# Multimodal Image Fusion and Its Applications

by

Yu-Hui Chen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering: Systems)
in The University of Michigan
2016

Doctoral Committee:

Professor Alfred O. Hero III, Chair
Assistant Professor Laura K. Balzano
Professor Selim Esedoglu
Professor Jeffrey A. Fessler

To my parents, my sister, and Hsiu-Jung

# ACKNOWLEDGEMENTS

The completion of my dissertation symbolizes the end of my five years study abroad life. The dissertation would not have been finished without the support of many people, to whom I would like to express my sincere gratitude.

For Professor Alfred Hero, my advisor for over five years, I do not have sufficient words to convey my thanks to him. He kindly accepted me to be in his research group the first semester I came to University of Michigan even though I have nearly zero research experience. Since then, he has always been supportive to me and guided me on different research projects. I am very much impressed by his depth and breath of knowledge in different areas and his insight to see through the problems. When I felt lost in research in my first few years, he did not give up on me but patiently lead me through it. I have learned a lot about research process from him and I wish I can have the enthusiasm as much as he has in exploring new research findings. My sincere thanks also goes to my other committee members, Professor Jeffrey Fessler, Professor Laura Balzano, and Professor Selim Esedoglu, for their valuable inputs to this dissertation.

Besides, I would like to thank Dr. Dennis Wei, a former postdoctoral researcher in Hero Research Group, who helped me a lot in my research projects. We worked together in the material microscopy image fusion project. He always discussed with me about the detail of our approach and gave me substantial suggestions. He also taught me how to organize and write research papers and patiently guided me on revising the papers. His efforts contributed a great deal to my publications and helped

improve my academic writing skill. I would also like to express my gratitude to co-operators from different associations, Professor Marc De Graef from Carnegie Mellon University, Professor Charles R. Meyer and Professor Boklye Kim from Radiology Department in University of Michigan, and Jeffrey Simmons from Air Force Research Lab.

I appreciate the time working and getting along with many graduate students and postdoctoral fellows in Hero Research Group and EECS department. It was a great experience working and discussing with Dr. Gregory Newstadt and Dr. SeUn Park about the research projects. I would like to acknowledge Dr. Chih-Chun Chia, Dr. Mark Hsiao, Dr. Hamed Firouzi, Dr. Zhaoshi Meng, Dr. Kevin Xu, Dr. Hung Nien, and Dr. Tzu-Yu Liu for their encouragements when I was down and their constructive advices as being senior students. I also thank Tianpei Xie, Kevin Moon, Brandon Oselio, Kristjan Greenewald, Pin-Yu Chen, and Joel Leblanc for their valuable discussions. I am grateful to Michele Feldkamp, Becky Turanski for helping me and answering my questions. To my good friends here at the University of Michigan, thank you all for enriching my life in Ann Arbor and helping me in those difficult times.

Last but not least, this dissertation is dedicated to my family. To my parents, Mr. Ko-Jen Chen and Ms. Si-Jing Hou, for their love and support to me. Without their support and encouragement, I would not have had the chance to study abroad to pursue my Ph.D. degree and expand my horizon. I am also thankful to my younger sister Yu-Chi Chen for accompanying and taking care of my parents when I was not in Taiwan. Thank my dearest Hsiu-Jung for accompanying me in finishing up the degree and fulfilling my life with her sweetest smile.

# TABLE OF CONTENTS

# LIST OF FIGURES

**Figure**

# LIST OF TABLES

**Table**

# LIST OF APPENDICES

**Appendix**

# ABSTRACT

Multimodal Image Fusion and Its Applications

by

Yu-Hui Chen

Chair: Alfred O. Hero III

Image fusion integrates different modality images to provide comprehensive information of the image content, increasing interpretation capabilities and producing more reliable results. There are several advantages of combining multi-modal images, including improving geometric corrections, complementing data for improved classification, and enhancing features for analysis...etc.

This thesis develops the image fusion idea in the context of two domains: material microscopy and biomedical imaging. The proposed methods include image modeling, image indexing, image segmentation, and image registration. The common theme behind all proposed methods is the use of complementary information from multi-modal images to achieve better registration, feature extraction, and detection performances. Specifically, this thesis makes the following contributions:

**Material Microscopy:**

(1) In Chapter III, we propose a statistical model to describe and analyze the crystal orientations which possess certain symmetry properties associ-

ated with the materials particular polycrystalline structure. The proposed model forms the basis of the following image indexing and segmentation. The experimental results demonstrate that the proposed model and the EM algorithm is able to provide accuracy parameters estimation in the presence of symmetry invariance.

(2) In Chapter IV, we propose a novel dictionary-based indexing method for Electron Backscatter Diffraction (EBSD) data. The proposed dictionary matching approach permits segmentation, anomaly detection, and indexing to be performed in a unified manner with the additional benefit of uncertainty quantification.

(3) In Chapter V, an anomaly-driven registration algorithm based on the statistical model is developed to address the correspondence problem between different modalities. This approach allows the registration to be done at the region level to facilitate data fusion while avoiding the need for interpolation. We demonstrate that our approach has significantly better registration and segmentation accuracy than the state-of-the-art registration and segmentation methods on microscopy images.

**Biomedical Imaging:**

(1) In Chapter VI, we develop a registration-based method which utilizes the information from different modalities to estimate and compensate the head movement during functional MRI scan. The proposed method possesses both the bias reduction properties of the slice-to-volume registration approach and the variance reduction properties of the volume-to-volume registration approach. The proposed approach show significant better head motion estimation accuracy as compared to previous registration-only methods.

# CHAPTER I

# Introduction

## 1.1 Motivation

In recent years, the development of imaging sensors and the improvement of computational power allow us to use different sensing modalities and process them together, giving us the ability to study the world from the smallest microparticles to the largest universe. The main advantage of combining multiple sensor modalities is that the complementary information carried by each modality together provide a better and more complete view of the image content. The task of intelligently combining multi-modal images to enhance the view of an object with extended information is referred to as **Image Fusion** (*Smith and Heather*, 2005). Image fusion has been widely applied and proven its powerfulness in different research areas, such as, computer vision (*Han et al.*, 2013), multimedia analysis (*Atrey et al.*, 2010), biomedical research (*Sui et al.*, 2012), material sciences (*Wortmann*, 2009).

One of the most challenging and necessary tasks in image fusion is image registration. The images in different modalities are acquired through different sensors, imaging mechanisms, or optical paths, resulting in images that have different resolutions, disparities and distortions. For example, in imaging microstructures of materials, Scanning Electron (SE) imaging have spatial resolution of $0.5nm$ per pixel while Energy Dispersive X-ray Spectroscopy (EDS) image only has spatial resolution of $2\mu m$

but carries information about the chemistry of the material for each pixel (*Bertin*, 1978). For such images to be accurately fused and analyzed simultaneously, it is important that they be properly aligned with each other spatially. This thesis introduces new methods for fusion of single and multimodality images, including multi-modal image registration, image classification, image parameter estimation, and image segmentation. These are developed in the context of two domains: *material microscopy* and *biomedical imaging* as described in the following sections.

## 1.2 Anomaly-driven Image Fusion in Materials

One important objective of materials data analysis is to discover anomalies. We define anomalous events in materials as unusual material properties having exceedingly low probability relative to some nominal distribution. For instance, unusually small or large grain size/shape distributions, unusual crystal orientation distribution over a grain, unusual carbide particle distribution. Detection of anomalies is the first step towards identifying defects that could possibly weaken the strength of the materials or cause fatigue failure while in operation. Accurate detection and identification of anomalies can lead to better safety and reliability of materials and can lower the cost of production. Image fusion enhances the power of anomaly detection since the anomalies may be noticeable only when the information across modalities is suitably combined.

The first topic developed in this thesis is anomaly-driven fusion in materials science. Specifically, we propose an anomaly-driven image fusion framework to perform the task of material microscopy image analysis and anomaly detection. This framework is based on a probabilistic model that enables us to index, process and characterize the data with systematic and well-developed statistical tools. We briefly introduce each of these chapters below.

### 1.2.1 Statistical Estimation and Clustering of Group-invariant Orientation Parameters

Estimation of crystal orientation is one of the fundamental objectives of polycrystalline materials science. We show in Chapter III how this problem can be cast as a parametric statistical estimation problem where the observed random variable (measured crystal orientation) has a distribution that is invariant to the actions of a finite spherical point symmetry group. In Chapter III, we first show that any such distribution must satisfy a restricted finite mixture representation and then specialize the case of distributions over the sphere that are invariant to the actions of a finite spherical symmetry group to obtain a representation of the class of group-invariant spherical distributions. One class of group-invariant distributions is the group invariance extension of the Von-Mises Fisher (VMF) or Watson distributions (*Mardia and Jupp*, 1999). These distributions are finite mixtures of VMF or Watson distributions that are parameterized by location and scale parameters that specify the distribution's mean orientation and its concentration about the mean, respectively. These parameters can be estimated using an Expectation Maximization (EM) maximum likelihood (ML) estimation algorithm. The EM-ML algorithm can be further used to cluster samples which come from mixtures of multiple group-invariant distributions having different parameters. It is shown that the proposed model is useful to model crystal orientations under the spherical point symmetry group associated with the crystal form, e.g., cubic or octahedral or hexahedral. We establish the advantages of this model through simulations and experiments for data acquired by EBSD microscopy of a polycrystalline Nickel alloy sample.

### 1.2.2 A Dictionary Approach to EBSD Indexing

Next, we propose a framework for indexing of grain and sub-grain structures in EBSD images of polycrystalline materials. In Chapter IV, we describe a novel

approach that is based on a previously introduced physics-based forward model by (*Callahan and De Graef*, 2013) relating measured patterns to grain orientations (Euler angle). The forward model is tuned to the microscope and to the crystal symmetry group. We discretize the domain of the forward model onto a dense grid of Euler angles and for each measured pattern we identify the most similar patterns in the dictionary. These patterns are used to identify boundaries, detect anomalies, and index crystal orientations. The statistical distribution of these closest matches is used in an unsupervised binary decision tree (DT) classifier to identify grain boundaries and anomalous regions. Indexing is accomplished by computing the mean orientation of the closest dictionary matches to each pattern. The mean orientation is estimated using our proposed group-invariant model of Chapter III with a maximum likelihood approach. The proposed dictionary matching approach permits segmentation, anomaly detection, and indexing to be performed in a unified manner with the additional benefit of uncertainty quantification. We demonstrate the proposed dictionary-based approach on a Ni-base IN100 alloy.

### 1.2.3 Coercive Region-level Registration for Multi-modal Images

After the microscopy images are indexed and segmented, the misalignment problem between modalities needs to be addressed before fusing their information. In Chapter V, we propose a coercive approach to simultaneously register and segment multi-modal images that share similar spatial structure. This novel approach allows the registration to be done at the region level to facilitate data fusion while avoiding the need for interpolation. The algorithm performs alternating minimization of an objective function defined by the statistical model specific to each modality. Hypothesis tests are developed to determine whether to refine segmentations by splitting regions. We demonstrate that our approach has significantly better registration and segmentation accuracy than the state-of-the-art registration and segmentation methods for

microstructure imaging in materials science.

## 1.3 Multi-modal Registration for Biomedical Images

The second part of the thesis focuses on the multi-modality biomedical image registration problem. Medical imaging has been an important component for biomedical research, clinical diagnosis, and surgical procedures planning and evaluating. Many different technologies have been developed for biomedical imaging, such as X-ray, Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and Ultrasound (US). It has been proven that integrating over multi-modal images is valuable for disease characterization (*Maintz and Viergever*, 1998; *Hill et al.*, 2001; *Oliveira and Tavares*, 2014). In this thesis we focus on the multi-modal registration problem for functional MRI (fMRI) brain images which helps us to analyze different modality images for the brain activation detection task.

### 1.3.1 Multimodality Motion Compensated fMRI Using Slice-to-Volume Registration

Head movement during fMRI image acquisition introduces artificial signal changes and misalignment between images that impedes activation detection and localization in brain studies. Previous work (*Park et al.*, 2004) registers the observed Echo Planar Imaging (EPI) slices onto an anatomical volume to estimate and compensate the head motion. Due to the fact that the objective function, commonly defined as Mutual Information (*Maes et al.*, 1997), used in the registration process is usually not convex and has many local maxima, incorporation of information concerning continuous head motion can improve the registration result. We explore the challenging case where the head motion is estimated from the fMRI data and an anatomical pre-scan MRI image taken prior to the activation study. In Chapter VI, we develop a Gaussian particle filter based head motion tracking algorithm, which uses a state space model

to enforce the smooth motion constraint. This results an algorithm that has improved registration and activation detection accuracy as compared to state-of-the-art approaches.

## 1.4 Outline of the Thesis

The thesis is composed of two main topics: ***I. Anomaly-driven Image Fusion in Material Microscopy Images*** and ***II. Multi-modal Registration for Biomedical Images***. In Chapter II, we review background of image fusion in these two domains and discuss different types of frameworks to fuse the images. The first topic consists of Chapter III, IV and V. In Chapter III. we propose a statistical model to describe and analyze the crystal orientations which posses certain symmetry properties associated with the materials particular polycrystalline structure; In Chapter IV, we propose a novel dictionary-based indexing method for Electron Backscatter Diffraction (EBSD) data; In Chapter V, an anomaly-driven registration algorithm based on the statistical model is developed to address the correspondence problem between different modalities. The second topic is discussed in Chapter VI, in which we develop a registration-based method which utilizes the information from different modalities to estimate and compensate the head movement during fMRI scan. In Chapter VII, we conclude this thesis and give possible directions for future works.

## 1.5 Publications

The publications that have come out of research presented in this thesis are listed as follows:

[**1**] Chen, Y.-H., D. Wei, G. Newstadt, M. De Graef, J. Simmons, and A. O. Hero III (2015c), Parameter Estimation in Spherical Symmetry Groups, in *International*

*Conference on Acoustic Speech and Signal Processing (ICASSP)* **Oral Presentation**

[**2**] Chen, Y.-H., D. Wei, G. Newstadt, M. De Graef, J. Simmons, and A. O. Hero III (2015d), Parameter Estimation in Spherical Symmetry Groups, *Signal Processing Letters, IEEE*, *22*(8), 1152–1155, doi: 10.1109/LSP.2014.2387206

[**3**] Chen, Y.-H., D. Wei, G. Newstadt, M. De Graef, J. Simmons, and A. O. Hero III (2015e), Statistical Estimation and Clustering of Group-invariant Orientation Parameters, in *18-th International Conference on Information Fusion (Fusion)* **Oral Presentation**

[**4**] Chen, Y.-H., S. U. Park, D. Wei, G. Newstadt, M. A. Jackson, J. P. Simmons, M. De Graef, and A. O. Hero (2015a), A Dictionary Approach to Electron Backscatter Diffraction Indexing, *Microscopy and Microanalysis*, *21*(03), 739–752

[**5**] Chen, Y.-H., D. Wei, G. Newstadt, M. De Graef, J. Simmons, and A. O. Hero III (2015b), Coercive Region-level Registration for Multi-modal Images, in *International Conference on Image Processing (ICIP)* **Top** 10% **paper**

[**6**] Chen, Y.-H., R. Mittelman, B. Kim, C. Meyer, and A. O. Hero III (2016b), Particle Filtering for Slice-to-volume Motion Correction in EPI Based Functional MRI, in *International Conference on Acoustic Speech and Signal Processing (ICASSP)*, p. submitted

[**7**] Chen, Y.-H., R. Mittelman, B. Kim, C. Meyer, and A. O. Hero III (2016a), Multimodal MRI Neuroimaging with Motion Compensation Based on Particle Filtering, *IEEE Transaction on Medical Image*, p. submitted

# CHAPTER II

# Overview of Image Fusion

## 2.1 Concepts of Image Fusion

A general description of image fusion is to *"intelligently combine two or more different image modalities by using a certain algorithm to achieve refined or improved information for decision making"* (*Van Genderen and Pohl*, 1994; *Hall and McMullen*, 2004). The integration of different modalities provides more information than one can derive from each of the single modality image alone, increasing interpretation capabilities and more reliable results. The idea of image fusion has been widely applied in different research areas including, biology, computer vision, remote sensing...etc. In materials science, due to the fact that the material's response and behavior involve a bunch of physical phenomena with no single overarching modeling approach, integrating signals over different modalities gives us comprehensive information for characterizing the materials (*Fish*, 2006). In biomedical imaging, fusion over multimodal images has been proposed and proven that the unified, integrated information is valuable for disease characterization (*Segal et al.*, 2007; *Viswanath et al.*, 2012; *Gevaert et al.*, 2012; *Rusu et al.*, 2013).

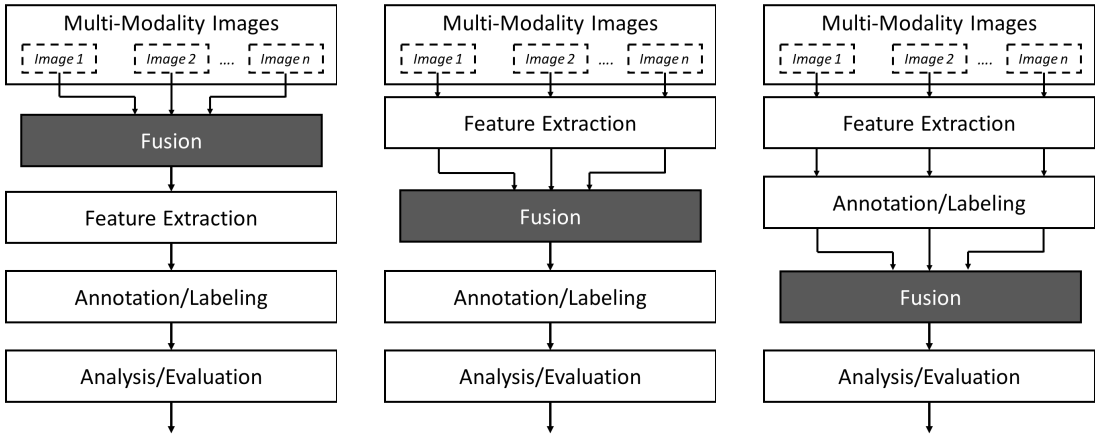
Image fusion is capable of leading to more accurate data, increased confidence, reduced ambiguity, improved reliability, and improved classification (*Keys et al.*, 1990; *Rogers and Wood*, 1990). There are several objectives of image fusion as pointed out

in (*Pohl and Van Genderen*, 1998). For example, image fusion can be used as a tool to increase the spatial resolution (*Franklin and Blodgett*, 1993; *Pellemans et al.*, 1993). Also, fusion of modalities can enhance certain features or substitute missing information not visible in either of the single modality (*Leckie and others*, 1990; *Aschbacher and Lichtenegger*, 1990). The goal of the first topic of this thesis (***Anomaly-driven Image Fusion in Material Microscopy Images***) is to utilize complementary information from different microscopy imaging modalities to detect anomaly/defect in materials. Besides, it can be used to improve geometric corrections and registration accuracy (*Strobl et al.*, 1990; *Ehlers*, 1991). The second topic of this thesis (***Multi-modal Image Registration in Biomedical Images***) is to take advantage of multi-modal functional MR images to improve the registration for head motion estimation and perform better motion compensated images reconstruction for brain activity analysis.

According to the stage at which the fusion takes place, image fusion can be categorized into three different types: *Pixel-level fusion*, *Feature-level fusion* and *Decision-level fusion* (*Pohl and Van Genderen*, 1998; *Mitchell*, 2007), described as follows:

### Pixel-level fusion

Image fusion at pixel level means fusion at the lowest processing level referring to the merging of measured physical parameters as illustrated in Fig 2.1(a). The alignment and geocoding of different modality data plays an important role because the image misalign may introduce artificial noise and falsify the interpretation after fusion. Therefore, pixel-level methods fuse two images by transforming one of the images in such a way to make it the most similar to the other with respect to some objective measure of similarity. Specifically, a parameterized class of transformations is specified, e.g., rotation, affine linear, rigid body transform (*Hill et al.*, 2001), or smooth non-linear (*Meyer et al.*, 1997; *Rueckert et al.*, 1999; *Rohr et al.*,




(a) Pixel-level fusion (b) Feature-level fusion (c) Decision-level fusion

Figure 2.1: Three types of image fusion: (a) Pixel-level fusion merges the measured physical parameters at the lowest processing level. (b) Feature-level fusion extracts objects of interest in the various image modalities and then fuses the objects for further assessment using statistical approaches. (c) Decision-level fusion uses the annotated/labeled data where the input images are processed individually for information extraction and applies decision rules to combine the information to reinforce common interpretation and provide a better understanding of the observed objects.

2001; *Zitova and Flusser*, 2003), and the parameters are varied to give the highest similarity between the images. For example, in material microscopy images, Latham et al. use the correlation coefficient as the similarity metric to register a BSE image to a X-ray micro-CT images (*Latham et al.*, 2008); Cao et al. take mutual information as the metric to register between BSE, SEM and TEM images for correlative microscopy analysis (*Cao et al.*, 2014). In biomedical images, image registration has been applied to many different modalities, e.g., X-ray, ultrasound, magnetic resonance imaging (MRI), computed tomography (CT), positron emission tomography (PET) images...etc (*Maintz and Viergever*, 1998; *Hill et al.*, 2001; *Oliveira and Tavares*, 2014).

We propose a new pixel level fusion algorithm in Chapter VI where the similarity measure is an empirical estimate of the mutual information and we utilize the complimentary information between modalities to better estimate and compensate the artificial noise introduced by head motion for brain MRI images.

**Feature-level fusion**

Fusion at feature level first extracts objects of interest in the various image modalities, e.g., using segmentation procedures. These similar objects (e.g., regions) from different modality images are then fused for further assessment using statistical approaches as illustrated in Fig 2.1(b). The extracted features are more informative than single pixel since they serve as meaningful units that describe the content. For example, the grain regions in material microscopy images determine the material mechanical properties (*Rice*, 2000). Many different segmentation algorithms have been developed for materials microscopy and biomedical images. For example, the watershed segmentation algorithm is proposed and applied on TEM and SEM images to find cleavage fractures in steel (*Beucher and others*, 1992); Voronoi-based approach (*Bachmann et al.*, 2011) has been widely used to segment the grains for EBSD image; The stabilized inverse diffusion equations (SIDEs) algorithm has been demonstrated to be effective on ion-induced secondary electrons (ISE) images to find the grain regions in Ni samples (*Huffman et al.*, 2008; *Chuang et al.*, 2008); Kapur et. al. (*Kapur et al.*, 1996) combined adaptive segmentation, binary morphology, and active contour methods to segment the brain tissues from MR images; An improved watershed transform algorithm which utilizes the information from atlas registration has been proposed to segment the knee cartilage and gray/white matter in MR images (*Grau et al.*, 2004); Yong et. al. ( *Yong et al.*, 2010) fused the MR and CT images through wavelet transform features and construct the fused image which is more robust to noise and provides more accurate clinical information. Image fusion after feature extraction requires the correspondences of the features or regions between modalities which is not easily obtained, especially when these extraction algorithms produce inconsistent results in different modalities.

**Decision-level fusion**

Decision-level fusion uses the annotated/labeled data where the input images are processed individually for information extraction as illustrated in 2.1(c). Decision rules are applied to combine the information to reinforce common interpretation and provide a better understanding of the observed objects (*Shen*, 1990). The main advantage of this approach is that the higher level representations of the data summarize the object's characteristics (*Moissinac et al.*, 1995; *Groeber et al.*, 2008, 2009; *Groeber*, 2011), making multi-modality fusion more robust and reliable. In many applications, decision-level fusion is considered as the primary multi-modal data fusion algorithm. For example, Movellan et al. applied this algorithm in automatic speech recognition to prevent catastrophic fusion (*Movellan and Mineiro*, 1998); An indirect tire pressure monitoring system using decision-level fusion is proposed in (*Persson et al.*, 2002); Inza et al. adopted similar approach in the feature selection task in DNA microarray experiments (*Inza et al.*, 2004). A diagnostic system which is composed of self-organizing map classifiers, and combines techniques based on a confidence measure has been proposed to improve diagnostic accuracy (*Christodoulou et al.*, 1999).

Decision-level fusion can also be fed back to the segmentation and annotation stages in an iterative manner in order to improve these up-stream operations, resulting in improved fusion. We propose the anomaly-driven image fusion which is a type of decision-level fusion with feedback. Our approach utilizes the decision-level annotation as feedback to the feature extraction/segmentation stage, significantly improving the segmentation result.

In Section 2.2, we introduce multi-modal image fusion in materials and our new anomaly-driven image fusion framework. In Section 2.3, we present multi-modal image fusion in biomedical images and our proposed multi-modal registration for head motion compensation approach.

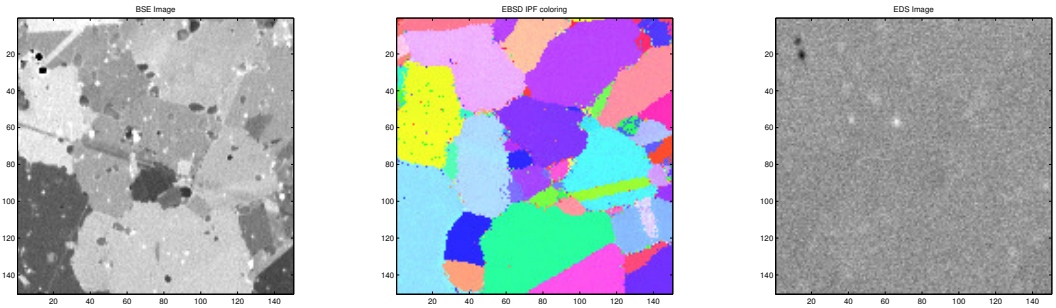
## 2.2 Multi-modal Image Fusion in Materials

The applications of advanced materials in national security, human health and clean energy are important to human well-being and to achieving global competitiveness. President Obama launched the Materials Genome Project (MGI) in 2011 to replicate the successes of Human Genome Project (HGP) in computational biology onto computational materials science, aiming to "discover, develop, manufacture, and deploy advanced materials at least twice as fast as possible today, at a fraction of the cost." To achieve the goal, MGI proposed to develop new integrated computational, experimental, and data informatics tools as an open source platform, which allow researchers to share their knowledge and collaborate conveniently. The developed tool could be used to predict materials behavior and construct the taxonomy of materials properties as well as engineering principles. In order to identify fundamental mechanisms and define the models that describe materials behavior, gathering different signals of the materials and analyzing the integrated data is an essential step.

One of the important purposes of material characterization is anomaly detection. Detecting the anomalies help us understand whether the produced material has enough strength, tensility or other physical properties permitting it to be safely operated under the situation for which it is designed. Early-stage anomaly detection further enables us to discover possible failures in production time that we could change or adjust the process, achieving higher product yield and lower production cost. In the detection process, two types of events are of prime interest: rare event and anomalous events. Rare events occur naturally in tails of nominal feature distribution and may presage anomalies. For example, crack formation of a sample may start as a rare event that it would be very helpful if we are able to identify potential crack before it develops. Anomalous events are the observations with exceedingly low probability in nominal distribution, indicating incipient failures. Fusion over multimodal data is able to provide more comprehensive information, resulting in higher

detection power. For example, large-scale carbides exist along the grain boundary can impair the material strength. This type of anomaly requires the detection of large carbide which is only obvious in Scanning Electron image and the detection of grain boundary which is distinct in Electron Backscatter Diffraction image.

### 2.2.1 Description of SEM modalities




(a) BSE Image (b) EBSD data in IPF coloring (c) EDS Image

Figure 2.2: (a) Backscattered Electron (BSE) image in gray scale. There are lots of surface textures (dots and lines) revealed by BSE. (b) Electron Backscatter Diffraction (EBSD) image in Inverse Pole Figure (IPF) coloring scheme. Notice that there are less textures as in BSE but the grain regions and boundaries are much clearer in EBSD. (c) Spectral unmixing result of Energy-Dispersive X-Ray (EDS) data. It has very little grain structure or precipitates information as BSE or EBSD but it has higher contrast for the region of the sample with different chemical composition (two black dots on the top left corner).

Scanning Electron Microscope produces microscopy images of a material by capturing the signals coming from the interaction between the focused electron beam and the atoms in the sample. By controlling the energy of the electron beam and detecting the electrons interacted with the material in different ways, the signals are able to reveal information of the sample in different aspects: Backscattered Electron (BSE) imaging is more sensitive to the atomic mass of the nuclei but suffers lower surface topographical characteristics resolution. As shown in Fig.2.2(a), the BSE image shows clear textures of the sample which corresponds to the precipitates; Electron Backscatter Diffraction (EBSD) is capable of determining the crystal orientation by

capturing the diffracted electrons from atomic layers in crystalline materials. Figure 2.2(b) is the Inverse Pole Figure (IPF) coloring scheme of the EBSD crystal orientation. We can see that it has less power to detect those precipitates but provides very clear grain regions and grain boundaries; Energy-Dispersive X-Ray Spectroscopy (EDS) records the energy spectrum of the sample, providing the information of the chemical composition of the materials. Figure 2.2(c) shows the robust spectral unmixing (*Newstadt et al.*, 2014) result for the EDS data with the same sample region that it has very little precipitates information or grain structure as BSE or EBSD but has higher contrast for the region in the sample with different chemical composition (two black dots on the top left corner).

One important fact of these modalities is that they are usually acquired asynchronously and have different spatial resolution. The actual resolution depends on the electron beam diameter and the type of electron source, probe and the specimen. In general, BSE image has the best spatial resolution of less than $0.5nm$ with scalar value (intensity) for each pixel. EBSD is able to identify the pixel as small as $10nm$, where each pixel contains a small (e.g. $80 \times 60$ pixels$^2$) gray scale image of diffraction pattern. The diffraction patterns can be further translated to crystal orientation, which has 3 degrees of freedom at each pixel location. EDS provides a spectrum (e.g. vector of size 2048) for each pixel but has lower spatial resolution as $2\mu m$. The difference in spatial resolution between different modality data complicates the process of information fusion. Furthermore, due to the different experiment settings of the machines, e.g. electron beam incident angles or recorders, the acquired images may have certain level of distortions. Many approaches have been proposed that use the fiducial marks with manual and automated alignment tools to alleviate the misalignment issue (*Fung et al.*, 1996; *Ress et al.*, 1999; *Brandt et al.*, 2001). Recently, acquisition methods which allow for simultaneous measurements of different signals have been developed. For example, Scanning Transmission Electron Microscopy (STEM) (*Nellist*,

2007; *Pennycook and Nellist*, 2011) can measure bright field, high angle dark field and EDS spectra signals simultaneously; The TriBeam technique (*Echlin et al.*, 2012) can provide simultaneously SE, BSE, EDS and EBSD on a single microscopy platform. However, even though these systems collect multi-modality data synchronously, the differences in the physics of the image formation process in each modality require compensation. For instance, ion-induced secondary electrons come from a shallower sub-surface region while X-rays come from much deeper region in the sample. Therefore, in order to fuse the multi-modal data, registration algorithms must be developed to handle the different resolutions and compensate the misalignment between modalities.

### 2.2.2 Anomaly-driven Fusion of Scanning Electron Microscopy Images

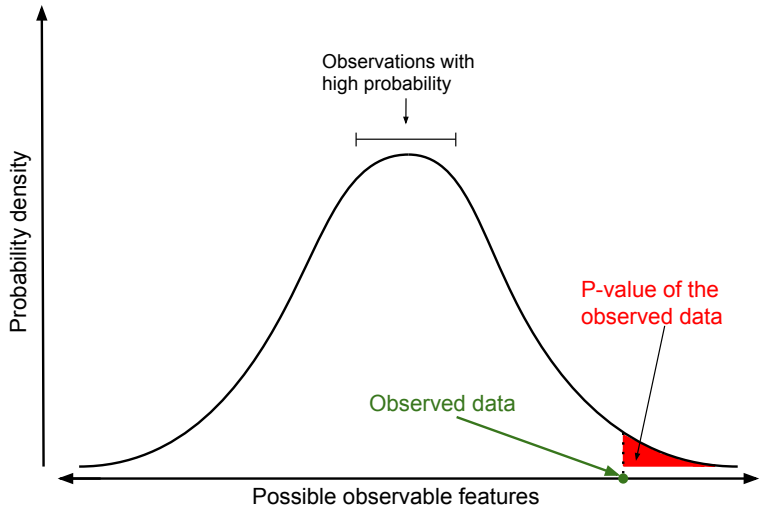



Figure 2.3: The figure shows the relation between the distribution which models the possible observations and the $P$-value of the actual observed data. The vertical coordinate is the probability density of each outcome, calculated from the null hypothesis. The $P$-value is the red area under the tail of the curve past the observed data point, which is associated with testing the null hypothesis that $T = 0$ vs $T > 0$.

Anomaly-driven data fusion utilizes the anomaly information as feedback to drive the image indexing, feature extraction and registration process. Here anomalies are

defined as properties of the images that are inconsistent with a null hypothesis, e.g., homogeneity of the pixel values within a segmented region (a grain). As such the theory of hypothesis testing (*Neyman and Pearson*, 1992) provides the framework for anomaly-driven fusion. In this framework, a null hypothesis is defined and this null hypothesis is tested using test statistics derived from the image samples, or homologous subregions of these samples, in each modality. Under mild distributional assumptions, the level of significance or $P$-value of the test statistic can be computed for each modality and these $P$-values can be fused. The $P$-value is defined as the probability, under the assumption of the null hypothesis, of obtaining a result equal to or more extreme than what was actually observed (*Fisher*, 1934). For the materials anomaly detection, the null hypothesis, $H_0$, and the test statistic, $T$, is designed according to our underlying knowledge about the normal event. Given the data, we can calculate the observed test statistic $t_{obs}$ and evaluate the $P$-value. Figure 2.3 illustrates the $P$-value computation. The vertical coordinate is the probability density of the test statistic under the null hypothesis. The $P$-value is the red area under the tail to the right of the computed test statistic $t_{obs}$ of a given sample. If the null hypothesis models the normal event accurately, then by thresholding the $P$-value at various levels one can distinguish between the normal event, rare event and anomaly event with prescribed level of statistical significance.

Hypothesis testing with $P$-value evaluation not only enables us to measure the anomaly level quantitatively but also provides a natural way to fuse the information from different modalities. Since the $P$-value represents the probability of observing a sample test statistic at least as extreme as observed, the $P$-value associated with $M$ statistically independent modalities is given by the exclusion/inclusion principle as:

$$P_{fuse} = 1 - \prod_{i=1}^{M}(1 - P_i), \tag{2.1}$$

where $P_i$ are the $P$-values for the $M$ different modalities and $P_{fuse}$ is the fused $P$-value. The simplicity of (2.1) facilitates the fusion process and motivates our statistical based anomaly-driven image fusion framework. The image modeling, indexing, registration, and fusion will be illustrated in Chapter III, IV, V.

## 2.3 Multi-modal Image Fusion in Biomedical Imaging

The main objective of medical image fusion is to increase the clinical applicability of medical images for diagnosis and assessment of medical problems, advancing the research and study of human body, organs, and cells (*Constantinos et al.*, 2001; *James and Dasarathy*, 2014). The process of medical image fusion involves registering and combining multiple imaging modalities to improve the imaging quality and reduce redundancy, providing more accurate and comprehensive information. The signals from multi-modal images offer a greater diversity of the features which are helpful for medical analysis applications, e.g., brain tissue mapping and volume identification (*Kok et al.*, 1996; *Barra and Boire*, 2000), breast cancer detection and surgery (*Raza et al.*, 2005, 2006; *Kirova et al.*, 2011), prostate localization and motion modeling (*Hervas et al.*, 2004; *Hu et al.*, 2011)...etc.

There are several medical imaging modalities which can serve as primary inputs to image fusion studies. Practically, it is impossible to capture every aspects of the object from single modality and therefore combining the information from different modalities ensure clinical accuracy and robustness of the analysis and resulting diagnosis. Some of the major modalities are listed as follows: (a) Computed Tomography (CT): an imaging method that utilizes multiple X-ray projections taken from different angles to construct detailed cross-sectional images; (b) Positron Emission Tomography (PET): a nuclear imaging technique which provides information about how tissues and organs are functioning; (c) Ultrasound: use high frequency sound waves' echo returned from body to create images of the inside of the body; (d) Magnetic Reso-

nance Imaging (MRI): use radio waves and a magnetic field to create detailed images of organs and tissues (*Bushberg and Boone*, 2011).

### 2.3.1 Multi-modal Head Motion Compensated Functional MRI Using Slice-to-Volume Registration

Magnetic resonance imaging (MRI) uses a magnetic field and pulses of radio wave energy to create images of organs and structures inside the body, providing different information from X-ray, ultrasound, or computed tomography scan. The advantage of MRI is its safety for the subject since it does not involve any exposure to radiation. Furthermore, it has high accuracy for the soft tissue structures such as brain, heart and eyes. However, the MRI images are relative sensitive to movement, increasing the difficulty of assessing organs that involve movement (*James and Dasarathy*, 2014). Therefore, the image fusion techniques can overcome this limitation in a multi-modal imaging environment, enabling reconstruction and prediction of the missing information (*Marshall and Matsopoulos*, 1993; *Erie et al.*, 1999; *Lindseth et al.*, 2001).

There are several modalities in MRI images depending on which emitted signal is captured. For example, T1-weighted image demonstrates differences in the T1 relaxation time which relies upon the longitudinal relaxation of a tissue's net magnetization vector. In T1-weighted image, fat appears bright while water has low signal and appears dark. On the other hand, T2-weighted image highlights differences in the T2 relaxation time, which relies upon the transverse relaxation of the net magnetization vector. Unlike T1-weighted image, water and fluid tend to be brighter in T2-weighted image.

#### 2.3.1.1 Functional MRI Image Acquisition

Functional MRI (fMRI) refers to the procedure which uses MRI technology to measure brain activity by detecting associated changes in blood flow, called blood-

oxygen-level dependent (BOLD) signal (*Ogawa et al.*, 1990). Echo planar imaging (EPI) (*DeLaPaz*, 1994) is a fast MR imaging technique captures T2-weighted signal that allows acquisition of single images in as little as $20 - 100$ msec. The excellent temporal resolution of EPI makes it a favorable technique for fMRI. However, it is more vulnerable to movement and geometric distortion (*Schmitt et al.*, 1998) due to its lower spatial resolution. Therefore, with the help of T1-weighted anatomical volume, we can better reconstruct and compensate the noise introduced by head motion in EPI images.

The image acquisition procedure could be separated into two parts. At first, the subject is taken a full scan to get a higher resolution T1-weighted anatomical volume. Secondly, the subject is asked to alternatively perform certain tasks (stimulation period), e.g., finger tapping, picture naming, speaking...etc, and rest (control period). During each of the period, we repeatedly take EPI images of the subject to capture possible signal changes in activated region due to those tasks. Generally, this on-off procedure is executed for several cycles until enough volumes in both stimulation and control period are collected as illustrated in Fig 2.4.

