## FACULTY OF SCIENCE UNIVERSITY OF COPENHAGEN



## **PhD Thesis**

Jacob Daniel Kirstejn Hansen

## Segmentation of Brains and Rocks from Tomographic Reconstructions

Supervisor: François Bernard Lauze

December, 2018

University of Copenhagen Faculty of Science Department of Computer Science

## Segmentation of Brains and Rocks from Tomographic Reconstructions

Jacob Daniel Kirstejn Hansen

Supervisor: François Lauze

This thesis has been submitted to the PhD School of The Faculty of Science, University of Copenhagen, on December 31, 2018

#### Abstract

Segmentation is an indispensable initial step in image analysis and computer vision. New advanced scanners and large scale imaging facilities have spiked the interest of researchers across fields to investigate the internal structures of objects in a noninvasive manner. However, with new machines come new artefacts and challenges that need to be addressed before subsequent analysis can be conducted.

This thesis presents six novel variational methods for the recovery of segments in tomographic reconstructions. Two primary types of volumetric datasets are used as target application; porous chalk rocks, from X-ray computerised microtomography (X-ray  $\mu$ CT) and rat cranial scans, acquired through magnetic resonance imaging (MRI). Several types of artefacts are addressed, with an emphasis on bias fields that corrupts both acquisition modalities.

The first two chapters of the thesis cope with noise and bias fields by extending the piecewise constant Mumford and Shah functional to utilise local instead of global means. A kernel function governs the locality and weight of the spatial dependency. Regularisation of label fields is achieved using a squared gradient and compared with a total variation variant. Oversmoothing of edges is prevented by integrating an edge sensitive halting function into the framework. Alternating optimisation of the local averages and label fields is used to minimise the energy functionals. Label fields are represented as posteriors by Hidden Markov Measure Field Models (HMMFM), to relax constraints and convexify the problem. Posterior labels additionally model uncertainty due to partial volume effects, owed to limited resolution of the imaging system. These presented methods are tested on both synthetically and experimental datasets of porous media. Numerous quantitative similarity and geophysical measures along with visual inspection demonstrate the usefulness of the methods.

The following three chapters focus primarily on the extraction of brains in bias field rich MRI scans of rodents crania. A pipeline for extracting brain matter, cerebrospinal fluid (CSF), and background serves as the stepping stone for further improvements. The framework utilises previously presented methods in succession as well as some classical approaches to yield promising results. The framework is extended to improve robustness and limit the amount of user interaction by introducing shape priors as regulariser for the label field evolution. The shape priors are inherently invariant to similarity transformations including translation, global scaling and rotations through alignment of shapes. Experiments demonstrate that the updated framework is able to satisfactorily segment challenging cases of rich bias, debris, as well as attached skull and brain matter.

The final chapter recognises the emphasis on precise background segments over foreground in a new approach, based on information theoretic measures for the data fidelity term. A special edge artefact is addressed by a novel halting function for a classical length regulariser. Preliminary experiments of individual energy terms show promising results, but holistic tests need to be performed.

#### Resumé

Segmentering er et uundværligt første skridt i billedanalyse og computer vision. Nye avancerede skannere og stor-skala billeddannelsesfaciliteter har, på tværs af felter, øget forskeres interesse i at undersøge indre strukturer af objekter, på en ikke-invasiv måde. Med større maskiner kommer der dog nye artefakter og udfordringer der skal adresseres, før efterfølgende analyser kan gennemføres.

Denne afhandling præsenterer seks nye variationelle metoder til genskabelse af segmenter i tomografiske rekonstruktioner. To primære typer af volumetrisk datasæt er brugt som fokusområde; porøse kalksten, fra røntgenstråle computermikrotomografi (X-ray  $\mu$ CT) og skanninger af rottekranier, anskaffet gennem magnetisk resonans skanning (MRI). Adskillige typer af artefakter er addresseret, med et fokus på biasfelter der korrumperer begge erhvervelsesmodaliteter.

Afhandlingens to første kapitler håndterer støj og biasfelter ved at udvide den stykvis konstante Mumford og Shah funktional til at benytte lokale, fremfor globale, gennemsnit. Regularisering af labelfelter er opnået ved brug af en kvadratisk gradient og sammenlignet med en total variation udgave. Overudglatning af kanter er forebygget ved integrering af en kantsensitiv stop-funktion i modelkonstruktionen. Skiftevis optimering af de lokale gennemsnit og labelfelterne er brugt til at minimere energy funktionalet. Labelfelterne er repræsenteret som posteriors af Skjulte Markov Målfeltmodeller (HMMFM), for at lempe begrænsningerne og gøre problemet konveks. Posterior labels modellerer yderliger usikkerhed fra delvise volumeneffekter, der skyldes begrænset opløsning af billeddannelsessystemet. Disse præsenterede modeller er testet på både syntetiske og eksperimentielle datasæt af porøse medier. Adskillige kvantitative similaritets- og geofysiske mål sammen med visuel inspektion demonstrerer brugbarheden af metoderne.

De efterfølgende tre kapitler fokuserer primært på ekstraktion af hjerner i biasfeltrige MRI skanninger af gnaveres kranier. En sekventiel model for udtrækning af hjernemateriale, cerebrospinalvæske (CSF) og baggrund bliver brugt som springbræt for yderligere forbedringer. Modelkonstruktionen udnytter tidligere præsenterede metoder i rækkefølge, såvel som nogle klassiske tilgange, til at give lovende resultater. Modelkonstruktionen er udvidet til at forbedre robusthed og begrænse mængden af brugerinteraktion, ved at introducere formpriorer som regularisering af labelfeltudviklingen. Formpriorerne er naturligt invariante overfor similaritetstransformationer som translation, global skalering og rotationer, gennem formtilpasning. Eksperimenter demonstrerer at den opdaterede modelkonstruktion er i stand til tilfredsstillende at segmente udfordrende tilfælde af biasfelter, vragrester, såvel som sammenhængende kranie- og hjernemateriale.

Det sidste kapitel anerkender vigtigheden af præcise baggrundssegmenter over forgrund i en ny tilgang, baseret på informationsteoretiske mål for datatroskabsbetingelsen. En speciel kantartefakt er adresseret af en ny stopfunktion for en klassisk længederegularisering. Indledende eksperimenter af individuelle energibetingelser viser lovende resultater, men holistiske tests skal udføres.

#### Acknowledgements

Doing an PhD is a lifechanging event and an emotional rollercoaster that is difficult to put into words. It has truly been an incredible journey in finding personal and academic limits and pushing them further than I thought imaginable. Naturally, such a journey has enabled me to meet and reconnect with amazing people over the course of the past three and a half years, who have helped shape me and my thesis into what it is today. Enourmous gratitude and acknowledgement is fit for the people who have enabled me to reach this milestone in my life.

I would foremost like to thank my supervisor, François for his extensive help, guidance, and patience through the years, starting from my Master's thesis. His ability to stay level headed in the face of immediate deadlines has been an inspiration to behold. I highly endorse his countless hours at the white- and blackboard cramming appropriate mathematics into my skull. It has truly been incredible to be able to collaborate on research with him. I am really grateful for having this opportunity.

I would like to thank Yiqiu Dong from the Danish Technical University for valuable scientific discussion. Thank you to Julia Huntenburg from Champalimaud Centre for the Unknown, and Sune Darkner from DIKU for collaboration as well as access to rat and mice data respectively. Thank you to Henning for valuable discussions and access to the P-cubed data.

Thank you to Jonas Bardino from the Niels Bohr Institute, for endless correspondences and repeating himself on explaining and setting up cluster computations, as well as integrating final code into the P-Cubed workflow pipeline. I would also like to thank the NanoGeoScience group at the University of Copenhagen for introducing me to the world of rocks, in particular thank you Arash, Stefan, and Dirk.

I acknowledge the funding provided to me by The Innovation Fund and the Department of Computer Science at the University of Copenhagen (DIKU).

A huge thanks also goes out to my fantastic colleagues and friends that I have met during my years as a PhD. In particular to my office-mates for being rubber ducks, spreading light when there seemed to be none, and enabling necessary procrastination (in order of appearance); Leise, Mateusz, Line, Sisi, Julian, Tobias, Wouter, and Mathias. An extended thank you goes out to Raghavendra, Niklas, Malte, Henrik, Kristoffer, Oswin, and Christian for countless social activities, chats, and welcome scientific discussions. The entire (old) image group at DIKU has been the ideal research environment for me and I will miss you all dearly.

Finally, I would like to wholeheartedly thank my parents Ole and Birgith for everything you have done for me. A huge thank you is in order for my brother, Martin for proofreading my thesis and always being there for me. And of course thank you, my beloved girlfriend Fangyue, for bearing with me and keeping me sane in periods of extreme pressure.

## Contents

Al	Abstract	iv	V
Re	Resumé	•	v
A	Acknowledgements	v	i
Li	List of Figures	i	x
Li	List of Tables	x	i
1	1 Introduction1.1 Segmentation methods in the literature1.2 Rocks and the P-Cubed project1.3 Rat brains1.4 Outline of the thesis	· · · · · · · · · · · · · · · · · · ·	<b>1</b> 1 7 9
2	<ul> <li>2 Local Mean Multiphase Segmentation</li> <li>2.1 Introduction</li> <li>2.2 Derivation of the model</li> <li>2.3 Optimisation</li> <li>2.3.1 Update for c.</li> <li>2.3.2 Update for v.</li> <li>2.4 Experimental Validation</li> <li>2.4.1 Synthetic datasets</li> <li>2.4.2 Performance criteria</li> <li>2.4.3 Methods</li> <li>2.4.4 Results</li> <li>2.5 Conclusion</li> </ul>	$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	5 3 9 0 0 2 2 2 4 5 7
3	<ul> <li>3 Local Mean Geodesic Active Regions</li> <li>3.1 Introduction</li> <li>3.2 Derivation of the model</li> <li>3.2.1 Weighted Perimeters and Relaxation.</li> </ul>	3 	1 2 5 5

		3.2.2 Non Stationarity and Proposed Models	36
	3.3	Optimisation	37
		3.3.1 Update for c	37
		3.3.2 Update for v	37
	3.4	Experimental Validation	38
		3.4.1 Performance of Segmentation	39
		3.4.2 Methods	40
		3.4.3 Results	41
	3.5	Conclusion	44
4	Rat	Brain Extraction and Segmentation	47
	4.1	Introduction	48
		4.1.1 Animals	48
		4.1.2 MRI	49
	4.2	Methods	50
		4.2.1 Bias field correction	50
		4.2.2 2-class segmentation	50
		4.2.3 GAC brain extraction	50
		4.2.4 3-class segmentation	50
	4.3	Results	51
	4.4	Discussion and Conclusion	51
5	Aut	omatic Brain Segmentation using Shape Priors	57
	5.1	Introduction	EO
	0.1		30
	0.1	5.1.1 Animals	58 59
	0.1	5.1.1       Animals	58 59 60
	5.2	5.1.1       Animals	58 59 60 60
	5.2	5.1.1       Animals	58 59 60 60 60
	5.2 5.3	5.1.1       Animals         5.1.2       MRI         Methods	58 59 60 60 60 61
	5.2 5.3 5.4	5.1.1Animals5.1.2MRIMethods5.2.1Brain shape prior segmentationResultsDiscussion and Conclusion	58 59 60 60 60 61 61
6	5.2 5.3 5.4 Inva	5.1.1       Animals         5.1.2       MRI         Methods	<ul> <li>58</li> <li>59</li> <li>60</li> <li>60</li> <li>60</li> <li>61</li> <li>61</li> <li>65</li> </ul>
6	5.2 5.3 5.4 Inva 6.1	5.1.1       Animals         5.1.2       MRI         Methods	<ul> <li>58</li> <li>59</li> <li>60</li> <li>60</li> <li>60</li> <li>61</li> <li>61</li> <li>65</li> <li>66</li> </ul>
6	5.2 5.3 5.4 Inv: 6.1 6.2	5.1.1       Animals         5.1.2       MRI         Methods	<ul> <li>58</li> <li>59</li> <li>60</li> <li>60</li> <li>60</li> <li>61</li> <li>61</li> <li>65</li> <li>66</li> <li>68</li> </ul>
6	5.2 5.3 5.4 <b>Inv</b> 6.1 6.2 6.3	5.1.1       Animals         5.1.2       MRI         Methods	<ul> <li>58</li> <li>59</li> <li>60</li> <li>60</li> <li>60</li> <li>61</li> <li>61</li> <li>65</li> <li>66</li> <li>68</li> <li>71</li> </ul>
6	5.2 5.3 5.4 Inva 6.1 6.2 6.3 6.4	5.1.1       Animals         5.1.2       MRI         Methods	<ul> <li>58</li> <li>59</li> <li>60</li> <li>60</li> <li>61</li> <li>61</li> <li>65</li> <li>66</li> <li>68</li> <li>71</li> <li>73</li> </ul>
6	<ul> <li>5.2</li> <li>5.3</li> <li>5.4</li> <li>Inva</li> <li>6.1</li> <li>6.2</li> <li>6.3</li> <li>6.4</li> <li>6.5</li> </ul>	5.1.1       Animals         5.1.2       MRI         Methods	<ul> <li>58</li> <li>59</li> <li>60</li> <li>60</li> <li>61</li> <li>61</li> <li>61</li> <li>65</li> <li>66</li> <li>68</li> <li>71</li> <li>73</li> <li>77</li> </ul>
6	5.2 5.3 5.4 <b>Inv</b> 6.1 6.2 6.3 6.4 6.5 <b>Seg</b>	5.1.1       Animals         5.1.2       MRI         Methods	<ul> <li>58</li> <li>59</li> <li>60</li> <li>60</li> <li>60</li> <li>61</li> <li>61</li> <li>65</li> <li>66</li> <li>68</li> <li>71</li> <li>73</li> <li>77</li> <li>81</li> </ul>
6	<ul> <li>5.2</li> <li>5.3</li> <li>5.4</li> <li>Inva</li> <li>6.1</li> <li>6.2</li> <li>6.3</li> <li>6.4</li> <li>6.5</li> <li>Seg</li> <li>7.1</li> </ul>	5.1.1       Animals         5.1.2       MRI         Methods	<ul> <li>58</li> <li>59</li> <li>60</li> <li>60</li> <li>61</li> <li>61</li> <li>65</li> <li>66</li> <li>68</li> <li>71</li> <li>73</li> <li>77</li> <li>81</li> <li>82</li> </ul>
6	<ul> <li>5.2</li> <li>5.3</li> <li>5.4</li> <li>Inv:</li> <li>6.1</li> <li>6.2</li> <li>6.3</li> <li>6.4</li> <li>6.5</li> <li>Seg</li> <li>7.1</li> <li>7.2</li> </ul>	5.1.1       Animals         5.1.2       MRI         Methods	<ul> <li>58</li> <li>59</li> <li>60</li> <li>60</li> <li>60</li> <li>61</li> <li>61</li> <li>65</li> <li>66</li> <li>68</li> <li>71</li> <li>73</li> <li>77</li> <li>81</li> <li>82</li> <li>84</li> </ul>
6	<ul> <li>5.2</li> <li>5.3</li> <li>5.4</li> <li>Inva</li> <li>6.1</li> <li>6.2</li> <li>6.3</li> <li>6.4</li> <li>6.5</li> <li>Seg</li> <li>7.1</li> <li>7.2</li> </ul>	5.1.1       Animals         5.1.2       MRI         Methods	<ul> <li>58</li> <li>59</li> <li>60</li> <li>60</li> <li>60</li> <li>61</li> <li>61</li> <li>65</li> <li>66</li> <li>68</li> <li>71</li> <li>73</li> <li>77</li> <li>81</li> <li>82</li> <li>84</li> <li>85</li> </ul>
6	<ul> <li>5.2</li> <li>5.3</li> <li>5.4</li> <li>Inva</li> <li>6.1</li> <li>6.2</li> <li>6.3</li> <li>6.4</li> <li>6.5</li> <li>Seg</li> <li>7.1</li> <li>7.2</li> <li>7.3</li> </ul>	5.1.1       Animals         5.1.2       MRI         Methods	58 59 60 60 61 61 65 66 68 71 73 77 <b>81</b> 82 84 85 86

		7.3.2 7.3.3 7.3.4	Entropy derivative	88 89 90
		7.3.5	Full descent direction	91
	7.4	Individ	dual experiments	91
	7.5	Conclu	asion	92
8	Disc	ussion		95
9	Con	clusion	L Contraction of the second	101
Bil	Bibliography 105			105

# List of Figures

1.1 1.2 1.3 1.4	Overview of the P-cubed pipeline Examples of P <sup>3</sup> reconstructed raw data Ground truth sketch and reconstruction of rat brain Example of bias in rat brains and its correction	8 9 10 11
<ol> <li>2.1</li> <li>2.2</li> <li>2.3</li> <li>2.4</li> <li>2.5</li> </ol>	Typical example tomograms for our applicationSynthetically generated data and real data segmentation resultSegmentation of synthetically generated dataSegmentation of a real full 3-class datasetSegmentation of a real full 3-class dataset, zoomed	17 23 24 28 29
<ol> <li>3.1</li> <li>3.2</li> <li>3.3</li> <li>3.4</li> <li>3.5</li> </ol>	Typical example tomograms for our applicationSynthetically generated dataSegmentation result for $SYN250_{global}^3$ Soft label segmentation of $SYN250_{local}^3$ , using the Tikhonov modelComparison segmentation of a real 3-class dataset	33 39 40 40 43
<ol> <li>4.1</li> <li>4.2</li> <li>4.3</li> <li>4.4</li> <li>4.5</li> </ol>	Illustration of the proposed framework	49 52 53 54 55
5.1 5.2 5.3 5.4 5.5	Illustration of the proposed pipelineBrain extraction comparisonStep 1-3 of the shape prior frameworkStep 4 of the pipelineComparison with AFNI's 3D skullstrip and the proposed method	59 61 62 63 64
6.1 6.2 6.3	Example slice of a rat cranium and 3D segmentation2D experiment with synthetic data3D experiment with synthetic data	67 74 75

## LIST OF FIGURES

6.4 6.5 6.6 6.7	Training shape masks for 2D experiment of synthetic data2D invariance experiment with synthetic data3D Segmentation evolution steps of real dataFinal 3D segmentation results of our proposed model	76 77 78 79
7.1 7.2	Edge artefacts and intensity profile	85 93
7.3 7.4	Individual tests of g function choices 2	94 94

## List of Tables

1.1	Publication status of manuscripts	13
2.1	Best parameter choices	25
2.2	Quality measures for the segmentation of SYN250 <sub>global</sub>	25
2.3	Quality measures for the segmentation of SYN250 <sub>local</sub>	26
2.4	Petrophysical measures for the segmentation of $SYN250_{global}$ .	26
2.5	Petrophysical measures for the segmentation of $SYN250_{local}$	27
3.1	Best parameter choices	41
3.2	Quality measures for the segmentation of SYN250 <sub>global</sub>	42
3.3	Quality measures for the segmentation of SYN250 <sub>local</sub>	42
3.4	Petrophysical measures for the segmentation of $SYN250_{global}$	44
3.5	Petrophysical measures for the segmentation of $SYN250_{local}$	44

## 1 Introduction

Image segmentation is of paramount importance for analysis across a multitude of fields. The process is characterised as the separation of data into multiple meaningful subsets or segments, enabling and simplifying subsequent data processing. Meaningful here depends on the specific goal of the segmentation procedure, but the general criteria is that data with the same segment label share certain characteristica or features. These attributes could be linked to individual element values, or overall modelled properties, such as neighbourhood information or statistical measures.

While the human visual system is able to process natural images for object detection at incredible precision and accuracy, with speeds below fractions of a second[Thorpe et al., 1996], we are still unsure which exact mechanisms are used in the process[DiCarlo et al., 2012]. With access to unfathomable amounts of data, humans and in particular machines and programs would undoubtedly benefit immensely from discovering which computations goes into this process. We have yet to uncover these algorithms, but incremental steps in improving and inventing new segmentation models will definitely push us in the right direction.

This thesis concerns itself with exactly that; to contribute to the field of variational segmentation methods for 2D and 3D images, with Rocks and Brains as the target application. Before we dive into the content of each chapter, a brief introduction of the literature on general segmentation methods is in order.

### **1.1** Segmentation methods in the literature

The list of various image segmentation methods in the literature is overwhelmingly long. In this brief literature review we have therefore chosen to omit the bulk of the methods, but provide the reader with an overview of some influential types.

With more than 30.000 citations, the work and influence of [Otsu, 1979] is difficult to overstate. Despite being almost 40 years old, the method is

still videly used today for its simplicity and excellent intermediate results. In this work we use Otsu's method primarily as initialisation of our more complex methods. The method works by minimising the intra-class variance and shows that it is equivalent to maximising the inter-class variance. It then searches for the threshold that fulfils this criteria. In the work of [Liu and Yu, 2009], they proved that the energy functional of Otsu is equivalent to that of k-means, in the multiphase setting. So while k-means is faster, because it does not have to compute histograms, Otsu makes use of an exhaustive search strategy, whereas k-means can end up in local minima. Another significant work in this context is that of [Lloyd, 1982], which proposed a k-means like energy computed on voronoi diagrams. These methods are particularly interesting for this thesis, as the majority of segmentation methods within the field of X-ray computed tomography of porous media are based on global thresholding [Iassonov et al., 2009]. For an excellent and more extensive review of thresholding methods in the literature, we refer to [Sezgin and Sankur, 2004], where they surveyed, implemented, and evaluated 40 two-phase thresholding methods.

In a related subfield, one has clustering approaches akin to the k-means method. These major categories can be divided into hierarchical clustering, density and distribution based approaches, as well as graph, and neural network based methods, which will be elaborated on in the following. There are additionally various variants of these types, based on strict and fuzzy labeling of pixels as well as different measures within the aforementioned categories. For a more in-depth discussion on general clustering methods for data partitioning, we refer to [Rokach and Maimon, 2005].

Other types of segmentation methods are those explicitly using graphs to represent images. This representation lends itself easily to images as both are discrete and simple mathematical representations. Additionally, graph theory is a well-developed field, so previous methods extends directly to images and proofs are readily available. The most prevalent incorporation of graph models for image modelling and segmentation as Markov random fields (MRF), was introduced in [Geman and Geman, 1984]. Among the most popular graph based methods for image segmentation are those based on graph-cuts[Boykov et al., 2001, Kolmogorov and Zabih, 2004, Rother et al., 2004], i.e. partitioning of the graph, such that the weights on edges between segments are small and high within segments - see [Yi and Moon, 2012] for an overview. A most influential method that introduced normalised cut is [Shi and Malik, 2000], which solved a generalised eigenvalue problem. Other methods sought to analyse spectral partitioning, which is a relaxed version of the NP-hard graph-cut minimisation, through studying eigenvectors and the associated graph laplacian and other graph matrix variants, such as the adjacency matrix. [Ng et al., 2002] more theoretically analysed the conditions

on which spectral graph clustering is expected to perform well, using tools from matrix pertubation theory. An excellent introduction to the field of spectral graph clustering, which uses the graph laplacian as the main tool, is additionally provided by [Von Luxburg, 2007]. Other noteworthy methods, that utilises the graph representation of images, are those that utilise random walks [Grady, 2006]. These types work by manually selecting seed labels for the various segments and then initialise random walkers at each unlabelled pixel. The pixels are then assigned the labels of the the seeds they are most likely to encounter first on their random walk. Additional higher order structural representations of graphs, such as minimum spanning trees have also been used in the literature [Zahn, 1970], for speeding up the segmentation of trees.

Machine Learning is an unavoidable mention in these days and the field of trainable segmentation models is vast and evergrowing. The subfield of machine learning that has gained a lot of attention in recent years is deep learning and specific architectures such as deep neural networks have been applied in numerous fields. The basic building blocks of an artificial neural network consist of the structure of multiple processors, so called neurons, that produce a sequence of activations based on connectivity, weights, and observation generating environments. Given a real number activation signal, a neuron computes some predefined nonlinear function of the sum of these ingoing signals. The strength of this new signal is then magnified or decreased, based on the weight of the edges connecting the neurons. The typical structure of artificial neural networks is formed by layers of neurons, that perform various different transformations of input signals. A threshold is usually used to determined whether or not a signal is passed on to the next layer. When we are talking about deep neural networks, it therefore refers to the number of layers. Various different versions and traversals of these structures are then used for different tasks, but common for them all is that they are based on initial learning or training steps. Given a sufficient number of pre-labeled data and a task, the network is able to learn specific optimal parameters or weights that solve the task, based on some cost function.

Deep learning have in particular been used with promising results in automatic speech recognition [Hinton et al., 2012, Graves et al., 2013, Dahl et al., 2012], AI for game theory [Silver et al., 2016, Mnih et al., 2013], natural language processing [Collobert and Weston, 2008, Collobert et al., 2011] as well as computer vision tasks like image classification [Cireşan et al., 2012, Krizhevsky et al., 2012], object detection [Szegedy et al., 2013, Erhan et al., 2014], semantic segmentation [Long et al., 2015], and many more. Generally deep learning relies on supervised learning practices, although unsupervised approaches exist and typically use restricted Boltzmann machines [Nair and Hinton, 2010, Hinton, 2007] or autoencoders [Hinton and Salakhutdinov, 2006, Poultney et al., 2007, Vincent et al., 2008, Masci et al., 2011] to learn features, probability distributions over inputs, or perform dimensionality reduction.

Two recent interesting approaches in machine learning are reinforcement learning [Watkins, 1989, Watkins and Dayan, 1992] and the so called U-net [Ronneberger et al., 2015, Cicek et al., 2016, Dong et al., 2017]. Reinforcement learning introduces the idea of delayed reward, through state transitions of an autonomous entity, called an agent. In order to reinforce optimal strategy, a reward can be given based on the expected value of the total reward of all previous steps to the current. The strategy is then continuously adjusted accordingly. A new step, through exploration of the solution space, while exploiting current strategy knowledge, is then taken. U-nets are an improvement of the fully convolutional network (FCN) presented in [Long et al., 2015] and rely in the same way on down- and upsampling to capure contextual information and recover spatial information respectively. U-nets are also symmetric in the number of down and upsampling layers and differ from FCNs in that the skip connections apply concatenation operations instead of summing, to preserve local information in the upsampling step. Skip connections are traditionally used to bypass at least one layer, to preserve fine grained spatial information that would otherwise have been lost in the 'pooling' or downsampling layers. U-nets, while still in the very early stages of development look promising, as they perform well even with very few training dataset, compared to deep learning approaches.

