifold structure of MR image description spaces has been explored in various works. A nonlinear extension to PCA is presented in [Fletcher 2004] and applied on analysis of DT-MRI data. In [Gerber et al. 2009, Gerber et al. 2010], a generative model is induced over a collection of brain MR images. In [Zhang et al. 2006], the manifold structure of cardiac MRI data is exploited in order to facilitate segmentation ; likewise in [Cao et al. 2011] for the processing of prostate MR images. In [Wolz et al. 2010], atlas-based hippocampus segmentation is performed over a set of brain MRI, with the choice of the proper atlas being guided by a two-dimensional embedding of the input brain images, in order to avoid registration bias. Using a similar argument, in [Hamm et al. 2010] large deformation registration is performed as a synthesis of small deformations that are constrained to correspond to image movements exclusively over the embedding surface. In [Wolz et al. 2010], a discriminative model is built on top of dimension-reduced data. Breathing gating is performed in [Wachinger et al. 2010], where a series of ultrasound images is assigned to corresponding breathing phases, using Laplacian eigenmaps to describe the image time-process. In [Wachinger & Navab 2010], manifold learning is used in multi-modal image registration. Registration targets are represented as manifolds, then are matched using an inexpensive $L2$-based energy term.

5. Conclusion

In this chapter we have briefly discussed elements of human brain anatomy and the relation of structural degeneration of the brain with the progression of Alzheimer's disease. Modelling the shape variability of AD-related volumes of interest, pathological and non-pathological alike, is hence an important issue leading ultimately to a better understanding of disease progression. In the next chapter we shall apply the proposed shape processing pipeline that we have presented in the previous parts of this thesis, to real anatomical data.

Chapter 8

Experiments on real medical data

In this chapter we evaluate the model that we have proposed in the previous parts of this thesis, over real medical data. We have discussed our case study of choice in the previous chapter, where we saw how structural degeneration of various brain tissues and Alzheimer's disease progression closely correlate. We have used two MRI datasets, comprising anatomies of real brain hippocampi and amygdalae. After reviewing the general characteristics of the used datasets, we generate discriminative trajectories in order to underpin significant structural difference between pathological and non-pathological tissues. We also evaluate our model's capacity as a classifier / anomaly detector.

1. Dataset description

Set name	Type	Anatomy	Number of subjects			
			controls	patients	int. group	total
Oasis-LH	real	left hippocampi	91	32	75	198
Oasis-RH	real	right hippocampi	91	32	75	198
UhS-LH	real	left hippocampi	11	15		26
UhS-RH	real	right hippocampi	11	15		26
UhS-LA	real	left amygdalae	11	15		26
UhS-RA	real	right amygdalae	11	15		26

TABLE 8.1 – Overview of real anatomical datasets used in the current work.

We have used two separate studies as a source of anatomical data, namely OASIS, a free-access database [MARCUS et al. 2007], and a study organized at the University Hospital of Strasbourg [PHILIPPI et al. 2012]. Both studies are comprised of structural MR scans of AD

patients and controls. We segmented the whole-brain MR scans to obtain tissues of interest. In table 8.1 we denote Oasis-xx and UhS-xx the real datasets studied in this thesis, corresponding to sets originating from OASIS and the University Hospital of Strasbourg study respectively. Next we present the two studies in more detail. See also figure 8.1 for samples of the studied datasets.

1.1 OASIS datasets

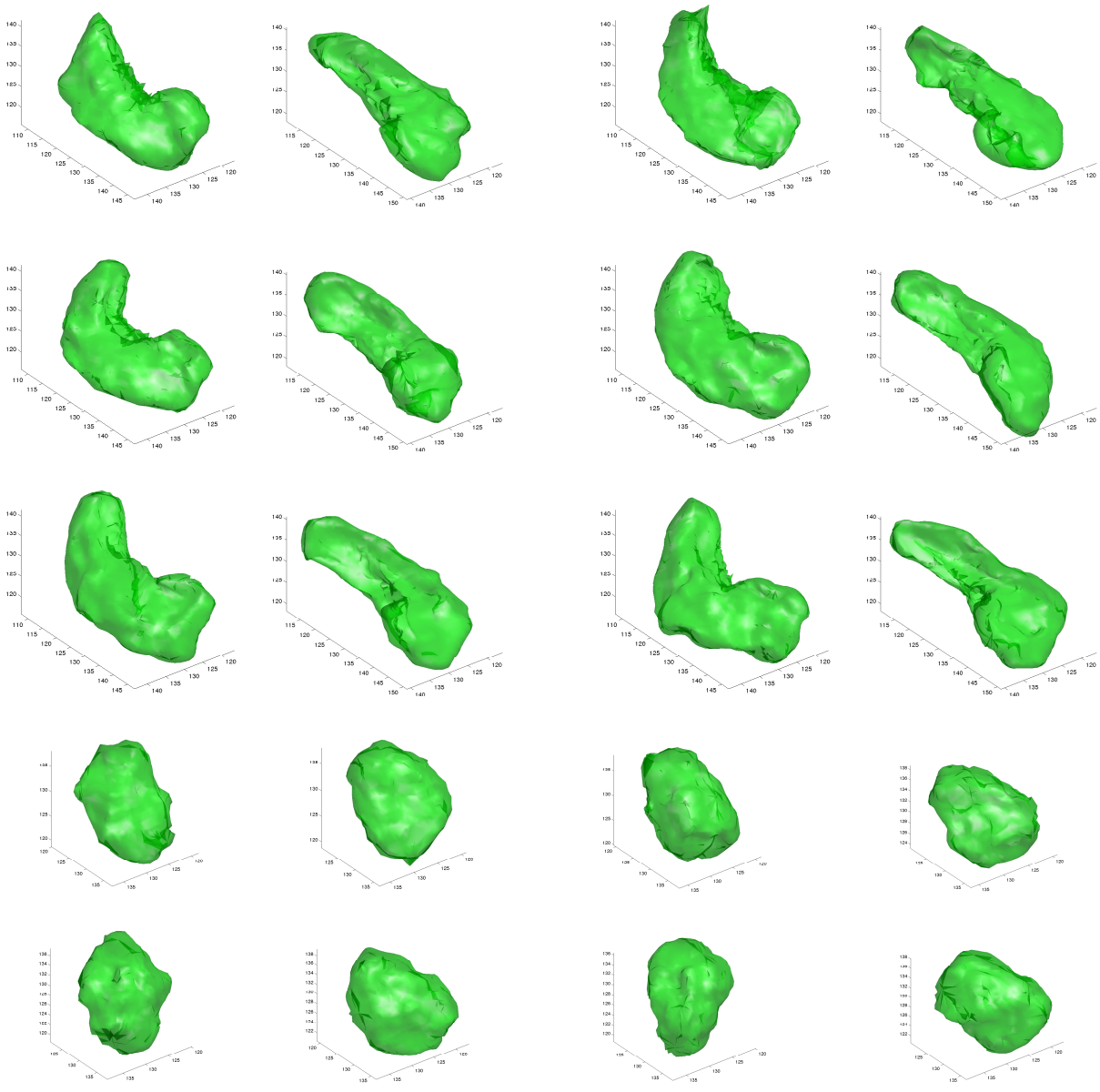


FIGURE 8.1 – Samples of right and left hippocampi and amygdalae from the Oasis and UhS datasets. In rows from top to bottom, we see : Hippocampi of AD patients, hippocampi of intermediate condition subjects, hippocampi of healthy control subjects, amygdalae of AD patients, amygdalae of healthy control subjects. Amygdalae of intermediate condition subjects were not available.

Two real datasets that we use in this thesis are related to the cross-sectional dataset of the OASIS database [MARCUS et al. 2007, GERBER et al. 2010]. *Oasis-LH* and *Oasis-RH* are

made up of OASIS T1-weighted MRI hippocampus volumes for right-handed subjects aged 60 or more. *Oasis-LH* includes left hippocampi and *Oasis-RH* includes right hippocampi. The OASIS datasets presented here include probable AD patients with dementia, and age-matched controls, 198 volumes in total. All OASIS participants with dementia have been diagnosed with probable AD [Marcus et al. 2007]. We chose to exclude younger subjects, as age is a known confounding factor in brain atrophy/dementia studies. We used the *Freesurfer* [Fischl et al. 2002] pipeline to process the OASIS data. Subjects are aligned to the Talairach atlas through an affine registration step. The required hippocampi forms are obtained with subcortical segmentation on the skull-stripped, bias-field treated input. We have partitioned the data into "normal control", "patient" and "intermediate" categories according to their respective clinical/cognitive assessment test results, captured in their MMSE (mini-mental state examination) [Cockrell & Folstein 2002] and CDR (clinical dementia rating) [Marcus et al. 2007, Morris 1997] scores. In our OASIS-based datasets, extreme MMSE and CDR scores, under or equal to 24 and over or equal to 0.5 respectively, were tagged as AD patients. Likewise, MMSE and CDR scores over or equal to 28 and equal to 0 respectively, were tagged as healthy controls. The remaining volumes were tagged as intermediate. We have based this partition on related clinical score statistics in the literature [Hua et al. 2008A]. Note that a partition of patient data based solely on clinical scores is not equivalent to diagnosis by an expert. While a CDR rating of 0.5 or more signifies a clinical diagnosis of AD, so is the case for MCI patients, which are included in the group of patients with a CDR rating equal to 0.5. Hence our partition of *Oasis-LH* and *Oasis-RH* into normal/ intermediate/patients should be perceived as loosely equivalent to a partition into healthy controls/MCI patients/AD patients. For this reason we will not present classification results run on the OASIS datasets using the SVDD model, as we do not in fact possess actual diagnosis ground truth data to differentiate MCI from AD, which would make unfair, to a certain extent, a comparison with classifier evalution scores of other works in the literature.

1.2 University Hospital of Strasbourg datasets

The UhS datasets refere to a study on a number of AD patients and controls conducted at the University Hospital of Strasbourg (Centre Hospitalo-Universitaire de Strasbourg, Service de Neurologie, Unité de Neuropsychologie) [Philippi et al. 2012]. The study includes 15 AD patients and 11 age-matched controls.

The 15 AD patients are aged 69 to 84 years, right-handed, French-native speakers. A diagnosis of probable AD was made according to the criteria of the NINCDS-ADRDA [McKhann et al. 1984] and to the more recent criteria proposed by [Dubois et al. 2007], whose specificity has recently been evaluated and proved to be reliable (i.e. 93%, see [deJager et al. 2010]). The 11 healthy elderly subjects matched for sex ratio, age distribution, education level and handedness were tested. To be included, subjects had to be free of major depression, central neurological disease, cognitive complaint and those with abnormal neurological examination were excluded. The control subjects underwent the Verbal Paired Associates test [Wechsler 1991] and the similarities subtest of the WAIS-R verbal scale in addition to the remote memory tests. Statistical analyses confirmed that the two groups were well matched for both quantitative and qualitative characteristics. The interested reader is referred to [Philippi et al. 2012] for a more detailed description of the study organization and participant characteristics. We used *Freesurfer* [Fischl et al. 2002], as in the OASIS datasets, in order to provide ourselves with segmentations of hippocampi and amygdalae. The resulting 4 datasets related to the University Hospital of Strasbourg study are *UhS-LH*, *UhS-RH*, *UhS-LA*, *UhS-RA*, corresponding to sets of left hippocampi, right hippocampi, left amygdalae, and right amygdalae respectively.

1.3 Estimating the embedding dimensionality

Isomap, which is our manifold learning algorithm of choice, works with anatomy-to-anatomy distances as its input. In order to evaluate the output embedding, we have computed the embedding residual variance [Tenenbaum et al. 2000], as a function of the embedding dimensionality d. We use the definition of residual variance of [Tenenbaum et al. 2000], $1 - \text{corr}^2(D, D_G)$, where corr is the linear correlation coefficient between input distances D and embedding Euclidean distances D_G, over the input dataset. This gives us an empirical way of deciding over the "best" embedding dimensionality. Residual variance captures the amount of divergence between the original distances and the Euclidean distances over the embedding, or put in other words, the distortion required to embed the original distances in a d-dimensional Euclidean space [Gerber et al. 2010]. As noted in [Tenenbaum et al. 2000], we can estimate the intrinsic data dimensionality to lie around the point where variance stops reducing significantly – the so-called "elbow" of the graph. The question of deciding the dimensionality of the manifold contains in itself an important trade-off. Choosing a dimension higher than the actual dimension leads to a non-compact result ; choosing a dimension too low may result in rejecting important data variability.

Embedding evaluation results over the *Oasis-LH* dataset using the deformation-induced distance metric (eq. (3.4)) against a voxel-to-voxel distance are shown in fig. 8.2. The latter distance is defined as the sum of the ℓ_1 norm of the intensity difference, per voxel. We have computed embeddings using a varying number of graph nearest neighbors, with $k \in [2, 20]$ and varying dimensionality, with $d \in [1, 20]$. In every case, we show error-bars corresponding to result variation over k for each dimensionality. A given distance metric induces a particular topology [Munkres 2000], in the sense of particular neighborhood structure in the reduced space, for the manifold. Different metrics induce different topologies, hence also a different manifold. Therefore, the choice of the distance metric affects fundamentally the characteristics and quality of the dimensionality reduction module, and of the model as whole.

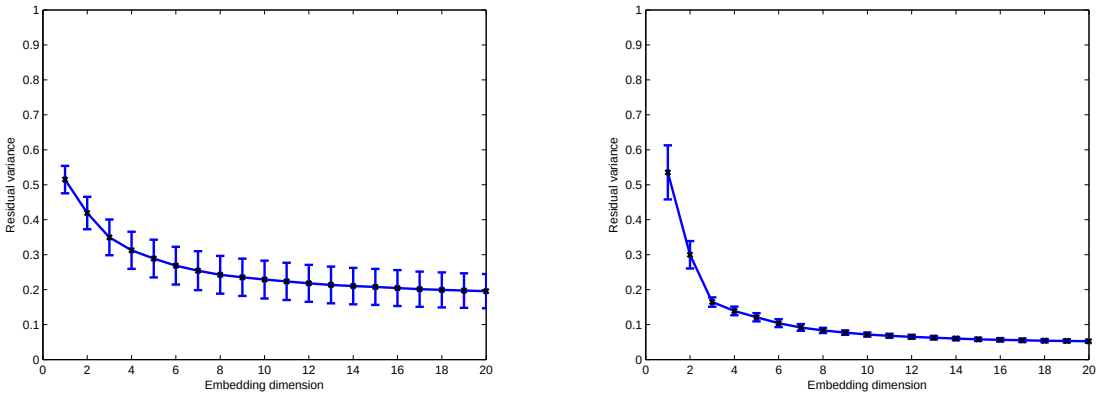


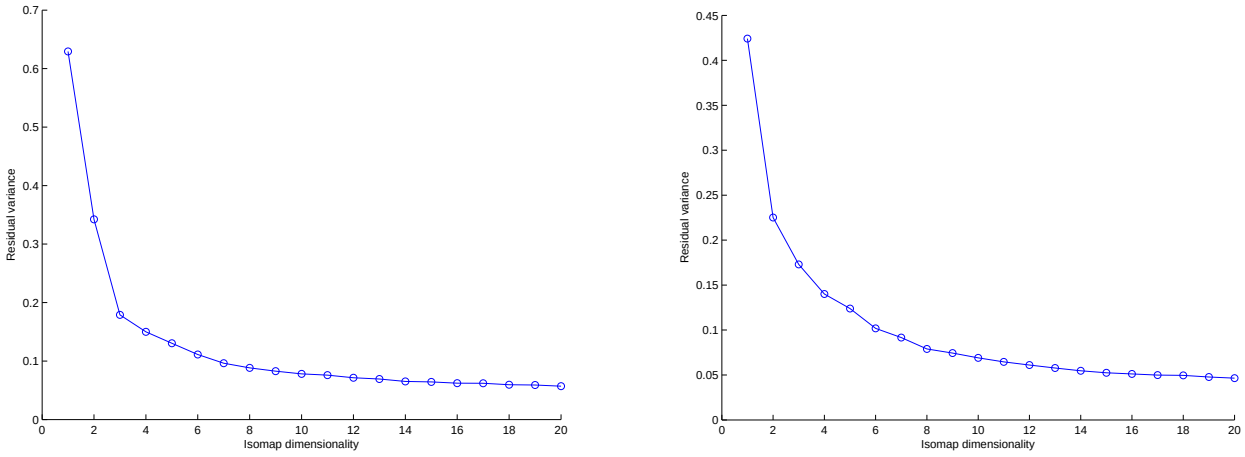
Figure 8.2 – Estimated embedding residual variance, given different distance metrics and different Isomap parameters. The horizontal axis corresponds to varying embedding dimension. The error-bars correspond to variance introduced by varying the number of neighbors in Isomap. The graph on the left shows results using a voxel-to-voxel ℓ_1 distance over registered intensities. The graph on the right shows results using the deformation-field induced distance (eq. (3.4)).

The results in fig. 8.2 show unstable –regarding the Isomap number of neighbors parameter– ℓ_1-induced embedding statistics compared to the ones for the shape-based distance. On the former we have no clear residual "elbow", while on the latter case we do, namely around $d = 3$. Shape-based distance Isomap is also almost invariable with regard to parameter k. These observations suggest that our data indeed form a low-dimensional embedding when using the

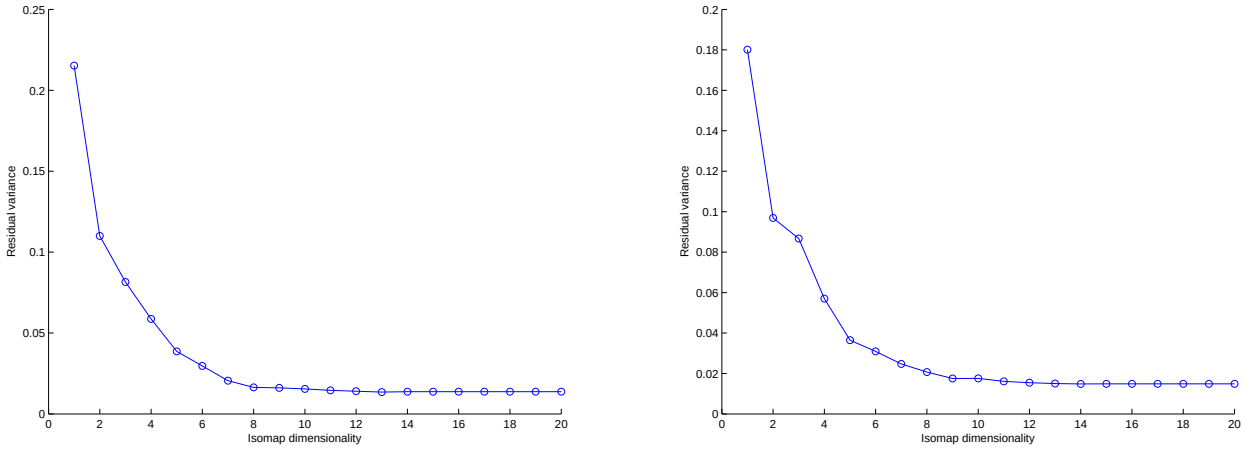
shape-based distance and induced topology, while this is relatively more uncertain when it comes to the voxel-to-voxel based distance case. The Isomap result –be it in terms of the residual graph or the embedding coordinates in each case– confirms that the two metrics do not capture the same topology. The choice of metric will thus depend entirely on the nature of application at hand. Each metric compares shapes in a different sense, capturing different qualities. In this work we shall be most interested in capturing shape variation and evolution, so we shall be using the deformation-field induced distance (eq. (3.4)).

Isomap parameters choice Following the "elbow" heuristic [Tenenbaum et al. 2000], we can estimate the true dimensionality of the embedded manifold to be most probably 3, for the deformation-field induced distance. In related work, the most useful results come for embeddings for a similar scale dimension magnitude [Wolz et al. 2010], with d as low as 2 [Gerber et al. 2010, Wolz et al. 2010]. In fig. 8.3 we see residual variance graphs for all real datasets used in this thesis. We note that the elbow is consistently around dimension 3 for residual graphs of the Oasis datasets. The UhS datasets show an elbow around dimension $1-3$. This implies that in all cases the main causes of variability are only a small number of elementary types of variation. We have chosen $d = 3$ for the Oasis datasets, and $d = 2$ for the UhS datasets.

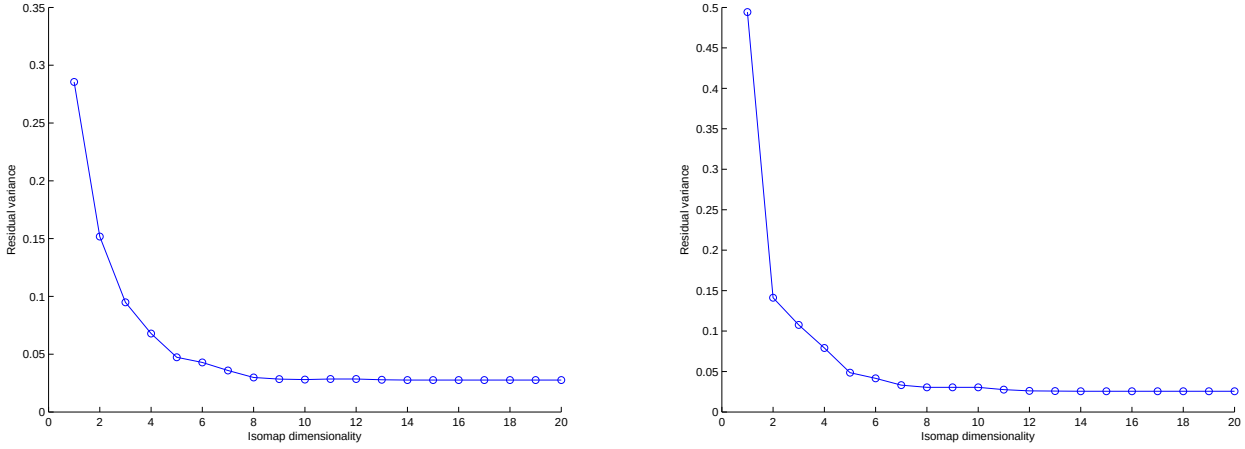
In the experiments performed in this work, unless otherwise specified, we use $k = 5$ nearest neighbors for the Dijkstra step of Isomap and the deformation-based induced distance metric as Isomap input.



(a)



(b)



(c)

FIGURE 8.3 – Residual variance graphs. (a) *Oasis-LH* and *Oasis-RH*. (b) *UhS-LH* and *UhS-RH*. (c) *UhS-LA* and *UhS-RA*.

2. Hippocampus discriminative trajectories and p-value significance maps

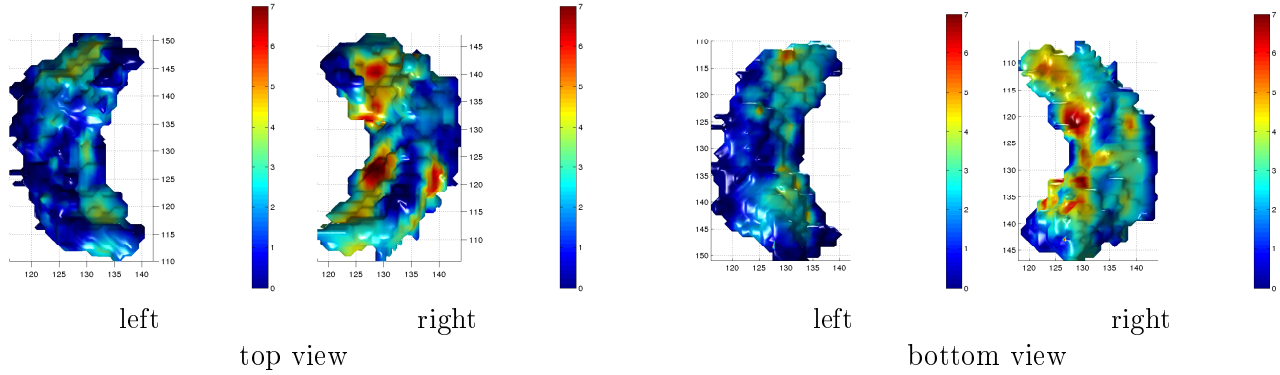


FIGURE 8.4 – p-value significance maps for the *Oasis-LH* and *Oasis-RH* dataset. Each couple of figures show a left hippocampus and a right hippocampus template with overlaid p-values. Minus log values are shown in all cases. Red corresponds to highly significant zones, blue corresponds to insignificant zones.

Under the hypothesis that Alzheimer's disease incurs a single deformational pattern on the hippocampus, i.e. the same hippocampal areas are more or less affected in all AD patients and under more or less the same pattern, we have computed a single discriminative trajectory. As with the synthetic sets *Synth-I* and *Synth-II* (see chapter 6), as starting point of the trajectory we set the point that lies the farthest from the normal support in the reduced space. Here this point corresponds to subject $OAS1\_0373\_MR1$, aged 80, female, with clinical scores MMSE = 20, CDR = 1. These scores are, as expected, evident of quite severe dementia. They describe a patient that has progressed through the major stages, if not the totality, of the illness, both clinically and anatomically (in the sense of hippocampal anatomy). While the discriminative trajectory does correspond to this subject in particular, in practice the reconstructed samples along the trajectory comprise anatomical information of neighboring subjects due to the regression model. This discriminative trajectory can thus be considered representative of the class difference between normal and abnormal subjects, to a certain extent.

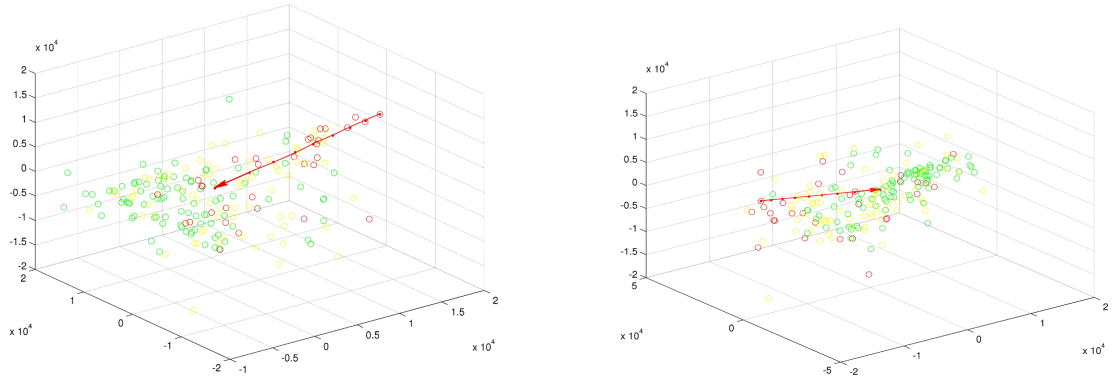
	Oasis-LH		Oasis-RH	
	corr. coefficient	p-value	corr. coefficient	p-value
SVDD function - MMSE	0.2717	1.08e-4	0.2200	1.0e-3
SVDD function - CDR	-0.2943	0.28e-4	-0.1926	6.6e-3

TABLE 8.2 – Correlation coefficients between the SVDD decision function and MMSE/CDR scores, as evaluated over all subjects of *Oasis-LH* and *Oasis-RH*. The p-value here is the probability of getting a correlation as large as the observed value by random chance, when the true correlation is zero.

We have computed correlation coefficients between clinical scores and the SVDD decision function (eq. (4.5)), over all subjects included in *Oasis-LH* and *Oasis-RH*. The SVDD decision function is positive for subjects lying inside the normal support and negative for the rest. The results of table 8.2 show a degree of correlation between the SVDD decision function and MMSE/CDR scores. Since movement over the discriminative trajectory is movement from negative SVDD values towards more positive ones, the discriminative trajectory correlates – albeit does not identify– with disease progress as reflected in MMSE/CDR scores.

In figures 8.6 and 8.7, we see reconstructed left hippocampi sampled along the discriminative

trajectory and following the direction from pathology towards the normal/healthy class. In figure 8.5(a) we see the Isomap $\mathbb{R}^3$ embedding for the *Oasis-LH* set, with the discriminative trajectory overlaid. As clearly observed in fig. 8.6, movement away from normality corresponds to volume loss. In particular, we note widespread inward deformation while moving towards pathology



(a) (b)

FIGURE 8.5 – Isomap embeddings for the (a) *Oasis-LH* and *Oasis-RH* datasets. Green points correspond to healthy controls. Yellow points correspond to intermediate subjects. Red points correspond to AD patients. The arrow corresponds to the discriminative trajectory towards the healthy elements support. Points along the arrow-trajectory correspond to samples of figures 8.6, 8.7 for *Oasis-LH*, and 8.8, 8.9 for *Oasis-RH*.

foremost at the hippocampus anterior and posterior surfaces (head and tail). This result is also captured by the significance map of fig. 8.4 (see 2.2 on how this map was computed). The areas showing significant deformation can also be largely identified with the cornu ammonis 1 (CA1) and subiculum regions (see also fig. **??** for a reference on hippocampus subfields).