#### 2.3.1.2 Particle Filtering for Slice-to-volume Motion Correction in EPI based fMRI

Head movement during scanning impedes activation detection in fMRI studies. Head motion in fMRI acquired using slice-based Echo Planar Imaging (EPI) can be estimated and compensated by aligning the images onto a reference volume through image registration. However, registering EPI images volume to volume fails to consider head motion between slices, which may lead to severely biased head motion estimates. Slice-to-volume registration can be used to estimate motion parameters for each slice by more accurately representing the image acquisition sequence. However, accurate slice to volume mapping is dependent on the information content of the

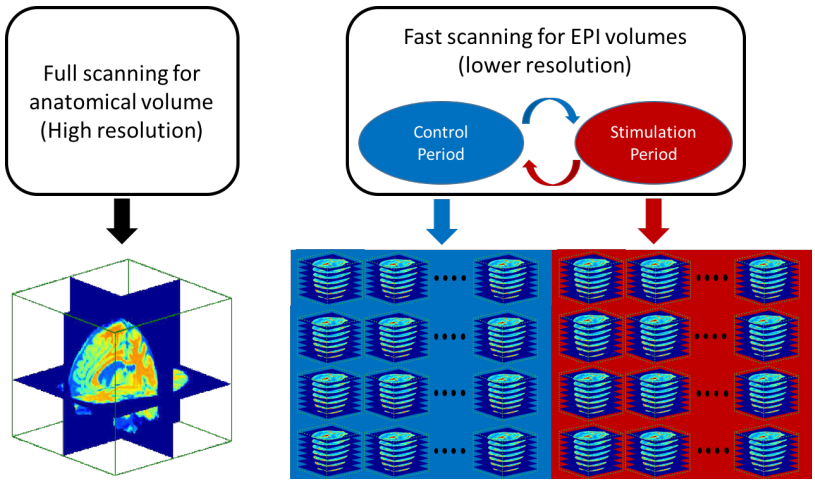



Figure 2.4: The subject is taken a full scan to get a higher resolution T1-weighted anatomical volume (left figure). The subject is asked to alternatively perform certain tasks (stimulation period) and rest (control period). During each of the period, the EPI images of the subject are repeatedly taken to capture possible signal changes in activated region due to those tasks. This on-off (red-blue) procedure is executed for several cycles until enough volumes in both stimulation and control period are collected (right figure)

slices: middle slices are information rich, while edge slides are information poor and more prone to distortion. In Chapter VI, we propose a Gaussian particle filter based head motion tracking algorithm to reduce the image misregistration errors. The algorithm uses a dynamic state space model of head motion with an observation equation that models continuous slice acquisition of the scanner. Under this model the particle filter provides more accurate motion estimates and voxel position estimates. We demonstrate significant performance improvement of the proposed approach as compared to registration-only methods of head motion estimation and brain activation detection.

# CHAPTER III

# Statistical Estimation and Clustering of Group-invariant Orientation Parameters

## 3.1 Introduction

In this chapter, we consider the estimation of parameters of distributions whose domain is a particular non-Euclidean geometry: a topological space divided into $M$ equivalence classes by actions of a finite spherical symmetry group. A well known example of a finite spherical symmetry group is the point group in 3 dimensions describing the soccer ball, or football, with truncated icosahedral symmetry that also corresponds to the symmetry of the Carbon-60 molecule. We formulates a general approach to parameter estimation in distributions defined over such domains. First we establish a restricted finite mixture representation for probability distributions that are invariant to actions of any topological group. This representation has the property that the number of mixture components is equal to the order of the group, the distributions in the mixture are all parameterized by the same parameters, and the mixture coefficients are all equal. This is practically significant since many reliable algorithms have been developed for parameter estimation when samples come from finite mixture distributions (*Dempster et al.*, 1977).

*This chapter is based on (*Chen et al.*, 2015c,d,e)

We illustrate the power of the representation for an important problem in materials science: analysis of mean orientation in polycrystals. Crystal orientation characterizes properties of materials including electrical conductivity and thermal conductivity. Polycrystalline materials are composed of grains, of varying size and orientation, where each grain contains crystal forms with similar orientations. The quality of the material is mainly determined by the grain structure i.e. the arrangement of the grains, their orientations, as well as the distribution of the precipitates. Thus accurate estimation of crystal orientation of the grains is useful for predicting how materials fail and what modes of failure are more likely to occur (*De Graef and McHenry*, 2007).

The mean orientation of the grain, characterized for example by its Euler angles, can only be specified modulo a set of angular rotations determined by the symmetry group associated with the specific type of crystal, e.g. hexagonal, cubic. This multiplicity of equivalent Euler angles complicates the development of reliable mean orientation estimators. The problem becomes even harder when the orientations are sampled from a region encompassing more than one grain such that the orientations cluster over different mean directions. In such case, we would like to identify whether the orientations are multi-modally distributed and also estimate the mean direction for each cluster.

To handle the above problems, we extend the Von Mises Fisher (VMF) and Watson models (*Mardia and Jupp*, 1999) under the proposed finite mixture representation, and apply the expectation maximization (EM) maximum likelihood (ML) algorithm for mixtures (*Dempster et al.*, 1977). The proposed approach provides an accurate iterative estimator of the mean Euler angle parameter and angular concentration parameter of the extended distribution. Specifically, the extension is accomplished as follows. We start with the standard spherical distribution model (VMF or Watson), which is a density parameterized by location (angle mean) and scale (angle

concentration) defined over the $p$-dimensional sphere. In this model, a point on the sphere is specified by its direction vector, and the angle between two vectors is the arc-cosine of the normalized inner product between them. The spherical symmetry group extension is accomplished by applying the mixture representation to the standard distribution using the group of quaternion rotation matrices $\mathcal{G}$, resulting in a $\mathcal{G}$-invariant distribution. Additionally, based on the $\mathcal{G}$-invariant model, we propose to use the Generalized Likelihood Ratio Test (GLRT) to detect the presence of multiple modes and a mixture of $\mathcal{G}$-invariant models with EM algorithm to estimate the parameters for each cluster.

The performance of the proposed EM orientation estimators is evaluated by simulation and compared to other angle estimators. The ML orientation estimator is then illustrated on Electron Backscatter Diffraction EBSD data collected from a Nickel alloy whose crystal form induces the $m\overline{3}m$ cubic point symmetry group. We establish that the ML orientation estimator results in significantly improved estimates of the mean direction in addition to providing an accurate estimate of concentration about the mean. Furthermore, with the extended mixture model, we are able to identify and cluster multi-modally distributed samples more accurately than a standard K-means algorithm.

The rest of this chapter is organized as follows. Section 3.2 describes group invariant random variables and gives the mixture representation for their densities. Section 3.3 specializes to random variables invariant relative to actions of the spherical symmetry group and develops the $\mathcal{G}$-invariant VMF and Watson distributions along with EM-ML parameter estimator. The clustering methods based on the $\mathcal{G}$-invariant distributions along with the GLRT are elaborated in Section 3.4. The crystallography application and data simulation are presented in Section 3.5 and the experiment results are shown in Section 3.6. Section 3.7 has concluding remarks.

## 3.2 Group-invariant random variables

Consider a finite topological group $\mathcal{G} = \{G_1, \ldots, G_M\}$ of $M$ distinct actions on a topological space $\mathcal{X}$, $G_i : \mathcal{X} \to \mathcal{X}$ and a binary operation "$*$" defining the action composition $G_i * G_j$, denoted $G_i G_j$. $\mathcal{G}$ has the properties that composition of multiple actions is associative, for every action there exists an inverse action, and there exists an identity action (*Birkhoff and Mac Lane*, 1963). A real valued function $f(\mathbf{x})$ on $\mathcal{X}$ is said to be invariant under $\mathcal{G}$ if: $f(G\mathbf{x}) = f(\mathbf{x})$ for $G \in \mathcal{G}$. Let $\mathbf{X}$ be a random variable defined on $\mathcal{X}$. We have the following theorem for the probability density $f(\mathbf{x})$ of $\mathbf{X}$.

**Theorem III.1.** *The density function $f : \mathcal{X} \to \mathbb{R}$ is invariant under $\mathcal{G}$ if and only if*

$$\begin{aligned} &\exists\ h : \mathcal{X} \to \mathbb{R}\ s.t. \\ &\quad f(\mathbf{x}) = \frac{1}{M} \sum_{i=1}^{M} h(G_i \mathbf{x}). \end{aligned} \tag{3.1}$$

*Proof:* If (3.1) holds then $f(G\mathbf{x}) = M^{-1} \sum_{i=1}^{M} h(G_i G\mathbf{x})$. Since $\mathcal{G}$ is a group $\mathcal{G}G = \mathcal{G}$ so that

$$f(G\mathbf{x}) = \frac{1}{M} \sum_{i=1}^{M} h(G_i G\mathbf{x}) = \frac{1}{M} \sum_{j=1}^{M} h(G_j \mathbf{x}) = f(\mathbf{x}) \tag{3.2}$$

On the other hand, if $f(G\mathbf{x}) = f(\mathbf{x}), \forall G \in \mathcal{G}$ then let $h(.) = f(.)$ we have

$$\frac{1}{M} \sum_{i=1}^{M} f(G_i \mathbf{x}) = \frac{1}{M} \sum_{i=1}^{M} f(\mathbf{x}) = f(\mathbf{x}) \quad \square \tag{3.3}$$

Theorem III.1 says that any density $f(\mathbf{x})$ that is invariant under group $\mathcal{G}$ can be represented as a finite mixture of some function's translates $h(G_i \mathbf{x})$ under the group's actions $G_i \in \mathcal{G}$. This simple result has important implications on $\mathcal{G}$-invariant density estimation and parameter estimation. In particular it can be used to construct

maximum likelihood estimators for parametric densities and kernel density estimators of non-parametric $\mathcal{G}$-invariant densities with finite sample guaranteed performance.

To illustrate the non-parametric case, assume that $\mathcal{X}$ has topological dimension $d$ with Lebesgue $\mathcal{G}$-invariant density $f(\mathbf{x})$. Define the symmetric non-negative second order kernel function $\phi : \mathcal{X} \to \mathbb{R}$, i.e., $\phi(\mathbf{x}) \geq 0$, $\phi(\mathbf{x}) = \phi(\|\mathbf{x}\|, 0, \ldots, 0)$, $\int \phi(\mathbf{x}) d\mathbf{x} = 1$, and $\int \|\mathbf{x}\|^2 \phi(\mathbf{x}) d\mathbf{x} < \infty$. For the finite group $\mathcal{G}$, define the $\mathcal{G}$-invariant kernel function $K(\mathbf{x}) = M^{-1} \sum_{i=1}^{M} \phi(G_i \mathbf{x})$. Given a realization $\{\mathbf{x}_i\}_{i=1}^n$ of $n$ i.i.d. samples from $f$ define the kernel density estimator $\hat{f}_h(\mathbf{x}) = n^{-1} \sum_{i=1}^{n} K\left(\frac{\mathbf{x}-\mathbf{x}_i}{h}\right)$. Assume that $h_n$ is a sequence of kernel widths that satisfies $\lim_{n\to\infty} h_n = 0$ while $\lim_{n\to\infty} h_n n^d = \infty$. Then, if $f$ is smooth, using Thm III.1 and concentration results from (*Devroye and Lugosi*, 2001), it can be shown that as $n$ goes to infinity

$$E[\|f - \hat{f}_{h_n}\|] = O(n^{-2/(4+d)}),$$

where $\|f - \hat{f}_{h_n}\| = \int |f(\mathbf{x}) - \hat{f}_{h_n}(\mathbf{x})| d\mathbf{x}$ is the $\ell_1$ norm difference.

For the parametric case, let $h(\mathbf{x}; \boldsymbol{\theta})$ be a density on $\mathcal{X}$ that is parameterized by a parameter $\boldsymbol{\theta}$ in a parameter space $\Theta$. We extend $h(\mathbf{x}; \boldsymbol{\theta})$ to a $\mathcal{G}$-invariant density $f$ by using Thm. III.1, obtaining:

$$f(\mathbf{x}; \boldsymbol{\theta}) = \frac{1}{M} \sum_{i=1}^{M} h_i(\mathbf{x}; \boldsymbol{\theta}), \tag{3.4}$$

where $h_i(\mathbf{x}; \boldsymbol{\theta}) = h(G_i \mathbf{x}; \boldsymbol{\theta})$. This density is of the form of a finite mixture of densities $h_i(\mathbf{x}; \boldsymbol{\theta})$ of known parametric form where the mixture coefficients are all identical and equal to $1/M$. Maximum likelihood (ML) estimation of the parameter $\boldsymbol{\theta}$ from an i.i.d. sample $\{\mathbf{x}_i\}_{i=1}^n$ from any $\mathcal{G}$-invariant density $f$ can now be performed using finite mixture model methods (*McLachlan and Peel*, 2004) such as the Expectation-Maximization (EM) algorithm (*Dempster et al.*, 1977) or the restricted Boltzman machine (RBM) (*Sohn et al.*, 2011).

## 3.3 ML within a Spherical Symmetry Group

In this section we specialize to estimation of parameters for the case that the probability density is on a sphere and is invariant to actions in a spherical symmetry group. In Section 3.5 this will be applied to a crystallography example under spherical distribution likelihood models for the mean crystal orientation. The measured and mean orientations can be represented in three equivalent ways.

**Euler angles $\mathcal{E}$:** The orientation is defined by a set of three successive rotations of a reference unit vector about the specified axes (*Eberly*, 2008). Denote the Euler angles as $\mathbf{e} = (\alpha, \beta, \gamma) \in \mathcal{E}$, where $\alpha, \gamma \in [0, 2\pi]$ and $\beta \in [0, \pi]$.

**Quaternion $\mathcal{Q}$:** The quaternion representation describes the orientation as a 4D vector on the 3D sphere (*Altmann*, 2005): $\mathbf{q} = (q_1, q_2, q_3, q_4) \in \mathcal{Q}$, where $\|\mathbf{q}\| = 1$. The main advantage of this representation is that any rotation of $\mathbf{q}$ is simply accomplished via left multiplication by a $4 \times 4$ orthogonal matrix $\mathbf{Q}$ called a quaternion matrix.

**Rodrigues Vector $\mathcal{D}$:** The Rodrigues vector describes the orientation by rotating a reference vector along one direction $\mathbf{v}$ by angle $\theta$ according to the right hand rule (*Rodrigues*, 1840). It is denoted as $\mathbf{d} = \mathbf{v} \tan w/2 = (d_1, d_2, d_3) \in \mathcal{D}$, where $\|\mathbf{v}\| = 1$ and $w \in [0, \pi]$.

In this thesis, we use the quaternion representation to enable orientations to be modeled by spherical distributions since the quaternion representation is a 4D vector on the 3-sphere $S^3$. Any of the aforementioned orientation representations have inherent ambiguity due to crystal symmetries. For example, if the crystal has cubic symmetry, its orientation is only uniquely defined up to a 24-fold set of proper rotations of the cube about its symmetry axes. These actions form a point symmetry group, called 432, a sub-group of $m\overline{3}m$. In quaternion space, since each orientation

corresponds to two quaternions with different sign $\{\mathbf{q}, -\mathbf{q}\}$, these rotations reflections, and inversions can be represented as a spherical symmetry group $\mathcal{G}$ of quaternionic matrices $\{\mathbf{P}_1, \ldots, \mathbf{P}_M\}$, with sign symmetry such that $\mathbf{P}_i = -\mathbf{P}_{i-M/2}\, \forall M/2 < i \leq M$, where $M = 48$ for cubic symmetry. Based on the symmetry group $\mathcal{G}$, we can define the distance between two quaternions under $\mathcal{G}$ as:

$$d_{\mathcal{G}}(\mathbf{q}_1, \mathbf{q}_2) = \min_{\mathbf{P} \in \mathcal{G}} \arccos\left(\mathbf{q}_1^T \mathbf{P} \mathbf{q}_2\right) \tag{3.5}$$

Two quaternions $\mathbf{q}_1, \mathbf{q}_2$ are called symmetry-equivalent to each other if they are mapped to an equivalent orientation under $\mathcal{G}$, i.e. $d_{\mathcal{G}}(\mathbf{q}_1, \mathbf{q}_2) = 0$. A fundamental zone (FZ), also called the fundamental domain, is a conic solid subset of the sphere that can be specified to disambiguate any particular orientation $\mathbf{x}$. However, as will be seen in Sec. 3.5, reduction of the entire data sample $\{\mathbf{x}_i\}_{i=1}^n$ to a FZ destroys information necessary for maximum likelihood estimation: the entire $\mathcal{G}$-invariant density (3.4) must be used. In the following two subsections, we introduce two spherical distributions: von Mises-Fisher and Watson distribution (*Mardia and Jupp*, 1999) to model the orientations in quaternion space.

### 3.3.1 $\mathcal{G}$-invariant von Mises-Fisher Distribution

The von Mises-Fisher (VMF) distribution arises in directional statistics (*Mardia and Jupp*, 1999) as a natural generalization of the multivariate Gaussian distribution to the $(p-1)$-dimensional sphere $S^{(p-1)} \subset \mathbb{R}^p$, where $p \geq 2$. The VMF distribution is parameterized by the mean direction $\boldsymbol{\mu} \in S^{(p-1)}$ and the concentration parameter $\kappa \geq 0$:

$$\phi(\mathbf{x}; \boldsymbol{\mu}, \kappa) = c_p(\kappa) \exp\left(\kappa \boldsymbol{\mu}^T \mathbf{x}\right), \tag{3.6}$$

where $c_p(\kappa) = \frac{\kappa^{p/2-1}}{(2\pi)^{p/2} I_{p/2-1}(\kappa)}$ and $I_p(\cdot)$ is the modified Bessel function of the first kind of order $p$. Given an i.i.d sample $\{\mathbf{x}_i\}_{i=1}^n$ from the VMF distribution the ML estimator

has the closed form expressions (*Mardia and Jupp*, 1999)

$$\hat{\boldsymbol{\mu}} = \frac{\boldsymbol{\gamma}}{\|\boldsymbol{\gamma}\|},\ \hat{\kappa} = A_p^{-1}\left(\frac{\|\boldsymbol{\gamma}\|}{n}\right), \tag{3.7}$$

where $\boldsymbol{\gamma} = \sum_{i=1}^{n} \mathbf{x}_i$ and $A_p(u) = \frac{I_{p/2}(u)}{I_{p/2-1}(u)}$.

Let $\mathcal{G}$ be a group of symmetric actions $\{\mathbf{P}_1, \ldots, \mathbf{P}_M\}$ acting on the quaternionic representation of orientation on the 3-dimensional sphere $S^3$. We extend the VMF distribution (3.6) using the mixture representation in Thm III.1:

$$f_v(\mathbf{x}; \boldsymbol{\mu}, \kappa) = \frac{1}{M}\sum_{m=1}^{M} \phi(\mathbf{P}_m\mathbf{x}; \boldsymbol{\mu}, \kappa) \tag{3.8}$$

$$= \frac{1}{M}\sum_{m=1}^{M} \phi(\mathbf{x}; \mathbf{P}_m\boldsymbol{\mu}, \kappa) \tag{3.9}$$

where in going from (3.8) to (3.9) we used the inner product form $\boldsymbol{\mu}^T\mathbf{x}$ in (3.6) and the symmetry of $\mathbf{P}_m$. The expression (3.9) for the extended VMF distribution is in the form of a finite mixture of standard VMF distributions on the same random variable $\mathbf{x}$ having different mean parameters $\boldsymbol{\mu}_m = \mathbf{P}_m\boldsymbol{\mu}$ but having the same concentration parameters $\kappa$.

The finite mixture (3.9) for the $\mathcal{G}$-invariant density $f_v(\mathbf{x}; \boldsymbol{\mu}, \kappa)$ is in a form for which an EM algorithm (*Dempster et al.*, 1977) can be implemented to compute the ML estimates of $\boldsymbol{\mu}$ and $\kappa$. Denoting the parameter pair as $\boldsymbol{\omega} = \{\boldsymbol{\mu}, \kappa\}$, the EM algorithm generates a sequence $\{\boldsymbol{\omega}^{(k)}\}$ of estimates that monotonically increase the likelihood. These estimates are given by $\boldsymbol{\omega}^{(k+1)} = \mathbf{amax}_{\boldsymbol{\omega}} E_{S|\mathbf{X},\boldsymbol{\omega}^{(k)}}[\log L(\boldsymbol{\omega}; \{\mathbf{x}_i, s_i\})]$, where $s_i$ is a latent variable assigning $\mathbf{x}_i$ to a particular mixture component in (3.9) and $L(\boldsymbol{\omega}; \{\mathbf{x}_i, s_i\})$ is the likelihood function of $\boldsymbol{\omega}$ given the complete data $\{\mathbf{x}_i, s_i\}_{i=1}^n$.

Specifically,

$$
\begin{aligned}
& E_{S|\mathbf{X},\boldsymbol{\omega}}[\log L(\boldsymbol{\omega}; \{\mathbf{x}_i, s_i\})] \\
= & \sum_{i=1}^{n} \sum_{m=1}^{M} r_{i,m} (\log c_p(\kappa) + \kappa (\mathbf{P}_m \boldsymbol{\mu})^T \mathbf{x}_i),
\end{aligned} \tag{3.10}
$$

where $r_{i,m} = P(s_i = m | \mathbf{x}_i, \boldsymbol{\omega})$. The EM algorithm takes the form where the detail derivation is given in Appendix B:

E-step:

$$
r_{i,m} = \frac{\phi(\mathbf{x}_i; \mathbf{P}_m \boldsymbol{\mu}, \kappa)}{\sum_{l=1}^{M} \phi(\mathbf{x}_i; \mathbf{P}_l \boldsymbol{\mu}, \kappa)}, m \in \{1, 2, \ldots, M\}. \tag{3.11}
$$

M-step:

$$
\hat{\boldsymbol{\mu}} = \frac{\boldsymbol{\gamma}}{\|\boldsymbol{\gamma}\|}, \ \hat{\kappa} = A_p^{-1} \left( \frac{\|\boldsymbol{\gamma}\|}{n} \right), \tag{3.12}
$$

$$
\boldsymbol{\gamma} = \sum_{i=1}^{n} \sum_{m=1}^{M} r_{i,m} \mathbf{P}_m^T \mathbf{x}_i. \tag{3.13}
$$

Notice that the proposed mixture has fixed mixing weights $1/M$ and shared concentration parameter $\kappa$, and the mean directions in different mixtures are connected through the symmetry operators $\mathbf{P}_m$. Therefore, the proposed model has much fewer parameters need to be estimated than usual EM algorithm.

The E and M-step can be further simplified by exploiting the aforementioned sign

symmetry of $\mathcal{G}$. Equation (3.13) in M-step can be re-written as:

$$\begin{aligned}\boldsymbol{\gamma} &= \sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}\mathbf{P}_m^T\mathbf{x}_i \\ &= \sum_{i=1}^{n}\left(\sum_{m=1}^{M/2} r_{i,m}\mathbf{P}_m^T\mathbf{x}_i - r_{i,\frac{M}{2}+m}\mathbf{P}_m^T\mathbf{x}_i\right) \\ &= \sum_{i=1}^{n}\sum_{m=1}^{M/2}\frac{\sinh\left(\kappa(\mathbf{P}_m\boldsymbol{\mu})^T x_i\right)}{\sum_{l=1}^{M/2}\cosh\left(\kappa(\mathbf{P}_l\mu)^T x_i\right)}\mathbf{P}_m^T\mathbf{x}_i,\end{aligned} \tag{3.14}$$

where $\sinh, \cosh$ are the hyperbolic sinusoidal functions. Equation (3.11) in E-step is simplified as:

$$r'_{i,m} = \frac{\sinh\left(\kappa(\mathbf{P}_m\boldsymbol{\mu})^T x_i\right)}{\sum_{l=1}^{M/2}\cosh\left(\kappa(\mathbf{P}_l\mu)^T x_i\right)}, m \in \left\{1, 2, \ldots, \frac{M}{2}\right\}. \tag{3.15}$$

The results in Section 3.6 will demonstrate the computational improvement due to the simplification.

### 3.3.2 $\mathcal{G}$-invariant Watson Distribution

As described at the beginning of this section each orientation corresponds to two quaternions with different sign, which is equivalent to an axis of the sphere. For axial data it is more natural to use the Watson distribution (*Watson*, 1965), which models the probability distribution of axially symmetric vectors on the $(p-1)$-dimensional unit sphere, i.e. $\pm\mathbf{x} \in S^{p-1}$ are equivalent. Similar to VMF, the distribution is parametrized by a mean direction $\pm\boldsymbol{\mu} \in S^{p-1}$, and a concentration parameter $\kappa \in \mathbb{R}$. Its probability density function is

$$W_p(\mathbf{x}; \boldsymbol{\mu}, \kappa) = \frac{1}{\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)}\exp\left(\kappa(\boldsymbol{\mu}^T\mathbf{x})^2\right), \tag{3.16}$$

where $\mathbb{M}$ is the Kummer confluent hypergeometric function defined in (*Bateman et al.*, 1955). According to (3.16), the positive-negative pair of group actions $\{\mathbf{P}_m, -\mathbf{P}_m\}$ contribute the same value in the density function. The set of the group action pairs $\mathcal{G}' = \{\{\mathbf{P}_m, -\mathbf{P}_m\}\}_{m=1}^{M/2}$ is the quotient group $\mathcal{G}/\mathcal{I}$, where $\mathcal{I} = \{\mathbf{I}_p, -\mathbf{I}_p\} \subset \mathcal{G}$ and $\mathbf{I}_p$ is the identity matrix of dimension $p$. Therefore, $\mathcal{G}'$ is also a group and we can use Thm III.1 to extend the Watson distribution to the mixture representation under $\mathcal{G}'$:

$$f_w(\mathbf{x}; \boldsymbol{\mu}, \kappa) = \frac{1}{M'} \sum_{m=1}^{M'} W_p(\mathbf{x}; \mathbf{P}_m\boldsymbol{\mu}, \kappa), \tag{3.17}$$

where $M' = M/2$. The ML estimates of $\boldsymbol{\mu}$ and $\kappa$ can also be calculated by EM algorithm. The E-step for the Watson mixture distribution is

$$r_{i,m} = \frac{\exp\left\{\kappa((\mathbf{P}_m\boldsymbol{\mu})^T\mathbf{x}_i)^2\right\}}{\sum_{l=1}^{M'} \exp\left\{\kappa((\mathbf{P}_l\boldsymbol{\mu})^T\mathbf{x}_i)^2\right\}}, m \in \left\{1, 2, ..., M'\right\}. \tag{3.18}$$

For the M-step, we take a similar approach as (*Mardia and Jupp*, 1999) as follows:

$$\begin{aligned} &E_{S|\mathbf{X},\boldsymbol{\omega}}[\log L(\boldsymbol{\omega}; \{\mathbf{x}_i, s_i\})] \\ =&n\left(\kappa\boldsymbol{\mu}^T\tilde{T}\boldsymbol{\mu} - \log\left(M'\mathbb{M}\left(\frac{1}{2}, \frac{p}{2}, \kappa\right)\right)\right), \end{aligned} \tag{3.19}$$

where $\tilde{T} = \frac{1}{n}\sum_{i=1}^{n}\sum_{m=1}^{M'} r_{i,m}(\mathbf{P}_m^T\mathbf{x}_i\mathbf{x}_i^T\mathbf{P}_m)$ is the scatter matrix of $\mathbf{x}_1, ..., \mathbf{x}_n$. Let $\tilde{t}_1, ..., \tilde{t}_p$ be the eigenvalues of $\tilde{T}$ with

$$\tilde{t}_1 \geq ... \geq \tilde{t}_p, \tag{3.20}$$

and let $\pm\mathbf{t}_1, ..., \pm\mathbf{t}_p$ be the corresponding unit eigenvectors. Since we want to find $\boldsymbol{\mu}$ which maximizes (3.19) such that $\boldsymbol{\mu}^T\boldsymbol{\mu} = 1$, the estimator of $\boldsymbol{\mu}$ for fixed $\kappa$ has the

following form:

$$\begin{aligned} \hat{\boldsymbol{\mu}} &= \mathbf{t}_1, \hat{\kappa} > 0, \\ \hat{\boldsymbol{\mu}} &= \mathbf{t}_p, \hat{\kappa} < 0. \end{aligned} \tag{3.21}$$

Similarly by fixing $\boldsymbol{\mu}$ and setting to zero the derivative of (3.19) with respect to $\kappa$, we have:

$$\begin{aligned} Y_p(\kappa) &= \frac{\mathbb{M}'(\frac{1}{2}, \frac{p}{2}, \kappa)}{\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)} = \frac{\sum_{i=1}^{n} \sum_{m=1}^{M'} r_{i,m} (\boldsymbol{\mu}^T \mathbf{P}_m^T \mathbf{x}_i)^2}{n} \\ \Rightarrow \hat{\kappa} &= Y_p^{-1} \left( \frac{\sum_{i=1}^{n} \sum_{m=1}^{M'} r_{i,m} (\boldsymbol{\mu}^T \mathbf{P}_m^T \mathbf{x}_i)^2}{n} \right), \end{aligned} \tag{3.22}$$

The final estimates of $\boldsymbol{\mu}$ and $\kappa$ are obtained by checking both cases ($\hat{\kappa} > 0$,$\hat{\kappa} < 0$) and choosing the one which is consistent for (3.21)(3.22). Detail derivation of the E, M-step is given in Appendix C.

## 3.4 Clustering with a Spherical Symmetry Group

In this section we extend the parameter estimation problem to the situation where there are multiple group-invariant distributions with different parameters that govern the samples. This problem arises, for example, when estimating the mean orientation over a region containing more than one grain (perhaps undetected). Usually, this problem can be solved by first applying some standard clustering methods, e.g. K-means(*Hartigan and Wong*, 1979), and then estimating the parameters for each cluster. However, clustering methods based on the distance relation between the samples are complicated by the presence of spherical symmetry because it is necessary to distinguish modes that are due only to symmetry. Therefore, we propose a model-based clustering algorithm which accommodates symmetry to handle this problem.

Consider the situation where the samples $\{\mathbf{x}_i\}_{i=1}^n$ follow a mixture of $\mathcal{G}$-invariant density functions. For VMF distribution, the mixture density has the following form:

$$g_v(\mathbf{x}; \{\boldsymbol{\mu}_c, \kappa_c, \alpha_c\}) = \sum_{c=1}^{C} \alpha_c \left( \sum_{m=1}^{M} \frac{1}{M} \phi(\mathbf{x}; \mathbf{P}_m \boldsymbol{\mu}_c, \kappa_c) \right), \tag{3.23}$$

where $C$ is the number of clusters assumed to be fixed a priori and $\boldsymbol{\mu}_c, \kappa_c$ are the parameters for the $c$-th cluster and $\alpha_c$ are the mixing coefficients where $\sum_{c=1}^{C} \alpha_c = 1$. The parameters of (3.23) can be estimated by the EM algorithm:

E-step:

$$r_{i,c,m} = \frac{\alpha_c \phi(\mathbf{x}_i; \mathbf{P}_m \boldsymbol{\mu}_c, \kappa_c)}{\sum_{h=1}^{C} \alpha_h \sum_{l=1}^{M} \phi(\mathbf{x}_i; \mathbf{P}_l \mu_h, \kappa_h)} \tag{3.24}$$

M-step:

$$\alpha_c = \sum_{i=1}^{n} \sum_{m=1}^{M} r_{i,c,m}, \tag{3.25}$$

$$\hat{\boldsymbol{\mu}}_c = \frac{\boldsymbol{\gamma}_c}{\|\boldsymbol{\gamma}_c\|}, \hat{\kappa}_c = A_p^{-1} \left( \frac{\|\boldsymbol{\gamma}_c\|}{n\alpha_c} \right), \tag{3.26}$$

$$\boldsymbol{\gamma}_c = \sum_{i=1}^{n} \sum_{m=1}^{M} r_{i,c,m} \mathbf{P}_m^T \mathbf{x}_i, \tag{3.27}$$

where $r_{i,c,m}$ is the probability of sample $\mathbf{x}_i$ belonging to the $c$-th cluster and the $m$-th symmetric component.

For Watson distribution, the mixture of $\mathcal{G}$-invariant Watson density is

$$g_w(\mathbf{x}; \{\boldsymbol{\mu}_c, \kappa_c\}) = \sum_{c=1}^{C} \alpha_c \left( \sum_{m=1}^{M'} \frac{1}{M'} W_p(\mathbf{x}; \mathbf{P}_m \boldsymbol{\mu}_c, \kappa_c) \right) \tag{3.28}$$

The E-step is similar to (3.24) with $\phi$ replaced by $W_p$ function. The M-step can

be computed with a similar approach with the following modifications:

$$\tilde{T}_c = \frac{1}{n\alpha_c} \sum_{i=1}^{n} \sum_{m=1}^{M'} r_{i,c,m} (\mathbf{P}_m^T \mathbf{x}_i \mathbf{x}_i^T \mathbf{P}_m), \tag{3.29}$$

$$\hat{\kappa}_c = Y_p^{-1} \left( \frac{\sum_{i=1}^{n} \sum_{m=1}^{M'} r_{i,c,m} (\boldsymbol{\mu}_c^T \mathbf{P}_m^T \mathbf{x}_i)^2}{n\alpha_c} \right), \tag{3.30}$$

where $\alpha_c = \sum_{i=1}^{n} \sum_{m=1}^{M} r_{i,c,m}$.

### 3.4.1 Multi-modality Tests on $\mathcal{G}$-invariant Spherical Distributions

Given sample set $\{\mathbf{x}_i\}_{i=1}^n$ on $S^{p-1}$, we want to determine whether the $n$ samples are drawn from one single distribution or a mixture of $C$ distributions. For polycrystalline materials, the solution to this problem can be used to discover undetected grains within a region. We propose to use a multi-modality hypothesis test based on the $\mathcal{G}$-invariant distributions to solve this problem. The two hypotheses are $H_0$: The samples are from a single $\mathcal{G}$-invariant distribution $f(\mathbf{x}; \{\boldsymbol{\mu}, \kappa\})$; and $H_1$: The samples are from a mixture of $C$ distributions $g(\mathbf{x}; \{\boldsymbol{\mu}_c, \kappa_c\}_{c=1}^C)$. The Generalized Likelihood Ratio Test (GLRT) (*Hero*, 2000) has the following form:

$$\Lambda_{GLR} = \frac{\max_{\{\boldsymbol{\mu}_c, \kappa_c\}_{c=1}^C \in \Theta_1} g(\{\mathbf{x}_i\}_{i=1}^n; \{\boldsymbol{\mu}_c, \kappa_c, \alpha_c\}_{c=1}^C)}{\max_{\{\boldsymbol{\mu}, \kappa\} \in \Theta_0} f(\{\mathbf{x}_i\}_{i=1}^n; \{\boldsymbol{\mu}, \kappa\})} \gtrless_{H_0}^{H_1} \eta \tag{3.31}$$

where $\Theta_0, \Theta_1$ are the parameter spaces for the two hypotheses. The $f$ and $g$ functions for VMF and Watson distributions are defined in (3.9), (3.17) and (3.23), (3.28) respectively and the test statistic, $\Lambda_{GLR}$, can be calculated by the proposed EM algorithm. According to Wilks's theorem (*Wilks*, 1938) as $n$ approaches $\infty$, the test statistic $2 \log \Lambda_{GLR}$ will be asymptotically $\chi^2$-distributed with degrees of freedom equal to $(p+1)(c-1)$, which is the difference in dimensionality of $\Theta_0$ and $\Theta_1$. Therefore, the threshold $\eta$ in (3.31) can be determined by a given significance level $\alpha$.

## 3.5 Application to Crystallographic Orientation

Crystal orientation and the grain distribution in polycrystalline materials determine the mechanical properties of the material, such as, stiffness, elasticity, and deformability. Locating the grain regions and estimating their orientation and dispersion play an essential role in detecting anomalies and vulnerable parts of materials.

Electron backscatter diffraction (EBSD) microscopy acquires crystal orientation at multiple locations within a grain by capturing the Kikuchi diffraction patterns of the backscatter electrons (*Saruwatari et al.*, 2007). A Kikuchi pattern can be translated to crystal orientation through Hough Transformation analysis (*Lassen*, 1994) or Dictionary-Based indexing (*Park et al.*, 2013). The process of assigning mean orientation values to each grain is known as indexing. Crystal forms possess point symmetries, e.g. triclinic, tetragonal, or cubic, leading to a probability density of measured orientations that is invariant over an associated spherical symmetry group $\mathcal{G}$. Therefore, when the type of material has known symmetries, e.g., cubic-type symmetry for nickel or gold, the $\mathcal{G}$-invariant VMF and Watson models introduced in Section 3.3 can be applied to estimate the mean orientation $\boldsymbol{\mu}_g$ and the concentration $\kappa_g$ associated with each grain. Furthermore, the clustering method along with the multi-sample hypothesis test in Section 3.4 can be used to detect the underlying grains within a region.

### 3.5.1 Simulation of Crystallographic Orientation

To simulate the crystallographic orientations, we first draw random samples from VMF and Watson distributions with $p = 4$. The random variable $\mathbf{x}$ in a spherical distribution can be decomposed as:

$$\mathbf{x} = t\boldsymbol{\mu} + \sqrt{1 - t^2} S_{\boldsymbol{\mu}}(\mathbf{x}), \tag{3.32}$$

where $t = \boldsymbol{\mu}^T\mathbf{x}$ and $S_{\boldsymbol{\mu}}(\mathbf{x}) = (I_p - \boldsymbol{\mu}\boldsymbol{\mu}^T)\mathbf{x}/\|(I_p - \boldsymbol{\mu}\boldsymbol{\mu}^T)\mathbf{x}\|$. Let $f(\mathbf{x};\boldsymbol{\mu})$ be the p.d.f. of the distribution where $\boldsymbol{\mu}$ is the mean direction. According to the normal-tangent decomposition property, for any rotationally symmetric distribution, $S_{\boldsymbol{\mu}}(\mathbf{x})$ is uniformly distributed on $S^{p-2}_{\boldsymbol{\mu}^\perp}$, the $(p-2)$-dimensional sphere normal to $\boldsymbol{\mu}$, and the density of $t = \mathbf{x}^T\boldsymbol{\mu}$ is given by:

$$t \mapsto cf(t)(1-t^2)^{(p-3)/2}. \tag{3.33}$$

For VMF distribution, substituting (3.32) into (3.6) and combining with (3.33), we have the density of the tangent component $t$ as:

$$\begin{aligned} f_v(t) &= C_v \exp\{\kappa t\}(1-t^2)^{(p-3)/2} \\ C_v &= \left(\frac{\kappa}{2}\right)^{(p/2-1)} \left(I_{p/2-1}(\kappa)\Gamma\left(\frac{p-1}{2}\right)\Gamma\left(\frac{1}{2}\right)\right)^{-1}. \end{aligned} \tag{3.34}$$

Similarly, the density of the tangent component of Watson distribution is:

$$\begin{aligned} f_w(t) &= C_w \exp\{\kappa t^2\}(1-t^2)^{(p-3)/2} \\ C_w &= \frac{\Gamma(\frac{p}{2})}{\Gamma(\frac{p-1}{2})\Gamma(\frac{1}{2})} \frac{1}{\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)}. \end{aligned} \tag{3.35}$$

Random samples from the density functions (3.34) and (3.35) can be easily generated by rejection sampling.

The generated quaternions from VMF and Watson distributions are then mapped into the Fundamental Zone (FZ) with the symmetric group actions to simulate the wrap-around problem we observe in real data, i.e. observations are restricted to a single FZ. For cubic symmetry, the FZ in quaternion space is defined in the following set of equations where the derivation is given in Appendix A:

$$
\begin{cases}
|q_2/q_1| \leq \sqrt{2}-1 \\
|q_3/q_1| \leq \sqrt{2}-1 \\
|q_4/q_1| \leq \sqrt{2}-1 \\
|q_2/q_1 + q_3/q_1 + q_4/q_1| \leq 1 \\
|q_2/q_1 - q_3/q_1 + q_4/q_1| \leq 1 \\
|q_2/q_1 + q_3/q_1 - q_4/q_1| \leq 1 \\
|q_2/q_1 - q_3/q_1 - q_4/q_1| \leq 1 \\
|q_2/q_1 - q_3/q_1| \leq \sqrt{2} \\
|q_2/q_1 + q_3/q_1| \leq \sqrt{2} \\
|q_2/q_1 - q_4/q_1| \leq \sqrt{2} \\
|q_2/q_1 + q_4/q_1| \leq \sqrt{2} \\
|q_3/q_1 - q_4/q_1| \leq \sqrt{2} \\
|q_3/q_1 + q_4/q_1| \leq \sqrt{2}
\end{cases}
\tag{3.36}
$$

where $q_i$ is the $i$-th component of quaternion $\mathbf{q}$.