A number of classical non-variational methods have been videly used, such as the mean-shift algorithm [Cheng, 1995], originally formulated in [Fukunaga and Hostetler, 1975] to locate maxima of density functions. Segmentation using mean-shift works by converting data into feature space and initialising nonoverlapping search windows over the space. Within each window the mean is computed and the center of the windows are relocated or shifted to the respective means. New means and shifts are computed until convergence and windows sharing centers are merged. Each feature is then labelled according to the mean of the potentially merged window originally covering it. The number of classes is based on window size and initial placement.

Another example of non-variational methods is the introduction of an initial oversegmentation, resulting in so called "superpixels"[Ren and Malik, 2003] and it has been used as an important preprocessing step and main algorithm in a number of segmentation algorithms [Liu et al., 2011, Achanta et al., 2012, Shen et al., 2014, Veksler et al., 2010]. The introduction of superpixels result in easier and faster feature extraction on regions that are more meaningful than regular pixels. Superpixels can be based on a variety of cues, like intensity level, neighbourhood continuation, similarity, edges, texture information, or pixel proximity.

A final example of non-variational methods are those based on the wa-

tershed transform, originally described in [Strahler, 1957]. The idea of using watershed by flooding, for segmentation, was described in [Beucher and Lantuéj, 1979] and popularised in [Meyer, 1994]. The algorithm by Meyer views the gray-level values of an image as a topographic relief, such that higher intensity values correspond to hills and lower ones to catchment bassins. Flooding the bassins will then result in watershed lines or ridges, that make up the edges of each respective class. A number of alterations and improvements have since been made to the classical algorithms, one example is the priority-flood method [Barnes et al., 2014] that proposed to initialise flooding at the edges and flood inwards, to avoid depressions or dams.

The variational setting is able to model problems in an infinite dimensional way. Variational optimisation is therefore able to accomodate very general problems, by accounting for all possible functions by use of the calculus of variations to solve the associated Euler-Legrange equations. Additionally, because the deriviation of models is often done in a continuous setting, which is very well understood, the bias of choosing correct discretisation representations is dependant on the end user and specific problem, not the model itself. One of the downsides of solving variational problems is the necessary computation power, but recent hardware advances and GPU optimised methods have popularised these techniques. Some classical examples of variational methods include [Geman and Geman, 1984] for introducing a Bayesian paradigm by use of Markov Random Fields for image analysis. The Active Contours model or 'snakes' [Kass et al., 1988] proposed three energy terms for 1) elasticity, based on a smoothness measure and continuity of the contour, 2) an edge based energy to adapt the fitting of the contour, whereas the elasticity energy controlled the deformation of the evolving contour, and 3) a constraint energy which is typically the initial control points. [Cohen, 1991] further improved on the model with the introduction of the classical "balloon" force, to further push the contour towards edges. This enables the original snakes model to converge to an optimal solution, even when the initial curve is not close to the solution. It further is only stopped by strong edges, but passes over weaker edges. Many variants of the snakes model have been introduced over the years, some of these include Geodesic Active Contours (GAC) [Caselles et al., 1997], which seeks to constrain the curve evolution by introducing a halting function on the edge detection energy term, such that stronger edges result in smaller energies. This addition makes the model especially adapt at following strong edges, even in the pressence of other strong forces. A similar approach was used in [Perona and Malik, 1990] for recovering of image segments and edge detection, in the presence of noise. Here the laplacian-based smoothing was constrained by a halting function, to only encourage smoothing of low strength edges, thus preserving significant image features. For an extensive overview of anisotropic diffusion in the field of image processing, we refer to [Weickert, 1998]. Additional statistical shape information have also been incorporated in the classical GAC method, by guiding the deformable shape evolution according to a probability distribution over the variances of a training set of shapes [Leventon et al., 2002] and using mean shape statistics to compute characteristic deformation modes to reveal shape variability [Charpiat et al., 2007]. Another alteration of the snakes model is Diffusion snakes [Cremers et al., 2002] that deals with occlusions, cluttering, and overall noise by the inclusion of statistical shape knowledge in an altered Mumford-Shah functional [Mumford and Shah, 1989].

The work of Mumford and Shah and its cartoon limit, the picewise constant Mumford-Shah model, has been the starting point of numerous recent models and the focus of optimisation [Pock et al., 2009, Chambolle and Pock, 2011, Chambolle et al., 2012]. The former works also deal with the optimisation of the classical total variation regularisation, also known as the ROF model as it was pioneered by Rudin, Osher, and Fatemi [Rudin et al., 1992]. The Mumford-Shah functional consist of three energy terms; 1) a data fidelity term, that makes sure that segments are close to their respective segment means, 2) a smoothness term to guarantee that segment variations are smooth within regions, but not on the discontinuity set, and 3) a length penalisation of the discontinuity set. The cartoon limit is then introducing the assumption that regions are piecewise constant, meaning that the second term disappears, by letting the weight of the second term go to infinity. If one assumes two regions, then a discrete version of the Mumford-Shah functional is the Ising model [Lenz, 1920, Ising, 1925], which was later generalised and studied in works such as [Heisenberg, 1928, Potts, 1952] all of which were influential in the field of ferromagnetism and statistics in physics. One of the popular models inspired by the Mumford-Shah functional is the two-phase Chan and Vese model [Chan and Vese, 2001], using the level-set method [Osher and Sethian, 1988] to represent the evolving contour, its extension to multiphases [Vese and Chan, 2002], and their many variants.

This concludes the brief literature review. In Sections 1.2 and 1.3 we introduce the various projects that I have been involved in. We highlight key challenges and provide motivation for why these problems are of particular interest to solve. The two sections cover two very different types of datasets, namely rocks and rodent brains from two different modalities. It is however clear from our results that methods developed from at least one area are general enough, that they are applicable in the other.

## 1.2 Rocks and the P-Cubed project

The P-Cubed (styled P<sup>3</sup>) project was a collaboration between The eScience group<sup>1</sup>, The NanoGeoScience group<sup>2</sup>, and The Image Group<sup>3</sup>, which was sponsored by The Innovation Fund DK, Maersk Oil & Gas, and The University of Copenhagen. P<sup>3</sup> is an abbreviation for Predicting Petrophysical Parameters and the project was established to build an automated data processing pipeline, that ultimatively would turn physical rock samples into feasible geophysical measures. These measures would be used both for commercial interest such as extraction of natural resources, but also give valuable insight into the internal microstructures of chalk rocks. The developed methods would in turn be used for different kinds of data, such as sea urchin shells to analyse efficient light but strong structural designs.

The overall P<sup>3</sup> pipeline can be seen in Figure 1.1. Each of the boxes represents a project in itself, responsible for taking the input from the previous step, process it, and sending it on to the next. A chalk rock sample would initially be dug up from the North Sea area, several kilometers down, then scanned in a synchrotron to produce the initial tomographic projection that make up the sinogram. After tomographic reconstruction, one is left with the raw noisy image volume to be denoised and segmented. This was the task assigned to us and we will present the proposed methods in Chapters 2 and 3 and as ongoing work in Chapter 7, where we deal with some particular observed artifacts. Following that, the workflow pipeline would then pick up the output, mesh it and simulate fluid flows to extract the parameters of interest.

The collection of data available were roughly 4.2 TB of raw reconstructed data, representing different acquisition depths and porosities, which was only a fraction of the entire data intended to be processed. The requirement of the models to be reliable and automated were essential, especially considering that one full dataset was 4 GB. Additionally, none of the datasets were professionally anotated, in part due to the sheer amount of data to go through in a single dataset, leaving out newer machine learning approaches, such as deep neural networks. A typical example of chalk datasets and its corresponding segmentation can be seen in Figure 1.2. Furthermore, X-ray microtomography is well known to harbour a wealth of artefacts, which explains an entire project dedicated to denoising and removing artefacts in Figure 1.1 - even anotaters would have a difficult time because of the artefacts. These artefacts, aside from noise and blurring, cover: ringing artefacts resulting from variation in detector element sensitivity, various streaking artefacts due to a reduced

<sup>&</sup>lt;sup>1</sup>Research group, under the Niels Bohr Institute, at The University of Copenhagen.

<sup>&</sup>lt;sup>2</sup>Research group, under the Nano-Science Center, at The University of Copenhagen.

<sup>&</sup>lt;sup>3</sup>Former research group under The Department of Computer Science, at The University of Copenhagen - now known as the IMAGE group

#### 1. INTRODUCTION



Figure 1.1: Overview of the P-cubed workflow pipeline

number of projections, double edges caused by slight motion of the sample, partial volume effect caused by averaging due to resolution, bias fields due to beam hardening caused by polychromatic radiation, and many more. We refer to [Davis and Elliott, 2006] and [Hansen, 2015] for a detailed description of X-ray micro-CT artefacts and for various reconstruction caused artefacts, we refer to [Hsieh, 2009] for an in-depth list and characterisation.

When the P<sup>3</sup> project launched, the most successful segmentation results were achived using Otsu [Otsu, 1979], watershed based segmentation using flood-ing [Meyer, 1994], and dual filtering [Müter et al., 2012]. We therefore generally compare with these classical but simple methods.

In Chapters 2 and 3 we target specifically the bias field artefact, as it is quite dominant and locally - as well as globally - observable in our datasets. The methods are furthermore robust to strong noise, while preserving valuable small edge details, that can have large consequences for subsequent analysis. The models are based on the classical Mumford-Shah functional and deal with the bias fields implicitely, instead of trying to explicitely model it.

Chapter 7 proposes a method fitted to deal with an artefact that in certain ways mimics that of beam hardening. Whereas we have generally proposed Mumford-Shah based segmentation methods, this approach relies on information theoretic measures. Specifically one of the goals of this idea is to cut down on the number of free parameters, particularly observing that often only the background segment is of interest in further analysis. The observed artefact is affecting material edges and is dealt with by explicitly modelling the intensity profile of these edges.



**Figure 1.2:** Examples of typical reconstructed raw  $P^3$  datasets. Surface mesh of a segmented dataset crop, dimensions =  $(500 \times 500 \times 500)$ , original dataset dimension =  $(1304 \times 1304 \times 1012)$  (a). A 3D slice representation of the same dataset, in its raw version (b).

## 1.3 Rat brains

The study of lab rodents, and their brains in particular, has been of the utmost importance in drug development, medicine, and biology research. One of the main advantages of mice in particular mice, is that their genome is very similar (99% [Vandamme, 2014]) to the human genome and due to their small size and cost, they are efficient to handle, transport, and scale to large cohort studies. Typical studies are those of obtaining functional and safe dosage information of drugs, to increase human patient safety, before clinical studies are initiated [Vandamme, 2014]. Another very important factor is that, with the advert of genetical engineering, scientists can breed so called transgenic mice, that are genetically altered mice whose genes mimic the human genomes susceptible to disease. In layman's terms: scientists are able to "turn on and off" specific genes depending on the study. This approach also makes studies highly reproducible and robust. Additionally, in the period of 2005-2007, rats and mice ammounted to 93% (of the nine most commonly used species) of all neurodevelopmental research article model mammals. Of these 93%, rats ammount to 54% and mice 39% [Clancy et al., 2007]. The general shape of the rat brain, seen from above, can be found in Figure 1.3a, where the different regions are also highlighted. Typically, brain extraction software will however obtain a more convex shape, due to exterior cerebrospinal fluid (CSF) being difficult to deal with in general between region transitions, but in particular in the gaps between the orange and teal regions, as well as the teal and green regions.

#### 1. INTRODUCTION



**Figure 1.3:** Sketch of a typical rat brain seen top-down, major brain regions are divided by different colors (a). Example of a reconstructed raw rat brain, notice also the short distance from brain to skull and accompanied debris, which complicate segmentation (b).

The work on rodent brains presented in this thesis, has not been associated with any project for our part. The work was initiated by a colleague who was collaborating with clinical scientists, but was having a difficult time segmenting the brains, due to strong bias fields, observed in Figure 1.4. Having previously worked on accounting for bias fields in X-ray CT imaging, we attempted our method on this new MRI data. Having no prior experience with the field of medical image analysis, nor any knowledge of the current best results, we sent an initial conference abstract to the ISMRM (International Society for Magnetic Resonance in Medicine). At the conference, the abstract presentation sparked a lot of interest that gave us not only ideas for improvement and insight in the data from a clinical perspective, but has also set up future collaborations. It is therefore planned that Chapters 5 and 6 will be extended and sent to the journal of Magnetic Resonance in Medicine (MRM) and the software made available.

The initial steps in rat brain analysis is computerised brain extraction from the rat cranium. Unfortunately, at least in this case, clinical research groups tend to use third party software for this step, which is often so general application based, that it fails to account for harder artefacts. Bias fields are one such artefact, often encountered in MRI of rodents, which has been studied immensely in the MRI community. Even though the literature on bias field correction algorithms is vast, in particular for human brains, it is not considered a completely solved problem[Vovk et al., 2007]. Bias fields are particularly problematic as many segmentation methods rely on relatively homogeneous intensity distributions in the different classes. Local CSF and brain matter consequently look so similar, based on their intensity ranges in



**Figure 1.4:** Example of a raw rat brain slice, affected by extreme levels of bias fields (a) and its bias field corrected version (b).

the face of bias fields, that they get misclassified. A dataset effected by heavy bias fields can be found in Figure 1.4. Notice that classical bias field correction techniques, here an opensource implementation of [Tustison et al., 2010], are unable to fully restore the original image. There are still observable dark smudges in the center of the brain and quite bright regions near the edges.

Another complication for segmentation is when the brain and skull are too close to each other, mainly because of the resolution of the image, but debris also play a significant role, which is also observable from the brain in Figure 1.3b. Motivated by these observations, we propose methods in Chapters 4 to 6 to handle these two major issues; Bias fields and closely positioned brain and skulls, with debris.

### **1.4** Outline of the thesis

This thesis is structured so that each chapter constitutes a paper, either as published or in review, with the exception of Chapter 7 that presents an ongoing work, to be published in the future. The only alterations to the orignal publications are the size of figures as well as a few typos or rewordings to fit the margin layout of the thesis. No contentual changes has been made to the papers.

The overall goal of this thesis was to be able to handle various artefacts that presented themselves in tomographic datasets such as porous media and brain scans. In the face of these artefacts, we aimed to be able to segment meaninful regions that would be used for postprocessing and analysis by a third party. It was therefore important not to bias our solutions unnecessarily, by making too many choices in regards to the output. Preferably, albeit not imperative, the methods should be able to generalise well to other datasets of different origin and modality.

In Chapter 2, we present our first method that deals with the observed bias field artefact. In that pursuit, the model attempts to implicitely deal with the bias, rather than explicitely trying to model it. In doing so, we don't enforce too many assumptions on the nature of the bias field, other than at small enough locality it is negligible. The model is based on the Mumford-Shah functional and introduces and compares both a squared and TV regularisation, for this setting.

Chapter 3 further extends the work of Chapter 2. The major goal of this chapter is to address the oversmoothing of edges, as a result of the quadratic regulariser. By introducing a classical halting function based on edge detection, in our Chambolle-Pock optimisation setting, we achieve better results and arrive at a richer more general formulation. The usefullness of the method is demonstrated on experimental, as well as the same synthetic data as Chapter 2, for comparison.

The method presented in Chapter 4 introduces the second dataset this thesis concerns itself with, MRI scans of rat crania. Inspired by the promising results of Chapters 2 and 3, we apply our previous method to bias field corrected rat crania scans, in a four step pipeline. Through a serious of steps of our pipeline, we demonstrate that the proposed framework is able to automatically segment brain matter, cerebrospinal fluid, and background satisfactorily.

One of the shortcomings of the method in Chapter 4 was the robustness to particular situations when the brain matter and skull would intersect heavily or large debris would be present in the gap. The Geodesic Active Contour based skull extraction step would oversegment the brain matter and include large parts of the skull. Therefore, in Chapter 5 we attempt to improve on this drawback, by introducing a powerful similarity invariant shape prior. We demonstrate that the inclusion of this shape prior in the framework enables us to deal with the difficult situations satisfactorily. Additionally, parameter choice robustness is also achieved.

Chapter 6 introduces the theoretical framework for the extension presented in Chapter 5. The chapter presents a general shape prior framework for robust and fully automated object segmentation. The method achieves this by making the shape priors intrinsically invariant to translation, rotation, and global scaling by alignment. This means that transformation parameters do not need to individually optimised for.

An ongoing work is presented in Chapter 7 dealing specifically with the problem of porous media segmentation. The chapter analysis a particularly challenging edge type akin to beam hardening and constructs a powerful novel edge detector, that is able to identify the target edges, as well as smaller normal edges. Rooted in the observation that the background segment is of particular interest in geophysical simulations, we propose an information

Title	Status	
Local Mean Multiphase Segmentation with HMMF Models.	Accepted as [Hansen and Lauze, 2017].	
Multiphase Local Mean Geodesic Active Regions.	Accepted as [Hansen and Lauze, 2018].	
Brain Extraction and Segmentation Framework for Bias Field Rich Cranial MRI Scans of Rats.	Accepted as [Hansen et al., 2018].	
Automatic Brain Segmentation Framework for Bias Field Rich Cranial MRI Scans of Rats and Mice via Similarity Invariant Shape Priors.	In review at ISMRM.	
Segmentation of 2D and 3D objects with intrinsically similarity invariant shape regularisers.	In review at SSVM.	
Information Theory-Driven Two-Phase Segmentation with Edge Penalty.	In preparation.	

Table 1.1: Publication status for the manuscripts included in this thesis.

theoretic based segmentation method, that divides a potentially multiphased problem into a simply behaving background and complex foreground. Experimental validation has been performed on the individual data and regularisation terms and the intermediate results are promising.

An overview of all the manuscripts used for this thesis and their publication status can be found in Table 1.1

## 2 Local Mean Multiphase Segmentation

The work presented in this chapter is based on [Hansen and Lauze, 2017]<sup>†</sup>.

<sup>†</sup> Hansen, J. D. K. and Lauze, F. (2017). Local Mean Multiphase Segmentation with HMMF Models. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 396–407. Springer.

## Local Mean Multiphase Segmentation with HMMF Models

J. D. K. Hansen, F. Lauze, {jdkh,francois@di.ku.dk}

Department of Computer Science, University of Copenhagen

Abstract. This paper presents two similar multiphase segmentation methods for recovery of segments in complex weakly structured images, with local and global bias fields, because they can occur in some X-ray CT imaging modalities. Derived from the Mumford-Shah functional, the proposed methods assume a fixed number of classes. They use local image average as discriminative features. Region labels are modelled by Hidden Markov Measure Field Models. The resulting problems are solved by straightforward alternate minimisation methods, particularly simple in the case of quadratic regularisation of the labels. We demonstrate the proposed methods' capabilities on synthetic data using classical segmentation criteria as well as criteria specific to geoscience. We also present a few examples using real data.

### 2.1 Introduction

Image segmentation remains a fundamental task in Image Analysis, often used as a mandatory preprocessing step for further analysis. The large variety of sources and contents has generated a myriad of approaches, from simple clustering to more sophisticated ones. Our field of application is the analysis of X-ray computerised micro or nanotomograph (X-ray  $\mu$ CT or nCT) of geological samples. While these samples are essentially made of homogeneous materials, with edges and flat surfaces, they are not too structured in terms of shapes. Figure 2.1 illustrates two typical examples, with different resolutions.

Such an image can be modelled as a function  $u : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ , with d = 2, 3

$$u = L\left(\sum_{i=1}^{n} \alpha_i \chi_{\Omega_i}\right) + \eta \tag{2.1.1}$$

where  $\Omega_1, \ldots, \Omega_n$  are the different segments,  $\alpha_i$  their intensity. *L* models blur and resolution effect, such as partial volume;  $\eta$  contains noise and bias field.



**Figure 2.1:** Two (slices) of tomograms with internal reference names (irn): Hod chalk #16[Müter et al., 2014] (a) and WIG1T #156 (b). A bias field is clearly present in (a).

In high-photon-count synchrotron imaging, the noise can be considered as Gaussian, but still depends on the underlying intensity[Buzug, 2008].

A wealth of methods for multiphase image segmentation are based on variational models, especially the classical Mumford-Shah (MS) functional [Mumford and Shah, 1989]. The MS model is region-based and copes with non stationarity in segments by proposing a cartoon-representation of the image data:

$$\mathcal{E}(c,\Gamma;u) = \frac{1}{2} \int_{\Omega} (c-u)^2 \, d\boldsymbol{x} + \frac{\nu}{2} \int_{\Omega \setminus \Gamma} |\nabla c|^2 \, d\boldsymbol{x} + \mu \ell(\Gamma), \tag{2.1.2}$$

with  $\Gamma$  a hypersurface in  $\Omega$  (i.e., 1D when d = 2, 2D when d = 3) such that  $\Omega \setminus \Gamma$  is a disjoint union of regions  $\Omega_1 \sqcup \cdots \sqcup \Omega_{n(\Gamma)}$ , the number of classes *n* being determined by  $\Gamma$ , and  $\ell(\Gamma)$  is the Hausdorff measure of  $\Gamma$ . At optimality, the model provides segment boundaries and a simplified representation *c* of the input image *u*.

The optimisation of  $\Gamma$  is a complex task and has led to many approaches and simplifications. The Chan and Vese formulation[Chan and Vese, 2001] assumes a two-segment model and piecewise constant *c*. Although the model has been extended to multiple classes, it cannot cope with local or global bias fields, or more complex non-stationarity in *u*. For complex non-stationarity and potential transparency-like effects, such as in X-ray CT imaging, an approach based on inpainting was proposed in [Lauze and Nielsen, 2008]. This approach, however, is limited to two phases and relatively simple content. Bias field problems are well known in the MRI community, see for instance [Pham and Prince, 1999]. In contrast to our additive representation, these bias fields are multiplicative.

An approach more relevant to our work is that of [Wang et al., 2010] which presents a global, a regularisation, and a local term model with level sets, but is limited to two phases. It is also worth mentioning the work of [Huang and Zeng, 2015] for two-phase segmentation and bias field estimation. Our methods do not attempt to estimate the bias field, but deal with it implicitely. Closest to our model is the work of [Brox and Cremers, 2009], that we will use in the next section.

To cope with a higher, but fixed number n of segments, binary labelling is used: let  $\mathcal{B}_n = \{e_1, \ldots, e_n\}$  be the standard basis vector of  $\mathbb{R}^n$ . Then a labelling function  $\mathbf{v} = (\mathbf{v}_1, \ldots, \mathbf{v}_n) : \Omega \to \mathcal{B}_n$  is defined by  $\mathbf{v}(\mathbf{x}) = e_i$  iff  $x \in \Omega_i$ , and optimisation is performed on  $\mathbf{v}$  instead of  $\Gamma$ :

$$\mathcal{E}(c, \mathbf{v}; u) = \frac{1}{2} \sum_{i=1}^{n} \int_{\Omega_{i}} (c - u)^{2} \mathbf{v}_{i} \, d\mathbf{x} + \frac{\nu}{2} \sum_{i=1}^{n} \int_{\Omega_{i}} |\nabla c|^{2} \, d\mathbf{x} + \mu \int |D\mathbf{v}|. \quad (2.1.3)$$

The presence of the operator *L* in the image model (2.1.1) makes soft labelling more appropriate in our setting. The concept of soft labelling is formalised in Marroquin's Hidden Markov Measure Field Models (HMMFM) [Marroquin et al., 2003]. This formalisation replaces  $\mathcal{B}_n$  by its convex hull  $\Delta_n = \{(s_1, \ldots, s_n) : \forall i, s_i \ge 0, \sum_i s_i = 1\}$ , also used in [Chambolle et al., 2012] for their relaxation approach, which we use in one of our proposed models.

The rest of the paper is organised as follows: In Section 2.2 we derive two models, differing in their label regularisation term. We propose minimisation algorithms in Section 2.3. Then we demonstrate their capabilities experimentally in Section 2.4. We conclude and discuss future work in the last section.

### 2.2 Derivation of the model

As noted by Brox et al.[Brox and Cremers, 2009], in their two-phase approach, the minimiser c of (2.1.2) on segment  $\Omega_i$  is the Tikhonov regularisation of u on  $\Omega_i$  and satisfies the equation

$$\frac{c-u}{\nu} = \Delta c$$

and approximates the solution at time  $t = \nu$  of the diffusion equation  $c_t = \Delta c$ with initial value  $c_0 = u$ . Following [Nielsen et al., 1997], one can replace the regularisation term  $\frac{\nu}{2} \int_{\Omega_i} |\nabla c|^2 dx$  by an infinite regulariser

 $\frac{1}{2}\sum_{n=1}^{\infty}\frac{\nu^n}{n!}\int_{\Omega_i}|D^nc|^2$ . The corresponding minimiser is the solution of the diffusion equation for all  $\nu > 0$ . These solutions can be approximated by Gaussian convolution, i.e.,  $c \approx g_{\sqrt{\nu}} * u$  On  $\Omega_i$ . This opens the way for other smoothing

kernels, in particular simple rotationally symmetric nearest neighbours (NN) kernels  $h_{\rho}$  (moving average) with radius  $\rho > 0$ . In the sequel g will denote one of these kernels. Kernel parameters, i.e., standard deviation  $\sigma$  for a Gaussian kernel, or the radius  $\rho$  for the NN-kernel, are linked to the variation of the unknown bias field. Following the discussion above, we replace c in (2.1.3) by a smoothing  $c_i$  of u on  $\Omega_i$  and attempt to minimise instead

$$\mathcal{E}(\mathbf{v};u) = \frac{1}{2} \sum_{i=1}^{n} \int_{\Omega} (c_{i} - u)^{2} \mathbf{v}_{i} \, d\mathbf{x} + \frac{\mu}{2} \sum_{i=1}^{n} \int_{\Omega} |D\mathbf{v}_{i}|$$
(2.2.1)

with

$$c_i(\boldsymbol{x}) = \frac{(u\mathbf{v}_i) * g(\boldsymbol{x})}{\mathbf{v}_i * g(\boldsymbol{x})}, \quad \boldsymbol{x} \in \operatorname{supp} \mathbf{v}_i.$$
(2.2.2)

Energy (2.2.1) depends only on the labelling function **v**, though in a complicated way in the local average (2.2.2). The gradient for the data term of (2.2.1) can be derived easily, but the resulting object is rather complex. This can be replaced by a "Chan-Vese"-like trick. Indeed, for  $x \in \Omega_i$ ,  $c_i(x)$  is the minimiser of  $d \mapsto g * [(u - d)^2 \mathbf{v}_i](x)$  as shown by a direct calculation. We can rewrite the variational segmentation problem 2.2.1 as the optimisation of

$$\mathcal{E}_{TV}(\mathbf{c}, \mathbf{v}) = \frac{1}{2} \sum_{i=1}^{n} \int_{\Omega} g * \left[ (u - c_i(\boldsymbol{x}))^2 \mathbf{v}_i \right] (\boldsymbol{x}) \, d\boldsymbol{x} + \frac{\mu}{2} \sum_{i=1}^{n} \int_{\Omega} |D\mathbf{v}_i| \quad (2.2.3)$$
$$\mathbf{v}(\boldsymbol{x}) \in \Delta_n \text{ (a.e)}$$

with  $\mathbf{c} = (c_1 \dots, c_n)$ , and  $c_i : D_i \supset \text{supp } \mathbf{v}_i \rightarrow \mathbb{R}$ . The relaxation of the perimeter term used in [Chambolle et al., 2012] is very tight and often produces almost binary label fields  $\mathbf{v}$ . While this may be desirable in many situations, this can be a drawback, especially when partial volume effects have strong influence, as this is often the case in our 3D micro-CT applications, where a posterior probability may be better than a hard assignment. To avoid this relaxation behaviour, we replace the perimeter term by a quadratic one:

$$\mathcal{E}_{Q}(\mathbf{c}, \mathbf{v}) = \frac{1}{2} \sum_{i=1}^{n} \int_{\Omega} g * \left[ (u - c_{i}(\boldsymbol{x}))^{2} \mathbf{v}_{i} \right] (\boldsymbol{x}) d\boldsymbol{x} + \frac{\mu}{2} \sum_{i=1}^{n} \int_{\Omega} |D\mathbf{v}_{i}|^{2} d\boldsymbol{x} \quad (2.2.4)$$

under the same HMMFM constraint on v.