Apparently the same areas are more or less affected, either on trajectory points near the normal support, i.e. mildly atrophic hippocampi (roughly early stages of the disease), or on trajectory points far from the support, i.e. severely atrophic hippocampi (roughly late stages of the disease). The findings above are consistent with well-known findings in related research (for example [CSERNANSKY et al. 2005, COSTAFREDA et al. 2011, MORRA et al. 2010, GREENE et al. 2012, THOMPSON et al. 2004, WEST et al. 2000]). Concerning the state of the deformation pattern with respect to the atrophy severity, given here by the discriminative trajectory, in related literature the head and the tail of the hippocampus [MORRA et al. 2010] and the CA1 region [WEST et al. 2000] are indeed hotspots in the progression of AD, acknowledged to show signs of atrophy early on. Hippocampal atrophy that is related with normal aging, is on the contrary more confined to the subiculum and the dentate gyrus [WEST 1993]. Atrophy on CA1 has been associated in general with impairment of late retrieval and consolidation [MUELLER et al. 2011], and is in general correlated with AD onset. We have in the same manner processed the right hippocampi dataset *Oasis-RH*. In figure 8.5(b) we see the Isomap $\mathbb{R}^3$ embedding for the *Oasis-RH* set, with the discriminative trajectory overlaid. The results are largely analogous to the results for the left hippocampi. The trajectory corresponds again to a visible volume loss, and inward deformation is present on roughly the same areas as those of the left hippocampi case (fig. 8.4, 8.8, 8.9).

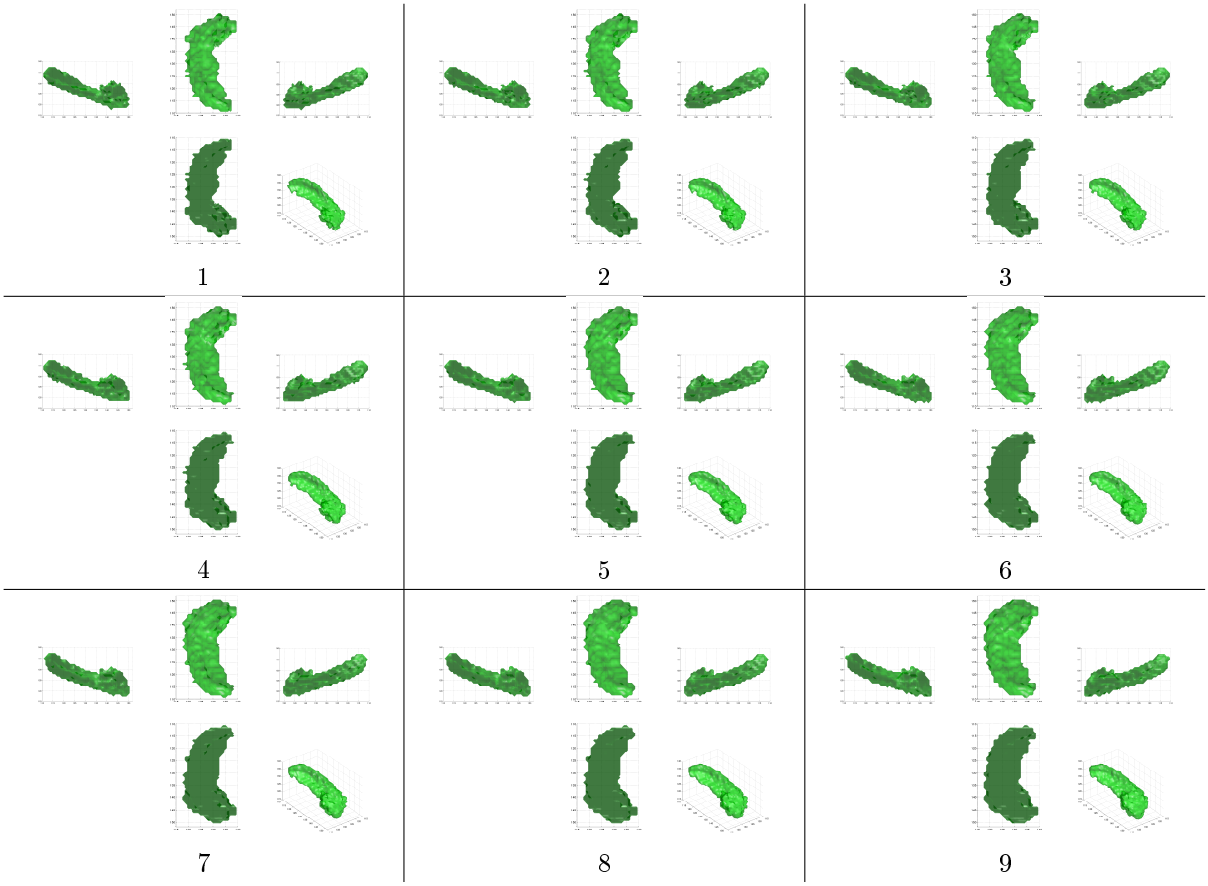


FIGURE 8.6 – *Oasis-LH* discriminative trajectory samples. From 1 to 9 the samples correspond to movement from extreme pathology gradually towards a healthy, non-atrophic anatomy.

3. Using the proposed model for hippocampus and amygdala classification

We have used the UhS datasets to test our model as a 1-class classifier. In figure 8.10 we see the reduced dimension embeddings for the University Hospital of Strasbourg datasets. We have evaluated classification as the percentage of correctly classified data. We remind the reader that in the *UhS-xx* datasets we have data partitioned in only two classes, which are AD patients and healthy controls. SVDD is trained assuming the healthy controls as normal, as also done previously. The evaluation mark was computed during the phase of cross-validating for model parameters C and kernel bandwidth σ. The mean for repeated leave-10%-out cross validation trials is shown in table 8.3. Note that state of the art classification results for AD patient-control hippocampus datasets are typically around the $80\% - 85\%$ correct classification ratio mark [CUINGNET et al. 2011, WOLZ et al. 2010]. We also note the corresponding results for a SVM model instead of SVDD. SVM/Isomap corresponds to using SVM over dimensionality reduced data, and SVM corresponds to using SVM directly on ambient space data. The same notation applies to SVDD, with SVDD/Isomap corresponding to the proposed model, and SVDD corresponding to a direct application on ambient space data.

As expected, SVM results are overall superior to those of SVDD. This is due to the SVM formulation, which exploits all information available in a 2-class setting, while the SVDD hypersphere is a direct function of only the normal training set. There is work in the literature that extends the SVDD so that it can produce a competitive classifier especially in multiclass (i.e. more than 2 classes) classification problems (for example [LEE & LEE 2007] or [BAN &

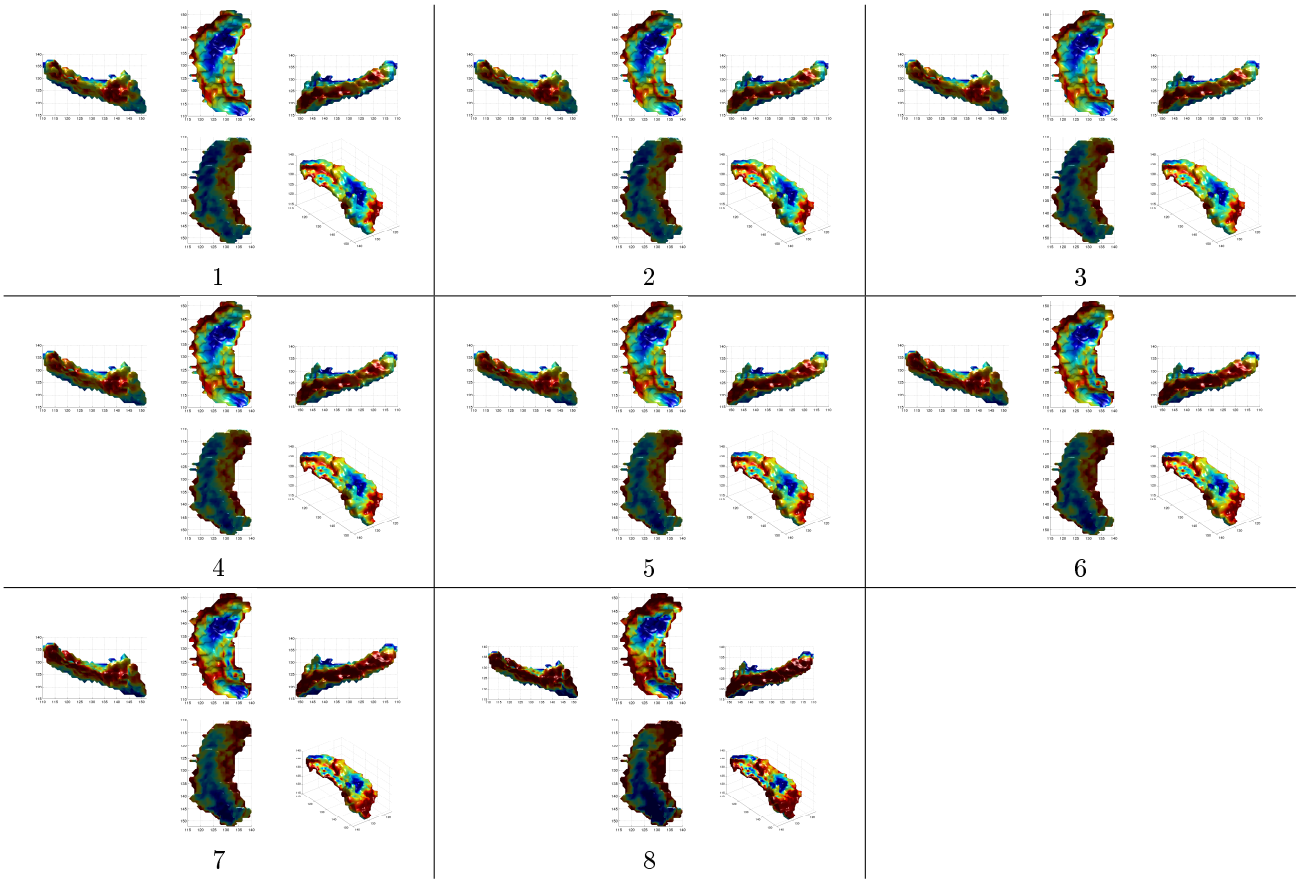


FIGURE 8.7 – *Oasis-LH* discriminative trajectory swelling/contraction maps. From 1 to 8 the samples correspond to movement from extreme pathology gradually towards a healthy, non-atrophic anatomy. Red areas correspond to outward deformation as we are moving towards the healthy controls support (from the first towards the last frame), or inward deformation as we are moving towards the opposite direction. Blue areas correspond to inward deformation as we are moving towards the healthy controls support (from the first towards the last frame), or outward deformation as we are moving towards the opposite direction. Green areas correspond to no change.

	SVDD	SVDD/Isomap	SVM	SVM/Isomap
UhS-LH	0.85	0.96	0.96	1.00
UhS-RH	0.81	0.86	0.88	0.92
UhS-LA	0.83	0.84	0.88	0.92
UhS-RA	0.79	0.87	0.92	0.96

TABLE 8.3 – Classification accuracy results. Figures show percentage of correctly classified data in the UhS datasets. SVDD and SVM were run directly over data on the ambient space, with no dimensionality reduction. SVDD/Isomap is the proposed model.

ABE 2006, KANG et al. 2006, YANG et al. 2011]). An extended discussion would be out of the scope set in this thesis, as using the proposed model as a classifier is not our primary goal. We shall limit ourselves to note the application of such multiclass extensions of the SVDD on practical problems similar to the ones discussed here (for example classification in 4 classes : AD, MCI converters, MCI non-converters, and healthy controls class).

Looking at the results of table 8.3 as classification without isomap versus classification with isomap, we conclude that working with the manifold hypothesis is always more beneficial than

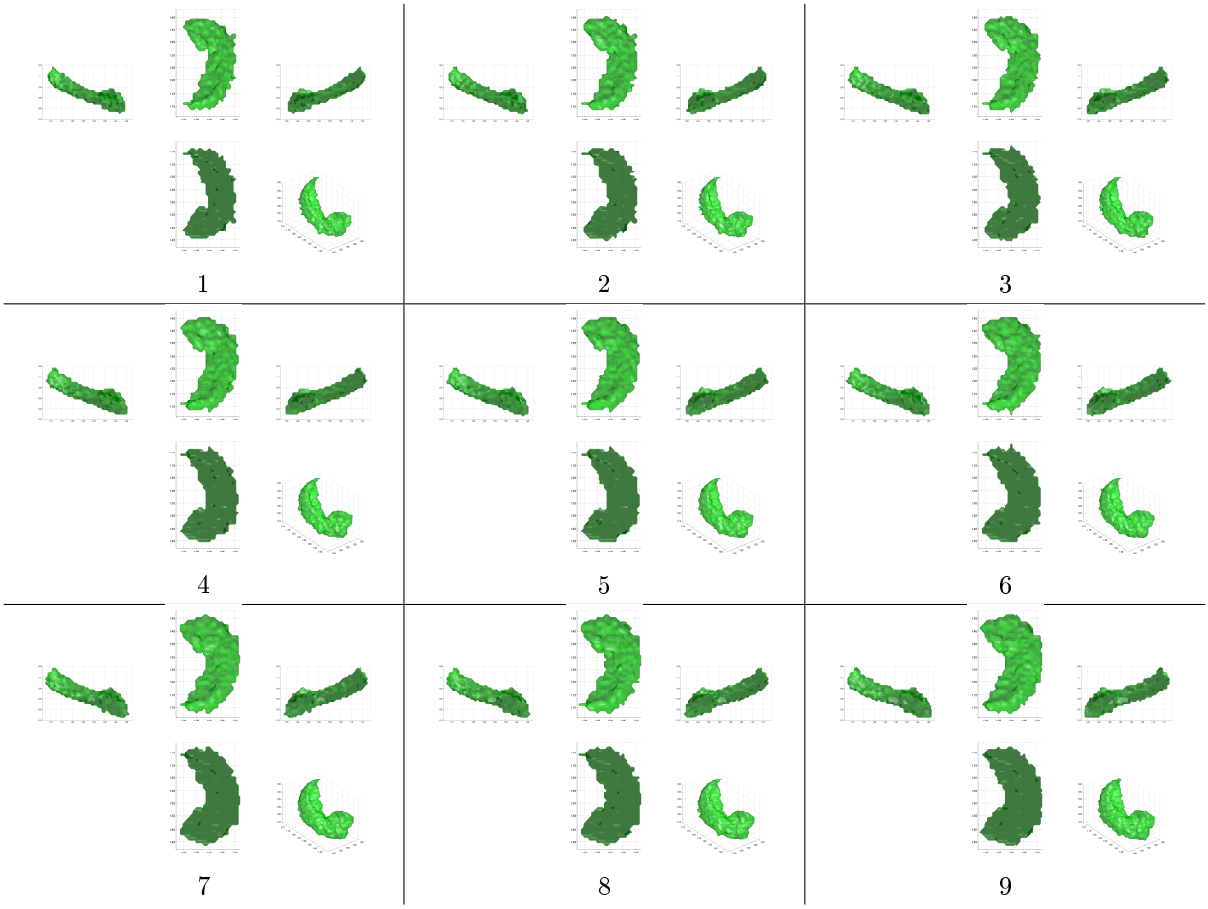


FIGURE 8.8 – *Oasis-RH* discriminative trajectory samples. From 1 to 9 the samples correspond to movement from extreme pathology gradually towards a healthy, non-atrophic anatomy.

not, regardless of the discriminative base model used. This confirms our premise set earlier in this thesis (chapter 3). It also confirms previous work in the literature where manifold learning is used to create a better SVM classifier [WOLZ et al. 2010], and extends it for the SVDD classifier case.

4. Conclusion

With this chapter we finish the last part and the main presentation theme of this thesis. We have applied the proposed shape processing pipeline in real data. Discriminative trajectory results over the hippocampus have shown a discriminative pattern that appears coherent with previous results in AD-related research.

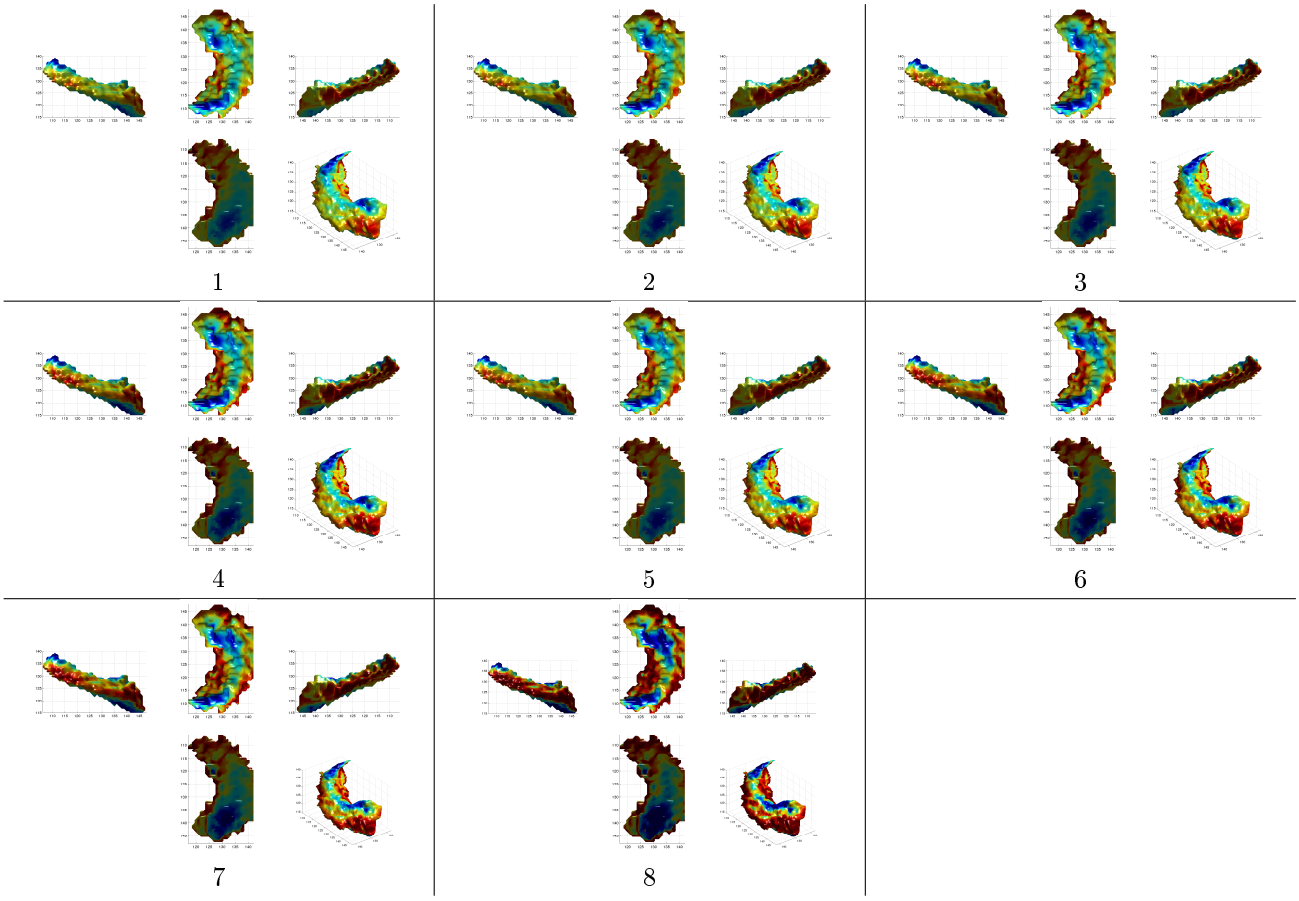


FIGURE 8.9 – *Oasis-RH* discriminative trajectory swelling/contraction maps. From 1 to 8 the samples correspond to movement from extreme pathology gradually towards a healthy, non-atrophic anatomy. Red areas correspond to outward deformation as we are moving towards the healthy controls support (from the first towards the last frame), or inward deformation as we are moving towards the opposite direction. Blue areas correspond to inward deformation as we are moving towards the healthy controls support (from the first towards the last frame), or outward deformation as we are moving towards the opposite direction. Green areas correspond to no change.

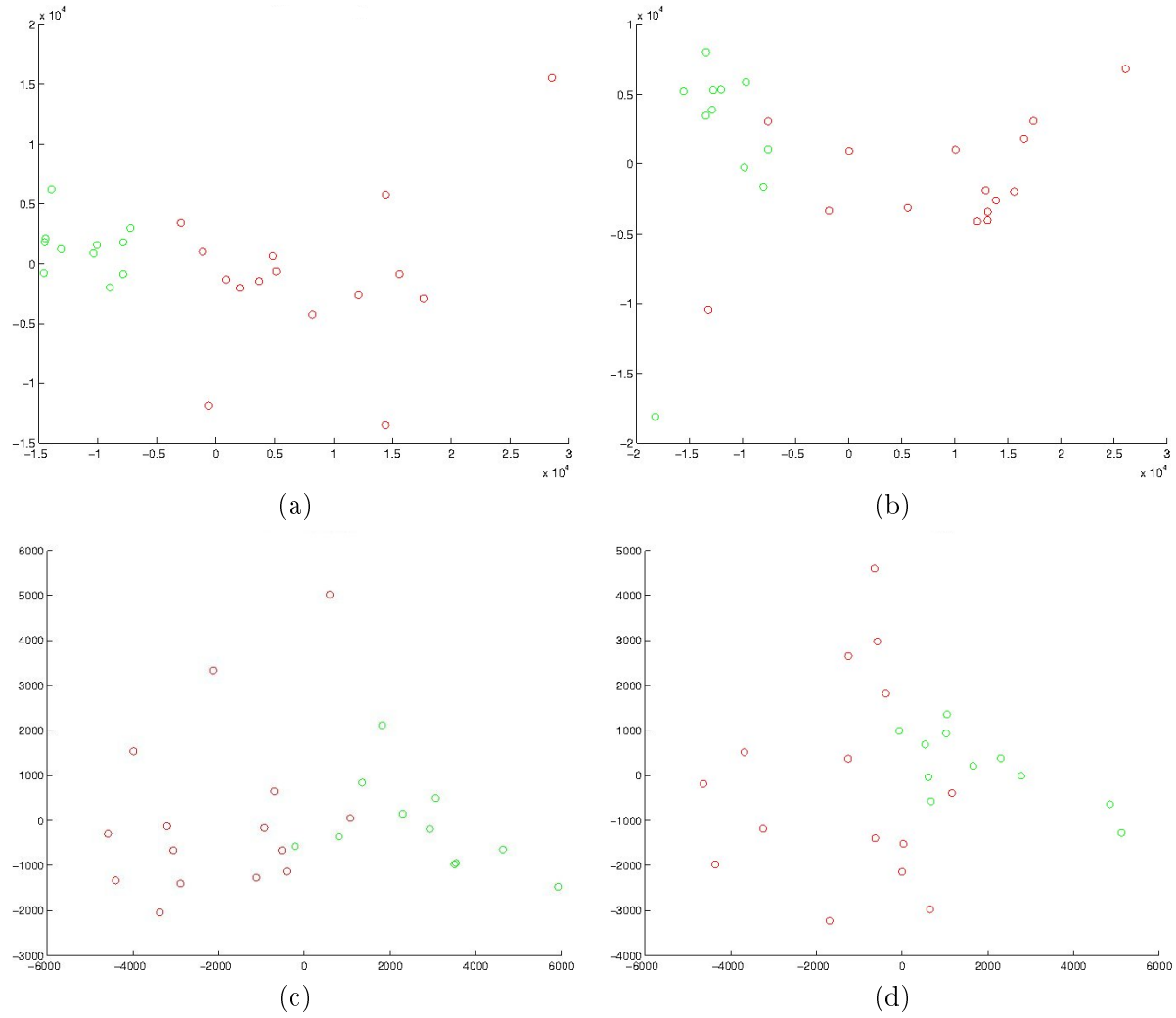


FIGURE 8.10 – Isomap embeddings for the UhS datasets. (a) *UhS-LH* (left hippocampi) (b) *UhS-RH* (right hippocampi) (c) *UhS-LA* (left amygdalae) (d) *UhS-RA* (right amygdalae). Green points correspond to healthy controls. Red points correspond to AD patients.

Conclusion and perspective

In this thesis, we have presented a non-linear, non-Gaussian compact model capable of describing shape. Given a set of training shapes labelled as normal, we have estimated the distribution of these elements in the shape description space, using SVDD theory. The estimated normal support was then used to discriminate normal data from abnormal, or put in the medical imaging context, healthy tissues from pathological ones. We have also introduced the discriminative trajectory, as a tool to estimate discriminative difference. Instrumental for the utility of these tools is always the underlying data distribution, modelled as a non-linear manifold.

To the extent that the manifold learning step captures the underlying shape data structure correctly, the resulting model is compact, and leads to a more accurate description. On the one hand, newly generated shapes are anatomically correct, since they are constrained to lie on the hypothesized manifold. On the other hand, discriminative trajectory moments are inherently also constrained to lie over the hypothesized manifold, leading to an accurate discriminative difference estimate. We have illustrated this point on geometrical shape and synthetic anatomical datasets. In the real anatomical data scenario, the results showed coherence with results previously presented in related literature.

We have confirmed that during Alzheimer's disease, particular tissue regions and subfields are damaged the most, like the the head and the tail of the hippocampus. This has been shown using the discriminative trajectory as well as p-value significance maps. Discriminative trajectories capture discriminative difference as a subject-to-class process, while the p-value map gives a class-to-non class difference summary. As far as the discriminative trajectory is concerned, let us note that the trajectory parameter is *not* identical to a time parameter. At best, it could be perceived as correlated to time. We have indeed seen that the SVDD decision function correlates to a certain extent with clinical assessment scores, which in turn reflect disease progress. We have not identified dramatical difference in the hippocampal subregions affected with respect to different discriminative trajectory moments, which would indicate towards a changing pattern in AD-related atrophy, as a function of disease progress. In perspective, a longitudinal dataset, and more importantly a model capable of handling such data, could shed more light and lead to modelling of shape as a direct function of time.

The SVDD model used as the discriminative component of the proposed model is essentially one-class based, however it can be in perspective extended to a multiclass model. Related experimental results have shown that multiclass SVDD extensions compete fairly against state of the art multiclass discriminative models [Lee & Lee 2007, Yang et al. 2011].

In the experiments run here we assumed that the observed, ambient space is the space of deformation vector fields given a single template. Combined with OLS regression, this constitutes a fast and practical solution. However, using a single template when dealing with complex populations introduces registration bias, which can have a damaging effect on the analysis [Blezek & Miller 2007, Sabuncu et al. 2009, Koikkalainen et al. 2011, Sfikas et al. 2010A]. A suitable representation that could fit the role of the ambient space of our model, could relax the related bias. Another perspective is to use input information other than shape to describe and

discriminate subjects, as done for example in [Aljabar et al. 2010], where different distance metrics are fused in order to produce a better embedding ; or integrate alternate shape matching frameworks such as the *currents* framework [Vaillant & Glaunes 2005]. Concerning again the proposed model, the regression mechanism could be replaced with a more efficient, robust approach such as the Relevance Vector Machine (RVM). Originally proposed for regression over scalar output data [Tzikas et al. 2006], they were recently extended to multivariate outputs in the context of pose estimation [Thayananthan et al. 2006]. The basic idea in RVM is to weigh down the effect that some of the basis functions have in the regression, or, in the limit, disregard them altogether. The inputs that correspond to the remaining basis vectors are the *relevance* vectors. Compared to the simple linear model, RVM results in a sparse represenation of the interpolation function while exhibiting comparable generalization error. The need to lighten computational cost makes RVM an attractive solution for biomedical data, which are typically very demanding in terms of both disk space and processing power. These issues could also be attenuated using compressive sensing [Lustig et al. 2007] on the volume representation part of the model, or as part of the regression mechanism.

