Spring 5-2014

Identifying Humans by the Shape of Their Heartbeats and Materials by Their X-Ray Scattering Profiles

Ikenna C. Odinaka
Washington University in St Louis

WASHINGTON UNIVERSITY IN ST. LOUIS

School of Engineering & Applied Science

Department of Electrical & Systems Engineering

Dissertation Examination Committee:
Joseph A. O'Sullivan, Chair
R. Martin Arthur
Arye Nehorai
David G. Politte
John W. Rohrbaugh
Kilian Q. Weinberger

Identifying Humans by the Shape of Their Heartbeats and Materials by Their X-ray Scattering Profiles

by

Ikenna C. Odinaka

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2014

St. Louis, Missouri

Table of Contents

List of Figures

List of Tables

List of Pseudocodes

List of Abbreviations

AFS	After Feature Selection
AM	Alternating Minimization
BFS	Before Feature Selection
CACSI	Coded Aperture Coherent Scatter Imaging
CAXI	Coded Aperture X-ray Imaging
CMC	Cumulative Match Characteristic
CT	Computed Tomography
DET	Detection Error Trade-off
ECG	Electrocardiogram
EER	Equal Error Rate
EM	Expectation-Maximization
FAR	False Acceptance Rate
FRR	False Rejection Rate
FWHM	Full Width Half Maximum
keV	kilo-electron volt
kVp	Peak kilo-voltage
LDV	Laser Doppler Vibrometry
LLR	Log-likelihood Ratio
mAs	milliampere second
ML	Maximum Likelihood
OSEM	Ordered Subset Expectation-Maximization
rad	Radians
ROC	Receiver Operating Characteristic
STFT	Short-Time Fourier Transform

Acknowledgments

I thank God for everything He has done and continues to do in my life.

I thank my loving wife, Yara, for her constant support and for her understanding during these past couple of months of writing my dissertation. I thank my daughter Trinity for not revolting, even as I spent less time with her while writing. I also wish to thank my daughter Sophia, who gave me a lot of time to write, even going several days past her due date, before gracing us with her presence.

I am thankful to my parents, Chief and Mrs. C.C. Odinaka and my cousin Engr. Uche Umeh for their financial support. I will like to acknowledge my older brother, Engr. Chikezie Odinaka, who taught me immense love for the sciences.

My in-laws, Dr. Rafael and Dr. Alda Massinga took the time to travel across continents to help with the upkeep of my children and the home, during the course of my writing. I could not have done it without them.

I thank all those who have been instrumental in teaching and guiding me in the past, including, but not limited to, Mrs. O.P. Odinaka (my mother and Mathematics teacher), Mr. Ese Ottuh, Prof. Narendra Jaggi, Prof. Gabriel Spalding, Prof. Tian Xiao He, and Prof. Zahia Drici.

A special thanks goes to the many graduate students and distinguished faculty who have helped support my research. This includes but is not limited to Prof. David Brady, Dr. Joel Greenberg, Mehadi Hassan, Andrew Holmgren, Prof. Anuj Kapadia, Dr. Alan D. Kaplan, Dr. Kalyani Krishnamurthy, Dr. Po-Hsiang Lai, Manu Lakshmanan, Dr. Kenneth MacCabe, Prof. Joseph O'Sullivan, Dr. David G. Politte, Prof. John W. Rohrbaugh, Pooyan Sahbaee, Dr. Erik J. Sirevaag, and Prof. Scott Wolter.

A few of the figures used in this dissertation were created by others and credits are given in the body of the dissertation. The creators include Dr. Politte, Prof. O'Sullivan, and Pooyan Sahbaee. The Monte Carlo and real data were obtained by Dr. Greenberg, Mehadi Hassan, Andrew Holmgren, Manu Lakshmanan, Dr. MacCabe, and Pooyan Sahbaee. The reference database of momentum transfer spectra was created by Prof. Scott Wolter.