## 2.3 Optimisation

To optimise segmentation functionals (2.2.3) and (2.2.4), we use an approach where updates on c and v are computed alternately. The same general framework is used for both approaches.

Algorithm 1 Sketch of the algorithms.

**Input:** Input image volume u, number of classes n, weight parameter  $\mu$ , kernel g and maximum number of iterations  $L_{RS}$ .

**Output:** Segmentation  $(\mathbf{v}, \mathbf{c})$  of u.

**Initialisation:** Run a *K*-means or Otsu clustering to produce  $(c^0, v^0)$  from *u*. For (2.2.3), an extra variable  $\bar{v}^0$  is initialised as  $v^0$  and dual variable  $\boldsymbol{\xi}$  is initialised as 0, see below.

```
for r = 0 to L_{RS} do

\triangleright Solution in c

Solve for \mathbf{c}^{r+1} from u and \mathbf{v}^r.

\triangleright Solution in \mathbf{v}

Solve for \mathbf{v}^{r+1} from u and \mathbf{c}^{r+1}.

end for
```

#### 2.3.1 Update for c.

Because of their very definitions, updates of the different local functions  $c_1, \ldots, c_n$  are computed via (2.2.2).

#### 2.3.2 Update for v.

To compute an update of the HMMFM variable v for (2.2.3), we use the framework of [Chambolle et al., 2012] and a proximal method for (2.2.4). In the sequel, we denote by  $\mathcal{L}_n := \mathcal{L}(\Omega, \Delta_n)$  the set of HMMFM  $\Omega \to \Delta_n$  with proper regularity:  $\mathcal{L}(\Omega, \Delta_n) \subset W^{1,2}(\Omega, \mathbb{R}^n)$  (resp.  $\subset BV(\Omega, \mathbb{R}^n)$ ) for functional (2.2.4) (resp. (2.2.3)).

We first compute the gradient (in the  $L^2(\Omega)$  sense) of the shared data term:

$$\mathbf{v} \mapsto \mathcal{E}_D(\mathbf{c}, \mathbf{v}) = \frac{1}{2} \sum_{i=1}^n \int_{\Omega} g * \left[ (u - c_i(\boldsymbol{x}))^2 \mathbf{v}_i \right] (\boldsymbol{x}) \, d\boldsymbol{x}.$$
(2.3.1)

This term is linear in  $\mathbf{v}$ , thus equal to its differential and the gradient is obtained by adjunction, via the easily shown rule

$$\langle g * u, v \rangle_{L^2(\Omega)} = \langle u, \check{g} * (v\chi_{\Omega}) \rangle_{L^2(\mathbb{R}^2)}$$
(2.3.2)

with  $\check{g}(t) = g(-t)$ . Both Gaussian and NN-kernels are even:  $\check{g} = g$ . Note also that if supp  $u \subset \Omega$ , the adjunction rule (2.3.2) simplifies to

$$\langle g * u, v \rangle_{L^2(\Omega)} = \langle u, \check{g} * (v\chi_{\Omega}) \rangle_{L^2(\Omega)}.$$
(2.3.3)

We use these properties to rewrite  $d_{\mathbf{v}} \mathcal{E}_D \cdot \mathbf{w} = \mathcal{E}_D(\mathbf{w}, \mathbf{c})$  as

$$\mathcal{E}_D(\mathbf{c}, \mathbf{w}) = \frac{1}{2} \sum_{i=1}^n \left\langle \mathbf{w}_i, u^2 g * \chi_\Omega - 2ug * c_i + g * c_i^2 \right\rangle_{L^2(\Omega)},$$

20

for a  $\mathbf{w} = (\mathbf{w}_1, \dots, \mathbf{w}_n)$  with support in  $\Omega$ . The sought gradient is thus

$$\nabla_{\mathbf{v}} \mathcal{E}_D = \frac{1}{2} \left( u^2 g * \chi_\Omega - 2ug * c_i + g * c_i^2 \right)_{i=1}^n.$$
(2.3.4)

#### Quadratic regularised functional.

Functional (2.2.4) is differentiable with respect to v, with a gradient given by

$$\nabla_{\mathbf{v}} \mathcal{E}_Q = \nabla_{\mathbf{v}} \mathcal{E}_D - \mu \Delta \mathbf{v} \tag{2.3.5}$$

with  $\Delta$  a vector Laplacian and assuming null Neumann boundary conditions on  $\partial\Omega$ . An update for a current value  $\mathbf{v}^r$  at iteration r is computed in two steps. In the sequel,  $t_r$  is a descent step that may or may not depend on r. We in fact decrease it in our implementation,  $t_r = C/r$  for a constant C.

1. Implicit descent step:  $\overline{\mathbf{v}^r} = \mathbf{v}^r - t_r (\nabla_{\mathbf{v}} \mathcal{E}_D - \mu \Delta \overline{\mathbf{v}^r})$ , i.e. we solve the following equation

$$\left(-\mu\Delta + t_r^{-1} \operatorname{id}\right) \overline{\mathbf{v}^{r+1}} = t_r^{-1} \mathbf{v}_r - \nabla_{\mathbf{v}} \mathcal{E}_D.$$
(2.3.6)

2. Projection:

$$\mathbf{v}^{r+1} = \mathcal{P}_{\mathcal{L}_n}\left(\overline{\mathbf{v}^r}\right). \tag{2.3.7}$$

This is a variation over an incremental proximal step, the reader is referred to [Bertsekas, 2011]. Numerically, a 4-points stencil is used to discretise the Laplacian and we perform one sweep of a Jacobi solver per update. The orthogonal projection operator  $\mathcal{P}_{\mathcal{L}_n}$  projects each  $\mathbf{v}(\mathbf{x})$  on the standard simplex. We use the classical algorithm from [Held and Crowder, 1974].

**Total-Variation regularised functional.** Functional (2.2.3) is not differentiable with respect to v and the optimisation in v from (2.3.6) is replaced by the primal dual step of [Chambolle et al., 2012]. We recall it for the reader's convenience. In a continuous setting, the total variation is defined by duality. The local convex envelope of [Chambolle et al., 2012]

$$J(\mathbf{v}) = \sup\left\{-\int_{\Omega}\sum_{i=1}^{n} \mathbf{v}_{i} \operatorname{div} \xi_{i} : \boldsymbol{\xi} \in \mathcal{C}^{\infty}_{c}(\Omega, \mathbb{R}^{n \times d}, \boldsymbol{\xi}(x) \in K, \forall x \in \Omega\right\}$$

with  $K = \{ \mathbf{q} \in \mathbb{R}^{n \times d} : |q_i - q_j| \le 1, \forall i < j \}$  gives in fact the proper representation of the TV-norm of the HMMFM  $\mathbf{v}$ , due to a convexity argument.

The primal variable at iteration r is  $\mathbf{v}^r$ , used with the extra one  $\bar{\mathbf{v}}^r$ , while the dual variable is  $\boldsymbol{\xi}^r$ . We update them as follows. Set  $\boldsymbol{\xi}_0^r = \boldsymbol{\xi}^r$ ,  $\mathbf{v}_0^r = \mathbf{v}^r$  and  $\bar{\mathbf{v}}_0^r = \bar{\mathbf{v}}^r$ . Then we run I iterations

$$\boldsymbol{\xi}_{i+1}^r = \mathcal{P}_K \left( \boldsymbol{\xi}_i^r + \tau_r \nabla \bar{\mathbf{v}}_i^r \right) \tag{2.3.8}$$

$$\mathbf{v}_{i+1}^r = \mathcal{P}_{\mathcal{L}_n} \left( \mathbf{v}_i^r + t_r \left( \operatorname{div} \boldsymbol{\xi}_{i+1}^r - \nabla_{\mathbf{v}} \mathcal{E}_D \right) \right)$$
(2.3.9)

$$\bar{\mathbf{v}}_{i+1}^r = 2\mathbf{v}_{i+1}^r - \mathbf{v}_i^r.$$
(2.3.10)

21

and set  $\boldsymbol{\xi}^{r+1} = \boldsymbol{\xi}_{I}^{r}$ ,  $\mathbf{v}^{r+1} = \mathbf{v}_{I}^{r}$  and  $\bar{\mathbf{v}}^{r+1} = \bar{\mathbf{v}}_{I}^{r}$ . We set I = 3 as it provides the best results. Numerically, forward differences are used for the gradient  $\nabla$ , and the numerical divergence is defined by duality. In contrast to the quadratic regularisation method, the descent step parameter  $t_{r}$  and ascent step parameter  $\tau_{r}$  are kept fixed.

### 2.4 Experimental Validation

There is no ground truth segmentation available on real tomograms. Therefore we present and evaluate results on a synthetic volume called SYN250 in the sequel. It simulates a porous medium, i.e., a medium which contains voids. It has a background class (void) and three non-background classes, simulating three different materials in a sample. Four segments are thus expected.

Noise is then added to it. A global bias field is added, resulting in the test volume  $SYN250_{global}$ . A per-segment bias field is also added to  $SYN250_{global}$ , resulting in the second test volume  $SYN250_{local}$ . We describe below how we have generated them, the evaluation methodology used, and the parameter selection.

All results have been obtained on a personal computer with a 4th generation Intel Core i7-4910MQ CPU, at 2.90 GHz, 32 GB DDR3 1600 Mhz RAM, and Ubuntu vivid, 15.04 operating system.

#### 2.4.1 Synthetic datasets

**Clean data.** It consists of randomly distributed balls on a 3D volume of size  $250 \times 250 \times 250$  voxels. Radii range from 3 to 10 voxels, uniformly distributed as well as the centre locations. Intensity values follow known material class distributions, and the sampling is stopped when a background to material ratio is reached. Ball overlap is allowed to complicate the geometry of the data.

**Bias field and noise.** Global bias fields are added as 3rd degree polynomials with random coefficients scaled to cover a given percentage of the intensity range. Per-segment bias fields are generated the same way. Gaussian white noise is added with a standard deviation up to 15% of the intensity range. Bias field and noise is added to the ground truth.

#### 2.4.2 Performance criteria

We report primarily Sørensen-Dice (DSC) index. True positive rate (TPR / sensitivity), true negative rate (TNR / specificity), and positive predictive value (PPV / precision) are computed. Also, porosity (amount of background in material), pore network connectivity, which summarises the background class / void inside the material: largest connected pore network divided by
#### 2.4. Experimental Validation



**Figure 2.2:** The synthetic datasets with global (a) and global + class bias fields (b), SYN250<sub>global</sub> and SYN250<sub>class</sub> respectively. Each red square highlights significant differences between the two. In the largest square, an entire class has been cancelled out, due to the class bias fields. (c) shows the segmentation result, using the proposed method with  $\mu = 25$  and kernel size = 256, of a real 2-phase tomogram (irn: Chalk P3 1.1.3C), of dimensions  $800 \times 800 \times 256$ , with added bias field, inducing intensity variations between -10% and 20% of the original intensity range.

the total pore size. Porosity ratio (Por. ratio) is the ratio between measured porosity and the ground truth porosity. Connectivity ratio (Con. ratio) does the same for connectivity.

These parameters describe properties of the pore space, and are essential for the correctness of future geophysical applications, like fluid simulation

#### 2. LOCAL MEAN MULTIPHASE SEGMENTATION



**Figure 2.3:** Segmentation of SYN250<sup>3</sup><sub>global</sub> (row 2) and SYN250<sup>3</sup><sub>local</sub> (row 3), using the proposed method. Row 1 is ground truth. Each column corresponds to a class.

studies. A detailed characterisation and explanation of the porosity and connectivity ratio parameters can be found in [Müter et al., 2012].

Measurements of TPR, TNR, and PPV have been computed for each class individually. To account for class distribution, a weighted sum is used, with weights corresponding to the ground truth class distributions. A unique DSC score formed by averaging the different class versus class is reported.

The running time T is an important factor for the method. All experiments run five iterations of Algorithm 1. We also experimented with up to 100 iterations for all the relevant iterative methods. They showed a slight improvement in segmentation results, but at the cost of a large running time.

#### 2.4.3 Methods

In the sequel, Equation (2.2.4) is referred to as Tikhonov and Equation (2.2.3) as TV. We also use the piecewise constant Mumford-Shah method of [Chambolle et al., 2012] as PCMS. We compare with Otsu and Dual filtering (Dual filter) methods used in [Müter et al., 2012].

Method	kernel parameter	$\left  \begin{array}{c} \frac{1}{\mu} \end{array} \right $
Tikhonov NN	53	91
Tikhonov Gaussian	55	89
TV NN	102	4
TV Gaussian	172	4

Table 2.1: Best parameters for the variants of the proposed method, run on  ${\rm SYN250_{global}}.$ 

**Table 2.2:** DSC, TPR, TNR, PPV, and T values for the segmentation results of  $SYN250_{qlobal}$ , using the selected methods.

Method	kernel	DSC	TPR	TNR	PPV	T (s)
Tikhonov	NN	0.991576	0.98018	0.99504	0.98606	90.10
Tikhonov	Gaussian	0.991422	0.97975	0.99497	0.98587	162.92
TV	NN	0.956096	0.89397	0.97732	0.93107	228.28
TV	Gaussian	0.956050	0.89389	0.97729	0.93097	311.28
Method	-	DSC	TPR	TNR	PPV	T (s)
PCMS	-	0.958090	0.88087	0.98259	0.94811	114.96
Otsu	-	0.894342	0.78868	0.93196	0.80911	2.02
Dual filter	-	0.954899	0.90980	0.96833	0.91485	16.51

#### 2.4.4 Results

Segmentation results of the simulated data are visualised in Figure 2.3, using Tikhonov and NN-kernels. It clearly illustrates the difficulty of a visual evaluation, due to the complexity of the data, and the necessity of objective criteria.

We started by estimating parameters of the methods to provide the best DSC score. Both Tikhonov and TV include kernel parameters and smoothness versus data fidelity weight parameters. For a Gaussian kernel, the kernel parameter is its standard deviation  $\sigma$ , while for the NN-kernel, it is the radius  $\rho$ . Table 2.1 shows the best parameters for the DSC score of the SYN250<sub>global</sub> dataset. The methods show comparable behaviours in terms of kernel parameters as well as smoothness weights. They are slightly suboptimal for SYN250<sub>local</sub>, but with a DSC score difference less than 0.05%.

Next, we compare DSC, TPR, TNR, PPV for the  $SYN250_{global}$  dataset in Table 2.2 for our methods and the other test methods. Running time is also reported in the same table. The same is done for the  $SYN250_{local}$  dataset in Table 2.3. We now report geological relevant criteria, in Table 2.4 for  $SYN250_{global}$  and in Table 2.5 for  $SYN250_{local}$ .

For both datasets, Tikhonov methods perform better than the others with

Method	kernel	DSC	TPR	TNR	PPV	T (s)
Tikhonov	NN	0.978750	0.95435	0.98884	0.96224	76.60
Tikhonov	Gaussian	0.977205	0.95039	0.98801	0.95997	160.85
TV	NN	0.940032	0.86449	0.96659	0.89872	224.71
TV	Gaussian	0.939879	0.86027	0.96764	0.90113	305.83
Method	-	DSC	TPR	TNR	PPV	T (s)
PCMS	-	0.941654	0.84959	0.97265	0.91403	108.18
Otsu	-	0.888680	0.77736	0.92772	0.79739	1.89
Dual filter	-	0.937939	0.87588	0.95904	0.88345	15.44

**Table 2.3:** DSC, TPR, TNR, PPV, and T values for the segmentation results of SYN250<sub>local</sub>, using the selected methods.

**Table 2.4:** Porosity, Connectivity and their respective ratios to the same measures forthe Ground truth, run on SYN250globalusing the selected methods.

Method	kernel	Porosity	Por. ratio	Connectivity	Con. ratio
Ground truth	-	0.249998	1.0	0.999617	1.0
Tikhonov	NN	0.248629	0.994523	0.998814	0.999197
Tikhonov	Gaussian	0.248567	0.994272	0.998851	0.999234
TV	NN	0.237072	0.948296	0.990781	0.991160
TV	Gaussian	0.237048	0.948198	0.990785	0.991164
Method	-	Porosity	Por. ratio	Connectivity	Con. ratio
PCMS	-	0.233670	0.934685	0.997779	0.998161
Otsu	-	0.228024	0.912102	0.980602	0.980978
Dual filter	-	0.228537	0.914155	0.996137	0.996519

a slight advantage on all the criteria used. For DSC, TPR, TNR, PPV, there is a slight advantage to the NN-kernel, while Gaussian seems slightly better for connectivity measurements, though, the difference is not significant. All methods present degraded performances when per-segment bias fields are applied. This is of course to be expected. Figures 2.2a and 2.2b illustrate that in that situation, it seems that some spatially closed segments merge.

Our proposed method has also been validated on real experimental data in Figure 2.2c for a high resolution 2-phased setting and Figures 2.4 and 2.5 for a low resolution 3-phased setting. In both cases a [-10, 20] % bias field was added. From visual inspection, the results are very satisfying in both cases.

Generally the smoothness to data fidelity weight  $\mu$  has the biggest impact on the results, and this is of course expected. We optimised for it directly, but other methods, such as Hansen's L-curve[Hansen and O'Leary, 1993] could be considered.

The extend of the kernel was related to the rate of convergence, larger

Regulariser	kernel	Porosity	Por. ratio	Connectivity	Con. ratio
Ground truth	-	0.249998	1.0	0.999617	1.0
Tikhonov	NN	0.250620	1.002488	0.998527	0.998909
Tikhonov	Gaussian	0.250724	1.002902	0.998541	0.998923
TV	NN	0.243557	0.974236	0.994584	0.994966
TV	Gaussian	0.242811	0.971251	0.995116	0.995498
Method	-	Porosity	Por. ratio	Connectivity	Con. ratio
PCMS	-	0.239993	0.959977	0.998053	0.998436
Otsu	-	0.234354	0.937422	0.979921	0.980297
Dual filter	-	0.236129	0.944522	0.996227	0.996609

**Table 2.5:** Porosity, Connectivity and their respective ratios to the same measures for<br/>the Ground truth, run on  $SYN250_{local}$ , using the selected methods.

kernels converging faster. Of course, choosing a kernel size approaching the dimensions of the image will result in a Chan-Vese / piecewise constant-like behaviour and not remove the bias field properly. Choosing too small a kernel produces similarly poor results because the underlying image structure is not represented well. Kernel extents should be related to the frequency / scale of the bias field. Priori knowledge of this scale could be obtained by considerations of the physics of the measurement device, as well as prior knowledge of the material content.

#### 2.5 Conclusion

In this paper we have proposed two multiphase segmentation methods that cope with noise and bias fields for complex low structured volumes, especially targeted to  $\mu$ CT and nanoCT used in geology. Based on the Mumford-Shah functional, they discriminate regions based on local averages of their contents. Using HMMFM and with Tikhonov regularisation or TV regularisation, they provide spatially coherent class posteriors of voxels. The methods are easy to implement and reasonably fast, though we have not optimised for running time. Their running time performances could be greatly improved by proper parallelisation.

We have investigated the methods on synthetic datasets via Gaussian- and NN-kernels. They appear robust to parameter variation. They show similar results, though with an advantage to NN-kernels and quadratic regularisation. We have validated them using some classical segmentation validation criteria as well as more specific criteria used in geoscience. In a future work, we will investigate means for estimating the kernel extent and develop parallel code.

#### 2. LOCAL MEAN MULTIPHASE SEGMENTATION



**Figure 2.4:** Segmentation of a real 3-class dataset full size, (irn: Hod chalk #16 from the North Sea Basin) using the proposed method. Size =  $256 \times 1025 \times 825$ .  $\nu = 300$ , kernel size =  $512^3$ .

#### Acknowledgements

J. Hansen and F. Lauze acknowledge funding from the Innovation Fund Denmark and Mærsk Oil and Gas A/S, for the P<sup>3</sup> Project. We thank Henning Osholm Sørensen for giving us access to the experimental tomography data.



Figure 2.5: Segmentation of a real 3-class dataset zoomed version, (irn: Hod chalk #16 from the North Sea Basin) using the proposed method. Size =  $256 \times 1025 \times 825$ .  $\nu = 300$ , kernel size =  $512^3$ .

# 3 Local Mean Geodesic Active Regions

The work presented in this chapter is based on [Hansen and Lauze, 2018]<sup>†</sup>.

<sup>†</sup> Hansen, J. D. K. and Lauze, F. (2018). Multiphase Local Mean Geodesic Active Regions. In *24th International Conference on Pattern Recognition*, pages 3031–3036. IEEE.

# Multiphase Local Mean Geodesic Active Regions

J. D. K. Hansen, F. Lauze, {jdkh,francois@di.ku.dk}

Department of Computer Science, University of Copenhagen

Abstract. This paper presents two variational multiphase segmentation methods for recovery of segments in weakly structured images, presenting local and global intensity bias fields, as often is the case in microtomography. The proposed methods assume a fixed number of classes. They use local image averages as discriminative features and binary labelling for class membership and their relaxation to per pixel/voxel posterior probabilities, Hidden Markov Measure Field Models (HMMFM). The first model uses a Total Variation weighted semi-norm (wTV) for label field regularisation, similar to Geodesic Active Contours, but with a different and possibly richer representation. The second model uses a weighted Dirichlet (squared gradient) regularisation. Both problems are solved by alternating minimisation on computation of local class averages and label fields. The quadratic problem is essentially smooth, except for HMMFM constraints. The wTV problem uses a Chambolle-Pock scheme for label field updates. We demonstrate the capabilities of the approaches on synthetic examples, and illustrate them on a real examples.

#### 3.1 Introduction

Image Segmentation is one of the very first steps necessary in Image Analysis and understanding. The extraordinary variety of image sources and settings has generated and continue to generate a very broad spectrum of approaches and techniques. Our application will be the analysis of geological samples through X-ray computerised micro or nanotomography (X-ray  $\mu$ CT or nCT). Figure 3.1 serves as an example of typical material reconstructions. While we visually observe, in the left example, essentially homogeneous materials with mostly well defined edges up to some scale, we also note a clear bias field. In the right image, coarser resolution and some CT reconstruction artifacts can make some of the edges less pronounced. In both images individual object sizes and shapes are not notably structured. While Deep Learning



**Figure 3.1:** Two (slices) of tomograms with internal reference names (irn): Hod chalk #16[Müter et al., 2014] (a) and WIG1T #156 (b). A bias field is clearly present in (a).

methods are generally state of the art in segmentation, they are extremely data intensive and unfortunately, ground truth segmentation of the type of material we are interested in is rare. Thus we instead resort to methods which attempt to model the signal that need processing. In a previous work [Hansen and Lauze, 2017], we introduced two multiphase variational algorithms capable of modelling non stationarity of the phases together with classical regularisation that encourages limited phase interface length/areas. They were intensity based and did not use the local edge structure information. In this work we add edge information to the regularisations, to improve robustness of the method. The images and volumes we are interested in have limited structure and we model such as a function  $u : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ , with d = 2, 3 as

$$u = L\left(\sum_{i=1}^{n} \alpha_i \chi_{\Omega_i}\right) + \eta \tag{3.1.1}$$

where  $\Omega_i$  is the *i*th segment and  $\alpha_i$  its intensity. *L* models blur and resolution effects, such as partial volume;  $\eta$  contains the noise and bias field, which generates non stationarity on the observed segments. While still depending on the underlying intensity, the noise in high-photon-count synchrotron imaging can be considered Gaussian[Buzug, 2008].

We choose to base our method on the classical region-based Mumford-Shah (MS) functional[Mumford and Shah, 1989], that models the non stationarity in segments by proposing the following data representation:

$$\mathcal{E}(c,\Gamma;u) = \frac{1}{2} \int_{\Omega} (c-u)^2 \, d\boldsymbol{x} + \frac{\nu}{2} \int_{\Omega \setminus \Gamma} |\nabla c|^2 \, d\boldsymbol{x} + \mu \ell(\Gamma), \tag{3.1.2}$$

33

here  $\Gamma$  is a hypersurface in  $\Omega$ , such that  $\Omega \setminus \Gamma$  is a disjoint union of regions  $\Omega_1 \sqcup \cdots \sqcup \Omega_{n(\Gamma)}$ , *n* is the number of classes determined by  $\Gamma$ , and  $\ell(\cdot)$  the Hausdorff measure. At optimality, the MS model outputs the sought decomposition of  $\Omega$  as simplified representations *c* of the input image *u*.

Numerous approaches have been proposed to deal with the optimisation of  $\Gamma$ , such as the Chan and Vese formulation[Chan and Vese, 2001] - a twophase approach, later extended to multiphase, where *c* is piecewise constant and equates to the average of its respective class. [Lauze and Nielsen, 2008] attempted to model complex non-stationarity and transparency-like effects, observed in X-ray imaging using an inpainting-based approach. Both of these methods however are unable to cope with local or global bias fields by their assumptions. Although additive bias field problems are not extensively studied, multiplicative ones are common in the MRI community, as for example seen in [Pham and Prince, 1999].

In our previous work [Hansen and Lauze, 2017], we proposed two methods sharing the same data attachment term based on local average discrimination features with different regularisers. They seek to implicitly deal with the bias fields, rather than explicitly estimating it, like the two-phase approach of [Huang and Zeng, 2015]. The work of [Wang et al., 2010] is essentially equivalent to our approach for the data term, though for a two-phases formulation based on a levelset implementation.