Part IV

Appendix

Appendix A

Bayesian models for medical, natural and polarimetric image processing

In this appendix, we present a number of works that we have co-authored during the first part of this thesis. The general theme of this set of works is using Bayesian spatially varying mixture models to perform image segmentation (chiefly) and image restoration. This subject is organically unrelated to as much as thematically different from the main part of this thesis, to quite a great extent. Hence, we have decided to present these works in an appendix, in a brief yet concise manner.

The models that we shall present here cover aspects of machine learning, different from the ones presented in the main part of the thesis. All models are hierarchical, generative models based on a special form of a Gaussian mixture model, which assumes different model statistics for different image pixels -hence the title spatially-varying. Basic theoretical tools that are used in the models that follow include Expectation-Maximization (EM), Variational Inference (VI), Majorization-Minimization (MM). We use a graphical model illustration in each subsection to describe the corresponding model structures concisely.

The work presented here partly covers work published in [Sfikas et al. 2008B, Sfikas et al. 2008A, Sfikas et al. 2010B, Sfikas et al. 2011A, Sfikas et al. 2011B].

1. Introduction

In this section we present an outline of the works in image processing that we discuss in this appendix. A common trait of all proposed models is that they are generative, in the sense that they describe a mechanism of image generation. In all cases the (only) observed quantity is the image itself, and solving each of the models amounts to estimating the exact values or distributions of the model parameters. The observed image in model terms is expressed as a set of N image cues, denoted as the set $X = \{x^n\}_{n=1..N}$.

The input image is perceived as a set of K segments (classes). Each of these segments is governed by different statistics over its features, centered around a mean cue value. This is expressed as a Gaussian distribution over X, given class information. The class each image pixel belongs to is encoded with the $Z = \{z_j^n\}_{n=1..N,j=1..K}$, $\Pi = \{\pi_j^n\}_{n=1..N,j=1..K}$ sets of vectors. Vector set Z encodes which pixel belongs to which segment directly, while each vector in Π encodes the probability of the given pixel being part of any given segment. This relation between sets Z and Π is expressed as a probability distribution of Z given Π.

In order to enforce spatial smoothness, we define a Markov random field (MRF) prior over Π, which effectively encodes our assertion that spatially neighboring pixels are more likely

to belong to the same segment. The MRF prior includes the possibility of not enforcing smoothness, if an edge is detected. Edge information about the image is captured by variables $U = \{u_j^{nk}\}_{n=1..N,j=1..K,k=1..\Gamma}$ (and $\ell = \{\ell^{nk}\}_{n=1..N,k=1..\Gamma}$ in section 2) and their respective parameters.

The required segmentation or restoration is obtained by solving the model, or in other words estimating the *a posteriori* distributions (*a posteriori* in the sense of after having observed the image data) of the aforementioned hidden variables and parameters. Depending on the envisaged application at hand, our objective goal translates here to estimating distributions of Z (segmentation) or $S = \{s^n\}_{n=1..N}$ (actual Stokes image). This task, due to the complexity of the models, is in all cases directly intractable. To this end, an approximation scheme is employed. Expectation-Maximization, Majorization-Minimization and Variational Inference are used, the choice of which depends on the particular structure and accompanying inference difficulties of each model. To the extent that it is computationally possible, the full posterior distribution is computed, or a MAP estimate for each model variable.

Each particular application discussed in sections 2, 3, 4, 5 of this appendix shares the core generative model outlined above. In section 2 the MRF spatial relationships are extended to a $3D$ grid, so as to handle $3D$ image inputs, such as structural MR images. In section 3 the model is applied on natural $2D$ image data.

The base generative model assumes a fixed number of segments. In section 4 the segmentation model is extended so as to allow for the automatic estimation of the number of segments. Alongside the set Π, which expresses mixing probabilities at a local spatial level, we add a set $\Omega = \{\omega_j\}_{j=1..K}$ of global mixing probabilites. This allows for a more efficient pruning of unnecessary components.

In section 5 we propose a model for segmentation and restoration of polarimetric images. The observed variables $G = \{g^n\}_{n=1..N}$ are distorted versions of a true Stokes image S. Each pixel of the Stokes image is constrained under a specific range of values. This is expressed as a non-linear relation between each of its four channels. Both the required restoration S and the segmentation Z appear as hidden variables in the model hierarchy, and are estimated by solving the model, as before.

A summary of the graphical models for the works that we discuss in this appendix is given in table A.1.

2. MR brain tissue classification using an edge-preserving spatially variant mixture model

2.1 Introduction

The segmentation of 3D brain magnetic resonance (MR) images into the three main types, namely, white matter (WM), gray matter (GM) and cerebro-spinal fluid (CSF) is of great importance in most neuroimaging applications. Although many research studies have been presented in this area, MRI brain segmentation still remains a challenging issue due to specific difficulties of MRI, such as intensity inhomogeneity, partial volume effect and acquisition noise. A first approach to the problem relied on the expectation-maximization (EM) algorithm [Kapur et al. 1996, Wells III et al. 1996] which led to an important category of methods resorting to Gaussian mixture models (GMM). Among them, many studies incorporate prior information (e.g. anatomical atlases) on tissue topology [Marroquin et al. 2002, Prastawa et al. 2004, Tasdizen et al. 2005, Pohl et al. 2007] or constrain the segmentation to be spatially smooth and take into account edge discontinuities (e.g. using Markov random field (MRF) priors) [Niessen et al. 1998, Laidlaw et al. 1998, Kapur et al. 1998, Van Leemput et al. 1999, Van Leemput

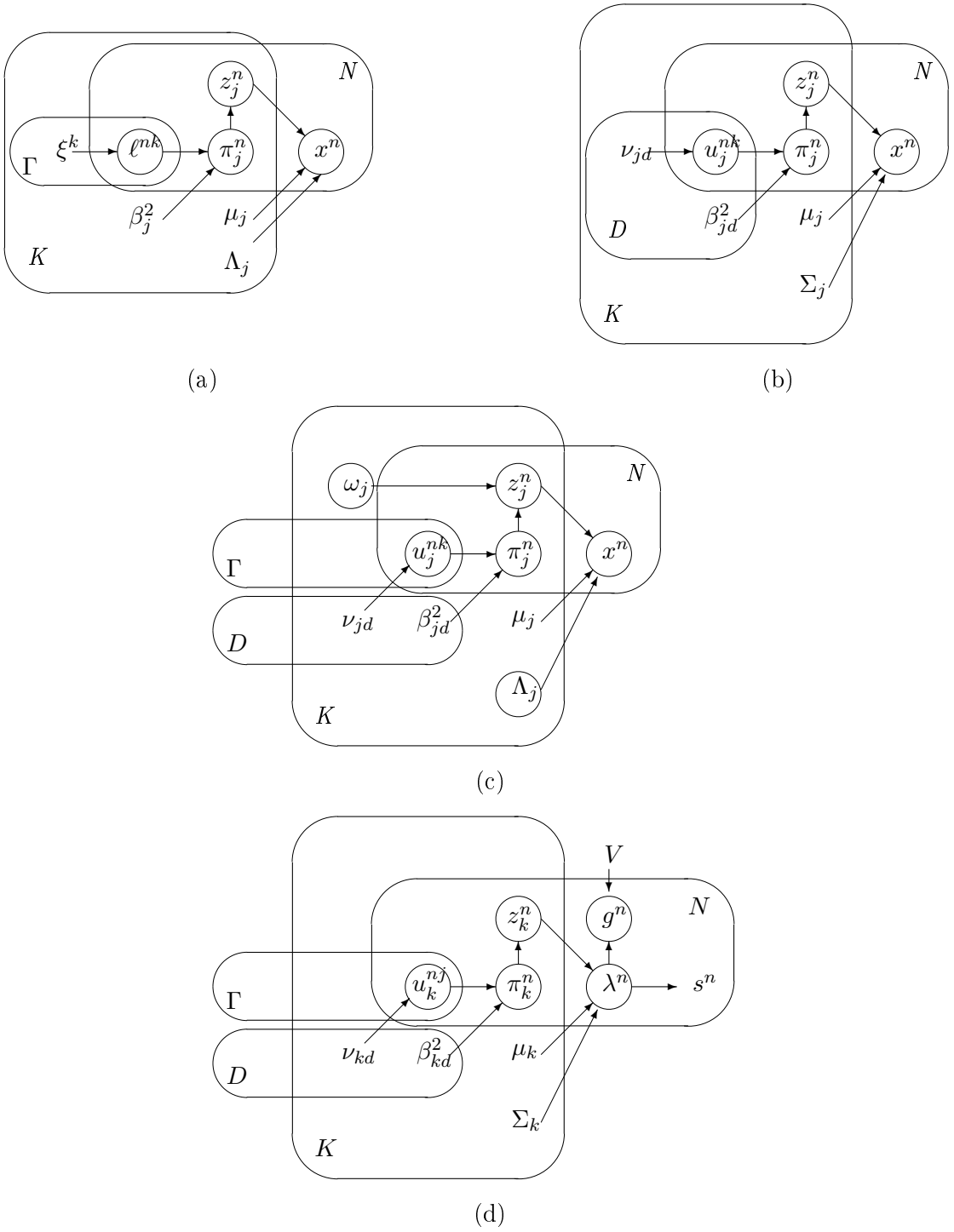


TABLE A.1 – Summary of graphical models for the Bayesian models for image processing, discussed in this chapter. The structure of all models treats the input image as a group of segments (classes) each with separate feature statistics, and enforces spatial smoothness through an edge-preserving MRF prior. (a) Model applied to $3D$ MR images for classification in 3 tissue types (see section 2). This is closely related to (b), a model applied to the segmentation natural images (see section 3). (c) Model extended to allow for the automatic determination of the number of image segments (see section 4). (d) Model applied to the joint segmentation and restoration of polarimetric images (see section 5).

et al. 2001, Greenspan et al. 2006, Peng et al. 2006, Awate et al. 2006, Awate et al. 2007].

Modeling the probability density function (pdf) of pixel or voxel attributes with a GMM [Bishop 2006] is a natural way to cluster data because it automatically provides a grouping based on the components of the mixture that generated them. Furthermore, the likelihood of a GMM is a rigorous metric for clustering performance. The parameters of the GMM can be estimated very efficiently through maximum likelihood (ML) estimation using the Expectation-Maximization (EM) algorithm [Bishop 2006].

The prior knowledge that adjacent pixels most likely belong to the same cluster is not used in standard GMM. To overcome this shortcoming, spatial smoothness constraints may be imposed, generally applying an MRF prior, like the spatially varying Gaussian mixture model (SVGMM) in [Nikou et al. 2007]. However, this model enforces smoothness between pixels belonging to different classes. Since the seminal work in [Geman & Geman 1984], *line processes* were also introduced in several applications, other than brain tissue classification, to respond to this drawback, see for instance [Molina et al. 2003] and [Kanemura et al. 2007A, Kanemura et al. 2007B] for image restoration and superresolution respectively.

In this section, we propose a Bayesian, spatially varying Gaussian mixture model for the classification of brain images to the three tissue types (WM, GM, CSF). The main contribution of the model is that it takes into account not only that adjacent voxels are more probable to belong to the same class but it also prohibits smoothing across boundary voxels. Motivated by the studies in brain image segmentation incorporating MRF-based prior knowledge [Tasdizen et al. 2005, Niessen et al. 1998, Laidlaw et al. 1998, Kapur et al. 1998, Van Leemput et al. 1999, Van Leemput et al. 2001, Greenspan et al. 2006, Peng et al. 2006, Awate et al. 2006, Awate et al. 2007] we impose proper hyperpriors to simultaneously address local smoothing and edge preservation. The main difference with other, state-of-the-art methods imposing MRF-type priors, is that the constraint is imposed on the probabilities of the voxel labels (generally known in mixture modeling as *contextual mixing proportions*) and not onto the labels themselves. By these means, closed form solutions are provided for the model parameters through variational inference.

2.2 Model description and solution

The K-kernel spatially varying GMM differs from the standard GMM in the definition of the mixing proportions. More precisely, in the SVGMM, each voxel x^n, $n = 1, ..., N$ has a distinct vector of mixing proportions denoted by π_j^n, $j = 1, ..., K$. We call these parameters *contextual mixing proportions* to distinguish them from the mixing proportions of a standard GMM. Hence, the probability of a distinct voxel is expressed by :

$$f(x^n; \pi, \mu, \Lambda) = \sum_{j=1}^{K} \pi_j^n \mathcal{N}(x^n; \mu_j, \Lambda_j), \tag{A.1}$$

where $0 \leq \pi_j^n \leq 1$, $\sum_{j=1}^{K} \pi_j^n = 1$ for $j = 1, 2, ..., K$ and $n = 1, 2, ..., N$, μ_j are the Gaussian kernel means and Λ_j are the Gaussian kernel precision (inverse covariance) matrices.

We now assume that the voxels $X = \{x^1, x^2, ..., x^N\}$ are independent and generated by the graphical model shown in figure A.1.

Note that a set $Z = \{z_j^n\}_{n=1..N, j=1..K}$ of $N \times K$ latent variables is introduced, in order to make inference tractable for the model. The Z variables are distributed multinomially :

$$p(Z|\Pi) = \prod_{j=1}^{K} \prod_{n=1}^{N} (\pi_j^n)^{z_j^n}, \tag{A.2}$$

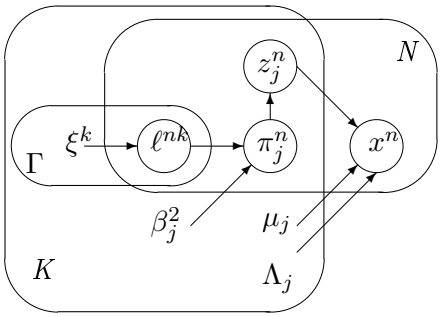


FIGURE A.1 – Graphical model for the edge preserving spatially variant Gaussian mixture model. Superscript $n \in [1, N]$ denotes voxel index, subscript $j \in [1, K]$ denotes kernel (segment) index and $\Gamma \in [1, 26]$ describes the neighborhood direction type.

where each z^n is a binary vector, with $z_j^n = 1$ if datum n is generated by the j-th kernel and $z_j^n = 0$ otherwise.

Considering the set of *contextual mixing proportions* Π as random variables and assuming a proper prior, we can incorporate the intuitive fact that neighboring voxels are more likely to share the same class label. We impose an edge preserving Gaussian prior on Π :

$$p(\Pi|\ell;\beta) = \prod_{j=1}^{K}\prod_{n=1}^{N}\prod_{k=1}^{\Gamma}\mathcal{N}(\pi_j^n|\pi_j^k;\beta_j^2)^{\ell^{nk}}, \tag{A.3}$$

where ℓ^{nk} is a binary random variable we call *line-process*. If $\ell^{nk} = 1$, then there is a link on the random field between the voxel indexed n and its k-th *possible* neighbor (we denote by π^k the k-th neighbor of specific voxel n). Otherwise, if $\ell^{nk} = 0$ there is no link between them, signifying the presence of an edge. We assume that two voxels can be possible neighbors when they are vertically, horizontally or diagonally adjacent with regard to their spatial location in the three-dimensional mesh, implying $\Gamma = 26$ neighbors per voxel.

Parameters $\beta = \{\beta_1^2, \beta_2^2, ..., \beta_K^2\}$ control the spatial smoothness of the *contextual mixing proportions*. The prior in eq. (A.3) implies that

$$\pi_j^n - \pi_j^k \sim \mathcal{N}(0, \beta_j^2), \;\; \forall k \in [1, \Gamma] \mid \ell^{nk} = 1,$$

reflecting the fact that the *contextual mixing proportions* which implicitly control voxel class membership are similar for neighboring voxels except in case there exists an edge.

We now regard the *line process* variables ℓ^{nk} as Bernoulli distributed random variables, governed by a parameter set $\xi = \{\xi^1, \xi^2, ..., \xi^\Gamma\}$:

$$p(\ell|\xi) = \prod_{n=1}^{N}\prod_{k=1}^{\Gamma} p(\ell^{nk}|\xi^k) = \prod_{n=1}^{N}\prod_{k=1}^{\Gamma} {\xi^k}^{\ell^{nk}}(1-\xi^k)^{(1-\ell^{nk})}. \tag{A.4}$$

The Beta distribution is the conjugate for the Bernoulli pdf, therefore, we impose it on the ξ parameters :

$$p(\xi;\alpha_{\xi 0},\beta_{\xi 0}) = \prod_{k=1}^{\Gamma}\frac{\Gamma(\alpha_{\xi k0}+\beta_{\xi k0})}{\Gamma(\alpha_{\xi k0})\Gamma(\beta_{\xi k0})}\xi^{k(\alpha_{\xi k0}-1)}(1-\xi^k)^{(\beta_{\xi k0}-1)}. \tag{A.5}$$

The main advantage of the model in fig. A.1 is that (i) it takes into account that neighboring voxels are more probably generated by the same Gaussian pdf and (ii) it does not smooth adjacent voxels separated by an edge.

To perform segmentation, the evidence with respect to model parameters has to be optimized :

$$\underset{\mu,\Lambda,\Pi,\beta}{\operatorname{argmax}} \ln p(X,\Pi;\mu,\Lambda,\beta).$$

This MAP solution cannot be computed directly, or even estimated using the EM algorithm, due to the Π prior distribution complexity. Therefore, we resort to variational inference [Bishop 2006]. This leads to an iterative scheme with one step for the computation of the stochastic parameters Z, ℓ, ξ and Π and one step for the deterministic parameters μ, Λ and β. Due to lack of space we present here the final expressions.

The expected values of the stochastic parameter are

$$<z_j^n>= \tilde{\pi}_j^n, \quad <l^{nk}>= \tilde{\xi}^{nk}, \quad <\ln \xi_k>= \psi(\alpha_{\xi k}) - \psi(\alpha_{\xi k}+\beta_{\xi k}),$$

$$<\ln(1-\xi_k)>= \psi(\beta_{\xi k}) - \psi(\alpha_{\xi k}+\beta_{\xi k}),$$

where $\psi(\cdot)$ is the digamma function and the expectations (denoted by a tilde) being as follows, with $sig(x) = (1+e^{-x})^{-1}$:

$$\tilde{\pi}_j^n = \frac{\pi_j^n \mathcal{N}(x^n;\mu_j,\Lambda_j)}{\sum_{l=1}^K \pi_l^n \mathcal{N}(x^n;\mu_l,\Lambda_l)},$$

$$\tilde{\xi}^{nk} = sig\left(\sum_{j=1}^K \ln \mathcal{N}(\pi_j^k|\pi_j^n;\beta_j^2) + <\ln \xi_k> - <\ln(1-\xi_k)>\right),$$

$$\tilde{\alpha}_{\xi k} = \alpha_{\xi 0} + \sum_{n=1}^N <l^{nk}>, \quad \tilde{\beta}_{\xi k} = \beta_{\xi 0} + \sum_{n=1}^N <1-l^{nk}> .$$

The *contextual mixing proportions* π_j^n are computed as the roots of a quadratic equation :

$$a_j^n \left(\pi_j^n\right)^2 + b_j^n \left(\pi_j^n\right) + c_j^n = 0, \tag{A.6}$$

with coefficients :

$$a_j^n = -\sum_{k=1}^\Gamma <l^{nk}>, \quad b_j^n = \sum_{k=1}^\Gamma <l^{nk}> \pi_j^k, \quad c_j^n = \frac{<z_j^n>\, \beta_j^2}{2}.$$

The form of the coefficients guarantees that there is always a non negative solution. The solutions of eq. (A.6) however will not in general satisfy the constraint $\sum_{j=1}^K \pi_j^n = 1, \pi_j \geq 0, \forall j \in [1..K]$ so we project the corresponding π^n vectors $\forall n \in [1..N]$ onto the constraints subspace ; this is done using the quadratic programming algorithm described in [Blekas et al. 2005].

Furthermore, the deterministic parameters are also obtained in closed form :

$$\tilde{\mu}_j = \frac{\sum_{n=1}^N <z_j^n> x^n}{\sum_{n=1}^N <z_j^n>}, \quad \tilde{\Lambda}_j^{-1} = \frac{\sum_{n=1}^N <z_j^n> (x^n-\mu_j)(x^n-\mu_j)^T}{\sum_{n=1}^N <z_j^n>}, \tag{A.7}$$

$$\tilde{\beta}_j^2 = \frac{\sum_{n=1}^N \sum_{k=1}^\Gamma <l^{nk}> (\pi_j^n - \pi_j^k)^2}{\sum_{n=1}^N \sum_{k=1}^\Gamma <l^{nk}>}. \tag{A.8}$$

The above update equations, for both the stochastic and deterministic parameters, are considered for the full range of each of the indices, namely n, j and k and are computed iteratively until convergence [Boyd & Vandenberghe 2004] of the variational lower bound.

2.3 Experimental results

We have evaluated the proposed model on simulated images with known ground truth from the BrainWeb database [Collins et al. 1998], [Kwan et al. 1999] using the voxel intensities as features. Prior to segmentation, we have preprocessed each volume so that only WM, GM and CSF are included (fig. A.2). Hence, we set the number of kernels in our model to $K = 3$. The hyperparameter values of the Beta prior distribution were set to $\alpha_{\xi k0} = \beta_{\xi k0} = 1$, $\forall k$, making the prior uninformative as the data size $N \gg 1$.

The algorithm was applied to a simulated T1-weighted data without any bias field and with intensity noise levels between 0% and 9%. The noise percentages were defined with respect to the mean intensity of each tissue class. We have compared our segmentation results with two of the state of the art methods of [Van Leemput et al. 1999] and [Tasdizen et al. 2005]. In both of these studies, the Dice metric was used for evaluation. Therefore we present our results using this performance measure. Figure A.2 summarizes the Dice metrics for the compared methods. In that figure, the curves for the state-of-the-art methods are reproduced from the respective publications [Van Leemput et al. 1999, Tasdizen et al. 2005]. As it can be observed, in all cases, our method provides better segmentations with respect to the method in [Van Leemput et al. 1999]. Also, for low level noise the Dice metric of the proposed method is higher with respect to the method proposed in [Tasdizen et al. 2005]. On the other hand, the method of [Tasdizen et al. 2005] performs better for noise levels of 7% and 9%. However, our method takes no more than 50 minutes to run on a 2.7 GHz standard PC whereas the method in [Tasdizen et al. 2005] requires at least *six* hours runtime for convergence.

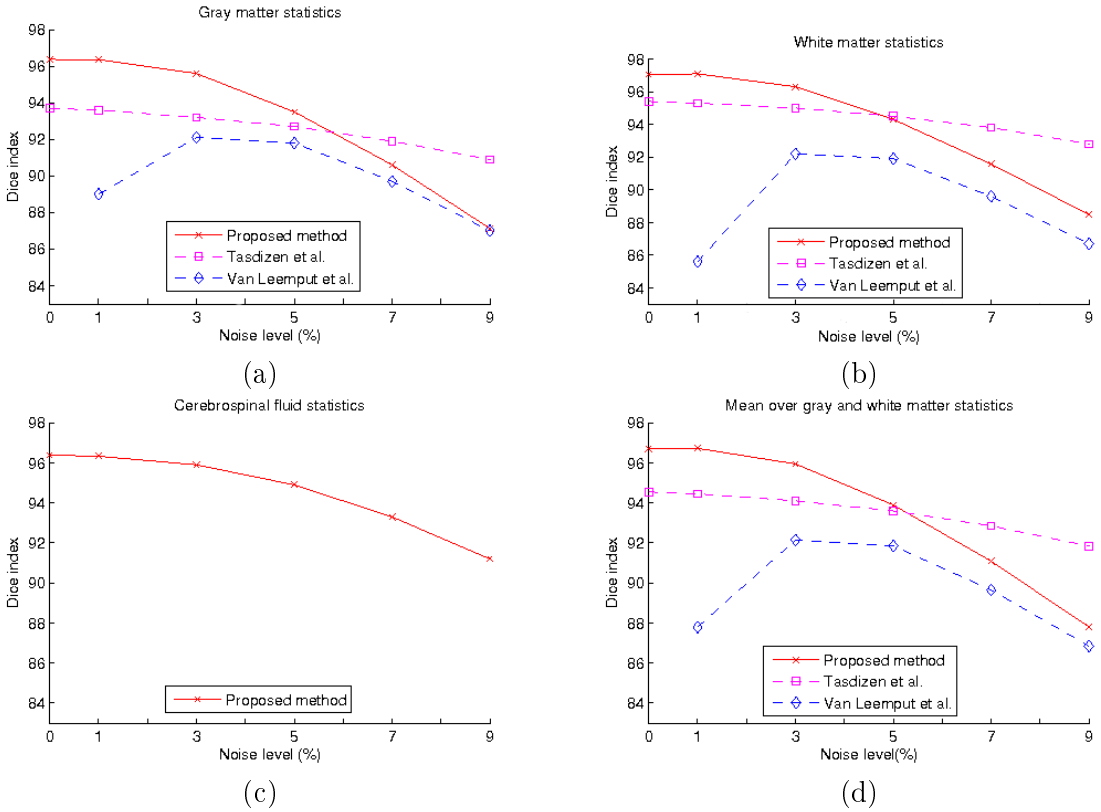


Figure A.2 – Dice metric as a function of noise level for data without bias field. (a) Gray matter, (b) white matter, (c) CSF, (d) mean over gray and white matter. The dashed lines plot the results for the same images as presented in [Tasdizen et al. 2005] and [Van Leemput et al. 1999]. Results are not provided by the respective publications for the case in (c).

We have also compared our Bayesian model to a standard GMM as well as to the spatially

varying GMM (SVGMM) proposed in [Nikou et al. 2007]. In all cases, our model performed better than both methods. Both of the spatially varying models provided Dice metrics significantly better than the standard GMM. Moreover, the difference between our model and the SVGMM [Nikou et al. 2007] is small but consistently in favor of our method. These differences are underpinned in table A.2. A representative segmentation example is presented in figure A.3.

Noise (%)	*Bayes-SVGMM*	*SVGMM*	*GMM*
0	96.6	96.5	77.2
1	96.6	96.5	89.8
3	96.0	95.9	95.3
5	94.2	94.2	94.1
7	91.8	91.7	90.8
9	88.9	88.8	86.8

Table A.2 – Mean values for the Dice metric as a function of noise level over the three tissue types. The compared methods are the proposed method (Bayes-SVGMM), the non edge preserving spatially varying GMM (SVGMM) proposed in [Nikou et al. 2007] and a standard GMM.

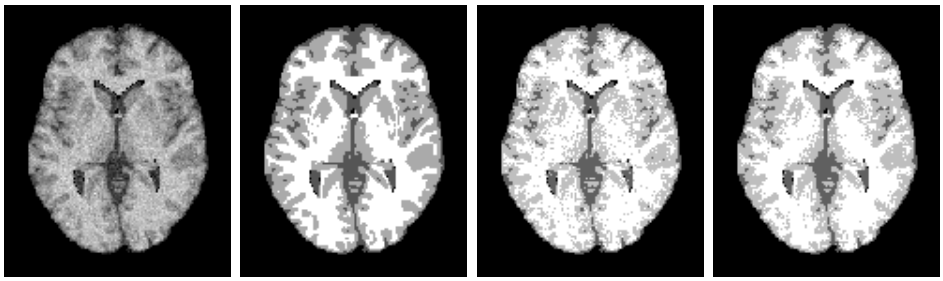


Figure A.3 – Axial slices of a $3D$ segmentation example. From left to right : original MR slice with (9%) noise, ground truth, segmentation using a GMM, and the proposed method.

2.4 Conclusion

We have presented a framework for segmenting the brain anatomy from 3D MRI. The proposed method relies on a Bayesian finite mixture model with a Gauss-Markov random field prior on the probabilities of the pixel labels. Also, the model incorporates a probabilistic line process for edge preservation. The quantitative evaluation reveals that the method not only improves the standard GMM and refines the SVGMM [Nikou et al. 2007] but also performs at least at the same level as other state-of-the-art methods. A perspective of this study is the extension of the model to include more brain tissues and to integrate bias field correction into the segmentation procedure.

Methodologically, the model described in this section is related to the natural image segmentation model presented in the next section.

3. Edge-preserving spatially varying models for image segmentation

Clustering-based image segmentation methods rely on arranging data into groups having common characteristics [Xu &Wunsch II 2005]. During the last decade, the main research directions in the relevant literature have been focused on mixture model [Sanjay-Gopal & Hebert 1998, Nikou et al. 2007], graph theoretic approaches [Shi & Malik 2000,

Felzenswalb & Huttenlocher 2004, Zabih & Kolmogorov 2004], methods based on the mean shift algorithm [Comaniciu & Meer 2002, Carreira-Perpinan 2006] and rate distortion theory techniques [Yang et al. 2007].