## 3.6 Experimental Results

### 3.6.1 $\mathcal{G}$-invariant EM-ML Parameter Estimation on Simulated Data

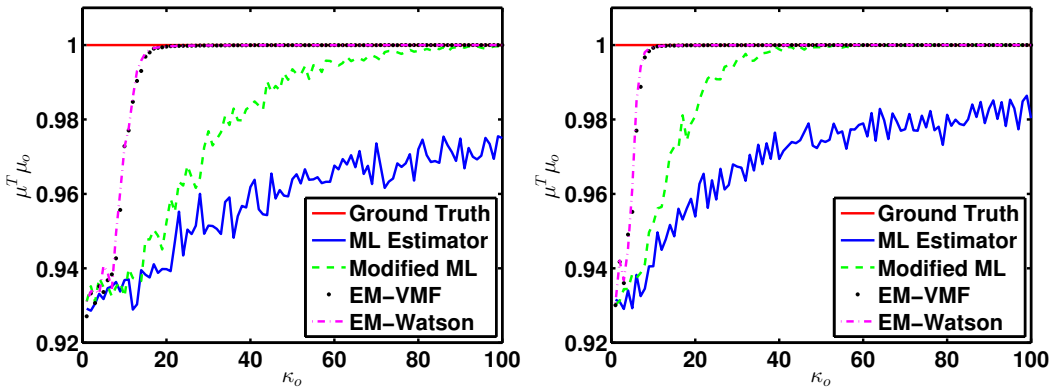
Sets of $n$ i.i.d. samples were simulated from the VMF or Watson distributions using the method described in Sec.3.5.1 with given $\boldsymbol{\mu} = \boldsymbol{\mu}_o, \kappa = \kappa_o$ for the $m\overline{3}m$ point symmetry group associated with the symmetries of cubic crystal lattice planes. The number of samples for each simulation was set to $n = 1000$ and $\kappa_o$ was swept from 1 to 100 while, for each simulation run, $\boldsymbol{\mu}_o$ was selected uniformly at random. The experiment was repeated 100 times and the average values of $\hat{\kappa}$ and the inner

product $\hat{\boldsymbol{\mu}}^T \boldsymbol{\mu}_o$ are shown in Fig. 3.1 and 3.2. In the figures we compare performance for the following methods: (1) the naive ML estimator for the standard VMF or Watson model that does not account for the point group structure (labeled "ML Estimator"). (2) Mapping the $n$ samples toward a reference direction $\mathbf{x}_r$ (randomly selected from $\{\mathbf{x}_i\}_{i=1}^n$) according to (3.37) followed by performing ML for the standard VMF or Watson distribution (labeled "Modified ML"). (3) Applying our proposed EM algorithm directly to the $n$ samples using the mixture of VMF distribution (3.11)-(3.13) (labeled "EM-VMF") (4) Applying our proposed EM algorithm to the mixture of Watson distribution (3.18)-(3.22) (labeled "EM-Watson").

$$\begin{aligned} \mathbf{x} &\mapsto \mathbf{P}_m \mathbf{x} \\ \mathbf{P}_m &= \arg\min_{\mathbf{P}\in\mathcal{G}} \arccos\left(\mathbf{x}_r^T \mathbf{P} \mathbf{x}\right) \end{aligned} \tag{3.37}$$

To prevent the EM algorithm from converging to local minimum, for each set of samples we performed the EM algorithm for $T$ trials with randomly initialized parameters. The final estimates of each set are obtained by choosing the estimates which give the largest likelihood out of the $T$ trials. In the experiment, we observed that for each trial the EM algorithm usually converges within 30 iterations and the likelihood function is monotonically increasing each iteration as expected. We found out that the estimation accuracy is not sensitive to initialization especially for samples which are generated from larger $\kappa$, and $T = 10$ is more than enough for obtaining stable results as we shown in the following discussion.

Figure 3.1 shows the inner product values $\boldsymbol{\mu}_o^T \hat{\boldsymbol{\mu}}$. The proposed EM-VMF and EM-Watson estimators have similar performance in that they achieve perfect recovery of the mean orientation ($\boldsymbol{\mu}_o^T \hat{\boldsymbol{\mu}} = 1$) much faster than the other methods as the concentration parameter $\kappa_o$ increases (lower dispersion of the samples about the mean) no matter whether the data is generated from VMF (Fig.3.1a) or Watson distribution (Fig.3.1b), indicating the robustness of the proposed approaches under model mis-

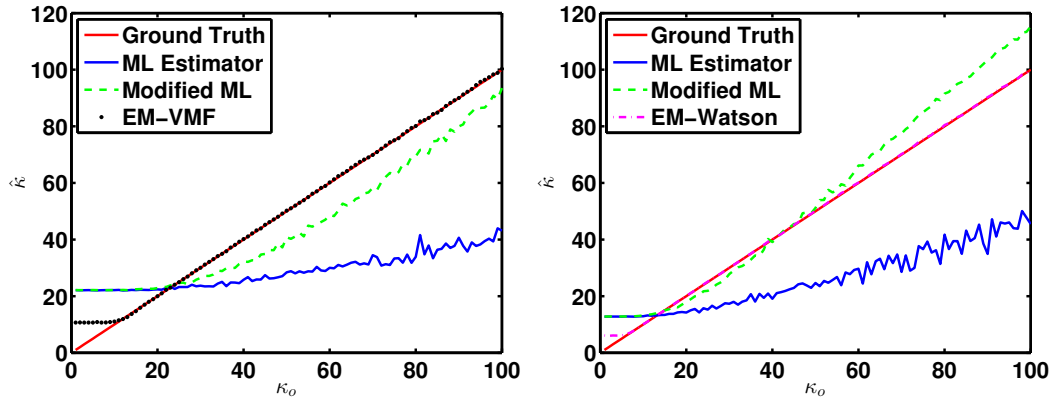



(a) VMF Simulated Data (b) Watson Simulated Data

Figure 3.1: Mean orientation estimator comparisons for $\mathcal{G}$-invariant densities when $\mathcal{G}$ is the $m\bar{3}m$ point symmetry group. Shown is the average inner product $\boldsymbol{\mu}_o^T\hat{\boldsymbol{\mu}}$ of four estimators $\hat{\boldsymbol{\mu}}$ when $\boldsymbol{\mu}_o$ is the true mean orientation as a function of the true concentration parameter $\kappa_o$ for the data simulated from VMF (Fig.3.1a) and from Watson (Fig.3.1b) distribution. Each estimator was implemented with $n = 1000$ i.i.d. samples from the $\mathcal{G}$-invariant density and the inner product shown is averaged over 100 trials. The naive estimator ("ML Estimator" in blue line) does not attain perfect estimation (inner product = 1) for any $\kappa_o$ since it does not account for the spherical symmetry group structure. A modified ML estimator ("modified ML" in green dashed line) achieves perfect estimation as $\kappa_o$ becomes large. The proposed EM-ML methods ("EM-VMF", "EM-Watson") achieve perfect estimation much faster than the other methods even under model mismatch (EM-VMF for Watson simulated data and vice versa).

match. Notice that when $\kappa_o$ is small ($\kappa_o < 20$ for VMF data and $\kappa_o < 10$ for Watson data), none of the methods can accurately estimate the mean orientation. The reason is that when $\kappa_o$ is small the samples become nearly uniformly distributed over the sphere. The threshold $\kappa_o$ value at which performance starts to degrade depends on the choice of point symmetry group and the distribution used to simulate the data. In Fig. 3.2 it is seen that the biases of the proposed EM-VMF (*Chen et al.*, 2015c) and EM-Watson $\kappa$ estimators are significantly lower than that of the other methods compared. While the modified ML performs better than the naive ML estimator, its bias is significantly worse than the proposed EM-VMF and EM-Watson approaches.

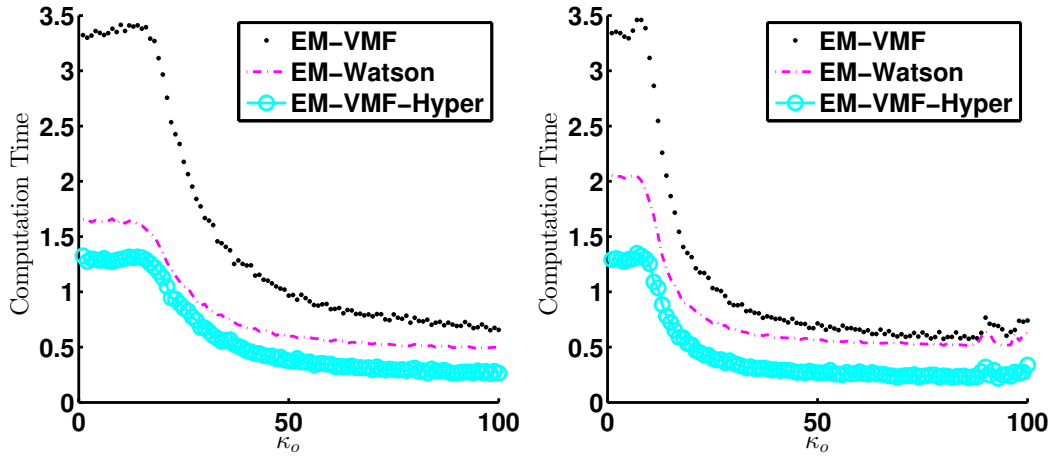
Figure 3.3 shows the computation time of the estimation algorithms presented in




(a) VMF Simulated Data (b) Watson Simulated Data

Figure 3.2: Concentration parameter estimator bias as a function of the true concentration $\kappa_o$ for data simulated from VMF (Fig.3.2a)(*Chen et al.*, 2015c) and from Watson (Fig.3.2b) distributions. The bias of the naive ML (blue solid line) is large over the full range of $\kappa_o$. The modified ML (green dashed line) estimates $\kappa$ more accurately when $\kappa_o$ is small. Our proposed EM-VMF and EM-Watson estimators (black dotted line and magenta dashed line) have lower bias than the other estimators.

Fig.3.1 and Fig.3.2. The computation time for all methods decreases as $\kappa_o$ becomes larger. When $\kappa_o$ is small ($\kappa_o < 20$ for VMF data and $\kappa_o < 10$ for Watson data), because the samples are almost uniformly distributed around the sphere, it is difficult for the EM algorithms to converge to the optimal solution and they therefore require maximum number of iterations to stop, forming the plateaus in Fig.3.3. Notice that EM-Watson requires less time than EM-VMF even though it has more complicated E and M-steps. The reason is that EM-Watson uses only half of the symmetry operators, which corresponds to the size of the quotient group $\mathcal{G}/\mathcal{I}$ as described in Section 3.3.2. By applying the hyperbolic sinusoidal simplification in Section 3.3.1 (labeled "EM-VMF-Hyper"), we can further reduce the computation time by more than a factor of 2 compared to the original EM-VMF.




(a) VMF Simulated Data (b) Watson Simulated Data

Figure 3.3: Computation time for calculating the result in Fig.3.1 and Fig.3.2. EM-Watson (magenta dashed line) has less computation time than EM-VMF (black dotted line) because it uses only half of the symmetry operators. EM-VMF-Hyper (cyan circle line) which uses the hyperbolic sinusoidal simplification of EM-VMF reduces the computation time by more than a factor of 2.

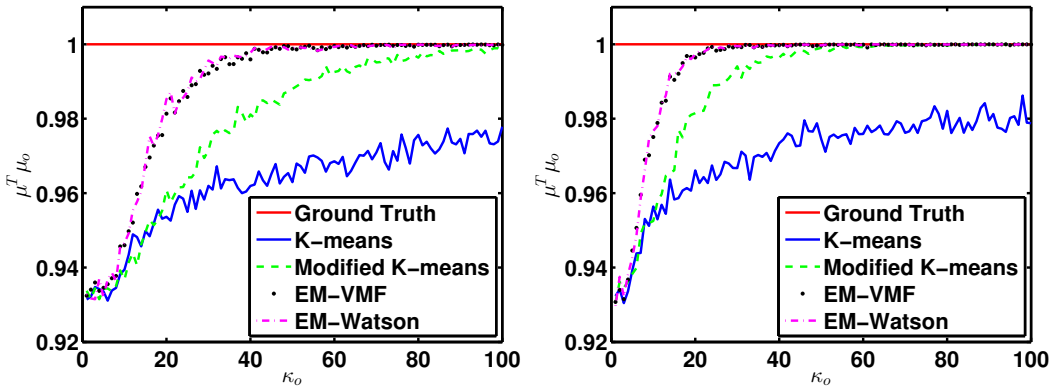
### 3.6.2 $\mathcal{G}$-invariant Clustering on Simulated Data

In this section, we demonstrate the performance of our proposed EM approaches for clustering. Sets of $n$ i.i.d. samples were simulated from the VMF or Watson distributions with $\kappa = \kappa_o$ and one of two mean directions $(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2)$ to generate two clusters of samples. The spherical symmetry group is $m\overline{3}m$ as before. The number of samples for each set was set to $n = 1000$ and $\kappa_o$ was swept from 1 to 100 while, for each set, $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ was selected uniformly at random. The experiment was repeated 100 times and the average values of the inner product $(\hat{\boldsymbol{\mu}}_1^T \boldsymbol{\mu}_1 + \hat{\boldsymbol{\mu}}_2^T \boldsymbol{\mu}_2)/2$ are shown in Fig. 3.4. In the figure we compare performances of the following methods: (1) Cluster the samples by standard K-means algorithm with the distance defined by the arc-cosine of the inner product and then use the naive ML within each cluster to estimate the mean directions (labeled "K-means"). (2) Cluster the samples by K-means with the distance defined as (3.5) and then use the aforementioned modified ML estimator (labeled "Modified K-means"). (3) Apply our proposed multi-cluster

EM-VMF algorithm to the $n$ samples directly (3.24)-(3.27) (labeled "EM-VMF") (4) Apply our multi-cluster EM-Watson algorithm to the $n$ samples directly (3.29)-(3.30) (labeled "EM-Watson").

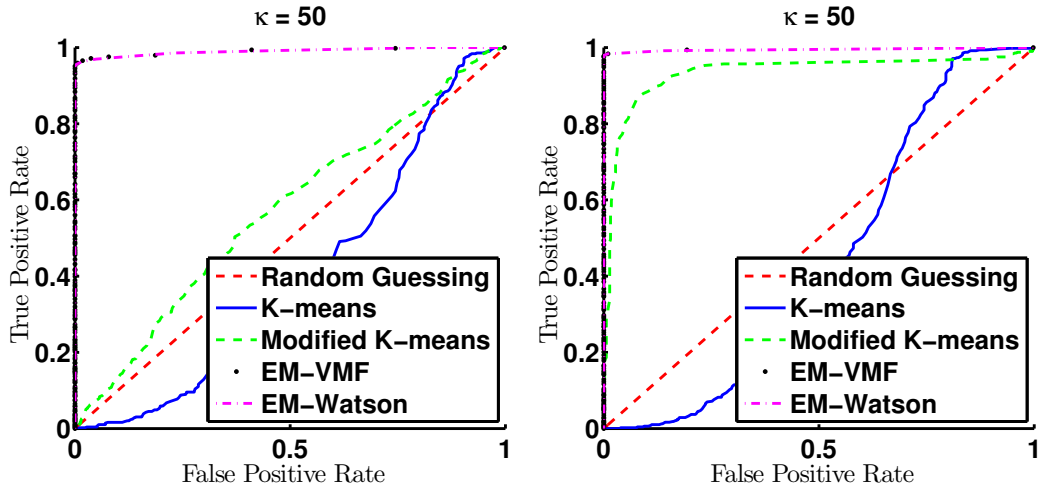
Figure 3.4 shows the average inner product values $(\hat{\boldsymbol{\mu}}_1^T \boldsymbol{\mu}_1 + \hat{\boldsymbol{\mu}}_2^T \boldsymbol{\mu}_2)/2$ from the mean direction estimation. The proposed EM-VMF and EM-Watson are able to correctly cluster the samples and achieve perfect recovery of the two mean orientations much faster than the other K-means approaches. Notice that the region where all the methods fail is larger than the single cluster case since multiple clusters increase the difficulty of parameter estimation. Again, no matter whether the samples are simulated from VMF or Watson distribution, our proposed approaches perform equally well under both cases.

To further test the ability to detect multiple clusters given a set of samples, we generate 1000 sets of samples. Each set has 1000 samples and is assigned randomly to label 0 or 1. If the set is labeled 0, the samples are generated from a single distribution; If the set is labeled 1, then the samples in the set are randomly generated from two distributions with different means. The GLRT is used with the four aforementioned clustering methods to test whether the samples in each set are uni-modal or multi-modal. The Receiver Operating Characteristic (ROC) curves of the test results are shown in Fig. 3.5. The naive K-means with ML estimator which does not consider the symmetry group actions fails to distinguish whether the multiple modes are from actual multiple distributions or due to the wrap-around effect from the fundamental zone mapping. Therefore, this approach tends to over-estimate the goodness of fit of the $H_1$ model for true negative cases and under-estimate it for true positive cases, resulting in a result that is even worse than random guessing. The modified K-means performs better than K-means but worse than our proposed EM-VMF and EM-Watson algorithms.




(a) VMF Simulated Data (b) Watson Simulated Data

Figure 3.4: Mean orientation estimator comparisons for samples generated from two different means. Shown is the average inner product $(\hat{\boldsymbol{\mu}}_1^T \boldsymbol{\mu}_1 + \hat{\boldsymbol{\mu}}_2^T \boldsymbol{\mu}_2)/2$ of four methods when $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ are the true mean orientations as a function of the true concentration parameter $\kappa_o$ for the data simulated from VMF (Fig.3.4a) and from Watson (Fig.3.4b) distributions. Each estimator was implemented with $n = 1000$ i.i.d. samples from the $\mathcal{G}$-invariant densities with two different mean directions and the inner product shown is averaged over 100 trials. The K-means with naive estimator ("K-means" in blue line) does not attain perfect estimation (inner product = 1) for any $\kappa_o$. A modified K-means with ML estimator ("modified K-means" in green dashed line) achieve perfect estimation as $\kappa_o$ becomes large. The proposed EM-VMF and EM-Watson methods ("EM-VMF", "EM-Watson") achieves perfect estimation much faster than the other methods no matter whether the data are generated from VMF or Watson distribution, indicating the robustness to model mismatch.




(a) VMF Simulated Data (b) Watson Simulated Data

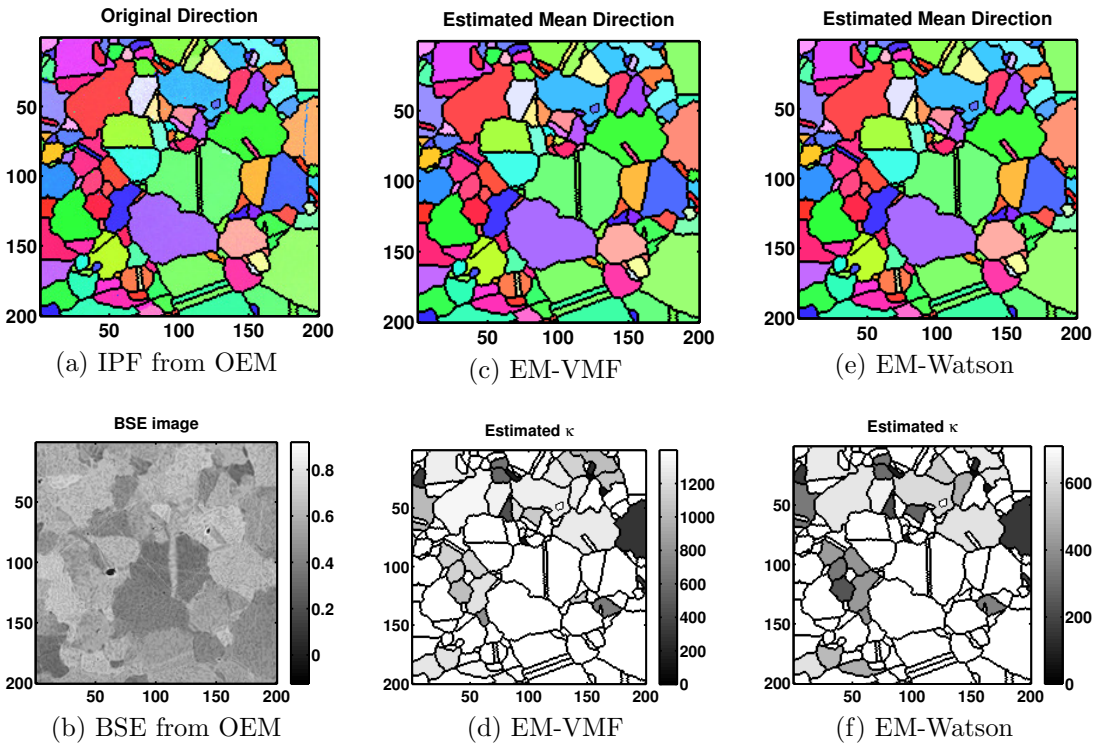
Figure 3.5: ROC curve for detecting bi-modally distributed samples. The samples are uni-modal or bi-modal distributed from VMF (Fig.3.5a) or Watson (3.5b) distributions with $\kappa_o = 50$. The naive K-means with ML estimator cannot cluster the samples well and estimate the mean directions accurately, resulting in poor detection which is even worse than random guessing. The modified K-means (green dashed line) performs better than K-means but is still unsatisfactory. Our proposed EM-VMF (black dots) and EM-Watson (magenta dashed line) methods have very good performance in this detection task.

### 3.6.3 EM-ML orientation estimation for IN100 Nickel Sample

We next illustrate the proposed EM-VMF and EM-Watson orientation estimators on a real IN100 sample acquired from US Air Force Research Laboratory (AFRL) (*Park et al.*, 2013). The IN100 sample is a polycrystalline Ni superalloy which has cubic symmetry in the $m\overline{3}m$ point symmetry group. EBSD orientation measurements were acquired on a $512 \times 384$ pixel grid, corresponding to spatial resolution of 297.7 nm. The Kikuchi diffraction patterns were recorded on a $80 \times 60$ photosensitive detector for each of the pixels.

Figure 3.6 (a) shows a $200 \times 200$ sub-region of the full EBSD sample where the orientations are shown in the inverse pole figure (IPF) coloring obtained from the OEM EBSD imaging software and (d) is the back-scattered electron (BSE) image. Note that the OEM-estimated orientations in some grain regions of the IPF image are very inhomogeneous, which is likely due to a fundamental zone wrap-around problem. Figure 3.6 (b)(c) shows the estimates of the mean orientations of the grains using the proposed EM-VMF and EM-Watson algorithms. Figures 3.6 (d)(e) show the estimated concentration parameter $\kappa$ for the grains using the proposed EM-VMF and EM-Watson algorithms.

Figure 3.7 shows the same sub-region of the full EBSD sample with an alternative approach. We apply a combination of the proposed EM estimators (EM-VMF or EM-Watson) and the GLRT (3.31) with $C = 2$ and significance level $\alpha = 0.05$ to detect multi-modal distributions within each OEM-segmented region. Figure 3.6 (a)(b) show the estimates of the mean orientations of the regions/sub-regions, where the sub-regions surrounded by white boundaries indicate those that have been detected as deviating from the distribution of the majority of samples from the same region. The multi-modally distributed regions may be due to undetected grains, inaccurate segmentation, or noisy orientation observations. To distinguish the latter situations from the first in which the region really consists of two grains, the misalignment/noise

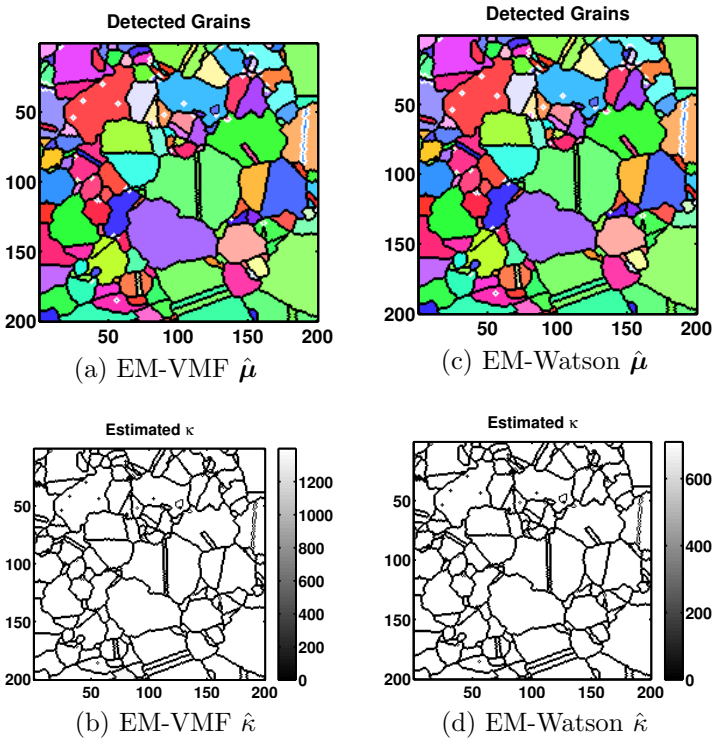



(a) IPF from OEM (c) EM-VMF (e) EM-Watson

(b) BSE from OEM (d) EM-VMF (f) EM-Watson

Figure 3.6: A $200 \times 200$ sub-region of the IN100 sample. (a) is the IPF image for the Euler angles extracted from EBSD by OEM imaging software. IPF coloring in some grains is not homogeneous, likely due to the ambiguity problem. (d) is the BSE image of the sample. (b) is the IPF image for the mean orientation of the grains estimated by EM-VMF algorithm and (c) is the concentration parameters $\kappa$ estimated by EM-VMF for the $\mathcal{G}$-invariant VMF density. (e) and (f) are the mean orientation and concentration parameters $\kappa$ estimated by EM-Watson algorithm. Our proposed EM-ML estimators have high concentration $\kappa$ even for those grains with inhomogeneous Euler angles.

test introduced in Chapter V can be used. Figures 3.6 (c)(d) show the estimated concentration parameter $\kappa$ for the regions/sub-regions. Note that the estimated $\kappa$ are large for most of the regions/sub-regions because those regions which have multi-modally distributed samples are detected and their concentration parameters are estimated separately for each sub-region.




(a) EM-VMF $\hat{\boldsymbol{\mu}}$ (c) EM-Watson $\hat{\boldsymbol{\mu}}$

(b) EM-VMF $\hat{\kappa}$ (d) EM-Watson $\hat{\kappa}$

Figure 3.7: A $200 \times 200$ sub-region of the IN100 sample. (a)(b) show the estimates of the mean orientations of the regions/sub-regions using a combination of the proposed EM estimators, EM-VMF and EM-Watson respectively, and the GLRT (3.31) to detect multi-modal distributions within each OEM-segmented region. The sub-regions surrounded by white boundaries indicate those that have been detected as deviating from the distribution of the majority of samples from the same region. (c)(d) show the estimated concentration parameter $\kappa$ for the regions/sub-regions. Note that the estimated $\kappa$ are large for most of the regions/sub-regions because those regions which have multi-modally distributed samples are detected and their concentration parameters are estimated separately for each sub-region.

## 3.7 Conclusion

We have obtained a general finite mixture representation for densities on domains whose topologies have group invariances. This representation was used to construct mixtures of the Von Mises-Fisher and Watson distributions that possess spherically symmetric group invariances. An efficient EM algorithm was derived for estimation of parameters of this mixture model. We further extended the model to consider the multi-cluster case where the samples are from several group-invariant distributions

with different parameters. The GLRT is proposed along with the mixture model to detect the presence of multiple modes. The mixture of VMF and Watson models were applied to the problem of estimation of mean grain orientation parameters in polycrystalline materials whose orientations lie in the $m\overline{3}m$ point symmetry group. Application of the finite mixture representation to other types of groups would be worthwhile future work.

# CHAPTER IV

# A Dictionary Approach to EBSD Indexing

## 4.1 Introduction

Electron backscatter diffraction, EBSD, is used to perform quantitative microstructure analysis of polycrystalline materials on a millimeter to nanometer scale (*Schwartz et al.*, 2009). Most current EBSD segmentation (i.e., delineation of individual grains by determination of the grain boundary locations) and indexing (i.e., orientation determination) methods extract orientations and widths of Kikuchi bands in a measured pattern by using a modified Hough transform, implemented using image processing tools such as butterfly convolution, Gaussian filtering, binning, peak detection, and image quality maps to gauge indexing and segmentation accuracy (*Tao and Eades*, 2005; *Wright and Nowell*, 2006). By comparing the measured diffraction line parameters to a pre-computed database, indexing yields the crystal orientation, commonly described by three Euler angles with respect to a reference frame, for the volume illuminated by the beam. By repeating the process on a grid of scanning locations on the sample, an orientation map or image is produced. The image is then segmented into grains by thresholding normed differences between the Euler angles (misorientations). The accuracy of the Hough approach to EBSD indexing depends to a large extent on the visibility of the Kikuchi bands, which is often represented in terms of

*This chapter is based on (*Chen et al.*, 2015a)

an image quality parameter.

In this chapter we introduce an alternative indexing approach that uses a physics-based forward model for the full diffraction patterns and does not require application of the Hough transform or other image processing tools. The proposed approach exploits the known physics of electron scattering phenomena that underlies the image formation process. With this additional information, grain boundaries and anomalous points can be detected as explicit classes at the same time as grains are segmented. Anomaly detection (i.e., the automated detection of abnormal or unexpected diffraction patterns) is an important capability, since anomalies may correspond to defects or contaminants that affect the material properties. In addition, unlike methods based on the Hough transform, the proposed pattern dictionary approach differentiates grain interiors from grain boundaries without requiring additional processing of the measured patterns.

Automated indexing of electron diffraction patterns by means of pattern matching techniques is not new, and was proposed in 1991 by Wright and coworkers (*Wright et al.*, 1991) for backscattered Kikuchi diffraction patterns, now commonly known as electron backscatter diffraction patterns or EBSPs. Their approach involved automated comparisons between a series of experimental patterns and idealized patterns, created from a set of orientations that uniformly covers the asymmetric region (or fundamental zone) of Euler space. While the technique showed promising results, computer limitations prevented the approach from gaining widespread acceptance. In the context of precession electron diffraction in the transmission electron microscope, Rauch and coworkers (*Rauch and Dupuy*, 2005; *Rauch et al.*, 2008) proposed a template-based pattern matching approach for the automatic orientation determination of quasi-kinematical diffraction patterns. The templates are computed from kinematical structure factors, and the proper asymmetric part of Euler space is sampled uniformly to generate a template collection, which is then compared quantita-

tively against the experimental patterns. In this template matching process, only the top match is considered in the determination of the crystal orientation. As we will explain in detail in the rest of this chapter, our dictionary appproach uses, in addition to a physics-based model for the generation of the dictionary, the statistical model introduced in Chapter III for the orientation distribution of all the highly ranked pattern matches to provide both a stable and robust estimate of orientation, as well as a quantitative statistical uncertainty; the method proposed in (*Rauch and Dupuy*, 2005) does not perform such a statistical analysis. The current commercially available EBSD indexing suites also lack a statistical determination of the uncertainties in the orientation determination.

The proposed indexing framework relies on two components: offline dictionary generation and online dictionary matching to the sample. Offline dictionary generation is accomplished as follows. First, a dictionary of raw diffraction patterns is generated using the forward model of (*Callahan and De Graef*, 2013). This dictionary is tuned to the parameters of the microscope and the crystal symmetry group(s) of the sample. Second, the singular value decomposition (SVD) of the dictionary is computed and a second dictionary, called the dictionary of background compensated patterns, is generated by projection of the raw dictionary onto the space orthogonal to the first principal component; this is essentially a background subtraction process.

The first step in the online dictionary matching algorithm is to compute normalized inner products between the uncompensated (raw) sample patterns and the uncompensated dictionary. This is repeated on the compensated sample patterns and the compensated dictionary patterns. The basis for the online dictionary matching algorithm is the construction of a pair of bipartite graphs, uncompensated and compensated, respectively, using these normalized inner products. A bipartite graph is a graph connecting vertices or nodes in two disjoint sets () (*Diestel*, 2005); in our case, the first set contains the experimental patterns at different locations on the sample,

the second the dictionary patterns for different orientations. For each spatial location on the sample, the top $k$ (normalized) inner products between the sample diffraction pattern and a dictionary determine the $k$-nearest-neighbor ($k$-NN) neighborhood of the pattern in the dictionary. These $k$-NN neighborhoods form the backbone of the proposed online dictionary matching algorithm.

From the bipartite graphs, sample patterns can be classified as grain interiors, grain boundaries, or anomalies, using an unsupervised decision tree (DT) classifier defined on the graphs. Specifically, the classifier uses the shapes of these $k$-NN neighborhoods to discriminate between these types of patterns. A tightly clustered and connected $k$-NN neighborhood indicates a grain interior sample pattern. A $k$-NN neighborhood that forms two or three clusters in the dictionary indicates a grain boundary sample pattern. $k$-NN neighborhoods that are very spread-out and scattered indicate anomalous sample patterns. The unsupervised DT classifier uses the uncompensated dictionary to distinguish unusual (anomalous) background patterns. The compensated dictionary is used to distinguish non-anomalous patterns as interior to a grain or on the boundary of a grain. A pixel is classified as anomalous, grain interior or grain boundary using an unsupervised decision tree (DT) classifier to test homogeneity of the pattern matches over a $3 \times 3$ spatial patch centered at the pixel. The effect of surface roughness, and the resulting shifts in the EBSD background intensity, is a topic of ongoing analysis.

Indexing of crystal orientation is performed using a maximum likelihood (ML) estimation strategy for determining the Euler angles. The ML strategy described in Chapter III fits a group-invariant density model to the observed distribution of the Euler angles of the top dictionary matches; the distribution is a probability distribution on a sphere in a $p$-dimensional space. An iterative expectation-maximization (EM) estimation algorithm is used to perform the fit. The use of the mixture model allows one to account for the symmetry-induced ambiguities of the crystal orientation

and produces an estimate of the Euler angles in the desired fundamental zone. A side benefit is that the algorithm yields an estimate of the spread of the VMFm model that can be used as an a posteriori confidence measure on the orientation estimate.

To the best of our knowledge, the framework proposed here is the first EBSD indexing approach that uses a dictionary generated by a physics-based forward model. Some advantages of our model-based approach are: 1) it incorporates the physics of dynamical electron scattering; 2) it unifies segmentation, indexing and anomaly detection; 3) it incorporates a statistical model that naturally generates both an estimate of orientation and a measure of confidence in the estimate; 4) it involves parallelizable operations relying on simple inner products and nearest neighbor search. At the same time, the large size of the dictionary, together with the high dimension of the diffraction patterns, create computational challenges as discussed in Section 4.5.

The outline of the chapter is as follows. Section 4.2 presents the proposed dictionary model for EBSD pattern classification and indexing. Section 4.3 develops the statistical algorithms for anomaly detection, segmentation, and indexing based on the dictionary model. Section 4.4 describes the dictionary generation process. Section 4.6 presents the experimental methods for generating the EBSD samples used in Section 4.7. Section 4.7 presents the results of applying the proposed dictionary-based classification and indexing to a Ni-base IN100 alloy. Finally, Section 4.8 summarizes this chapter and points to future directions.

## 4.2 Dictionary Model

This section describes the non-linear forward model of (*Callahan and De Graef*, 2013). It then shows how a physics-based dictionary of diffraction patterns can be used to approximately linearize the forward model. For descriptive economy, throughout this chapter we denote by a pixel a particular scan location on the sample surface; with each pixel, there is an associated EBSD pattern acquired at the sample location.

### 4.2.1 Forward model

When the beam is focused on a grain within the sample, the measured backscatter diffraction pattern, $\mathbf{Y}$, on the detector surface can be expressed as a function of the crystal orientation $\boldsymbol{\theta}$, parameterized by an Euler angle triplet $(\alpha, \beta, \gamma)$, the incident electron energy, $E$, and interaction depth, $z_0$, as:

$$\mathbf{Y} = \mathbf{H}(\mathcal{P}(\boldsymbol{\theta}, E, z_0)) + \mathbf{N}, \tag{4.1}$$

where $\mathbf{H}(\mathcal{P}(\boldsymbol{\theta}, E, z_0))$ is a forward model for the backscatter process and $\mathbf{N}$ is detector noise. Here $\mathbf{H}$ is a measurement operator that accounts for the instrument geometry and sensitivity, and $\mathcal{P}$ represents the thickness and energy averaged mean backscatter yield. An accurate forward model for this yield was proposed by (*Callahan and De Graef*, 2013). The model employs Monte Carlo simulations to obtain the energy, spatial, and exit depth distributions of the backscattered electrons. This statistical information is then used to compute a series of dynamical EBSD master patterns as a function of the electron exit energies and directions. The master pattern represents all possible EBSD patterns for a given exit energy, and specification of the grain orientation $\boldsymbol{\theta}$ along with the detector geometry then leads to an actual EBSD pattern for that orientation. These three steps constitute the forward model $\mathcal{P}$. The measurement operator $\mathbf{H}$ includes the scintillator-to-CCD conversion process in the form of a point spread function for the coupling optics, as well as detector quantum efficiency and CCD binning mode. In the remainder of this chapter, we will refer to the depth and energy averaged diffraction pattern generated by the forward model described above as the mean diffraction pattern.

As the electron beam is scanned across the sample surface, the diffraction pattern will change due to changes in the local crystal orientation $\boldsymbol{\theta}$ between homogeneous grains. In grain interiors, the forward model (4.1) can be used to produce estimates of

crystal orientation at each scan location. Elsewhere in the sample, e.g., near boundary regions or near locations of anomalous features, the model (4.1) will no longer be a good fit to the measured patterns. Thus, the goodness of fit of the forward model can be used to classify grains, grain boundaries, and anomalies in the sample. This forms the basis for our proposed use of the forward-model to perform classification and indexing.

### 4.2.2 Sparse dictionary-based forward model

In principle one could formulate classification and indexing as a non-linear inverse problem using the full forward model (4.1). For example, given a noise model for $\mathbf{N}$ one could perform maximum likelihood estimation to solve the indexing problem and likelihood-ratio testing to solve the classification problem. In the special case of a Gaussian noise model both solutions would require solving the non-linear least-squares problem $\min_{\boldsymbol{\theta}} \|\mathbf{Y} - \mathbf{H}(\mathcal{P}(\boldsymbol{\theta}, E, z_0))\|^2$, in which the Euclidean norm squared of the residual fitting errors is to be minimized with respect to the orientation $\boldsymbol{\theta}$. Here, we take a simpler approach to the inverse problem that leads directly to tractable indexing and classification algorithms.

Let $\mathcal{D}$ denote a precomputed dictionary of mean diffraction patterns obtained by densely sampling the function $\mathbf{H}(\mathcal{P}(\boldsymbol{\theta}, E, z_0))$ over the range of orientations $\boldsymbol{\theta}$, keeping $E$ and $z_0$ fixed. Assume that the size of $\mathcal{D}$ is $d$. Then, for sufficiently dense samples, the model (4.1) can be approximated by

$$\mathbf{Y} = \sum_{i=1}^{d} x_i \phi_i + \mathbf{N}, \tag{4.2}$$

where $\phi_i \in \mathcal{D}$ are dictionary elements and $x_i$ are coefficients. When $\mathbf{Y}$ corresponds to the measured pattern at a location within a grain, one might expect that only a few $x_i$'s will be non-zero, i.e., the representation (4.2) is sparse. In particular,

as the dictionary becomes increasingly dense the sparsity of the representation will also increase and, in the limit, $x_i$ will become a delta function $x_i = \Delta(i - j)$, where $j$ is the index of the true orientation at that location. Note that when $\mathbf{N} = 0$, in this limiting case (of a fully dense dictionary) errorless estimation of $\boldsymbol{\theta}$ can be accomplished by finding the index $i$ which yields the largest normalized inner product $\rho(\mathbf{Y}, \phi_i) = \frac{(\mathbf{Y}, \phi_i)}{\|\mathbf{Y}\|\|\phi_i\|}$, where, for two diffraction pattern $\mathbf{A} = ((A_{lm}))$ and $\mathbf{B} = ((B_{lm}))$, $\langle \mathbf{A}, \mathbf{B} \rangle = \sum_{l,m} A_{lm} B_{lm}$, where $l$ and $m$ index the vertical and horizontal locations on the photodetector. The significance of this fact is that we have simplified the solution of a complicated non-linear least squares problem to the solution of a linear least squares problem followed by a table lookup (matching an index of the dictionary to the associated Euler angle).