We instead write the segmentation problem as a binary labeling problem, with regularisation of the label fields. In order to make the labeling sensitive to the local image structure, we modify the regularisation to incorporate a measure  $h(\mathbf{x})$  of the local structure around a point  $\mathbf{x} \in \Omega$ , in a way similar to [Paragios and Deriche, 2002a], but with our labeling. To optimise it, by continuous methods, we convexify it and show how the local tight relaxation of [Chambolle et al., 2012] can be easily extended by incorporating the measure of the structure. As in the seminal work of [Caselles et al., 1997], the structure measure h(x) can be interpreted as an isotropic Riemannian metric on the image / volume domain.

The convexification of the problem also leads to a natural reformulation in terms of the Hidden Markov Measure Field models (HMMFM) of Marroquin [Marroquin et al., 2003], as implicitly done in [Chambolle et al., 2012] and our previous work [Hansen and Lauze, 2017]. Because of the operator L in (3.1.1), we may want smoother label fields and we also study a weighted quadratic regularisation.

The rest of the paper is organised as follows: In Section 3.2 we extend our two previously described models, introducing an image structure guided regularisation of the label field. We propose updated minimisation algorithms in Section 3.3. Their capabilities are experimentally demonstrated in Section 3.4 and compared with previous results. In the last section we conclude and discuss future work.

#### 3.2 Derivation of the model

Before proceeding to the introduction of our model, we need a few notations. In the sequel, the standard basis of  $\mathbb{R}^n$  is denoted  $\mathcal{B}_n = (e_1, \ldots, e_n)$ . A binary n-labeling function  $\mathbf{v}$  is a vector valued function  $\mathbf{v} = (v_1 \ldots, v_n) : \Omega \to \mathcal{B}_n$ . The convex hull of  $\mathcal{B}_n$  is the standard simplex  $\Sigma_n = \{(s_1, \ldots, s_n) | \forall i, s_i \geq 0, \sum_{i=1}^n s_i = 1\}$  and HMMFM are simply  $\Sigma_n$ -valued functions. The interior of a subset  $H \subset \Omega$  is denoted  $\mathring{H}$ . In the sequel we follow closely the notations of [Chambolle et al., 2012].

#### 3.2.1 Weighted Perimeters and Relaxation.

Recall that the total variation of a function  $u: \Omega \to \mathbb{R}$  is defined as

$$\mathbf{J}(u) = \inf\left\{\int_{\Omega} u \operatorname{div} \varphi d\mathbf{x} | \varphi \in \mathcal{C}_{c}^{1}\left(\Omega, \mathbb{R}^{d}\right), \|\varphi(\mathbf{x})\| \leq 1\right\}$$
(3.2.1)

and if  $\int_{\Omega} |Du| := \mathbf{J}(u) < \infty$ , u has bounded variation. The space of functions of bounded variations is denoted  $BV(\Omega)$ . When  $u = \chi_H$ ,  $H \subset \Omega$  is the characteristic function of H,  $\mathbf{J}(\chi_H) = \operatorname{Per}(H)$  the perimeter of H in  $\Omega$ . This extends to vector valued functions and the reader is referred to [Chambolle et al., 2012]. Given a function  $h : \Omega \to \mathbb{R}^+$ , there is a notion of *weighted total variation* of u,  $\mathbf{J}_h(u) = \int_{\Omega} h |Du|$ , by replacing the constraint  $\|\varphi(\mathbf{x})\| \leq 1$ by  $\|\varphi(\mathbf{x})\| \leq h(\mathbf{x})$  in (3.2.1). Call the corresponding space  $BV_h(\Omega)$ . The h-perimeter of  $H \subset \Omega$  is  $\operatorname{Per}_h(H) = \mathbf{J}_h(\chi_H)$ . Under sufficient regularity conditions on h, this can be interpreted as the Hausdorff measure of the boundary  $\partial H$  of H on the Riemannian manifold  $\Omega$  equipped with the isotropic metric  $H(\mathbf{x}) = h^2(\mathbf{x})Id$ . Given a binary labeling  $\mathbf{v} : \Omega \to \mathcal{B}_n$ , each  $v_i$  is the characteristic function of an  $\Omega_i \subset \Omega$  and the measure of  $\Gamma = \Omega - \sqcup \mathring{\Omega}_i$  is given by  $\frac{1}{2} \sum_{i=1^n} \operatorname{Per}_h \Omega_i$ . In the sequel we set

$$BV_h(\Omega, \mathcal{B}_n) = \{ \mathbf{v} : \Omega \to \mathcal{B}_n, v_i \in BV_h(\Omega) \},\$$

$$BV_h(\Omega, \Sigma_n) = \{ \mathbf{v} : \Omega \to \Sigma_n, v_i \in BV_h(\Omega) \}.$$

Define

$$\mathcal{F}_{h}(\mathbf{v}) = \begin{cases} \frac{1}{2} \sum_{i=1^{n}} \operatorname{Per}_{h} \Omega_{i} & \mathbf{v} \in BV_{h}(\Omega, \mathcal{B}_{n}) \\ +\infty & \mathbf{v} \notin BV_{h}(\Omega, \mathcal{B}_{n}). \end{cases}$$
(3.2.2)

As  $BV_h(\Omega, \mathcal{B}_n)$  is not convex,  $\mathcal{F}_h(\mathbf{v})$  is not. Computing its convex relaxation does not seem possible but a tightest *local* relaxation, under general enough assumptions, can be obtained by setting

$$\mathcal{J}_{h}(\mathbf{v}) = \begin{cases} \int_{\Omega} \psi^{**}(x, D\mathbf{v}) & \mathbf{v} \in BV_{h}(\Omega, \Sigma_{n}) \\ +\infty & \mathbf{v} \notin BV_{h}(\Omega, \Sigma_{n}). \end{cases}$$
(3.2.3)

35

where  $\psi^{**}$  is the convex relaxation ([Ekeland and Téman, 1999]) w.r.t. to its second variable  $\mathbf{p} \in \mathbb{R}^{n \times d}$  of  $\psi$ , defined by

$$\psi(\mathbf{x}, \mathbf{p}) = \begin{cases} h(\mathbf{x})|p| & \text{if } \mathbf{p} = (e_i - e_j) \otimes p \\ +\infty & \text{otherwise.} \end{cases}$$

Setting

$$K_{\alpha} = \{ \mathbf{q} = (q_1, \dots, q_n)^T \in \mathbb{R}^{n \times d}, |q_i - q_j| \le \alpha, \forall i < j \}$$
(3.2.4)

one has  $\mathbf{p} \mapsto \psi^{**}(\mathbf{x}, \mathbf{p}) = \sum_{\mathbf{q} \in K_{h(\mathbf{x})}} \mathbf{q} \cdot \mathbf{p}.$ 

Proofs are very similar to the non weighted case and ommitted in this work.

#### 3.2.2 Non Stationarity and Proposed Models.

In our previous work [Hansen and Lauze, 2017], we proposed to minimise

$$\mathcal{E}(\mathbf{v};u) = \frac{1}{2} \sum_{i=1}^{n} \int_{\Omega} (c_i - u)^2 \mathbf{v}_i \, d\mathbf{x} + \frac{\mu}{2} \sum_{i=1}^{n} \int_{\Omega} |D\mathbf{v}_i|^2$$
(3.2.5)

with

$$c_i(\boldsymbol{x}) = \frac{(u\mathbf{v}_i) * g(\boldsymbol{x})}{\mathbf{v}_i * g(\boldsymbol{x})}, \quad \boldsymbol{x} \in \operatorname{supp} \mathbf{v}_i.$$
(3.2.6)

arguing that it can be seen as an approximation of (3.1.2), after replacing the quadratic regulariser  $\int |\nabla \mathbf{v}|^2$  with an appropriate infinite order one. Energy (3.2.5) depends only on the labeling function  $\mathbf{v}$ , though in a complicated way in the local average (3.2.6). The gradient for the data term of (3.2.5) can be easily derived, but the resulting object is rather complex. This can be replaced by a "Chan-Vese"-like trick. Indeed, for  $\mathbf{x} \in \Omega_i$ ,  $c_i(\mathbf{x})$  is the minimiser of  $d \mapsto g * [(u - d)^2 \mathbf{v}_i](\mathbf{x})$  as shown by a direct calculation. Replacing the flat regularisation by the weighted one introduced in the previous paragraph, we propose the following generalised Geodesic Active Regions functional

$$\mathcal{E}_{TV}(\mathbf{c}, \mathbf{v}) = \frac{1}{2} \sum_{i=1}^{n} \int_{\Omega} g * \left[ (u - c_i(\boldsymbol{x}))^2 \mathbf{v}_i \right] (\boldsymbol{x}) \, d\boldsymbol{x} + \mu \mathcal{J}_h(\mathbf{v})$$
(3.2.7)

with  $\mathbf{v}(\mathbf{x}) \in \Sigma_n$  (a.e),  $\mathbf{c} = (c_1 \dots, c_n)$ , and  $c_i : D_i \supset \text{supp } \mathbf{v}_i \rightarrow \mathbb{R}$ . The relaxation of the perimeter term used in (3.2.7) is very tight and often produces almost binary label fields  $\mathbf{v}$ . While this may be desirable in many situations, this can be a drawback, especially when partial volume effects have strong influence, as this is often the case in our 3D micro-CT applications, where a posterior probability may be better than a hard assignment. To avoid this relaxation behavior, we can replace the perimeter term by a quadratic one:

$$\mathcal{E}_Q(\mathbf{c}, \mathbf{v}) = \frac{1}{2} \sum_{i=1}^n \int_{\Omega} g * \left[ (u - c_i(\boldsymbol{x}))^2 \mathbf{v}_i \right] (\boldsymbol{x}) \, d\boldsymbol{x} + \frac{\mu}{2} \| D \mathbf{v} \|_h^2$$
(3.2.8)

with  $||D\mathbf{v}||_h^2 = \sum_{i=1}^n \int_{\Omega} h |D\mathbf{v}_i|^2$  under the same HMMFM constraint on  $\mathbf{v}$ .

#### 3.3 Optimisation

To optimise segmentation functionals (3.2.7) and (3.2.8), we use an approach where updates on c and v are computed alternately. The same general framework is used for both approaches.

Algorithm 2 Sketch of the algorithms.

**Input:** Input image volume u, number of classes n, weight parameter  $\mu$ , kernel g and maximum number of iterations  $L_{RS}$ .

**Output:** Segmentation  $(\mathbf{v}, \mathbf{c})$  of u.

**Initialisation:** Run a *K*-means or Otsu clustering to produce  $(c^0, v^0)$  from *u*. For (3.2.7), an extra variable  $\bar{v}^0$  is initialised as  $v^0$  and dual variable  $\boldsymbol{\xi}$  is initialised as 0, see below.

for r = 0 to  $L_{RS}$  do Solve for  $\mathbf{c}^{r+1}$  from u and  $\mathbf{v}^r$ . Solve for  $\mathbf{v}^{r+1}$  from u and  $\mathbf{c}^{r+1}$ . end for

#### 3.3.1 Update for c.

Because of their very definitions, updates of the different local functions  $c_1, \ldots, c_n$  are computed via (3.2.6).

#### 3.3.2 Update for v.

To compute an update of the HMMFM variable v for (3.2.7), we use the framework of [Chambolle et al., 2012] modified to incorporate the weight function h and a simple proximal method for (3.2.8). We only highlight the parts that differ from our previous work [Hansen and Lauze, 2017] due to the presence of the weight function h.

Both formulations share their data term

$$\mathbf{v} \mapsto \mathcal{E}_D(\mathbf{c}, \mathbf{v}) = \frac{1}{2} \sum_{i=1}^n \int_{\Omega} g * \left[ (u - c_i(\boldsymbol{x}))^2 \, \mathbf{v}_i \right] (\boldsymbol{x}) \, d\boldsymbol{x}.$$
(3.3.1)

Quadratic regularised functional.

Functional (3.2.8) is differentiable with respect to v, with a gradient given by

$$\nabla_{\mathbf{v}} \mathcal{E}_Q = \nabla_{\mathbf{v}} \mathcal{E}_D - \mu \nabla \cdot (h D \mathbf{v}) \tag{3.3.2}$$

with  $\nabla \cdot$  a vector divergence, i.e. divergence of the lines of the matrix  $hD\mathbf{v}$ , with  $D\mathbf{v}$  the Jacobian matrix of  $\mathbf{v}$ , assuming null Neumann boundary conditions on  $\partial\Omega$ . The implicit descent step of [Hansen and Lauze, 2017] is replaced by  $\overline{\mathbf{v}^r} = \mathbf{v}^r - t_r (\nabla_{\mathbf{v}} \mathcal{E}_D - \mu \nabla \cdot (h \nabla \overline{\mathbf{v}^r}))$ , i.e. we solve the following equation

$$\left(-\mu\nabla\cdot h\nabla + t_r^{-1}\operatorname{id}\right)\mathbf{v}^{r+1} = t_r^{-1}\mathbf{v}_r - \nabla_{\mathbf{v}}\mathcal{E}_D.$$
(3.3.3)

37

Numerically, a 4-points stencil is used to discretise the Laplacian and we perform one sweep of a Jacobi solver per update.

#### Total-Variation regularised functional.

Functional (3.2.7) is not differentiable with respect to v and the optimisation in v from (3.3.3) is replaced by the primal dual step of [Chambolle et al., 2012] adapted to the weighted total variation. Define

$$\mathcal{K}_{h} = \left\{ \xi \in \mathcal{C}_{c}^{1}(\Omega, \mathbb{R}^{n \times d}), \xi(\mathbf{x}) \in K_{h(\mathbf{x})}, \forall \mathbf{x} \in \Omega \right\}.$$

with  $K_{\alpha}$ ,  $\alpha > 0$  defined in (3.2.4). Modifying the optimisation scheme in [Chambolle et al., 2012] to incorporate the function h is straightforward. The orthogonal projection  $\mathcal{P}_{\mathcal{K}_h}$  on  $\mathcal{K}_h$  is the local version of the projection on K, and can as in [Chambolle et al., 2012], be computed by Dykstra projection algorithm. Numerically, forward differences are used for the gradient  $\nabla$ , and the numerical divergence is defined by duality. We run only three iterations of the Chambolle-Pock algorithm for each update of **v**.

#### 3.4 Experimental Validation

Ground truth annotated segmentations are not available for real tomograms. We use the same synthetic volumes as used in [Hansen and Lauze, 2017] for this framework. They are created as follow: a first volume of size  $250^3$  voxels, called SYN250 is generated. It consists of a background and randomly distributed balls with radii between 3 and 10 voxels. Intensity values for these balls follow known material class distribution and sampling is done until a selected background to material ratio is reached. To allow for more complicated geometry, balls can overlap. Global and per segment bias fields are added as 3rd degree polynomials with randomised coefficients. Gaussian distributed noise of standard deviation corresponding to 15% of the intensity range is also added. The volume with only global bias field is called SYN250<sub>global</sub>, the one with added segment bias fields is called SYN250<sub>class</sub>. The two datasets can be seen in Figure 3.2.

The weight / structure function used in our experiments is a classical edge detector one,

$$h(\mathbf{x}) = \frac{1}{1 + \left(\frac{|\nabla u_{\sigma}(\mathbf{x})|}{\kappa}\right)^2}$$

where  $u_{\sigma} = g_{\sigma} * u$  is the convolution of u with a isotropic Gaussian of standard deviation  $\sigma$  and  $\kappa$  is an edge / contrast parameter. In all our experimentations,  $\sigma = 0.5$  voxels and  $\kappa = 0.05$ .

#### 3.4. Experimental Validation



**Figure 3.2:** The synthetic datasets  $SYN250_{global}$  and  $SYN250_{class}$  with global (a) and global + per segment bias fields (b), respectively. The red squares illustrate the important differences between the two. The largest square show an entire segment has been canceled out, merged due to the per segment bias fields.

#### 3.4.1 Performance of Segmentation

We primarily report the Sørensen-Dice (DSC) coefficient, true positive rate (TPR / sensitivity), true negative rate (TNR / specificity), and positive predictive value (PPV / precision). Because our goal is the study of porous rocks, we report two geophysical parameters: porosity (the lack of material) and pore network connectivity, which describes the pore class inside the material: specifically, it is the largest connected pore network divided by the total amount of pore. Pore space properites like these are important for future geophysical applications, like fluid flow analysis. For a more in-depth explanation of the two parameters we refer to [Müter et al., 2012].

Since the measurements of TPR, TNR, and PPV are based on individual classes, we report a weighted sum version, using the ground truth class distributional weights. DSC scores are disclosed as averaing the individual class vs class scores.

All listed results of the proposed method, are based on 5 gradient descent steps, but even for 100 descent steps the similarity measures are still slightly improving. A convergence criteria is therefore advised, based on the change in mean squared error of the label field, in each iteration.

#### 3. LOCAL MEAN GEODESIC ACTIVE REGIONS



**Figure 3.3:** A segmentation result for  $SYN250^3_{global}$  and 4 classes



**Figure 3.4:** Soft labeling segmentation of SYN250<sup>3</sup><sub>local</sub> using the proposed Tikhonov method. The two classes illustrate the difficulty of visual interpretation of the results.

#### 3.4.2 Methods

In the sequel, W-Tikhonov refers to the proposed weighted quadratic method and W-TV refers to the proposed weighted total variation method. We compare with method (2.4) and (2.5) from [Hansen and Lauze, 2017] by setting h = 1, and refer to these as Tikhonov and TV respectively. We also compare with the piecewise constant Mumford-Shah proposed method of [Chambolle

Method	kernel parameter	$\frac{1}{\mu}$
Tikhonov NN	45	20
TV NN	65	45
W-Tikhonov NN	45	20
W-TV NN	65	45

Table 3.1: Best parameters for the variants of the proposed method, run on  ${\rm SYN250_{global}}.$ 

et al., 2012] using alias PCMS. We also use Otsu thresholding and the Dual filtering method of [Müter et al., 2012].

#### 3.4.3 Results

Figure 3.3 and Figure 3.4 show segmentation results of the synthetically generated data, using the weighted Tikhonov method. The figures highlight the complications of visually evaluating results, and necessitates objective quantitative criteria, due to the complexity of the data.

For both the Tikhonov and TV case, kernel scale and smoothness versus data fidelity weight parameters were estimated first, with respect to highest DSC score of the SYN250<sub>global</sub> dataset. The kernel parameter for our selected nearest neighbour (NN)-kernel is the radius  $\rho$ . We do not include Gaussian kernels as NN-kernels seemed to slightly outperform them. We intentionally keep the parameters fixed to illustrate the sole contribution of adding the *h* function. Optimal parameters, based on DSC scores, can be found in Table 3.1.

DSC, TPR, TNR, and PPV are compared in Table 3.2 for  $SYN250_{global}$ , for the proposed and selected test methods. Fixing the parameters we found to be optimal from segmenting  $SYN250_{global}$ , we now segment  $SYN250_{local}$  and report the results in Table 3.3. Table 3.4 and 3.5 contains the relevant geological criteria for  $SYN250_{global}$  and  $SYN250_{local}$  respectively.

For all measurements used, the proposed weighted Tikhonov method variant performs slightly better than the others, except for the connectivity scores - which gives the edge to the regular Tikhonov method. As we expect, all used criteria show decreased scores for the more complex dataset SYN250<sub>local</sub>, as different class segments appear to merge, as illustrated in Figures 3.2a and 3.2b. However these differences are only around 1.5%. We note a slight advantage to regular TV regularisation over the proposed weighted version. This could be parameter or data specific, as visual results on experimental data look preferable for the weighted variant of the regulariser. The choice of *h* function and its parameters should also be adjusted to the regulariser and dataset.

We also compared the proposed W-Tikhonov and the regular Tikhonov

Method	kernel	DSC	TPR	TNR	PPV
Tikhonov	NN	0.989286	0.97406	0.99336	0.98296
TV	NN	0.984218	0.96703	0.98956	0.97013
W-Tikhonov	NN	0.990615	0.97918	0.99376	0.98325
W-TV	NN	0.980704	0.92985	0.97880	0.94153
Method	-	DSC	TPR	TNR	PPV
PCMS	-	0.958090	0.88087	0.98259	0.94811
Otsu	-	0.894342	0.78868	0.93196	0.80911
Dual filter	-	0.954899	0.90980	0.96833	0.91485

Table 3.2: DSC, TPR, TNR, PPV, and T values for the segmentation results of  $SYN250_{global}$ , using the selected methods.

**Table 3.3:** DSC, TPR, TNR, PPV, and T values for the segmentation results of  $SYN250_{local}$ , using the selected methods.

Method	kernel	DSC	TPR	TNR	PPV
Tikhonov	NN	0.973969	0.94232	0.98584	0.95473
TV	NN	0.967062	0.93240	0.98052	0.93869
W-Tikhonov	NN	0.975307	0.94816	0.98604	0.95477
W-TV	NN	0.963770	0.92598	0.97820	0.93244
Method	-	DSC	TPR	TNR	PPV
PCMS	-	0.941654	0.84959	0.97265	0.91403
Otsu	-	0.888680	0.77736	0.92772	0.79739
Dual filter	-	0.937939	0.87588	0.95904	0.88345

methods on 3-phased experimental dataset with low resolution, as seen in Figure 3.5. Inspecting the result visually we conclude that they are very satisfying for the proposed weighted variant, as pore throat branches and crooks are segmented more satisfiably. These features are essential to petrophysical parameter extraction from fluid flow simulations and further analysis.

As expected, the biggest influence on the segmentation results stems from the data fidelity to smoothness weight  $\mu$ . Instead of manually optimising this parameter, we could employ Hansen's L-curve method[Hansen and O'Leary, 1993]. The rate of convergence is naturally influenced by the quality of the initialisation, but also the extend of the kernels used. We observe that larger kernels faster. The optimal scale of these is connected to the severity of the bias field and the intensity frequency of these. If these parameters are not easily accessible from the material content, prior physics knowledge could be acquired from the measurement devices. Choosing a kernel that covers the entire image domain will naturally result in a piecewise constant-



#### 3.4. Experimental Validation

**Figure 3.5:** 10 iteration segmentation of a crop of a real 3-class dataset (irn: WIG1T #156 from the North Sea Basin) using the unweighted and proposed weighted Tikhonov method. Size =  $1025 \times 1025$ .  $\frac{1}{\mu} = 10$ , kernel size =  $3^2$  for both methods,  $\kappa = 0.05$  and  $g_{\sigma} = 1$  in the weighted case.

#### 3. LOCAL MEAN GEODESIC ACTIVE REGIONS

Method	kernel	Porosity	Connectivity
Ground truth	-	0.249998	0.999617
Tikhonov	NN	0.245389	0.999065
TV	NN	0.244769	0.996200
W-Tikhonov	NN	0.247282	0.998602
W-TV	NN	0.242451	0.995784
Method	-	Porosity	Connectivity
PCMS	-	0.233670	0.997779
Otsu	-	0.228024	0.980602
Dual filter	-	0.228537	0.996137

 Table 3.4: Porosity and Connectivity measurements compared to the Ground truth measures, run on SYN250global, using the selected methods.

 Table 3.5: Porosity and Connectivity measurements compared to the Ground truth measures, run on SYN250<sub>local</sub>, using the selected methods.

Regulariser	kernel	Porosity	Connectivity
Ground truth	-	0.249998	0.999617
Tikhonov	NN	0.248905	0.999047
TV	NN	0.249490	0.996240
W-Tikhonov	NN	0.250266	0.998517
W-TV	NN	0.248083	0.995940
Method	-	Porosity	Connectivity
PCMS	-	0.239993	0.998053
Otsu	-	0.234354	0.979921
Dual filter	-	0.236129	0.996227

like behavior for the data term and consequently not properly remove bias fields. If a too small kernel is employed, similar poor results can be observed, as the representation of the underlying image structure is not acceptable. Additionally, the convergence rates of smaller kernel model selections is heavily influenced by a good initialisation.

#### 3.5 Conclusion

We proposed two variational multiphase segmentation methods in this paper, designed to handle complex, low structured data contaminated by noise and bias field. These methods are especially targeted to  $\mu$ CT and nanoCT used in geology. We have done it by desinging functionals with a data term robust to potential bias field and Gaussian noise, and a label field regulariser which

is adpated to the underlying image structure. We have shown how, in the weighted TV case, this can be easily optimised by a Chambolle-Pock scheme, while it leads to a standard elliptic problem in the quadratic case. We have inspected performance of our methods on synthetically generated datasets as well as a real dataset. The proposed methods are straightforward to implement and run fairly quick, without being optimised for it. Running time could be notably improved by parallelisation. In future works, we plan to investigate techniques for kernel extent estimation and implement parallelisation. For the regularisers, we will also look into general anisotropic TV terms.

#### Acknowledgements

J. Hansen and F. Lauze acknowledge funding from the Innovation Fund Denmark and Mærsk Oil and Gas A/S, for the  $P^3$  Project. We thank Henning Osholm Sørensen for giving us access to the experimental tomography data.

# 4 Rat Brain Extraction and Segmentation

The work presented in this chapter is based on [Hansen et al., 2018]<sup>†</sup>.

<sup>†</sup> Hansen, J. D. K. et al. (2018). Brain Extraction and Segmentation Framework for Bias Field Rich Cranial MRI Scans of Rats. In *Book of Abstracts, Annual meeting ISMRM*, page 3249, Paris, France.

# Brain Extraction and Segmentation Framework for Bias Field Rich Cranial MRI Scans of Rats

Jacob Daniel Kirstejn Hansen<sup>1</sup>, François Lauze<sup>1</sup>, Sune Darkner<sup>1</sup>, Kristian Nygaard Mortensen<sup>2</sup>, Simon Sanggaard<sup>2</sup>, Helene Benveniste<sup>3</sup>, and Maiken Nedergaard<sup>2,4</sup>

 <sup>1</sup>Department of Computer Science, University of Copenhagen, <sup>2</sup>Center for Translational Neuromedicine, University of Copenhagen, <sup>3</sup>Anesthesiology, Yale School of Medicine, Yale University,
 <sup>4</sup>Center for Translational Neuromedicine, University of Rochester

**Abstract.** This abstract presents a framework to extract brain tissue and internal Cerebrospinal fluid in cranial magnetic resonance imaging of rats with strong bias fields. Desired segments are obtained through bias field correction and several passes of segmentation. A refinement procedure is proposed to remove brain surface CSF. Promising planar and 3D visualisations of results are presented and demonstrate the capabilities of the framework.

#### 4.1 Introduction

Magnetic resonance imaging employing surface receive coils often faces the problem of strong image bias fields in data that makes relevant tissue extractions very challenging, as bias field correction algorithms cannot fully correct for it. We present a framework for the extraction of rat brain tissue and the cerebrospinal fluid (CSF) networks, in the case where MR volumes are affected by strong bias fields. We demonstrate that our framework obtains very satisfactory segmentation results.