Research efforts in imposing spatial smoothness for image segmentation can be grouped into two categories. In the methods of the first category, spatial smoothness is imposed on the discrete *hidden* variables of the FMM that represent class labels [Marroquin et al. 2003, Zabih & Kolmogorov 2004]. These approaches may be categorized in a more general area involving simultaneous image recovery and segmentation which is better known as *image modeling* [Rajagopalan & Chaudhuri 1999]. More specifically, spatial regularization is achieved by imposing a discrete MRF on the classification labels of neighboring pixels that penalizes solutions where neighboring pixels belong to different classes. Inference in this category of models is non trivial and generally performed through EM-like alternating optimization, Markov Chain Monte Carlo (MCMC) or inexact variational inference.

In the second category of methods, the MRF-based smoothness constraint is not imposed on the labels but on the contextual mixing proportions, that is on the probabilities of the pixel labels. This model is called spatially variant finite mixture model (SVFMM) [Sanjay-Gopal & Hebert 1998] and avoids the inference problems of DMRFs. For instance, in [Nikou et al. 2007], a new family of smoothness priors was proposed for the contextual mixing proportions based on the Gauss-Markov random fields that takes into account cluster statistics, thus enforcing different smoothness strength for each cluster. The model was also refined to capture information in different spatial directions. In these models, maximum *a posteriori* (MAP) estimation of the contextual mixing proportions via the MAP-EM algorithm is possible. However, the main disadvantage is that smoothness is imposed in the neighborhood of each pixel without taking into account that the respective pixel may be an edge pixel or its neighborhood consists of edge pixels.

In this section, we present a new hierarchical Bayesian model for mixture model-based image segmentation with spatial constraints. This model assumes that the local differences of the *contextual mixing proportions* follow a Student's t-distribution. The generative model of Student's t-distribution contains two levels. The lower level is a Gaussian pdf with precision (inverse variance) varying with each pixel and the higher level a Gamma pdf. The varying precisions with each pixel of the Gaussians of this model capture the *local image variations* and thus allow the smoothness constraints to incorporate the image edge structure.

A MAP-EM algorithm was used for Bayesian inference with this model. An important feature of this algorithm is that *all the necessary parameters* are estimated from the data. Thus, the proposed segmentation algorithm is *automatic* in the sense that it does not require empirical selection of parameters like other state-of-the-art methods (n-cuts [Shi & Malik 2000], mean-shift [Comaniciu & Meer 2002]).

The model was extensively evaluated on the 300 images of the Berkeley image data base and was compared with other GMM based methods that do not require parameter selection. More specifically, it compared favorably to standard GMM and to GMM with "standard" spatial smoothness constraints [Nikou et al. 2007].

3.1 Model description and solution

The K-kernel spatially varying GMM [Sanjay-Gopal & Hebert 1998, Nikou et al. 2007] differs from the standard GMM [Bishop 2006] in the definition of the mixing proportions. More precisely, in the SVGMM, each pixel x^n, $n = 1, ..., N$ has a distinct vector of mixing proportions denoted by π_j^n, $j = 1, ..., K$, with K being the number of Gaussian kernels. We call these parameters *contextual mixing proportions* to distinguish them from the mixing proportions

of a standard GMM. Hence, the probability of a distinct pixel is expressed by :

$$f(x^n;\pi,\mu,\Sigma)=\sum_{j=1}^{K}\pi_j^n\mathcal{N}(x^n;\mu_j,\Sigma_j), \tag{A.9}$$

where $0\leq\pi_j^n\leq 1$, $\sum_{j=1}^{K}\pi_j^n=1$ for $j=1,2,...,K$ and $n=1,2,...,N$, μ_j are the Gaussian kernel means and Σ_j are the Gaussian kernel covariance matrices.

Generally, in image processing and computer vision, we assume that, conditioned on a hidden variable Z, pixels $X=\{x^1,x^2,...,x^N\}$ are independent and Gaussian-distributed :

$$p(X|Z)=\prod_{j=1}^{K}\prod_{n=1}^{N}\mathcal{N}(x^n|\mu_j,\Sigma_j)^{z_j^n}, \tag{A.10}$$

where the set of $N\times K$ latent variables $Z=\{z_j^n\}_{n=1..N,k=1..K}$ is introduced to make inference tractable for the model. The Z variables are distributed multinomially :

$$p(Z|\Pi)=\prod_{j=1}^{K}\prod_{n=1}^{N}(\pi_j^n)^{z_j^n}, \tag{A.11}$$

where each z_n is a binary vector, with $z_j^n=1$ if datum n is generated by the j-th kernel and $z_j^n=0$ otherwise.

Considering the set of *contextual mixing proportions* Π as random variables and assuming a proper prior, we can incorporate the intuitive fact that neighboring pixels are more likely to share the same class label. We assume a Markov random field on Π, which equivalently means that Π is governed by a Gibbs distribution [Geman & Geman 1984] generally expressed by :

$$p(\Pi)\propto\prod_{C}e^{-\psi_c(\Pi)},$$

where ψ_c is a function on clique c, called *clique potential* function in the literature, and the product is over all minimal cliques of the Markov random field.

In the herein proposed model, we consider clique potential functions imposing local differences of *contextual mixing proportions* to follow a univariate Student's t-distribution.

A d-dimensional random variable X follows a multivariate t-distribution, $X\sim\mathcal{S}t(\mu,\Sigma,\nu)$, with mean μ, positive definite, symmetric and real $d\times d$ covariance matrix Σ and has $\nu\in[0,\infty)$ degrees of freedom when [Bishop 2006], given the weight u, the variable X has the multivariate normal distribution with mean μ and covariance Σ/u :

$$X|\mu,\Sigma,u\sim\mathcal{N}(\mu,\Sigma/u), \tag{A.12}$$

and the weight u follows a Gamma distribution parameterized by ν :

$$u\sim\mathcal{G}(\nu/2,\nu/2). \tag{A.13}$$

Integrating out the weights from the joint density leads to the density function of the marginal distribution :

$$p(x;\mu,\Sigma,\nu)=\frac{\Gamma\left(\frac{\nu+d}{2}\right)|\Sigma|^{-\frac{1}{2}}}{(\pi\nu)^{\frac{d}{2}}\Gamma\left(\frac{\nu}{2}\right)\left[1+\nu^{-1}\delta(x,\mu;\Sigma)\right]^{\frac{\nu+d}{2}}}, \tag{A.14}$$

where $\delta(x,\mu;\Sigma)=(x-\mu)^T\Sigma^{-1}(x-\mu)$ is the Mahalanobis squared distance and Γ is the Gamma function [Bishop 2006]. It can be shown that for $\nu\to\infty$ the Student's t-distribution tends to a Gaussian distribution with covariance Σ. Also, if $\nu>1$, μ is the mean of X and if $\nu>2$,

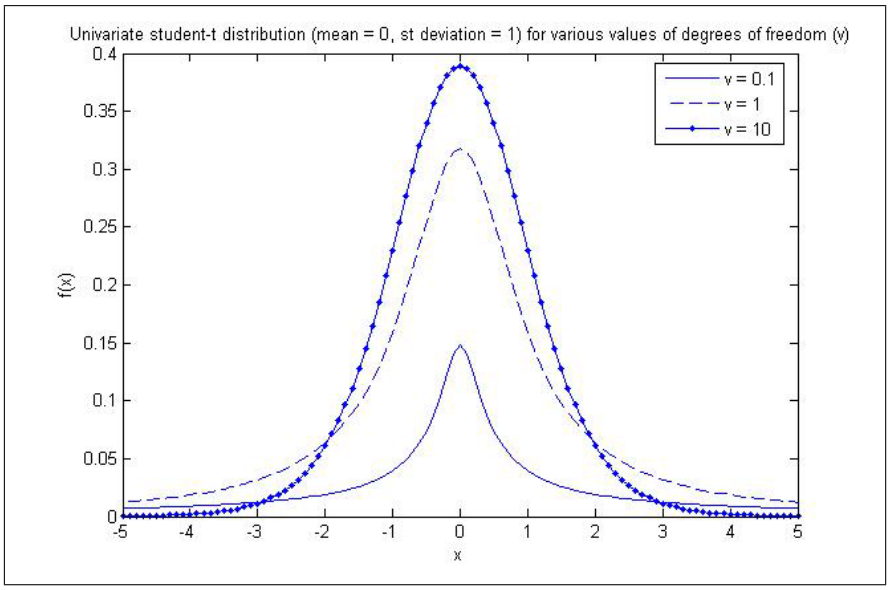


FIGURE A.4 – The Student's t-distribution for various degrees of freedom. As $\nu \to \infty$ the distribution tends to a Gaussian. The Student's t-distribution has heavier tails than a Gaussian.

$\nu(\nu-2)^{-1}\Sigma$ is the covariance matrix of X. Therefore, the family of t-distributions provides a heavy-tailed alternative to the normal family with mean μ and covariance matrix that is equal to a scalar multiple of Σ, if $\nu > 2$ (fig. A.4) [BISHOP 2006].

Therefore, the clique potential functions are properly defined in order to impose :

$$\pi_j^n - \pi_j^k \sim \mathcal{S}t(0, \beta_{jd}^2, \nu_{jd}), \quad \forall n, j, d, \forall k \in \gamma_d, (n). \tag{A.15}$$

As it can be observed in eq. (A.15), we introduce $K \times D$ different t-distributions, amounting to an equal number of parameter sets, $\{\beta_{jd}, \nu_{jd}\}_{j=1..K, d=1..D}$. In eq. (A.15), D stands for the number of a pixel's neighborhood adjacency types, and $\gamma_d(n)$ is the set of neighbors of pixel indexed n, with respect to the d^{th} adjacency type. In our model, we assume 4 neighbors for each pixel, and partition the corresponding adjacency types into horizontal and vertical, thus, setting $D = 2$. This variability of parameter sets aims to capture the fact that smoothness statistics may vary along clusters and spatial directions [NIKOU et al. 2007]. Therefore, the joint distribution on Π is given by :

$$p(\Pi; \beta, \nu) = \prod_{d=1}^{D} \prod_{j=1}^{K} \prod_{n=1}^{N} \prod_{k \in \gamma_d(n)} \mathcal{S}t(\pi_j^n | \pi_j^k; \beta_{jd}^2, \nu_{jd}). \tag{A.16}$$

Following the definition of the t-distribution in eq. (A.12) and eq. (A.13) we introduce the latent variables $U = \{u_j^{nk}\}_{n=1..N, j=1..K, k=1..D}$ and the distribution of the differences of local *contextual mixing proportions* becomes :

$$\pi_j^n - \pi_j^k \sim \mathcal{N}(0, \beta_{jd}^2 / u_j^{nk}),$$
$$u_j^{nk} \sim \mathcal{G}(\nu_{jd}/2, \nu_{jd}/2), \quad \forall n, j, d, \quad \forall, k \in \gamma_d(n). \tag{A.17}$$

This generative model (fig. A.5), apart from being tractable using the EM algorithm, as will be demonstrated in what follows, allows better insight on our assumption of Student-t cliques. Since u_j^{nk} depends on datum indexed by n, each weight difference in the MRF can be described by a different instance of a Gaussian distribution. Therefore, as $u_j^{nk} \to +\infty$ the distribution tightens around zero, and forces neighboring *contextual mixing proportions* to be smooth. On the other hand, when $u_j^{nk} \to 0$ the distribution tends to be uninformative, and forces no smoothness.

This is a desirable property when there exists an edge between neighboring pixels. Thus, the U-variable maps provide a very detailed description of the image edge structure. Furthermore, they may be considered as a continuous generalization of the binary line-process variable idea in [Blake & Zisserman 1987, Geman & Geman 1984].

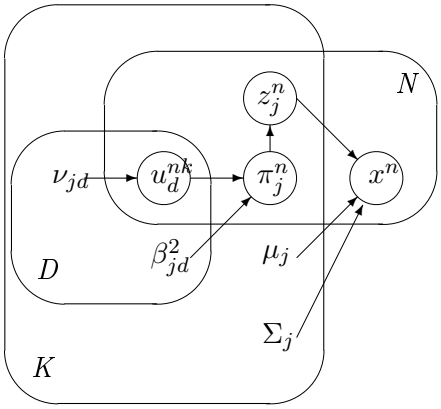


Figure A.5 – Graphical model for the edge preserving spatially variant Gaussian mixture model. Superscript $n \in [1, N]$ denotes pixel index, subscript $j \in [1, K]$ denotes kernel (segment) index and $d \in [1, D]$ describes the neighborhood direction type.

As shown in the graphical model in figure A.5, the unknowns $\Psi = \{\mu, \Sigma, \beta, \nu\}$ are considered as *parameters* and will be estimated in the M-step of the EM algorithm that follows. The Z, U are *hidden* random variables and will be inferred in the E-step of the same algorithm. The unknown quantities Π, although being random variables, are treated as parameters and are estimated in the M-step. This is the reason we refer to this algorithm as MAP-EM [Bishop 2006].

To perform model inference, the evidence with respect to the model parameters $\Psi = \{\mu, \Sigma, \beta, \nu\}$ and the *contextual mixing proportions* Π has to be optimized :

$$\ln p(X, \Pi; \Psi). \tag{A.18}$$

In EM terminology [Dempster et al. 1977] this is the incomplete data log-likelihood while the complete log-likelihood is expressed by :

$$\ln p(X, \Pi, Z, U; \Psi). \tag{A.19}$$

The conditional expectation of the complete likelihood is an important quantity in EM - it is defined as

$$\mathcal{E}_{Z^{(t)}, U^{(t)} | X, \Pi^{(t)}} \Big\{ \ln p(X, \Pi, Z, U; \Psi) \Big\}. \tag{A.20}$$

By optimizing this expectation with respect to Ψ and Π, given the observed variables and some initial estimate $\Psi^{(0)}$,$\Pi^{(0)}$, we can produce a new estimate $\Psi^{(1)}$,$\Pi^{(1)}$. In the same way, the new estimate can be used to produce a next one and so on. It can be proved that these estimates converge to a local optimum for the incomplete likelihood, eq. (A.18) ; this is the main idea in the EM algorithm. Formally this iteration scheme is split in two steps, the Expectation and the Maximization step.

The E-step consists in computing the joint expectation of the hidden variables Z and U, with respect to current iteration parameters $\Psi^{(t)}$ where t denotes the number of current iteration. Observing the graphical model in fig. A.5, we can see that given X and Π, Z and U are

conditionally independent ; therefore $\mathcal{E}_{Z,U|X,\Pi}(\cdot) = \mathcal{E}_{Z|X,\Pi}\{(\mathcal{E}_{U|X,\Pi}(\cdot)\}$ and we can compute these expectations separately. So we have the updates $\forall n, j, d, \forall k \in \gamma_d(n)$:

$$<z_j^n>^{(t)} = \frac{\pi_j^{n(t)} \mathcal{N}(x^n; \mu_j^{(t)}, \Sigma_j^{(t)})}{\sum_{l=1}^K \pi_l^{n(t)} \mathcal{N}(x^n; \mu_l^{(t)}, \Sigma_l^{(t)})}$$

$$<u_j^{nk}>^{(t)} = \zeta_j^{nk(t)} / \eta_j^{nk(t)}$$

$$<\ln u_j^{nk}>^{(t)} = \psi(\zeta_j^{nk(t)}) - \ln \eta_j^{nk(t)}$$

where $\psi(\cdot)$ stands for the digamma function, and parameters ζ, η being :

$$\zeta_j^{nk(t)} = \frac{1}{2}\left(\nu_{jd}^{(t)} + 1\right),$$

$$\eta_j^{nk(t)} = \frac{1}{2}\left(\nu_{jd}^{(t)} + \frac{(\pi_j^{n(t)} - \pi_j^{k(t)})^2}{\beta_{jd}^{2(t)}}\right).$$

Maximization of the current complete likelihood (A.20) must be driven with respect to the model parameters Ψ and Π. With some manipulation, we can rewrite it equivalently as

$$\mathcal{E}_{Z|X,\Pi}\{\ln p(X|Z;\mu,\Lambda)\} + \mathcal{E}_{Z|X,\Pi}\{\ln p(Z|\Pi)\} +$$

$$+\mathcal{E}_{U|\Pi}\{\ln p(\Pi|U;\beta)\} + \mathcal{E}_{U|\Pi}\{\ln p(U;\nu)\}.$$

In this form, parameter optimization is straightforward. The resulting update equations make up the M-step :

$$\mu_j^{(t+1)} = \frac{\sum_{n=1}^N <z_j^n>^{(t)} x^n}{\sum_{n=1}^N <z_j^n>^{(t)}},$$

$$\Sigma_j^{(t+1)} = \frac{\sum_{n=1}^N <z_j^n>^{(t)} (x^n - \mu_j^{(t+1)})(x^n - \mu_j^{(t+1)})^T}{\sum_{n=1}^N <z_j^n>^{(t)}},$$

$$\beta_{jd}^{2(t+1)} = \frac{\sum_{n=1}^N \sum_{k\in\gamma_d(n)} <u_j^{nk}>^{(t)} (\pi_j^{n(t)} - \pi_j^{k(t)})^2}{\sum_{n=1}^N |\gamma_d(n)|}.$$

Moreover, the *contextual mixing proportions* π_j^n are computed as the roots of a quadratic equation :

$$a_j^n \left(\pi_j^{n(t+1)}\right)^2 + b_j^n \left(\pi_j^{n(t+1)}\right) + c_j^{n(t+1)} = 0, \tag{A.21}$$

with coefficients :

$$a_j^n = -\sum_{d=1}^D \left\{ \beta_{jd}^{-2(t)} \sum_{k\in\gamma_d(n)} <u_j^{nk}>^{(t)} \right\},$$

$$b_j^n = \sum_{d=1}^D \left\{ \beta_{jd}^{-2(t)} \sum_{k\in\gamma_d(n)} <u_j^{nk}>^{(t)} \pi_j^{k(t)} \right\},$$

$$c_j^n = \frac{1}{2} <z_j^n>^{(t)} .$$

The form of the coefficients guarantees that there is always a real non negative solution. However, the solutions of eq. (A.21) for a given pixel, indexed by n, will not in general satisfy the constraint $\sum_{j=1}^K \pi_j^n = 1$. In order to get proper mixing weight vectors we perform a projection step onto

the constraints subspace using the quadratic programming algorithm described in [Blekas et al. 2005].

Finally, setting the derivative of (A.20) with respect to the degrees of freedom equal to zero we obtain $\nu_{jd}^{(t+1)}$ as the solutions of the equation :

$$\ln(\nu_{jd}^{(t+1)}/2) - \psi(\nu_{jd}^{(t+1)}/2) + \\ + \left[\frac{\sum_{n=1}^{N} \sum_{k \in \gamma_d(n)} (<\ln u_j^{nk}>^{(t)} - <u_j^{nk}>^{(t)})}{\sum_{n=1}^{N} |\gamma_d(n)|} \right] + 1 = 0,$$

with $\psi(\cdot)$ being again the digamma function.

3.2 Experimental results

In our model, parameters U play a very important role in the preservation of the boundaries between image regions. The U-variable maps for the j^{th} kernel can be considered as the edges that separate the j^{th} segment of the image from the remaining segments. To demonstrate this point we show an example in figure A.6. In this example, an image is segmented into $K = 3$ segments thus 6 U-variable maps are shown. The first row of this figure shows the original and the segmented images. Then, moving from top to bottom, the U-variable maps for the three image segments, namely *sky, roof and shadows, building* are shown, respectively. The left column highlights vertical edges and the right column underpins horizontal edges. Notice that in the second row of fig. A.6, where the U-variable maps for segment *sky* are shown, the edges between the segment *sky* and the rest (*roof and shadows, building*) are mainly highlighted. The edges between the other segments, (*roof and shadows* and *building*) are mainly highlighted in the remaining two maps. Similarly, the edges between the segments *sky* and *building* are not highlighted in the third row of images as the U-variable maps for *roof and shadows* are underpinned.

In our implementation, we have used a 4-dimensional feature vector to describe the image data. It is comprised by the *Lab* color space features and the *Blobworld* contrast texture descriptor as described in [Carson et al. 2002]. Prior to segmentation, each variate has been separately normalized in order not to have dominating features. We have evaluated the proposed Student's t-based SVGMM (St-SVGMM) segmentation scheme on the 300 images of the Berkeley image database [Martin et al. 2001]. We have applied our algorithm for different values of the number of segments ($K = \{3, 5, 7, 10, 15, 20\}$). For comparison purposes, we have also experimented with the standard GMM [Bishop 2006] and the GMM based segmentation with "standard" smoothness constraints [Nikou et al. 2007] with the same number of components.

The obtained segmentations were quantitatively evaluated with two performance measures : the Rand index (RI) [Unnikrishnan et al. 2007] and the boundary displacement error (BDE) [Freixenet et al. 2002]. The RI measures the consistency between the ground truth and the computed segmentation map while the BDE measures error in terms of boundary displacement with respect to the ground truth. The statistics for these measures are presented in tables A.3 and A.4.

Based on the theoretical properties of the Student's t-model one might have expected that the St-SVGMM introduced erroneous boundaries that did not agree with human segmentation. Therefore it would provide a worse RI as compared to the "classical" non preserving algorithm (SVGMM) [Nikou et al. 2007]. However, as observed in the statistics of the RI (table A.3), the St-SVGMM outperforms the standard GMM in all cases and the SVGMM in the overwhelming majority of the different number of components.

Also, in terms of correct region boundary estimation, expressed by the BDE (table A.4), the St-SVGMM outperforms the SVGMM, as theoretically expected. However, it also outperforms

FIGURE A.6 – U-variable maps : the first row shows the original image and the segmentation for $K = 3$ clusters ; the rows below show U-variable maps (expected values of u_j^{nk} variables). Brighter values represent lower values of u. In each row, the U-variable maps for kernel indexed by $j = 1$ *(sky)*, $j = 2$ *(roof and shadows)* and $j = 3$ *(building)*, are shown respectively. The left column corresponds to u values computed for horizontal adjacencies, and the right column for vertical adjacencies.

standard GMM and the difference in performance increases with the number of segments. The explanation for this behavior is that the standard GMM generally computes correctly the boundaries between segments since it does not integrate a smoothing step (it also outperforms the SVGMM in the same median values). However, as the number of segments increases, the complexity of the image cannot be captured by a simple GMM and smoothness constraints that model the image edge structure become increasingly beneficial.

Overall, the St-SVGMM not only preserves region boundaries but also improves the correct classification rates with respect to the standard methods. Some representative segmentation examples are shown in figure A.7.

3.3 Conclusion

A segmentation algorithm based on clustering with GMM has been proposed here. The main novelty of this model is a smoothness prior which apart from constraining adjacent pixels to belong in the same cluster captures the image edge structure. Thus, it does not enforce smoothness across segment boundaries. Another important feature of the herein proposed segmentation algorithm is that all required parameters are estimated from the data. Thus, this algorithm is *automatic* and does not require empirical parameter selection like many recent state-of-the-art segmentation algorithms.

An important issue remains the estimation of the number of components K. In the next

FIGURE A.7 – Segmentation examples using the proposed edge preserving spatially variant mixture. From left to right, the columns show : the original image, segmentation with $K = 5$, $K = 10$ and $K = 15$ kernels.

section, we shall discuss a different Bayesian model that attempts to answer this problem.

TABLE A.3 – *Statistics on the Rand Index (RI) over the 300 images of the Berkeley image data base for the compared methods. Higher values represent better segmentations.*

	GMM			SVGMM			St-SVGMM		
K	Mean	Median	St. dev.	Mean	Median	St. dev.	Mean	Median	St. dev.
3	0.6754	0.6796	0.0853	0.6860	0.6902	0.0851	**0.6871**	**0.6911**	0.0857
5	0.7101	0.7346	0.1025	0.7175	0.7451	0.1071	**0.7183**	**0.7460**	0.1071
7	0.7171	0.7528	0.1191	0.7228	**0.7589**	0.1214	**0.7233**	0.7586	0.1212
10	0.7166	0.7590	0.1329	0.7209	**0.7598**	0.1355	**0.7209**	0.7592	0.1357
15	0.7125	0.7545	0.1431	**0.7159**	0.7579	0.1465	0.7157	**0.7589**	0.1469
20	0.7092	0.7489	0.1473	0.7062	0.7452	0.1529	**0.7110**	**0.7525**	0.1524

TABLE A.4 – *Statistics on boundary displacement error (BDE) over the 300 images of the Berkeley image data base for the compared methods. Lower values represent better segmentations.*

	GMM			SVGMM			St-SVGMM		
K	Mean	Median	St. dev.	Mean	Median	St. dev.	Mean	Median	St. dev.
3	4.7893	**4.1636**	2.3856	4.7870	4.2057	2.3967	**4.7633**	4.2109	2.3820
5	4.3862	**3.7571**	2.1734	4.3938	3.8141	2.1745	**4.3666**	3.7939	2.1713
7	4.2438	3.7077	2.0950	4.2118	3.6831	2.0554	**4.1902**	**3.6451**	2.0556
10	4.1370	3.6025	2.0088	4.0963	3.5038	1.9863	**4.0929**	**3.4969**	2.0033
15	4.0996	3.6353	1.9760	4.0341	3.5036	1.9401	**4.0119**	**3.4486**	1.9345
20	4.1283	3.6784	2.0110	4.1913	3.6549	1.9084	**4.0049**	**3.4735**	1.9286

4. Automatic number of components determination

Selecting the appropriate number of clusters for a given data set is an important issue, on which several approaches have been proposed. The most straightforward model selection approach is fitting a number of models with varying number of components, and evaluating the solutions using a suitable criterion. Such penalty terms, inspired by coding theory and minimum description length, try to avoid data overfitting by penalizing solutions with high number of components. Examples include Akaike's information criterion, the Bayesian information criterion and the Minimum message length criterion [BISHOP 2006].

In the more specific context of assuming the data being generated by a mixture model, methodologies include notably the Bayesian approaches [CORDUNEANU & BISHOP 2001, UEDA & GHAHRAMANI 2002] where the number of kernels and the model parameters are estimated simultaneously. In this family of methods, the model is initialized on a large number of components, and progressively removes those components that reside in the same region of the data space. On the contrary, in [CONSTANTINOPOULOS & LIKAS 2007], the model starts with a low number of components and more kernels are progressively added by splitting existing kernels when necessary.

It is straightforward to adapt these Bayesian models into image segmentation, simply by assuming our image feature vectors to be the data to be clustered. However, the important feature of an image's spatial structure would not be accounted for. Natural images have a spatial smoothness property which is neglected by standard mixture model approaches. Approaches like the method proposed in [BLEKAS et al. 2005, NIKOU et al. 2007, MARROQUIN et al. 2003] and the model discussed in the previous section of this thesis Appendix (see also [SFIKAS et al. 2008A]), are based on MRF priors [GEMAN & GEMAN 1984] to account for spatial characteristics.

However, they assume an *a priori* known number of segments.

Here we present a Bayesian model for image segmentation that enables the estimation of the number of segments during the training process while accounting for image spatial smoothness. We assume that the distribution of the hidden class labels is controlled by two distinct sets of probability weight vectors, tagged correspondingly as *local* and *global* weights.

The *local* weights vary with pixel location. Local differences in these weights follow a Student's-t distribution. The Student's-t distribution decomposes on two levels : the lower level is a Gaussian pdf with precision (inverse variance) that is spatially variant, while the higher level is a Gamma pdf that generates the aforementioned precision values. This precision variability of the Gaussians allows the model to incorporate elegantly the image edge structure along with imposing smoothness constraints.

The *global* weights control the number of image segments that are active in the model by imposing a Dirichlet prior on them. In this way, more probable solutions, which otherwise exhibit high model complexity by comprising many kernels, are penalized as low probability states. This allows the model to estimate the number of classes in the segmentation process, by starting from an initial high number of classes estimate and pruning mixing kernels gradually during the model training process.

The variational inference framework [Bishop 2006, Sudderth & Jordan 2008] is used to train the model. Variational inference involves iteratively optimizing a lower bound of the model evidence with regard to the posterior distribution of the hidden variables, and the model parameters. The mean field approximation is employed on the posterior distribution of the hidden variables, so as to render its estimation tractable. Let us note that the proposed model is different with respect to standard Dirichlet priors imposed on the mixing proportions of a mixture [Bishop 2006]. In our model, the hidden variables depend on two priors (*local* and *global* weights) and model inference is not trivial with standard inference techniques. Therefore, we optimize the variational lower bound by making use of the Majorization-Minimization (MM) methodology [Lange 2004].