I thank Prof. Arye Nehorai for organizing study groups on Linear Algebra, Statistics, and Linear Models during my first year of graduate studies. I have also benefited from attending and presenting at a Machine Learning seminar organized by Prof. Kilian Weinberger. I cherish the discussions I had with previous and current members of our group including Yaqi Chen, Soysal Degirmenci, Dr. Sean Kristjansson, Daheng Li, Ke Li, Jingwei Lu, Aswin J. Matthews, Homayoon Ranjbar, and Amanda K. Sheffield.

I thank the members of my committee for their time and invaluable comments and critique of my work: Prof. Joseph A. O'Sullivan, Prof. R. Martin Arthur, Prof. Arye Nehorai, Dr. David G. Politte, Prof. John W. Rohrbaugh, and Dr. Kilian Q. Weinberger.

Lastly, I wish to thank my advisor and mentor Prof. Joseph A. O'Sullivan for taking me on as a graduate student, sharing his interests with me, and forming me into the scholar I am today.

The work on Biometrics has been supported in part by the Technical Support Working Group (TSWG), contract number W91CRB-04-C-0030. The work on X-ray Scatter Imaging has been supported by the U.S. Department of Homeland Security Science and Technology Directorate, contract number HSHQDC-11-C-00083.

Ikenna C. Odinaka

Washington University in Saint Louis
May 2014

To my lovely wife, Yara.

ABSTRACT OF THE DISSERTATION

Identifying Humans by the Shape of Their Heartbeats and

Materials by Their X-ray Scattering Profiles

by

Ikenna C. Odinaka

Doctor of Philosophy in Electrical Engineering

Washington University in St. Louis, 2014

Professor Joseph A. O'Sullivan, Chair

Security needs at access control points presents itself in the form of human identification and/or material identification. The field of Biometrics deals with the problem of identifying individuals based on the signal measured from them. One approach to material identification involves matching their x-ray scattering profiles with a database of known materials.

Classical biometric traits such as fingerprints, facial images, speech, iris and retinal scans are plagued by potential circumvention they could be copied and later used by an impostor. To address this problem, other bodily traits such as the electrical signal acquired from the brain (electroencephalogram) or the heart (electrocardiogram) and the mechanical signals acquired from the heart (heart sound, laser Doppler vibrometry measures of the carotid pulse) have been investigated. These signals depend on the physiology of the body, and require the individual to be alive and present during acquisition, potentially overcoming circumvention.

We investigate the use of the electrocardiogram (ECG) and carotid laser Doppler vibrometry (LDV) signal, both individually and in unison, for biometric identity recognition. A parametric modeling approach to system design is employed, where the system parameters

are estimated from training data. The estimated model is then validated using testing data. A typical identity recognition system can operate in either the authentication (verification) or identification mode. The performance of the biometric identity recognition systems is evaluated using receiver operating characteristic (ROC) or detection error tradeoff (DET) curves, in the authentication mode, and cumulative match characteristic (CMC) curves, in the identification mode.

The performance of the ECG- and LDV-based identity recognition systems is comparable, but is worse than those of classical biometric systems. Authentication performance below 1% equal error rate (EER) can be attained when the training and testing data are obtained from a single measurement session. When the training and testing data are obtained from different measurement sessions, allowing for a potential short-term or long-term change in the physiology, the authentication EER performance degrades to about 6 to 7%.

Leveraging both the electrical (ECG) and mechanical (LDV) aspects of the heart, we obtain a performance gain of over 50%, relative to each individual ECG-based or LDV-based identity recognition system, bringing us closer to the performance of classical biometrics, with the added advantage of anti-circumvention.

We consider the problem of designing combined x-ray attenuation and scatter systems and the algorithms to reconstruct images from the systems. As is the case within a computational imaging framework, we tackle the problem by taking a joint system and algorithm design approach. Accurate modeling of the attenuation of incident and scattered photons within a scatter imaging setup will ultimately lead to more accurate estimates of the scatter densities of an illuminated object. Such scattering densities can then be used in material classification.