#### 4.1.1 Animals

Unaffected male rats of the Wistar Kyoto stain (WKY) and spontaneously hypertensive rats (SHR) were obtained from Charles River, Germany. Separate groups of rats were scanned at two age-ranges: young (7-9 weeks old; WKY7, n = 11; and SHR7, n = 8) and young adults (19-21 weeks WKY19, n = 8; and SHR19, n = 9). All treatments and imaging were performed according to protocols approved by the IACUC, and according to a protocol approved by the University of Copenhagen animal experimentation committee.



**Figure 4.1:** The proposed framework, highlighting each step as a box inside the dotted blue box. Arrows indicate that the output of the origin box is used as input in the destination box.

#### 4.1.2 MRI

All MRI investigations was conducted on a 9.4 T magnet (Bruker Biospec 9.4/30 USR) interfaced to a Bruker Advance III console and controlled by Paravision 5.1 software (Bruker BioSpin) at the Panum Institute, University of Copenhagen. Imaging was performed with an 86 mm3 volume resonator and a surface quadrature array receiver coil. Images were acquired with the spoiled gradient FLASH3D sequence (TE: 4 ms, TR: 15 ms, NA: 3; matrix:  $128 \times 128 \times 128$ , voxel size:  $0.24 \times 0.24 \times 0.26$  mm, FA: 15°, scan time: 4:05 min).

#### 4.2 Methods

The proposed framework consists of 4 major steps and is illustrated in Figure 4.1. Each step is explained below.

#### 4.2.1 Bias field correction

A standard bias field correction method[Tustison et al., 2010] is applied. This alone does not completely remove the bias and standard segmentation methods fail to deal with it.

#### 4.2.2 2-class segmentation

The local-means method proposed in [Hansen and Lauze, 2017, Brox and Cremers, 2009] is used and seeks to minimise the functional

$$\mathcal{E}_{Q}(\mathbf{c},\mathbf{v}) = \frac{1}{2} \sum_{i=1}^{n} \int_{\Omega} g * \left[ (u - c_{i}(\boldsymbol{x}))^{2} \mathbf{v}_{i} \right] (\boldsymbol{x}) d\boldsymbol{x} + \frac{\mu}{2} \sum_{i=1}^{n} \int_{\Omega} |D\mathbf{v}_{i}|^{2} d\boldsymbol{x} \quad (4.2.1)$$

where  $\mathbf{v}_i \in [0, 1]$  is the label field describing the probability of belonging to class *i*. *u* is the bias corrected volume, *g* a moving average kernel and

$$c_i(\boldsymbol{x}) = \frac{(u\mathbf{v}_i) * g(\boldsymbol{x})}{\mathbf{v}_i * g(\boldsymbol{x})}, \quad \boldsymbol{x} \in \operatorname{supp} \mathbf{v}_i.$$
(4.2.2)

This method produces a clear contour of the brain tissue, with strong edges, that makes rough brain extractions possible.

#### 4.2.3 GAC brain extraction

A full brain mask is retrieved from the skull, by initialising a Geodesic Active Contour (GAC) based model[Álvarez et al., 2010, Marquez-Neila et al., 2014] in the center of the segmented brain. Any 'holes' that would occur in the resulting brain mask are filled, using mathematical morphology operations.

#### 4.2.4 3-class segmentation

The retrieved brain mask is used to extract the brain tissue segment of the binary segmentation result and is in turn segmented by a tighter 3-class localmeans segmentation. The three classes are background, CSF, and brain tissue. A single connected brain tissue component is expected, so only the largest is kept.

To refine the segmentation of the internal CSF network, we process as follows. A run of the GAC method on the final brain tissue class provides a more accurate brain mask. It is used to extract the internal CSF network from the final CSF class, removing surface CSF. Small objects are removed in the CSF class to highlight larger networks in the brain.

### 4.3 Results

The framework is run on a sample rat cranium volume affected by significant bias. In Figure 4.2 the result of step 2 and 4 are shown for different brain slices.

Figure 4.3 presents a 3D surface view of the combined brain tissue matter (pink) and CSF (blue) for different viewing angles. In Figures 4.4 and 4.5 respectively, the 3D surfaces of the isolated brain tissue and CSF network can be seen.

## 4.4 Discussion and Conclusion

A framework that satisfactorily segments and extracts brain tissue and internal CSF networks in rat brain MRI scans affected by strong image bias fields has been proposed. The framework combines bias field correction with several passes of segmentation to adequately extract the relevant classes. 3D surface results for brain tissue and the major CSF network as well as axial, sagittal, and coronal planes have been presented.

To further improve on the results, a distance based feature could be incorporated, as the severity of the bias field is correlated with the distance to the coils. Additionally, strong priors on the connectivity of the CSF network will be integrated in the model. While the framework introduces more parameters, we believe the improvements in extracting high quality tissues from medical imaging affected by strong bias fields are highly beneficial.

#### 4. RAT BRAIN EXTRACTION AND SEGMENTATION



**Figure 4.2:** The rows highlight: The original image, showing a severe bias field; the bias field corrected image, which still has quite a bit of bias left; The final segmentation and extraction of the example MRI using the proposed framework. Each column provides the axial, sagittal, and coronal planes for the same brain. Brain tissue is light gray while CSF is is dark gray.



**Figure 4.3:** The columns show different viewing angles of the brain by rotating 180 degrees around the corresponding axis of the resulting brain. The rows from top to bottom show the x, y, and z direction for the same brain segmentation. Brain tissue is pink, while CSF is light blue.

#### 4. RAT BRAIN EXTRACTION AND SEGMENTATION



**Figure 4.4:** The columns from left to right show the negative and positive direction of the brain tissue (pink) segmentation, resulting from step 4. The rows from top to bottom show the x, y, and z direction for the same brain segmentation.



**Figure 4.5:** Each column shows two different angles of the extracted major CSF network (light blue) segmentation, from running our framework. The rows from top to bottom presents: The raw CSF segmentation; The internal CSF network after removing surface CSF; The internal CSF network after having removed small CSF objects. When rotating the volume in a 3D visualiser the network structure and symmetry is more apparent.

# 5 Automatic Brain Segmentation using Shape Priors

The work presented in this chapter is based on a manuscript<sup>†</sup> submitted and currently under review at the *International Society for Magnetic Resonance in Imaging*.

<sup>†</sup> Hansen, J. D. K. et al. (2019). Automatic Brain Segmentation Framework for Bias Field Rich Cranial MRI Scans of Rats and Mice via Similarity Invariant Shape Priors.

# Automatic Brain Segmentation Framework for Bias Field Rich Cranial MRI Scans of Rats and Mice via Similarity Invariant Shape Priors

Jacob Daniel Kirstejn Hansen<sup>1</sup>, François Lauze<sup>1</sup>, Sune Darkner<sup>1</sup>, Julia M. Huntenburg<sup>2</sup>, Kristian Nygaard Mortensen<sup>3</sup>, Simon Sanggaard<sup>3</sup>, Helene Benveniste<sup>4</sup>, and Maiken Nedergaard<sup>3,5</sup>

<sup>1</sup>Department of Computer Science, University of Copenhagen,
 <sup>2</sup>Champalimaud Centre for the Unknown, Champalimaud Research,
 <sup>3</sup>Center for Translational Neuromedicine, University of Copenhagen,
 <sup>4</sup>Anesthesiology, Yale School of Medicine, Yale University,
 <sup>5</sup>Center for Translational Neuromedicine, University of Rochester

**Abstract.** This abstract presents an extension to our previous work for the extraction of rat brain tissue and internal cerebrospinal fluid networks in MR imaging of rat crania that display severe bias fields. This work contributes automation and robustness for the skull extraction module by introducing an automatic similarity invariant shape prior segmentation method. We demonstrate the capabilities of our framework on both rat brain as well as mouse brain data, using the same minimal number of rat brain priors.

#### 5.1 Introduction

MR imaging using surface receiver coils are in general affected by strong image bias fields, for which only partial correction is available. This makes subsequent feature extraction and analysis a difficult task. In a previous work 1 we presented a semi-automatic framework that was able to segment and extract rat brain tissue and the cerebrospinal fluid (CSF) networks. In this work we improve on the robustness of the method and make it fully automatic by introducing segmentation shape priors which, by construction, are invariant to rotation, scale, and translation. We show that our improved method obtains very good segmentation results, while automating the process and limiting the number of free parameters.


**Figure 5.1:** Flowchart diagram of the proposed pipeline, with shape prior segmentation. Input and output data is indicated by the arrows.

#### 5.1.1 Animals

Source 1 (identical to our previous work[Hansen et al., 2018]): Unaffected male rats of the Wistar Kyoto stain (WKY) and spontaneously hypertensive rats (SHR) were obtained from Charles River, Germany. Separate groups of rats were scanned at two age-ranges: young (7-9 weeks old; WKY7, n = 11; and SHR7, n = 8) and young adults (19-21 weeks WKY19, n = 8; and SHR19, n= 9). All treatments and imaging were performed according to protocols approved by the IACUC, and according to a protocol approved by the University of Copenhagen animal experimentation committee.

Source 2: Male C57Bl6 mice (7-8 weeks old) were induced into deep anesthesia with 5% isoflurane and maintained under 2% isoflurane during the MRI experiments. All animal experiments were preapproved by the institutional and national authorities and were carried out according to European Directive 2010/63.

#### 5.1.2 MRI

Source 1: All MRI investigations was conducted on a 9.4 T magnet (Bruker Biospec 9.4/30 USR) interfaced to a Bruker Advance III console and controlled by Paravision 5.1 software (Bruker BioSpin) at the Panum Institute, University of Copenhagen. Imaging was performed with an 86 mm3 volume resonator and a surface quadrature array receiver coil. Images were acquired with the spoiled gradient FLASH3D sequence (TE: 4 ms, TR: 15 ms, NA: 3; matrix:  $128 \times 128 \times 128$ , voxel size:  $0.24 \times 0.24 \times 0.26$  mm, FA: 15°, scan time: 4:05 min).

Source 2: MRI was performed on a 9.4T Bruker BioSpec scanner controlled by Paravision 6.0.1 software, using a 86 mm3 quadrature resonator for transmittance and a 4-element cryoprobe for reception. Images were acquired with a 3D multi-gradient echo sequence (16 echos, TE1: 2.2 ms,  $\Delta$ TE: 2.2 ms, TR: 100 ms, matrix:  $180 \times 90 \times 115$ , voxel size:  $0.11 \times 0.1 \times 0.1$  mm, FA: 60°, scan time: 12:16 min). Processing was performed on the mean across all 16 echo times.

### 5.2 Methods

Our pipeline is modified from [Hansen et al., 2018] and can be seen in Figure 5.1. It consists of four steps: 1) bias correction, 2) a preliminary 2-classes segmentation, robust to remaining bias, 3) a segmentation with a proposed kernel density shape prior to automatically detect and segment the brain / CSF segments, and 4) a 3-classes (background, brain tissue and CSF) segmentation. Step 1, 2, and 4 are identical to [Hansen et al., 2018].

#### 5.2.1 Brain shape prior segmentation

In this step, we combine a Chan-Vese data term and a similarity invariant shape prior, built from n training / template shapes. We have only used n = 4 training shapes here. We use the following shape prior based segmentation energy

$$\mathcal{E}(\mathbf{v}) = \frac{1}{2} \int_{\Omega} ((u - c_1)^2 \mathbf{v} + (u - c_2)^2 (1 - \mathbf{v}) d\mathbf{x} + S(\mathbf{v})$$
(5.2.1)

where

$$S(\mathbf{v}) = -\log \sum_{i=1}^{n} e^{-\frac{d^2(\mathbf{v}, \mathbf{v}_i)}{2\rho^2}}$$
(5.2.2)

is the contribution from the kernel density estimated collection of training shapes  $\mathbf{v}_i$ .  $d^2$  is a pseudo squared distance measure.



**Figure 5.2:** Brain extraction comparison between the former model 1 (left) and our proposed shape prior segmentation (middle, data from Figure 5.3). The rightmost figure shows the left brain overlaid the middle one. Here red corresponds to agreement, yellow is an over-segmentation and teal is an under-segmentation with respect to our shape prior segmentation. The shape prior provides a much better match, while being automated and improving robustness to parameter choices as well.

# 5.3 Results

We ran our brain segmentation on rat brain data from source 1 and on mouse data from source 2. The results for step 1-3 are illustrated in Figure 5.3, while a comparison with our former brain extraction routine can be found in Figure 5.2, The final segmentation result is shown in Figure 5.4. Figure 5.5 highlights our brain extraction method using rat data for training, but running on mouse data and compares it to that of AFNI's 3D skullstrip.

# 5.4 Discussion and Conclusion

We have proposed an automatic framework for rat brain and internal CSF network segmentation for images affected by strong bias fields, using a brain shape prior. It is shown to work well for data coming from other sources than the one used for training our shape prior. To improve the results further, extra training shapes representing better shape variability should be incorporated.

#### 5. Automatic Brain Segmentation USING Shape Priors



**Figure 5.3:** Shape prior contribution: The first row is the process and result of the bias field correction [Tustison et al., 2010] and 2-classes local mean segmentation of [Hansen and Lauze, 2017]. The second row illustrates the shape prior segmentation process from left to right: 1) initialisation of the shape gradient, 2) shape gradient at iteration 1, 3) final shape gradient, and 4) final shape prior segmentation as a posterior. The last row shows the segmentation improvement of running a Geodesic active contour method - this step helps capture shapes previously unseen in the priors.



**Figure 5.4:** Step 4 of the pipeline: Final 3-classes segmentation for the brain tissue and CSF network results. The columns indicate background (left), brain tissue (middle left), CSF (middle right), and cropped data by the shape mask from the previous step (right). First row: initialisation from the 2-classes segmentation step. Second row: Final segmentation results. All figures show labels as posteriors, where red indicates belonging to the class and vice versa for blue. The rightmost figure in the second row shows thresholded labels for brain tissue and overlaid CSF.

#### 5. Automatic Brain Segmentation USING Shape Priors



**Figure 5.5:** Data from a bias field corrected MR mouse brain volume from source 2 (left), segmentation obtained by AFNI's 3D skullstrip (middle), and our shape prior segmentation result (right). Both segmentation masks have been overlaid on the original data to highlight the results.

# 6 Invariant Shape Regularisers for Object Segmentation

The work presented in this chapter is based on a manuscript<sup>†</sup> submitted and currently under review at the *International Conference on Scale Space and Variational Methods in Computer Vision*.

<sup>†</sup> Hansen, J. D. K. and Lauze, F. (2019). Segmentation of 2D and 3D objects with intrinsically similarity invariant shape regularisers.

# Segmentation of 2D and 3D objects with intrinsically similarity invariant shape regularisers

J. D. K. Hansen, F. Lauze, {jdkh,francois@di.ku.dk}

Department of Computer Science, University of Copenhagen

**Abstract.** This paper presents a 2D and 3D variational segmentation approach based on a similarity invariant, i.e., translation, scaling and rotation invariant shape regulariser. Indeed, shape moments of order up to 2 for shapes with limited symmetries can be combined to provide a shape normalisation for the group of similarities. In order to obtain a segmentation objective function, a two-means or two-local-means data term is added to it. Segmentation is then obtained by a standard gradient descent on it. We demonstrate the capabilities of the approach on a series of experiments, of different complexity levels. We specifically target rat brain shapes in MR scans, where the setting is complex, because of bias field and complex anatomical structures. Our last experiments show that our approach is indeed capable of recovering brain shapes automatically.

#### 6.1 Introduction

With the advent of 3D imaging devices, especially X-ray computerised tomography and magnetic resonance imaging, there has been a growing need for 3D segmentation methods that can handle a large variety of signals. Images produced from these modalities can show content with various degrees of complexity and structures, from almost fully random phases to highly structured data, for example in medical imaging when imaging different types of tissues and organs. In this work, we are interested in the latter, where segmentation targets specific structures characterised by a shape distribution.

In this paper, We develop a variational segmentation–matching approach which incorporates shape priors, and is at the same time robust to specific acquisition problems. In addition to noise, MR images are in general corrupted by a bias field coming from combinations of dropoff effects, as seen in Figure 6.1. For this we will use a data fidelity term robust to noise and bias



Figure 6.1: Slice of an MRI scan of a rat cranium (a), 3D brain segmentation (b).

field. We will discuss it in Section 6.2. Our emphasis in this paper, is however more on shape. We will first briefly review related ideas of shape priors in the next paragraph.

In a series of works, [Kendall, 1984] defined shape spaces based on equivalence classes of finite dimensional collections of points, so as to factor out scale, position, and pose. For shape segmentation, the Active Shapes Models [Cootes et al., 1995] is the seminal work on representation of shape distributions and shape priors and since then a large body of literature has been generated. A complete review is however out-of-scope here. An operational representation of shape distribution requires in general two main ingredients: representation of geometric objects in  $\mathbb{R}^2$  or  $\mathbb{R}^3$  and identifications of classes of such objects as a unique shape, in general via specific group actions. Of course these two ingredients are rarely independent of each others. Typically, shape priors and regularisers are built from similarity criteria between forms or some of their extracted features, which are invariant by the group operation. One general way is to introduce explicit minimisation over the transformation group in the prior definition. In this line of work, [Chan and Zhu, 2005] explicitly minimise a shape term in 2D over the similarity group S(2) of scaling, rotation and translation in  $\mathbb{R}^2$ . A similar idea, within a levelset framework, and competing priors, was used in [Fussenegger et al., 2006]. [Mezghich et al., 2014] use properties of the Fourier-Mellin transform cross-spectrum w.r.t. rotation and scaling to segment-match shapes from one or more training shapes. [Wang et al., 2015] define an affine matching and segmentation, based on the action of the affine group on a class of shape representing functions

called interior-points-to-shape relations.

Another consists of finding *canonical forms* i.e., special representatives or features. On special interest are the representations or features which exhibit invariance by group action. Such representatives might be complex to define, however that is our approach in this work. 2D rotations were already present in the Diffusion Snakes of [Cremers et al., 2002], which incorporates elements of active shapes. Our work goes in a slightly different direction and extends the work of [Cremers et al., 2003] to incorporate rotations. [Foulonneau et al., 2006] extract affine canonical moments, i.e., normalised moments which are invariant by an affine transformation in 2D, and define invariant shape dissimilarity from them.

The rest of this paper is organised as follows. In Section 6.2, we start with mathematical preliminaries. We look at invariant shape terms, and deduce a shape prior / regulariser term. As mentioned above, we will also discuss the chosen data fidelity term. Together they form the cost function of the method. Our algorithm is essentially a gradient descent on the cost function and we derive its first variation and gradient in Section 6.3. Validation of the capabilities and inherent transform invariance of our proposed model is done experimentally in Section 6.4. Finally, Section 6.5 discusses and concludes the proposed model.

# 6.2 Derivation of the model

In this section we start by introducing notations and elementary points about shapes and some types of canonical forms, that allows us to define invariant shape priors. We then introduce our objective, containing both data fidelity terms and shape regularisers.

**Normalisers and Canonical Forms.** If *G* is a group and  $\mathcal{T}$  a set with a (left) *G*-action, a *G*-normaliser is a mapping  $\alpha : \mathcal{T} \to G$  such that  $\alpha(g \cdot A) = \alpha(A)g^{-1}$ . It clearly satisfies  $\alpha(\alpha(A) \cdot A) = \operatorname{id}_G$ , thus the name. Its *group*-inverse (not to confuse with inverse mapping)  $\alpha(-)^{-1} : A \mapsto \tau(A) = \alpha(A)^{-1} \in G$  is just a *G*-covariant mapping. Given a mapping  $F : \mathcal{T} \to \mathcal{U}$ , the mapping  $\tilde{F} : A \mapsto F(\alpha(A) \cdot A)$  is *G*-invariant. Indeed, if  $B = g \cdot A$ , clearly  $\alpha(B) \cdot B = \alpha(A) \cdot A$ . For this reason, normalisation provides *canonical forms* for the *G*-set  $\mathcal{T}$ .

In the sequel we exhibit normalisers for shape related sets and certain Lie subgroups G of the affine group, Aff(n) of invertible linear transforms and translations, with n = 2 or 3. The main subgroup of interest is the subgroup of similarities,  $\mathbf{S}(n) \simeq (\mathbb{R}^*_+ \times SO(n)) \ltimes \mathbb{R}^n$ , of (positive) scalings, rotations and translations. To us, an *object* is a compact set of  $\mathbb{R}^n$ , with non-empty interior and regular boundary so as to be able to compute shape derivatives (see [Delfour and Zolésio, 2001]). We will denote this set by S. The affine group acts naturally on S and a G-shape is the orbit of an object under the action of G. More generally, we will consider the set  $\mathcal{F}$  of integrable functions with

compact support, non-empty interior, and values in [0, 1], so that an object A is naturally represented by its characteristic function,  $\chi_A$ . G acts on  $\mathcal{F}$  by  $g \cdot v = x \mapsto v(g^{-1}x)$ .

For the group of translation and/or scaling,  $G = \mathbb{R}^*_+ \ltimes \mathbb{R}^n$ , normalisers are easy to construct from centred moments up to order 1 or 2. For instance, set  $\mu_0(A) = \int_A dx$  its volume and  $\mu_1(A) = \mu_0(A)^{-1} \int_A x \, dx$  and define the transformation  $\tau(A) = (\mu_0(A), \mu_1(A)) : x \mapsto \mu_0(A)x + \mu_1(A)$ . Then  $v \mapsto$  $\tau(A)^{-1}$  is a normaliser. Scale normalisation can also be performed via the standard deviation form

$$\sigma(A) = \left(\frac{1}{\mu_0(A)} \int_A |x - \mu_1(A)|^2 \, dx\right)^{\frac{1}{2}}.$$
(6.2.1)

Setting  $\gamma(A) = (\sigma(A), \mu_1(A)), \gamma^{-1}$  is the normaliser used in [Cremers et al., 2003]. Extending this low-order moments approach to similarities necessitates a few restrictions.

To extend it to the group of similarities, we use the object's centred second order moment, i.e., the mapping

$$A \to SPD(n), \quad A \mapsto \Sigma(A) = \frac{1}{\mu_0(A)} \int_A (x - \mu_1(A))(x - \mu_1(A))^T dx$$
 (6.2.2)

with SPD(n) the space of symmetric positive-definite matrices of order n, and  $\sigma(A)^2 = \text{Tr}(\Sigma(A))$ . A general eigenvalue/eigenvector decomposition  $\Sigma(A) = R_A \Lambda_A R_A^T$  with  $R_A \in O(n)$  and  $\Lambda_A$  diagonal, is the solution of a system of analytic equations. Restricted to the open subset  $U \subset SPD(n)$ with *distinct eigenvalues*, there are finitely many solutions for  $R_A$ ,  $(n!2^n)$ . The choice of one of these solutions, still denoted by  $R_A$ , extends to U via the implicit function theorem [Magnus, 1985]. Choose one of the branches for which  $R_A \in SO(n)$  (there are  $n!2^{n-1}$  such choices, and can be reduced to  $2^{n-1}$ if eigenvalues are sorted). Restrict also the object space to  $S_d = \Sigma^{-1}(U)$ , it is a  $\mathbf{S}(n)$ -invariant subspace of S and  $A \mapsto \tau_A = (\sigma(A), R_A, \mu_1(A))$  is well-defined from what precedes, and  $\mathbf{S}(n)$ -covariant; it is straightforward to check that for  $g = (s, S, t) \in \mathbf{S}(n)$   $\tau_{g \cdot A} = g\tau_A$  as

$$\Sigma(g \cdot A) = s^2 S \Sigma(A) S^T = s^2 S R_A \Lambda_A R_A^T S^T, \quad \mu_1(g \cdot A) = g \mu_1(A).$$

Thus  $\alpha(A) = \tau_A^{-1}$  is a  $\mathbf{S}(n)$ -normaliser on  $\mathcal{S}_d$ .

Note that this construction based on moments of order up to two cannot, in some sense, extend further to the whole affine group due to "too many square roots" for SPD matrices, but we will not discuss it in this paper.

**Shape invariant dissimilarities and shape regulariser.** Given a *model object*  $A_0$ , which we assume normalised, we choose dissimilarities of the form

$$d^{2}(A, A_{0}) = \int_{\mathbb{R}^{n}} L(\chi_{\alpha(A) \cdot A}(x), \chi_{A_{0}}(x)) dx$$
  
= 
$$\int_{\mathbb{R}^{n}} L(\chi_{A}(\tau_{A}x), \chi_{A_{0}}(x)) dx$$
 (6.2.3)

69

with L(r, s) a positively  $C^1$ -function with L(r, s) = 0 if r = s so that the above integral is in fact an integral over  $(\alpha(A) \cdot A) \cup A_0$  and is finite. In the sequel we work with with  $L(r, s) = (r - s)^2$ . From the discussion in the previous paragraph, this is indeed an S(n)-invariant measure. To build a shape prior from it, we follow [Cremers et al., 2003] and build a kernel density estimator from training data made of normalised model shapes  $A_1, \ldots, A_N$ ,

$$p(A|A_1,...,A_N) \propto F(A,A_1,...,A_N) = \sum_{i=1}^N e^{-\frac{d^2(A,A_i)}{2\rho^2}}, \text{ for a } \rho > 0$$

. This density is well defined if we restrict it to a finite dimensional subspace of  $S_d$ , otherwise the normalising constant

$$c(A_1,\ldots,A_N) = \int_{A_d} F(A,A_1,\ldots,A_N) dA$$

is in general undefined. However we do not need a definite prior, we need the well defined regulariser

$$\mathcal{E}_{S}(A) = -\log F(A, A_{1}, \dots, A_{N}) = -\log \left(\sum_{i=1}^{N} e^{-\frac{d^{2}(A, A_{i})}{2\rho^{2}}}\right).$$
(6.2.4)

**Data fidelity term and proposed formulation.** Different data fidelity terms can be used in conjunction with the regulariser. Two data terms are used in this work. The first one is a classical global 2-means term, while the second is a local means term, better suited to MR data. In both cases, the signal is represented by a function  $u : \Omega \in \mathbb{R}^n \to \mathbb{R}$  and we want to partition the image domain  $\Omega$  into two regions, A and  $\Omega \setminus A$ . The two-means term is

$$\mathcal{E}_{D_g}(A, c_1, c_2) = \frac{1}{2} \int_{\Omega} \left( (u - c_1)^2 \chi_A + (u - c_2)^2 \chi_{\Omega \setminus A} \right) \, dx \tag{6.2.5}$$

with  $c_1, c_2 \in \mathbb{R}$ . MR image data suffers among others from bias fields, that are only partially corrected for by standard techniques. We use ideas from [Hansen and Lauze, 2017] and propose the two-local-means term

$$\mathcal{E}_{D_l}(A, c_1, c_2) = \frac{1}{2} \int_{\Omega} g * \left[ (u - c_1(x))^2 \chi_A + (u - c_2(x))^2 \chi_{\Omega \setminus A} \right] (x) \, dx \quad (6.2.6)$$

with  $c_1$  a smoothed version of u on A,  $c_2$  a smoothed version of u on  $\Omega \setminus A$ and g a smoothing kernel, e.g. a Gaussian or nearest neighbours (NN) kernel, which we assume to be even symmetric. Finally we propose to minimise the following criterion

$$\mathcal{E}(A, c_1, c_2) = \mathcal{E}_D(A, c_1, c_2) + \kappa \mathcal{E}_S(A)$$
(6.2.7)

with  $\kappa > 0$  a trade-off parameter between data fidelity and shape.