Thus, unlike state-of-the-art methods in image segmentation like normalized cuts [Shi & Malik 2000] and standard or spatially varying mixtures [Sfikas et al. 2010B], the proposed model can produce an estimate of the number of image segments while at the same time ensuring a smooth segmentation result. Methods supporting automatic determination of the number of classes typically depend on a scalar parameter, or a small set of parameters, that more or less directly control the number of classes / fit likelihood trade-off (e.g. the bandwidth in mean shift [Comaniciu & Meer 2002]). Such parameters are meant to be empirically adjusted beforehand. Concerning the rest of the parameters, affecting the quality of the segmentation itself, the proposed model determines them automatically. We consider this issue as an advantage in comparison with other methods. Graph-cut based methods fall under this latter category [Sfikas et al. 2010B], as well as recently proposed extensions that can handle number of components determination [Feng et al. 2010].

Producing a smooth segmentation result while determining the number of classes has also been addressed in the Dirichlet process prior models proposed in [DaSilva 2007, Orbanz & Buhmann 2007], among others. However such approaches rely on sampling techniques which are notoriously computationally expensive, in contrast to the Majorization-Minimization iterative scheme we propose in this section.

4.1 Model description and solution

Let $X = \{x^n\}_{n=1}^N$ be the observed set of the image intensities. Consider also that there exist *at most* K classes in our segmentation. Each datum x^n is governed by different statistics, according to which class it belongs to. Let us assume a hidden variable set $Z = \{z_j^n\}_{n=1..N,\, j=1..K}$,

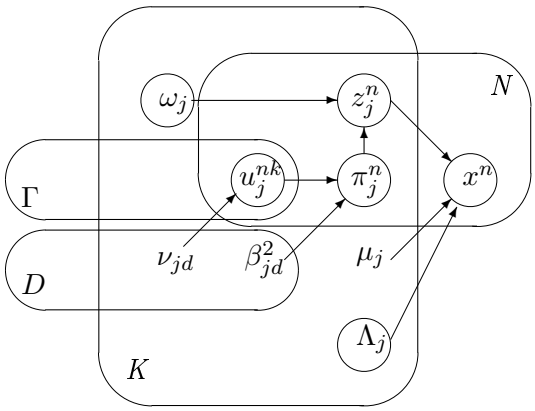


FIGURE A.8 – Graphical representation of the proposed model. Superscript $n \in [1, N]$ denotes pixel index, subscript $j \in [1, K]$ denotes kernel (segment) index, $d \in [1, D]$ describes the neighborhood direction type and $k \in [1, \Gamma]$ denotes neighbor index.

grouped as N one-zero $K \times 1$ vectors that control pixel class membership.

It is a popular choice in computer vision to choose the data to be Gaussian and *i.i.d* distributed, assuming knowledge of class memberships Z :

$$x^n | z_j^n = 1 \ \sim \ \mathcal{N}(\mu_j, \Lambda_j), \tag{A.22}$$

with $\mathcal{N}$ representing a Gaussian distribution with μ_j and Λ_j being the mean vector and the precision (inverse covariance) matrix respectively.

The distribution choice of Z plays a drastic role on the model behavior. Let us recall that under a multinomial and *i.i.d* assumption on the z^n, the model is essentially a Gaussian mixture [BISHOP 2006] governed by a set of K weights. In [SFIKAS et al. 2010B] this idea is extended by using spatially varying weights with a smoothness prior on them.

With the presented model we make a design choice that is meant to lie between the spatially and non-spatially varying hypotheses for the mixing proportions. Let z^n, $\forall n \in [1, N]$ be independently distributed with

$$p(z^n | \Omega, \Pi) = \frac{\prod_{j=1}^{K} (\pi_j^n \omega_j)^{z_j^n}}{\sum_{j=1}^{K} \pi_j^n \omega_j}. \tag{A.23}$$

In the equation above, we have introduced $\Pi = \{\pi_j^n\}_{n=1..N, j=1..K}$ and $\Omega = \{\omega_j\}_{j=1..K}$ weight variable sets, which are constrained by $\sum_{j=1}^{K} \pi_j^n = 1$, $\forall n \in [1, N]$ and $\sum_{j=1}^{K} \omega_j = 1$. In view that the denominator in (A.23) is the probability distribution normalization constant, eq. (A.23) acts closely like a multinomial distribution with weights given by the set $[\pi_1^n \omega_1, \pi_2^n \omega_2, ..., \pi_K^n \omega_K]$. Thus, for each pixel $n \in [1, N]$, a class membership is determined by two sets of weight vectors. At first, it depends on the set $\Pi = [\pi_1^n, \pi_2^n, ..., \pi_K^n]$ whose components are vectors, now called *local weights*. The local weights are spatially varying as they depend on the position (indexed by n). Secondly, it depends on the set $\Omega = [\omega_1, \omega_2, ..., \omega_K]$ whose components are scalars and are now called *global weights*.

For further insight, note that, if in eq. (A.23), for a given pixel indexed by n, we treat its local weights π_j^n as parameters with probabilities equal to the uninformative $1/K$, the distribution of the hidden variables z^n is a multinomial distribution. This is also true for the global weights ω_j. Thus, in eq.(A.23), for a given a pixel, both the local and global weights should have high values for a certain class in order to be dominant.

Considering the set of *local weights* Π as random variables and assuming a proper prior, we can incorporate the spatial smoothness trait by forcing neighboring vectors to be more likely to

share the same class label. We assume a Markov random field on Π, which equivalently means that Π is governed by a Gibbs distribution [GEMAN & GEMAN 1984], generally expressed by :

$$p(\Pi) \propto \prod_{C} e^{-\Psi_c(\Pi)}, \tag{A.24}$$

where Ψ_c is a function on clique c, called *clique potential* function in the literature, and the product is over all minimal cliques of the Markov random field.

An appropriate clique distribution choice would be to assume that the differences of *local weights* Π follow a Student's-t distribution with its peak set at zero. This setting, proposed previously in [SFIKAS et al. 2010B], also provides our model with the properties of an edge-preserving line-process [GEMAN & GEMAN 1984]. The probability law for local differences is thus expressed by

$$\pi_j^n - \pi_j^k \sim \mathcal{S}t(0, \beta_{jd}^2, \nu_{jd}), \quad \forall n, j, d, \forall k \in \gamma_d(n). \tag{A.25}$$

The parameters β_{jd} control how tightly smoothed we need the vectors of segment j to be. In eq. (A.25), D stands for the number of a pixel's neighborhood adjacency types and $\gamma_d(n)$ is the set of neighbors of pixel indexed by n, with respect to the d^{th} adjacency type, where $d \in [1, D]$. In our model, we assume 4 neighbors for each pixel (*first-order* neighborhood), and partition the corresponding adjacency types into horizontal and vertical, thus, setting $D = 2$. This variability of parameter aims to capture the intuitive property that smoothness statistics may vary along clusters and spatial directions.

It can be observed that the assumption in (A.25) is equivalent to

$$\pi_j^n - \pi_j^k \sim \mathcal{N}(0, \beta_{jd}^2 / u_j^{nk}),$$

$$u_j^{nk} \sim \mathcal{G}(\nu_{jd}/2, \nu_{jd}/2), \quad \forall n, j, d, \quad \forall k \in \gamma_d(n),$$

where $\mathcal{N}$ and $\mathcal{G}$ represent a Gaussian and a Gamma distribution respectively. This breaking-down of the Student's-t distribution allows clearer insight on how our implicit edge-preserving line-process works. Since u_j^{nk} depends on datum indexed by n, each weight difference in the MRF can be described by a different instance of a Gaussian distribution. Therefore, as $u_j^{nk} \to +\infty$ the distribution tightens around zero, and forces neighboring *local weights* to be smooth. On the other hand, $u_j^{nk} \to 0$ signifies the existence of an edge and consequently no smoothing.

The *global weights* Ω are introduced in the model in order to cover our model's second important property which is the automatic estimation of the number image segments. The idea is that starting from a predefined maximum segments figure K, during the training process some segments "fade out" eventually to zero-weights [CORDUNEANU & BISHOP 2001]. While it is possible that for a certain class j all local weights π_j^n, $\forall n \in [1, N]$ may attain negligible values, in practice this is difficult. This is due to the fact that updating each π_j^n $\forall n \in [1, N]$ individually must account for MRF local dependencies, which will not lead to an update far from each site's neighbors.

Thus, assigning for each class a single *global weight* scalar allows us to conveniently treat each class as a distinct entity during model training. Therefore, Ω is considered to be a random vector governed by a Dirichlet distribution :

$$p(\Omega; \alpha_0) \propto \prod_{j=1}^{K} \omega_j^{\alpha_0 - 1}. \tag{A.26}$$

By these means, we can penalize solutions with numerous non-zero components. As hyperparameter $\alpha_0 \to -\infty$, solutions with less segments are encouraged, and as $\alpha_0 \to +\infty$ all K initial segments tend to be preserved. While $\alpha_0 < 0$ may enforce the prior in (A.26) to be improper, in

practice negative α_0 values are applicable since it is not necessary to compute the normalizing constant of eq. (A.26) during inference as it will be explained in the next section.

Finally we impose a Wishart prior on precision matrices Λ_j, $\forall j \in [1, K]$:

$$p(\Lambda_j; W_0, \eta_0) \propto |\Lambda_j|^{(\eta_0 - \Delta - 1)/2} e^{-\frac{1}{2} Tr(W_0^{-1} \Lambda_j)}, \tag{A.27}$$

where the matrix W_0 and the scalar η_0 are such that $\mathcal{E}\{\Lambda_j\} = W_0 \eta_0$. Parameter Δ stands for the number of variates of the feature vectors x^n. Imposing this prior on precision matrices avoids degenerate cases, for instance, when the corresponding covariance matrix Λ_j^{-1} has zero eigenvalues or equivalently $|\Lambda_j| \to +\infty$ [Bishop 2006]. For an overview of the proposed model, see fig.(A.8).

To perform inference and consequently segmentation, the model likelihood with respect to model parameters has to be optimized :

$$\underset{\mu, \Pi, \Omega, \beta}{\operatorname{argmax}} \ln p(X, \Pi, \Omega; \mu, \beta, \nu).$$

Due to the functional form of the involved distributions the above optimization problem is practically intractable. Therefore, we resort to variational inference [Bishop 2006]. This involves calculating approximations of the posterior distributions $q(\cdot)$ of the hidden variables Z, U, Λ, then using them to find parameter estimates that maximize a lower bound of the model likelihood.

Adapting the standard variational methodology [Bishop 2006] to our problem, the lower bound to be optimized is

$$\begin{aligned}
&\mathcal{L}(q, \Pi, \Omega, \mu, \beta, \nu) \triangleq \\
&\sum_Z \int_{U,\Lambda} q(Z, U, \Lambda) \ln \frac{p(X, \Pi, Z, U, \Omega; \mu, \beta, \nu)}{q(Z, U, \Lambda)} dU d\Lambda \\
&= <\ln p(X, \Pi, Z, U, \Omega; \mu, \beta, \nu)>_{Z,U,\Lambda} - \\
&\quad <\ln q(Z, U, \Lambda)>_{Z,U,\Lambda} \\
&= <\ln p(X|Z, \Lambda; \mu)>_{Z,\Lambda} + <\ln p(\Lambda)>_{\Lambda} + \\
&<\ln p(Z|\Pi, \Omega)>_Z + <\ln p(\Pi|U; \beta)>_U + \ln p(\Omega) + \\
&\quad <\ln p(U; \nu)>_U - <\ln q(Z, U, \Lambda)>_{Z,U,\Lambda} .
\end{aligned} \tag{A.28}$$

Model evidence is decomposed to the lower bound $\mathcal{L}$ and the Kullback-Leibler distance between the approximation of the posterior and the posterior itself :

$$\ln p(X, \Pi, \Omega; \mu, \beta, v) = \mathcal{L}(q, \Pi, \Omega, \mu, \beta, \nu) + KL(q||p).$$

To proceed with the computation of the optimal distribution q on $\mathcal{L}$, we recur to the *mean field approximation* which stems from statistical physics [Bishop 2006] :

$$q(Z, U, \Lambda) = q(Z) q(U) q(\Lambda). \tag{A.29}$$

Note that in the proposed model, we only need to assume $q(Z, U, \Lambda) = q(Z, U) q(\Lambda)$, as $q(Z, U) = q(Z) q(U)$ is induced from the model structure (fig. A.8). Thence we can obtain update equations for the expected values of hidden variables Z,U,Λ :

$$<z_j^n>^{(t)} = \frac{\pi_j^{n(t)} \omega_j^{(t)} \mathcal{N}(x^n; \mu_j^{(t)} | <\Lambda_j>^{(t)})}{\sum_{l=1}^K \pi_l^{n(t)} \omega_j^{(t)} \mathcal{N}(x^n; \mu_l^{(t)} | <\Lambda_l>^{(t)})}, \tag{A.30}$$

$$<u_j^{nk}>^{(t)}= \zeta_j^{nk(t)}/\theta_j^{nk(t)},$$
$$<\ln u_j^{nk}>^{(t)}= \psi(\zeta_j^{nk(t)}) - \ln\theta_j^{nk(t)},$$

where $\psi(\cdot)$ stands for the digamma function, and parameters ζ, θ being :

$$\zeta_j^{nk(t)} = \frac{1}{2}\left(\nu_{jd}^{(t)} + 1\right),$$

$$\theta_j^{nk(t)} = \frac{1}{2}\left(\nu_{jd}^{(t)} + \frac{(\pi_j^{n(t)} - \pi_j^{k(t)})^2}{\beta_{jd}^{2(t)}}\right).$$

The required moment for variables Λ_j are given by

$$<\Lambda_j>^{(t)}= W_j^{(t)}\eta_j^{(t)} \qquad \eta_j^{(t)} = \eta_0 + \sum_{n=1}^{N} <z_j^n>^{(t)},$$

$$W_j^{-1(t)} = W_0^{-1} + \sum_{n=1}^{N} <z_j^n>^{(t)} (x^n - \mu_j^{(t)})(x^n - \mu_j^{(t)})^T.$$

Estimation of the deterministic parameters μ, Π, Ω, β, ν is achieved by maximization of (A.28) with respect to them. However, optimizing (A.28) with respect to Ω is difficult due to the normalizing factor in (A.23). We can work around this obstacle and find a closed form update for Ω as well, by making use of the Majorization-Minimization (MM) methodology [Lange 2004]. MM in its philosophy is quite close to variational inference and the EM algorithm [Bishop 2006], in the sense that the problem of minimizing a given objective function is transformed to successive minimizations of surrogate functions, i.e. majorizers of the original objective function that can be minimized in closed-form.

For the term of (A.28) involving $\ln p(Z|\Pi,\Omega)$, we note the following inequality :

$$<\ln p(Z|\Pi,\Omega)>_Z=$$
$$-\sum_{n=1}^{N}\ln\sum_{j=1}^{K}\pi_j^n\omega_j + \sum_{j=1}^{K}\ln\omega_j\sum_{n=1}^{N}\pi_j^n <z_j^n> \geqslant$$
$$-\sum_{n=1}^{N}\ln y^n - \sum_{n=1}^{N}\frac{1}{y^n}(\sum_{j=1}^{K}\pi_j^n\omega_j - y^n) + \sum_{j=1}^{K}\ln\omega_j\sum_{n=1}^{N}\pi_j^n <z_j^n>$$
$$\triangleq \varphi(Z,\Pi,\Omega,y), \tag{A.31}$$

where we have introduced $y = \{y^1, y^2, ..., y^N\}$ as a new set of auxiliary real parameters. In eq. (A.31), we made use of the linear *minorization*

$$f(x) \geqslant f(y) + \frac{df(y)}{dy}(x-y),$$

which holds for any convex function f. Here $f(x) = -\ln x$ and $x = \sum_{j=1}^{K}\pi_j^n\omega_j$.

Consequently, we define our *minorant* as

$$\mathcal{L}^{MM}(q,\mu,\beta,\nu,\Pi,\Omega,y) \triangleq \tag{A.32}$$
$$\mathcal{L}(q,\mu,\beta,\nu,\Pi,\Omega) - <\ln p(Z|\Pi,\Omega)>_Z + \varphi(Z,\Pi,\Omega,y),$$

which according to (A.31) is easily confirmed to be a lower bound of (A.28). Therefore, we have an MM approach in a *Minorization-Maximization* sense.

Optimization of (A.32) leads to parameter value updates for μ, y, β, ν, Π, Ω. For the first three parameter sets, the updates are

$$\mu_j^{(t+1)} = \frac{\sum_{n=1}^{N} <z_j^n>^{(t)} x^n}{\sum_{n=1}^{N} <z_j^n>^{(t)}}, \quad y^{n(t+1)} = \sum_{j=1}^{K} \omega_j \pi_j^n,$$

$$\beta_{jd}^{2(t+1)} = \frac{\sum_{n=1}^{N} \sum_{k \in \gamma_d(n)} <u_j^{nk}>^{(t)} (\pi_j^{n(t)} - \pi_j^{k(t)})^2}{\sum_{n=1}^{N} |\gamma_d(n)|}.$$

Setting the derivative of the lower bound (A.32) with respect to the degrees of freedom of the Student's -t distributions equal to zero we obtain $\nu_{jd}^{(t+1)}$ as the solutions of the equation :

$$\ln(\nu_{jd}^{(t+1)}/2) - \psi(\nu_{jd}^{(t+1)}/2) +$$

$$+ \left[\frac{\sum_{n=1}^{N} \sum_{k \in \gamma_d(n)} (<\ln u_j^{nk}>^{(t)} - <u_j^{nk}>^{(t)})}{\sum_{n=1}^{N} |\gamma_d(n)|} \right] + 1 = 0,$$

The solution for parameter ν is obtained using the bisection method [Peel & McLachlan 2000].

In order to estimate Π and Ω we have the difficulty that the optimization is under positivity and sum-to-unity constraints as they have to be probability vectors defined by (A.23). The *local weights* π_j^n are computed as the roots of a quadratic equation :

$$a_j^n \left(\pi_j^{n(t+1)} \right)^2 + b_j^n \left(\pi_j^{n(t+1)} \right) + c_j^{n(t+1)} = 0, \tag{A.33}$$

with coefficients :

$$a_j^n = -\sum_{d=1}^{D} \left\{ \beta_{jd}^{-2(t)} \sum_{k \in \gamma_d(n)} <u_j^{nk}>^{(t)} \right\},$$

$$b_j^n = \sum_{d=1}^{D} \left\{ \beta_{jd}^{-2(t)} \sum_{k \in \gamma_d(n)} <u_j^{nk}>^{(t)} \pi_j^{k(t)} \right\} - \frac{\omega_j}{2y^n},$$

$$c_j^n = \frac{1}{2} <z_j^n>^{(t)} .$$

The solutions of (A.33) for a given pixel, indexed by n, will not in general satisfy the constraints $\pi_j^n \geqslant 0$, $\sum_{j=1}^{K} \pi_j^n = 1$. In order to get proper mixing weight vectors we perform a projection step onto the constraints subspace using the quadratic programming algorithm described in [Sfikas et al. 2010B].

Motivated by the form of the objective function to be optimized, we follow a different strategy for the estimation of the *global weights*. At first, the unconstrained optimizers are computed :

$$\tilde{\omega}_j = \frac{\sum_{n=1}^{N} <z_j^n> + \alpha_0 - 1}{\sum_{n=1}^{N} \pi_j^n / y^n}. \tag{A.34}$$

If $\tilde{\omega}_j < 0$, we fix the corresponding constrained solution to $\omega_j = 0$, so that they comply with the positivity constraint $\omega_j \geqslant 0$.

We carry on to the second step with the remaining $J \leqslant K$ non-zero components after relabelling them as $\{\omega_1, ..., \omega_J\}$ and the $K - J$ zero components as $\{\omega_{J+1}, ..., \omega_K\}$. Solving the corresponding equation subject to the constraint $\sum_{j=1}^{J} \omega_j = 1$, we obtain

$$\omega_j^{(t+1)} = \frac{\sum_{n=1}^{N} <z_j^n>^{(t)} + \alpha_0 - 1}{\lambda + \sum_{n=1}^{N} \pi_j^{n(t)} / y^{n(t)}}, \tag{A.35}$$

where λ is the Lagrange multiplier. Substituting ω_j, $j = 1, ..., J$ from (A.35) to the sum-to-unity constraint yields

$$\sum_{j=1}^{J} \frac{\sum_{n=1}^{N} <z_j^n>^{(t)} + \alpha_0 - 1}{\lambda + \sum_{n=1}^{N} \pi_j^{n(t)} / y^{n(t)}} - 1 = 0. \tag{A.36}$$

Note that the left-hand side in (A.36) is continuous and monotonically decreasing function of :

$$\lambda \in \left[\max_j \left\{ -\sum_{n=1}^{N} \pi_j^{n(t)} / y^{n(t)} \right\}, \ +\infty \right)$$

Also, as $\lambda \to \max_j \left\{ -\sum_{n=1}^{N} \pi_j^{n(t)} / y^{n(t)} \right\}$, the left hand of (A.36) goes to $+\infty$ and as $\lambda \to +\infty$ the left hand side of (A.36) goes to -1. Thus, we can determine the solution for λ using the bisection method. Substituting it into (A.35) yields the updates for the constrained global weights.

Summing up, the updates presented previously for the posterior distribution and the updates for the model deterministic parameters constitute an iterative model training scheme. During the training process, some of the K ω_j global weight coefficients may gradually go down to zero. In view of update (A.30), no pixels will any longer be assigned to the corresponding class, and effectively these classes are pruned from the model.

The iterations terminate with lower bound convergence. Since bound convergence is guaranteed for both *MM* [Lange 2004] and variational inference [Bishop 2006], the proposed MM-derived lower bound $\mathcal{L}^{MM}$ will also converge in a finite number of iterations.

4.2 Experimental results

At first, we have applied the proposed model to the segmentation of a piecewise constant image slightly corrupted by white Gaussian noise at SNR of 20 *dB* (*Mondrian* [Orbanz & Buhmann 2007], fig. A.9). As this is a relatively easy segmentation problem, we use this example to show our algorithm results over varying Dirichlet hyperparameter values α_0. For convenience, we express α_0 in (A.26) as a function of the image size N and the maximum number of classes K, with $\alpha_0 = -\epsilon^{-1} N K^{-1}$. In view of the global weights updates (A.35), the value of α_0 is compared to the sum of the expected values z_j^n which is of the order of NK^{-1}. Thus, values of ϵ close to one will encourage pruning of kernels. On the other hand, as ϵ approaches $N^{-1}K$ kernel pruning is progressively less encouraged. For the *Mondrian* image (fig. A.9) we have set $\epsilon = 1, 5, 50, 1000$, starting from an initial number of kernels $K = 12$. Low values for parameter ϵ lead to underestimation of the true number of segments. High values of ϵ yield over-fitting problems.

Figure A.9 – Segmentation results for the *Mondrian* image, over various Dirichlet hyperparameter values. From left to right : the original image after degradation by additive white Gaussian noise (SNR of 20 *dB*), segmentations using $\epsilon = 1$, $\epsilon = 5$, $\epsilon = 20$, $\epsilon = 1000$. For too low ϵ, the number of segments is undervalued. For too high ϵ, extra segments are formed erroneously out of image noise.

Furthermore, to test the dependency of the estimated number of kernels on the initial number of segments K and the hyperparameter ϵ (and consequently the Dirichlet parameter α_0), we have run tests with varying parameter values on the *Mondrian* (fig. A.9) and *Church* (fig. A.10). The results presented in figure A.11 show that for low values of ϵ, that is, penalizing configurations with high number of kernels, the final number of segments are almost invariant with regard to its initial value K, as it would be desired.

FIGURE A.10 – Natural image segmentation using Lab features for $\epsilon = 5$, $K = 7$ initial number of segments. The algorithm converged to four segments.

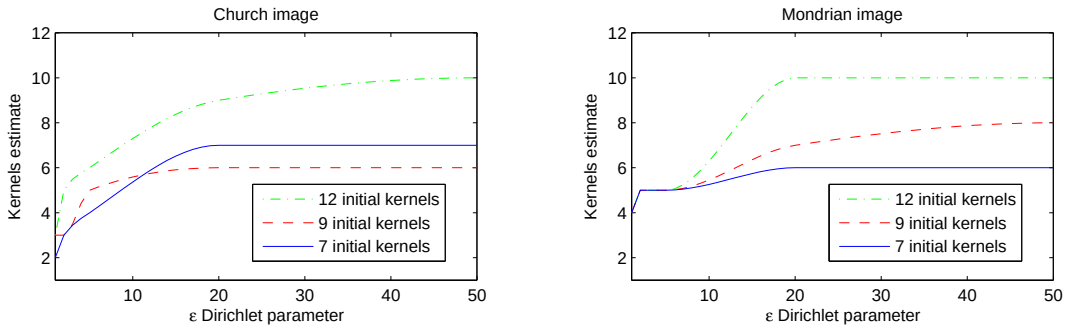


FIGURE A.11 – Estimate of the number of kernels for varying values of initial segments K and Dirichlet hyperparameter ϵ. Curves are interpolants for values at points $\epsilon = 1, 2, 5, 20, 50$.

In order to evaluate the combination of both the smoothing MRF, and the automatic determination of the number of components, we have compared the proposed algorithm to two other models. The first model comprises an MRF prior without any automatic component number selection. In that case, the *global weights* Ω are inactive. The second model consists of estimating the number of segments but incorporates no smoothing prior. Therefore, in this case, the *local weights* Π are inactive. Models close to these two may be found in [SFIKAS et al. 2010B] and [CORDUNEANU & BISHOP 2001] respectively.

The experiment was run on a test 3-class piecewise constant image degraded by white additive Gaussian (SNR of 18 dB) as presented in figure A.12. The initial number of segments was set to $K = 7$ and the Dirichlet parameter $\epsilon = 5$. The model with no kernel number selection fails as it identifies erroneously the noise as separate classes. Both models with kernel number selection successfully prune the extra kernels to the correct number of three. However, the proposed model succeeds also to deal with the noise due to its smoothing property.

Finally, we have tested our algorithm on the Berkeley natural image database [MARTIN et al. 2001]. We have used a superpixels initialization [MORI 2005] as described in [YANG et al. 2007]. We start by oversegmenting the images, all at full resolution of 480×320, to typically around 200 superpixels each. Then, we associate to each superpixel the medoid of the color feature vectors of the pixels that belong to the superpixel in question. These superpixels medoids play the role of the X observed set for our algorithm and represent the whole set of pixels belonging

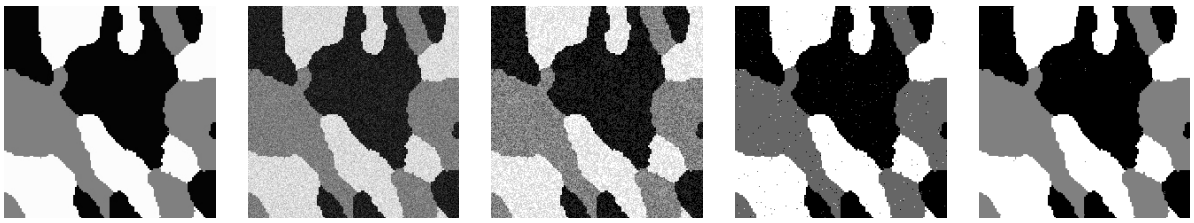


FIGURE A.12 – Segmentation results for a 3-class piecewise constant image. From left to right : original image, image degraded by additive Gaussian noise (SNR of 18 dB), segmentation using the proposed model without global weights (Rand index = 82.14%), segmentation using the proposed model with $\epsilon = 5$ and without local weights (Rand index = 99.2%), segmentation using the proposed model with $\epsilon = 5$ (Rand index = 99.86%).

to the corresponding superpixel. A region adjecency graph is also computed to keep track of the superpixel neighborhoods for the MRF imposed on the *local* weights. To this end, we consider for the MRF prior, those medoids that represent spatially adjacent superpixels.