In the practical case of a finite dictionary there will not be an exact match to the true diffraction pattern and (4.2) is interpreted as a model that interpolates over the patterns in the dictionary, with interpolation coefficients $\{x_i\}_{i=1}^{d}$. For the purposes of indexing and classification we will restrict ourselves to sparse models, where only a few ($k$) of these coefficients are non-zero, i.e., using only a small number of dictionary elements to fit (4.1). This sparse approximation problem is a well studied mathematical problem with many different iterative algorithms available for identifying the few non-zero coefficients. The brute force algorithm that tries to find the best fit over any set of $k$ dictionary elements is intractable except for very small values of $k$. Alternatives include basis pursuit methods such as orthogonal matching pursuit (OMP), stepwise OMP (stOMP), compressive sampling OMP, iterative soft thresholding (IST), and convex optimization relaxation methods such as $l_1$ minimization using active set, interior point, or sub-gradient methods (*Tropp and Wright*, 2010). In the EBSD application, the sparse approximation has to be performed for each and every pattern measured on the sample. Even with a relatively modest dictionary size and low sample resolution, e.g., $d = 100,000$ elements and $n = 100,000$ scan

locations, these methods are computationally heavy.

We have adopted a simpler correlation matching approach with significantly reduced computation requirements. Instead of fitting the model (4.2) through least squares we simply use inner products to find the $k$ top matches between the dictionary and the observations. Specifically, for each measured pattern $\mathbf{Y}$ we compute the normalized inner products (correlation) $\rho(\mathbf{Y}, \phi_1), ..., \rho(\mathbf{Y}, \phi_d)$ between $\mathbf{Y}$ and the dictionary and rank them in order of decreasing magnitude. The top correlation matches are the $k$ dictionary patterns having the highest inner products. These $k$ patterns constitute and pixel's $k$-nearest-neighbor ($k$-NN) neighborhood in the dictionary. The collections of $k$-NN's define a bipartite graph connecting measured patterns to patterns in the dictionary. In general, these connections will be one-to-many, i.e., for each experimental pattern there will be a small number of near matches in the dictionary. The $k$-NN neighborhoods will be used to perform classification and indexing as described in Section 4.3.

In addition to the dictionary $\mathcal{D}$, referred to as the uncompensated dictionary, we will use a derived dictionary of compensated patterns $\mathcal{D}_c$ to cluster the observed patterns into classes that can then be used for anomaly detection, segmentation and indexing. The compensated dictionary $\mathcal{D}_c$ consists of patterns in $\mathcal{D}$ after projecting away the background. This is performed by applying the singular value decomposition to determine the principal component, which closely resembles the population mean of the patterns in $\mathcal{D}$, and projecting the dictionary onto the space orthogonal to the principal component (i.e., removing the mean pattern from each individual pattern). This compensation process, which simply removes the background common to all patterns, improves the dictionarys ability to discriminate between diffraction patterns. However, the uncompensated dictionary will also be used since it can better discriminate anomalies that are primarily manifested in the background.

## 4.3 Classification and Indexing

The proposed correlation matching approach to classification proceeds in two steps. First inner products between the observed patterns and patterns in the dictionaries $\mathcal{D}$ and $\mathcal{D}_c$ are computed. For each observed pattern we only store its $k$ closest dictionary matches, i.e., those dictionary patterns with the $k$ highest inner products with respect to the observed pattern. Then a pixel is classified as a grain interior, a grain boundary, or an anomaly using a decision tree classifier applied to the set of top pattern matches in the dictionary. The proposed correlation matching approach to indexing pixel orientation computes an estimate of the mean orientation over the top matching patterns in the dictionary. In order to account for noise and the non-euclidean nature of the Euler sphere, the mean angles are estimated using a specially adapted maximum likelihood estimator, introduced in the indexing subsection below. Pixel classification and indexing are performed independently and are discussed separately.

### 4.3.1 Classification

Classification of a pixel is performed by evaluating the inner products between the pixels uncompensated and compensated patterns and patterns stored in the dictionaries $\mathcal{D}$ and $\mathcal{D}_c$, respectively. Anomalies are detected as abnormally low average inner products between an uncompensated pixel pattern and patterns in $\mathcal{D}$. Specifically the average inner product similarity measure is defined as

$$\bar{\rho}(i) = \frac{1}{d} \sum_{j=1}^{d} \frac{\mathbf{Y}_i^T \phi_j}{\|\mathbf{Y}_i\| \|\phi_j\|}, \tag{4.3}$$

where is the average of the normalized inner products between pattern $\mathbf{Y}_i$ at pixel $i$ and the $d$ pattern $\{\phi_j\}_{j=1}^{d}$ in the dictionary $\mathcal{D}$.

Boundaries between grains are detected based on the lack of homogeneity of the

matches to the dictionary $\mathcal{D}_c$ of a pixel and its 8 adjacent pixels, which we call a $3 \times 3$ spatial patch. Specifically, for pixel $i$ we define the neighborhood similarity measure $\rho_{\mathcal{N}_c}(i)$ as the average amount of overlap between the $k$-NN neighborhood in $\mathcal{D}_c$ of the pixel and the $k$-NN neighborhoods, with $k = 40$, in $\mathcal{D}_c$ of the adjacent pixels on the patch:

$$\rho_{\mathcal{N}_c}(i) = \frac{1}{8k} \sum_{j \in \mathcal{I}_{3\times3}(i)} card\{\mathcal{N}_{kNN}(j) \cap \mathcal{N}_{kNN}(i)\}, \tag{4.4}$$

where $\mathcal{I}_{3\times3}(i)$ are the indices of the 8 neighbors of pixel $i$. The neighborhood similarity $\rho_{\mathcal{N}_c}(i)$ will have value close to 1 when the image patch is located in a grain. Its value will be close to zero when the image patch is centered on an anomaly. Its value will be between zero and one when the image patch is at a grain boundary (See Fig. 4.1).

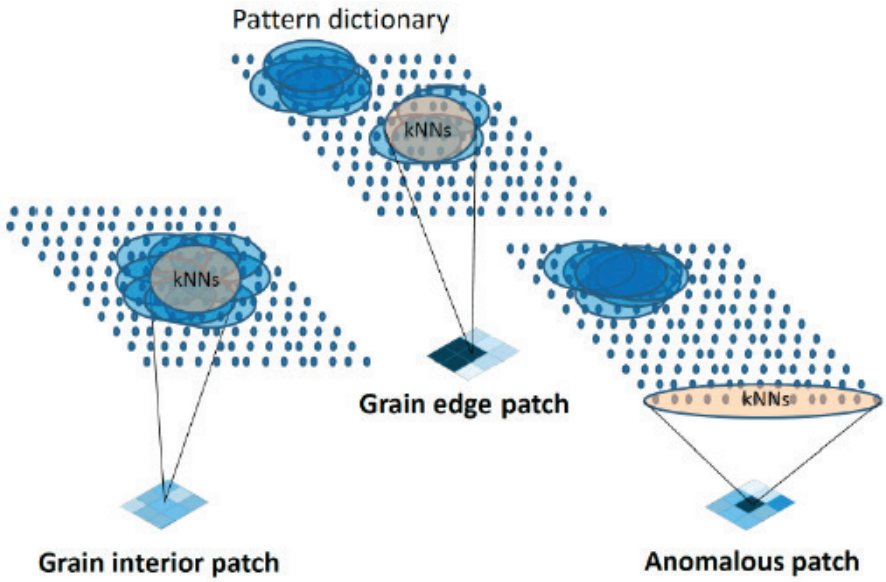



Figure 4.1: Illustration of a neighborhood similarity measure that quantifies the overlap between $k$-NN neighborhoods in an image patch. When the patch is inside a grain, the center of the patch will have a $k$-NN neighborhood that overlaps with the $k$-NN neighborhoods of the adjacent pixels. A patch that straddles a boundary will have the center pixel $k$-NN neighborhood overlapping with the neighborhoods of a small number of other pixels. When the patch is centered at an anomalous pixel there is little or no overlap between the $k$-NN neighborhoods of the center and adjacent pixels.

The inner product similarities (for anomalies) and the neighborhood similarities

(for grain boundaries) are used by a pattern classifier to assign each pixel to one of four classes. While many different types of unsupervised classifiers could be used, e.g., $k$-means, linear discriminant analysis (*Hastie et al.*, 2005) or deep learning networks (*Hinton et al.*, 2006), here we propose to cluster patterns using an unsupervised Decision Tree (DT) classifier (*Hastie et al.*, 2005) whose classification boundaries are determined so that they separate the modes (regions of concentration) of the histograms of inner-product similarity and neighborhood similarity over the sample. The non-overlapping modes can easily be separated by thresholding of the similarity value while the others can be estimated using a mode decomposition method such as Gaussian mixture modeling, also called mixture of Gaussian (MoG) modeling (*Figueiredo and Jain*, 2002). This is illustrated in Fig. 4.2. The unsupervised DT classifier is illustrated in Fig. 4.3 in Section 4.7 for the IN100 sample considered. Four clusters were discovered by the model: anomalous pixels, which divided into two subclusters of shifted background and noisy background, and normal pixels, divided into grain boundary and grain interior.

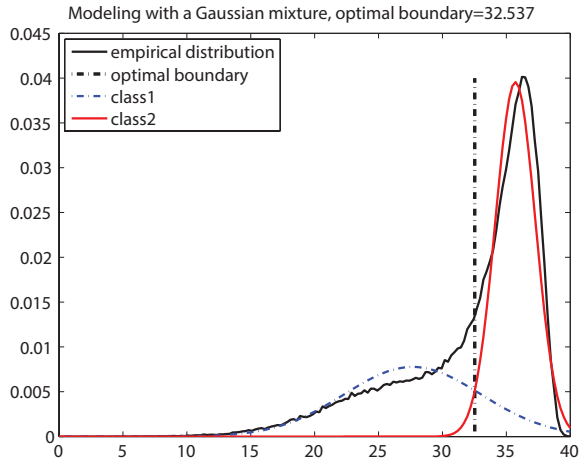



Figure 4.2: A two component Gaussian mixture model has a good fit to the neighborhood similarity histogram in right panel of Fig. 4.8. The point where the two Gaussian components cross (dotted vertical line) determines the threshold for the right lower branch of the unsupervised decision tree classifier in Fig. 4.3.

Decision tree classifiers have been previously applied to many imaging applica-

tions, e.g., land cover classification in remote sensing (*Friedl and Brodley*, 1997; *Pal and Mather*, 2003). However, there are significant differences between the proposed DT classifier and those previously applied. First, the proposed classifier is a hybrid DT that uses special features (background compensated and uncompensated patterns) and similarity measures (inner products and neighborhood intersections) specific to materials microanalysis. Second, unlike standard non-parametric DT classifiers, the proposed DT is informed by a physics model through the generated dictionary of diffraction patterns. Third, our use of unsupervised Gaussian mixture models to determine the classification thresholds means that the DT classifier threshold parameters are determined by the Gaussian mixture models and do not need to be tuned, thus eliminating the need for labeled training data and time consuming cross-validation.

### 4.3.2 Indexing

The proposed pixel indexing method is formulated as a statistical estimation problem. The pixel's crystal orientation is estimated via maximum likelihood under a Von Mises-Fisher mixture density model for the quaternions of the $k$ top dictionary matches. Note that the $k$ used in the maximum likelihood model may be different from the $k$ used to compute the neighborhood similarity for the DT classifier. We motivate the Von Mises-Fisher model as follows. Recall that the dictionary is generated for a set of predetermined orientations $\boldsymbol{\theta}$. Hence using simple table lookup, the indices of dictionary patterns found in the $k$-NN neighborhood of a pixel can be mapped to a set of $k$ orientations $\{\boldsymbol{\theta}_j\}_{j=1}^{k}$. If the pixel is in a grain then these orientation will be clustered around a true crystal orientation, that we call $\boldsymbol{\theta}$, at the pixel location. We extract a maximum likelihood estimate of this orientation using a statistical model for the variation of $\boldsymbol{\theta}_j$'s.

We assume that the $k$ best matching orientations $\{\boldsymbol{\theta}_j\}_{j=1}^{k}$ of a pixel form a ran-

dom sample from an underlying marginal density $f(;\boldsymbol{\theta})$, supported on the orientation sphere. Then the maximum likelihood estimator of $\boldsymbol{\theta}$ is

$$\hat{\boldsymbol{\theta}} = \mathbf{amax}_{\boldsymbol{\theta}} \prod_{j=1}^{k} f(\boldsymbol{\theta}_j; \boldsymbol{\theta}). \tag{4.5}$$

As is customary in the theory of directional statistics (*Mardia and Jupp*, 1999) we parameterize the density of 3-dimensional orientations by their equivalent 4-dimensional unit length quaternions $\{\mathbf{q}_j\}_{j=1}^{k}$ ($\|\mathbf{q}_j\|_2 = 1$). As described in Chapter III, due to crystal symmetry, there are many ($M$) orientations that are equivalent to each other, i.e., the representation is not unique. For any quaternion $\mathbf{q}$, the set of symmetry equivalent quaternions can be represented as $\{\mathbf{P}_m\mathbf{q}\}_{m=1}^{2M}$ where $\mathbf{P}_m$ is a $4 \times 4$ quaternion (rotation) matrix and $\mathbf{P}_1$ is the identity matrix. The matrices $\{\mathbf{P}_m\}_{m=1}^{48}$ to establish a 4D representation of the $m\overline{3}m$ point group (*De Graef and McHenry*, 2007).

The proposed model for orientations is based on a generalization of the Von Mises-Fisher density (*Mardia and Jupp*, 1999) to group structured domains on the sphere. The standard Von Mises-Fisher density over the 3-dimensional sphere with location parameter $\boldsymbol{\mu}$ and precision parameter $\kappa$ is defined as

$$\phi(\mathbf{x}; \boldsymbol{\mu}, \kappa) = c_4(\kappa) \exp(\kappa \boldsymbol{\mu}^T \mathbf{x}), \|\mathbf{x}\|_2 = 1 \tag{4.6}$$

where $c_p(\kappa) = \frac{\kappa^{p/2-1}}{(2\pi)^{p/2} I_{p/2-1}(\kappa)}$ and $I_\nu$ is the modified Bessel function of the first kind. Here $\|\boldsymbol{\mu}\|_2 = 1$ and $\kappa > 0$ control the location of the mode (maximum) and spread of the density over the sphere, respectively. The natural generalization to the group structured domain is the periodic mixture density

$$f_v(\mathbf{x}; \boldsymbol{\mu}, \kappa) = \frac{1}{2M} \sum_{m=1}^{2M} \phi(\mathbf{x}; \mathbf{P}_m\boldsymbol{\mu}, \kappa). \tag{4.7}$$

This Von Mises-FIsher mixture (VMFm) density contains $2M$ replicates of the Von Mises-Fisher density over the sphere centered at all the symmetry-equivalent values of the location parameters $\boldsymbol{\mu}$.

Substitution of (4.7) into (4.5) and use of the well-known invariance property of maximum likelihood estimation (*Lehmann and Casella*, 1998) gives a form for the maximum likelihood estimator of grain orientation $\hat{\boldsymbol{\theta}}$ in terms of the joint maximum likelihood estimators $\hat{\boldsymbol{\mu}}$ and $\hat{\kappa}$:

$$\{\hat{\boldsymbol{\mu}}, \hat{\kappa}\} = \mathbf{amax}_{\boldsymbol{\mu},\kappa} \prod_{j=1}^{k} \sum_{m=1}^{2M} \gamma_m f_v(\mathbf{q}_j; \mathbf{P}_m \boldsymbol{\mu}, \kappa). \tag{4.8}$$

where $\gamma_m = (2M)^{-1}$. Even though this maximization problem appears daunting, it can be iteratively and efficiently computed by applying the well known expectation-maximization (EM) procedure for constrained parameter estimation in mixture models (*McLachlan and Peel*, 2004). For a full account of this procedure has been elaborated in Chapter III. Note that $1/\hat{\kappa}$ gives an empirical estimate of the degree of spread of the density about the orientation estimate $\hat{\boldsymbol{\mu}}$. Thus, $\hat{\kappa}$ is a measure of confidence of this estimate.

## 4.4 Generation of the Dictionary

The dictionary approach requires a uniform sampling of orientation space $SO(3)$. Several sampling schemes are available in the literature; among the most popular schemes are a deterministic sampling method based on the Hopf fibration (*Yershova and LaValle*, 2004; *Yershova et al.*, 2009) and the HEALPix framework (Hierarchical Equal Area isoLatitude Pixelization) (*Gorski et al.*, 2005). Neither of these approaches is easily adaptable for integration with crystallographic symmetry. Instead, we employ a recently developed strategy that starts from a simple 3D cubic grid which is mapped uniformly onto $SO(3)$ (*Roca et al.*, 2014). This cubochoric mapping

is uniform, refinable, and isolatitudinal, and consists of three steps:

1. a uniform cubic grid of $(2N+1)^3$ grid points is generated inside a cube of edge length $a = \pi^{2/3}$;

2. the cube is divided into six pyramids with apex at the cube center and the cube faces as base, and each pyramid is mapped uniformly onto a sextant of a ball, using a generalization of the mapping of a square onto a curved square (*Roca and Plonka*, 2011);

3. all points inside the ball are then transformed, using a generalized inverse equal-area Lambert mapping, to the unit quaternion Northern hemi-sphere, which is isomorphous with $SO(3)$.

From the quaternion representation one can readily derive other orientation parameterizations; the Rodrigues parameterization is most suitable for the determination of the orientations that belong to the fundamental zone (FZ) for a given crystal symmetry, because the boundaries of the FZ are planar. The more conventional Euler angle representation typically has curved surfaces as the boundaries of the FZ, so that Euler angles are less useful for uniform sampling approaches. It should be noted that lower crystal symmetry implies a larger dictionary, since the fundamental zone size increases with a reduction in symmetry; acceleration of the dot product calculations by means of a GPU (graphical processing unit) is a topic of ongoing research. The use of the Rodrigues representation to determine the dictionary elements has computational advantages, but care must be taken in the case of crystal symmetries with a single diad axis, for which the Rodrigues fundamental zone becomes infinite in the direction normal to the axis. Our approach still produces a uniform sampling of orientation space, although all rotations by an angle of 180° are represented by points at infinity (which correspond to points on the outer cube surface in the cubochoric representation).

The dictionary used for the remainder of this chapter was generated by setting $N = 100$, and keeping only those orientations for which the corresponding Rodrigues vector lies inside the fundamental zone for the octahedral $m\bar{3}m$ point group. The patterns in the dictionary were down sampled to 60 by 80 pixel images. This results in a dictionary $\mathcal{D}$ with $d = 333226$ elements of dimension 4800. A representative (random) selection of 9 dictionary elements in $\mathcal{D}$ and $\mathcal{D}_c$ is shown in Fig. 4.4. The left panel of Fig. 4.5 shows how the sampling points are distributed inside the octahedral Rodrigues FZ. The right panel of the figure shows the rate of drop-off of the top 200 inner products in the compensated dictionary $\mathcal{D}_c$ for 4000 randomly selected reference elements. This decay rate is used to select the number of nearest neighbors ($k = 40$) used for the classifier described in Section 4.3 and implemented in Section 4.7.

## 4.5 Computational Considerations

The online dictionary matching algorithm requires inner product evaluation between the measured sample diffraction patterns and the dictionary diffraction patterns. Let $d$ denote the number of patterns in the dictionary (dictionary size), $L$ denote the number of pixels on the photodetector (pattern size), and $n$ denote the number of pixels on the sample (sample size). The time complexity of calculating the mean inner product over the entire measured sample is $O(L(d + n))$.

For the indexing method, to determine the $k$-NN dictionary patterns for a given pixel the $k$ largest inner products need to be determined from the set of all inner products between the pixel and patterns in the dictionary. The time complexity of the whole process is $O(Ldn)$, assuming $k \ll d$. The computation time and space grow rapidly when the dictionary size and sample size become large.

Computational challenges can be addressed in several ways. The simplest approach is to use parallelization and distributed computation. All of the algorithms introduced here can be parallelized over the spatial domain since they involve local

operations. For example, ML orientation estimation is applied independently to each pixel and DT classification is applied to each spatial patch in the sample. To speed up the $k$-NN dictionary search one can use methods from information retrieval such as dictionary caching and KD trees to accelerate the inner product evaluation and ranking process. These methods rely on the similarity of diffraction patterns over the $k$-NN neighborhood. However, as they rely on approximation, these methods may also introduce indexing errors. A study of these and other computational trade-offs is important and will be discussed in the final thesis.

## 4.6 Experimental Methods

To test the dictionary approach against an experimental data set, a polycrystalline IN100 Nickel-based super-alloy sample was selected. The sample was polished using a multi-platen Robomet.3D, using a grit of 1 micron diamond slurry on a TexMet cloth and finished with a 40 nm colloidal silica slurry on a ChemoMet cloth. Between polishing steps, a water clean was used on a ChemoMet cloth.

A backscattered electron (BSE) image was recorded using a Tescan Vega 3 XMH scanning electron microscope outfitted with a $LaB_6$ filament. An EBSD map was obtained with the same SEM and a Bruker e-Flash1000 system. A tilt angle of 70°, voltage of 30kV, working distance of 15mm, and emission current of about 1nA were used to collect the EBSD map. The spatial resolution in both $x$ and $y$ directions was 297.7nm, and a Kikuchi pattern of $80 \times 60$ pixels was acquired and stored at every point in a $512 \times 382$ map.

## 4.7 Results

A dictionary was designed as described in Section 4.4 for the octahedral $m\bar{3}m$ crystal symmetry group to match the known characteristics of the IN100 sample and

the SEM system, as described in Section 4.6. Dictionary inner-products and $k$-NN neighborhoods were computed from the detected pattern at each scan location.

We indicate four different scan locations (pixels) with qualitatively different patterns in Fig. 4.6. These locations are representative of the four different clusters of pattern, described below (see Fig. 4.3), that were discovered by the unsupervised DT classifier. In Fig. 4.7 the histograms of the dictionary inner-products are shown for each of these pixels. The shifted background and noisy background pixels have inner-products that are well separated from each other in addition to being separated from the inner-products of the grain interior and grain boundary pixels. Thus one might expect that the group of anomalous pixels, represented by the former two, could be easily separated from the group of normal pixels, represented by the latter two, using any reasonable clustering technique based on the inner-products. On the other hand, the inner-product histograms for the grain interior pixel and the grain boundary pixel are overlapping. This overlap makes it difficult to distinguish these two types of pixels and justifies the need for the more sensitive neighborhood similarity measure that is better able to separate them.

Figure 4.8 shows the full-sample histograms of inner-products $\bar{\rho}$ with respect to $\mathcal{D}$ and neighborhood similarities $\rho_{\mathcal{N}_c}$ with respect to $\mathcal{D}_c$, respectively, over all pixels and over all patterns in the dictionaries. The left panel of Fig. 4.8 can be interpreted as the addition of all other inner-product histogram to the four histograms shown in Fig. 4.7. Similarly to Fig. 4.7, the full-sample inner-product histogram exhibits three well separated modes, which confirms that anomalous pixels can easily be separated from the normal pixels on the basis of thresholding each pixel's average inner product measure. The three modes correspond to anomalous pixels with inner-products clustered around 0.7 and 0.97 (not visible in the range of $\bar{\rho}$ plotted) and normal pixels with inner-products clustered around 0.997.

The right panel of Fig. 4.8 shows the histogram of all neighborhood similarities

computed with neighborhood size $k = 40$. The latter histogram is bimodal and asymmetric about its mean. The higher mode located near 37 corresponds to pixels whose $k$-NN neighborhood in $\mathcal{D}_c$ has high overlap with the $k$-NN neighborhoods of its 8 adjacent pixels. Such pixels are likely to be interior to a grain. The lower mode located near 26 corresponds to patches of pixels near grain boundaries, patches that have less similar $k$-NN neighborhoods than in-grain pixels. To separate these tow modes, we fitted a two component mixture-of-Gaussian model to the histogram in the middle panel using the MoG EM algorithm (*McLachlan and Peel*, 2004). The result of this fit is shown in Fig. 4.2. The point of intersection of each of the fitted Gaussian densities (shown in the Figure by vertical dotted line) is used as the DT classification threshold for discriminating between grain boundaries and grain interiors.

Figure 4.3 shows the unsupervised DT classifier used to cluster observed patterns. The lower four nodes are leaf nodes while the upper three nodes are decision nodes for which thresholds are used to assign labels. These thresholds were determined from the observed histograms shown in Fig. 4.8 as described above. The DT classifies each pixel on the sample based on the pattern matches in the dictionaries. The top node, labeled "observation" classifies pixel as a "anomaly" or as "normal" by thresholding the average inner-product $\bar{\rho}$ between the pattern at the center pixel of the patch and the patterns of the dictionary. Any threshold between $\bar{\rho} = 0.97$ and $\bar{\rho} = 0.99$ would separate the normal pair (grain interior, grain boundary) from the anomalous pair (noisy background, shifted background) and we selected the midpoint. The anomalous patterns are further subclassified on the left branch of the DT by applying a threshold between 0.7 and 0.95 to $\bar{\rho}$. The DT classifies normal pixels as either grain-interior or grain-boundary pixels by applying the threshold 32/40 to the neighborhood similarities $\rho_{\mathcal{N}_c}$. Representative patterns are shown at the bottom of Fig. 4.3 that have been identified as belonging to each of the respective four clusters.

Figure 4.9 shows the pixel neighborhood similarities (left panel) and the pixel

classifications (right panel) as images as determined by the unsupervised DT classifier. In the classification image the blue/red regions and black regions respectively correspond to pixels classified as anomalies and boundaries. These class labels can be used for segmentation of the grains and identification of the anomalies. A blowup of these images in Fig. 4.9 is shown in Fig 4.10 for a small region.

Next we illustrate the use of the dictionary for estimation of the Euler angles in the sample. For the same subregion as in Fig. 4.10, Fig. 4.11 illustrates the OEM (original equipment manufacturer) orientation estimates (top left image), an estimate equal to the orientation of the element of the dictionary having largest normalized inner product with the pixel pattern (top right image), the proposed ML estimator $\hat{\boldsymbol{\mu}}$ of the orientation (VMFm location parameter) computed from the top $k = 4$ matches in the dictionary (bottom left image), the proposed ML VMFm estimator $\hat{\boldsymbol{\mu}}$ computed with $k = 10$ (bottom right image). The ML estimates $\hat{\kappa}$ of the VMFm scale parameters were computed for each pixel. This $\hat{\kappa}$ parameter is inversely proportional to the width of the estimated VMFm model over the 3 dimensional quaternion sphere. Figure 4.12 shows images of these ML estimates translated into angular uncertainty (in degrees) by using the transformation $\Delta\theta = \arccos(1 - 1/\kappa)\frac{180}{\pi}$, which is the $1/e$-width of the VMFm distribution in any fundamental zone. Note that the OEM image in Fig. 4.11 has many spurious orientation estimates within grains unlike the proposed dictionary based methods. Note also that the ML orientation estimates produce smoother in-grain orientations. The $k = 4$ and $k = 10$ ML orientation estimates have low confidence (high variance) at some locations on grain boundaries and in anomalous region at bottom right. This low confidence is quantified by the ML estimator of the scale parameter $\kappa$ of the VMFm model, shown in Fig. 4.12 for $k = 4$ and $k = 10$.

## 4.8 Conclusion and Future Work

We have introduced a novel method for indexing polycrystalline materials that uses both mathematical-physics modeling and mathematical-statistics modeling. The physics-based forward model is discretized into a dense dictionary of diffraction patterns that are indexed by Euler angle triplets. The dictionary is fixed for each crystal symmetry group and each SEM instrument configuration. The statistical model is based on the group symmetry of quaternion representation of the Euler angles on the 3-sphere in 4 dimensions. A feature of this method is that it performs classification, segmentation, and indexing in the unified framework of dictionary matching. A feature of the indexing method is that it incorporates a concentration parameter that can be estimated jointly with the Euler angles of a pixel or of a grain. This concentration parameter can be used to report the degree of confidence one can have in the Euler estimates. An iterative maximum likelihood estimator was proposed for estimating the orientation and associated confidence parameters in a statistical Von MisesFisher mixture model. The method was illustrated on a single sectional slice of a Nickel alloy sample.

As the proposed indexing method is pixel driven it is directly applicable to indexing over 3 dimensional volumes. Future work will include algorithm acceleration to make full volumetric indexing fast enough to be practical. Potential acceleration methods include multi-resolution and multi-scale trees, fast coordinate ascent ML optimization, and parallelization. Other areas for future work include robustification of the dictionary to model mismatch, sensitivity to reductions in detector image resolution, and extensions of the dictionary approach to other electron diffraction modalities, such as electron channeling patterns and precession electron diffraction. Preliminary investigations indicate that, while dictionary-based classification appears to be robust to model mismatch, the proposed dictionary-based indexing algorithm is somewhat sensitive to model mismatch. This suggests that the dictionary design

may need to be fine tuned to the SEM instrument in addition to the samples crystal symmetry group. The extension to other SEM modalities such as EDS is also possible but would require development of dictionaries that capture other types of data (e.g., spectra).

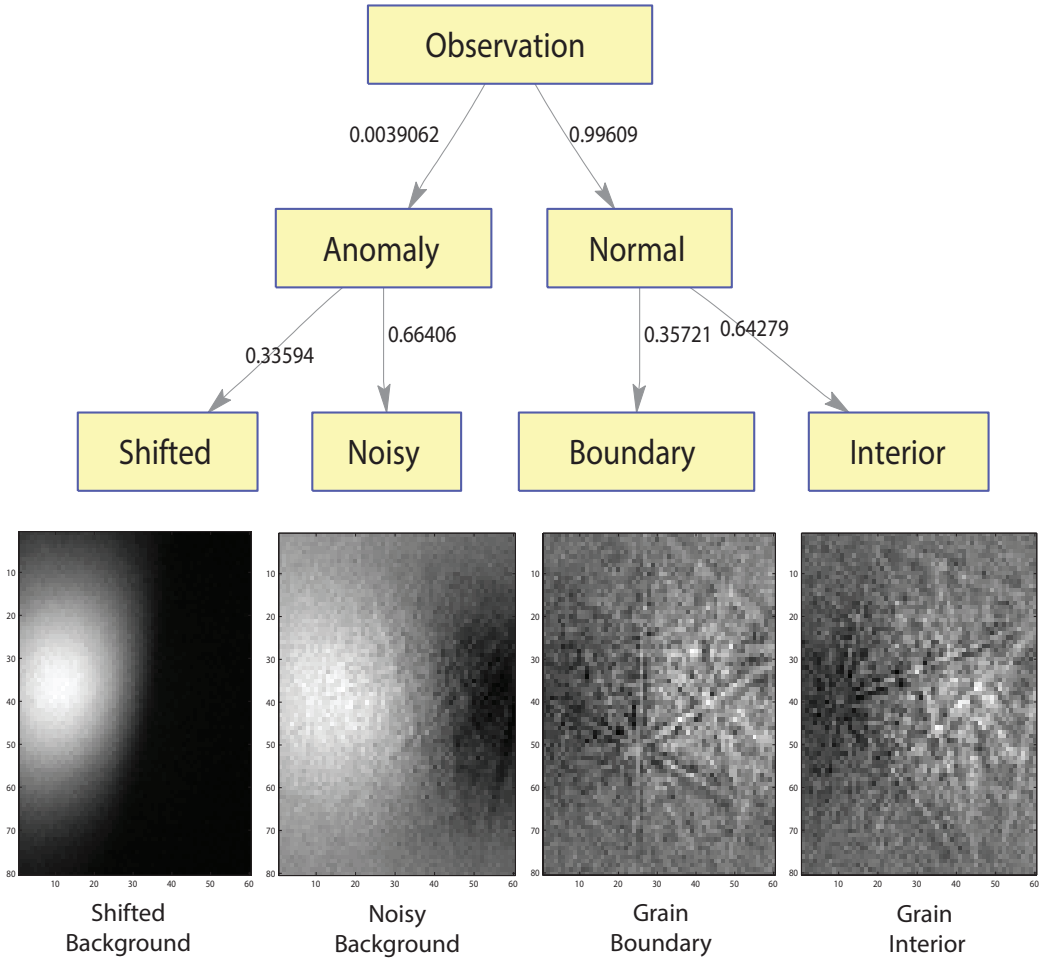



Figure 4.3: Decision tree for clustering detected patterns on the IN100 sample with examples of patterns in each cluster below the leaf nodes at bottom. Physical locations of these patterns on the sample are shown in Fig. 4.6. The classifier uses the uncompensated pattern matches of a pixel to decide between shifted and noisy background at lower left. It uses the homogeneity of the compensated pattern matches over a $3 \times 3$ patch to decide between grain boundary and grain interior on the right. The number on each decision tree branch is the proportion of patterns at the parent node that were classified with label of child node.

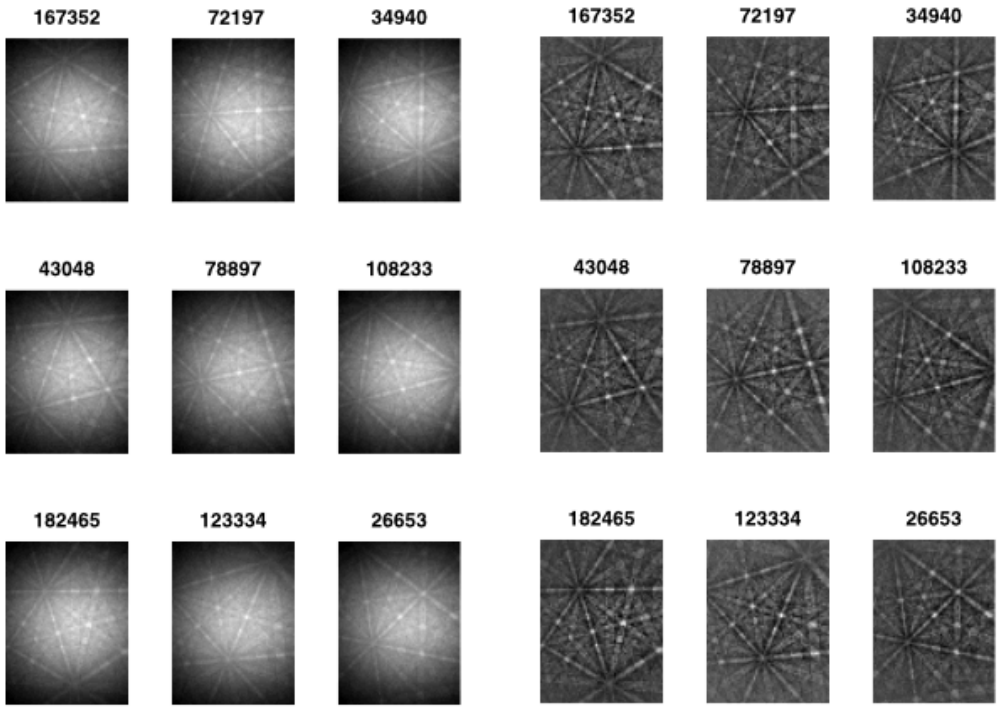



Figure 4.4: A random subset of the 333226 elements in the dictionary generated for the IN100 sample. Shown are 9 representative patterns, each 6080 pixels, in the uncompensated (Left) and compensated (Right) versions of the dictionary.

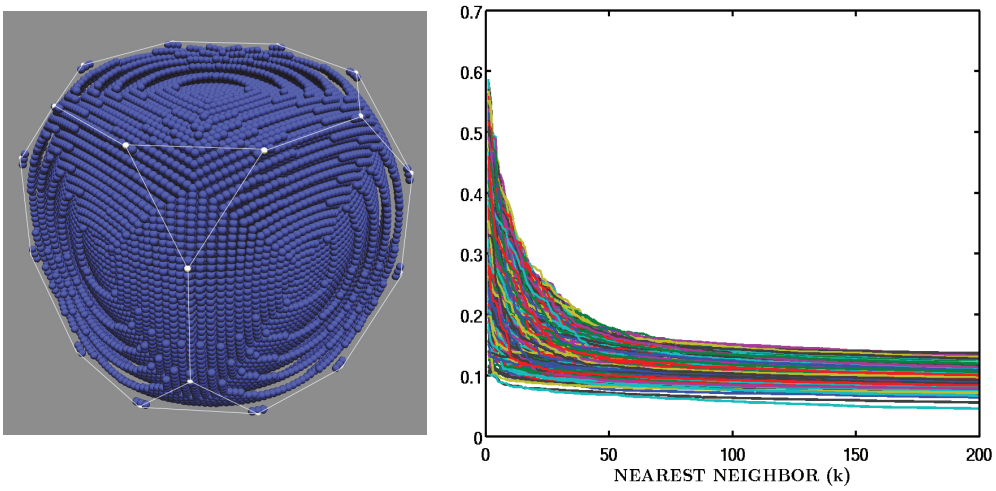



Figure 4.5: Left: The sampling pattern (at 1/8 density) of dictionary Rodrigues vectors in the fundamental zone (solid lines) of the cubic symmetry point group $m\bar{3}m$. Right: Graph of the top 200 normalized inner products between the entire compensated IN100 dictionary and a randomly selected set of 4000 reference elements in the IN100 dictionary. For each of the reference elements the top 200 inner products have been rank ordered in decreasing order and plotted. A knee occurs in vicinity of $k = 40$ for which the normalized inner product drops by at least 1/3 of the maximum value.

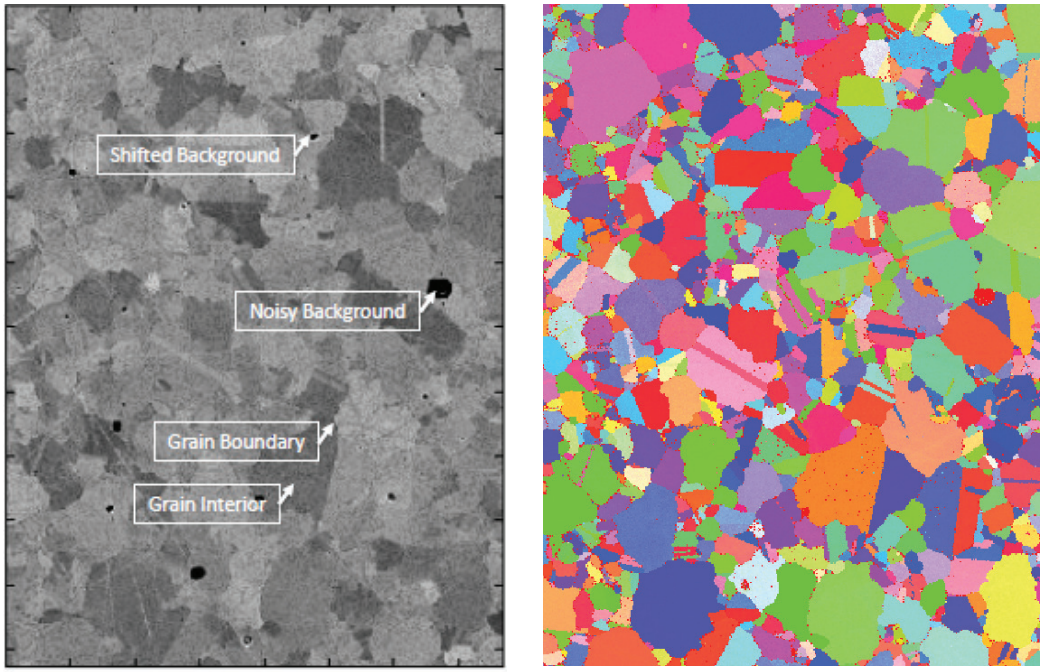



Figure 4.6: Raw SE and EBSD images of IN100 sample generated by the Tescan Vega SEM with native OEM software. Left: SE image of the IN100 sample showing physical locations of the four patterns shown at bottom of DT classifier in Fig. 4.3. The inner-product histograms for the diffraction patterns at these locations are shown in Fig. 4.7. Right: IPF colored EBSD pixel orientation image.

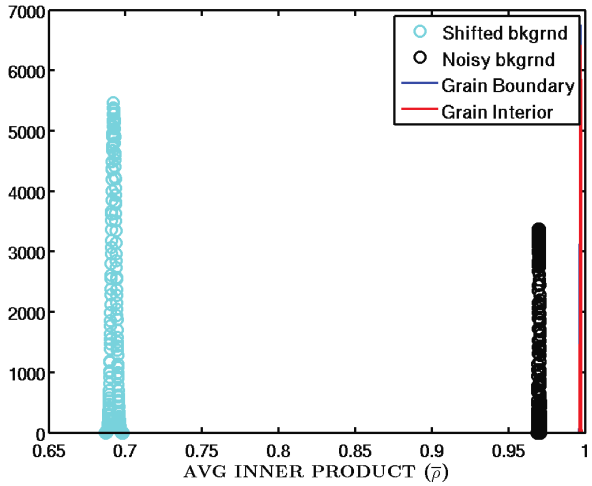



Figure 4.7: Histograms of the inner products between patterns in the dictionary and the patterns of the four EBSD scan locations (pixels) shown in Fig. 4.6. The histograms for the shifted background and the noisy background are well separated from each other and from the histograms for the grain boundary and grain interior pixels in Fig. 4.6. These latter two histograms are very concentrated near 1 and overlap each other (not distinguishable at this scale).