**From shapes to label fields.** In all the previous constructions, *A* can be replaced by  $\chi_A$  or a v in  $\mathcal{F}$ . Moments are extended trivially to  $\mu_0(v)$ ,  $\mu_1(v)$ ,  $\Sigma(v)$ , and  $\sigma(v)$ . This case and  $\mathcal{F}_d$  represent the subspace of functions in  $\mathcal{F}_d$  for which  $\Sigma(v)$  has distinct eigenvalues.

# 6.3 First Variations and Optimisation

In this section, we compute the first variation for (6.2.7) with a bit more focus for the term (6.2.4) as we follow [Hansen and Lauze, 2017] for the first term.

**Derivative of shape regulariser.** We start by stating moment derivatives  $d_v \mathbf{m}.w$  of the moment  $\mathbf{m}(v), v \in \mathcal{F}$  for a "variation" w of v, i.e., functions  $\mathbf{v}(t)$  for which  $\mathbf{v}(0) = v$  and  $\dot{\mathbf{v}}(0) = w$ . The moment derivatives are defined as  $\frac{d}{dt} \mid_{t=0} \mathbf{m}(\mathbf{v}(t))$ , for the moments  $\mu_0(v), \mu_1(v)$  and  $\Sigma(v)$ . Then we compute  $d_v \tau.w$  instead of the derivative of the normaliser  $d_v \alpha.w$ . We use the notation  $\langle f, g \rangle$  for the integral  $\int_{\mathbb{R}^n} f(x)g(x) dx$ . Straightforward calculations give

$$d_{v}\mu_{0}w = \langle 1, w \rangle, \quad d_{v}\mu w = \left\langle \frac{\cdot - \mu_{1}(v)}{\mu_{0}(v)}, w \right\rangle, \quad (6.3.1)$$

$$d_{v}\Sigma.w = \left\langle \frac{(\bullet - \mu_{1}(v))(\bullet - \mu_{1}(v))^{T}\Sigma(v)}{\mu_{0}(v)}, w \right\rangle, \qquad (6.3.2)$$

$$d_v \sigma.w = \left\langle \frac{|\cdot - \mu_1(v)|^2 - \sigma^2(v)}{2\sigma(v)\mu_0(v)}, w \right\rangle$$
(6.3.3)

To compute  $d_v R.w$ , we use classical formulas for the derivatives of eigenvectors and eigenvalues of a symmetric matrix. If  $e_1, \ldots, e_n$  and  $\lambda_1, \ldots, \lambda_n$  are the eigenvectors and associated eigenvalues of  $\Sigma(v)$ ,

$$d_{v}R.w = \frac{1}{\mu_{0}(v)} \left\langle \left[ \sum_{j \neq i} \frac{e_{i}^{T}(\cdot - \mu_{1}(v))e_{j}^{T}(\cdot - \mu_{1}(v))e_{j}}{\lambda_{i} - \lambda_{j}} \right]_{i=1}^{n}, w \right\rangle$$
(6.3.4)

Putting it together, one get

$$(d_v \tau . w)(x) = ((d_v \sigma . w)R(v) + \sigma(v)d_v R . w) x + d_v \mu_1 . w.$$
(6.3.5)

We set  $v_i = \chi_{A_i}$  and provide the first variation of the dissimilarity measure  $\mathcal{L}^i(v) = \int L(v(\tau_v x), v_i(x)) dx$ . A complete computation is not feasible within the page limit and of limited interest anyway.

$$d_{v}\mathcal{L}^{i}.w = \int L'(v(\tau_{v}x), x) \nabla_{\tau_{v}x} v \cdot (d_{v}\tau.w)(x) \, dx + \int L'(v(\tau_{v}x), x) w(\tau_{v}x) \, dx$$
(6.3.6)

71

The first integral can be rewritten as

$$\frac{1}{\mu_0(v)} \left[ \int_{\mathbb{R}^n} \left\{ \left\langle \int S(y,v) y^T \, dy, \frac{|x-\mu_1(v)|^2 - \sigma(v)^2}{2\sigma(v)} R(v) + \sigma(v) \Xi(x,v) \right\rangle_F + \left( \int S(y,v) \, dy \right)^T (x-\mu_1(v)) \right\} w(x) \, dx \right]$$
(6.3.7)

where we have set

$$S(x,v) = L_r(v(\tau_v x), x) \nabla_{\tau_v x} v,$$
  
$$\Xi(x,v) = \left[ \sum_{j \neq i} \frac{e_i^T(x - \mu_1(v))e_j^T(x - \mu_1(v))e_j}{\lambda_i - \lambda_j} \right]_{i=1}^n.$$

The second integral in (6.3.6) becomes, after a mere change of variables,

$$\frac{1}{\sigma(v)} \int L_r(v(x), \tau_v^{-1}x) w(x) \, dx.$$
(6.3.8)

From (6.3.7) and (6.3.8), one obtain  $d_v \mathcal{L}^i . w = \langle G_S^i(v), w \rangle$  where

$$G_{S}^{i}(v) = \frac{1}{\mu_{0}(v)} \left\langle \int S(y,v)y^{T} dy, \frac{|x-\mu_{1}(v)|^{2} - \sigma(v)^{2}}{2\sigma(v)} R(v) + \sigma(v)\Xi(x,v) \right\rangle_{F} + \frac{1}{\mu_{0}(v)} \left( \int S(y,v)dy \right)^{T} (x-\mu_{1}(v)) + \frac{1}{\sigma(v)} L_{r} \left( v(x), \frac{R(v)^{T}(x-\mu_{1}(v))}{\sigma(v)} \right).$$
(6.3.9)

Finally, using standard differentiations rules, one get that

$$d_{v}\mathcal{E}_{S}.w = -\frac{1}{2\rho^{2}} \left\langle \frac{\sum_{i=1}^{N} e^{-\frac{\mathcal{L}^{i}(v)}{2\rho^{2}}} G_{S}^{i}(v)}{\sum_{i=1}^{N} e^{-\frac{\mathcal{L}^{i}(v)}{2\rho^{2}}}}, w \right\rangle = \langle G_{S}(v), w \rangle$$
(6.3.10)

**Derivatives of data fidelity terms.** Using a variation w over  $\Omega$ , a trivial calculation gives, for (6.3.12)

$$d_v \mathcal{E}_{D_g} . w = \left\langle 2u(c_2 - c_1) + (c_1 - c_2)^2, w \right\rangle_{L^2(\Omega)} = \left\langle G_{D_g}, w \right\rangle_{L^2(\Omega)}$$
(6.3.11)

and is a bit more complex for (6.2.6)

$$d_{v}\mathcal{E}_{D_{l}}.w = \left\langle u^{2}g * \chi_{\Omega} - 2u * g(c_{2} - c_{1}) + g * (c_{1}^{2} - c_{2}^{2}), w \right\rangle_{L^{2}(\Omega)} = \left\langle G_{D_{l}}, w \right\rangle_{L^{2}(\Omega)}.$$
(6.3.12)

**Optimisation.** We can be tempted to identify the gradient as  $-\kappa G_s(v) - G_D(v)$  with  $G_D = G_{D_q}$  or  $G_{D_l}$ . Unfortunately,  $G_S$  contains non local terms, this

means that constrains on the support of w must be enforced. When v is a characteristic functions and shape derivatives [Delfour and Zolésio, 2001] are used, these variations are distributions supported along  $\partial A$  and have the form  $\mathbf{v}\delta_{\partial A}$  or its normal component along  $\partial A$ ,  $(\mathbf{v} \cdot \mathbf{n}_{\partial A})\mathbf{n}_{\partial A}\delta_{\partial A}$  with  $\mathbf{v}$  a vector field defined at least in a neighbourhood of  $\partial A$ . Here, instead of imposing a special form to our w we use a simple narrow band approach: we restrict their support to  $(\operatorname{supp} v)_{\varepsilon} = \operatorname{supp} v \oplus B(0, \varepsilon)$ , the dilation of  $\operatorname{supp} v$  by a ball of radius  $\varepsilon$ . In practice  $\varepsilon$  is one or two pixels (voxels) and we implement the projected gradient descent step. The algorithm is described in Algorithm 3. Note that because of the narrow band implementation, the initial estimate (or

Algorithm 3 Sketch of the full algorithm.

- **Input:**  $A_1, \ldots, A_n$  training shapes,  $\rho > 0$  the deviation parameter,  $\kappa > 0$  the data vs. shape trade-off parameter,  $\delta t > 0$  the descent step parameter,  $\eta > 0$  the convergence threshold,  $\zeta \in (0, 1)$  the binarisation threshold,  $v^0$  initial shape function.
- **Output:** Segmentation function / object closed to the training shapes. **repeat**

$$v^{n+1} = \mathcal{P}_{\mathcal{F}}\left(v^n - \delta t\left(\left(G_D(v^n) + \kappa G_S(v^n)\right)|_{(\operatorname{supp} v)_{\varepsilon}}\right)\right)$$
  
until  $\frac{\|v^{n+1} - v^n\|^2}{|\Omega|} < \eta$   
return  $\mathbf{v} := (v^n > \zeta).$ 

its dilation) should overlap the object we want to segment. This may limit this type of approach, however, for the application that we have in mind, rat brain segmentation, we usually start we a very large overlap due to the MR measurement setting.

# 6.4 Experimental Validation

We target MR scans of rats, but there is no ground truth available for these volumes. Therefore we first present two synthetic experiments where only one shape is used, then two experiments on an MR scan where we have segmented five rat brains by a combination of classical variational methods and postprocessing so as to extract and learn brain shapes. We use them to build a shape regulariser and proceed to segment an unseen MR volume. To evaluate segmentation when we have ground truth, we have used the Dice-Sørensen Coefficient (DSC) score:

$$DSC(X,Y) = \frac{2|X \cap Y|}{|X| + |Y|}$$

where  $|\cdot|$  is the cardinality of a set. The binarisation threshold in Algorithm 3 has been fixed to 0.5. There are classical techniques for computing a good

#### 6. INVARIANT SHAPE REGULARISERS FOR OBJECT SEGMENTATION



**Figure 6.2:** 2D experiment. (a): Shape prior. (b): Data to be segmented. (c): Initial shape estimate. (d): Final segmentation.

value for  $\rho$  in (6.2.4), though, in this work we fixed it to  $\rho = 1.5$  and 1 for the experimental and synthetic respectively, as they always provided satisfactory results.

A simple 2D experiment. In this experiment, as a simple validation of the framework, we are given one training shape and an image obtained by a similarity transform, with added noise and some small level of occlusion. The data term used is the two-means term (6.2.5). We illustrate our prior, input data, initial estimate and final segmentation in Figure 6.2. The values for  $\kappa$ ,  $\eta$  and  $\delta t$  in Algorithm 3 are respectively 5.0,  $3 \times 10^{-5}$  and 0.1. The reported DSC is 98.02%. We observe a spurious small contour in (d), which would have

disappeared with a slightly lower threshold. Scaling and rotations necessitate interpolation, which is also a source of inaccuracies. This is likewise present in the other experiments.

**A 3D experiment.** In this experiment, we use a simple training shape, a binarised 3D rendering of the letter "F", while we generated a more complex 3D image and added noise and complex structures to it. Here too, the data term is the two-means term (6.2.5). This is illustrated in Figure 6.3. The values





**Figure 6.3:** Synthetic 3D example. The ground truth (a) is a crenellated 3D 'F' shape. The training shape (b) is a simple 3D 'F' shape with flat faces. (c) is a noisy and cluttered 3D scene (at one of the corners and side of the letter), (d) is the initial shape for segmentation and (e) the final shape.

for  $\kappa$ ,  $\eta$  and  $\delta t$  in Algorithm 3 are respectively 10.0,  $1.5 \times 10^{-6}$  and 0.1. The reported DSC score is 90.7%, which may seem a bit low. Among other things, this is due to the relatively important deviation between the training shape (very flat) and the somewhat crenellated ground truth, as can be observed from Figure 6.3.

**Rat brain segmentation.** First, we start by a 2D experiment, mainly because it makes visualisation easier. Medically annotated data is not available for the

#### 6. INVARIANT SHAPE REGULARISERS FOR OBJECT SEGMENTATION



Figure 6.4: Training shape masks used in 2D segmentation of rotated rat cranium scan.

specific dataset series we worked on, and to our knowledge, no automatic segmentation exists for MR brain scans of rodents. Validation is therefore based on qualitative assessment of the achieved segmentation. Without annotated data, we have used the work in [Hansen et al., 2018] to obtain the brain segmentations. Figure 6.5 shows the effect of the shape regularisation in 2D. The rotated MR slice in Figure 6.5b shows Figure 6.1a after the bias field correction procedure of [Tustison et al., 2010]. It is clear that the bias field has not been fully eliminated. We therefore use the local means data term (6.2.6). In the next subfigure we illustrate the different terms in the objective function. Evolving only the regulariser term should result in a sort of mean training shape. This is illustrated in Figure 6.5a. We use data in the form of a rotated slice from a MR brain, which is represented in Figure 6.5b. The red square contour inside represent the initial segmentation guess. Optimising using only the local two-means data term results in an over-segmentation where part of the skull and other anatomical elements are segmented: this can be seen in Figure 6.5c. Finally, Figure 6.5d shows the obtained segmentation with our approach.

We now illustrate the segmentation in 3D. Using the pipeline of variational and image cleaning methods of [Hansen et al., 2018], we segmented five rat brains from MR scans. With careful postprocessing, three of the segmentations were refined and added to our set of training shapes, providing ten final training shapes. No initial alignment of the training data was performed. As initial guess, we use one of the the training shapes, slightly eroded, with its original pose and scale parameters. The physical setup indeed guarantees that pose and scale parameters for the different scans should vary moderately from scan to scan. Two-dimensional slices for 4 of the original shapes have already been shown in Figure 6.4. We tested our approach on an MR volume which could not be segmented by the approach of [Hansen et al., 2018] (it



**Figure 6.5:** (a): "Mean" training shape computed by minimisation of regulariser. (b): Rotated slice. (c): Local two-means clustering. (d): Segmentation with both terms.

would either incorporate part of the skull or severely under-segment the brain). The values for  $\kappa$ ,  $\eta$  and  $\delta t$  in Algorithm 3 are respectively 2.0,  $1 \times 10^{-5}$  and 2.0. Moreover, the two-local-means term (6.2.6) was used with a very simple  $3 \times 3 \times 3$  local mean kernel g. Some evolution steps are displayed in Figure 6.6. Small angular variations can be seen by observing the relative position of the shape and the reference black plane. In Section 6.4 we visualise the results for different viewing angles.

# 6.5 Conclusion

In this paper, we have proposed a 2D and 3D approach for the segmentation of objects whose distribution of shapes can be approximated via training shapes. This allowed us to construct a regularisation term, within a variational framework, which is invariant under similarity transforms. We coupled the regulariser with a classic or robust data attachment term, to obtain segmentation objective functions and computed their gradients analytically. We have demonstrated this approach in several scenarios, particularly on MRI scans of rat crania. We are working on a more objective evaluation of our approach for rat brains, which is somewhat complicated in the absence of ground truth or accepted gold standard. We have used projected gradient descent on simple shape representations via relaxed characteristic functions. More efficient optimisation methods could be considered, and other types of shape representation could be used efficiently, especially for shape distributions where topology does not change.

## 6. INVARIANT SHAPE REGULARISERS FOR OBJECT SEGMENTATION





Figure 6.6: Segmentation evolution in 3D after 0 (a), 24 (b), 49 (c), and 75 (d) iterations.

# Acknowledgements

J. Hansen and F. Lauze acknowledge Sune Darkner<sup>1</sup>, Kristian Nygaard Mortensen<sup>2</sup>, Simon Sanggaard<sup>2</sup>, Helene Benveniste<sup>3</sup>, and Maiken Nedergaard<sup>24</sup> for providing the data.

<sup>&</sup>lt;sup>1</sup>Department of Computer Science, University of Copenhagen

<sup>&</sup>lt;sup>2</sup>Center for Translational Neuromedicine, University of Copenhagen

<sup>&</sup>lt;sup>3</sup>Anesthesiology, Yale School of Medicine, Yale University

<sup>&</sup>lt;sup>4</sup>Center for Translational Neuromedicine, University of Rochester



**Figure 6.7:** Final segmentation results of our proposed model for difference viewing angles. (a) and (d) shows the *x* and -x direction respectively, *y* and -y directions in (b) and (e) and *z* and -z in (c) and (f).

# 7 Segmentation by Information Theory and Edge Penalty

The work presented in this chapter is based on an ongoing work by J. D. K. Hansen, F. Lauze, and Y. Dong.

# Information Theory-Driven Two-Phase Segmentation with Edge Penalty

J. D. K. Hansen<sup>†</sup>, F. Lauze<sup>†</sup>, {jdkh,francois@di.ku.dk}, Y. Dong<sup>‡</sup> {yido@dtu.dk}

†Department of Computer Science, University of Copenhagen ‡Department of Applied Mathematics and Computer Science, Technical University of Denmark

**Abstract.** This chapter presents a variational two-phase segmentation approach for recovery of background and material segments in weakly structured images. We present data fidelity terms based on Shannon entropy and symmetrised Kullback-Leibler divergence and a specialised weight function of classical length penalisation. The derivation of the length term is classical and the data term is done by computing the shape gradients. We demonstrate promising preliminary capabilities by individual experiments, for the data and weighted length terms, on synthetic and real examples.

# 7.1 Introduction

Image segmentation is an unavoidable step in most smage processing tasks. With the increasing amount of various sources and content, naturally the artefacts that follow need to be addressed. The field of X-ray computerised microand nanotomography (X-ray  $\mu$ CT and nCT) of various samples is growing very fast with the introduction of new more powerful synchrotron facilities, such as MAX IV. Our application for this work will consist of the analysis of reconstructed geological tomogram samples, with no expert annotations available.

Observed rock samples are typically not too structured in terms of shapes, but are in large assumed piecewise constant and homogenious with sharp edges. This, however, is not always the case. It is common to observe multiple material classes in the form of crystals and different sediments or different artefacts that effectively produce two background modes or a new material phase. For optimal segmentation results, choosing the correct number of classes is necessary but tedious work to do manually, in particular when faced with an abundance of data and in 3D. Estimating the number of classes is, on the other hand, an old well-understood problem and the amount of work in this field is plentyful, see for instance [Silverman, 1981, Chao, 1984, Cheng, 1995, Fukunaga and Hostetler, 1975]. These methods however are not always appropriate and often introduce additional parameters to tune for. For example, In the case of approaches based on [Parzen, 1962, Rosenblatt, 1956], the bandwidth parameter needs to be chosen, and there are no general rules to do so. In our application, hidden smaller classes and significant artefacts make the process of estimating classes even more challenging.

For most cases in computing fluid dynamics and geophysical parameters of porous structures, only the segmentation accuracy of the void region is essential to the task. In this work, we therefore propose a method that compares the background region to all others, effectively transforming a complicated multiphase problem into a two-phase problem, without neglecting essential accuracy, but hopefully improve on it. The main observation here, is that that while material classes are relatively close to each other, in terms of intensity values, the background class is more well-separated and less complex in terms of intensity variation.

Edges constitute an important segmentation cue in X-ray n-CT modalities, but they may present a challenging artifact as seen in Figure 7.1. This artefact affects edges, effectively producing a significant increment following a dip in the intensity profile over an edge from the background to foreground. Additionally, this artefact does not affect all edges in the data, only a subset. Although akin to, this artefact differs from classical beam hardening effects of X-ray CT tomography, as the datasets here have been collected from monochromatic beams. We believe the artefacts might stem from either X-ray spreadding from the surface of the crystal monochromator, a component of refraction due to a coherrent X-ray beam, or because the synograms have been reconstructed as if sampled from a parallel beam, which is not the case.

In this work, we present a model that is able to satisfactorily deal with the observed edge artefacts, while still being able to accomodate unaffected and finer in detail edges. We use Shannon entropy[Shannon, 1948] and a symmetrised version of Kullback-Leibler divergence[Kullback and Leibler, 1951] as data term and a special weighted length regularisation. The work that we draw inspiration from is that of [Aubert et al., 2003, Herbulot et al., 2004], where they present a general framework for naturally incorporating shape gradients into active contour methods. They also use an entropy-based information theoretic data term.

The rest of the chapter is organised as follows: In Section 7.2, we introduce the foundations for the proposed model and argue for our individual terms. Section 7.3 presents the associated Euler Legrange equations for each of the terms making up the objective function. Finally Section 7.4 presents individual experiments for our proposed data term and regularisation term and Section 7.5 concludes the work.

# 7.2 Derivation of the model

In the following we provide an overview of the individual components of our proposed objective function. The model consists of two parts; first, a histogram matching term that implicitely serves as a data fidelity term and could take on any image feature in theory. For this work, we assume histograms of gray-level image intensities. Secondly we introduce a weighted length term that favours true edges over false (artefact prone) edges.

An obvious application of the histogram matching term for image segmentation is that of video sequences, where a user wishes to e.g. track object movements in time. Matching histograms over the time scale would be the natural choice as we don't expect objects to dissappear or deviate much from their previous location one frame back in time. The same inference can be applied in our field of porous media, as we assume a significant degree of connectivity of the pore- and material space. Consequently, we expect small deviations in histograms of neighbouring slices, for any directional cut.

As pointed out in Section 7.1, since rock tomography can be complex in the foreground phase, containing a multitude of different materials, our method focuses on the background phase relative to all other phases. The purpose of this work is not a multiphase segmentation, but an accurate two-phase segmentation as the background class often is more well behaved and the target of further analysis. Furthermore, the background class must necessarily consist of, at least, the smallest intensity value mode, effectivily bounding the phase from the left.

We therefore choose to model our background phase as the entropy of a kernel density estimate of the region. Minimising this function alone will converge to a Dirac delta peak, maximising it will consequently result in a uniform distribution. Like [Herbulot et al., 2004], we could model the remaining phases by the entropy as well, but that would result in an oversegmentation of the background phase. Instead we propose to use the negative symmetrised Kullback and Leibler (SKL) divergence, to keep the two phases well-separated, but at the same time avoiding that one class disappears. Minimising the negative SKL divergence, we force the classes to be different, which fits our initial analysis of the background being simple, while the foreground is usually quite complex.

For our application, edges are for the most part nice and regions assumed piecewise constant. However, because of relatively strong edge artefacts, traditional edge sensitive active contour models like [Caselles et al., 1997] are prone to stop before the true edge, producing inaccurate pore space representations.



**Figure 7.1:** Experimental dataset with edge artefacts (a), a crop of an area with clear edge artefacts (b), and the intensity profile of a column perpendicular to an edge (c).

#### 7.2.1 Edge penalised length term

Inspired by the work on geodesic active contours [Caselles et al., 1997] and active shapes [Paragios and Deriche, 2002b] with a modified length term which takes small values along edges, we model our regularisation as the curve formulation

$$E_{reg}(\Omega) = \oint g(\nabla u(c(p))) |c_p| \, dp, \qquad (7.2.1)$$

where c(p) is a parametrisation of  $\partial\Omega$ . Equation (7.3.32) can be considered as a weighted length regularisation, favouring whatever the *g* function models. *C* is a parametrisation of a curve and we propose *g* to be a powerful edge detector, that is still able to respond to traditional edges, but also trickier ones that are prone to amplification / absorption artefacts. A traditional choice is when *g* is derived from the magnitude of the Gaussian smoothed gradient

$$g = \frac{1}{1 + \beta |\nabla_{\sigma_0} u|,}$$

where  $\nabla_{\sigma_0}$  means taking the gradient of the convolution of x with a Gaussian of standard deviation  $\sigma_0$ . But while its robustness to noise and ability to enhance edges is well understood, it fails to destinquish edge artefacts from real edges and to tune the parameters  $\beta$  and  $\sigma_0$  to adapt to the artefacts would mean losing finer detail weak edges, or include noise.

To argue for our choice of g function, we look at the dataset in Figure 7.1a and in particular a crop thereof over the artefacts as seen in Figure 7.1b. Analysing the intensity profile at a perpendicular crossing of an artefact affected edge, we observe the pattern portrayed in Figure 7.1c. We then proceed to design our g-function around this pattern, while still being able to pick up regular edges.

We model our *g* function based on classical filtering techniques for edge enhancement:

$$g(u) = \frac{1}{1 + \beta |\nabla(w_{(\sigma_1)}(2u - k_{(\sigma_2)} * u))|}.$$
(7.2.2)

Here  $\beta$  is a weight parameter,  $w_{(\sigma_2)}$  is a median filtering function, and  $k_{(\sigma_2)}$  is a Gaussian kernel, governed by  $\sigma_1$  and  $\sigma_2$  respectively. The essential part of the functional (7.2.2) is the unsharp masking of the original image u. This edge enhancement filter produces indeed the same edge intensity profile as the artefact we seek to model. Additionally it also enhancing less clear edges that would otherwise have been smoothed out. Since unsharpening the image is a high pass procedure, noise is also amplified greatly, we therefore subject it to a median filtering  $w_{(\sigma_1)}$  step, to maintain edge contrasts, but remove noise. Alternatively, we could alter  $\sigma_2$  in the Gaussian gradient magnitude step, but that would greatly blur edges and we would lose finer details. This approach is similar to that of [Müter et al., 2012], where they use a dual filtering approach to improve image thresholding.

# 7.3 Optimisation

In the following, we will derive the classic Euler Legrange equations for our proposed objective function. For completeness, we start by presenting the general derivation framework we use throughout this section.

#### 7.3.1 Derivation framework

We focus on general region based energy functionals on the form

$$J(\Omega(\tau)) = \int_{\mathbb{R}^N} \varphi(k(x, \Omega(\tau))) \, dx, \tag{7.3.1}$$

 $\psi$  being a vector field used to (at least infinitesimally) transport points and  $\Omega$  an open set that represents the region. Here we let  $\tau$  be the evolution parameter in a dynamic scheme, so that  $\Omega$  is continuously dependent on  $\tau$ . k in our case is a function  $\mathbb{R}^+ \to \mathbb{R}^+$  that describes some probabilistic measure of the region  $\Omega$ .  $\varphi$  is any information theoretic function. We use the shape derivative method to finally obtain the velocity vector **v**that minimises our objective function.