We have quantitatively evaluated the segmentations using four performance measures [YANG et al. 2007] : the Rand index (RI), the variation of information (VOI), the global consistency error (GCE) and the Boundary displacement error (BDE). The RI measures the consistency between human segmentations and the computed segmentation map. VOI measures the amount of information one segmentation conveys about the other. GCE measures the degree of refinement between two segmentations. Finally, BDE measures the average chamfer distance between the boundaries of two segmentation maps. The mean values of the results over the 300 images of the data base are summarized in table A.5. As it can be observed, all of the indices have values comparable to the ones obtained by state of the art techniques (for example [YANG et al. 2007]). It should be noted that the BDE depends highly on the image size. In figure

TABLE A.5 – *Segmentation evaluation of the algorithm on the 300 images of the Berkeley image data base. The mean values of the corresponding indices are presented (see text for abbreviations). The model was applied to the original images of size* 480×320 *pixels and it was initialized to* $K = 15$ *starting number of classes. Results are computed over two different segmentation scales,* $\epsilon = 3$ *(lower number of classes) and* $\epsilon = 10$ *(higher number of classes).*

Index	*RI*	*BDE*	*GCE*	*VOI*
$\epsilon = 3$	0.71	15.6	0.27	2.27
$\epsilon = 10$	0.72	14.4	0.31	2.52

A.13, we present some representative results using an initial number of kernels $K = 15$ and $\epsilon = \{3, 10\}$. As a final remark concerning the parameter ϵ (or equivalently α_0), let us stress that (a) its choice amounts to how coarse or fine we want the segmentation to be, (b) user-defined parameters with a similar role, affecting the resulting number of classes are used in other number-of-class-determining segmentation algorithms to our knowledge (e.g. the Dirichlet hyperparameter in [ORBANZ & BUHMANN 2007] or the bandwidth in [COMANICIU & MEER 2002]). Nonetheless, in the experiments on natural images described here values of $\epsilon \simeq 5$ seem to have given the best results.

Algorithm runtime is in the order of a few minutes. On a 2 Ghz workstation each algorithm iteration took around 20 seconds for a 480×320 color image (MATLAB code), converging at $20 - 30$ iterations.

FIGURE A.13 – Segmentation results on images from the Berkeley database, using a superpixels initialization. All images were initialized at $K = 15$ classes. From left to right, each column shows the original image, the segmentation for $\epsilon = 3$ and the segmentation for $\epsilon = 10$.

4.3 Conclusion

We proposed a segmentation algorithm based on a Bayesian model. The main novelty of this work is the use of a smoothness MRF prior along with automatic selection of the number of segmentation classes. Updates for the model training are obtained in an efficient manner by variational inference and the Majorization-Minimization (MM) methodology. Recently proposed algorithms that combine a smoothness prior with automatic selection of the number of kernels have to resort to computationally expensive Monte Carlo sampling instead. As future work, novel MRF energy minimization techniques such as proposal-based fusion [LEMPITSKY et al. 2010] could be integrated to our model, as MRF optimization is a critical step. The evaluation of the model to image data bases using more sophisticated features for natural images such as the MRF texture features [REN & MALIK 2003] and the Blobworld features [CARSON et al. 2002] is also envisaged.

5. Recovery and segmentation of polarimetric images

Exploiting the polarization of light has been shown to be a useful and powerful technique, overcoming many limitations that arise in radiance measurement-based classical imagery. Indeed, there is increasing evidence that recording the polarization properties of inhomogeneous objects provides a rich set of information about their local nature. This imaging modality requires the development of efficient imaging systems that can record spatially distributed polarization patterns across a scene and appropriate techniques of handling and processing the issued multicomponent images while preserving the physical integrity of the data.

Stokes imaging consists in estimating the four Stokes parameters of each pixel of an image. This is traditionally achieved by placing a Polarization State Analyzer (PSA) in front of a camera. This configuration allows acquiring polarized radiance images G that can be used to

calculate the multi-component Stokes image S. Images G and S are linked by the Polarization Measurement Matrix (matrix H) that depends on the PSA configuration ($g = Hs$, pixelwise). Classically, the Stokes parameters are obtained using a pseudo-inverse approach, which is sensitive to noises that degrade the acquired intensity images. Moreover, at each location, the Stokes parameters must satisfy physical constraints that can be infringed when using pseudo-inversion.

Here we present a Bayesian model based on a mixture of Gaussians, capable of performing recovery simultaneously with segmentation of the polarimetric image. Intuitively we conjecture that the problem of recovering the degraded image should be intertwined with the problem of segmenting the image. In other words, a good restoration should lead to a good segmentation and vice versa, which is the motivation of performing the two operations jointly.

Polarimetric imaging gives rise to intricate estimation problems because of the associated underlying physical admissibility conditions [Zallat & Heinrich 2007, Zallat et al. 2008]. Stokes images exhibit the particularity that, while they are comprised of 4 separate channels, only a subset of $\mathbb{R}^4$ constitutes admissible Stokes 4-variate vectors. In order to work around this problem, instead of assuming directly that the input Stokes vectors follow a certain distribution, we impose a mixture of Gaussian distributions on a suitable transformation of the Stokes image. Furthermore, with a properly chosen prior set on the probabilities of the underlying segmentation class labels, we achieve to produce smooth edge-preserving segmentations that in turn produce correspondingly reasonably smooth image restorations.

Polarimetric image restoration methodologies have already been presented in [Zallat & Heinrich 2007, Zallat et al. 2008], but under the hypothesis that the variables of interest were spatially piecewise constant. Our present model uses no such constraining assumption. We have tested the proposed model succesfully on sets of noise-degraded artificial Stokes images and on real data, and present both numerical and visual results.

5.1 Model description and solution

Each image to be restored consists of 4 channels. For pixel n, we group the scalars of each observed channel into vector $g^n = [g_1^n \; g_2^n \; g_3^n \; g_4^n]^T$ and let $G = \{g^n\}_{n=1}^N$ denote the set of vectors g^n. These vectors are considered as stemming from N corresponding Stokes vectors $S = \{s^n\}_{n=1}^N$, $s^n = [s_1^n \; s_2^n \; s_3^n \; s_4^n]^T$. Thus G is an indirect and noise corrupted version of S. The 4×4 observation matrix H is supposed to be known. Every Stokes channel corresponds to a specific function of the covariance between orthogonal electric vector components [Zallat et al. 2004]. Consequently the Stokes vectors are subject to the following constraints :

$$s_1^n \geqslant 0, \quad (s_1^n)^2 \geqslant (s_2^n)^2 + (s_3^n)^2 + (s_4^n)^2. \tag{A.37}$$

We model the noise on each Stokes channel as zero mean, additive, white Gaussian. Formally this translates to

$$g^n | s^n, V \sim \mathcal{N}(H \; s^n, V), \tag{A.38}$$

where the covariance matrix is $V = \sigma^2 I$. [1]

We take advantage of our probabilistic generative model formulation and assume a prior distribution on the real Stokes vector set S. Such prior knowledge is the intuitive fact that vectors with spatially neighboring coordinates are more likely to have values close to one another. In order to implement this hypothesis, we first assume the existence of an underlying segmentation of the polarimetric image in K segments. The segmentation is defined by the set of $K \times 1$ vectors $z = \{z^n\}_{n=1}^N$. Each member z^n is defined as a vector with its k^{th} variate set to 1 if the corresponding n^{th} Stokes vector s^n belongs to the k^{th} segment ; otherwise, it is set to zero. Also, every Stokes vector is assumed to belong to exactly one segment.

1. The noise model is straightforwardly extensible to consider colored noise.

For each image segment k we define a probability distribution (*kernel*) generating the Stokes vectors belonging to the corresponding segment. While in image segmentation problems the chosen distribution is typically Gaussian [McLachlan 2000], in the present problem we need a kernel choice that will assign zero probability mass to vectors not complying with the constraints given in (A.37). To this end, the Gaussian distribution is not convenient. To work around this particularity, we can instead consider a probability distribution on a special parametrization of the Stokes vectors. This parametrization is one of the contributions of the present model.

Let $\lambda = [\lambda_1\ \lambda_2\ \lambda_3\ \lambda_4]^T$ be a parametrization φ of the Stokes vector $s = [s_1\ s_2\ s_3\ s_4]^T$, with $s = \varphi(\lambda)$. Let matrices Φ and Λ be

$$\Phi = \begin{bmatrix} s_1 + s_4 & s_2 - i\ s_3 \\ s_2 + i\ s_3 & s_1 - s_4 \end{bmatrix},\ \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ \lambda_3 + i\ \lambda_4 & \lambda_2 \end{bmatrix}.$$

The transformation φ is defined to comply with

$$\Phi = \Lambda\Lambda^H. \tag{A.39}$$

It can be easily seen that any real vector λ will yield, according to transformation (A.39), a Stokes vector S that will necessarily satisfy the constraints (A.37). We can conveniently assume a Gaussian *i.i.d.* distribution hypothesis on the λ vectors :

$$\lambda^n | z_k^n = 1, \mu_k, \Sigma_k \ \sim\ \mathcal{N}(\mu_k, \Sigma_k). \tag{A.40}$$

In turn, the label vectors z are multinomially *i.i.d* distributed. This distribution is parametrized by the *contextual mixing proportions* set Π :

$$z^n | \boldsymbol{\pi}^n \sim Mult(\boldsymbol{\pi}^n).$$

The prior probability vectors $\Pi = \{\boldsymbol{\pi}^n\}_{n=1}^N$ are subject to the positiveness $\pi_k^n \geqslant 0$, $\forall k \in [1, \dots K]$, $\forall n \in [1, \dots N]$ and sum-to-unity constraints $\sum_{k=1}^K \pi_k^n = 1$, $\forall n \in [1, \dots N]$.

Considering the set of *contextual mixing proportions* Π as random variables and assuming a proper prior, we can incorporate the spatial smoothness trait, of which we have referred to earlier in the section, in an indirect way by forcing neighboring Stokes vectors to be more likely to share the same class label. We assume a Markov random field on Π, which equivalently means that Π is governed by a Gibbs distribution [Geman & Geman 1984], generally expressed by :

$$p(\Pi) \propto \prod_C e^{-\psi_c(\Pi)}, \tag{A.41}$$

where ψ_c is a function on clique c, called *clique potential* function in the literature, and the product is over all minimal cliques of the Markov random field.

An appropriate clique distribution choice would be to assume that the local differences of *contextual mixing proportions* follow a Student-t distribution, with its peak set at zero. This choice, proposed in a natural image segmentation context in [Sfikas et al. 2008B], also provides our model the properties of an edge-preserving line-process [Geman & Geman 1984]. The probability law for local differences is thus set to the Student's-t distribution :

$$\pi_k^n - \pi_k^j \sim \mathcal{S}t(0, \beta_{kd}^2, \nu_{kd}),$$
$$\forall n \in [1, \dots N], k \in [1, \dots K], d \in [1, \dots D], \forall j \in \gamma_d(n). \tag{A.42}$$

The parameters β_{kd} control how tightly smoothed we need the Stokes vectors of segment k to be. In (A.42), D stands for the number of a pixel's neighborhood adjacency types and $\gamma_d(n)$ is the set of neighbors of pixel indexed by n, with respect to the d^{th} adjacency type. In

our model, we assume 4 neighbors for each pixel (*first-order* neighborhood), and partition the corresponding adjacency types into horizontal and vertical, thus setting $D = 2$. This variability of parameter aims to capture the intuitive property that smoothness statistics may vary along clusters and spatial directions [Nikou et al. 2007].

One can see that the assumption in (A.42) is equivalent to

$$\pi_k^n - \pi_k^j \sim \mathcal{N}(0, \beta_{kd}^2/u_k^{nj}),$$

$$u_k^{nj} \sim \mathcal{G}(\nu_{kd}/2, \nu_{kd}/2), \quad \forall n, k, d, \quad \forall j \in \gamma_d(n),$$

where $\mathcal{N}$ and $\mathcal{G}$ represent a Gaussian and a Gamma distribution respectively. This breaking-down of the Student's-t distribution allows clearer insight on how our implicit edge-preserving line-process works. Since u_k^{nj} depends on datum indexed by n, each weight difference in the MRF can be described by a different instance of a Gaussian distribution. Therefore, as $u_k^{nj} \to +\infty$ the distribution tightens around zero, and forces neighboring *contextual mixing proportions* to be smooth. On the other hand, $u_k^{nj} \to 0$ signifies the existence of an edge and consequently no smoothing. This generative model can be examined in detail in figure A.14.

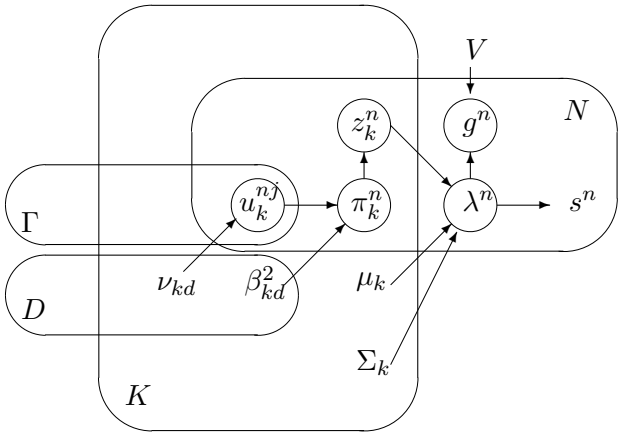


Figure A.14 – Graphical model for the proposed Stokes image restoration model. Stokes vectors s^n constitute the estimated restoration, produced by observations g^n. The rest of the model, namely random variable sets u_k^{nj}, π_k^n, z_k^n, constitute the prior for the proposed Stokes vector transformation λ^n. Superscript $n \in [1, \dots N]$ denotes pixel index, subscript $k \in [1, \dots K]$ denotes kernel (segment) index, subscript $d \in [1, \dots D]$ describes the neighborhood direction type and superscript $j \in [1, \dots \Gamma]$ denotes neighbor index.

The problem of inference formulates in our case as, given observations G, estimate the true Stokes image S. To achieve this, we need to find estimates jointly for all unknown parameters $\Psi = \{\mu, \Sigma, \beta, \nu\}$, Π and λ ; thence the restoration S can be computed given the λ estimate. The rest of the variables are considered as *hidden*, and are namely the labels z and the edge map u. While we do not need to determine estimates for the hidden variables, they play an important role in the model inference in an indirect way.

Hence, we need to optimize the model evidence, given by

$$\ln p(g, \lambda, \Pi; \Psi), \tag{A.43}$$

with respect to parameters Ψ, λ, Π. In this sense, our method is a maximum *a posteriori* (MAP) estimation, with λ and Π being the conditioned variables. As we cannot find MAP estimates in closed form for (A.43), we employ the EM algorithm [Tzikas et al. 2008]. In EM terminology, eq.(A.43) is referred to as the *incomplete* likelihood, while the complete log-likelihood is expressed by

$$\ln p(g, \lambda, \Pi, z, u; \Psi). \tag{A.44}$$

The conditional expectation of the complete likelihood is an important quantity in EM. It is defined as

$$\mathcal{E}_{z^{(t)},u^{(t)}|gp,\lambda p^{(t)},\Pi^{(t)}}\Big\{\ln p(g,\lambda,\Pi,z,u;\Psi)\Big\}. \tag{A.45}$$

By optimizing this expectation with respect to Ψ, Π and λ given the observed variables and some initial estimates $\Psi^{(0)}$, $\Pi^{(0)}$, $\lambda^{(0)}$ we can produce a new estimate $\Psi^{(1)}$, $\Pi^{(1)}$, $\lambda^{(1)}$. In the same way, estimates are computed iteratively. It can be proved that these estimates converge to a local optimum for the incomplete likelihood of eq. (A.43). This is the main idea in the EM algorithm. The iteration scheme is split in two steps, the Expectation and the Maximization step.

The E-step consists in computing the joint expectation of the hidden variables z and u, with respect to current iteration parameters $\Psi^{(t)}$, $\lambda^{(t)}$, $\Pi^{(t)}$ where t denotes the number of current iteration. Observing the graphical model in fig.A.14, we can see that given G, Π and λ, z and u are conditionally independent. Therefore $\mathcal{E}_{z,u|g,\lambda,\Pi}(\cdot) = \mathcal{E}_{z|g,\lambda,\Pi}\{(\mathcal{E}_{u|gp,\lambda,\Pi}(\cdot)\}$ and we can compute these expectations separately.

Due to lack of space, we point the reader to [Sfikas et al. 2008A] for the analytical expressions of the updates for parameters μ, Σ, β, ν, Π, and for the expected values of z and u, in a similar model proposed for natural image segmentation.

The update for the noise covariance matrix estimate V is given by

$$V^{(t+1)} = (4N)^{-1}\sum_{n=1}^{N}(Hg^n - s^{n(t)})^T(Hg^n - s^{n(t)})I. \tag{A.46}$$

Optimization with respect to the constraint-free parameters λ involves the following expression, after dropping constant terms from (A.45) :

$$(g^n - H\varphi(\lambda^{n(t)}))^T V^{-1}(g^n - H\varphi(\lambda^{n(t)}))+$$

$$\sum_{k=1}^{K}(\lambda^{n(t)} - \mu_k)^T\Sigma_k^{-1}(\lambda^{n(t)} - \mu_k) <z_k^n>,$$

which after some manipulation boils down to :

$$h^T\Omega^1 h + \lambda^T\Omega^2\lambda + \omega^3 h + \omega^4\lambda, \tag{A.47}$$

where we have omitted the data and iteration indices n and t for brevity. Parameters h, Ω^1, Ω^2, ω^3, ω^4 are given by

$$h \equiv H\varphi(\lambda), \Omega^1 \equiv V^{-1}, \Omega^2 \equiv \sum_{k=1}^{K} <z_k> \Sigma_k^{-1},$$

$$\omega^3 \equiv -2g^T V^{-1}, \omega^4 \equiv -2\sum_{k=1}^{K} <z_k> \mu_k^T\Sigma_k^{-1}.$$

In view of (A.39), eq. (A.47) is a fourth-order polynomial over each of the variates of λ. Setting the derivative of (A.47) with respect to each of the λ variates to zero, we can obtain optimizers for λ by solving the resulting third-order polynomial equations. Thus for each $n \in [1,\dots N]$, we solve iteratively four third-order polynomial equations, and repeat the operation until convergence of λ.

Finally, in order compute the Stokes estimates s^n we simply make use of the φ transformation definition (A.47) to obtain the update

$$\begin{cases} s_1^n = \frac{1}{2}\left[(\lambda_1^n)^2 + (\lambda_2^n)^2 + (\lambda_3^n)^2 + (\lambda_4^n)^2\right], \\ s_2^n = \lambda_1^n\ \lambda_3^n, \\ s_3^n = \lambda_1^n\ \lambda_4^n, \\ s_4^n = \frac{1}{2}\left[(\lambda_1^n)^2 - (\lambda_2^n)^2 - (\lambda_3^n)^2 - (\lambda_4^n)^2\right]. \end{cases} \tag{A.48}$$

5.2 Experimental results

We have applied the proposed recovery algorithm to two test Stokes images, one artificial image of size 64×64 and one real image of size 256×256. On the artificial image, the experiment was conducted by reproducing the blurring / noise model of (A.38) and applying varying levels of noise variance σ^2. We also used different assumed numbers of underlying segments K. The obtained results are shown in table A.6. These are computed as the improvement over SNR for the degraded image, given by

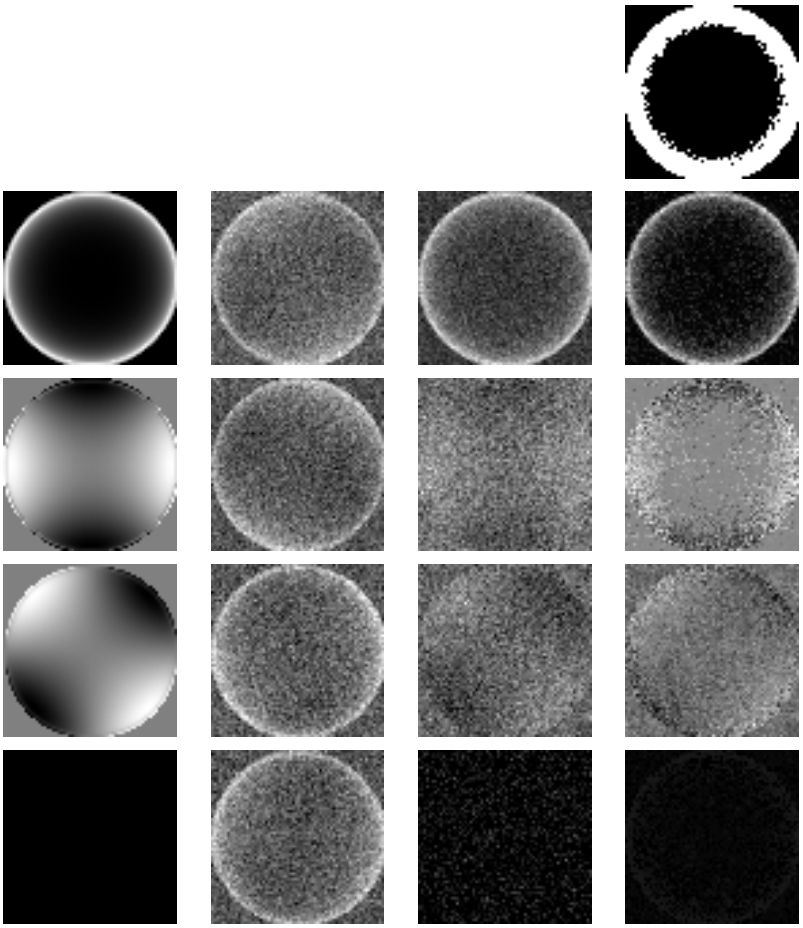
$$ISNR = 20 \log_{10} \frac{\|s^\star - g\|}{\|s^\star - \hat{s}\|}, \tag{A.49}$$

where $s^\star, \hat{s}, g$ represent the ground truth, the estimate, and the degraded (observed) image respectively. The pseudo-inverse estimates are as well computed for the restoration using $\hat{s}^n = (H^T H)^{-1} H^T g^n, \forall n \in [1, \dots N]$. The results clearly demonstrate that the proposed method gives a consistently better restoration compared to the general purpose pseudo-inverse estimator. Note also that the pseudo-inverse estimate will not necessarily yield values that satisfy the constraints (A.37). The main advantage of our method is that it takes explicitly into account the Stokes admissibility constraints (A.37), which is not the case for standard restoration methods. Representative visual results for the artificial image under $5dB$ noise are shown in figure A.15. Corresponding results for a real Stokes image are shown in figure A.16. The pseudo-inverse estimate of the real Stokes image contains 285 inadmissible vectors, out of a total of 65535 vectors; under the heavy-noise scenario of fig.A.15, this figure goes up to 1316 inadmissible vectors out of a total of 4096. Our method on the other hand, ensures always admissibility for *all* recovered vectors.

Runtimes for our algorithm were approximately 7 and 400 seconds respectively for the artificial and the much larger and more complex real Stokes image, for each EM iteration. We found that our algorithm converged in each case at around 10 EM iterations. Computations were done on a dual core 1.8 GHz PC workstation.

TABLE A.6 – *Restoration error results on the simulated Stokes data of fig. A.15. The image was degraded by varying noise levels. The presented values are the restoration ISNR (A.49); higher values correspond to better restorations. Results are shown for various numbers of classes K of the underlying segmentation, as well as the result of the pseudo-inverse estimate (PI).*

	PI	Proposed method		
SNR		$K=3$	$K=5$	$K=7$
$20dB$	21.4	22.4	22.3	22.7
$10dB$	11.4	14.7	14.5	14.5
$5dB$	6.3	10.9	9.9	9.9
$1dB$	2.3	7.3	7.6	7.6

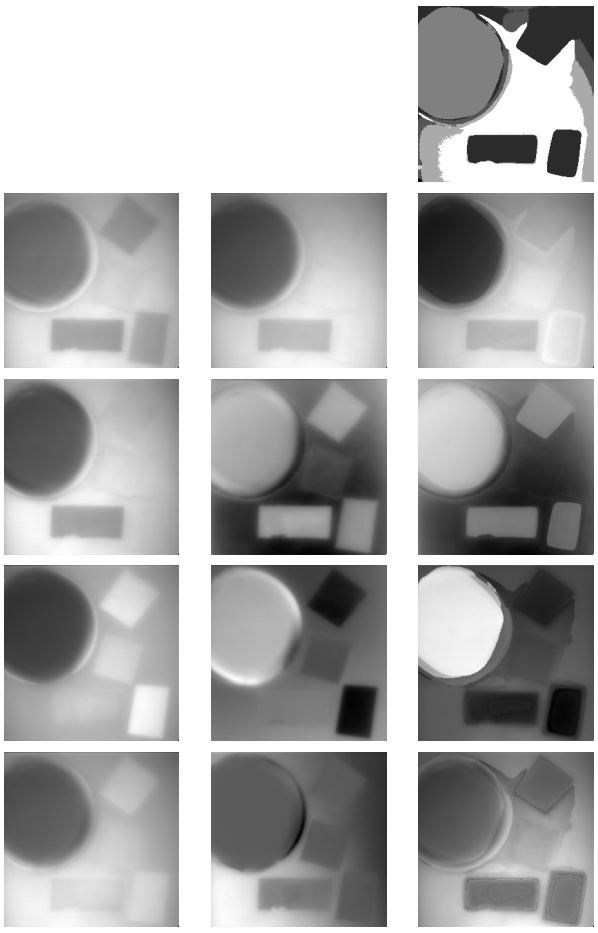


(a) (b) (c) (d)

FIGURE A.15 – Recovery result for simulated Stokes data under significant degradation. From left to right, each column shows the four channels of (a) the original Stokes image $s^\star$, (b) the degraded image G (SNR of 5 dB) (c) the non-complying to Stokes constraints pseudo-inverse recovery estimate and (d) the recovered image $\hat{s}$ obtained with our method. The corresponding segmentation of the degraded image into $K = 2$ classes is shown at the top of the (d) column.

5.3 Conclusion

We have presented an image recovery methodology suitable for Stokes images. Making use of a smoothing prior which assumes an underlying image segmentation and a suitable Stokes vector parametrization, we are able to produce a good estimate of the real image that at the same time satisfies the Stokes vector constraints (A.37). Also, we simultaneously produce a segmentation of the input image due to the model structure. Future work could be directed towards more sophisticated prior models, adapted properly in order to handle the distinctive difficulties of the Stokes image recovery problem.



(a) (b) (c)

FIGURE A.16 – Recovery result for real Stokes data. From left to right, each column shows the four channels of (a) the observed image G (b) the non-complying to Stokes constraints pseudo-inverse recovery estimate and (c) the recovered image $\hat{s}$ obtained with our method. The corresponding segmentation of the observed image into $K = 7$ classes is shown at the top of the (c) column.

Part V

References

Author publications

Articles in peer-reviewed international journals

[1] Sfikas G., Renard F., Heinrich C., Nikou C., « Estimating pathology discriminative deformation using trajectories over low-dimensional manifolds. Application to hippocampus MRI discrimination for Alzheimer's disease », *Medical Image Analysis*, submitted, 2012.

[2] Philippi N., Noblet V., Botzung A., Després O., Renard F., Sfikas G., Cretin B., Kremer S., Manning L., Blanc F., « MRI-Based Volumetry Correlates of Autobiographical Memory in Alzheimer's Disease », *PLoS ONE*, vol. 7, num. 10, e46200, 2012.

[3] Sfikas G., Heinrich C., Zallat J., Nikou C., Galatsanos N.P., « Recovery of polarimetric Stokes images by spatial mixture models », *Journal of the Optical Society of America A* , vol. 28, num. 5, pp. 465-474, 2011.

[4] Sfikas G., Nikou C., Galatsanos N.P., Heinrich C., « Spatially varying mixtures incorporating line processes for image segmentation », *Journal of Mathematical Imaging and Vision*, vol. 36, num. 2, pp. 91-110, 2010.

Communications in peer-reviewed international conferences with proceedings

[5] Sfikas G., Nikou C., Galatsanos N.P., Heinrich C., « Majorization-Minimization mixture model determination in image segmentation », Proceedings of *IEEE Computer Vision and Pattern Recognition conference* (CVPR), Colorado Springs, USA, pp. 2169-2176, June 2011.

[6] Sfikas G., Heinrich C., Nikou C., « Multiple atlas inference and population analysis with spectral clustering », Proceedings of *IEEE International Conference on Pattern Recognition* (ICPR), Istanbul, Turkey, pp. 2500-2503, August 2010.

[7] Sfikas G., Heinrich C., Zallat J., Nikou C., Galatsanos N.P., « Joint recovery and segmentation of polarimetric images using a compound MRF and mixture modeling », Proceedings of *IEEE International Conference on Image Processing* (ICIP), Cairo, Egypt, pp. 3901-3904, November 2009.

[8] Sfikas G., Nikou C., Heinrich C., Galatsanos N.P., « On optimization of probability vector random fields used for image segmentation », Proceedings of *IEEE International workshop on Machine Learning for Signal Processing* (MLSP), Grenoble, France, pp. 1-6, September 2009.