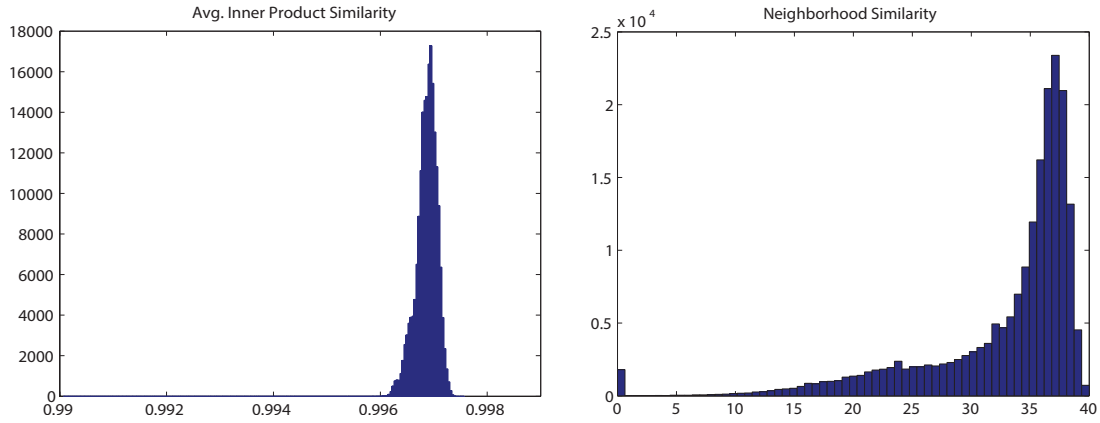



Figure 4.8: Left: histogram of normalized inner products between detected patterns on the sample and dictionary patterns restricted to the range $\bar{\rho} = [0.99, 0.999]$ to reveal the modes associated with grain interior and grain boundary patterns. Two other modes (not shown) are located near $\bar{\rho} = 0.7$ and $\bar{\rho} = 0.97$ corresponding to background shift and noisy background pixels, respectively (see Fig. 4.7). Right: histogram of neighborhood similarity measures between dictionary neighborhoods over a $3 \times 3$ patch centered at each pixel in the sample for neighborhood size $k = 40$.

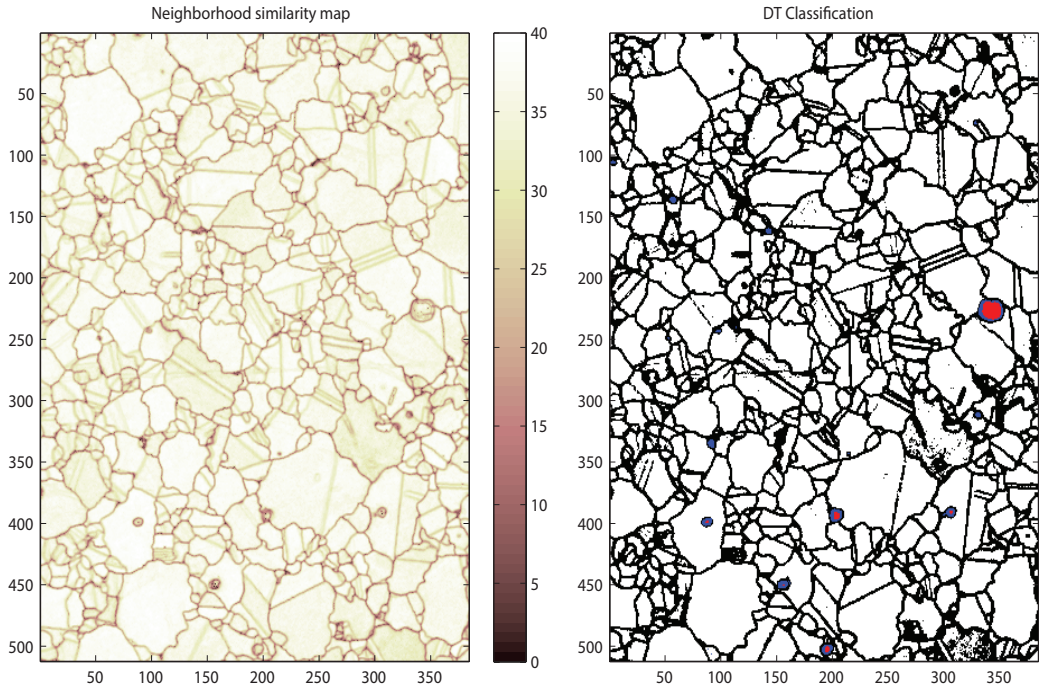



Figure 4.9: Left: An image rendering of the (un-normalized) neighborhood similarity measure ($k = 40$ nearest neighbors in dictionary) used in the right branch of the DT classifier in Fig. 4.3. Right: A map of the pattern classes in the IN100 sample as determined by the DT classifier in Fig. 4.3. The colors encode the four classes as follows: white=grain interior, black=grain boundary, red=noisy background, and blue=shifted background. Note that the black boundaries effectively segment the sample according to crystal orientation.

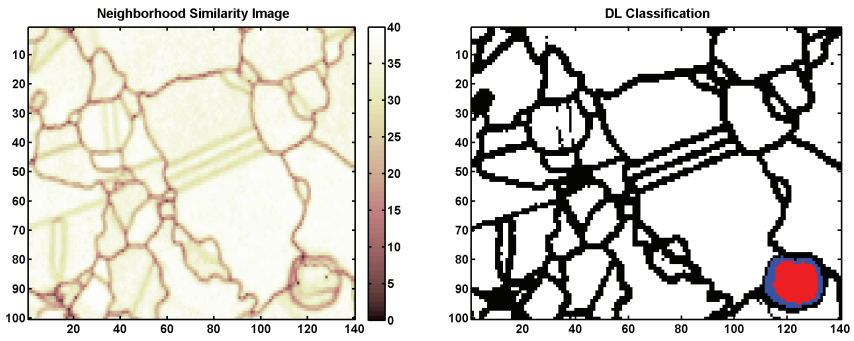



Figure 4.10: Blowup of a small region right of center in each of the images of Fig. 4.9.

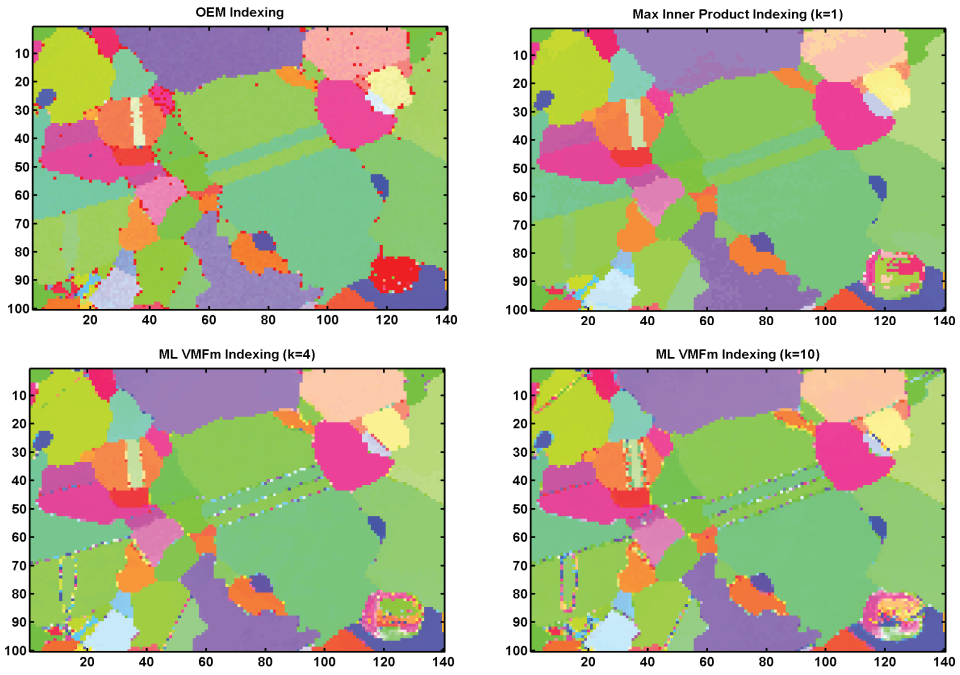



Figure 4.11: Comparison of orientation indexing. Top left: IPF images generated by OEM software. Top right: IPF image obtained by rendering the top matching patterns in the dictionary (this is identical to the ML estimator of the orientation using VMFm model with $k = 1$). Bottom left: Image of ML estimates of orientation using VMFm model on the orientations of the $k = 4$ top dictionary matches. Bottom right: Same as bottom left except that $k = 10$. Note that the OEM image has many spurious orientation estimates within grains unlike the other dictionary based methods. Note also that the ML orientation estimates produce smoother in-grain orientations. The $k = 4$ and $k = 10$ ML orientation estimates have low confidence (high variance) at some locations on grain boundaries and in anomalous region at bottom right. This low confidence is quantified by the ML estimator of the scale parameter $\kappa$ of the VMFm model, shown in Fig. 4.12 for $k = 4$ and $k = 10$.

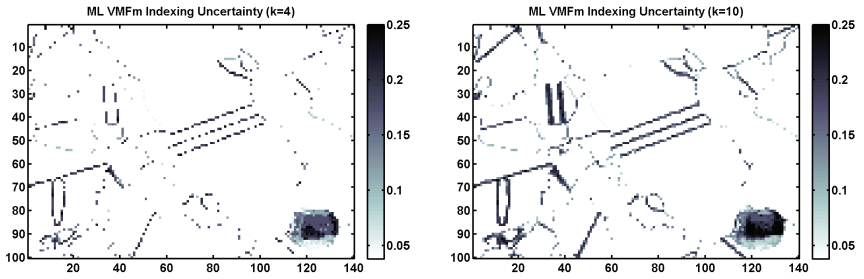



Figure 4.12: Images of the ML estimator of the orientation standard deviation (in degrees) obtained by ML estimation of the scale parameter $\kappa$ of the VMFm model corresponding to the bottom two sub-figures of Fig. 4.11. The angular standard deviation ranges from 0.05 degrees to 0.5 degrees but those values above 0.25 have been hard-limited for ease of visualization (only 1% of all values are above 0.25 degrees). Note that the areas of least confidence are in the vicinity of boundaries and anomalies. The highest standard deviations occur at pixels that straddle boundaries between grains having the highest misorientation.

# CHAPTER V

# Coercive Region-level Registration for Multi-modal Images

## 5.1 Introduction

In previous chapters, we introduced the statistical model which incorporates the crystal symmetry property and how this model can be applied to the mean crystal orientation estimation and the dictionary-based indexing approach for EBSD data. In this chapter, we will addresses the two problems of multi-modal image registration and image segmentation, which are inevitable steps before fusion, within a single framework. Multi-modal registration refers to registration of images acquired by different sensor/scanner types. It has been applied to many areas, e.g. medical images, microscopy images, and radar images, to combine information from different modalities and provide more comprehensive understanding about the true object. Image segmentation, the partitioning of an image into meaningful regions, is an important step in image analysis and understanding.

In this chapter, we focus on multi-modal registration and segmentation as applied to scanning electron microscope (SEM) images of materials; the methods to be discussed are equally applicable to other multi-modal images that share spatial structure.

*This chapter is based on (*Chen et al.*, 2015b)

SEM techniques are widely used in materials science for material characterization, for example detection of defects that may cause fatigue when operating. Segmentation is of interest to map locations of *grains*, uniform regions occupied by continuous crystal lattices, since grain structure is a principal factor in determining the properties of a polycrystalline material such metallic or ceramic materials (*Shah and Duhl*, 1988). Multi-modal registration is desired because different scanning electron modalities carry complementary information (*Wang et al.*, 2005; *Nordmark et al.*, 2008). For example, Backscattered Electrons (BSE) provide information about topography and local fine-scale surface texture (*Goldstein*, 2003) while Electron Backscatter Diffraction (EBSD) measures crystal orientation which is useful in locating grains and grain boundaries (*Schwartz et al.*, 2009).

Multi-modal registration is made challenging by the fact that images from different modalities may have different resolutions, values that lie in different spaces (e.g., scalars vs. vectors), and different levels of distortion. In SEM for instance, these differences are due to different electron beam geometries, sensors, and recording electronics. Furthermore, there is no complete forward imaging model that jointly characterizes the multi-modal signals. Without the complete image formation model, pixel-level registration methods (*Bonnet and Liehn*, 1988; *Glasbey and Martin*, 1996; *Tang et al.*, 2007), i.e., those that establish correspondences between pixels, usually resort to interpolation, a somewhat ad-hoc solution that may bias the resulting images toward excessive smoothness. On the other hand, image registration has to assume certain transformation model, such as rigid body displacement (*Hill et al.*, 2001), local deformations (*Rueckert et al.*, 1999; *Szeliski and Coughlan*, 1997), or other relative differences (*Meyer et al.*, 1997; *Rohr et al.*, 2001) between the reference and target images. Due to the fact that there is no transformation model that adequately describes the image distortion between different microscopy modalities, the transformation model mismatch may also introduce artificial errors in the registration

result.

Image segmentation is a process of partitioning an image into multiple segments (i.e., sets of connected pixels) which is useful in identifying and locating the grains in different SEM modalities of polycrystal materials, such as BSE (*Chuang et al.*, 2008) or EBSD (*Schwartz et al.*, 2009) images. Kass et al. formulated image segmentation as an optimization problem with an edge-based model, called Snakes (*Kass et al.*, 1988), which has an enormous impact in image segmentation community. Yet, it has the drawbacks as sensitive to initialization and non-trivial extension to color or texture images (*Cremers et al.*, 2007). Region-based models which utilize level set functions with variational methods are proposed to segment images into 2-phases (*Amadieu et al.*, 1999; *Chan and Vese*, 2001) or multi-phases (*Vese and Chan*, 2002). It has been shown that these approaches are more robust to noise and to varying initialization. However, one has to determine the number of segments a priori and they do not have the flexibility of adding or merging regions during the process, which is an essential step in our grain identification problem. Most of all, since all of the above approaches segment the image in each modality independently, finding the correspondences of the segmented regions between different modalities as required in image fusion becomes another difficult problem, especially when the segmentation results are not consistent across modalities due to noise or low contrast.

In this work, we propose a coercive region-level approach to simultaneously register and segment images of different modalities that share similar spatial structure. The algorithm is initialized by segmenting one image by a standard method and coarsely mapping the result onto the other image. Then the two images are registered at the region level and further segmented through alternating minimization of a statistically-based objective function which is adopted from region growing and mering techniques (*Zhu et al.*, 1995; *Nock and Nielsen*, 2004). There are several advantages of our approach. First, the region-level approach is free of pixel value

interpolation and its inherent assumptions. Second, it takes advantage of modalities with better discriminative power, improving the overall segmentation result. The approach also preserves region correspondences to facilitate data fusion. Lastly, both registration and refinement of segmentation are driven by statistical models. In particular, we propose hypothesis tests to detect boundaries that are missed by the initial segmentation due to low contrast.

The rest of this chapter is organized as follows. In Section 5.2, we describe the objective function, statistical models for data from different modalities, and optimization methods for the region-level registration algorithm. In Section 5.3, we focus on hypothesis testing for detecting missing boundaries. Section 5.4 shows experimental results for synthetic and real materials images and compares several different approaches. Section 5.5 concludes this chapter.

## 5.2 Algorithm Framework and Models

### 5.2.1 Objective Function

We assume that there are two images from different modalities. A pixel location $(x, y) \in \mathcal{I}_i, i \in \{1, 2\}$ is represented as vector $\mathbf{p}$, where $\mathcal{I}_i$ is the spatial domain for the $i$-th modality. The pixel value at $\mathbf{p}$ is given by a function $I_i(\mathbf{p})$. Note that the values $I_1(\mathbf{p}_1)$ and $I_2(\mathbf{p}_2)$ may lie in different spaces. The region-level registration problem is to find partitions of $\mathcal{I}_i$, $S_i = \{R_{i_j}\}_{j=1}^{N_i}$, where each segment $R_{i_j}$ is a collection of connected pixel locations and $N_i$ is the number of segments, to minimize the following objective function:

$$U(S_1, S_2) = J(S_1, I_1) + J(S_2, I_2) + \lambda D(S_1, S_2), \tag{5.1}$$

where $J(S_i, I_i)$ is the intra-modal energy function that measures how well the segmentation fits the image data and $D(S_1, S_2)$ is the inter-modal energy function that

coerces the segmentation results to be topologically similar, motivated by the fact that they share the same underlying physical structure. The parameter $\lambda$ controls the relative importance of the two terms.

In this chapter, we define the inter-modal energy $D(S_1, S_2)$ to be the number of segment boundaries that are present in one modality but not the other. This number is easily tracked because our algorithm maintains tight correspondences between segments in the two images. More generally, segment structure can be represented by a connected adjacency graph and the inter-modal energy can be any function which measures the topological distance between two graphs. The intra-modal energy $J(S_i, I_i)$ is defined by the statistical models described in the following subsection.

### 5.2.2 Statistical Models for Pixel Values

In the materials context, each segment $R_{ij}$ corresponds to a grain, a continuous crystal lattice. Motivated by this, we assume that the observed values within a segment are similar and can be modeled by i.i.d. random variables following a distribution with the same parameters. In the sequel, the image modality subscript $i$ is suppressed for simplicity. Let the probability density function (PDF) of the distribution for one modality be denoted by $f(I(\mathbf{p})|\boldsymbol{\alpha})$, where $\boldsymbol{\alpha}$ represents the parameters specifying the model. The intra-modal energy function in (5.1) given a set of segments $S = \{R_1, R_2, ..., R_N\}$ is defined as:

$$J(S, I) = \sum_{j=1}^{N} \left[ -\sum_{\mathbf{p} \in R_j} \log f(I(\mathbf{p})|\hat{\boldsymbol{\alpha}}_j) + \epsilon \int_{\partial R_j} dl \right] \tag{5.2}$$

where $\partial R_j$ is the boundary of region $R_j$ with counter-clockwise definition and $\hat{\boldsymbol{\alpha}}_j$ is the maximum-likelihood (ML) estimate for the parameters of region $R_j$. The first term is the negative log-likelihood of observations which penalizes grain inhomogeneity and the second term penalizes the boundary length, where $\epsilon$ controls the level of

smoothness.

In this chapter, we consider the EBSD and BSE images of one section of a material as our input. Note that other image types can be used directly given properly defined statistical models. For BSE images, since the pixel values are scalars, the intensities in the same grain region are modeled by a univariate Gaussian $\mathcal{N}(\mu_j, \sigma_j^2)$, where $\mu_j, \sigma_j^2$ are the mean and variance of $R_j$. Notice that $\mu_j$ and $\sigma_j^2$ are unknown parameters to be estimated from image data.

For EBSD images, the pixel values characterize the local crystal orientation, which can be represented by Euler angles (*Eberly*, 2008), Rodrigues vectors (*Morawiec and Field*, 1996) or quaternions (*Altmann*, 2005). We choose the unit-quaternion representation, i.e. a $\mathbf{q} \in S^3$, the 3-dimensional unit sphere in $\mathbb{R}^4$. This allows use of the von Mises-Fisher (VMF) distribution in directional statistics (*Mardia and Jupp*, 1999), a natural generalization of the multivariate Gaussian distribution to the sphere $S^{(p-1)} \subset \mathbb{R}^p$ (here $p = 4$). Again, due to the crystal symmetry which causes there to be more than one quaternion representation corresponding to a single crystal orientation, we adopt the mixture of VMF distributions as proposed in Chapter III to model the orientations. To briefly describe the model, let $\mathcal{G}$ be a group of quaternion matrices $\{\mathbf{P}_1, \ldots, \mathbf{P}_M\}$ which define the symmetry actions that map a quaternion $\mathbf{q}$ to its symmetric equivalents. The PDF of the pure VMF distribution is $\phi(\mathbf{x}; \boldsymbol{\mu}, \kappa) = c_p(\kappa) \exp(\kappa \boldsymbol{\mu}^T \mathbf{x})$, where $\mathbf{x}, \boldsymbol{\mu} \in S^{(p-1)}$, $\boldsymbol{\mu}$ is the mean direction, $\kappa$ is the concentration parameter, $c_p(\kappa) = \frac{\kappa^{p/2-1}}{(2\pi)^{p/2} I_{p/2-1}(\kappa)}$ and $I_p(.)$ is the modified Bessel function of the first kind with order $p$. The density function of the VMF mixture distribution is then given by

$$f_v(\mathbf{x}; \boldsymbol{\mu}, \kappa) = \sum_{m=1}^{M} \frac{1}{M} \phi(\mathbf{x}; \mathbf{P}_m \boldsymbol{\mu}, \kappa) \tag{5.3}$$

The parameters $\boldsymbol{\mu}$ and $\kappa$ can be estimated from image data through the Expectation-

Maximization algorithm derived in Chapter III.

### 5.2.3 Optimization

We minimize the objective function (5.1) by alternately fixing $S_1^{(k)}, S_2^{(k)}$ and solving (5.4) and (5.5),

$$S_2^{(k+1)} = \arg\min_S J(S, I_2) + \lambda D(S_1^{(k)}, S) \tag{5.4}$$

$$S_1^{(k+1)} = \arg\min_S J(S, I_1) + \lambda D(S, S_2^{(k+1)}) \tag{5.5}$$

where $k$ is the iteration index. Typically 2-3 iterations suffice.

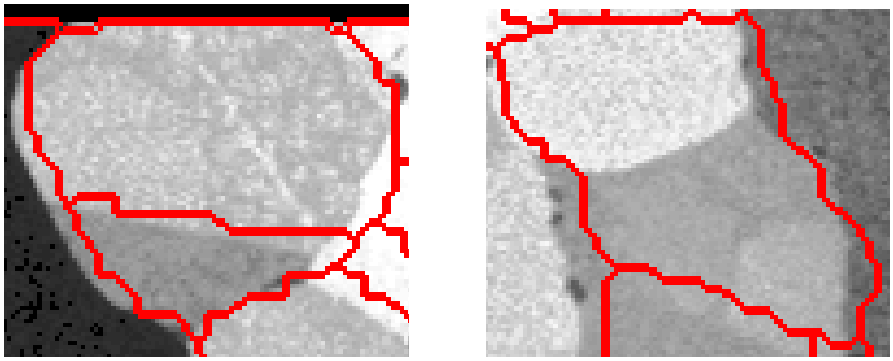
To initialize the algorithm, the initial segmentation of the first modality, $S_1^{(0)}$, is obtained by using a suitable image segmentation method. For example, the Voronoi-based method in (*Bachmann et al.*, 2011) or our proposed dictionary-based approach presented in Chapter IV can be applied to EBSD images and the Stabilized Inverse Diffusion Equation method to BSE images (*Chuang et al.*, 2008). Since EBSD data provides crystal orientation which defines grain regions more accurately, we choose to start with EBSD segmentation in this work. Next, to account for global misalignment and any resolution difference between the modalities, we determine an affine transformation by registering the outer boundary of the material sample in one modality to the other. The transformation is then used to map $S_1^{(0)}$ onto the other modality, yielding the initial segmentation $S_2^{(0)}$. Note that due to localized distortions between the modalities, the initial segmentation $S_2^{(0)}$ may be misaligned with the image values as shown in Fig.5.1a and therefore needs to be registered.

Optimizing (5.4) and (5.5) is done in two steps. The first step is to consider splitting regions in the current segmentation by adding boundaries. In Section 5.3, we propose a hypothesis testing approach for this purpose based on the statistical model

(5.2). The second step is to register the misaligned boundaries. Due to the fact that adjusting boundary positions does not change the topology of the segment structure, the inter-modal energy function $D(S_1, S_2)$ is not changed in this step, reducing (5.4) and (5.5) to the intra-modal energy function $J(S, I)$ alone, which is given by our statistical model. We use the Region Competition algorithm (*Zhu et al.*, 1995) to minimize $J(S, I)$. This algorithm applies gradient descent to move pixels comprising the boundaries $\partial R_j$ along their respective normal directions. There are two forces driving the movement corresponding to the two terms in (5.2): the statistics force which comes from the distribution model for the pixel values, and the smoothing force which drives the boundary to have smaller curvature. More details are given in (*Zhu et al.*, 1995).

## 5.3 Hypothesis Tests for Missing Boundaries

This section elaborates upon the first step in solving (5.4) and (5.5), namely hypothesis testing to determine whether a region $R \in S$ should be split into two based on the observed image values. We refer to this as the missing boundary problem. Recall that $R$ may come from the initial segmentation result from another modality and may not fit the current image data. Figure 5.1 shows examples of misalignment and a missing boundary. One can see that both of the situations have multi-modal distributions of pixel values within the initially defined regions but only Fig.5.1b shows a missing edge that should be identified. Therefore, a region $R$ is declared as having a missing boundary if and only if it satisfies the following two conditions: (1) The pixel values are multi-modally distributed. (2) The multi-modal distribution is unlikely to be caused by misalignment. We develop two hypothesis tests for the two criteria. The first hypothesis test uses the Generalized Likelihood Ratio Test (GLRT) (*Neyman and Pearson*, 1992) to test whether the pixel values are multi-modally distributed. The second hypothesis test differentiates misalignment from a



(a) Misalignment (b) Missing Edge

Figure 5.1: (a) shows a misaligned boundary while (b) indicates a missing boundary. Notice that in both situations, pixel values within the regions outlined in red are multi-modally distributed.

missing boundary.

### 5.3.1 Hypothesis Test for Multi-Modality

Recall from Section 5.2.2 that the set of pixel values within a region $R$ are modeled by a distribution $f(I(R);\boldsymbol{\alpha})$ with unknown parameters $\boldsymbol{\alpha}$, where $I(R) = \{I(\mathbf{p})\}_{\mathbf{p}\in R}$ are the observed pixel values in $R$. Assume there exists a boundary $\psi$ which partitions $R$ into two sub-regions $R_+, R_-$ with parameters $\boldsymbol{\alpha}_+, \boldsymbol{\alpha}_-$. The two hypotheses are $H_0$: region $R$ is indeed a single region, i.e. $\boldsymbol{\alpha}_+ = \boldsymbol{\alpha}_-$; and $H_1$: $R$ consists of two regions. The GLRT has the following form:

$$\begin{aligned}\log \Lambda_{GLR} &= \log \max_{\psi} \frac{\max_{\{\boldsymbol{\alpha}_+,\boldsymbol{\alpha}_-\}} f(I(R);\boldsymbol{\alpha}_+,\boldsymbol{\alpha}_-,\psi)}{\max_{\{\boldsymbol{\alpha}_+=\boldsymbol{\alpha}_-\}} f(I(R);\boldsymbol{\alpha}_+,\boldsymbol{\alpha}_-,\psi)} \\ &= \max_{\psi} \sum_{p\in R_+} \log f(I(p)|\hat{\boldsymbol{\alpha}}_+) + \sum_{p\in R_-} \log f(I(p)|\hat{\boldsymbol{\alpha}}_-) \\ &\quad - \sum_{p\in R} \log f(I(p)|\hat{\boldsymbol{\alpha}}) \gtrless_{H_0}^{H_1} \lambda \end{aligned} \tag{5.6}$$

where $\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\alpha}}_+, \hat{\boldsymbol{\alpha}}_-$ are the ML estimates of the parameters under the null and alternative hypotheses and $\lambda$ is the coefficient in (5.1). The GLRT can be viewed as a

trade-off between the improvement in the intra-modal energy and the penalty of $\lambda$ paid in the inter-modal energy when inserting a boundary. The boundary length penalty is neglected here for simplicity but can be included easily.

In the following subsections, we derive the GLRT for univariate Gaussian and VMF distributions given the boundary $\psi$. We only discuss the equal variance (concentration parameter) case due to the paper length constraint. These expressions supply the objective function, denoted as $\Lambda_{GLR|\psi}$, to be maximized with respect to $\psi$ in (5.6). We use the Region Growing algorithm (*Zhu et al.*, 1995) to locate the optimal boundary $\psi$. The algorithm partitions a region starting from two seed pixels and greedily adds neighboring pixels until all pixels in the region are chosen.

#### 5.3.1.1 Multi-modality Test for Univariate Gaussian Distribution

The GLR $\Lambda_{GLR|\psi}$ given boundary $\psi$ for testing mean equality between two Gaussian distributions has the following form (*Seize*, 1977):

$$\Lambda_{GLR|\psi} = \left( \frac{\hat{\sigma}_0^2}{\frac{n_+}{n}\hat{\sigma}_+^2 + \frac{n_+}{n}\hat{\sigma}_-^2} \right)^{n/2}, \tag{5.7}$$

where $\hat{\sigma}_0^2, \hat{\sigma}_+^2, \hat{\sigma}_-^2$ are the ML estimators of the variances under the null and alternative hypothesis and $n, n_+, n_-$ are the numbers of pixels in regions $R, R_+, R_-$. The optimization of the boundary $\psi$ then takes the form

$$\hat{\psi} = \arg\min_{\psi} \sum_{\mathbf{p}\in R_+} (I(\mathbf{p}) - \hat{\mu}_+)^2 + \sum_{\mathbf{p}\in R_-} (I(\mathbf{p}) - \hat{\mu}_-)^2 \tag{5.8}$$

where $\hat{\mu}_\pm$ is the ML estimate of the mean in $R_\pm$.

#### 5.3.1.2 Multi-modality Test for von Mises-Fisher Distribution

The VMF mixture distribution is reduced to single VMF through transoforming the samples by the symmetry operator towards the mean direction estimated by the EM algorithm introduce in ChapterIII. According to the derivation of the ML estimators in (*Dhillon and Sra*, 2003), $\Lambda_{GLR|\psi}$ has the following form:

$$\Lambda_{GLR|\psi} = \frac{c_p(\hat{\kappa}_1)^n}{c_p(\hat{\kappa}_0)^n} \exp\left(\hat{\kappa}_1(\|\mathbf{r}_+\| + \|\mathbf{r}_-\|) - \hat{\kappa}_0\|\mathbf{r}_0\|\right) \tag{5.9}$$

where $\mathbf{r}_+ = \sum_{\mathbf{p}\in R_+} I(\mathbf{p}), \mathbf{r}_- = \sum_{\mathbf{p}\in R_-} I(\mathbf{p}), \mathbf{r}_0 = \mathbf{r}_+ + \mathbf{r}_-$ and $\hat{\kappa}_1 = A_p^{-1}((\|\mathbf{r}_+\| + \|\mathbf{r}_-\|)/n), \hat{\kappa}_0 = A_p^{-1}(\|\mathbf{r}_0\|/n)$, $A_p(x) = I_{p/2}(x)/I_{p/2-1}(x)$. The optimization over $\psi$ is

$$\begin{aligned} \hat{\psi} &= \arg\max_{\psi} n(\log c_p(\hat{\kappa}_1) + \hat{\kappa}_1 A_p(\hat{\kappa}_1)) \\ &= \arg\max_{\psi} \|\sum_{\mathbf{p}\in R_+} I(\mathbf{p})\| + \|\sum_{\mathbf{p}\in R_-} I(\mathbf{p})\| \end{aligned} \tag{5.10}$$

The last equality comes from the fact that $n(\log c_p(x) + xA_p(x))$ and $A_p(x)$ are monotonically increasing functions of $x$. The proof of (5.10) is elaborated in Appendix D.

### 5.3.2 Hypothesis Test for Misalignment

For regions that pass the previous multi-modality test ($H_1$ declared in (5.6)), we perform a second hypothesis test to determine whether the multi-modal distribution is due to $H_0$: boundary misalignment, or $H_1$: a missing boundary. Since in most cases, misalignment causes only a small portion of pixels to differ from the majority, one naive test is to set a threshold on the ratio of the size of the smaller region to the whole region:

$$T = \frac{\min\{|R_+|, |R_-|\}}{|R|} \gtrless_{H_0}^{H_1} \eta \tag{5.11}$$

However, since region size can vary over several orders of magnitude, the same

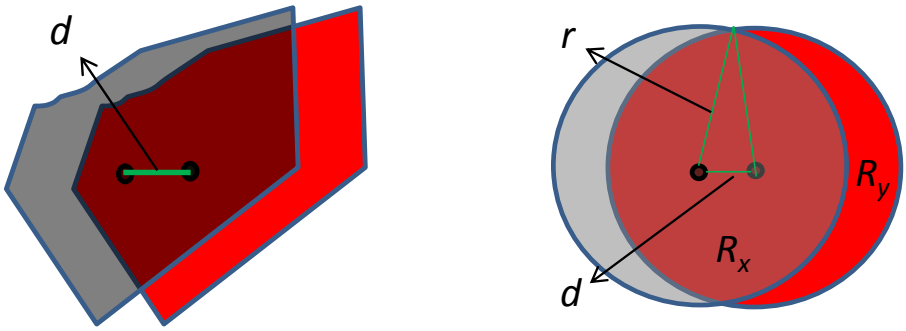



Figure 5.2: Misalignment caused by displacement for (a) a realistic region shape; (b) a simplified circular model.

absolute amount of misalignment (in pixels) can result in very different size ratios, making it hard to set a universal threshold. Therefore, we propose an adaptive threshold which incorporates region size. Boundary misalignment is modeled by a displacement in position (see Fig.5.2a), where the displacements $(d_x, d_y)$ are bivariate Gaussian with zero mean and covariance $\mathbf{\Sigma}_d = \sigma_d^2 \mathbf{I}_2$ and $\mathbf{I}_2$ is the $2 \times 2$ identity matrix. For simplicity, the region is modeled as a circle with radius $r$ (Fig. 5.2b), where $r = \sqrt{|R|/\pi}$ is the equivalent radius of region $R$. Based on these assumptions, the test statistic in (5.11) can be formulated as the following function of $d = \sqrt{x^2 + y^2}$ given $r$:

$$T = f_r(d) = 1 - \frac{2}{\pi} \arccos(\frac{d}{2r}) + \frac{d}{\pi r^2} \sqrt{r^2 - d^2/4}$$
$$\Rightarrow d = f_r^{-1}(T) \gtrless_{H_0}^{H_1} f_r^{-1}(\eta) = \eta'$$

The second line follows because $f_r$ is an increasing function. Since the displacement $d$ follows a Rayleigh$(\sigma_d)$ distribution, given the user specified false positive rate $\alpha$, we set

$$\alpha = P(d > \eta' | H_0) = Q(\eta') = P(T > f_r(\eta') | H_0)$$
$$\Rightarrow \eta = f_r(Q^{-1}(\alpha))$$

where $Q(.)$ is the Rayleigh tail distribution. As a result, the threshold is adaptively determined by $\alpha$ and the equivalent radius $r$.

## 5.4 Experiments

### 5.4.1 Boundary Detection Accuracy on Simulated Data

In this section, we compare grain boundary detection performance on simulated EBSD and BSE images using three different approaches: A. Segment the BSE and EBSD images separately by suitable existing segmentation algorithms (*Nock and Nielsen*, 2004; *Bachmann et al.*, 2011); B. Segment EBSD and register the boundaries onto BSE using a B-Spline deformation model and the mutual information criterion (*Maes et al.*, 1997; *Rueckert et al.*, 1999); C. The proposed coercive registration/segmentation algorithm with $\lambda = 0.15, \epsilon = 25, \alpha = 0.05$.

The grain shapes in the testing data are taken from real microscopy images downloaded from BlueQuartz (*BlueQuartz*, 2003) and segmented by their Dream3D toolbox. For each slice, some of the grains are randomly selected and displaced to produce boundary misalignment according to the Gaussian displacement model with $\sigma_d = 3$ (pixels). This creates the ground truth boundaries for evaluation. The pixel values for BSE and EBSD are generated from Gaussian and VMF distributions with random mean and variance/concentration for each grain region. More sophisticated methods which simulate more realistic grain shapes evolution and orientation distribution are worth further investigation (*Elsey et al.*, 2013; *Esedo Lu and Otto*, 2015).

To evaluate the boundary detection accuracy, we use the "overlapping rate". Let $B(w)$ be the set of boundary pixel locations with boundary width $w$, which is obtained by image dilation with filter disk radius $w/2$. The overlapping rate is defined as $O(w) = |B_T(w) \cap \hat{B}(w)|/|B_T(w) \cup \hat{B}(w)|$, where $B_T(w), \hat{B}(w)$ are the ground truth and estimated boundary.

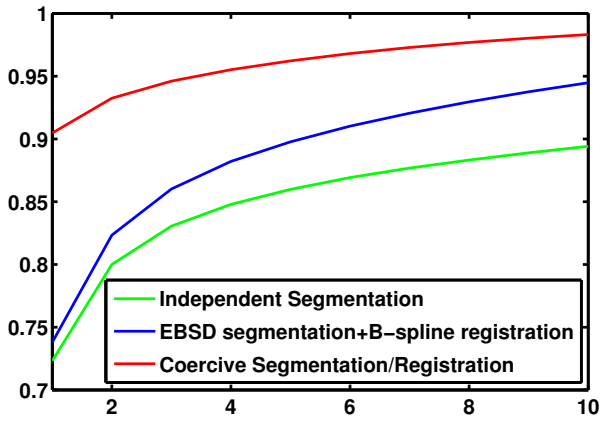



Figure 5.3: The proposed coercive approach (red line) has much higher boundary overlapping rate than other approaches since it is able to detect missing boundaries and register misaligned boundaries.

Figure 5.3 shows the overlapping rate of the three approaches for different boundary widths. Independent segmentation has the worst performance since it does not make use of shared sub-structure between modalities. With B-spline registration, there is some improvement but it is still not satisfactory, especially for small $w$. The proposed coercive registration approach with hypothesis testing is able to accurately register misaligned boundaries and detect missing edges. Therefore, it has much better boundary detection performance.

### 5.4.2 Results on Real Microscopy Data

We apply the proposed method to the IN100 data set which contains 170 slices of EBSD and BSE images of a Ni-base alloy. Figure 5.4 shows one registration/segmentation result overlaid on the BSE image. The red lines are the initial boundaries obtained by the EBSD segmentation and affine-transformed to match BSE. The blue lines are the realigned boundaries and the green lines are the missing boundaries detected by the hypothesis tests. The initial red lines are misaligned with the BSE image values but are corrected by our registration algorithm. Using statistical hypothesis tests, we

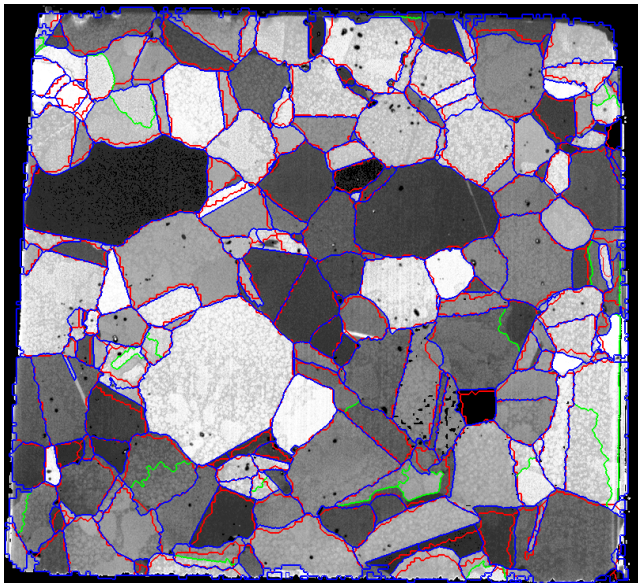


Figure 5.4: The registered boundaries (blue lines) fit the BSE image values much better than the initial boundaries from EBSD (red lines). The proposed approach is also able to detect and locate missing boundaries within grain regions (green lines).

are also able to detect and locate missing boundaries in some grain regions. These results in real data demonstrate that the proposed approach can accurately register boundaries and segment grain regions at the same time.