In x-ray scatter imaging, tomographic measurements of the forward scatter distribution are used to infer scatter densities within a volume. A mask placed between the object and the detector array provides information about scatter angles. An efficient computational

implementation of the forward and backward model facilitates iterative algorithms based upon a Poisson log-likelihood. The design of the scatter imaging system influences the algorithmic choices we make. In turn, the need for efficient algorithms guides the system design.

We begin by analyzing an x-ray scatter system fitted with a fanbeam source distribution and flat-panel energy-integrating detectors. Efficient algorithms for reconstructing object scatter densities from scatter measurements made on this system are developed. Building on the fanbeam source, energy-integrating flat-panel detection model, we develop a pencil beam model and an energy-sensitive detection model. The scatter forward models and reconstruction algorithms are validated on simulated, Monte Carlo, and real data.

We describe a prototype x-ray attenuation scanner, co-registered with the scatter system, which was built to provide complementary attenuation information to the scatter reconstruction and present results of applying alternating minimization reconstruction algorithms on measurements from the scanner.

Chapter 1

Introduction

1.1 Motivation

Security needs at access points such as airports, border crossings, and hospitals have been the impetus for my work in Biometrics and Computational Imaging. Biometrics deals with identifying humans by means of any of their traits while Coherent Scatter Imaging, an area of Computational Imaging, deals with estimating the coherent scattering densities of materials from measured scatter data. These scatter densities can be compared with a database of known materials for material identification. These two aspects of security are important for a well rounded system, since loopholes in either human or material identification could lead to a potential security threat.

1.2 Background

Several biometric traits have been used for access control and human recognition including fingerprint, iris scans, retina scans, face, and voice. The discovery that biometric systems based on these traits can be circumvented by counterfeiting has led to investigation of several potential biometric traits that could potentially avoid this problem. A few biometric traits include the electrocardiogram (ECG), electroencephalogram (EEG), and the laser Doppler vibrometry (LDV) signal acquired at the neck region overlying the carotid artery. These signals are functional in nature and depend on the state of the individual, in addition to encoding stable traits.

The functional nature of the ECG, EEG and LDV signals suggests that they can only be acquired from a living person. However, this advantage of liveness and robustness to

counterfeiting is offset to some extent by the recognition performance degradation that can result from a substantive change of state from the time the signal is first recorded to a future time.

In addition to the use of biometric systems for access control, several other systems have been deployed for detecting security threats, such as explosives. These systems include X-ray attenuation imaging, X-ray scatter imaging, and imaging approaches based on other forms of radiation, such as gamma rays. Each view of an X-ray attenuation imaging system represents a set of lines through the object, with an object-integrated measurement obtained at the detector. X-ray attenuation imaging, in the form of computed tomography (CT), relies on multiple views around the object, to delineate the contribution of each point within the object [98]. This can be achieve most commonly by a single source and an array of detectors on a rotating gantry [98]. The result of reconstructing data from X-ray CT machines is a volumetric attenuation map. A dual-energy X-ray attenuation imaging system is able to recover both the electron density and effective atomic number effective Z).

There are two forms of X-ray scatter imaging namely, incoherent (Compton) and coherent (Rayleigh). Incoherent scatter imaging has been proposed for estimating the electron density of an object volume [32, 22]. However, several materials share the same electron density and effective atomic number (effective Z), making material identification based solely on X-ray attenuation imaging or incoherent scatter imaging very limited.

In contrast, coherent scatter imaging has been proposed for material characterization using X-ray scattering signatures such as the momentum transfer profile or the coherent scattering differential cross section, which are material-specific [45, 28, 79]. Being able to accurately reconstruct such properties of a material could mean successful material recognition using a dictionary of known materials.