The boundary evolution equation and the first variation is defined as the Eulerian derivative of *J*, with respect to  $\tau$ , in the direction  $\psi$ :

$$\frac{d}{d\tau}J(\Omega(\tau);\psi) \coloneqq \left\langle \frac{d}{d\tau}J(\Omega(\tau);\psi) \right\rangle \coloneqq \lim_{\tau \searrow 0} \frac{J(\Omega(\tau)) - J(\Omega(0))}{\tau}.$$
 (7.3.2)

In the following derivative calculations we will be using the technique of rewriting the integration domain of a functional by use of a characteristic function

$$\chi(x,\Omega(\tau)) = \begin{cases} 1 & \text{if } x \in \Omega(\tau) \\ 0 & \text{if } x \in D \in \mathbb{R}^n \setminus \Omega(\tau) \end{cases}$$
(7.3.3)

Due to the nature of characteristic functions being constant everywhere, except for the boundary  $\partial \Omega(\tau)$  of the specified domain  $\Omega(\tau)$ , the integration domain of characteristic functions are therefore limited to the contour of  $\Omega(\tau)$  and we get that  $\chi$ 's variation is given as the (outward) normal component of the vector field  $\psi$ :

$$\frac{d}{d\tau}\Big|_{\tau=0}\int_{D}\chi(x,\Omega(\tau))\,dx = \int_{D}\frac{d}{d\tau}\Big|_{\tau=0}\chi(x,\Omega(\tau))\,dx = \int_{\partial\Omega(\tau)}(\mathbf{n}\cdot\psi)\,da,\tag{7.3.4}$$

where *a* is the arc-length of the boundary. We now compute the shape gradient of our general region based energy formulation *J*, i.e.  $\frac{d}{d\tau}\varphi(q(\alpha, \Omega(\tau)))$ , where  $\alpha$  is an intensity value in the image

$$\frac{d}{d\tau}\varphi(q(\alpha,\Omega(\tau))) = \frac{\partial}{\partial q}\varphi(q(\alpha,\Omega(\tau)))\left\langle\frac{\partial}{\partial\tau}q(\alpha,\Omega(\tau)),\psi\right\rangle =$$

$$\frac{1}{|\Omega(\tau)|}\int_{\partial\Omega(\tau)}\frac{\partial}{\partial q}\varphi(q(\alpha,\Omega(\tau)))\left[-q(\alpha,\Omega(\tau)) + K(\alpha - I(a))\right](\mathbf{n}\cdot\psi)\,da.$$
(7.3.5)

For a more general form of the shape semi-derivative we refer the reader to [Aubert et al., 2003].

We can also write this as a function of  $\mathcal{G}$ :

$$\frac{d}{d\tau}J(\Omega(\tau);\psi) = \int_{\partial\Omega(\tau)} \mathcal{G}(\alpha,\Omega(\tau)) \cdot (\mathbf{n}\cdot\psi) \, da, \qquad (7.3.6)$$

where

$$\mathcal{G}(\alpha, \Omega(\tau)) = \frac{1}{|\Omega(\tau)|} \frac{\partial}{\partial q} \varphi(q(\alpha, \Omega(\tau))) \left[-q(\alpha, \Omega(\tau)) + K(\alpha - I(a))\right].$$
(7.3.7)

Rewriting (7.3.6) as an inner product, we get

$$\frac{d}{d\tau}J(\Omega(\tau);\psi) = \langle \mathcal{G}(\alpha,\Omega(\tau))\cdot\mathbf{n},\psi\rangle_{L^2(\partial\Omega(\tau))}.$$
(7.3.8)

Now, Cauchy-Schwartz inequality theorem states that for all vectors u and v of an inner product space, it holds that

$$||v \cdot u|| \le ||v|| \cdot ||u||, \tag{7.3.9}$$

and in our case

$$\left\| \left| \frac{d}{d\tau} J(\Omega(\tau); \psi) \right\| = \left\| \left| \left\langle \mathcal{G}(\alpha, \Omega(\tau)) \cdot \mathbf{n}, \psi \right\rangle_{L^2(\partial\Omega(\tau))} \right\| \le \left\| \mathcal{G}(\alpha, \Omega(\tau)) \cdot \mathbf{n} \right\| \cdot \left\| \psi \right\|$$
(7.3.10)

87

To achieve a maximal change in  $\partial\Omega$ , our velocity vector  $\psi$  of steepest descent must necessarily be the one that causes equality of the Cauchy-Schwartz equation. Equality is achieved exactly only when  $\psi$  is parallel to the normalised gradient. We therefore conclude that the velocity vector of steepest descent is

$$\frac{\partial(\partial\Omega)}{\partial\tau} = \psi = -\frac{d}{d\tau}J(\Omega(\tau)) = -\mathcal{G}(\alpha,\Omega(\tau)) \cdot \mathbf{n} =$$

$$\frac{1}{|\Omega(\tau)|} \left(\frac{\partial}{\partial q}\varphi(q(\alpha,\Omega(\tau)))\left[q(\alpha,\Omega(\tau)) - K(\alpha - I(\hat{x}))\right]\right) \cdot \mathbf{n} = \qquad (7.3.11)$$

$$G(\alpha,\Omega(\tau)) \cdot \mathbf{n}$$

#### 7.3.2 Entropy derivative

The differential entropy of a continuous random variable with probability density p is defined as

$$D_E(p) = \int_{-\infty}^{\infty} -\mathbf{p}(\alpha) \log \mathbf{p}(\alpha) \, d\alpha \tag{7.3.12}$$

where  $\alpha$  is an intensity value and **p** denotes the density. In our case, our Parzen window estimate represents **p**:

$$\varphi(q(\alpha, \Omega)) = -q(\alpha, \Omega) \log q(\alpha, \Omega). \tag{7.3.13}$$

We therefore seek to minimise

$$E_E(\Omega) = \int_{\mathbb{R}} -q(\alpha, \Omega) \log q(\alpha, \Omega) \, d\alpha \tag{7.3.14}$$

To obtain the gradient descent direction, we use the derivation framework developed in Section 7.3.1 and compute the functional:

$$G(\alpha, \Omega) = \frac{1}{|\Omega|} \left( \frac{\partial}{\partial q} \varphi(q(\alpha, \Omega)) \left[ q(\alpha, \Omega) - K(\alpha - I(\hat{x})) \right] \right).$$
(7.3.15)

We start by computing the q dependent derivation term

$$\frac{\partial}{\partial q}\varphi(q(\alpha,\Omega)) = \frac{\partial}{\partial q} - q(\alpha,\Omega)\log q(\alpha,\Omega)$$
$$= \frac{\partial}{\partial q} - [q(\alpha,\Omega)]\log q(\alpha,\Omega) - q(\alpha,\Omega)\frac{\partial}{\partial q}[\log q(\alpha,\Omega)] \quad (7.3.16)$$
$$= -\log q(\alpha,\Omega) - 1.$$

Inserting  $\frac{\partial}{\partial q}\varphi(q(\alpha,\Omega))$  in *G*, we arrive at the descent direction

$$\psi_E = \frac{1}{|\Omega|} \left( \int_{\mathbb{R}} (-\log q(\alpha, \Omega) - 1) \left[ q(\alpha, \Omega) - K(\alpha - I(\hat{x})) \right] d\alpha \right) \cdot \mathbf{n}, \quad (7.3.17)$$

which is the direction of steepest descent for our entropy based term.

#### 7.3.3 Symmetrised Kullback-Leibler divergence derivative

For distributions P and Q of a continuous random variable, The Kullback-Leibler (KL) divergence is given as

$$D_{KL}(P||Q) = \int_{-\infty}^{\infty} \mathbf{p}(\alpha) \log\left(\frac{\mathbf{p}(\alpha)}{\mathbf{q}(\alpha)}\right) d\alpha, \qquad (7.3.18)$$

where  $\alpha$  is an intensity value and **p** and **q** denote the respective densities. The symmetric KL divergence is given as

$$D_{SKL}(P,Q) = D_{KL}(P||Q) + D_{KL}(Q||P) = \int_{-\infty}^{\infty} \mathbf{p}(\alpha) \log\left(\frac{\mathbf{p}(\alpha)}{\mathbf{q}(\alpha)}\right) + \mathbf{q}(\alpha) \log\left(\frac{\mathbf{q}(\alpha)}{\mathbf{p}(\alpha)}\right) d\alpha.$$
(7.3.19)

Our distributions *P* and *Q* are defined via our Parzen window function *q*:

$$\mathbf{p}(\alpha) = q(\alpha, \Omega) = \frac{1}{|\Omega|} \int_{\Omega} K(\alpha - I(\hat{x})) \, d\hat{x}$$
(7.3.20)

$$\mathbf{q}(\alpha) = q(\alpha, \Omega^c) = \frac{1}{|\Omega^c|} \int_{\Omega^c} K(\alpha - I(\hat{x})) \, d\hat{x}.$$
 (7.3.21)

where  $\Omega^c = D \setminus \Omega$  is the complement of  $\Omega$ , *D* the entire image domain. Knowing that the smoothed intensity level line *L*, for a given intensity  $\alpha$ , must necessarily be given as

$$L_K(\alpha) = I(I \simeq \alpha) = |\Omega| p(\alpha) + |\Omega^c| q(\alpha) = \int_{\Omega} K(\alpha - I(\hat{x})) d\hat{x} + \int_{\Omega^c} K(\alpha - I(\hat{x})) d\hat{x},$$
(7.3.22)

we can therefore express q as a function purely of  $\Omega$ :

$$\mathbf{q}(\alpha) = \frac{1}{|\Omega^c|} \int_{\Omega^c} K(\alpha - I(\hat{x})) \, d\hat{x}$$
(7.3.23)

$$= \frac{1}{|D| - |\Omega|} \int_{\mathbb{R}} (K(\alpha - I(\hat{x})))(1 - \chi(x, \Omega)) \, d\hat{x}$$
 (7.3.24)

$$=\frac{1}{|D|-|\Omega|}\int_{\mathbb{R}}K(\alpha-I(\hat{x}))\,d\hat{x}-\frac{|\Omega|}{|D|-|\Omega|}q(\alpha,\Omega) \tag{7.3.25}$$

$$=C(\alpha,\Omega) \tag{7.3.26}$$

This allows us to conveniently replace every occurrency of integrals over the complementary domain  $\Omega^c$  for this formulation of C.

Notice that the derivative of C, with respect to q is then simply

$$\frac{\partial}{\partial q}C(\alpha,\Omega) = \frac{|\Omega|}{|D| - |\Omega|}$$
(7.3.27)

89

We therefore seek to minimise

$$E_{SKL}(\Omega) = \int_{\mathbb{R}} q(\alpha, \Omega) \log\left(\frac{q(\alpha, \Omega)}{C(\alpha, \Omega)}\right) + C(\alpha, \Omega) \log\left(\frac{C(\alpha, \Omega)}{q(\alpha, \Omega)}\right) \, d\alpha.$$
(7.3.28)

Using the derivation framework developed in Section 7.3.1, to get the gradient descent direction, we need to compute the functional:

$$G(\alpha, \Omega) = \frac{1}{|\Omega|} \left( \frac{\partial}{\partial q} \varphi(q(\alpha, \Omega)) \left[ q(\alpha, \Omega) - K(\alpha - I(\hat{x})) \right] \right)$$
(7.3.29)

To simplify the calculations, I compute the derivative with respect to q first:

$$\begin{split} \frac{\partial}{\partial q}\varphi(q(\alpha,\Omega)) &= \frac{\partial}{\partial q} \left[ q(\alpha,\Omega) \log\left(\frac{q(\alpha,\Omega)}{C(\alpha,\Omega)}\right) + C(\alpha,\Omega) \log\left(\frac{C(\alpha,\Omega)}{q(\alpha,\Omega)}\right) \right] \\ &= \left(\frac{|\Omega|}{|D| - |\Omega|} + 1\right) \cdot \log\left(\frac{q(\alpha,\Omega)}{C(\alpha,\Omega)}\right) + 1 \\ &+ (C(\alpha,\Omega) - q(\alpha,\Omega)) \cdot \frac{|D| - |\Omega|}{|\Omega|} - \frac{C(\alpha,\Omega)}{q(\alpha,\Omega)} \end{split}$$
(7.3.30)

Inserting  $\frac{\partial}{\partial q}\varphi(q(\alpha,\Omega))$  into  $G(\alpha,\Omega)$  we get the descent direction

$$\psi_{SKL} = \frac{1}{|\Omega|} \left( \int_{\mathbb{R}} \left( \left( \frac{|\Omega|}{|D| - |\Omega|} + 1 \right) \cdot \log \left( \frac{q(\alpha, \Omega)}{C(\alpha, \Omega)} \right) + 1 + (C(\alpha, \Omega) - q(\alpha, \Omega)) \cdot \frac{|D| - |\Omega|}{|\Omega|} - \frac{C(\alpha, \Omega)}{q(\alpha, \Omega)} \right) + \left[ q(\alpha, \Omega) - K(\alpha - I(\hat{x})) \right] d\alpha \right) \cdot \mathbf{n},$$
(7.3.31)

which is our direction of greatest descent for the Symmetrised Kullback-Leibler term.

#### 7.3.4 Edge penalised length term

Our proposed regularisation term has the classical form of [Caselles et al., 1997] and is therefore simply given as

$$E_{reg}(\Omega) = \oint g(\nabla u(c(p))) |c_p| dp, \qquad (7.3.32)$$

which has the steepest descent-direction of

$$\psi_{reg} = -(\nabla g \cdot \mathbf{n}) \cdot \mathbf{n} - g\kappa \cdot \mathbf{n}. \tag{7.3.33}$$

(7.3.33) is then the negative gradient of  $E_{reg}$  and the classical result (with a sign difference) obtained by [Caselles et al., 1997] using a parametrised curve.

#### 7.3.5 Full descent direction

We combine the terms derived in the previous section and arrive at the direction of steepest descent for our full energy formulation:

$$\psi = \lambda_{1}\psi_{E} - \lambda_{2}\psi_{SKL} + \lambda_{3}\psi_{reg}$$

$$= \frac{\lambda_{1}}{|\Omega|} \left( \int_{\mathbb{R}} (-\log q(\alpha, \Omega) - 1) \left[ q(\alpha, \Omega) - K(\alpha - I(\hat{x})) \right] d\alpha \right) \cdot \mathbf{n}$$

$$- \frac{\lambda_{2}}{|\Omega|} \left( \int_{\mathbb{R}} \left( \left( \frac{|\Omega|}{|D| - |\Omega|} + 1 \right) \cdot \log \left( \frac{q(\alpha, \Omega)}{C(\alpha, \Omega)} \right) + 1 \right) + (C(\alpha, \Omega) - q(\alpha, \Omega)) \cdot \frac{|D| - |\Omega|}{|\Omega|} - \frac{C(\alpha, \Omega)}{q(\alpha, \Omega)} \right) [q(\alpha, \Omega) - K(\alpha - I(\hat{x}))] d\alpha$$

$$\cdot \mathbf{n} - \lambda_{3}((\nabla g \cdot \mathbf{n}) \cdot \mathbf{n} - g\kappa \cdot \mathbf{n})$$

$$(7.3.35)$$

# 7.4 Individual experiments

Synthetic data was generated by first initialising a background intensity and foreground intensity squares. Darker and lighter edges, with varying thickness, were then added to a subset of the squares to capture all the different observed edge types.

The experiments that have been done so far consist of individual tests of the data term and the length regularisation term, but no combined tests. In order to test the choice of g function and argue for the results, we used the opensource implementation from github.com/pmneila/morphsnakes - the Morphological Geodesic Active Contours method of [Marquez-Neila et al., 2014] - and altered their choice of g function manually. The three different g functions are

$$g_{\text{mag}} = \frac{1}{1 + \beta |\nabla_{\sigma_0} u|} \tag{7.4.1}$$

$$g_{\text{mag-unsharp}} = \frac{1}{1 + \beta |\nabla_{\sigma_0} (2u - k_{(\sigma_2)} * u)|}$$
(7.4.2)

$$g = \frac{1}{1 + \beta |\nabla_{\sigma_0} w_{(\sigma_1)}(2u - k_{(\sigma_2)} * u)|}.$$
 (7.4.3)

The results of testing the three different choices of g function can be seen in Figures 7.2 and 7.3. We notice that the classical gradient magnitude edge detector (7.4.1) is not able to satisfactorily pickup the true edge, when the dark edge is slightly thicker. Figure 7.2b shows that while (7.4.2) satisfactorily deal with exterior edges, it does not allow segmentation of the overlapping squares and is inherently more prone to noise because of the unsharp masking.

Figure 7.3 satisfactorily segments the squares using the proposed g function (7.4.3).

For the preliminary tests of the data terms, we used an opensource implementation from github.com/kevin-keraudren/.../chanvese3d.py. The Chan-Vese data term was replaced with our own, but we kept their adopted narrow band based level-set approach, which was originally formulated in [Chopp, 1991, Chopp, 1993] and then analysed in detail and optimised in [Adalsteinsson and Sethian, 1995]. The individual result of our data fidelity experiment can be seen in Figure 7.4. We notice that the data term is able to pickup the darker regions of the cameraman's jacket satisfactorily, without deteriorating into only black, thanks to the symmetrised Kulback Leibler term.

# 7.5 Conclusion

In this chapter, we have proposed a mode that is able to satisfactorily deal with a specific complex type of edge, by utilising unsharp masking and median filtering, before taking the gradient magnitude. This weight function matches the general intensity profile of an observed edge, but is also able to accomodate weaker edges in the image, while being robust to noise. The data fidelity term utilises the information theoretic terms entropy and a symmetrised version of Kulback Leibler divergence. Individual preliminary tests have been run on synthetic and real data for both the data and regularisation term. The two terms need to be combined in a single model and run on experimental data, to draw any definitive conclusions about the useability, but preliminary results are very promising.

#### Acknowledgements

The authors thank Henning Osholm Sørensen of the NanoGeoScience group at the University of Copenhagen, for access to the data and insight in the nature of the targeted artefact.



**Figure 7.2:** Comparison of different choices of the *g* weight function. The top figure corresponds to the best results for the classical gradient magnitude *g*, the bottom one the gradient magnitude of the unsharp-masked image.

(b)

# 7. Segmentation by Information Theory and Edge Penalty



**Figure 7.3:** The proposed *g* function model.



**Figure 7.4:** Individual test of the proposed entropy and symmetrised Kulback-Leibler data terms, with classic length penalisation.
## 8 Discussion

In the following, a number of key topics have been selected to elaborate and improve on, they range from model choices and details to meaningful extensions.

- 1. Empty scans for model calibration: Datasets from a variety of sources have been obtained and processed. These sources however each introduce different noise levels and types, introduce various artefacts, and suffer from defects and other anomalies. This complicates any automatic selection of parameters, as each dataset is not only structually different, but also has a large variance of noise and artefacts. One approach to remedy or improve on manual parameter selection methods, other than analysing and estimating the noise of each dataset, is to use the same real phantom objects as an initial scan for each acquisition facility. This object should be known beforehand, so a groundtruth segmentation would be available, and noise and artefact analysis would consequently be straightforward. Knowing the exact type and extent of the errors in the signal can greatly help calibrating or even selecting subsequnt algorithms, from reconstruction to segmentation.
- 2. Quadratic regularisation vs TV: Chapter 2 showed that our proposed method with a Dirichlet energy based regularisation outperformed the Total Variation variant. To be able to conduct the experiments we build complicated images from the ground truth, as a necessary verification test. One could argue that the synthetically generated dataset in particular favoured the quadratic method over the TV based. Additionally a nonexchaustive grid search was used for parameter selection, but this was also shared for all tested methods. Furthermore, comparisons and similarity measures were computed based on thresholded label fields, not the probabilistic labels, so overly smoothed edges from the quadratic regulariser were cut off.
- 3. **Posterior vs thresholded labels:** A number of different segmentation methods are used in the P<sup>3</sup>-project, whereas only our proposed methods

rely on probabilistic labels. Having binary labels introduce a very heavy bias on the corresponding meshing procedure and subsequent fluid flow simulations. Oversegmentations of the material phase often occurs at edges, in particular narrow edges, where partial volume effects are most prone. A single voxel misclassification can therefore result in cutting off an entire pore network, which can have detrimental effects on the fluid flow analyses. Additionally, if one decidedes to threshold the labels at 50%, one makes a choice that might not be straightforwardly justified. Additionally that choice might have been made at a premature stage, considering so many modules in the project still remain. Conversely, outputting the posterior label field allows the meshing software and group to make the decision, based on their expertice. Smaller thresholds are more well-founded if it means including a pore-subnetwork to the main network, as sedimentary rock pore networks are assumed to have few 'lakes', due to their origin. Finally, the meshing software can more lenient select a preferred resolution to accomodate the probabilistic labels, while we decided to constrain ourselves to voxel scale.

- 4. Generalised methods: In Chapters 2 and 3 two methods for segmenting bias field rich datasets were presented. One noteworthy feature is that the TV variant inherently generalises [Chambolle et al., 2012], thus providing a richer representation of their method, but with added free parameters. For the method proposed in Chapter 3, setting the kernel extent to the entire image and the weight function to zero, we arrive exactly at the method proposed by [Chambolle et al., 2012]. This generalisation enables our proposed methods to at least address the same problems as their method and more.
- 5. Number of classes: The number of classes is a fixed, manually selected, parameter for all methods. The choice is solely based on visual inspection of rather complex 3D tomographic volumes, which is tedious and prone to errors. An alternative, rather successful, heuristic method used was choosing a definite higher *k* than needed and combining classes in the end. Having an automatised selection method is however more beneficial and more work should be put into incorporating classical *k*-selection methods into our proposed models. Alternatively, if only the background is of interest, one could use methods inspired by Chapter 7.
- 6. Fuzzy segmentation initialisation: The proposed methods of Chapters 2 and 3 use either a classical Otsu's thresholding method or a dual filtering approach for initialisation. This is however prone to introduce bias in the methods and could mean slower convergence or eronous results. Alternatively one could implement a fuzzy Otsu's thresholding method, that produces posterior label fields, instead of binary. Intensity values close to the thresholds from Otsu's method would then be

weighted by some smoothly decaying function, to indicate that the probability of a given intensity belonging to either class is approaching 0.5. Having such a probabilistic initialisation seem more appropriate and correct, not only for our methods, but for thresholding in general.

- 7. Shape and type of local mean kernels: In Chapters 2 and 3 the local mean kernel shapes are implemented as squared entities instead of being circular / ball shaped. A radial basis function based kernel should be used, but was avoided due to fast implementations of squared local mean kernels. Coincidentally, this might explain the slight dice-score difference between using Gaussian and uniform kernels. Although, it is not a given that Gaussian kernels should be preferred, as values further away from the voxel of interest should not necessarily contribute less. This is especially true when working with smaller kernels, where capturing a significant amount of image information is essential for the local mean to be meaningful.
- 8. Utilising rat injection needle data: The proposed rodent brain segmentation methods presented in Chapters 4 to 6 all utilise a classical bias field correction method as preprocessing, in an attempt to clean the data. Through presenting and discussing the method at the ISMRM conference in Paris 2018, new information was obtained on a visible needle in the scans. Precise knowledge of the needle and its content could provide valuable information on the nature of the bias field and its strength. Knowing this information, a more specialised bias field correction method could be developed.
- 9. Edge penalising prior Chapter 6 presented a segmentation method especially fitted for object segmentation, by use of shape prior invariant to similarity transformations. A further improvement would be to include a classical edge length penalising term to further regularise the label field. This would complement the shape prior by constraining the vast majority of the label field to the brain region and make sure that strong edges are kept as part of the result. Generally, for brains, the only sharp edges are found at the transition between brain and the background. This would hopefully further increase the robustness of the method and is straightforward to implement.
- 10. Extending shape priors to multiphase: The shape prior method presented in Chapter 6 has been tested for two classes; background and combined brain matter / CSF. CSF refers to the cerebrospinal fluid of the brain which was segmented in Chapter 5, but not a focus in Chapter 6. While this works satisfactorily, an additional step would be required to segment the CSF from the brain matter. A straightforward idea would be to extend the framework to two (or more) classes of shape priors, this

seems easy to achieve from our current formulation. However satisfactory results are not a guarantee, as individual CSF network locations in rat brains seem to vary more than the overall surface of a brain and CSF network shape variance are not necessarily as well-behaved. Additionally, removing internal CSF regions from the full brain training shapes might jeopardise the eigenvector directions, making them less robust. Another possibility is to run a specific CSF segmentation after brain mass extraction. Experiments need to be conducted to clarify this concern.

- 11. Initialisation of the shape prior method: The method based on shape priors presented in Chapter 6 currently uses the first training shape as initialisation, but eroded to stay inside the target brain, while keeping the general shape of a brain. While the strength and iterations of the erosion procedure can be controlled, the method still relies on the initialisation label field staying inside the brain we are trying to segment. While this is not a problem for our target application, it is not robust to general object extraction. One way to completely automatise this initialisation procedure is to incorporate the idea of [Wang et al., 2015]. They proposed to compute salient points in both the training data and the real data and then perform feature matching. The matched points would in turn create a shape decision kernel function, from which the most likely shape region can be determined. Using this concept, we can simply align a training shape to this shape region for an automatic and robust initialisation.
- 12. **Speed considerations:** All presented methods have been implemented purely as proof of concepts. This means that although the model description is inherently optimised to be efficient, the implementation is not. Processing up to 4 GB for a single dataset where thousands are available, and hundreds of iterations are usually required, means that running times are essential, even when large computational clusters are available. Assuming that RAM is readibly available an obvious optimisation target is the individual class updates, which for all presented methods is done sequentially in each iteration. A rudimentary parallisation implementation of this step would mean a significant factor k speedup, where k is the number of classes.

Naturally the choice of programming language has a big impact as well, where currently all methods have been implemented using a combination of Python scientific computing libraries and a few C++ backend procedures. One could consider porting the framework to newer languages that focus on automatic hardware acceleration of computations, such as Bohrium [Kristensen et al., 2013, Kristensen et al., 2014]. Another option is Futhark [Henriksen and Oancea, 2014], which is a high-performance functional language that utilises data-parallel array programming on the GPU, that could be considered for the most computation intense updates.

In regards to particularly the shape prior method, presented in detail in Chapter 6, each iteration needs to align all training shapes to the current label field estimate, which is the heaviest computation. Naturally, this could be done by naive parallelisation of each training shape. Another consideration is to compute a single mean training shape, of all aligned training shapes. While this is not entirely correct or equal to our current model, it could serve as an approximation that would greatly reduce computation time as well as space requirements and is worth investigating.

## 9 Conclusion

This thesis presented six novel variational methods for segmentation of tomographic data. The datasets cover X-ray micro- and nano-CT acquired porous geological samples, as well as MRI scans of rodent crania.