## 5.5 Conclusion

In this chapter, we proposed a coercive registration/segmentation algorithm for multi-modal images. The algorithm alternately utilizes information from one modality to help segment the image in the other modality, resulting in significant performance improvement in both modalities. The proposed hypothesis test based on statistical models of pixel values can accurately detect and locate missing boundaries between regions. Furthermore, our approach identifies and preserves all of the correspondences between regions in different modalities, which is important for fusing information after registration. The experiment results on simulated and real microscopy images show that our approach is able to effectively correct misaligned

grain boundaries and detect missing boundaries within grain regions.

# CHAPTER VI

# Multimodality Motion Compensated fMRI Using Slice-to-Volume Registration

In this chapter, we move our discussion from material microscopy image to biomedical image. We propose a multi-modal image registration approach based on particle filter to handle the head motion correction problem in functional Magnetic Resonance Imaging (fMRI). We demonstrate that this approach is able to accurately register the images to compensate for head motion, improving the following analysis and the activation detection result.

## 6.1 Introduction

Brain activation studies aim to identify specific regions in the brain that are associated with particular tasks. Detection of such functional regions is commonly performed by acquiring functional magnetic resonance imaging (fMRI) data using echo planar imaging (EPI) where the signal contrast is caused by the change of oxygenation in blood flow associated with local upstream neural activity. To detect brain activation in this noisy environment, one typically averages responses over several identical stimuli. Repeated scanning that is synchronous with the onset of the

*This chapter is based on (*Chen et al.*, 2016a,b)

required task (e.g., finger tapping, picture naming, etc.) is used to support signal averaging to improve the signal to noise ratio (SNR) for detecting the blood oxygen level de-saturation (BOLD) response (*Turner et al.*, 1998). By synchronously averaging the series of brain image volumes over the course of an fMRI study, the BOLD signal contrast can be significantly enhanced.

Ideally, each voxel in the volume time series records the signal evolving over time for a specific position. However, if the head of the subject moves during the scanning process, the time variation of voxel locations results in blurring or loss of signal and severe degradation of the fMRI image. This effect accumulates additional noise in the activation signal, impairing activity analysis accuracy. In experiments that require verbalized activation studies the head cannot be immobilized as the subject is required to speak during scanning. Therefore, some degree of head motion is inevitable even with cooperative subjects.

To deal with the above problem, the head motion can first be estimated and then used to correctly place fMRI image slices into the fMRI volume. Stereo optical tracking systems have been proposed to provide good real-time motion estimation with reasonable accuracy (*Zaitsev et al.*, 2006; *Qin et al.*, 2009). However, these systems require complicated and time-consuming system calibration. Other works use micro radio-frequency coils, called "active markers", for real-time prospective correction (*Ooi et al.*, 2009, 2011). Although such approaches can achieve good performance, they require additional equipment, incurring additional expense and adding complexity to the experimental protocol. Besides, there is a time lag between the actual instantaneous position of the subject's head and its computation from the sensors. Furthermore, the markers are mounted on the skin whose elasticity can introduce errors in head motion estimation.

In this chapter we take an image registration approach to head motion estimation, which does not require additional equipment except for a computer. We model the

head motion by rigid body motion and the motion is directly estimated from the parameters of a rigid body transformation that maps the target image into a reference image. Specifically, the motion parameters are estimated by optimizing pre-defined image similarity measures, e.g., cross-correlation or mutual information (*Maintz and Viergever*, 1998), between functional and anatomical reference images. In (*Friston et al.*, 1995), the head motion is estimated for each functional volume by registering the volumes to a reference volume. However, since the EPI images are taken slice by slice, stacking the slices directly and treating them as volumes neglects the head motion between consecutive slices within the same volume, i.e. inter-slice motion. Figure 6.1 shows the inter-slice motion with respect to the scanner caused by head nodding during the scan. Note that, in the interleaved acquisition (*Butts et al.*, 1994) as shown in the figure, the time interval between adjacent slices can be multiple of the nominal slice acquisition interval. The second figure from the right demonstrates the mismatch between the slice-stacked volume and the true human brain due to head motion. The most right figure shows how the volume is reconstructed by correcting the motion for each slice, which captures the brain activity signal more accurately.

Mapping-slice-to-volume (MSV)(*Kim et al.*, 1999) proposed by Kim et al. was the first work to address the slice-to-volume registration approach. As compared to volume-to-volume registration, the slice-to-volume approach is capable of estimating and correcting the head motion for each slice by more accurately following the EPI acquisition sequence slice by slice. However, the images at the top apex of the head have fewer image features, and are more prone to geometric distortions (*Schmitt et al.*, 1998) than the slices from the mid brain. This may negatively affect the performance of slice-to-volume registration approaches to motion estimation. A main disadvantage of the slice-to-volume approach is computational complexity as the image similarity measure may have many local maxima in the presence of noise and inadequate image features. As usual, choosing suitable initialization for the optimization process is

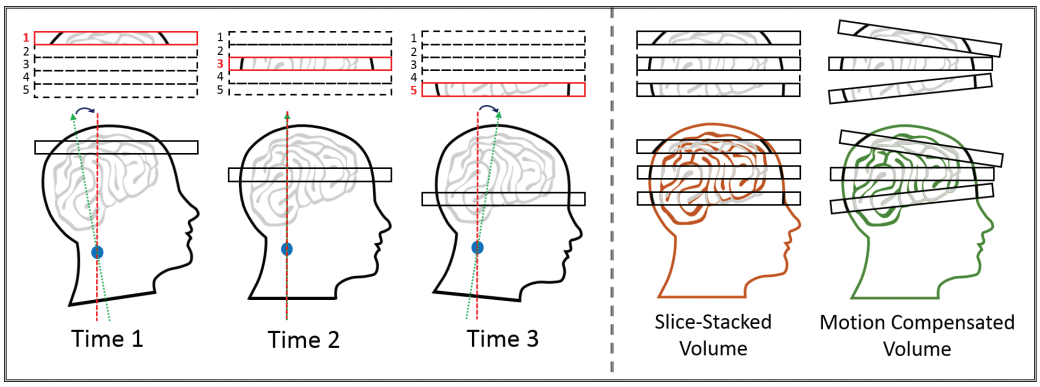



Figure 6.1: The inter-slice motion with respect to the scanner caused by head nodding during the scan. Note the interleaved acquisition (*Butts et al.*, 1994) shown in the figure that the time interval between adjacent slices is large compared to the slice acquisition interval. The second figure from the right demonstrates the mismatch between the slice-stacked volume and the true human brain due to head motion. The most right figure shows the motion corrected volume, which geometrically instantiates the original brain signal more accurately.

essential for accurate registration. While the focus of this work is fMRI we wish to acknowledge work focused on fetal anatomical imaging in utero by other authors (*Kim et al.*, 2010; *Kainz et al.*, 2015).

In this work, we propose a head motion tracking (HMT) algorithm based on a dynamic state space model (SSM) that tracks and estimates the head motion for each slice. The head motion parameters are modeled by a random walk, and the Gaussian particle filter (*Kotecha and Djuric*, 2003) is used to estimate the head motion given the observed sequence of EPI slices. The main advantage of our proposed approach is that it utilizes the information from previous acquired slices to provide a good starting point and effectively reduces the parameter search space in the optimization process, improving registration accuracy. The experimental results in Section 6.4 show that our approach outperforms other methods in terms of head motion parameter estimation, and in terms of activation detection accuracy for both synthetic and noisy real data.

This chapter is organized as follows. In Section 6.2, we review background of

the general image registration problem as well as the existing head motion correction methods. In Section 6.3, we describe our Head Motion Tracking (HMT) algorithm and how it is used to estimate the motion parameters. Section 6.4 shows the experimental results for synthetic and real data, and provides comprehensive comparisons between different approaches. Section 6.5 concludes this chapter.

## 6.2 Head Motion Estimation by Image Registration

The aim of image registration is to find a one-to-one transformation $T_{\boldsymbol{\theta}}$ that maps a reference image $I_R$ onto a target image $I_T$; The two images which may come from different imaging modalities. The transformation parameter $\boldsymbol{\theta}$ is found by optimizing an image similarity measure $\mathbb{M}(.)$ between the target image and the transformed image $T_{\boldsymbol{\theta}}(I_R)$ with respect to $\boldsymbol{\theta}$:

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \mathbb{M}(I_T, T_{\boldsymbol{\theta}}(I_R)), \tag{6.1}$$

where $T_{\boldsymbol{\theta}}(.)$ is the transformation function parameterized by $\boldsymbol{\theta}$. The parameterization of $T_{\boldsymbol{\theta}}$ could account for rigid body displacement (*Hill et al.*, 2001), local deformations (*Rueckert et al.*, 1999), or other relative differences between the reference and target image volumes (*Meyer et al.*, 1997; *Rohr et al.*, 2001). For head motion a rigid body displacement parameterization is adequate: $\boldsymbol{\theta} = [\alpha, \beta, \gamma, \delta x, \delta y, \delta z]$, where $\alpha, \beta, \gamma$ are spherical Euler angles and $\delta_x, \delta_y, \delta_z$ are spatial positions defining the origin of the spherical coordinate system.

The image similarity measure used in this chapter is the mutual information (MI), which has been widely applied to multi-modal biomedical image registration (*Maes et al.*, 1997). Mutual information between the images can be evaluated by first estimating the marginal and joint distributions $p(X), p(Y), p(X, Y)$ and then substituting

into:

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}, \tag{6.2}$$

where $X, Y$ are the random variables of the target and reference images' pixel intensity, respectively.

The image acquisition process starts by collecting an anatomical volume $V_{\text{anat}}$ of the subject's head using $T_1$-weighted MRI (*McRobbie et al.*, 2006), which serves as the reference $I_R$ for a functional MR image. The functional MR images are acquired via multislice single-shot echo-planar imaging (EPI) sequences acquired by $T_2^*$-weighted MRI, which has significantly lower spatial resolution than the $T_1$-weighted MRI. Let $\mathcal{V} = \{V_m\}_{m=1}^{M}$ denote the set of collected EPI volumes, where $M$ is the total number of volumes acquired during the brain scan session. Each of the EPI volumes is composed of a set of EPI slices $V_m = \{S_{mn}\}_{n=1}^{N}$, where $N$ is the number of slices per volume. The head motion is estimated by registering the set of EPI images $\mathcal{V}$ onto the anatomical volume $V_{\text{anat}}$. There are two main approaches that are commonly used to perform this multi-modality registration:

#### 6.2.0.1 Volume-to-volume Registration

Friston et al. (*Friston et al.*, 1995) proposed to estimate the head motion for each volume by registering the EPI images volume by volume via the optimization:

$$\hat{\boldsymbol{\theta}}_m = \arg\max_{\boldsymbol{\theta}} \mathbb{MI}(V_m, T_{\boldsymbol{\theta}}(V_{\text{anat}})). \tag{6.3}$$

The advantage of this approach is that the 3D volume contains abundant image features. However, since the EPI images are acquired slice by slice, this approach is not able to track significant movement occurring between each EPI slice. As EPI slices are commonly acquired in interleaved fashion, the typical time elapsed between adjacent slices can be as large as 1 second (*Turner et al.*, 1998; *Butts et al.*, 1994).

Therefore, inter-slice head motion can be significant.

#### 6.2.0.2 Slice-to-volume Registration

Slice-to-volume registration maps each individual slice into the anatomical reference volume space as proposed in (*Kim et al.*, 1999). The motion parameters are estimated for slices instead of volumes via the optimization:

$$\hat{\boldsymbol{\theta}}_{mn} = \arg\max_{\boldsymbol{\theta}} \mathbb{MII}(S_{mn}, T^*_{\boldsymbol{\theta}}(V_{\text{anat}})), \tag{6.4}$$

where $T^*_{\boldsymbol{\theta}}(.)$ is the function that interpolates the anatomical volume into 2D section with the motion parameter $\boldsymbol{\theta}$. This approach is capable of estimating and recovering the inter-slice head motion. However, because each 2D EPI slice $S_{mn}$ carries less information than the 3D volume $V_m$, using (6.4) can be sensitive to noise. Thus it is important to couple together the registration of successive EPI slices. The coupling of successive EPI slices in the registration process constitutes the main contribution of this chapter.

## 6.3 Head Motion Tracking

### 6.3.1 Coordinate Transformation

Our head motion tracking algorithm adopts the slice-to-volume approach to estimate the head motion for each EPI slice. As in (6.3) and (6.4) we formulate this problem as an optimization. We use a Gaussian particle filter to initialize and track the rigid body motion parameters across EPI slices. Let $\mathcal{S} = \{S_t\}_{t=1}^{T}$ denote the set of acquired EPI slices re-arranged in order of acquisition time, where $T = MN$ is the total number of slices in the experiment. Given the acquired EPI slices $\mathcal{S}$ and the anatomical volume $V_{\text{anat}}$, the aim of the tracking algorithm is to estimate the head motion parameters at each time $\{\boldsymbol{\theta}_t\}_{t=1}^{T}$. Since we model the head motion as a

rigid body transformation, the parameter $\boldsymbol{\theta}_t$ has six degrees of freedom and can be represented as a $3 \times 3$ rotation matrix $\mathbf{R}_t$ and a translation vector $\mathbf{q}_t$. Let $\mathbf{x}_r$, $\mathbf{x}_o$ denote the 3D-coordinates in the reference and observation coordinate systems. The conversion between the two coordinate systems can be described as:

$$(\mathbf{x}_r - \mathbf{c}) = \mathbf{R}_t((\mathbf{R}_s\mathbf{x}_o + \mathbf{q}_s) - \mathbf{c}) + \mathbf{q}_t, \tag{6.5}$$

where $\mathbf{R}_s, \mathbf{q}_s$ are fixed transformations introduced by coordinate mismatch between the two MRI scanners, e.g., due to initial head position difference, and $\mathbf{c}$ is the head rotation center that ideally corresponds to the location of the cervical vertebrae. Note that $\mathbf{R}_s, \mathbf{q}_s, \mathbf{c}$ are constant over time and only need to be estimated once in the whole experiment. The proposed method to estimate these parameters is discussed in Section 6.3.3.

### 6.3.2 Head Motion Tracking Algorithm

We use a state space model (SSM) (*Durbin and Koopman*, 2012) to describe the head motion, where $\boldsymbol{\theta}_t$ denotes the rigid body parameters at time $t$. The state equation is modeled using a Gaussian random walk with covariance matrix $\boldsymbol{\Sigma}_d$:

$$\boldsymbol{\theta}_{t+1} = \boldsymbol{\theta}_t + \mathbf{u}_t, \ \mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_d) \tag{6.6}$$

Note that our HMT algorithm can also be applied with more general head motion model, e.g., a kinematic model (*Han et al.*, 2009). The acquired EPI slices, called the observation in the sequel, is related to the state through the quasi-likelihood function:

$$p(\mathbf{S}_t|\boldsymbol{\theta}_t) = \frac{1}{Z}\mathbb{L}(\mathbb{M}(\mathbf{S}_t, T^*_{\boldsymbol{\theta}_t}(V_{\text{anat}}))), \tag{6.7}$$

where $\mathbb{L}(.)$ can be chosen as any function such that it is positive and monotonically increasing (i.e. $\mathbb{L}(x) \geq 0, \ \forall -\infty < x < \infty, x > y \Rightarrow \mathbb{L}(x) > \mathbb{L}(y)$) and $Z$ is

a normalization coefficient that turns the objective function $\mathbb{L}(.)$ into a conditional probability, which is denoted $p(\mathbf{S}_t|\boldsymbol{\theta}_t)$ and is called the quasi-likelihood function of $\boldsymbol{\theta}_t$. Here $\mathbf{S}_t = \{S_j\}_{j=t-h}^{t+h}$ denotes the stack of slices over a length $2h+1$ time interval centered at time $t$. If $h = 0$, $\mathbf{S}_t$ is reduced to a single EPI slice $S_t$; If $h = N/2$, then the optimization becomes similar to the usual volume-to-volume registration. The parameter $h$ controls the trade-off between parameter estimator bias and variance. In the analysis reported below we have used $h = 1$, which was found sufficient to incorporate enough features to achieve robustness without too much smoothing and biasing the estimation result.

The Kalman Filter (*Kalman*, 1960) is the optimal minimum mean squared error estimator for a linear SSM when both the state dynamics and the measurement equations are linear in the state vector and the driving noise vectors. In non-linear cases, one has to resort to some form of approximation to the minimum mean squared error estimator, such as the extended Kalman filter (EKF)(*Julier and Uhlmann*, 1997) or the unscented Kalman filter (UKF)(*Wan and Van der Merwe*, 2000). These approaches require explicit state and observation equations, which are not readily available in the fMRI problem treated here. Alternatively, one can approximate the posterior distribution of the state using sequential importance sampling, i.e., the particle filter (*Doucet et al.*, 2000).

The Gaussian particle filter (GPF)(*Kotecha and Djuric*, 2003) approximates the posterior using a set of weighted samples, known as particles, and uses importance sampling and Monte-Carlo integration methods to approximate the state and observation distributions. The main advantage of GPF compared to other particle filtering approaches is its lower computational complexity and amenability to parallel implementation. In GPF algorithm, the posterior at time $t$ is approximated by a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$, and then resampling follows by drawing $P$ particles from the Gaussian distribution. The particles are weighted according to the observation and

are used to form the distribution for the next time step.

Our Head Motion Tracking (HMT) algorithm is based on the GPF framework which utilizes the posterior distribution to provide a good initialization to the optimization process. The algorithm is summarized in Algorithm *HMT* (page 109). Initially slice-to-volume registration (*Kim et al.*, 1999) is used to generate an initial head motion estimate $\hat{\boldsymbol{\theta}}_0$. As in the GPF, for each slice at time $t$, the algorithm has two stages: *Measurement update* and *Time update*. In the *Measurement update* stage, we use $P$ particles $\{\boldsymbol{\theta}_t^{(j)}\}_{j=1}^{P}$ drawn at the last time step to evaluate the particle weights using the quasi-likelihood function $p(\mathbf{S}_t|\boldsymbol{\theta}_t)$ defined in (6.7). The quasi-likelihood function should have two properties: (1) It is monotonically increasing with the image similarity $\mathbb{M}(\mathbf{S}_t, T_{\boldsymbol{\theta}}^*(V_{\text{anat}}))$; (2) The weighted particles are approximately distributed according to a multivariate Gaussian density. To satisfy the two properties, we propose to use a histogram equalization approach to evaluate the particle weights. The target density is the distribution of $z = f(\mathbf{x})$ where $\mathbf{x}$ and $f(.)$ are the 6-dimension multivariate Gaussian random variable and density, respectively. Letting $g_Z(z)$ denote the density of $z$, we can equalize the histogram to obtain the particle weights.

$$g_Z(z) = \pi^3 \left(-2 \log (2\pi)^3 z\right)^2, \ z \in (0, (2\pi)^{-3}]. \tag{6.8}$$

The detailed derivation of (6.8) is given in Appendix E. The particle weights are normalized to sum to 1 and then used to calculate the weighted mean and covariance. Since the weighted mean incorporates abundant information about the image similarity distribution in neighboring regions, it is a good starting point for the optimizer. In this chapter, we use the Nelder-Mead (*Nelder and Mead*, 1965) optimizer to perform the maximization:

$$\hat{\boldsymbol{\theta}}_t = \arg\max_{\boldsymbol{\theta}} \ \mathbb{M}(\mathbf{S}_t, T_{\boldsymbol{\theta}}^*(V_{\text{anat}})). \tag{6.9}$$

Nelder-Mead is a simplex method used to iteratively find the optimum of an

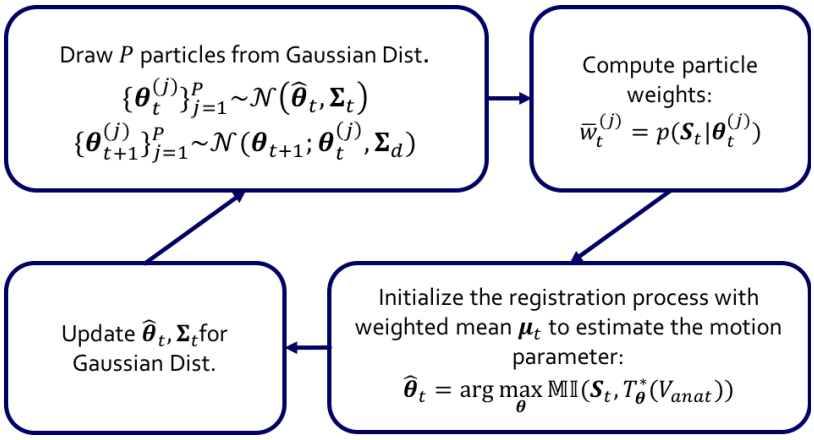



Figure 6.2: The state and time update for each iteration in the proposed HMT algorithm.

objective function in a multi-dimensional space. Note that the proposed histogram equalization approach is not restricted to any particular definition of image similarity. Therefore MI can be replaced by any other image similarity measure, e.g., Normalized MI (*Studholme et al.*, 1999), localized MI (*Klein et al.*, 2008), graph-based MI (*Staring et al.*, 2009), or feature-based measures (*Oliveira and Tavares*, 2014)...etc. The transformation parameter $\hat{\boldsymbol{\theta}}_t$ that maximizes (6.9) is the estimated head motion at time $t$. After the motion parameter is estimated, we perform a standard re-sampling step to estimate the covariance matrix of the posterior distribution, which is then used to establish the prior distribution of the next slice in the *Time Update* stage using (6.6). The diagram of each iteration in the proposed HMT algorithm is summarized in Fig. 6.2.

Often the MRI acquired images are very noisy and difficult to register, especially for slices near the lower and upper apex of the head. Figure 6.3 shows an example of the images of the middle head (a) and top apex (b). We can see that the top apex head image has much less information content than the middle head that can be used for registration. To reduce the effect of these noisy slices, we screen the slices for adequate signal strength. Specifically, we reject all EPI slices for which fewer than

15% of the pixels are above a certain threshold value. For these rejected slices, we skip the optimization step and estimate the motion parameters through interpolation of the estimates from neighboring slices. We use 2nd-order interpolation, which is accurate when the head motion has approximately constant angular and translational accelerations (*Park et al.*, 2004).

### 6.3.3 System Parameters Setting

In the proposed Head Motion Tracking algorithm there are several parameters that need to be set: $\mathbf{R}_s, \mathbf{q}_s, \mathbf{c}, \mathbf{\Sigma}_d$:

#### 6.3.3.1 Fixed Coordinate Transformation $\mathbf{R}_s, \mathbf{q}_s$

Since $\mathbf{R}_s, \mathbf{q}_s$ are constant over the entire experiment, they can be estimated by first taking the average of all EPI volumes over time, and then registering the averaged EPI volume to the anatomical volume.

#### 6.3.3.2 Head Rotation Center c

To estimate the head rotation center, we run the HMT algorithm on the first $K$ image slices (we used $K = 70$ in our experiment) by assuming $\mathbf{c} = \mathbf{0}$ as the origin. Let $\{\hat{\boldsymbol{\theta}}_t\}_{t=1}^K$ denote the estimate of the motion parameters for these $K$ image slices. Here we assume that the patient's body position is stable during the scan (the subject is often immobilized and lying in the machine) and therefore the amount of translation should be small, i.e. $\|\mathbf{q}_t\| \approx 0$. Based on this assumption, the rotation center can be estimated by solving the least squares problem:

$$\hat{\mathbf{c}} = \arg\min_{\mathbf{c}} \sum_{t=1}^{K} \|\mathbf{q}_t - (\mathbf{I}_3 - \mathbf{R}_t)\mathbf{c}\|_2^2, \tag{6.10}$$

where $\mathbf{I}_3$ is the $3 \times 3$ identity matrix.

#### 6.3.3.3 Head Motion Covariance $\mathbf{\Sigma}_d$

The estimate of the head motion covariance matrix is generated in two steps. We initially set $\mathbf{\Sigma}_d$ to the identity matrix (rotation in degree and translation in $mm$) and run the HMT algorithm over $K$ image slices to obtain the estimates $\{\hat{\boldsymbol{\theta}}_t\}_{t=1}^K$. Subsequently, the matrix $\mathbf{\Sigma}_d$ is estimated as the covariance matrix of the consecutive parameter differences:

$$\begin{aligned} \hat{\mathbf{\Sigma}}_d &= Cov(\hat{\boldsymbol{\theta}}_t - \hat{\boldsymbol{\theta}}_{t-1}) \\ &= \frac{1}{K-1}\sum_{t=2}^{K}(\hat{\boldsymbol{\theta}}_t - \hat{\boldsymbol{\theta}}_{t-1})(\hat{\boldsymbol{\theta}}_t - \hat{\boldsymbol{\theta}}_{t-1})^T \end{aligned} \tag{6.11}$$

**Algorithm** *HMT*

**Input:** EPI slices $\{S_t\}_{t=1}^T$ and anatomical volume $V_{\text{anat}}$

1. Estimate the parameters for the first slice $\hat{\boldsymbol{\theta}}_0$ using slice-to-volume registration.
2. Draw $P$ particles $\{\boldsymbol{\theta}_0^{(j)}\}_{j=1}^P$ from $\mathcal{N}(\hat{\boldsymbol{\theta}}_0, \mathbf{\Sigma}_d)$.
3. **for** $t \leftarrow 1$ **to** $T$
4. ($*$ *Measurement update* $*$)
5. Equalize the histogram of $\mathbb{M}(\mathbf{S}_t, T^*_{\boldsymbol{\theta}_t^{(j)}}(V_{\text{anat}}))$ to (6.8) to get $\bar{w}_t^{(j)}$ and then normalize to sum to 1

$$w_t^{(j)} = \bar{w}_t^{(j)} / \sum_{j=1}^{P} \bar{w}_t^{(j)}$$

6. Estimate the sample mean and covariance

$$\begin{aligned} \boldsymbol{\mu}_t &= \sum_{j=1}^{P} w_t^{(j)} \boldsymbol{\theta}_t^{(j)} \\ \mathbf{\Sigma}_t &= \sum_{j=1}^{P} w_t^{(j)} (\boldsymbol{\theta}_t^{(j)} - \boldsymbol{\mu}_t)(\boldsymbol{\theta}_t^{(j)} - \boldsymbol{\mu}_t)^T \end{aligned}$$

7. Initialize the registration process with $\boldsymbol{\mu}_t$ to estimate the motion param-

eter:

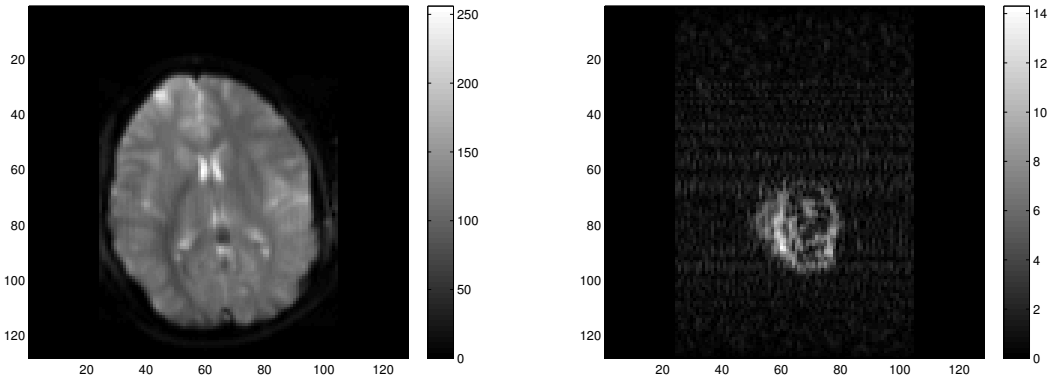
$$\hat{\boldsymbol{\theta}}_t = \arg\max_{\boldsymbol{\theta}} \mathbb{M}(\mathbf{S}_t, T^*_{\boldsymbol{\theta}}(V_{\text{anat}}))$$

8. ($*$ *Time update* $*$)

9. Draw samples $\{\boldsymbol{\theta}_t^{(j)}\}_{j=1}^P$ from $\mathcal{N}(\boldsymbol{\theta}_t, \boldsymbol{\Sigma}_t)$.

10. For $j = 1, ..., P$, sample from $p(\boldsymbol{\theta}_{t+1}|\boldsymbol{\theta}_t = \boldsymbol{\theta}_t^{(j)})$ to obtain $\{\boldsymbol{\theta}_{t+1}^{(j)}\}_{j=1}^P$.

11. **return** $\{\hat{\boldsymbol{\theta}}_t\}_{t=1}^T$



(a) middle slice of head

(b) top apex slice of head

Figure 6.3: The middle head (a) and top apex (b) of the real human data are shown in gray scale. Notice that the top apex image has very little useful features and the signal strength (pixel intensity) is much lower than the middle head image.

## 6.4 Experimental Results

### 6.4.1 Synthetic Data Generation

We downloaded high resolution $T_1$, $T_2$-weighted MRI volumes from the International Consortium of Brain Mapping (ICBM) (*Cocosco et al.*, 1997). The high resolution $T_1$ MRI brain volume was used as the anatomical reference volume with voxel size $0.78 \times 0.78 \times 1.5mm^3$. The EPI slices were emulated by interpolating the $T_2$-weighted volume under artificial motion induced by applying a sequence of trans-

formations to the image with smoothly varying motion parameters as used in (*Kim et al.*, 1999). The voxel size of the EPI slices is $1.56 \times 1.56 \times 6mm^3$, a blurring Gaussian low-pass kernel with $\sigma = 2$ was applied, and 3% Gaussian noise was added to simulate real EPI slices. The activation signal was introduced by adding 5% intensity to selected voxels in manually drawn regions of interest at various locations in the volume as in (*Kim et al.*, 2008). This produced a synthetic EPI data set consisting of $M = 120$ volumes with $N = 14$ slices per volume. Figure 6.4(a) shows the ground truth motion parameter of the three rotation angles (in degree) from slice 1 to 200. The simulated time series in a block design paradigm consists of 120 volumes with 6 activation cycles. There are 20 volumes per cycle which contains 10 stimulation and 10 control volumes.

### 6.4.2 Performance Measures

In the following comparison, we evaluate the performance quantitatively with respect to misregistration error, activation detection accuracy and reliability:

#### 6.4.2.1 Average Voxel Distance

The misregistration error is measured by *average voxel distance*, which is the average distance between the registered voxel coordinate and the true voxel coordinate. Let $\mathbf{x}_t^{reg}(i)$ and $\mathbf{x}_t^{true}(i)$ denote the coordinates of voxel $i$ transformed using the estimated motion parameter $\hat{\boldsymbol{\theta}}_t$ and true motion parameter $\boldsymbol{\theta}_t$ of slice $t$. The average voxel distance is defined as:

$$D_t = \frac{1}{N_v} \sum_{i=1}^{N_v} \|\mathbf{x}_t^{reg}(i) - \mathbf{x}_t^{true}(i)\|, \tag{6.12}$$

where $N_v$ is the total number of voxels in a single EPI slice.

#### 6.4.2.2 Activation Detection ROC Curve

The estimated motion parameters $\{\hat{\boldsymbol{\theta}}_t\}_{t=1}^{T}$ are used to reconstruct the motion corrected EPI volumes $\tilde{\mathcal{V}} = \{\tilde{V}_m\}_{m=1}^{M}$. To identify the activated brain region, the non-parametric random permutation test(*Nichols and Holmes*, 2002) is performed on the intensities in the EPI volumes. Let $\{u_m(i)\}_{m=1}^{M}$ to be the set of intensities for voxel $i$ of the $M$ reconstructed volumes. The null hypothesis $H_0$ of the activation test is: "The mean of the voxel intensities under each of the conditions, stimulation or control, are equal." Under this hypothesis, any re-ordering of $\{u_m(i)\}_{m=1}^{M}$ should give the same statistic, which we used the two-sample $t$-test statistic. Let $N_r$ denote the number of re-ordering, $t_j$ be the two-sample $t$-test statistic corresponding to ordering $j$ and $\tilde{t}$ be the statistic of actual ordering. The $P$-value is then calculated by counting the proportion of the test statistics $\{t_j\}_{j=1}^{N_r}$ which are more extreme than $\tilde{t}$. By taking a threshold on the $P$-value, we can determine which voxels are activated in this experiment. In this work, we set $N_r$ equal to 2000 and the threshold for $P$-value is 0.5%. When a ground truth activation map is available as in the synthetic data, the detection performance can be evaluated by the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC).

#### 6.4.2.3 Activation Detection Reliability

We use the Activation Test-retest Reliability (ATR) measure (*Noll et al.*, 1997; *Genovese et al.*, 1997) to compare the performance when the ground truth of motion parameters and activation map are unknown. This approach assumes that each voxel is either truly active or truly inactive. We use the random permutation test with two-sample $t$-test statistic to generate the activation maps as describe in Section 6.4.2.2. The reliability of the test is measured in terms of true active and false active probability, $p_A = p(\{v \text{ is classified as active}\}|\{v \text{ is truly active}\})$ and $p_I = p(\{v \text{ is classified as active}\}|\{v \text{ is truly inactive}\})$, respectively. Ideally, $p_A$

should be 1 and $p_I$ should be 0. Therefore, higher $p_A$ and lower $p_I$ indicate more reliable testing result.

To estimate $p_A, p_I$, we need to replicate the fMRI experiments $L$ times, where $L \geq 3$. In this work, we obtain the replications by splitting the acquired volumes into $L = 4$ disjoint sets randomly as suggested in (*Noll et al.*, 1997) to ensure statistical independence accross voxels and replications. We use the random permutation test to generate $L$ activation maps for each of the sets. Let $r(i) \in \{0, 1, ..., L\}$ be the number of replications out of $L$ in which the voxel $i$ is classified active. We model $r(i)$ as a mixture of two binomial distributions:

$$\lambda \mathbf{B}(L, p_A) + (1 - \lambda)\mathbf{B}(L, p_I), \tag{6.13}$$

where $\mathbf{B}$ is the binomial distribution and $\lambda$ represents the proportion of truly active voxels. We estimated the parameters by maximizing the likelihood function.

### 6.4.3 Evaluation Using Synthetic Data

The simulated EPI slices described in Section 6.4.1 are registered to the anatomical volume to estimate the motion parameters by using the following three methods (implemented in MATLAB R2015a): (1) volume-to-volume registration (*Friston et al.*, 1995) (V2V); (2) slice-to-volume registration (*Kim et al.*, 1999) (S2V), where the optimization process is initialized by the V2V result; (3) the proposed Head Motion Tracking algorithm (HMT) with $P = 4000$ particles. Figures 6.4(b)-(d) show the estimated motion parameters for the first 200 slices, where the black solid lines denote ground truth and the color dashed lines denote estimated motion parameters. Figure 6.4(b) demonstrates that the volume-to-volume registration method can accurately estimate motion for each volume but cannot accurately track the motion over the slices. On the other hand, S2V (Fig. 6.4(c)) can better track the head mo-

tion over different slices but has high bias, especially for slices near the apex of the head where slice image intensity and contrast are low. Our proposed HMT algorithm (Fig. 6.4(d)) is able to track the head motion much more accurately than the other two approaches. Figure 6.5(a) shows the boxplot of the average voxel distance after registration for different methods. The whiskers are the outliers outside the inner fence (defined by $1.5 \times F$-spread (*Hoaglin et al.*, 1986)). All of these methods reduced a fair amount of the voxel misregistration errors compared to no motion correction case (NoCorr). Notice that our HMT algorithm has significantly lower misregistration error, as measured by voxel distance, and is much more stable (fewer outliers) than the other methods. The mean of $D_t$ over all slices are listed in the first column of Table 6.1.

The estimated parameters are used to reconstruct the motion corrected EPI volumes, and activated voxels are identified by the random permutation test. The ROC curves of the activation detection result of different approaches are compared in Fig. 6.5(b). Note that the volumes that are reconstructed using ground truth motion parameters achieve perfect detection (red solid line). Again, our HMT algorithm (blue dashed line) outperforms other methods and is closest to the ground truth. The Area under Curve (AUC) for each approach is listed in the second column of Table 6.1. The comparison of activation detection reliability is listed in the first two columns in Table 6.2. It can be seen that all of the three methods have similar $p_I$, but the proposed HMT has significantly higher $p_A$ than the other two methods.

### 6.4.4 Evaluation Using Real Data

We further validate the performance of the proposed HMT algorithm on real fMRI experimental data. We used two datasets that are denoted "Run1" and "Run2", and that were acquired from two normal volunteers. The study was approved by the Institutional Review Board at the University of Michigan Medical School and informed

| | Avg. $D_t$ | AUC |
|---|---|---|
| Truth | 0.000 | 1.000 |
| No Corr. | 4.497 | 0.732 |
| V2V | 2.426 | 0.855 |
| S2V | 1.225 | 0.924 |
| HMT | **0.393** | **0.953** |

Table 6.1: Estimation and Activation Result Comparison: As compared to the other motion compensation algorithms (No Corr., V2V, S2V), the proposed HMT algorithm attains lower average misregistration error $D_t$ and better Area Under the Curve (AUC) detection performance.

consent was obtained from each subject prior to participation. The subjects performed a simple motor task, uni-lateral sequential finger tapping, in the experiment. We asked the subject to do their best to minimize head motion for Run1 dataset and asked the subject to intentionally nod his head for Run2 dataset. The head was scanned 126 times with 14 slices in each volume for these two datasets. The anatomical voxel size is $1 \times 1 \times 1.5mm^3$ and the EPI voxel size is $2 \times 2 \times 6mm^3$.

Figure 6.6 shows the three Euler angles estimated by S2V (first column color dashed lines) and HMT (second column color dashed lines) overlaid with the V2V result (black solid lines) for the first 200 slices. Notice that the estimated rotation in Run2 (second row) is larger than Run1 (first row), which matches our expectations given the experimental protocol. Similarly to the experiments with synthetic data, reported in Section 6.4.3, S2V can be used to estimate the motion for each slice but is noisy. The abrupt changes in the motion parameters demonstrated by S2V represent unlikely head movement, which suggests incorrect estimation. On the other hand, the proposed HMT algorithm produced much more stable and smoother motion estimates, which more accurately reflects real head motion. The superior tracking performance of HMT is a consequence of the dynamical modeling that couples together estimates from successive slices leading to smoother and less noisy tracking performance.

The improvement in the head tracking translates into better activation detection

performance, Fig 6.7 shows colorized activation maps overlaid on the anatomical MRI, which is used as an additional reference volume for registration. These selected slices (denoted as slice A, B, and C) displayed in different rows, show representative activated regions. Significant voxels are marked in red and blue to indicate the temporal positive and negative correlations, respectively.

Figure 6.7(a) shows the activation maps for V2V, S2V, and the proposed HMT algorithms applied to the Run1 dataset. For this easier dataset (less head motion), we can see that all methods are able to produce active regions that are near the motor cortex related to finger moves (*Beisteiner et al.*, 2001). However, the volume-based (first column) approach produced much more spread out active regions, which may be due to small amounts of head motion. S2V (second column) did produce more clustered active regions, however, it also has some active voxels which are scattered in the white matter and are therefore likely to be false positive detections. Our proposed HMT (third column) generated active regions along the gray matter and has the least false positive voxels in the white matter. For the more challenging Run2 dataset (larger head motion), shown in Fig. 6.7(b), the activation maps of V2V and S2V (left two columns) have very few active voxels that are scattered across the volume. In contrast, the proposed HMT algorithm (third column) produced clean and well clustered active regions on the gray matter, which are more likely to correspond to real brain activity responses. A quantitative measure of the activation detection reliability is summarized in Table 6.2. We can see that the three methods have the same level of $p_I$ values but HMT has significantly higher $p_A$, especially for the harder Run2 dataset.