Publications related to research realized before the current thesis

Communications in peer-reviewed international conferences with proceedings

[9] Sfikas G., Nikou C., Galatsanos N.P., Heinrich C., « MR Brain Tissue classification using an edge-preserving spatially variant Bayesian mixture model », Proceedings of *Medical Image Computing and Computer Assisted Intervention* (MICCAI), New York, USA, pp. 43-50, September 2008.

[10] Sfikas G., Nikou C., Galatsanos N.P., « Edge-preserving spatially-varying mixtures for image segmentation », Proceedings of *IEEE Computer Vision and Pattern Recognition* (CVPR), Anchorage, USA, pp. 1-7, June 2008.

[11] Sfikas G., Nikou C., Galatsanos N.P., « Robust image segmentation with mixtures of student's t-distributions », Proceedings of *IEEE International Conference on Image Processing* (ICIP), San Antonio, USA, pp. 273-276, September 2007.

[12] Sfikas G., Constantinopoulos C., Likas A., Galatsanos N.P., « An analytic distance metric for Gaussian mixture models with application in image retrieval », Proceedings of *International conference on Artificial Neural Networks* (ICANN), Warsaw, Poland, part II, pp. 835-840, September 2005.

Bibliography

[Aljabar et al. 2010] P. Aljabar, R. Wolz, L. Srinivasan, S. Counsell, J. Boardman, M. Murgasova, V. Doria, M. Rutherford, A. Edwards, J. Hajnal & D. Rueckert (2010), «Combining morphological information in a manifold learning framework : application to neonatal MRI», in *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, Beijing, China.

[Allen et al. 2004] P. Allen, S. Feiner, A. Troccoli, H. Benko, E. Ishak & B. Smith (2004), «Seeing into the past : creating a 3D modeling pipeline for archaeological visualization», in *International Symposium on 3D Data Processing, Visualization and Transmission (3DPVT)*, pages 751 – 758, Sept. 2004.

[Ashburner & Friston 2000] J. Ashburner & K. J. Friston (2000), «Voxel-based morphometry–the methods», *NeuroImage*, **11**, n° 6, pages 805 - 821.

[Ashburner et al. 1998] J. Ashburner, C. Hutton, R. Frackowiak, I. Johnsrude, C. Price & K. Friston (1998), «Identifying global anatomical differences : deformation-based morphometry», *Human Brain Mapping*, **6**, n° 5-6, pages 348–357.

[Awate et al. 2006] S. P. Awate, T. Tasdizen, N. Foster & R. Whitaker (2006), «Adaptive Markov modeling for mutual information based, unsupervised MRI brain tissue classification», *Medical Image Analysis*, **10**, pages 726-739.

[Awate et al. 2007] S. P. Awate, H. Zhang & J. C. Gee (2007), «A fuzzy, nonparametric segmentation framework for DTI and MRI analysis : with applications to DTI tract extraction», *IEEE Transactions on Medical Imaging*, **26**, pages 1525-1536.

[Ban & Abe 2006] T. Ban & S. Abe (2006), «Implementing multi-class classifiers by one-class classification methods», in *International Joint Conference on Neural Networks (IJCNN)*, pages 327 –332, 2006.

[Beg et al. 2005] M. F. Beg, M. I. Miller, A. Trouvé & L. Younes (2005), «Computing large deformation metric mappings via geodesic flows of diffeomorphisms», *International Journal of Computer Vision*, **61**, pages 139-157.

[Belkin & Niyogi 2003] M. Belkin & P. Niyogi (2003), «Laplacian eigenmaps for dimensionality reduction and data representation», *Neural Computation*, **15**, pages 1373–1396.

[Bellman 1961] R. Bellman (1961), *Adaptive control processes.*, Princeton University Press, Princeton, NJ.

[Benjamini & Hochberg 1995] Y. Benjamini & Y. Hochberg (1995), «Controlling the false discovery rate : a practical and powerful approach to multiple testing», *Journal of the Royal Statistical Society B*, **57**, n° 1, pages 289–300.

[Bishop 2006] C. M. Bishop (2006), *Pattern recognition and machine learning*, Springer.

[Blake & Zisserman 1987] A. Blake & A. Zisserman (1987), *Visual reconstruction*, The MIT Press.

[Blekas et al. 2005] K. Blekas, A. Likas, N. Galatsanos & I. Lagaris (2005), «A spatially constrained mixture model for image segmentation», *IEEE Transactions on Neural Networks*, **16**, n° 2, pages 494-498.

[Blezek & Miller 2007] D. Blezek & J. Miller (2007), «Atlas stratification», *Medical Image Analysis*, **11**, pages 443–457.

[Blum 1967] H. Blum (1967), «A transformation for extracting descriptors of shape», in *Models for the Perception of Speech and Visual Forms*, pages 362–380, 1967.

[Blum & Nagel 1978] H. Blum & R. N. Nagel (1978), «Shape description using weighted symmetric axis features», *Pattern Recognition*, **10**, n° 3, pages 167 - 180.

[Bookstein 1996] F. Bookstein (1996), «Landmark methods for forms without landmarks : localizing group differences in outline shape», in *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis*, pages 279 –289, June 1996.

[Boyd & Vandenberghe 2004] S. Boyd & L. Vandenberghe (2004), *Convex optimization*, Cambridge University Press.

[Bozzali et al. 2002] M. Bozzali, A. Falini, M. Franceschi, M. Cercignani, M. Zuffi, G. Scotti, G. Comi & M. Filippi (2002), «White matter damage in Alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging», *Journal of Neurology, Neurosurgery and Psychiatry*, **72**, n° 6, pages 742-746.

[Brechbühler et al. 1995] C. Brechbühler, G. Gerig & O. Kübler (1995), «Parametrization of closed surfaces for 3-d shape description», *Computer Vision and Image Understanding*, **61**, n° 2, pages 154 - 170.

[Brucher 2008] M. Brucher (2008), *Représentations compactes et apprentissage non supervisé de variétés non linéaires. Application au traitement d'images.*, PhD thesis, University Louis Pasteur - Strasbourg I.

[Cao et al. 2011] Y. Cao, Y. Yuan, X. Li & P. Yan (2011), «Putting images on a manifold for atlas-based image segmentation», in *Proceedings of the IEEE International Conference on Image Processing*, pages 721–724, Brussels, Belgium.

[Carreira-Perpinan 2006] M. A. Carreira-Perpinan (2006), «Fast nonparametric clustering with Gaussian blurring mean-shift», in *Proceedings of the International Conference on Machine Learning (ICML'06)*, pages 153–160, 2006.

[Carson et al. 2002] C. Carson, S. Belongie, H. Greenspan & J. Malik (2002), «Blobworld : image segmentation using Expectation-Maximization and its application to image querying», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, n° 8, pages 1026-1038.

[Chen et al. 2011] K. Chen, N. Ayutyanont, J. B. Langbaum, A. S. Fleisher, C. Reschke, W. Lee, X. Liu, D. Bandy, G. E. Alexander, P. M. Thompson, L. Shaw, J. Q. Trojanowski, C. R. Jack Jr., S. M. Landau, N. L. Foster, D. J. Harvey, M. W. Weiner, R. A. Koeppe, W. J. Jagust & E. M. Reiman (2011), «Characterizing Alzheimer's disease using a hypometabolic convergence index», *NeuroImage*, **56**, n° 1, pages 52–60.

[Chung et al. 2010] M. K. Chung, K. J. Worsley, B. M. Nacewicz, K. M. Dalton & R. J. Davidson (2010), «General multivariate linear modeling of surface shapes using SurfStat», *NeuroImage*, **53**, pages 491–505.

[Cockrell & Folstein 2002] J. R. Cockrell & M. F. Folstein (2002), *Mini-mental state examination*, pages 140–141, John Wiley and Sons, Ltd.

[Collins et al. 1998] D. Collins, A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes & A. Evans (1998), «Design and construction of a realistic digital brain phantom», *IEEE Transactions on Medical Imaging*, **17**, n° 3, pages 463–468.

[Comaniciu & Meer 2002] D. Comaniciu & P. Meer (2002), «Mean shift : a robust approach toward feature space analysis», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, n° 5, pages 603-619.

[Constantinopoulos & Likas 2007] C. Constantinopoulos & A. Likas (2007), «Unsupervised learning of Gaussian mixtures based on variational component splitting», *IEEE Transactions on Neural Networks*, **18**, n° 3, pages 745–755.

[Cootes & Taylor 1995] T. Cootes & C. Taylor (1995), «Combining point distribution models with shape models based on finite element analysis», *Image and Vision Computing*, **13**, n° 5, pages 403 - 409.

[Cootes et al. 2008] T. Cootes, C. Twining, K. Babalola & C. Taylor (2008), «Diffeomorphic statistical shape models», *Image and Vision Computing*, **26**, pages 326–332.

[Corduneanu & Bishop 2001] A. Corduneanu & C. M. Bishop (2001), «Variational Bayesian model selection for mixture distributions», *Artificial Intelligence and Statistics*, pages 27–34.

[Cortes & Vapnik 1995] C. Cortes & V. Vapnik (1995), «Support-vector networks», *Machine Learning*, **20**, pages 273-297.

[Costafreda et al. 2011] S. G. Costafreda, I. D. Dinov, Z. Tu, Y. Shi, C.-Y. Liu, I. Kloszewska, P. Mecocci, H. Soininen, M. Tsolaki, B. Vellas, L.-O. Wahlund, C. Spenger, A. W. Toga, S. Lovestone & A. Simmons (2011), «Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment», *NeuroImage*, **56**, n° 1, pages 212 - 219.

[Cox et al. 2000] T. F. Cox, M. A. A. Cox & T. F. Cox (2000), *Multidimensional Scaling, Second Edition*, Chapman & Hall/CRC.

[Csernansky et al. 2005] J. Csernansky, L. Wang, J. Swank, J. Miller, M. Gado, D. McKeel, M. Miller & J. Morris (2005), «Preclinical detection of Alzheimer's disease : hippocampal shape and volume predict dementia onset in the elderly», *NeuroImage*, **25**, n° 3, pages 783 - 792.

[Csernansky et al. 2004] J. G. Csernansky, L. Wang, S. C. Joshi, J. T. Ratnanather & M. I. Miller (2004), «Computational anatomy and neuropsychiatric disease : probabilistic assessment of variation and statistical inference of group difference, hemispheric asymmetry, and time-dependent change», *NeuroImage*, **23**, pages S56 - S68 (Supplement 1).

[Cuingnet et al. 2011] R. Cuingnet, E. Gerardin, J. Tessieras, G. Auzias, S. Lehéricy, M.-O. Habert, M. Chupin, H. Benali & O. Colliot (2011), «Automatic classification of patients with Alzheimer's disease from structural MRI : a comparison of ten methods using the ADNI database», *NeuroImage*, **56**, n° 2, pages 766 - 781.

[DaSilva 2007] A. R. F. DaSilva (2007), «A Dirichlet process mixture model for brain MRI tissue classification», *Medical Image Analysis*, **11**, pages 169–182.

[Davatzikos 2004] C. Davatzikos (2004), «Why voxel-based morphometric analysis should be used with great caution when characterizing group differences», *NeuroImage*, **23**, n° 1, pages 17 - 20.

[Davies et al. 2008] R. Davies, C. Twining & C. Taylor (2008), *Statistical models of shape : optimisation and evaluation*, Springer, Dordrecht.

[Davies et al. 2010] R. Davies, C. Twining, T. Cootes & C. Taylor (2010), «Building 3-D statistical shape models by direct optimization», *IEEE Transactions on Medical Imaging*, **29**, n° 4, pages 961 -981.

[Davies et al. 2001] R. H. Davies, T. F. Cootes, C. J. Twining & C. J. Taylor (2001), «An information theoretic approach to statistical shape modelling», in *Proceedings of the British Machine Vision Conference (BMVC)*, 2001.

[deJager et al. 2010] C. deJager, T. Honey, J. Birks & G. Wilcock (2010), «Retrospective evaluation of revised criteria for the diagnosis of Alzheimer's disease using a cohort with post-mortem diagnosis», *International Journal of Geriatric Psychiatry*, **25**, pages 988–997.

[Dempster et al. 1977] P. Dempster, N. M. Laird & D. B. Rubin (1977), «Maximum likelihood from incomplete data via the EM algorithm», *Journal of the Royal Statistical Society B*, **39**, n° 1, pages 1-38.

[Donoho & Grimes 2003] D. Donoho & C. Grimes (2003), «Hessian eigenmaps : Locally linear embedding techniques for high-dimensional data», *Proceedings of the National Academy of Sciences of the USA*, **100**, n° 10, pages 5591–5596.

[Donoho & Grimes 2006] D. Donoho & C. Grimes (2006), «Image manifolds which are isometric to Euclidean space», *Journal of Mathematical Imaging and Vision*, **23**, n° 1, pages 5–24.

[Driscoll et al. 2003] I. Driscoll, D. A. Hamilton, H. Petropoulos, R. A. Yeo, W. M. Brooks, R. N. Baumgartner & R. J. Sutherland (2003), «The aging hippocampus : cognitive, biochemical and structural findings», *Cerebral Cortex*, **13**, n° 12, pages 1344-1351.

[Dubois et al. 2007] B. Dubois, H. Feldman, C. Jacova, S. Dekosky, P. Barberger-Gateau, D. A. G. D. Cummings, J., S. Gauthier, G. Jicha, K. Meguro, J. O'Brien, F. Pasquier, P. Robert, M. Rossor, S. Salloway, Y. Stern, P. Visser & P. Scheltens (2007), «Research criteria for the diagnosis of Alzheimer's disease : revising the NINCDS-ADRDA criteria», *Lancet Neurology*, **6**, pages 734–746.

[Durrleman et al. 2008] S. Durrleman, X. Pennec, A. Trouvé & N. Ayache (2008), «Sparse approximation of currents for statistics on curves and surfaces», *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pages 390–398.

[Duvernoy 1988] H. Duvernoy (1988), *The human hippocampus : an atlas of applied anatomy*, J.F. Bergmann Verlag, Munich.

[Eichenbaum 1999] H. Eichenbaum (1999), «The hippocampus and mechanisms of declarative memory», *Behavioural Brain Research*, **103**, n° 2, pages 123 - 133.

[Fan et al. 2007] Y. Fan, D. Shen, R. Gur, R. Gur & C. Davatzikos (2007), «COMPARE : Classification of morphological patterns using adaptive regional elements», *IEEE Transactions on Medical Imaging*, **26**, n° 1, pages 93 -105.

[Felzenswalb & Huttenlocher 2004] P. F. Felzenswalb & D. Huttenlocher (2004), «Efficient graph-based image segmentation», *International Journal of Computer Vision*, **59**, n° 2, pages 167-181.

[Feng et al. 2010] W. Feng, J. Jia & Z.-Q. Liu (2010), «Self-validated labeling of Markov random fields for image segmentation», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**, n° 10, pages 1871–1887.

[Filipovych & Davatzikos 2011] R. Filipovych & C. Davatzikos (2011), «Semi-supervised pattern classification of medical images : application to mild cognitive impairment (MCI)», *NeuroImage*, **55**, n° 3, pages 1109 - 1119.

[Fischl et al. 2002] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness,

A. Montillo, N. Makris, B. Rosen & A. M. Dale (2002), «Whole brain segmentation : automated labeling of neuroanatomical structures in the human brain», *Neuron*, **33**, n° 3, pages 341 - 355.

[Fletcher 2004] T. Fletcher (2004), *Statistical variability in nonlinear spaces*, PhD thesis, University of North Carolina.

[Fletcher et al. 2001] T. Fletcher, S. M. Pizer, A. Thall & A. G. Gash (2001), «Shape modeling and image visualization in 3d with m-rep object models», Technical report TR01-004, Dept. of Computer Science, Univ. of North Carolina, 2001.

[Freixenet et al. 2002] J. Freixenet, X. Munoz, D. Raba, J. Marti & X. Cuff (2002), «Yet another survey on image segmentation : region and boundary information integration», in *Lecture Notes in Computer Science. Proceedings of the European Conference on Computer Vision (ECCV'02)*, pages 408–422, 2002.

[Frisoni et al. 2010] G. B. Frisoni, N. C. Fox, C. R. Jack Jr., P. Scheltens & P. M. Thompson (2010), «The clinical use of structural MRI in Alzheimer disease», *Nature Reviews Neurology*, **6**, pages 67–77.

[Geman & Geman 1984] S. Geman & D. Geman (1984), «Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, n° 6, pages 721-741.

[Gerber et al. 2009] S. Gerber, T. Tasdizen, S. Joshi & R. Whitaker (2009), «On the manifold structure of the space of brain images», in *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2009.

[Gerber et al. 2010] S. Gerber, T. Tasdizen, P. T. Fletcher, S. Joshi & R. Whitaker (2010), «Manifold modeling for brain population analysis», *Medical Image Analysis*, **14**, n° 5, pages 643 - 653.

[Goate & Holtzman 2010] A. Goate & D. M. Holtzman (2010), «Biomarkers will revolutionize the way we diagnose and treat Alzheimer's disease», *Biomarkers in Medicine*, **4**, n° 1, pages 1-2.

[Golland 2000] P. Golland (2000), *Statistical shape analysis of anatomical structures*, PhD thesis, Massachusetts Institute of Technology.

[Golland 2001] P. Golland (2001), «Discriminative direction for kernel classifiers», *Advances in Neural Information Processing Systems (NIPS)*, pages 745–752.

[Golland et al. 2000] P. Golland, W. Eric & L. Grimson (2000), «Fixed topology skeletons», in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 10–17, 2000.

[Golland et al. 2005] P. Golland, W. E. L. Grimson, M. E. Shenton & R. Kikinis (2005), «Detection and analysis of statistical differences in anatomical shape», *Medical Image Analysis*, **9**, n° 1, pages 69 - 86.

[Greene et al. 2012] S. J. Greene, R. J. Killiany &the Alzheimer's Disease Neuroimaging Initiative (2012), «Hippocampal subregions are differentially affected in the progression to Alzheimer's disease», *The Anatomical Record : Advances in Integrative Anatomy and Evolutionary Biology*, **295**, n° 1, pages 132–140.

[Greenspan et al. 2006] H. Greenspan, A. Ruf & J. Goldberger (2006), «Constrained Gaussian mixture model framework for automatic segmentation of MR brain images», *IEEE Transactions on Medical Imaging*, **25**, pages 1233-1245.

[Hamm et al. 2010] J. Hamm, D. H. Ye, R. Verma & C. Davatzikos (2010), «GRAM : A framework for geodesic registration on anatomical manifolds», *Medical Image Analysis*, **14**, n° 5, pages 633 - 642.

[Heinrich 2008] C. Heinrich (2008), *Problèmes inverses, analyse statistique de formes et problèmes connexes*, Habilitation dissertation, University of Strasbourg.

[Hsu et al. 2010] C.-W. Hsu, C.-C. Chang & C.-J. Lin (2010), «A practical guide to support vector classification», Technical report, Department of Computer science, National Taiwan University.

[Hua et al. 2008A] X. Hua, A. D. Leow, S. Lee, A. D. Klunder, A. W. Toga, N. Lepore, Y.-Y. Chou, C. Brun, M.-C. Chiang, M. Barysheva, C. R. Jack Jr., M. A. Bernstein, P. J. Britson, C. P. Ward, J. L. Whitwell, B. Borowski, A. S. Fleisher, N. C. Fox, R. G. Boyes, J. Barnes, D. Harvey, J. Kornak, N. Schuff, L. Boreta, G. E. Alexander, M. W. Weiner, P. M. Thompson &the Alzheimer's Disease Neuroimaging Initiative (2008), «3D characterization of brain atrophy in Alzheimer's disease and mild cognitive impairment using tensor-based morphometry», *NeuroImage*, **41**, n° 1, pages 19 - 34.

[Hua et al. 2008B] X. Hua, A. D. Leow, N. Parikshak, S. Lee, M.-C. Chiang, A. W. Toga, C. R. Jack Jr., M. W. Weiner & P. M. Thompson (2008), «Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease : an MRI study of 676 AD, MCI, and normal subjects», *NeuroImage*, **43**, n° 3, pages 458–469.

[Hua et al. 2009] X. Hua, S. Lee, I. Yanovsky, A. D. Leow, Y.-Y. Chou, A. J. Ho, B. Gutman, A. W. Toga, C. R. Jack Jr., M. A. Bernstein, E. M. Reiman, D. J. Harvey, J. Kornak, N. Schuff, G. E. Alexander, M. W. Weiner & P. M. Thompson (2009), «Optimizing power to track brain degeneration in Alzheimer's disease and mild cognitive impairment with tensor-based morphometry : an ADNI study of 515 subjects», *NeuroImage*, **48**, n° 4, pages 668 - 681.

[Jack Jr. et al. 2003] C. R. Jack Jr., M. Slomkowski, S. Gracon, T. M. Hoover, J. P. Felmlee, K. Stewart, Y. Xu, M. Shiung, P. C. O'Brien, R. Cha, D. Knopman & R. C. Petersen (2003), «MRI as a biomarker of disease progression in a therapeutic trial of milameline for AD», *Neurology*, **60**, n° 2, pages 253-260.

[Jack Jr. et al. 2008] C. R. Jack Jr., M. A. Bernstein, N. C. Fox, P. Thompson, G. Alexander, D. Harvey, B. Borowski, P. J. Britson, J. L. Whitwell, C. Ward, A. M. Dale, J. P. Felmlee, J. L. Gunter, D. L. Hill, R. Killiany, N. Schuff, S. Fox-Bosetti, C. Lin, C. Studholme, C. S. DeCarli, G. Krueger, H. A. Ward, G. J. Metzger, K. T. Scott, R. Mallozzi, D. Blezek, J. Levy, J. P. Debbins, A. S. Fleisher, M. Albert, R. Green, G. Bartzokis, G. Glover, J. Mugler & M. W. Weiner (2008), «The Alzheimer's disease neuroimaging initiative (ADNI) : MRI methods», *Journal of Magnetic Resonance Imaging*, **27**, n° 4, pages 685–691.

[Jellinger 2006] K. A. Jellinger (2006), «Alzheimer 100 - highlights in the history of Alzheimer research», *Journal of Neural Transmission*, **113**, pages 1603-1623.

[Joshi et al. 1997] S. Joshi, U. Grenander & M. Miller (1997), «The geometry and shape of brain submanifolds», *International Journal Pattern Recognition Artificial Intelligence Special Issue on Processing of MR Images of the Human*, **11**, pages 1317–1343.

[Kandel et al. 1995] E. R. Kandel, J. H. Schwartz & T. M. Jessel (1995), *Essentials of neural science and behavior*, Prentice-Hall.

[Kanemura et al. 2007A] A. Kanemura, S. Maeda & S. Ishii (2007), «Edge-preserving Bayesian image superresolution based on compound Markov random fields», in *Proceedings of the 17th Int'l conference on Artificial Neural Networks, 2007*, Porto, Portugal.

[Kanemura et al. 2007B] A. Kanemura, S. Maeda & S. Ishii (2007), «Hyperparameter estimation in Bayesian image superresolution with a compound Markov random field

prior», in *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing*, Thessaloniki, Greece.

[Kang et al. 2006] W.-S. Kang, K. Im & J. Choi (2006), «SVDD-based method for fast training of multi-class support vector classifier», in J. Wang, Z. Yi, J. Zurada, B.-L. Lu & H. Yin, editors, *Advances in Neural Networks - ISNN 2006*, volume 3971 of *Lecture Notes in Computer Science*, pages 991–996, Springer Berlin / Heidelberg.

[Kapur et al. 1996] T. Kapur, W. E. L. Grimson, W. M. Wells III & R. Kikinis (1996), «Segmentation of brain tissue from magnetic resonance images», *Medical Image Analysis*, **1**, pages 109-127.

[Kapur et al. 1998] T. Kapur, W. E. L. Grimson, R. Kikinis &W. M. Wells III (1998), «Enhanced spatial priors for segmentation of magnetic resonance imagery», in *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'98)*, pages 457–468, 1998.

[Klöppel et al. 2008] S. Klöppel, C. M. Stonnington, C. Chu, B. Draganski, R. I. Scahill, J. D. Rohrer, N. C. Fox, C. R. Jack Jr., J. Ashburner & R. S. Frackowiak (2008), «Automatic classification of MR scans in Alzheimer's disease», *Brain*, **131**, pages 681–689.

[Koikkalainen et al. 2011] J. Koikkalainen, J. Lftjfnen, L. Thurfjell, D. Rueckert, G. Waldemar & H. Soininen (2011), «Multi-template tensor-based morphometry : application to analysis of Alzheimer's disease», *NeuroImage*, **56**, n° 3, pages 1134 - 1144.

[Koutsouleris et al. 2009] N. Koutsouleris, E. M. Meisenzahl, C. Davatzikos, R. Bottlender, T. Frodl, J. Scheuerecker, G. Schmitt, T. Zetzsche, P. Decker, M. Reiser, H.-J. Moller & C. Gaser (2009), «Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition», *Archives of General Psychiatry*, **66**, n° 7, pages 700-712.

[Kwan et al. 1999] R.-S. Kwan, A. Evans & G. Pike (1999), «MRI simulation-based evaluation of image-processing and classification methods», *IEEE Transactions on Medical Imaging*, **18**, pages 1085–1097.

[Laidlaw et al. 1998] D. H. Laidlaw, K. W. Fleischer & A. H. Barr (1998), «Partial-volume Bayesian classification of material mixtures in MR volume data using voxel histograms», *IEEE Transactions on Medical Imaging*, **17**, n° 1, pages 74–86.

[Lange 2004] K. Lange (2004), *Optimization*, Springer.

[Lee & Lee 2007] D. Lee & J. Lee (2007), «Domain described support vector classifier for multi-classification problems», *Pattern Recognition*, **40**, n° 1, pages 41 - 51.

[Lemm et al. 2011] S. Lemm, B. Blankertz, T. Dickhaus & K.-R. Müller (2011), «Introduction to machine learning for brain imaging», *NeuroImage*, **56**, n° 2, pages 387 - 399.

[Lempitsky et al. 2010] V. Lempitsky, C. Rother, S. Roth & A. Blake (2010), «Fusion moves for Markov random field optimization», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **32**, pages 1392-1405.

[Leventon et al. 2000] M. Leventon, W. Grimson & O. Faugeras (2000), «Statistical shape influence in geodesic active contours», in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 316–323, 2000.

[Leymarie & Levine 1992] F. Leymarie & M. Levine (1992), «Simulating the grassfire transform using an active contour model», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14**, n° 1, pages 56 -75.

[Lu et al. 1998] H. Lu, Y. Fainman & R. Hecht-Nielsen (1998), «Image manifolds», in *Proceedings of the SPIE Symposium on Electronic Imaging : Science and Technology ; Conference on Artificial Neural Networks in Image Processing III*, San Jose, California.

[Lustig et al. 2007] M. Lustig, D. Donoho & J. M. Pauly (2007), «Sparse MRI : the application of compressed sensing for rapid MR imaging», *Magnetic Resonance in Medicine*, **58**, n° 6, pages 1182–1195.

[Malykhin et al. 2010] N. Malykhin, R. Lebel, N. Coupland, A. Wilman & R. Carter (2010), «In vivo quantification of hippocampal subfields using 4.7T fast spin echo imaging», *NeuroImage*, **49**, n° 2, pages 1224 - 1230.

[Marcus et al. 2007] D. Marcus, T. Wang, J. Parker, J. Csernansky, J. Morris & R. Buckner (2007), «Open Access Series of Imaging Studies (OASIS) : cross-sectional MRI data in young, middle aged, nondemented, and demented older adults», *Journal of Cognitive Neuroscience*, **19**, pages 1498–1507.

[Marroquin et al. 2002] J. Marroquin, B. Vemuri, S. Botello, F. Calderon & A. Fernandez-Bouzas (2002), «An accurate and efficient Bayesian method for automatic segmentation of brain MRI», *IEEE Transactions on Medical Imaging*, **21**, pages 934-945.

[Marroquin et al. 2003] J. Marroquin, E. Arce & S. Botello (2003), «Hidden Markov measure field models for image segmentation», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **25**, n° 11, pages 1380-1387.

[Martin et al. 2001] D. Martin, C. Fowlkes, D. Tal & J. Malik (2001), «A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics», in *Proceedings of the 8th International Conference on Computer Vision (ICCV '01)*, volume 2, pages 416–423, July 2001.

[McKhann et al. 1984] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price & E. Stadlan (1984), «Clinical diagnosis of Alzheimer's disease : report of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services task force on Alzheimer's disease», *Neurology*, **34**, pages 939-944.