In Chapter 2 a method to deal with intensity inhomogeneity and noise in complex weakly structured images was presented. It focused on accounting for local and global bias fields without explicitedly modelling them, but instead inherently deal with them. The classical global mean based data fidelity term was constrained to local regions, governed by a kernel choice and its extent, thus making the mean spatially dependent. Bias field inherently violate the usual assuptions that classes are piecewise constant, so by controlling the extend of the kernel, the method is able to accomodate various variation strength of the intensity inhomogeneity. Two methods with different regularisers were proposed and tested experimentally on synthetic and real data; one regulariser based on the Dirichlet energy, the other a tight Total Variation. Both methods were derived from the Mumford-Shah functionali, uses a fixed number of classes and Hidden Markov Measure Field Models, modelling the region labels, to account for partial volume effects. Both of the methods are solved by alternatingly minimising the label fields and the local class averages. Through reporting both segmentation quality and geophysical parameters the Dirichlet energy was found to have the edge on the synthetic data.

Chapter 3 presented an extension to the work in Chapter 2, by addressing the observed sharp edges in geological samples. The foundation of the model handling bias fields is shared with that of Chapter 2 and binary labels are similarly modelled as posterior probabilities as a relaxation and to account for partial volume effects. The extended method presented a weighting term to the two regularisers. One was a Total Variation weighted semi-norm akin to that of Geodesic Active Contours, but with a richer and different representation. The second being a weighted Dirichlet regularisation energy, i.e. weighted squared gradient. Since the weight terms are based on first order information of the constant image, not the label fields, the optimisation problem is only affected by an image dependent constant, compared to the non-weighted formulation. Similar experiments on synthetic and real data were performed and compared with those from Chapter 2, which improved the previous best results, except for pore network connectivity measurements.

The framework presented in Chapter 4 aimed to extract cerebrospinal fluid (CSF) and brain matter from MRI scans of rat crania. The main segmentation method adopted in this framework was the Dirichlet regularised energy presented in Chapter 2. The extractions were achieved through a process of four major steps, namely 1) bias field correction, 2) 2-class segmentation, 3) Geodesic Active Contours (GAC) brain extraction, and 4) 3-class segmentation. Additionally, some postprocessing was applied to filter out exterior CSF as well as very small CSF occurances. For bias field correction, a classical correction method was used to recover more of the underlying data. The 2-class segmentation was used both as target image for the GAC method as well as initialisation of the brain and csf region of the extracted brain. The GAC method was initialised at the center of the brain and would expand until stopping at the sharp edges achieved by the 2-class segmentation. Finally the 3-class segmentation would focus on segmenting the masked brain of the original bias field corrected image, using the 2-class masked segmentation as initialisation. This framework was very successful in segmenting and extracting the CSF network as well as brain matter of highly bias field affected rat crania and will be of great help in the medical community.

Chapter 5 outlined an extension to the framework presented in Chapter 4 by introducing a shape prior term to guide the brain extraction process. The data term for this new brain extraction term was the classical Chan-Vese one, while the shape prior term replaced the length penalisation. The new term is invariant to translation, global scaling, and rotations by alignment, making the method fully automatic and greatly limits the number of free parameters, while still improving robustness. Experiments were conducted by training satisfiably segmented rat brains from Chapter 4 and testing on ones that were not possible to segment with the old framework. Additionally an experiment was performed on a mouse brain, while training on rat priors with great success, especially compared to third party software.

Chapter 6 provided a more detailed and cleaner formulation of the shape prior extension of Chapter 5. The model utilised a narrow-band like version of the data term presented in Chapter 2 and works directly on the bias field corrected data. While experiments were conducted on just two classes, the method already incompasses multiple classes in its current formulation and the implementation straightforwardly extends to multiple classes as well. Multiple classes simply requires the corresponding shape prior training sets, by splitting the segmented brains from Chapter 4 into their brain matter and CSF components. In order to initialise the model, a training shape subjected to mathematical erosion was used to give a good initial rotational orientation and shape variance, while staying within the brain of interest. The chapter demonstrated experimentally the invariance to translation, scaling, and rotation, in addition to being able to segment datasets previously impossible by the framework proposed in Chapter 4.

In Chapter 7 an ongoing work for segmentation of tomographic geological material was presented. The model encompassed information theoretic data terms by use of Shannon entropy and symmetrised Kullback-Leibler energies using shape gradients. To accomodate a special edge artefact, a novel weigh function was proposed that mimics the intensity profile of such edges. It is designed using an unsharp mask to enhance edges, while a median filter limits the impact of enhanced noice. The gradient magnitude of the resulting image is then used as a weight to guide the length penalisation. The length term was compared with classical weight functions and showed to outperform them on representative synthetic data. Individual experiments were conducted, by adding the proposed terms to open source implementations of classical methods, to assess the usefulness of each of the terms and preliminary results were very promising.

This thesis has demonstrated that we are able to model and account for various artefacts related to tomographc reconstructions of chalk rocks and rat crania and satisfactorily segment relevant regions of interest. The presented methods have rich representations that lends themselves well to various types of datasets and problems. Through numerous types of synthetic and real data experiments, this thesis has validated and established the usefulness and versatility of the presented models.

## Bibliography

- [Achanta et al., 2012] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Süsstrunk, S., et al. (2012). Slic superpixels compared to state-of-the-art superpixel methods. *IEEE transactions on pattern analysis and machine intelli*gence, 34(11):2274–2282.
- [Adalsteinsson and Sethian, 1995] Adalsteinsson, D. and Sethian, J. A. (1995). A fast level set method for propagating interfaces. *Journal of computational physics*, 118(2):269–277.
- [Álvarez et al., 2010] Álvarez, L., Baumela, L., Henríquez, P., and Márquez-Neila, P. (2010). Morphological snakes. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2197–2202. IEEE.
- [Aubert et al., 2003] Aubert, G., Barlaud, M., Faugeras, O., and Jehan-Besson, S. (2003). Image segmentation using active contours: Calculus of variations or shape gradients? *SIAM Journal on Applied Mathematics*, 63(6):2128–2154.
- [Barnes et al., 2014] Barnes, R., Lehman, C., and Mulla, D. (2014). Priorityflood: An optimal depression-filling and watershed-labeling algorithm for digital elevation models. *Computers & Geosciences*, 62:117–127.
- [Bertsekas, 2011] Bertsekas, D. (2011). Incremental proximal methods for large scale convex optimization. *Mathematical Programming, Ser. B.*, 129:163– 195.
- [Beucher and Lantuéj, 1979] Beucher, S. and Lantuéj, C. (1979). Use of watersheds in contour detection. In *Proceedings of the International Workshop on Image Processing*. CCETT.
- [Boykov et al., 2001] Boykov, Y., Veksler, O., and Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Transactions on pattern analysis and machine intelligence*, 23(11):1222–1239.
- [Brox and Cremers, 2009] Brox, T. and Cremers, D. (2009). On Local Regions Models and a Statistical Interpretation of the Piecewise Smooth Mumford-Shah Functional. *Int'l J. Computer Vision*, 84:184–193.

- [Buzug, 2008] Buzug, T. N. (2008). Computed Tomography: from Photon Statistics to Modern Cone Beam CT. Springer.
- [Caselles et al., 1997] Caselles, V., Kimmel, R., and Sapiro, G. (1997). Geodesic active contours. *International journal of computer vision*, 22(1):61–79.
- [Chambolle et al., 2012] Chambolle, A., Cremers, D., and Pock, T. (2012). A convex approach to minimal partitions. *SIAM Journal on Imaging Sciences*, 5(4):1113–1158.
- [Chambolle and Pock, 2011] Chambolle, A. and Pock, T. (2011). A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of mathematical imaging and vision*, 40(1):120–145.
- [Chan and Zhu, 2005] Chan, T. and Zhu, W. (2005). Level set based shape prior segmentation. In *Computer Vision and Pattern Recognition*, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 2, pages 1164–1170. IEEE.
- [Chan and Vese, 2001] Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *Trans. Img. Proc.*, 10(2):266–277.
- [Chao, 1984] Chao, A. (1984). Nonparametric estimation of the number of classes in a population. *Scandinavian Journal of statistics*, pages 265–270.
- [Charpiat et al., 2007] Charpiat, G., Faugeras, O., and Keriven, R. (2007). Shape statistics for image segmentation with prior. In *Computer Vision and Pattern Recognition*, 2007. *CVPR'07. IEEE Conference on*, pages 1–6. IEEE.
- [Cheng, 1995] Cheng, Y. (1995). Mean shift, mode seeking, and clustering. *IEEE transactions on pattern analysis and machine intelligence*, 17(8):790–799.
- [Chopp, 1991] Chopp, D. L. (1991). *Computing minimal surfaces via level set curvature flow*. PhD thesis, Mathematics Department, Physics Division, Lawrence Berkely Laboratory, University of California.
- [Chopp, 1993] Chopp, D. L. (1993). Computing minimal surfaces via level set curvature flow. *Journal of Computational Physics*, 106(1):77 91.
- [Çiçek et al., 2016] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016). 3d u-net: learning dense volumetric segmentation from sparse annotation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 424–432. Springer.
- [Cireşan et al., 2012] Cireşan, D., Meier, U., and Schmidhuber, J. (2012). Multicolumn deep neural networks for image classification. *arXiv preprint arXiv:*1202.2745.

- [Clancy et al., 2007] Clancy, B., Finlay, B. L., Darlington, R. B., and Anand, K. (2007). Extrapolating brain development from experimental species to humans. *Neurotoxicology*, 28(5):931–937.
- [Cohen, 1991] Cohen, L. D. (1991). On active contour models and balloons. *CVGIP: Image understanding*, 53(2):211–218.
- [Collobert and Weston, 2008] Collobert, R. and Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. In *Proceedings of the 25th international conference on Machine learning*, pages 160–167. ACM.
- [Collobert et al., 2011] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., and Kuksa, P. (2011). Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, 12(Aug):2493–2537.
- [Cootes et al., 1995] Cootes, T., Taylor, C., Cooper, D., and Graham, J. (1995). Active shape models-their training and application. *CVIU*, 61(1):38–59.
- [Cremers et al., 2003] Cremers, D., Kohlberger, T., and Schnörr, C. (2003). Shape statistics in kernel space for variational image segmentation. *Pattern Recognition*, 36(9):1929–1943.
- [Cremers et al., 2002] Cremers, D., Tischhäuser, F., Weickert, J., and Schnörr, C. (2002). Diffusion snakes: Introducing statistical shape knowledge into the mumford-shah functional. *International journal of computer vision*, 50(3):295– 313.
- [Dahl et al., 2012] Dahl, G. E., Yu, D., Deng, L., and Acero, A. (2012). Contextdependent pre-trained deep neural networks for large-vocabulary speech recognition. *IEEE Transactions on audio, speech, and language processing*, 20(1):30–42.
- [Davis and Elliott, 2006] Davis, G. and Elliott, J. (2006). Artefacts in x-ray microtomography of materials. *Materials science and technology*, 22(9):1011–1018.
- [Delfour and Zolésio, 2001] Delfour, M. and Zolésio, J.-P. (2001). *Shapes and Geometries*. Advances in Design and Control. Siam.
- [DiCarlo et al., 2012] DiCarlo, J. J., Zoccolan, D., and Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron*, 73(3):415–434.
- [Dong et al., 2017] Dong, H., Yang, G., Liu, F., Mo, Y., and Guo, Y. (2017). Automatic brain tumor detection and segmentation using u-net based fully convolutional networks. In *Annual Conference on Medical Image Understanding and Analysis*, pages 506–517. Springer.

- [Ekeland and Téman, 1999] Ekeland, I. and Téman, R. (1999). *Convex Analysis and Variational Problems*. SIAM, Philadelphia, PA, USA.
- [Erhan et al., 2014] Erhan, D., Szegedy, C., Toshev, A., and Anguelov, D. (2014). Scalable object detection using deep neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2147–2154.
- [Foulonneau et al., 2006] Foulonneau, A., Charbonnier, P., and Heitz, F. (2006). Affine-invariant geometric shape priors for region-based active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(8):1352–1357.
- [Fukunaga and Hostetler, 1975] Fukunaga, K. and Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on information theory*, 21(1):32–40.
- [Fussenegger et al., 2006] Fussenegger, M., Deriche, R., and Pinz, A. (2006). A multiphase level set based segmentation framework with pose invariant shape priors. In *Asian Conference on Computer Vision*, pages 395–404. Springer.
- [Geman and Geman, 1984] Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, 6(6):721–741.
- [Grady, 2006] Grady, L. (2006). Random walks for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 28(11):1768–1783.
- [Graves et al., 2013] Graves, A., Mohamed, A.-r., and Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In *Acoustics, speech and signal processing (icassp), 2013 ieee international conference on,* pages 6645–6649. IEEE.
- [Hansen, 2015] Hansen, J. D. K. (2015). Volume segmentation of synthetically generated data of porous chalk rocks. Master's thesis, Department of Computer Science, University of Copenhagen, Denmark.
- [Hansen et al., 2018] Hansen, J. D. K. et al. (2018). Brain extraction and segmentation framework for bias field rich cranial mri scans of rats. In *Book of Abstracts*, Annual meeting ISMRM, page 3249, Paris, France.
- [Hansen and Lauze, 2017] Hansen, J. D. K. and Lauze, F. (2017). Local mean multiphase segmentation with hmmf models. In *International Conference* on Scale Space and Variational Methods in Computer Vision, pages 396–407. Springer.

- [Hansen and Lauze, 2018] Hansen, J. D. K. and Lauze, F. (2018). Multiphase local mean geodesic active regions. In 24th International Conference on Pattern Recognition (ICPR), pages 3031–3036. IEEE.
- [Hansen and O'Leary, 1993] Hansen, P. C. and O'Leary, D. P. (1993). The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems. *SIAM Journal on Scientific Computing*, 14(6):1487–1503.
- [Heisenberg, 1928] Heisenberg, W. (1928). Zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 49(9):619–636.
- [Held and Crowder, 1974] Held, M. and Crowder, P. W. W. (1974). Validation of Subgradient Optimization. *Math. Program.*, 6:62–88.
- [Henriksen and Oancea, 2014] Henriksen, T. and Oancea, C. E. (2014). Bounds checking: An instance of hybrid analysis. In *Proceedings of ACM SIGPLAN International Workshop on Libraries, Languages, and Compilers for Array Programming*, page 88. ACM.
- [Herbulot et al., 2004] Herbulot, A., Jehan-Besson, S., Barlaud, M., and Aubert, G. (2004). Shape gradient for multimodal image segmentation using mutual information. In *Image Processing*, 2004. ICIP'04. 2004 International Conference on, volume 4, pages 2729–2732. IEEE.
- [Hinton et al., 2012] Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., et al. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal processing magazine*, 29(6):82–97.
- [Hinton, 2007] Hinton, G. E. (2007). Learning multiple layers of representation. *Trends in cognitive sciences*, 11(10):428–434.
- [Hinton and Salakhutdinov, 2006] Hinton, G. E. and Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786):504–507.
- [Hsieh, 2009] Hsieh, J. (2009). *Computed tomography: principles, design, artifacts, and recent advances.* SPIE Bellingham, WA.
- [Huang and Zeng, 2015] Huang, C. and Zeng, L. (2015). An active contour model for the segmentation of images with intensity inhomogeneities and bias field estimation. *PloS one*, 10(4):e0120399.
- [Iassonov et al., 2009] Iassonov, P., Gebrenegus, T., and Tuller, M. (2009). Segmentation of x-ray computed tomography images of porous materials: A crucial step for characterization and quantitative analysis of pore structures. *Water Resources Research*, 45(9).

- [Ising, 1925] Ising, E. (1925). Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258.
- [Kass et al., 1988] Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331.
- [Kendall, 1984] Kendall, D. (1984). Shape manifolds, procrustean metrics and complex projective spaces. *Bulletin of London Mathematical Society*, 16:81–121.
- [Kolmogorov and Zabih, 2004] Kolmogorov, V. and Zabih, R. (2004). What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 26(2):147–159.
- [Kristensen et al., 2013] Kristensen, M. R., Lund, S. A., Blum, T., Skovhede, K., and Vinter, B. (2013). Bohrium: unmodified numpy code on cpu, gpu, and cluster. *Python for High Performance and Scientific Computing (PyHPC 2013)*.
- [Kristensen et al., 2014] Kristensen, M. R., Lund, S. A., Blum, T., Skovhede, K., and Vinter, B. (2014). Bohrium: a virtual machine approach to portable parallelism. In *Parallel & Distributed Processing Symposium Workshops (IPDPSW)*, 2014 IEEE International, pages 312–321. IEEE.
- [Krizhevsky et al., 2012] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105.
- [Kullback and Leibler, 1951] Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86.
- [Lauze and Nielsen, 2008] Lauze, F. and Nielsen, M. (2008). From Inpainting to Active Contours. *Int'l Journal of Computer Vision*, 79(1):31–43.
- [Lenz, 1920] Lenz, W. (1920). Beitrage zum verstandnis der magnetischen eigenschaften in festen körpern. *Physikalische Zeitschrift*, 21:2.
- [Leventon et al., 2002] Leventon, M. E., Grimson, W. E. L., and Faugeras, O. (2002). Statistical shape influence in geodesic active contours. In *Biomedical Imaging*, 2002. 5th IEEE EMBS International Summer School on, pages 8–pp. IEEE.
- [Liu and Yu, 2009] Liu, D. and Yu, J. (2009). Otsu method and k-means. In *Hybrid Intelligent Systems*, 2009. *HIS'09*. *Ninth International Conference on*, volume 1, pages 344–349. IEEE.
- [Liu et al., 2011] Liu, M.-Y., Tuzel, O., Ramalingam, S., and Chellappa, R. (2011). Entropy rate superpixel segmentation. In *Computer Vision and Pattern Recognition (CVPR)*, 2011 IEEE Conference on, pages 2097–2104. IEEE.

- [Lloyd, 1982] Lloyd, S. (1982). Least squares quantization in pcm. *IEEE transactions on information theory*, 28(2):129–137.
- [Long et al., 2015] Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440.
- [Magnus, 1985] Magnus, J. (1985). On Differentiating Eigenvalues and Eigenvectors. *Econometric Theory*, 1:179–191.
- [Marquez-Neila et al., 2014] Marquez-Neila, P., Baumela, L., and Alvarez, L. (2014). A morphological approach to curvature-based evolution of curves and surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(1):2–17.
- [Marroquin et al., 2003] Marroquin, J. L., Santana, E. A., and Botello, S. (2003). Hidden Markov Measure Field Models for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(11):1380–1387.
- [Masci et al., 2011] Masci, J., Meier, U., Cireşan, D., and Schmidhuber, J. (2011). Stacked convolutional auto-encoders for hierarchical feature extraction. In *International Conference on Artificial Neural Networks*, pages 52–59. Springer.
- [Meyer, 1994] Meyer, F. (1994). Topographic distance and watershed lines. *Signal processing*, 38(1):113–125.
- [Mezghich et al., 2014] Mezghich, M. A., Slim, M., and Ghorbel, F. (2014). Invariant shape prior knowledge for an edge-based active contours invariant shape prior for active contours. In 2014 International Conference on Computer Vision Theory and Applications (VISAPP), pages 454–461. IEEE.
- [Mnih et al., 2013] Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., and Riedmiller, M. (2013). Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602*.
- [Mumford and Shah, 1989] Mumford, D. and Shah, J. (1989). Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on pure and applied mathematics*, 42(5):577–685.
- [Müter et al., 2012] Müter, D., Pedersen, S., Sørensen, H. O., Feidenhans'L, R., and Stipp, S. L. S. (2012). Improved segmentation of x-ray tomography data from porous rocks using a dual filtering approach. *Comput. Geosci.*, 49:131–139.
- [Müter et al., 2014] Müter, D., Sørensen, H. O., Jha, D., Harti, R., Dalby, K. N., Suhonen, H., Feidenhans'l, R., Engstrøm, F., and Stipp, S. L. S. (2014). Resolution dependence of petrophysical parameters derived from x-ray tomography of chalk. *Applied Physics Letters*, 105(4).

- [Nair and Hinton, 2010] Nair, V. and Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814.
- [Ng et al., 2002] Ng, A. Y., Jordan, M. I., and Weiss, Y. (2002). On spectral clustering: Analysis and an algorithm. In *Advances in neural information processing systems*, pages 849–856.
- [Nielsen et al., 1997] Nielsen, M., Florack, L., and Deriche., R. (1997). Regularization, scale-space, and edge detection filters. *Journal of Mathematical Imaging and Vision*, 7:291–307.
- [Osher and Sethian, 1988] Osher, S. and Sethian, J. A. (1988). Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. *Journal of computational physics*, 79(1):12–49.
- [Otsu, 1979] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66.
- [Paragios and Deriche, 2002a] Paragios, N. and Deriche, R. (2002a). Geodesic active regions: a new paradigm to deal with frame partition problems in computer vision. Journal of Visual Communication and Image Representation, Special Issue on Partial Differential Equations in Image Processing, Computer Vision and Computer Graphics, 13(1/2):249–268.
- [Paragios and Deriche, 2002b] Paragios, N. and Deriche, R. (2002b). Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision*, 46(3):223–247.
- [Parzen, 1962] Parzen, E. (1962). On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076.
- [Perona and Malik, 1990] Perona, P. and Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on pattern analysis and machine intelligence*, 12(7):629–639.
- [Pham and Prince, 1999] Pham, D. L. and Prince, J. L. (1999). Adaptive fuzzy segmentation of magnetic resonance images. *IEEE transactions on medical imaging*, 18(9):737–752.
- [Pock et al., 2009] Pock, T., Cremers, D., Bischof, H., and Chambolle, A. (2009). An algorithm for minimizing the mumford-shah functional. In *Computer Vision*, 2009 IEEE 12th International Conference on, pages 1133–1140. IEEE.
- [Potts, 1952] Potts, R. B. (1952). Some generalized order-disorder transformations. *Mathematical Proceedings of the Cambridge Philosophical Society*, 48(1):106–109.

- [Poultney et al., 2007] Poultney, C., Chopra, S., Cun, Y. L., et al. (2007). Efficient learning of sparse representations with an energy-based model. In *Advances in neural information processing systems*, pages 1137–1144.
- [Ren and Malik, 2003] Ren, X. and Malik, J. (2003). Learning a classification model for segmentation. In *null*, page 10. IEEE.
- [Rokach and Maimon, 2005] Rokach, L. and Maimon, O. (2005). Clustering methods. In *Data mining and knowledge discovery handbook*, pages 321–352. Springer.
- [Ronneberger et al., 2015] Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer.
- [Rosenblatt, 1956] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, pages 832–837.
- [Rother et al., 2004] Rother, C., Kolmogorov, V., and Blake, A. (2004). Grabcut: Interactive foreground extraction using iterated graph cuts. In *ACM transactions on graphics (TOG)*, volume 23, pages 309–314. ACM.
- [Rudin et al., 1992] Rudin, L. I., Osher, S., and Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268.
- [Sezgin and Sankur, 2004] Sezgin, M. and Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic imaging*, 13(1):146–166.
- [Shannon, 1948] Shannon, C. E. (1948). A mathematical theory of communication. Bell system technical journal, 27(3):379–423.
- [Shen et al., 2014] Shen, J., Du, Y., Wang, W., and Li, X. (2014). Lazy random walks for superpixel segmentation. *IEEE Transactions on Image Processing*, 23(4):1451–1462.
- [Shi and Malik, 2000] Shi, J. and Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on pattern analysis and machine intelligence*, 22(8):888–905.
- [Silver et al., 2016] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., et al. (2016). Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484.

- [Silverman, 1981] Silverman, B. W. (1981). Using kernel density estimates to investigate multimodality. *Journal of the Royal Statistical Society. Series B* (*Methodological*), pages 97–99.
- [Strahler, 1957] Strahler, A. N. (1957). Quantitative analysis of watershed geomorphology. *Eos, Transactions American Geophysical Union*, 38(6):913– 920.
- [Szegedy et al., 2013] Szegedy, C., Toshev, A., and Erhan, D. (2013). Deep neural networks for object detection. In Advances in neural information processing systems, pages 2553–2561.
- [Thorpe et al., 1996] Thorpe, S., Fize, D., and Marlot, C. (1996). Speed of processing in the human visual system. *nature*, 381(6582):520.
- [Tustison et al., 2010] Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., and Gee, J. C. (2010). N4itk: improved n3 bias correction. *IEEE transactions on medical imaging*, 29(6):1310–1320.
- [Vandamme, 2014] Vandamme, T. F. (2014). Use of rodents as models of human diseases. *Journal of pharmacy & bioallied sciences*, 6(1):2.
- [Veksler et al., 2010] Veksler, O., Boykov, Y., and Mehrani, P. (2010). Superpixels and supervoxels in an energy optimization framework. In *European conference on Computer vision*, pages 211–224. Springer.
- [Vese and Chan, 2002] Vese, L. A. and Chan, T. F. (2002). A multiphase level set framework for image segmentation using the mumford and shah model. *International journal of computer vision*, 50(3):271–293.
- [Vincent et al., 2008] Vincent, P., Larochelle, H., Bengio, Y., and Manzagol, P.-A. (2008). Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning*, pages 1096–1103. ACM.
- [Von Luxburg, 2007] Von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and computing*, 17(4):395–416.
- [Vovk et al., 2007] Vovk, U., Pernus, F., and Likar, B. (2007). A review of methods for correction of intensity inhomogeneity in mri. *IEEE transactions* on medical imaging, 26(3):405–421.
- [Wang et al., 2015] Wang, J., Yeung, S.-K., and Chan, K. L. (2015). Matchingconstrained active contours with affine-invariant shape prior. *Computer Vision and Image Understanding*, 132:39–55.
- [Wang et al., 2010] Wang, X.-F., Huang, D.-S., and Xu, H. (2010). An efficient local chan–vese model for image segmentation. *Pattern Recognition*, 43(3):603–618.

- [Watkins and Dayan, 1992] Watkins, C. J. and Dayan, P. (1992). Q-learning. *Machine learning*, 8(3-4):279–292.
- [Watkins, 1989] Watkins, C. J. C. H. (1989). *Learning from delayed rewards*. PhD thesis, King's College, Cambridge.
- [Weickert, 1998] Weickert, J. (1998). *Anisotropic diffusion in image processing*, volume 1. Teubner Stuttgart.
- [Yi and Moon, 2012] Yi, F. and Moon, I. (2012). Image segmentation: A survey of graph-cut methods. In *Systems and Informatics (ICSAI), 2012 International Conference on*, pages 1936–1941. IEEE.
- [Zahn, 1970] Zahn, C. T. (1970). Graph theoretical methods for detecting and describing gestalt clusters. *IEEE Trans. Comput.*, 20(SLAC-PUB-0672-REV):68.