## 6.5 Conclusion

In this work, we have proposed a head motion tracking (HMT) algorithm that uses an image registration objective function combined with a Gaussian particle filter

| Method | Simulated | | Run1 | | Run2 | |
|---|---|---|---|---|---|---|
| | $p_A$ | $p_I$ | $p_A$ | $p_I$ | $p_A$ | $p_I$ |
| Truth | 1.000 | 0.000 | - | - | - | - |
| V2V | 0.128 | 0.003 | 0.521 | 0.003 | 0.047 | 0.001 |
| S2V | 0.248 | 0.002 | 0.614 | 0.003 | 0.048 | 0.001 |
| HMT | **0.662** | 0.003 | **0.623** | 0.003 | **0.087** | 0.002 |

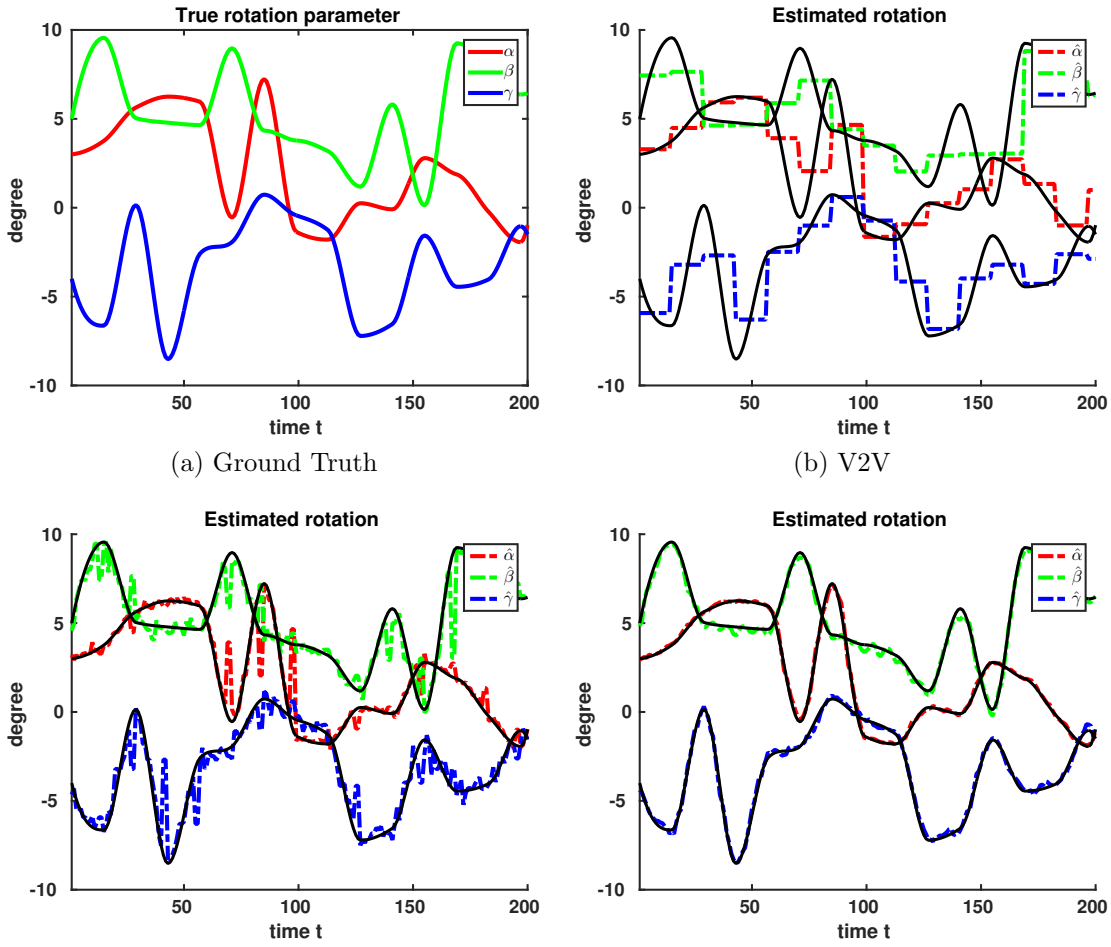
Table 6.2: Activation Detection Reliability: The proposed HMT algorithm attains significantly higher $p_A$, especially for Run2 dataset, while keeps the same level of $p_I$ compared to the other motion compensation algorithms (V2V, S2V).

to couple motion estimates from successive EPI slices, resulting in improved performance. Due to the fact that the proposed algorithm utilizes the information from consecutive slices in the fMRI scan volume, it effectively combines the bias reduction properties of the S2V approach and the variance reduction properties of the V2V approach.

Evaluation based on synthetic data demonstrated that the proposed HMT algorithm can significantly improve accuracy over the volume-to-volume and slice-to-volume approaches in terms of motion parameter estimation and activation detection accuracy. Using real human experimental data we demonstrated that the proposed algorithm is able to produce more stable estimates of head motion and brain activation maps. Unlike previous approaches to head motion compensation, the activation maps of the HMT produce more reliable active regions even when the head motion is large during the fMRI scan.

Improvements in robustness and accuracy of the proposed HMT algorithm may permit scientists to analyze more complex brain activation patterns. This can be especially beneficial for experiments that involve a wider spatial distribution activation regions, and are more likely to have motion artifacts, e.g., in working memory or speech experiments. Furthermore, our HMT approach might allow fMRI to be reliably applied to patients having significant motion disorders, e.g., Parkinson's disease,

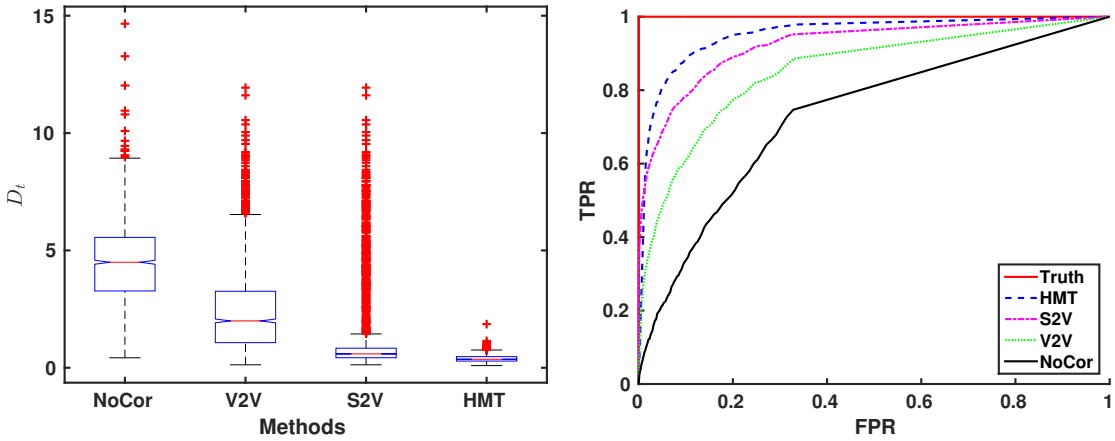
who currently do not benefit from fMRI examinations.




(a) Ground Truth (b) V2V

(c) S2V (d) HMT

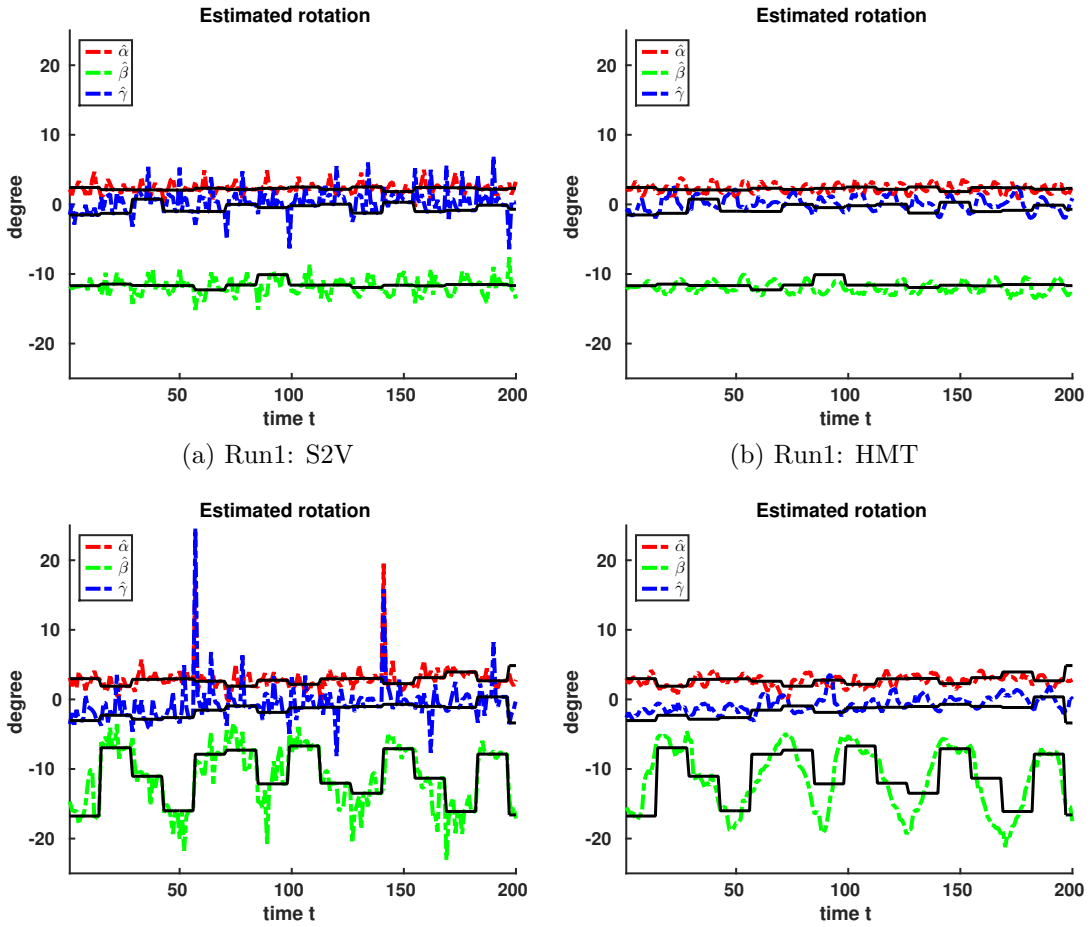
Figure 6.4: (a) shows the ground truth of head motion parameters in three Euler angles for the first 200 slices. (b)(c)(d) show the motion parameters estimated by volume-to-volume (V2V), slice-to-volume (S2V) and the proposed head motion tracking (HMT) algorithm. The black solid lines are the ground truth and the color dashed lines are the estimated motion parameters. (b) demonstrates that the volume-to-volume registration method can accurately track the average motion for each volume but does not accurately track motion for each slice in the volume. S2V (c) can estimate the head motion for each slice but suffers from large tracking errors. The proposed HMT algorithm (d) is able to track the head motion much accurately than the other two approaches.




(a) Avg. Voxel Distance

(b) Activation ROC Curve

Figure 6.5: (a) is the boxplot of the average voxel distance after registration for different methods. The whiskers are the outliers outside the inner fence (defined by $1.5 \times F$-spread). The proposed HMT algorithm has significantly lower voxel misregistration errors and is more stable (fewer outliers) than the other methods. (b) shows the ROC curves for activation detection. Note that the volumes that are reconstructed using ground truth motion parameters achieve perfect detection (red solid line). Our proposed HMT algorithm (blue dashed line) outperforms other methods (S2V, V2V, No Correction) and is closest to the ground truth.



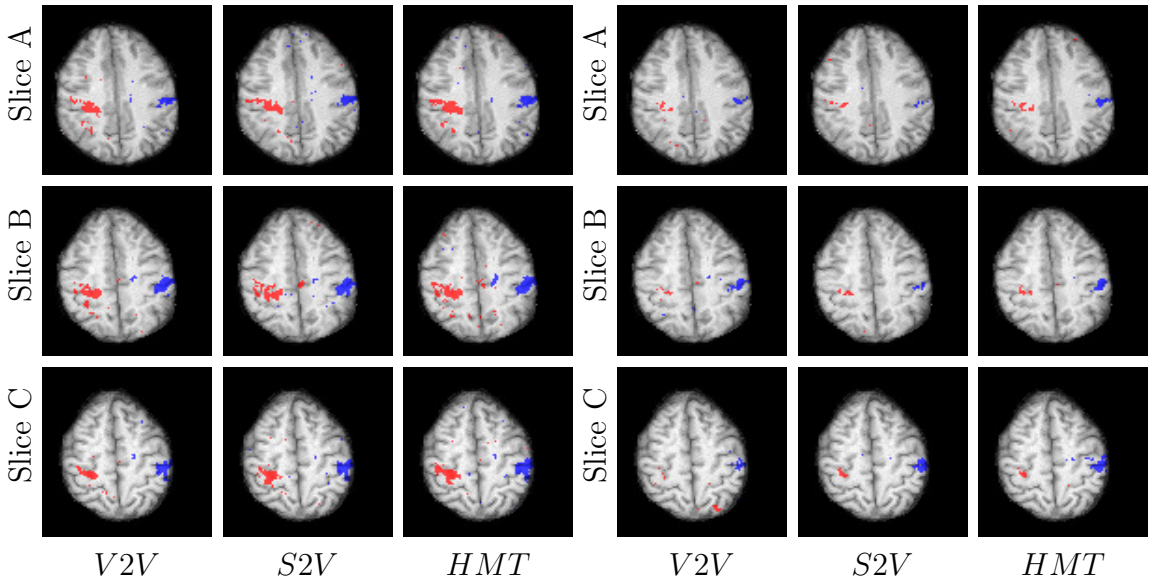

(a) Run1: S2V

(b) Run1: HMT

(c) Run2: S2V

(d) Run2: HMT

Figure 6.6: The three Euler angles estimated by S2V (first column color dashed lines) and HMT (second column color dashed lines) overlaid with V2V result (black solid lines) for the first 200 slices. Note that the estimated rotation in Run2 (second row) is larger than that of Run1 (first row). Similarly to the experiments with synthetic data summarized in Fig. 6.4, S2V can estimate the motion for each slice but is noisy. The proposed HMT algorithm produces more stable and continuous head motion estimates which is more convincing in describing real head motion.




(a) Run1 Activation Map (b) Run2 Activation Map

Figure 6.7: The colorized activation maps overlaid on the anatomical MRI images for Run1 (a) and Run2 (b) datasets. The results of the three methods: (1) V2V registration; (2) S2V registration; (3) proposed HMT algorithm are listed in order from left to right column. In (a), we can see that the V2V (first column) approach produced a more dispersed set of active regions due to the inter-slice head motion. S2V (second column) produced more clustered active regions but has lots of false positive voxels scattered in the white matter. The proposed HMT (third column) generated the least dispersed active regions and had the least false positive voxels in the white matter. In (b), the activation maps from V2V and S2V (left two columns) had few and scattered active voxels due to the effect of head motion. The proposed HMT (third column) produced clean and well clustered active regions.

# CHAPTER VII

# Conclusion and Future Work

## 7.1 Conclusion

Image fusion provides comprehensive information which advances the analysis and characterization of the image content in many different areas. This thesis focused on multi-modal image registration and fusion problems in two main domains: *I. materials microscopy* and *biomedical imaging*. For each of the domains, we have proposed a novel and effective algorithm and provided numerical experiments on both simulated and real world data to demonstrate the performance improvement.

In Chapter II, we have reviewed background of image fusion, described the advantages and difficulties of image fusion, and introduced different types of fusion frameworks. We have motivated and elaborated our proposed approaches of image fusion in material microscopy and biomedical images.

In Chapter III, we have presented a general finite mixture representation for densities on domains whose topologies have group invariances. An efficient EM algorithm was derived for estimation of parameters of this mixture model. We have further extended the model to consider the multi-cluster case where the samples are from several group-invariant distributions with different parameters. We have demonstrated the effectiveness of this model in both simulated and real dataset. This model forms an important component in the following EBSD indexing, region-level registration and

uncertainty quantification.

In Chapter IV, we have presented a novel method for indexing polycrystalline materials that uses both mathematical-physics modeling and mathematical-statistics modeling introduced in Chapter III. The dictionary generated from the physics-based forward model is fixed for each crystal symmetry group and each SEM instrument configuration. A feature of this method is that it performs classification, segmentation, and indexing in the unified framework of dictionary matching. A feature of the indexing method is that it incorporates a concentration parameter that can be estimated jointly with the orientations of a pixel or of a grain. This concentration parameter can be used to report the degree of confidence one can have in the orientation estimates. The segmentation result can further be utilized as the input of our proposed coercive region-level registration.

In Chapter V, we have proposed a coercive registration/segmentation algorithm for multi-modal images. The algorithm alternately utilizes information from one modality to help segment the image in the other modality, resulting in significant performance improvement in both modalities. The proposed hypothesis test based on statistical models of pixel values can accurately detect and locate missing boundaries between regions. Furthermore, our approach identifies and preserves all of the correspondences between regions in different modalities, which is important for fusing information after registration.

In Chapter VI, we have proposed a head motion tracking (HMT) algorithm that uses an image registration objective function and the Gaussian particle filter to couple motion estimates from successive EPI slices. With the help of the information from multi-modal images and the prior knowledge from consecutive observations, our proposed method combines both the bias reduction properties of the S2V approach and the variance reduction properties of the V2V approach, resulting in significantly improved performance. The robustness and accuracy of the proposed HMT algorithm

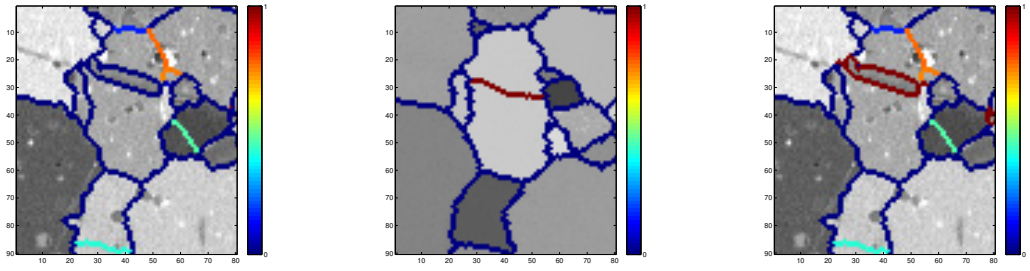
can allow scientists to analyze more complex brain activation patterns. Especially for experiments which may involve a wider spatial distribution of functional signal, and are more likely to have motion artifacts, e.g., working memory or speech.

## 7.2 Future Work

There are many interesting directions that are worthy of future study:

First, since we have finished the works of anomaly-driven fusion framework based on the statistical model to perform indexing, segmentation and anomaly-level evaluation, a future direction is to focus on quantifying and fusing the anomaly-level with $P$-values. Figure 7.1 shows an example of the boundary $P$-value of BSE and EBSD images. The boundary $P$-value of BSE image is calculated by the grain-level Gaussian assumption with the hypothesis test on mean difference between neighboring grains. Lower $P$-value exhibits that the existence of the boundary is significant while higher $P$-value shows that the boundary does not have obvious mean difference. On the other hand, the boundary $P$-value for EBSD crystal orientations is computed by the proposed mixture model with the mean distance hypothesis test. Notice that some of the boundaries which are significant in BSE image are not significant in EBSD data and vice versa. The $P$-values for the two modalities are fused together by Eq.(2.1) to generate Fig 7.1(c). Large $P$-values in the fused image shows that the boundaries are not significant either in BSE or EBSD or both, which might be caused by false boundary detection, phase transition or other possible anomalies.

Furthermore, as discussed in Chapter II that the material's mechanical properties are mainly determined by the grain structure (*Rice*, 2000), how to model the nominal distribution of grain structure and infer the abnormal structure of grain arrangement combining the multi-modal microscopy images is an important question for detecting the defects and incipient fatigue failures. One possible approach is to use the graph representation for the grain regions extracted by our proposed method in Chapter V



(a) BSE boundary $P$-value (b) EBSD boundary $P$-value (c) Fused boundary $P$-value

Figure 7.1: (a) The boundary $P$-value is overlaid with BSE image. The boundary $P$-value is calculated by the grain-level Gaussian assumption with the hypothesis test on mean difference between neighboring grains. The $P$-values represent the significance of the existence of the boundaries in BSE. (b) The boundary $P$-value is overlaid with EBSD image. The boundary $P$-value is computed by the von Mises-Fisher distribution with the mean distance hypothesis test. The $P$-values represent the significance of the existence of the boundaries in EBSD. Notice that some boundaries which are significant in BSE image are not significant in EBSD data. (c) The $P$-value fused by Eq.(2.1) for the two modalities. Notice that the boundaries with higher $P$-values are either obscure in BSE or EBSD or both.

and apply the tools in graph theory (*Bondy and Murty*, 1976; *West and others*, 2001; *Diestel*, 2005) to analyze the anomaly of grain structure.

Another interesting direction of the dictionary-based indexing approach is that we can exploit the information of the observed and dictionary diffraction patterns to perform superresolution of grain boundary. As we discussed in Chapter IV, a grain boundary diffraction pattern is a mixture of the diffraction patterns from the spatially neighboring grains. By analyzing the mixture component or the $k$NN patterns in the dictionary of a grain boundary pixel, we should be able to estimate the boundary in sub-pixel level. One of the difficulty we might face of this approach is the computational complexity. The size of dictionary is fairly large (more than 300000 elements in our experiment). How to efficiently find the two or more clusters of $k$NN of a diffraction pattern and perform the analysis is an important challenge of this approach. Also, incorporating the information of misorientation between neighboring grains to improve the segmentation accuracy is another direction worth exploration.

In Chapter VI, we have proposed a robust and accurate registration approach

which uses mutual information as image similarity metric to estimate head movement. Although MI has been proven to be effective in many multi-modal registration applications, its assumption of image pixel distribution may be violated due to the geometric distortion in MR images, resulting in unstable registration result. Other similarity metrics which have been proven to be more robust to the geometric distortion, such as localized mutual information (LMI) (*Russakoff et al.*, 2004; *Klein et al.*, 2008), or graph-based MI (*Neemuchwala et al.*, 2006), may improve the performance. Other than the head motion parameter estimation problem as we focused in this thesis, how to reconstruct the EPI volume using the estimated motion parameters also plays an important role in the performance of brain activation detection. Due to the fact that the EPI images are not parallel and have much coarser resolution than the anatomical volume, how to interpolate and reconstruct the original BOLD signal is worth our further investigation. On the other hand, since the rigid body transformation parameters are composed of Euler angles and translations, it is more natural to use spherical distributions, e.g., Von Mises-Fisher distribution, than multivariate Gaussian to model the change of head rotation. Also, more general head motion model, e.g., a kinematic model (*Han et al.*, 2009), can be adopted to further improve the head motion estimation accuracy.

# APPENDICES

# APPENDIX A

# Fundamental Zone for Cubic Symmetry

In (*Morawiec and Field*, 1996) a set of conditions defining the fundamental zone are given in terms of Rodrigues space:

$$\bigcap_{i=2}^{N}\{r; tan(w_i/4) \pm r\mathbf{l}_i \geq 0\}. \tag{A.1}$$

Here $w_i \in [0, \pi]$ and $\mathbf{l}_i$ are the rotation angle and the unit vector of the rotation axis of the $i$-th element of the rotation symmetry group. $M$ is the order of the group and $i = 1$ corresponds to the identity operator.

For cubic symmetry corresponding to the $m\overline{3}m$ point symmetry group, there are 24 Rodrigues symmetry operators as shown in Table A.1. The transformation equations between the Rodrigues vector $\mathbf{d} = (d_1, d_2, d_3)$ and the quaternion $\mathbf{q} = (q_1, q_2, q_3, q_4)$ are as follows:

$$q_1 = \frac{1}{\sqrt{1 + \|\mathbf{d}\|^2}}, q_i = \frac{d_{i-1}}{\sqrt{1 + \|\mathbf{d}\|^2}}, i \in [2, 3, 4]. \tag{A.2}$$

By applying (A.1) and (A.2), the fundamental zone for the cubic structure in

| Symmetry Operator | Rodrigues Vector |
|---|---|
| 2-fold on $<1,0,0>$ | $\infty(1,0,0)$<br>$\infty(0,1,0)$<br>$\infty(0,0,1)$ |
| 4-fold on $<1,0,0>$ | $\pm(1,0,0)$<br>$\pm(0,1,0)$<br>$\pm(0,0,1)$ |
| 2-fold on $<1,1,0>$ | $\infty(1,\pm1,0)$<br>$\infty(1,0,\pm1)$<br>$\infty(0,1,\pm1)$ |
| 3-fold on $<1,1,1>$ | $\pm(1,1,1)$<br>$\pm(1,-1,1)$<br>$\pm(1,1,-1)$<br>$\pm(-1,-1,1)$ |

Table A.1: Cubic symmetry operators and their Rodrigues Vectors.

quaternion space obeys the following set of equations:

$$\begin{cases} |q_2/q_1| \leq \sqrt{2}-1 \\ |q_3/q_1| \leq \sqrt{2}-1 \\ |q_4/q_1| \leq \sqrt{2}-1 \\ |q_2/q_1 - q_3/q_1| \leq \sqrt{2} \\ |q_2/q_1 + q_3/q_1| \leq \sqrt{2} \\ |q_2/q_1 - q_4/q_1| \leq \sqrt{2} \\ |q_2/q_1 + q_4/q_1| \leq \sqrt{2} \\ |q_3/q_1 - q_4/q_1| \leq \sqrt{2} \\ |q_3/q_1 + q_4/q_1| \leq \sqrt{2} \\ |q_2/q_1 + q_3/q_1 + q_4/q_1| \leq 1 \\ |q_2/q_1 - q_3/q_1 + q_4/q_1| \leq 1 \\ |q_2/q_1 + q_3/q_1 - q_4/q_1| \leq 1 \\ |q_2/q_1 - q_3/q_1 - q_4/q_1| \leq 1 \end{cases} \tag{A.3}$$

# APPENDIX B

# Derivation of EM Algorithm for Mixture of VMF Distributions

The Expectation-Maximization algorithm (*Dempster et al.*, 1977) is an iterative approach for obtaining maximum-likelihood parameter estimates in models where there are hidden latent variables and, in particular, finite mixture models. The algorithm alternates between performing two steps: "Expectation step (E-step)" and "Maximization step (M-step)" in each iteration. In the E-step, one calculates the expectation of the complete data log-likelihood function assuming the parameters of the model are fixed. In the M-step, the updated parameters are estimated by maximizing the expectation function. The process is repeated until the objective function converges.

Here we assume that all the observed quaternions $\{\mathbf{x}_i\}_{i=1}^n$ are realizations from the $M$-fold finite mixture of Von Mises-Fisher (VMF) distributions which has the PDF (B.1).

$$f_v(\mathbf{x}; \boldsymbol{\mu}, \kappa) = \frac{1}{M} \sum_{m=1}^{M} c_p(\kappa) \exp\left(\kappa (\mathbf{P}_m \boldsymbol{\mu})^T \mathbf{x}\right), \tag{B.1}$$

Notice that The only parameters we need to estimate are the mean parameter $\boldsymbol{\mu}$ and the concentration parameter $\kappa$. The latent variable $s_i \in [1, 2, ..., M]$ indicates the

index of the distribution a sample $\mathbf{x}_i$, $i = 1, \ldots, n$, belongs to. Based on the model, the expectation of the log-likelihood given the complete data $\{\mathbf{x}_i, s_i\}_{i=1}^n$, called the $Q$ function, has the following form:

$$\begin{aligned} & Q(\boldsymbol{\omega}; \boldsymbol{\omega}^{(k)}) \\ =& E_{S|X,\boldsymbol{\omega}^{(k)}}[\log L(\boldsymbol{\omega}; \{\mathbf{x}_i, s_i\})] \\ =& \sum_{i=1}^{n} \sum_{m=1}^{M} r_{i,m} (\log c_p(\kappa) + \kappa (\mathbf{P}_m \boldsymbol{\mu})^T \mathbf{x}_i), \end{aligned} \tag{B.2}$$

where $\boldsymbol{\omega} = \{\boldsymbol{\mu}, \kappa\}$ is the set of parameters and $r_{i,m}$ is the posterior probability $P(s_i = m | \mathbf{x}_i, \boldsymbol{\omega})$.

For the E-step, denote $\boldsymbol{\omega}^{(k)} = \{\boldsymbol{\mu}^{(k)}, \kappa^{(k)}\}$ to be the estimated parameters at the $k$-th iteration, $r_{i,m}$ can be calculated by:

$$\begin{aligned} r_{i,m} &= E[\mathbb{1}(s_i = m) | \mathbf{x}_i; \boldsymbol{\omega}^{(k)}] \\ &= P(s_i = m | \mathbf{x}_i; \boldsymbol{\omega}^{(k)}) \\ &= \frac{c_p(\kappa^{(k)}) \exp\left(\kappa^{(k)} (\mathbf{P}_m \boldsymbol{\mu}^{(k)})^T \mathbf{x}_i\right)}{\sum_{l=1}^{M} c_p(\kappa^{(k)}) \exp\left(\kappa^{(k)} (\mathbf{P}_l \boldsymbol{\mu}^{(k)})^T \mathbf{x}_i\right)}. \end{aligned} \tag{B.3}$$

In the M-step the parameters are updated by maximizing the $Q$ function. Taking the partial derivative w.r.t $\boldsymbol{\mu}$ of .(B.2) subject to the constraint $\|\boldsymbol{\mu}\| = 1$ we have:

$$\begin{aligned} \frac{\partial}{\partial \boldsymbol{\mu}} Q(\boldsymbol{\omega}; \boldsymbol{\omega}^{(k)}) &= \sum_{i=1}^{n} \sum_{m=1}^{M} r_{i,m} \kappa \mathbf{Q}_m^T \mathbf{x}_i = 2\lambda \boldsymbol{\mu} \\ \Rightarrow \hat{\boldsymbol{\mu}} = \frac{\boldsymbol{\gamma}}{\|\boldsymbol{\gamma}\|}, \boldsymbol{\gamma} &= \sum_{i=1}^{n} \sum_{m=1}^{M} r_{i,m} \mathbf{Q}_m^T \mathbf{x}_i, \end{aligned} \tag{B.4}$$

where $\lambda$ is a Lagrange multiplier. By taking the partial derivative w.r.t $\kappa$, the ML

estimator of $\kappa$ is:

$$\begin{aligned} \frac{\partial}{\partial \kappa} Q(\boldsymbol{\omega}; \boldsymbol{\omega}^{(k)}) &= N \frac{c_p^{'}(\kappa)}{c_p(\kappa)} + \kappa \boldsymbol{\gamma}^T \boldsymbol{\mu} \\ \Rightarrow \hat{\kappa} &= A_p^{-1}(\frac{\|\boldsymbol{\gamma}\|}{N}), \end{aligned} \tag{B.5}$$

where $A_p(u) = \frac{I_{p/2}(u)}{I_{p/2-1}(u)}$.

# APPENDIX C

# Derivation of EM Algorithm for Mixture of Watson Distributions

Here we assume that all the observed quaternions $X = \{\mathbf{x}_i\}_{i=1}^n$ are realizations from the $M$-fold finite mixture of Watson distributions which has the PDF.

$$f_w(\mathbf{x}; \boldsymbol{\mu}, \kappa) = \frac{1}{M} \sum_{m=1}^{M} \mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)^{-1} \exp \kappa((\mathbf{P}_m \boldsymbol{\mu})^T \mathbf{x})^2, \tag{C.1}$$

where $\mathbb{M}(1/2, p/2, \dot{)}$ denotes the Kummer function

$$\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa) = B(\frac{p-1}{2}, \frac{1}{2})^{-1} \int_{-1}^{1} \exp \kappa t^2 (1-t^2)^{(p-3)/2} dt, \tag{C.2}$$

Let the latent variable $s_i \in [1, 2, ..., M]$ indicates the index of the distribution a sample $\mathbf{x}_i$, $i = 1, \ldots, n$, belongs to. Based on the model, the expectation of the complete likelihood with the hidden variables has the following form:

$$
\begin{aligned}
& Q(\boldsymbol{\omega}; \boldsymbol{\omega}^{(k)}) \\
=& E_{S|X,\boldsymbol{\omega}^{(k)}}[\log L(\boldsymbol{\omega}; \{\mathbf{x}_i, s_i\})] \\
=& \sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}(\log \frac{1}{M\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)} + \kappa((\mathbf{P}_m\boldsymbol{\mu})^T\mathbf{x}_i)^2) \\
=& \sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}\kappa(\boldsymbol{\mu}^T\mathbf{P}_m^T\mathbf{x}_i)(\boldsymbol{\mu}^T\mathbf{P}_m^T\mathbf{x}_i)^T - n\log\left(M\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)\right) \\
=& \kappa\boldsymbol{\mu}^T\left(\sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}(\mathbf{P}_m^T\mathbf{x}_i\mathbf{x}_i^T\mathbf{P}_m)\right)\boldsymbol{\mu} - n\log\left(M\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)\right) \\
=& n\left(\kappa\boldsymbol{\mu}^T\tilde{T}\boldsymbol{\mu} - \log\left(M\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)\right)\right),
\end{aligned}
\tag{C.3}
$$

where $\boldsymbol{\omega} = \{\boldsymbol{\mu}, \kappa\}$ is the set of parameters and $r_{i,m}$ is the posterior probability that $s_i = m$. For the E-step, $r_{i,m}$ can be calculated by:

$$
\begin{aligned}
r_{i,m} &= E[\mathbb{1}(s_i = m)|\mathbf{x}_i; \boldsymbol{\omega}] \\
&= P(s_i = m|\mathbf{x}_i; \boldsymbol{\omega}) \\
&= \frac{\exp\kappa((\mathbf{P}_m\boldsymbol{\mu})^T\mathbf{x}_i)^2}{\sum_{l=1}^{M}\exp\kappa((\mathbf{P}_l\boldsymbol{\mu})^T\mathbf{x}_i)^2}.
\end{aligned}
\tag{C.4}
$$

In the M-step the parameters are updated by maximizing the $Q$ function. From (C.3) we know that $\tilde{T} = \frac{1}{n}\sum_{i=1}^{n}\sum_{m=1}^{M'} r_{i,m}(\mathbf{P}_m^T\mathbf{x}_i\mathbf{x}_i^T\mathbf{P}_m)$ is the scatter matrix of $\mathbf{x}_1, ..., \mathbf{x}_n$. Let $\tilde{t}_1, ..., \tilde{t}_p$ be the eigenvalues of $\tilde{T}$ with

$$
\tilde{t}_1 \geq ... \geq \tilde{t}_p, \tag{C.5}
$$

and let $\pm\mathbf{t}_1, ..., \pm\mathbf{t}_p$ be the corresponding unit eigenvectors. Since we want to find $\boldsymbol{\mu}$ which maximizes (C.3) such that $\boldsymbol{\mu}^T\boldsymbol{\mu} = 1$, the estimator of $\boldsymbol{\mu}$ for fixed $\kappa$ has the

following form:

$$\begin{aligned} \hat{\boldsymbol{\mu}} &= \mathbf{t}_1, \hat{\kappa} > 0, \\ \hat{\boldsymbol{\mu}} &= \mathbf{t}_p, \hat{\kappa} < 0. \end{aligned} \tag{C.6}$$

Similar to the approach in (*Bijral et al.*, 2007), we have the Lagrange function of (C.3) as the following:

$$L = \sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}(\log \frac{1}{M\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)} + \kappa((\mathbf{P}_m\boldsymbol{\mu})^T\mathbf{x}_i)^2) + \lambda\boldsymbol{\mu}^T\boldsymbol{\mu}, \tag{C.7}$$

where $\lambda$ is the Lagrange multiplier. By taking partial derivative w.r.t. $\kappa$ of (C.7) and set to zero, we have:

$$\sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}(\boldsymbol{\mu}^T\mathbf{P}_m^T x_i)^2 = \frac{\mathbb{M}'(\frac{1}{2}, \frac{p}{2}, \kappa)}{\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)} \sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}. \tag{C.8}$$

Let

$$\begin{aligned} Y_p &= \frac{\mathbb{M}'(\frac{1}{2}, \frac{p}{2}, \kappa)}{\mathbb{M}(\frac{1}{2}, \frac{p}{2}, \kappa)} = \frac{\sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}(\boldsymbol{\mu}^T\mathbf{P}_m^T x_i)^2}{\sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}} \\ \Rightarrow \hat{\kappa} &= Y_p^{-1}(\frac{\sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}(\boldsymbol{\mu}^T\mathbf{P}_m^T x_i)^2}{\sum_{i=1}^{n}\sum_{m=1}^{M} r_{i,m}}), \end{aligned} \tag{C.9}$$

The final estimates of $\boldsymbol{\mu}$ and $\kappa$ are obtained by checking both cases ($\hat{\kappa} > 0$, $\hat{\kappa} < 0$) and choosing the one which is consistent for (C.6)(C.9).

# APPENDIX D

# Proof of Equation (5.10)

To proof the second equality of (5.10), we need the following two properties:

**Proposition D.1.** *$A_p(\kappa) = \frac{I_{p/2}(\kappa)}{I_{p/2-1}(\kappa)}$ and $A_p^{-1}(\kappa)$ are strictly increasing functions $\forall \kappa > 0$.*

*Proof.* As shown in (*Amos*, 1974), we have:

$$\frac{\mathrm{d}I_{v+1}(x)/I_v(x)}{\mathrm{d}x} > 0, \forall x > 0. \tag{D.1}$$

This proposition comes directly by letting $v = p/2 - 1$ and $x = \kappa$. The inverse function $A_p^{-1}(\kappa)$ is also a strictly increasing function. □

**Proposition D.2.** *$n \log c_p(\kappa) + n\kappa A_p(\kappa)$ is a strictly increasing function of $\kappa$ for any $n$.*

*Proof.* Let $f(\kappa) = n(\log c_p(\kappa) + \kappa A_p(\kappa))$. Take the derivative of $f$ with respect to $\kappa$ we have:

$$\frac{\mathrm{d}f(\kappa)}{\mathrm{d}\kappa} = n(\frac{1}{c_p(\kappa)}\frac{\mathrm{d}c_p(\kappa)}{\mathrm{d}\kappa} + A_p(\kappa) + \kappa\frac{\mathrm{d}A_p(\kappa)}{\mathrm{d}\kappa}) \tag{D.2}$$

From the derivation of (*Dhillon and Sra*, 2003), we have the formula:

$$A_p(\kappa) = -\frac{1}{c_p(\kappa)} \frac{\mathrm{d}c_p(\kappa)}{\mathrm{d}\kappa} \tag{D.3}$$

Substitute it back to (D.2). We have:

$$\frac{\mathrm{d}f(\kappa)}{\mathrm{d}\kappa} = n(\kappa \frac{\mathrm{d}A_p(\kappa)}{\mathrm{d}\kappa}) \tag{D.4}$$

According to Proposition D.1, we have $\mathrm{d}A_p(\kappa)/\,\mathrm{d}\kappa > 0$. As a result, we can conclude that $\mathrm{d}f(\kappa)/\,\mathrm{d}\kappa > 0$ and $f(\kappa)$ is a strictly increasing function with respect to $\kappa$. □

With the monotonic increasing Propositions D.1 and D.2, (5.10) can be simplified as:

$$\hat{\psi} = \arg\max_{\psi} n \log c_p(\hat{\kappa}_1) + n\hat{\kappa}_1 A_p(\hat{\kappa}_1) \tag{D.5}$$

$$= \arg\max_{\psi} \hat{\kappa}_1 \tag{D.6}$$

$$= \arg\max_{\psi} A_p^{-1}(\frac{\|\mathbf{r}_+\| + \|\mathbf{r}_-\|}{n}) \tag{D.7}$$

$$= \arg\max_{\psi} \|\mathbf{r}_+\| + \|\mathbf{r}_-\| \tag{D.8}$$

$$= \arg\max_{\psi} \|\sum_{\mathbf{p}\in\psi^+} I(\mathbf{p})\| + \|\sum_{\mathbf{p}\in\psi^-} I(\mathbf{p})\| \tag{D.9}$$

# APPENDIX E

# Particle Weights Evaluation

The particle weights are evaluated through the quasi-likelihood function $p(\mathbf{S}_t|\boldsymbol{\theta}_t)$. The quasi-likelihood function should have two properties: (1) It is monotonically increasing with the image similarity $\mathbb{M}(\mathbf{S}_t, T^*_{\boldsymbol{\theta}}(V_{anat}))$; (2) The weighted particles are distributed approximately to multivariate Gaussian. To satisfy the two properties, we propose to use a histogram equalization approach to evaluate the particle weights. The multivariate Gaussian density is shown below:

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^d|\boldsymbol{\Sigma}|}} \exp -\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}). \tag{E.1}$$

The goal here is to find the distribution of $z = f(\mathbf{x})$ where $\mathbf{x}$ is the random variable following (E.1). Let $g_Z(z)$ denote the density of $z$. We can equalize the histogram of image similarity to $g_Z(z)$ to obtain the particle weights.