[McLachlan 2000] G. McLachlan (2000), *Finite mixture models*, Wiley-Interscience.

[Medical Dictionary 2012] Merriam-Webster's Medical Dictionary (2012), «Merriam-Webster's medical dictionary, lemma "biomarker"», http ://www.merriam-webster.com/medical/biomarker (Accessed 29/3/2012).

[Molina et al. 2003] R. Molina, J. Mateos, A. Katsaggelos & M. Vega (2003), «Bayesian multichannel image restoration using compound Gauss-Markov random fields», *IEEE Transactions on Image Processing*, **12**, pages 1642–1654.

[Mori 2005] G. Mori (2005), «Guiding model search using segmentation», in *ICCV'05 : Proceedings of the Tenth IEEE International Conference on Computer Vision*, volume 2, pages 1417–1423, 2005.

[Morra et al. 2010] J. Morra, Z. Tu, L. Apostolova, A. Green, A. Toga & P. Thompson (2010), «Comparison of adaboost and support vector machines for detecting Alzheimer's disease through automated hippocampal segmentation», *IEEE Transactions on Medical Imaging*, **29**, n° 1, pages 30 -43.

[Morris 1997] J. C. Morris (1997), «Clinical Dementia Rating : A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type», *International Psychogeriatrics*, **9**, pages 173-176.

[Morris et al. 2001] J. C. Morris, M. Storandt, J. P. Miller, D. W. McKeel, J. L. Price, E. H. Rubin & L. Berg (2001), «Mild cognitive impairment represents early-stage Alzheimer disease», *Archives of Neurology*, **58**, n° 3, pages 397-405.

[Mourão-Miranda et al. 2011] J. Mourão-Miranda, D. R. Hardoon, T. Hahn, A. F. Marquand, S. C. Williams, J. Shawe-Taylor & M. Brammer (2011), «Patient classification as an outlier detection problem : an application of the one-class support vector machine», *NeuroImage*, **58**, n° 3, pages 793 - 804.

[Mueller et al. 2011] S. Mueller, L. Chao, B. Berman & M. Weiner (2011), «Evidence for functional specialization of hippocampal subfields detected by MR subfield volumetry on high resolution images at 4T», *NeuroImage*, **56**, n° 3, pages 851 - 857.

[Mueller et al. 2006] S. G. Mueller, N. Schuff & M. W. Weiner (2006), «Evaluation of treatment effects in Alzheimer's and other neurodegenerative diseases by MRI and MRS», *NMR in Biomedicine*, **19**, n° 6, pages 655–668.

[Munkres 2000] J. R. Munkres (2000), *Topology*, Prentice Hall, 2nd edition.

[Nain et al. 2007] D. Nain, M. Styner, M. Niethammer, J. Levitt, M. Shenton, G. Gerig, A. Bobick & A. Tannenbaum (2007), «Statistical shape analysis of brain structures using spherical wavelets», in *Proceedings of the IEEE International symposium on Biomedical Imaging : From Nano to Macro*, pages 209 –212, April 2007.

[Nascimento & Silva 2010] J. Nascimento & J. Silva (2010), «Manifold learning for object tracking with multiple motion dynamics», in *Proceedings of the European Conference on Computer Vision (ECCV)*, volume 6313, pages 172–185, 2010.

[Niessen et al. 1998] W. J. Niessen, K. L. Vincken, J. Weickert & M. A. Viergever (1998), «Three dimensional MR brain segmentation», in *Proceedings of the* 6^{th} *International Conference on Computer Vision (ICCV'98)*, pages 53–58, 1998.

[Nikita 2009] I. Nikita (2009), *Comparative study of the rythmic activity induced by achetylcholine receptor agonist in dorsal and ventral hippocampal slices of the rat*, PhD thesis, University of Patras.

[Nikou et al. 2007] C. Nikou, N. Galatsanos & A. Likas (2007), «A class-adaptive spatially variant mixture model for image segmentation», *IEEE Transactions on Image Processing*, **16**, n° 4, pages 1121–1130.

[Nordberg 1999] A. Nordberg (1999), «PET studies and cholinergic therapy in Alzheimer's disease», *Revue Neurologique (Paris)*, **155**, pages S53-63.

[Orbanz & Buhmann 2007] P. Orbanz & J. M. Buhmann (2007), «Nonparametric Bayesian image segmentation», *International Journal of Computer Vision*, **77**, pages 25-45.

[Park et al. 2007] J. Park, D. Kang, J. Kim, J. T. Kwok & I. W.Tsang (2007), «SVDD-based pattern denoising», *Neural Computation*, **19**, n° 7, pages 1919-1938.

[Peel & McLachlan 2000] D. Peel & G. J. McLachlan (2000), «Robust mixture modeling using the *t*-distribution», *Statistics and Computing*, **10**, pages 339-348.

[Peng et al. 2006] Z. Peng, W. Wee & J. H. Lee (2006), «Automatic segmentation of MR brain images using spatial-varying Gaussian mixture and Markov random field approach», in *Proceedings of the Computer Vision and Pattern Recognition Workshop (CVPRW'06)*, 2006.

[Pennec et al. 2000] X. Pennec, N. Ayache & J.-P. Thirion (2000), «Landmark-based registration using features identified through differential geometry», in I. Bankman, editor, *Handbook of Medical Imaging - Processing and Analysis. I.*, pages 499–513, Academic Press.

[Petersen et al. 2001] R. C. Petersen, R. Doody, A. Kurz, R. C. Mohs, J. C. Morris, P. V. Rabins, K. Ritchie, M. Rossor, L. Thal & B. Winblad (2001), «Current concepts in mild cognitive impairment», *Archives of Neurology*, **58**, n° 12, pages 1985–1992.

[Petersen et al. 2009] R. C. Petersen, R. O. Roberts, D. S. Knopman, B. F. Boeve, Y. E. Geda, R. J. Ivnik, G. E. Smith & C. R. Jack Jr. (2009), «Mild cognitive impairment : Ten years later», *Archives of Neurology*, **66**, n° 12, pages 1447-1455.

[Philippi et al. 2012] N. Philippi, V. Noblet, A. Botzung, O. Després, F. Renard, G. Sfikas, B. Cretin, S. Kremer, L. Manning & F. Blanc (2012), «Neuronal correlates of autobiographical memory : an MRI-based volumetry study in Alzheimer's disease», Technical report.

[Pizer et al. 2003] S. M. Pizer, P. T. Fletcher, S. Joshi, A. Thall, J. Z. Chen, Y. Fridman, D. S. Fritsch, A. G. Gash, J. M. Glotzer, M. R. Jiroutek, C. Lu, K. E. Muller, G. Tracton, P. Yushkevich & E. L. Chaney (2003), «Deformable m-Reps for 3D medical image segmentation», *International Journal of Computer Vision*, **55**, pages 85-106.

[Pless & Souvenir 2009] R. Pless & R. Souvenir (2009), «A survey of manifold learning for images», *IPSJ Transactions on Computer Vision and Applications*, **1**, pages 83–94.

[Pohl et al. 2007] K. Pohl, S. Bouix, M. Nakamura, T. Rohfling, R. McCarley, R. Kikinis, L. Grimson, M. Shenton & W. Wells (2007), «A hierarchical algorithm for MR brain image parcellation», *IEEE Transactions on Medical Imaging*, **26**, pages 1201-1212.

[Prastawa et al. 2004] M. Prastawa, J. Gilmore, W. Lin & G. Gerig (2004), «Automatic segmentation of neonatal brain MRI», in *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'04)*, 2004.

[Ptucha et al. 2011] R. Ptucha, G. Tsagkatakis & A. Savakis (2011), «Manifold learning for simultaneous pose and facial expression recognition», in *Proceedings of the IEEE International Conference on Image Processing*, pages 3082–3085, Brussels, Belgium.

[Rajagopalan & Chaudhuri 1999] A. N. Rajagopalan & S. Chaudhuri (1999), «An MRF model-based approach to simultaneous recovery of depth and restoration from defocused images», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **21**, n° 17, pages 577-589.

[Ren & Malik 2003] X. Ren & J. Malik (2003), «Learning a classification model for segmentation», in *ICCV'03 : Proceedings of the IEEE International Conference on Computer Vision*, volume 1, pages 10–17, 2003.

[Renard 2011] F. Renard (2011), *Création et utilisation d'atlas en IRM de diffusion. Application à l'étude des troubles de la conscience*, PhD thesis, University of Strasbourg.

[Rosner 2000] B. Rosner (2000), *Fundamentals of Biostatistics*, Brooks/Cole, 5th edition.

[Rothenberg et al. 2011] K. G. Rothenberg, S. L. Siedlak, H.-G. Lee, X. Zhu, G. Perry & M. A. Smith (2011), «Neurodegenerative processes in Alzheimer's disease : an overview of pathogenesis with strategic biomarker potential», *Future Neurology*, **6**, n° 2, pages 173-185.

[Roweis & Saul 2000] S. Roweis & L. Saul (2000), «Nonlinear dimensionality reduction by locally linear embedding», *Science*, **290**, pages 2323-2326.

[Rueckert et al. 2003] D. Rueckert, A. Frangi & J. Schnabel (2003), «Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration», *IEEE Transactions on Medical Imaging*, **22**, n° 8, pages 1014–1025.

[Sabuncu et al. 2009] M. Sabuncu, S. Balci, M. Shenton & P. Golland (2009), «Image-driven population analysis through mixture modelling», *IEEE Transactions on Medical Imaging*, **28**, n° 9, pages 1473 -1487.

[Sakla et al. 2011] W. Sakla, A. Chan, J. Ji & A. Sakla (2011), «An SVDD-based algorithm for target detection in hyperspectral imagery», *IEEE Geoscience and Remote Sensing Letters*, **8**, n° 2, pages 384–388.

[Sanjay-Gopal & Hebert 1998] S. Sanjay-Gopal & T. Hebert (1998), «Bayesian pixel classification using spatially variant finite mixtures and the generalized EM algorithm», *IEEE Transactions on Image Processing*, **7**, n° 7, pages 1014-1028.

[Schmand et al. 2011] B. Schmand, P. Eikelenboom, W. A. van Gool &the Alzheimer's Disease Neuroimaging Initiative (2011), «Value of neuropsychological tests, neuroimaging, and biomarkers for diagnosing Alzheimer's disease in younger and older age cohorts», *Journal of the American Geriatrics Society*, **68**, n° 10, pages 1257–1266.

[Schönefeld 2005] V. Schönefeld (2005), «Spherical harmonics», Technical report, Computer Graphics and Multimedia Group, Technical Note. RWTH Aachen University.

[Schölkopf et al. 2001] B. Schölkopf, J. C. Platt, J. Shawe-taylor, A. J. Smola & R. C. Williamson (2001), «Estimating the support of a high-dimensional distribution», *Neural Computation*, **13**, n° 7, pages 1443–1471.

[Selkoe 1997] D. J. Selkoe (1997), «Alzheimer's Disease—Genotypes, Phenotype, and Treatments», *Science*, **275**, n° 5300, pages 630-631.

[Sfikas et al. 2008A] G. Sfikas, C. Nikou & N. Galatsanos (2008), «Edge preserving spatially varying mixtures for image segmentation», in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, Alaska, USA.

[Sfikas et al. 2008B] G. Sfikas, C. Nikou, N. Galatsanos & C. Heinrich (2008), «MR brain segmentation using an edge-preserving spatially variant Bayesian mixture model», in *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, New York, USA.

[Sfikas et al. 2010A] G. Sfikas, C. Heinrich & C. Nikou (2010), «Multiple atlas inference and population analysis with spectral clustering», in *Proceedings of the IEEE International Conference on Pattern Recognition*, pages 2500–2503, Istanbul, Turkey.

[Sfikas et al. 2010B] G. Sfikas, C. Nikou, N. Galatsanos & C. Heinrich (2010), «Spatially varying mixtures incorporating line processes for image segmentation», *Journal of Mathematical Imaging and Vision*, **36**, n° 2, pages 91–110.

[Sfikas et al. 2011A] G. Sfikas, C. Nikou, N. Galatsanos & C.Heinrich (2011), «Majorization-Minimization mixture model determination in image segmentation», in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, Colorado springs, USA.

[Sfikas et al. 2011B] G. Sfikas, C. Heinrich, J. Zallat, C. Nikou & N. Galatsanos (2011), «Recovery of polarimetric stokes images by spatial mixture models», *Journal of the Optical Society of America A*, **28**, n° 3, pages 465–474.

[Shawe-Taylor & Cristianini 2004] J. Shawe-Taylor & N. Cristianini (2004), *Kernel methods for pattern analysis*, Cambridge University Press.

[Shenton et al. 2002] M. E. Shenton, G. Gerig, R. W. McCarley, G. Székely & R. Kikinis (2002), «Amygdala–hippocampal shape differences in schizophrenia : the application of 3D shape models to volumetric MR data», *Psychiatry Research : Neuroimaging*, **115**, n° 1–2, pages 15 - 35.

[Shi & Malik 2000] J. Shi & J. Malik (2000), «Normalized cuts and image segmentation», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**, n° 8, pages 888-905.

[Sjöstrand et al. 2007] K. Sjöstrand, M. S. Hansen, H. B. Larsson & R. Larsen (2007), «A path algorithm for the support vector domain description and its application to medical imaging», *Medical Image Analysis*, **11**, n° 5, pages 417 - 428.

[Souvenir & Pless 2007] R. Souvenir & R. Pless (2007), «Image distance functions for manifold learning», *Image and Vision Computing*, **25**, n° 3, pages 365–373.

[Sparks & Madabhushi 2011] R. Sparks & A. Madabhushi (2011), «Out-of-sample extrapolation using semi-supervised manifold learning (OSE-SSL) : Content-based image retrieval for prostate histology grading», in *Proceedings of the IEEE International symposium on Biomedical Imaging : From Nano to Macro*, pages 734 –737, 2011.

[Staib & Duncan 1992] L. H. Staib & J. S. Duncan (1992), «Boundary finding with parametrically deformable models», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **14**, n° 11, pages 1061–1075.

[Stonnington et al. 2010] C. M. Stonnington, C. Chu, S. Klöppel, C. R. Jack Jr., J. Ashburner & R. S. Frackowiak (2010), «Predicting clinical scores from magnetic resonance scans in Alzheimer's disease», *NeuroImage*, **51**, n° 4, pages 1405 - 1413.

[Storey 2003] J. Storey (2003), «The positive false discovery rate : A Bayesian interpretation and the q-value», *The Annals of Statistics*, **31**, n° 6, pages 2013–2035.

[Studholme et al. 1999] C. Studholme, D. Hill & D. Hawkes (1999), «An overlap invariant entropy measure of 3D medical image alignment», *Pattern Recognition*, **32**, pages 71-86.

[Sudderth & Jordan 2008] E. Sudderth & M. Jordan (2008), «Shared segmentation of natural scenes using dependent Pitman-Yor processes», *Proceedings of Neural Information Processing Systems (NIPS)*, **21**.

[Tarawneh & Holtzman 2010] R. Tarawneh & D. M. Holtzman (2010), «Biomarkers in translational research of Alzheimer's disease», *Neuropharmacology*, **59**, n° 4-5, pages 310 - 322.

[Tasdizen et al. 2005] T. Tasdizen, S. Awate, R. Whitaker & N. Foster (2005), «MR tissue classification with neighborhood statistics : a nonparametric entropy-minimizing approach», in *Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI'05)*, pages 517–525, 2005.

[Tax & Duin 2004] D. Tax & R. Duin (2004), «Support vector data description», *Machine Learning*, **54**, pages 45-66.

[Tax & Duin 1999] D. M. Tax & R. P. Duin (1999), «Support vector domain description», *Pattern Recognition Letters*, **20**, n° 11-13, pages 1191–1199.

[Tenenbaum et al. 2000] J. B. Tenenbaum, V. de Silva & J. C. Langford (2000), «A global geometric framework for nonlinear dimensionality reduction», *Science*, **290**, pages 2319–2323.

[Thayananthan et al. 2006] A. Thayananthan, R. Navaratnam, B. Stenger, P. H. S. Torr & R. Cipolla (2006), «Multivariate Relevance Vector Machines for tracking», *Proceedings of the European Conference on Computer Vision (ECCV)*.

[Thompson 1917] D. A. W. Thompson (1917), *On growth and form*, Cambridge University Press.

[Thompson et al. 2000] P. Thompson, J. Giedd, R. Woods, D. MacDonald, A. Evans & A. Toga (2000), «Growth patterns in the developing brain detected by using continuum mechanical tensor maps», *Nature*, **404**, pages 190-193.

[Thompson et al. 2004] P. M. Thompson, K. M. Hayashi, G. I. de Zubicaray, A. L. Janke, S. E. Rose, J. Semple, M. S. Hong, D. H. Herman, D. Gravano, D. M.

Doddrell & A. W. Toga (2004), «Mapping hippocampal and ventricular change in Alzheimer disease», *NeuroImage*, **22**, n° 4, pages 1754 - 1766.

[Trouvé 1998] A. Trouvé (1998), «Diffeomorphisms groups and pattern matching in image analysis», *International Journal of Computer Vision*, **28**, n° 3, pages 213–221.

[Turk & Pentland 1991] M. Turk & A. Pentland (1991), «Eigenfaces for recognition», *Journal of Cognitive Neuroscience*, **3**, n° 1, pages 71–86.

[Twining & Marsland 2008] C. Twining & S. Marsland (2008), «Constructing an atlas for the diffeomorphism group of a compact manifold with boundary with application to the analysis of image registrations», *Journal of Computational Applied Mathematics*, **222**, pages 411-428.

[Tzikas et al. 2006] D. Tzikas, A. Likas & N. Galatsanos (2006), «A tutorial on Relevance Vector Machines for regression and classification with applications», *EURASIP newsletter*, **17**, n° 2, pages 4–23.

[Tzikas et al. 2008] D. Tzikas, A. Likas & N. Galatsanos (2008), «The variational approximation for Bayesian inference», *IEEE Signal Processing Magazine*, **25**, n° 6, pages 131-146.

[Ueda & Ghahramani 2002] N. Ueda & Z. Ghahramani (2002), «Bayesian model search for mixture models based on optimizing variational bounds», *Neural Networks*, **15**, n° 10, pages 1223–1241.

[Unnikrishnan et al. 2007] R. Unnikrishnan, C. Pantofaru & M. Hebert (2007), «Toward objective evaluation of image segmentation algorithms», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**, n° 6, pages 929-944.

[Vaillant & Glaunes 2005] M. Vaillant & J. Glaunes (2005), «Surface matching via currents», in *Information Processing in Medical Imaging*, pages 1–5, Springer, 2005.

[Van der Maaten et al. 2009] L. Van der Maaten, E. Postma & H. van den Herik (2009), «Dimensionality reduction : a comparative review», Technical report 2009-005, Tilburg centre for Creative Computing, Tilburg University, Tilburg, The Netherlands.

[Van Leemput et al. 1999] K. Van Leemput, F. Maes, D. Vandermeulen & P. Suetens (1999), «Automated model-based tissue classification of MR images of the brain», *IEEE Transactions on Medical Imaging*, **18**, pages 897-908.

[Van Leemput et al. 2001] K. Van Leemput, F. Maes, D. Vandermeulen, A. Colchester & P. Suetens (2001), «Automated segmentation of multiple sclerosis lesions by model outlier detection», *IEEE Transactions on Medical Imaging*, **20**, pages 677-688.

[Vercauteren et al. 2008] T. Vercauteren, X. Pennec, A. Perchant & N. Ayache (2008), «Symmetric log-domain diffeomorphic registration : a demons-based approach», in *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 5241, pages 754–761.

[Vercauteren et al. 2009] T. Vercauteren, X. Pennec, A. Perchant & N. Ayache (2009), «Diffeomorphic demons : efficient non-parametric image registration», *NeuroImage*, **45**, n° 1, Supplement 1, pages S61 - S72.

[Vereecken et al. 1994] T. Vereecken, O. Vogels & R. Nieuwenhuys (1994), «Neuron loss and shrinkage in the amygdala in Alzheimer's disease», *Neurobiology of Aging*, **15**, n° 1, pages 45 - 54.

[Vik 2004] T. Vik (2004), *Modèles statistiques d'apparence non gaussiens. Application à la création d'un atlas probabiliste de perfusion cérébrale en imagerie médicale*, PhD thesis, University Louis Pasteur - Strasbourg I.

[Vik et al. 2007] T. Vik, F. Heitz & P. Charbonnier (2007), «Robust pose estimation and recognition using non-Gaussian modeling of appearance subspaces», *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **29**, n° 5, pages 901 -905.

[Wachinger & Navab 2010] C. Wachinger & N. Navab (2010), «Manifold learning for multi-modal image registration», in *Proceedings of the British Machine Vision Conference*, pages 1–12, 2010.

[Wachinger et al. 2010] C. Wachinger, M. Yigitsoy & N. Navab (2010), «Manifold learning for image-based breathing gating with application to 4D ultrasound», in *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, volume 6362, pages 26–33, 2010.

[Wechsler 1991] D. Wechsler (1991), *Echelle clinique de mémoire Wechsler-révisée*, Les Editions du Centre de Psychologie Appliquée.

[Weinberger & Saul 2006] K. Q. Weinberger & L. K. Saul (2006), «Unsupervised learning of image manifolds by semidefinite programming», *International Journal of Computer Vision*, **70**, pages 77-90.

[Wells III et al. 1996] W. Wells III, W. Grimson, R. Kikinis & F. Jolesz (1996), «Adaptive segmentation of MRI data», *IEEE Transactions on Medical Imaging*, **15**, pages 429-442.

[West 1993] M. West (1993), «Regionally specific loss of neurons in the aging human hippocampus», *Neurobiology of Aging*, **14**, n° 4, pages 287 - 293.

[West et al. 2000] M. J. West, C. H. Kawas, L. J. Martin & J. C. Troncoso (2000), «The CA1 region of the human hippocampus is a hot spot in Alzheimer's disease», *Annals of the New York Academy of Sciences*, **908**, n° 1, pages 255–259.

[Wolz et al. 2010] R. Wolz, P. Aljabar, J. V. Hajnal, A. Hammers & D. Rueckert (2010), «LEAP : Learning embeddings for atlas propagation», *NeuroImage*, **49**, n° 2, pages 1316 - 1325.

[Wolz et al. 2010] R. Wolz, P. Aljabar, J. Hajnal & D. Rueckert (2010), «Manifold learning for biomarker discovery in MR imaging», in F. Wang, P. Yan, K. Suzuki & D. Shen, editors, *Machine Learning in Medical Imaging*, volume 6357 of *Lecture Notes in Computer Science*, pages 116–123, Springer Berlin / Heidelberg.

[Wright 2005] M. H. Wright (2005), «The interior-point revolution in optimization : history, recent developments, and lasting consequences», *Bulletin of the American Mathematical Society (N.S)*, **42**, pages 39–56.

[Xu &Wunsch II 2005] R. Xu & D. Wunsch II (2005), «Survey of clustering algorithms», *IEEE Transactions on Neural Networks*, **16**, n° 3, pages 645-678.

[Yang et al. 2007] A. Yang, J. Wright, S. Sastry & Y. Ma (2007), «Unsupervised segmentation of natural images via lossy data compression», *Computer Vision and Image Understanding*, **110**, n° 2, pages 212–225.

[Yang et al. 2011] L. Yang, W.-M. Ma & B. Tian (2011), «New multi-class classification method based on the SVDD model», in D. Liu, H. Zhang, M. Polycarpou, C. Alippi & H. He, editors, *Advances in Neural Networks – ISNN 2011*, volume 6676 of *Lecture Notes in Computer Science*, pages 103–112, Springer Berlin / Heidelberg.

[Younes et al. 2009] L. Younes, F. Arrate & M. Miller (2009), «Evolutions equations in computational anatomy», *NeuroImage*, **45**, n° 1, pages S40–S50.

[Yu et al. 2007] P. Yu, P. Grant, Y. Qi, X. Han, F. Segonne, R. Pienaar, E. Busa, J. Pacheco, N. Makris, R. Buckner, P. Golland & B. Fischl (2007), «Cortical surface shape analysis based on spherical wavelets», *IEEE Transactions on Medical Imaging*, **26**, n° 4, pages 582 -597.

[Zabih & Kolmogorov 2004] R. Zabih & V. Kolmogorov (2004), «Spatially coherent clustering using graph cuts», in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 437–444, 2004.

[Zallat & Heinrich 2007] J. Zallat & C. Heinrich (2007), «Polarimetric data reduction : a Bayesian approach», *Optics Express*, **15**, n° 1, pages 83-96.

[Zallat et al. 2004] J. Zallat, C. Collet & Y. Takakura (2004), «Clustering of polarization-encoded images», *Applied Optics*, **43**, n° 2, pages 283–292.

[Zallat et al. 2008] J. Zallat, C. Heinrich & M. Petremand (2008), «A Bayesian approach for polarimetric data reduction : the Mueller imaging case», *Optics Express*, **16**, n° 10, pages 7119-7133.

[Zhang et al. 2006] Q. Zhang, R. Souvenir & R. Pless (2006), «On manifold structure of cardiac MRI data : Application to Segmentation», in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2006.

[Zhou et al. 2008] L. Zhou, R. Hartley, L. Wang, P. Lieby & N. Barnes (2008), «Regularized discriminative direction for shape difference analysis», in *Proceedings of Medical Image Computing and Computer Assisted Intervention (MICCAI)*, New York, USA.

[Zhou et al. 2009] L. Zhou, R. Hartley, L. Wang, P. Lieby & N. Barnes (2009), «Identifying anatomical shape difference by regularized discriminative direction», *IEEE Transactions on Medical Imaging*, **28**, n° 6, pages 937–950.

[Zhou et al. 2011] L. Zhou, Y. Wang, Y. Li, P.-T. Yap & D. Shen (2011), «Hierarchical anatomical brain networks for MCI prediction by partial least square analysis», in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1073–1080, June 2011.

UNIVERSITÉ DE STRASBOURG

Giorgos SFIKAS

Modèles statistiques non linéaires pour l'analyse de formes. Application à l'imagerie cérébrale.

Résumé

Cette thèse a pour objet l'analyse statistique de formes, dans le contexte de l'imagerie médicale. Dans le champ de l'imagerie médicale, l'analyse de formes est utilisée pour décrire la variabilité morphologique de divers organes et tissus. Nous nous focalisons dans cette thèse sur la construction d'un modèle génératif et discriminatif, compact et non-linéaire, adapté à la représentation de formes. Ce modèle est evalué dans le contexte de l'étude d'une population de patients atteints de la maladie d'Alzheimer et d'une population de sujets contrôles sains. Notre intérêt principal ici est l'utilisation du modèle discriminatif pour découvrir les différences morphologiques les plus discriminatives entre une classe de formes donnée et des formes n'appartenant pas à cette classe. L'innovation théorique apportée par notre modèle réside en deux points principaux : premièrement, nous proposons un outil pour extraire la différence discriminative dans le cadre Support Vector Data Description (SVDD) ; deuxièmement, toutes les reconstructions générées sont anatomiquement correctes. Ce dernier point est dû au caractère non-linéaire et compact du modèle, lié à l'hypothèse que les données (les formes) se trouvent sur une variété non-linéaire de dimension faible. Une application de notre modèle à des données médicales réelles montre des résultats cohérents avec les connaisances médicales.

Mots-clés : analyse statistique de formes, apprentissage de variétés, support vector data description, différence discriminative, maladie d'Alzheimer, hippocampe.

Résumé en anglais

This thesis addresses statistical shape analysis, in the context of medical imaging. In the field of medical imaging, shape analysis is used to describe the morphological variability of various organs and tissues. Our focus in this thesis is on the construction of a generative and discriminative, compact and non-linear model, suitable to the representation of shapes. This model is evaluated in the context of the study of a population of Alzheimer's disease patients and a population of healthy controls. Our principal interest here is using the discriminative model to discover morphological differences that are the most characteristic and discriminate best between a given shape class and forms not belonging in that class. The theoretical innovation of our work lies in two principal points : first, we propose a tool to extract discriminative difference in the context of the Support Vector Data description (SVDD) framework ; second, all generated reconstructions are anatomically correct. This latter point is due to the non-linear and compact character of the model, related to the hypothesis that the data (the shapes) lie on a low-dimensional, non-linear manifold. The application of our model on real medical data shows results coherent with well-known findings in related research.

Keywords: statistical shape analysis, manifold learning, support vector data description, discriminative difference, Alzheimer's disease, hippocampus.