Coherent scatter imaging can be roughly classified into three categories: selected volume tomography (SVT) [28], coherent scatter computed tomography (CSCT) [45], and coded aperture coherent scatter imaging (CACSI) [77]. SVT utilizes special collimators to isolate a subset of object voxels at a time. This makes it easy to invert measured data, but usually results in larger X-ray doses, longer scan times for an entire volume, and longer acquisition times (or larger X-ray flux) to get a usable signal for a given voxel, since most of the scatter signal is suppressed by the collimators. CSCT is based on the idea of computed tomography in X-ray attenuation imaging. Several views around the object help delineate the contributions of different voxels to the measured data. This results in large X-ray doses, long

scan times, and a rotating gantry. More recently, snapshot coherent scatter imaging using a coded aperture was proposed [79]. This approach solves the problem of signal suppression by collimators and a rotating gantry, at the price of more computation to disentangle the contributions of different voxels to the measured signal. The later chapters of this work describe our contributions towards system and algorithm design for a joint coded aperture X-ray scatter and attenuation imaging system. This is joint work with our collaborators from Duke University and the University of North Carolina.

1.3 Outline

Chapter 2 considers our approach to ECG biometrics based on a log-normal model applied to the magnitude of short-time Fourier transform (STFT) features. There we show performance degradation when training and testing occur on different recording sessions. We consider an approach to ameliorate this performance loss by the use of training data from more than one recording session.

There have been several techniques proposed in peer-reviewed literature for ECG biometrics in the past 15 years. These methods have been applied to different data sets extracted from public or in-house databases. Moreover, the sample size and demographics of the population vary considerably. This makes a direct comparison of all the methods difficult. In Chapter 3, we perform a comprehensive study of all the methods that have been proposed for ECG biometrics prior to 2013, and test them on our in-house database. Our database is comprised of about 300 individuals from diverse populations. We show that our method is one of the best for ECG biometrics.

When a single biometric trait does not reach a target performance level in certain applications, an attractive option is to combine multiple traits to achieve that goal. The question of which biometric traits should be combined is an important one and depends on the performance level of each individual biometric system, and how correlated the signals are. When one of the biometric systems suffers from very poor performance level, it can be the limiting system in the fusion. Also, highly correlated traits may not lead to improved performance, if naively combined. Chapter 4 considers our work in fusing an ECG biometric system and an LDV biometric system. This work represents the first time these two signals have been fused for biometric purposes. We show that the two signals are highly uncorrelated and

fusing them at different levels of the biometric pipeline leads to a substantial improvement in biometric performance.

It is of interest to determine the total number of people that can be recognized by a biometric system with a low probability of error. That is, what is the system's biometric capacity? Chapter 5 sets out to address this question using LDV data acquired from a single session. A hierarchical model, based on conjugate distributions, for the mean and variance of the log-normal distribution of the magnitude STFT features, is developed for this purpose.

The last chapter on Biometrics, Chapter 6, considers an analytical framework that captures random hidden states that occur in functional biometrics such as ECG and LDV biometrics. We propose an analytical basis for the use of more robust models for biometrics and speculate on the impact of these hidden states for biometric system design.

Chapter 7 considers an X-ray forward scatter system that utilizes a fan beam source distribution, a flat-panel energy-integrating detector, and a secondary aperture placed between the object and the detector to reconstruct object scattering densities, which include momentum transfer profiles, from measured forward scatter data. There we address techniques for speeding up the computation of the forward and backward operator, and the overall reconstruction algorithm. The forward model is then validated using simulated and Monte Carlo scatter data. The fan beam model assumes a non-attenuating path for the incident and scattered rays.

The fan beam model developed in Chapter 7 is extended in Chapter 8 to two other systems, namely: a pencil beam system and an energy-sensitive detection system. For the pencil beam system, we validate our model using Monte Carlo data from a suitcase containing a target material surrounded by clutter. Real data acquired from a modified Smith's [30] machine is used in validating the fan beam energy-sensitive detection system. In addition, we consider two noise reduction techniques that can be utilized by any of the scatter models.