Without loss of generality and for simplicity, in the following derivation, we assume the covariance to be identity matrix and $\boldsymbol{\mu} = \mathbf{0}$. The density function and its inverse

can be re-written as:

$$\begin{aligned} h(r) &= \frac{1}{\sqrt{(2\pi)^d}} \exp -\frac{1}{2} r^2, \\ h^{-1}(z) &= \sqrt{-2 \log \left( \sqrt{(2\pi)^d} z \right)}, \ z \in (0, \sqrt{(2\pi)^{-d}}], \end{aligned} \tag{E.2}$$

where $r = \|\mathbf{x}\|$. Define $G_Z(z)$ as cumulative density function of $g_Z(z)$ (i.e. $G_Z(z) = p(\{f(\mathbf{x}) \leq z\})$), where $\mathbf{x}$ is the random variable following the multivariate Gaussian density (E.1). According to the spherical symmetry, $G(z)$ has the following form by integration along the radial direction:

$$\begin{aligned} G(z) &= 1 - p(\{f(\mathbf{x}) \geq z\}) \\ &= 1 - \int_0^{h^{-1}(z)} \mathcal{S}_{d-1} u^{d-1} h(u) du \\ &= 1 - \mathcal{S}_{d-1} \left( H^*(h^{-1}(z)) - H^*(0) \right), \end{aligned} \tag{E.3}$$

$$H^*(u) = \int_{-\infty}^{u} v^{d-1} h(v) dv, \tag{E.4}$$

where $\mathcal{S}_{d-1}$ is the surface area of unit $(d-1)$-sphere, e.g., $\mathcal{S}_0 = 2, \mathcal{S}_1 = 2\pi$. To obtain $g(z)$, we need to take the derivative of $G(z)$ with respect to $z$:

$$\begin{aligned} g(z) = \frac{dG(z)}{dz} &= -\mathcal{S}_{d-1} \frac{dH^*(h^{-1}(z))}{dz} \\ &= -\mathcal{S}_{d-1} \frac{H^*(h^{-1}(z))}{dh^{-1}(z)} \frac{dh^{-1}(z)}{dz}, \end{aligned} \tag{E.5}$$

$$\frac{H^*(h^{-1}(z))}{dh^{-1}(z)} = z \left( h^{-1}(z) \right)^{d-1}, \tag{E.6}$$

$$\frac{dh^{-1}(z)}{dz} = -\frac{1}{zh^{-1}(z)}. \tag{E.7}$$

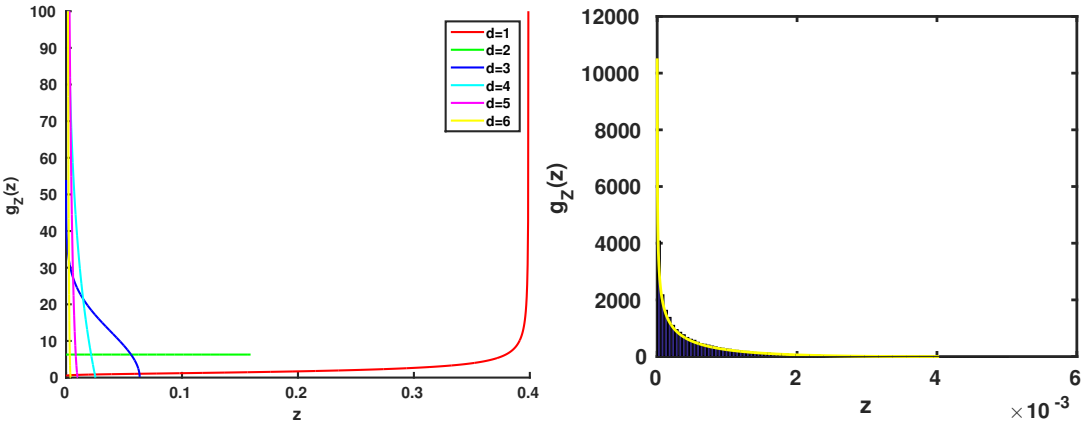
By substituting (E.6)(E.7) into (E.5), we have:

$$\begin{aligned} g_Z(z) &= \mathcal{S}_{d-1}\left(-2\log\left(\sqrt{(2\pi)^d}z\right)\right)^{(d-2)/2} \\ &= \frac{d\pi^{d/2}}{\Gamma(\frac{d}{2}+1)}\left(-2\log\left(\sqrt{(2\pi)^d}z\right)\right)^{(d-2)/2}. \end{aligned} \tag{E.8}$$

Figure E.1(a) shows the density $g_Z(z)$ for different dimension $d$. Notice that in this work, the multivariate Gaussian is used to model the rigid body head motion which has 6 dimensions and therefore $g_Z(z)$ has the following form:

$$g_Z(z) = \pi^3\left(-2\log(2\pi)^3 z\right)^2, \ z \in (0, (2\pi)^{-3}]. \tag{E.9}$$

Figure E.1(b) plots (E.9) with the simulated histogram. The histogram of image similarity is equalized to (E.9) to obtain the weights of each particle.




(a) (b)

Figure E.1: (a) shows the density $g_Z(z)$ for different $d$. (b) shows the simulated histogram compared with theoretical $g_Z(z)$ for $d = 6$.

# BIBLIOGRAPHY

Altmann, S. L. (2005), Rotations, quaternions, and double groups.

Amadieu, O., E. Debreuve, M. Barlaud, and G. Aubert (1999), Inward and outward curve evolution using level set method, in *Image Processing, 1999. ICIP 99. Proceedings. 1999 International Conference on*, vol. 3, pp. 188–192, IEEE.

Amos, D. E. (1974), Computation of Modified Bessel Functions and Their Ratios, *Mathematics of Computation*, *28*(125), 239–251, doi: 10.2307/2005830.

Aschbacher, J., and J. Lichtenegger (1990), Complementary nature of SAR and optical data: a case study in the Tropics, *Earth Observation Quarterly*, *31*, 4–8.

Atrey, P., M. Hossain, A. El Saddik, and M. Kankanhalli (2010), Multimodal fusion for multimedia analysis: a survey, *Multimedia Systems*, *16*(6), 345–379, doi: 10.1007/s00530-010-0182-0.

Bachmann, F., R. Hielscher, and H. Schaeben (2011), Grain detection from 2d and 3d EBSD dataSpecification of the MTEX algorithm, *Ultramicroscopy*, *111*(12), 1720–1733.

Barra, V., and J.-Y. Boire (2000), Quantification of brain tissue volumes using MR/MR fusion, in *Engineering in Medicine and Biology Society, 2000. Proceedings of the 22nd Annual International Conference of the IEEE*, vol. 2, pp. 1451–1454, IEEE.

Bateman, H., A. Erdlyi, W. Magnus, F. Oberhettinger, and F. G. Tricomi (1955), *Higher transcendental functions*, vol. 3, McGraw-Hill New York.

Beisteiner, R., et al. (2001), Finger somatotopy in human motor cortex, *Neuroimage*, *13*(6), 1016–1026.

Bertin, E. P. (1978), Energy-Dispersive X-Ray Spectrometry, in *Introduction to X-Ray Spectrometric Analysis*, pp. 231–253, Springer.

Beucher, S., and others (1992), The watershed transformation applied to image segmentation, *SCANNING MICROSCOPY-SUPPLEMENT-*, pp. 299–299.

Bijral, A. S., M. Breitenbach, and G. Z. Grudic (2007), Mixture of watson distributions: a generative model for hyperspherical embeddings, in *International Conference on Artificial Intelligence and Statistics*, pp. 35–42.

Birkhoff, G., and S. Mac Lane (1963), *A brief survey of modern algebra*, Macmillan.

BlueQuartz (2003), BlueQuartz Software.

Bondy, J. A., and U. S. R. Murty (1976), *Graph theory with applications*, vol. 290, Macmillan London.

Bonnet, N., and J.-C. Liehn (1988), Image registration in electron microscopy: application of a robust method, *Journal of electron microscopy technique*, *10*(1), 27–33.

Brandt, S., J. Heikkonen, and P. Engelhardt (2001), Multiphase method for automatic alignment of transmission electron microscope images using markers, *Journal of structural biology*, *133*(1), 10–22.

Bushberg, J. T., and J. M. Boone (2011), *The essential physics of medical imaging*, Lippincott Williams & Wilkins.

Butts, K., S. J. Riederer, R. L. Ehman, R. M. Thompson, and C. R. Jack (1994), Interleaved echo planar imaging on a standard MRI system, *Magnetic resonance in medicine*, *31*(1), 67–72.

Callahan, P. G., and M. De Graef (2013), Dynamical electron backscatter diffraction patterns. Part I: Pattern simulations, *Microscopy and Microanalysis*, *19*(05), 1255–1265.

Cao, T., C. Zach, S. Modla, D. Powell, K. Czymmek, and M. Niethammer (2014), Multi-modal registration for correlative microscopy using image analogies, *Medical image analysis*, *18*(6), 914–926.

Chan, T., and L. Vese (2001), Active contours without edges, *IEEE Transactions on Image Processing*, *10*(2), 266–277, doi: 10.1109/83.902291.

Chen, Y.-H., S. U. Park, D. Wei, G. Newstadt, M. A. Jackson, J. P. Simmons, M. De Graef, and A. O. Hero (2015a), A Dictionary Approach to Electron Backscatter Diffraction Indexing, *Microscopy and Microanalysis*, *21*(03), 739–752.

Chen, Y.-H., D. Wei, G. Newstadt, M. De Graef, J. Simmons, and A. O. Hero III (2015b), Coercive Region-level Registration for Multi-modal Images, in *International Conference on Image Processing (ICIP)*.

Chen, Y.-H., D. Wei, G. Newstadt, M. De Graef, J. Simmons, and A. O. Hero III (2015c), Parameter Estimation in Spherical Symmetry Groups, in *International Conference on Acoustic Speech and Signal Processing (ICASSP)*.

Chen, Y.-H., D. Wei, G. Newstadt, M. De Graef, J. Simmons, and A. O. Hero III (2015d), Parameter Estimation in Spherical Symmetry Groups, *Signal Processing Letters, IEEE*, *22*(8), 1152–1155, doi: 10.1109/LSP.2014.2387206.

Chen, Y.-H., D. Wei, G. Newstadt, M. De Graef, J. Simmons, and A. O. Hero III (2015e), Statistical Estimation and Clustering of Group-invariant Orientation Parameters, in *18-th International Conference on Information Fusion (Fusion)*.

Chen, Y.-H., R. Mittelman, B. Kim, C. Meyer, and A. O. Hero III (2016a), Multimodal MRI Neuroimaging with Motion Compensation Based on Particle Filtering, *IEEE Transaction on Medical Image*, p. submitted.

Chen, Y.-H., R. Mittelman, B. Kim, C. Meyer, and A. O. Hero III (2016b), Particle Filtering for Slice-to-volume Motion Correction in EPI Based Functional MRI, in *International Conference on Acoustic Speech and Signal Processing (ICASSP)*, p. submitted.

Christodoulou, C., C. S. Pattichis, and others (1999), Medical diagnostic systems using ensembles of neural SOFM classifiers, in *Electronics, Circuits and Systems, 1999. Proceedings of ICECS'99. The 6th IEEE International Conference on*, vol. 1, pp. 121–124, IEEE.

Chuang, H.-C., L. M. Huffman, M. L. Comer, J. P. Simmons, and I. Pollak (2008), An automated segmentation for nickel-based superalloy, in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, pp. 2280–2283, IEEE.

Cocosco, C., V. Kollokian, Kwan, and A. Evans (1997), BrainWeb: Online Interface to a 3d MRI Simulated Brain Database, *NeuroImage*, *5*(4).

Constantinos, S., M. S. Pattichis, and E. Micheli-Tzanakou (2001), Medical imaging fusion applications: An overview, in *Signals, Systems and Computers, 2001. Conference Record of the Thirty-Fifth Asilomar Conference on*, vol. 2, pp. 1263–1267, IEEE.

Cremers, D., M. Rousson, and R. Deriche (2007), A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape, *International journal of computer vision*, *72*(2), 195–215.

De Graef, M., and M. E. McHenry (2007), *Structure of materials: an introduction to crystallography, diffraction and symmetry*, Cambridge University Press.

DeLaPaz, R. L. (1994), Echo-planar imaging., *Radiographics*, *14*(5), 1045–1058.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977), Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 1–38.

Devroye, L., and G. Lugosi (2001), *Combinatorial methods in density estimation*, Springer.

Dhillon, I. S., and S. Sra (2003), Modeling data using directional distributions, *Tech. rep.*, Technical Report TR-03-06, Department of Computer Sciences, The University of Texas at Austin. URL ftp://ftp. cs. utexas. edu/pub/techreports/tr03-06. ps. gz.

Diestel, R. (2005), Graph theory (Graduate texts in mathematics).

Doucet, A., S. Godsill, and C. Andrieu (2000), On sequential Monte Carlo sampling methods for Bayesian filtering, *Statistics and Computing*, *10*(3), 197–208, doi: 10.1023/A:1008935410038.

Durbin, J., and S. J. Koopman (2012), *Time series analysis by state space methods*, 38, Oxford University Press.

Eberly, D. (2008), Euler angle formulas, *Geometric Tools, LLC, Technical Report*.

Echlin, M. P., A. Mottura, C. J. Torbet, and T. M. Pollock (2012), A new TriBeam system for three-dimensional multimodal materials analysis, *Review of Scientific Instruments*, *83*(2), 023,701.

Ehlers, M. (1991), Multisensor image fusion techniques in remote sensing, *ISPRS Journal of Photogrammetry and Remote Sensing*, *46*(1), 19–30.

Elsey, M., S. Esedog, P. Smereka, and others (2013), Simulations of anisotropic grain growth: Efficient algorithms and misorientation distributions, *Acta Materialia*, *61*(6), 2033–2043.

Erie, M. C., C. H. Chu, and R. D. Sidman (1999), Visualization of the cortical potential field by medical imaging data fusion, in *Visual Information and Information Systems*, pp. 815–822, Springer.

Esedo Lu, S., and F. Otto (2015), Threshold dynamics for networks with arbitrary surface tensions, *Communications on Pure and Applied Mathematics*, *68*(5), 808–864.

Figueiredo, M. A., and A. K. Jain (2002), Unsupervised learning of finite mixture models, *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *24*(3), 381–396.

Fish, J. (2006), Bridging the scales in nano engineering and science, *Journal of Nanoparticle Research*, *8*(5), 577–594.

Fisher, R. A. (1934), Statistical methods for research workers.

Franklin, S. E., and C. F. Blodgett (1993), An example of satellite multisensor data fusion, *Computers & Geosciences*, *19*(4), 577–583.

Friedl, M. A., and C. E. Brodley (1997), Decision tree classification of land cover from remotely sensed data, *Remote sensing of environment*, *61*(3), 399–409.

Friston, K. J., J. Ashburner, C. D. Frith, J.-B. Poline, J. D. Heather, and R. S. J. Frackowiak (1995), Spatial registration and normalization of images, *Human Brain Mapping*, *3*(3), 165–189, doi: 10.1002/hbm.460030303.

Fung, J. C., W. Liu, W. De Ruijter, H. Chen, C. K. Abbey, J. W. Sedat, and D. A. Agard (1996), Toward fully automated high-resolution electron tomography, *Journal of structural biology*, *116*(1), 181–189.

Genovese, C. R., D. C. Noll, and W. F. Eddy (1997), Estimating test-retest reliability in functional MR imaging I: Statistical methodology, *Magnetic Resonance in Medicine*, *38*(3), 497–507.

Gevaert, O., J. Xu, C. D. Hoang, A. N. Leung, Y. Xu, A. Quon, D. L. Rubin, S. Napel, and S. K. Plevritis (2012), Nonsmall cell lung cancer: identifying prognostic imaging biomarkers by leveraging public gene expression microarray datamethods and preliminary results, *Radiology*, *264*(2), 387–396.

Glasbey, C., and N. Martin (1996), Multimodal microscopy by digital image processing, *Journal of Microscopy*, *181*(3), 225–237.

Goldstein, J. (2003), *Scanning electron microscopy and x-ray microanalysis*, Kluwer Academic/Plenum Publishers, New York.

Gorski, K. M., E. Hivon, A. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann (2005), HEALPix: a framework for high-resolution discretization and fast analysis of data distributed on the sphere, *The Astrophysical Journal*, *622*(2), 759.

Grau, V., A. Mewes, M. Alcaniz, R. Kikinis, and S. K. Warfield (2004), Improved watershed transform for medical image segmentation using prior information, *Medical Imaging, IEEE Transactions on*, *23*(4), 447–458.

Groeber, M., S. Ghosh, M. D. Uchic, and D. M. Dimiduk (2008), A framework for automated analysis and simulation of 3d polycrystalline microstructures.: Part 1: Statistical characterization, *Acta Materialia*, *56*(6), 1257–1273.

Groeber, M. A. (2011), Digital representation of materials grain structure, in *Computational Methods for Microstructure-Property Relationships*, pp. 53–97, Springer.

Groeber, M. A., D. J. Rowenhorst, and M. D. Uchic (2009), Collection, Processing, and Analysis of Three-Dimensional EBSD Data Sets, in *Electron Backscatter Diffraction in Materials Science*, pp. 123–137, Springer.

Hall, D. L., and S. A. McMullen (2004), *Mathematical techniques in multisensor data fusion*, Artech House.

Han, B., Y. Zhu, D. Comaniciu, and L. S. Davis (2009), Visual tracking by continuous density propagation in sequential bayesian filtering framework, *IEEE transactions on pattern analysis and machine intelligence*, *31*(5), 919–930, doi: 10.1109/TPAMI.2008.134.

Han, J., L. Shao, D. Xu, and J. Shotton (2013), Enhanced computer vision with microsoft kinect sensor: A review, *Cybernetics, IEEE Transactions on*, *43*(5), 1318–1334.

Hartigan, J. A., and M. A. Wong (1979), Algorithm AS 136: A k-means clustering algorithm, *Applied statistics*, pp. 100–108.

Hastie, T., R. Tibshirani, J. Friedman, and J. Franklin (2005), The elements of statistical learning: data mining, inference and prediction, *The Mathematical Intelligencer*, *27*(2), 83–85.

Hero, A. O. (2000), Statistical methods for signal processing, *MATRIX*, *2*, 6.

Hervas, A., et al. (2004), Image-fusion based CT pre and post-implant in seed implantation: a useful tool for accurate prostate definition in the post-planning setting, *Radiotherapy and Oncology*, *71*(Suppl. 2), S107.

Hill, D. L., P. G. Batchelor, M. Holden, and D. J. Hawkes (2001), Medical image registration, *Physics in medicine and biology*, *46*(3), R1.

Hinton, G., S. Osindero, and Y.-W. Teh (2006), A fast learning algorithm for deep belief nets, *Neural computation*, *18*(7), 1527–1554.

Hoaglin, D. C., B. Iglewicz, and J. W. Tukey (1986), Performance of some resistant rules for outlier labeling, *Journal of the American Statistical Association*, *81*(396), 991–999.

Hu, Y., T. J. Carter, H. U. Ahmed, M. Emberton, C. Allen, D. J. Hawkes, and D. C. Barratt (2011), Modelling prostate motion for data fusion during image-guided interventions, *Medical Imaging, IEEE Transactions on*, *30*(11), 1887–1900.

Huffman, L. M., J. Simmons, and I. Pollak (2008), Segmentation of digital microscopy data for the analysis of defect structures in materials using nonlinear diffusions, in *Electronic Imaging 2008*, pp. 68,140B–68,140B, International Society for Optics and Photonics.

Inza, I., P. Larraaga, R. Blanco, and A. J. Cerrolaza (2004), Filter versus wrapper gene selection approaches in DNA microarray domains, *Artificial intelligence in medicine*, *31*(2), 91–103.

James, A. P., and B. V. Dasarathy (2014), Medical image fusion: a survey of the state of the art, *Information Fusion*, *19*, 4–19.

Julier, S. J., and J. K. Uhlmann (1997), New extension of the Kalman filter to nonlinear systems, pp. 182–193, doi: 10.1117/12.280797.

Kainz, B., et al. (2015), Fast Volume Reconstruction from Motion Corrupted Stacks of 2d Slices.

Kalman, R. E. (1960), A New Approach to Linear Filtering and Prediction Problems, *Journal of Fluids Engineering*, *82*(1), 35–45, doi: 10.1115/1.3662552.

Kapur, T., W. E. L. Grimson, W. M. Wells, and R. Kikinis (1996), Segmentation of brain tissue from magnetic resonance images, *Medical image analysis*, *1*(2), 109–127.

Kass, M., A. Witkin, and D. Terzopoulos (1988), Snakes: Active contour models, *International journal of computer vision*, *1*(4), 321–331.

Keys, L., N. Schmidt, and B. Phillips (1990), A prototype example of sensor fusion used for a siting analysis, in *1990 ACSM-ASPRS Annual Convention, Denver, CO*, pp. 238–249.

Kim, B., J. L. Boes, P. H. Bland, T. L. Chenevert, and C. R. Meyer (1999), Motion correction in fMRI via registration of individual slices into an anatomical volume, *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *41*(5), 964–972.

Kim, B., D. T. B. Yeo, and R. Bhagalia (2008), Comprehensive mathematical simulation of functional magnetic resonance imaging time series including motion-related image distortion and spin saturation effect, *Magnetic Resonance Imaging*, *26*(2), 147–159, doi: 10.1016/j.mri.2007.05.007.

Kim, K., P. Habas, F. Rousseau, O. Glenn, A. J. Barkovich, C. Studholme, and others (2010), Intersection based motion correction of multislice MRI for 3-D in utero fetal brain image formation, *Medical Imaging, IEEE Transactions on*, *29*(1), 146–158.

Kirova, Y. M., V. Servois, F. Reyal, D. Peurien, A. Fourquet, and N. Fournier-Bidoz (2011), Use of deformable image fusion to allow better definition of tumor bed boost volume after oncoplastic breast surgery, *Surgical oncology*, *20*(2), e123–e125.

Klein, S., U. A. van der Heide, I. M. Lips, M. van Vulpen, M. Staring, and J. P. Pluim (2008), Automatic segmentation of the prostate in 3d MR images by atlas matching using localized mutual information, *Medical physics*, *35*(4), 1407–1417.

Kok, C., Y. Hui, and T. Nguyen (1996), Medical image pseudo coloring by wavelet fusion, in *Engineering in Medicine and Biology Society, 1996. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE*, vol. 2, pp. 648–649, IEEE.

Kotecha, J. H., and P. Djuric (2003), Gaussian particle filtering, *IEEE Transactions on Signal Processing*, *51*(10), 2592–2601, doi: 10.1109/TSP.2003.816758.

Lassen, N. C. K. (1994), *Automated determination of crystal orientations from electron backscattering patterns*, Institut for Matematisk Modellering, Danmarks Tekniske Universitet.

Latham, S., T. Varslot, and A. Sheppard (2008), Image registration: enhancing and calibrating X-ray micro-CT imaging, in *International Symposium of the Society of Core Analysts, Abu Dhabi*, vol. 29.

Leckie, D., and others (1990), Synergism of synthetic aperture radar and visible/infrared data for forest type discrimination., *PE&RS, Photogrammetric Engineering & Remote Sensing*, *56*(9), 1237–1246.

Lehmann, E. L., and G. Casella (1998), *Theory of point estimation*, vol. 31, Springer Science & Business Media.

Lindseth, F., S. Ommedal, J. Bang, G. Unsg\a ard, and T. N. Hernes (2001), Image fusion of ultrasound and MRI as an aid for assessing anatomical shifts and improving overview and interpretation in ultrasound-guided neurosurgery, in *International Congress Series*, vol. 1230, pp. 254–260, Elsevier.

Maes, F., A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens (1997), Multimodality image registration by maximization of mutual information, *Medical Imaging, IEEE Transactions on*, *16*(2), 187–198.

Maintz, J. A., and M. A. Viergever (1998), A survey of medical image registration, *Medical image analysis*, *2*(1), 1–36.

Mardia, K. V., and P. E. Jupp (1999), Directional statistics.

Marshall, S., and G. Matsopoulos (1993), Morphological data fusion in medical imaging, in *Nonlinear Digital Signal Processing, 1993. IEEE Winter Workshop on*, pp. 6–1, IEEE.

McLachlan, G., and D. Peel (2004), Finite Mixture Models.

McRobbie, D. W., E. A. Moore, M. J. Graves, and M. R. Prince (2006), *MRI from Picture to Proton*, Cambridge university press.

Meyer, C. R., J. L. Boes, B. Kim, P. H. Bland, K. R. Zasadny, P. V. Kison, K. Koral, K. A. Frey, and R. L. Wahl (1997), Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations, *Medical image analysis*, *1*(3), 195–206.

Mitchell, H. B. (2007), *Multi-sensor data fusion: an introduction*, Springer Science & Business Media.

Moissinac, H., H. Maitre, and I. Bloch (1995), Markov random fields and graphs for uncertainty management and symbolic data fusion in an urban scene interpretation, in *Satellite Remote Sensing II*, pp. 298–309, International Society for Optics and Photonics.

Morawiec, A., and D. Field (1996), Rodrigues parameterization for orientation and misorientation distributions, *Philosophical Magazine A*, *73*(4), 1113–1130.

Movellan, J. R., and P. Mineiro (1998), Robust sensor fusion: Analysis and application to audio visual speech recognition, *Machine Learning*, *32*(2), 85–100.

Neemuchwala, H., A. Hero, S. Zabuawala, and P. Carson (2006), Image registration methods in high-dimensional space, *International Journal of Imaging Systems and Technology*, *16*(5), 130–145, doi: 10.1002/ima.20079.

Nelder, J. A., and R. Mead (1965), A Simplex Method for Function Minimization, *The Computer Journal*, *7*(4), 308–313, doi: 10.1093/comjnl/7.4.308.

Nellist, P. D. (2007), Scanning transmission electron microscopy, in *Science of microscopy*, pp. 65–132, Springer.

Newstadt, G. E., A. O. Hero III, and J. Simmons (2014), Robust spectral unmixing for anomaly detection, in *Statistical Signal Processing (SSP), 2014 IEEE Workshop on*, pp. 109–112, IEEE.

Neyman, J., and E. S. Pearson (1992), *On the problem of the most efficient tests of statistical hypotheses*, Springer.

Nichols, T. E., and A. P. Holmes (2002), Nonparametric permutation tests for functional neuroimaging: a primer with examples, *Human brain mapping*, *15*(1), 1–25.

Nock, R., and F. Nielsen (2004), Statistical region merging, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *26*(11), 1452–1458, doi: 10.1109/TPAMI.2004.110.

Noll, D. C., C. R. Genovese, L. E. Nystrom, A. L. Vazquez, S. D. Forman, W. F. Eddy, and J. D. Cohen (1997), Estimating test-retest reliability in functional MR imaging II: application to motor and cognitive activation studies, *Magnetic Resonance in Medicine*, *38*(3), 508–517.

Nordmark, H., M. Di Sabatino, M. Acciarri, J. Libal, S. Binetti, E. Ovrelid, J. Walmsley, and R. Holmestad (2008), EBIC, EBSD and TEM study of grain boundaries in multicrystalline silicon cast from metallurgical feedstock, in *Photovoltaic Specialists Conference, 2008. PVSC'08. 33rd IEEE*, pp. 1–6, IEEE.

Ogawa, S., T. M. Lee, A. R. Kay, and D. W. Tank (1990), Brain magnetic resonance imaging with contrast dependent on blood oxygenation, *Proceedings of the National Academy of Sciences*, *87*(24), 9868–9872.

Oliveira, F. P., and J. M. R. Tavares (2014), Medical image registration: a review, *Computer methods in biomechanics and biomedical engineering*, *17*(2), 73–93.

Ooi, M. B., S. Krueger, W. J. Thomas, S. V. Swaminathan, and T. R. Brown (2009), Prospective real-time correction for arbitrary head motion using active markers, *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *62*(4), 943–954, doi: 10.1002/mrm.22082.

Ooi, M. B., S. Krueger, J. Muraskin, W. J. Thomas, and T. R. Brown (2011), Echoplanar imaging with prospective slice-by-slice motion correction using active markers, *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *66*(1), 73–81, doi: 10.1002/mrm.22780.

Pal, M., and P. M. Mather (2003), An assessment of the effectiveness of decision tree methods for land cover classification, *Remote sensing of environment*, *86*(4), 554–565.

Park, H., C. R. Meyer, and B. Kim (2004), Improved Motion Correction in fMRI by Joint Mapping of Slices into an Anatomical Volume, in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2004*, edited by C. Barillot, D. R. Haynor, and P. Hellier, no. 3217 in Lecture Notes in Computer Science, pp. 745–751, Springer Berlin Heidelberg.

Park, S. U., D. Wei, M. De Graef, M. Shah, J. Simmons, and A. O. Hero (2013), EBSD image segmentation using a physics-based forward model., in *ICIP*, pp. 3780–3784.

Pellemans, A., R. Jordans, and R. Allewijn (1993), Merging multispectral and panchromatic SPOT images with respect to the radiometric properties of the sensor, *Photogrammetric Engineering and Remote Sensing*, *59*(1), 81–87.

Pennycook, S. J., and P. D. Nellist (2011), *Scanning transmission electron microscopy: imaging and analysis*, Springer.

Persson, N., F. Gustafsson, and M. Drev (2002), Indirect tire pressure monitoring using sensor fusion.

Pohl, C., and J. L. Van Genderen (1998), Review article multisensor image fusion in remote sensing: concepts, methods and applications, *International journal of remote sensing*, *19*(5), 823–854.

Qin, L., P. van Gelderen, J. A. Derbyshire, F. Jin, J. Lee, J. A. de Zwart, Y. Tao, and J. H. Duyn (2009), Prospective head-movement correction for high-resolution MRI using an in-bore optical tracking system, *Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine / Society of Magnetic Resonance in Medicine*, *62*(4), 924–934, doi: 10.1002/mrm.22076.

Rauch, E., and L. Dupuy (2005), Rapid spot diffraction patterns idendification through template matching, *Archives of Metallurgy and Materials*, *50*, 87–99.

Rauch, E., M. Veron, J. Portillo, D. Bultreys, Y. Maniette, and S. Nicolopoulos (2008), Automatic crystal orientation and phase mapping in TEM by precession diffraction, *Microscopy and Analysis-UK*, (128), S5.

Raza, M., I. Gondal, D. Green, and R. Coppel (2005), Classifier fusion to predict breast cancer tumors based on microarray gene expression data, in *Knowledge-Based Intelligent Information and Engineering Systems*, pp. 907–907, Springer.

Raza, M., I. Gondal, D. Green, and R. L. Coppel (2006), Classifier fusion using dempster-shafer theory of evidence to predict breast cancer tumors, in *TENCON 2006. 2006 IEEE Region 10 Conference*, pp. 1–4, IEEE.

Ress, D., M. L. Harlow, M. Schwarz, R. M. Marshall, and U. J. McMahan (1999), Automatic acquisition of fiducial markers and alignment of images in tilt series for electron tomography, *Journal of electron microscopy*, *48*(3), 277–287.

Rice, R. W. (2000), *Mechanical properties of ceramics and composites: grain and particle effects*, vol. 17, CRC Press.

Rodrigues, O. (1840), *Des lois gomtriques qui rgissent les dplacements d'un systme solide dans l'espace: et de la variation des cordonnes provenant de ces dplacements considrs indpendamment des causes qui peuvent les produire*, publisher not identified.

Rogers, R., and L. Wood (1990), The history and status of merging multiple sensor data- An overview, in *1990 ACSM-ASPRS Annual Convention, Denver, CO*, pp. 352–360.

Rohr, K., H. S. Stiehl, R. Sprengel, T. M. Buzug, J. Weese, and M. Kuhn (2001), Landmark-based elastic registration using approximating thin-plate splines, *Medical Imaging, IEEE Transactions on*, *20*(6), 526–534.

Roca, D., and G. Plonka (2011), Uniform spherical grids via equal area projection from the cube to the sphere, *Journal of Computational and Applied Mathematics*, *236*(6), 1033–1041.

Roca, D., A. Morawiec, and M. De Graef (2014), A new method of constructing a grid in the space of 3d rotations and its applications to texture analysis, *Modelling and Simulation in Materials Science and Engineering*, *22*(7), 075,013.

Rueckert, D., L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. Hawkes (1999), Nonrigid registration using free-form deformations: application to breast MR images, *IEEE Transactions on Medical Imaging*, *18*(8), 712–721, doi: 10.1109/42.796284.

Russakoff, D. B., C. Tomasi, T. Rohlfing, C. R. Maurer, and Jr. (2004), Image Similarity Using Mutual Information of Regions, in *8th European Conference on Computer Vision (ECCV*, pp. 596–607, Springer.

Rusu, M., H. Wang, T. Golden, A. Gow, and A. Madabhushi (2013), Multiscale multimodal fusion of histological and MRI volumes for characterization of lung inflammation, in *SPIE Medical Imaging*, pp. 86,720X–86,720X, International Society for Optics and Photonics.

Saruwatari, K., J. Akai, Y. Fukumori, N. Ozaki, H. Nagasawa, and T. Kogure (2007), Crystal orientation analyses of biominerals using Kikuchi patterns in TEM, *Journal of Mineralogical and Petrological Sciences*, *103*(1), 16–22.

Schmitt, F., M. K. Stehling, and R. Turner (1998), *Echo-planar imaging*, Springer Science & Business Media.

Schwartz, A. J., M. Kumar, B. L. Adams, and D. P. Field (2009), *Electron backscatter diffraction in materials science*, vol. 2, Springer.

Segal, E., et al. (2007), Decoding global gene expression programs in liver cancer by noninvasive imaging, *Nature biotechnology*, *25*(6), 675–680.

Seize, T. K. (1977), Student's t-test, *Southern Medical Journal*, *70*(11), 1299.

Shah, D., and D. Duhl (1988), Effect of Minor Elements on the Deformation Behavior of Nickel-Base Superalloys, *Superalloys 1988*, pp. 693–702.

Shen, S. (1990), Summary of types of data fusion methods utilized in workshop papers, in *Multisource Data Integration in Remote Sensing, Proceedings of Workshop*, pp. 145–149.

Smith, M. I., and J. P. Heather (2005), A review of image fusion technology in 2005, in *Defense and Security*, pp. 29–45, International Society for Optics and Photonics.

Sohn, K., D. Y. Jung, H. Lee, and A. O. Hero (2011), Efficient learning of sparse, distributed, convolutional feature representations for object recognition, in *Computer Vision (ICCV), 2011 IEEE International Conference on*, pp. 2643–2650, IEEE.

Staring, M., U. A. van der Heide, S. Klein, M. A. Viergever, and J. P. Pluim (2009), Registration of cervical MRI using multifeature mutual information, *Medical Imaging, IEEE Transactions on*, *28*(9), 1412–1421.

Strobl, D., J. Raggam, and M. Buchroithner (1990), Terrain correction geocoding of a multi-sensor image data set, in *Proceedings 10th EARSeL Symposium, Toulouse, France (Paris: European Space Agency)*, pp. 98–107.

Studholme, C., D. L. G. Hill, and D. J. Hawkes (1999), An overlap invariant entropy measure of 3d medical image alignment, *Pattern Recognition*, *32*(1), 71 – 86, doi: http://dx.doi.org/10.1016/S0031-3203(98)00091-0.

Sui, J., T. Adali, Q. Yu, J. Chen, and V. D. Calhoun (2012), A review of multivariate methods for multimodal fusion of brain imaging data, *Journal of Neuroscience Methods*, *204*(1), 68 – 81, doi: http://dx.doi.org/10.1016/j.jneumeth.2011.10.031.

Szeliski, R., and J. Coughlan (1997), Spline-based image registration, *International Journal of Computer Vision*, *22*(3), 199–218.

Tang, G., L. Peng, P. R. Baldwin, D. S. Mann, W. Jiang, I. Rees, and S. J. Ludtke (2007), EMAN2: an extensible image processing suite for electron microscopy, *Journal of structural biology*, *157*(1), 38–46.

Tao, X., and A. Eades (2005), Errors, artifacts, and improvements in EBSD processing and mapping, *Microscopy and Microanalysis*, *11*(01), 79–87.

Tropp, J. A., and S. J. Wright (2010), Computational methods for sparse solution of linear inverse problems, *Proceedings of the IEEE*, *98*(6), 948–958.

Turner, R., A. Howseman, G. E. Rees, O. Josephs, and K. Friston (1998), Functional magnetic resonance imaging of the human brain: data acquisition and analysis, *Experimental Brain Research*, *123*(1-2), 5–12.

Van Genderen, J., and C. Pohl (1994), Image fusion: Issues, techniques and applications, in *Intelligent Image Fusion, Proceedings EARSeL Workshop, Strasbourg, France*, vol. 11, pp. 18–26.

Vese, L. A., and T. F. Chan (2002), A multiphase level set framework for image segmentation using the Mumford and Shah model, *International journal of computer vision*, *50*(3), 271–293.

Viswanath, S. E., N. B. Bloch, J. C. Chappelow, R. Toth, N. M. Rofsky, E. M. Genega, R. E. Lenkinski, and A. Madabhushi (2012), Central gland and peripheral zone prostate tumors have significantly different quantitative imaging signatures on 3 tesla endorectal, in vivo T2-weighted MR imagery, *Journal of Magnetic Resonance Imaging*, *36*(1), 213–224.

Wan, E., and R. Van der Merwe (2000), The unscented Kalman filter for nonlinear estimation, in *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000*, pp. 153–158, doi: 10.1109/ASSPCC.2000.882463.

Wang, Y., H. Kimura, Y. Akiniwa, and K. Tanaka (2005), EBSD-AFM hybrid analysis of fatigue slip system and crack initiation in polycrystalline metal under cyclic torsional loading, in *Micro-NanoMechatronics and Human Science, 2005 IEEE International Symposium on*, pp. 223–228, IEEE.

Watson, G. (1965), Equatorial distributions on a sphere, *Biometrika*, pp. 193–201.

West, D. B., and others (2001), *Introduction to graph theory*, vol. 2, Prentice hall Upper Saddle River.

Wilks, S. S. (1938), The large-sample distribution of the likelihood ratio for testing composite hypotheses, *The Annals of Mathematical Statistics*, *9*(1), 60–62.

Wortmann, T. (2009), Fusion of AFM and SEM scans, in *Optomechatronic Technologies, 2009. ISOT 2009. International Symposium on*, pp. 40–45, doi: 10.1109/ISOT.2009.5326096.

Wright, S. I., and M. M. Nowell (2006), EBSD image quality mapping, *Microscopy and Microanalysis*, *12*(01), 72–84.

Wright, S. I., J.-W. Zhao, and B. L. Adams (1991), Automated determination of lattice orientation from electron backscattered Kikuchi diffraction patterns, *Texture, Stress, and Microstructure*, *13*(2-3), 123–131.

Yershova, A., and S. M. LaValle (2004), Deterministic sampling methods for spheres and SO (3), in *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, vol. 4, pp. 3974–3980, IEEE.

Yershova, A., S. Jain, S. M. Lavalle, and J. C. Mitchell (2009), Generating uniform incremental grids on SO (3) using the Hopf fibration, *The International journal of robotics research*.

Yong, Y., P. Dongsun, H. Shuying, and R. Nini (2010), Medical image fusion via an effective wavelet-based approach, *EURASIP Journal on Advances in Signal Processing*, *2010*(1), 579,341.

Zaitsev, M., C. Dold, G. Sakas, J. Hennig, and O. Speck (2006), Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system, *NeuroImage*, *31*(3), 1038–1050, doi: 10.1016/j.neuroimage.2006.01.039.

Zhu, S., T. S. Lee, and A. Yuille (1995), Region competition: unifying snakes, region growing, energy/Bayes/MDL for multi-band image segmentation, in *, Fifth International Conference on Computer Vision, 1995. Proceedings*, pp. 416–423, doi: 10.1109/ICCV.1995.466909.

Zitova, B., and J. Flusser (2003), Image registration methods: a survey, *Image and vision computing*, *21*(11), 977–1000.