The models developed in Chapters 7 and 8 assume the incident and scattered photons do not get attenuated. A step towards lifting this simplifying assumption is to model the attenuation of the object, by supplementing the scatter system with a co-registered attenuation system. Chapter 9 considers a prototype X-ray attenuation system. This system was built by hand, and although steps were taking to properly align the detectors along a tunnel, it is clear that some misalignments still exist. In that chapter, we consider the problem of calibrating the detectors using a phantom and some prior information about the location and orientation of

the detector boards. Preliminary reconstruction results are shown for data obtained using the prototype system.

The last chapter of this dissertation considers avenues for extending the work described here. In particular, for biometrics, ways for fusing the ECG and LDV signals at the raw and feature levels are explored. The next step for the X-ray scatter system is the utilization of the attenuation information in the scatter model. The derivation and efficient implementation of this joint attenuation and scatter model is one of our future goals.

This work contains co-authored material that has previously been published and is yet to be published.

1.4 Contributions

The contributions of this work are:

- Performed a large scale ECG biometric experiment on about 300 individuals from three sessions using a novel method for feature extraction, feature selection, and classification.
- Performed a comprehensive and comparative review of all the methods that have been proposed for ECG biometrics using our large multi-session data set and showed that our method is one of the best.
- Performed a large scale cardiovascular biometric experiment on about 300 individuals from three sessions. This represents a significant extension of the results of uni-variate modeling [20, 88] to a multidimensional case.
- Showed that ECG and LDV signals provide uncorrelated information for human recognition.
- Evaluated standard fusion methods for combining the information provided by ECG and LDV signals.
- Showed a 60% authentication performance improvement can be achieved by combining both modalities.

- Developed key elements of a computational imaging approach to a joint attenuation and coherent scatter imaging system.
- Developed and tested an efficient forward model for an X-ray coherent scatter system with a fan beam source distribution and flat-panel energy-integrating detectors.
- Extended the model to other source distributions and detection systems.
- Developed a software package for reconstructing X-ray coherent scatter data.
- Developed a software package for reconstructing co-registered X-ray attenuation and scatter data.
- Implemented and improved a trust region modification of Newton's method for one-dimensional root-finding.
- Developed a method for estimating the position and orientation of detector boards.
- Contributed to the analysis of new system design through performance analysis as a function of system design choices.

Chapter 2

ECG Biometrics

This chapter appeared in a previous IEEE conference publication [87].

2.1 Highlights

We present the results of an analysis of the electrocardiogram as a biometric using a novel short-time frequency method with robust feature selection. Our proposed method incorporates heartbeats from multiple days and fuses information. Single channel ECG signals from a comparatively large sample of 269 individuals that were sampled from the general population were collected on three separate occasions over a seven-month period. We studied the impact of long-term variability, health status, data fusion, the number of training and testing heartbeats, and database size on ECG biometric performance. The proposed method achieves 5.58% equal error rate (EER) in verification, 76.9% accuracy in rank-1 recognition, and 93.5% accuracy in rank-15 recognition when training and testing heartbeats are from different days. If training and testing heartbeats are collected on the same day, we achieve 0.37% EER and 99% recognition accuracy for decisions based on a single heartbeat.

2.2 Introduction

Since the work of Biel *et al.* [9, 10] and Kyoso [67], there has been increased interest in biometrics based on the Electrocardiogram (ECG). ECG as a biometric has three key properties:

- ECG signals are difficult to counterfeit.

- The ECG signal is present in all living individuals.
- It provides additional information related to psychological states and physiological status, which may be of interest for certain applications.

A number of methods and analyses have been proposed under several operational scenarios. For example, fiducial-based methods, autocorrelation, and time-frequency methods using 12-channel ECG signals from healthy individuals under normal resting conditions have been examined extensively [106, 53, 96, 16, 17, 6, 118, 48]. ECG as a biometric based on a single channel has also been studied [16, 17]. Chan *et al.* studied the biometric performance of one channel ECG signals recorded from individuals' thumbs. Israel *et al.* [53] studied the effects of mental stress on biometric performance by requiring healthy individuals to perform various tasks while the ECG signal was recorded. Wübbeler *et al.* specifically studied the effect of variability over time by using training and testing ECG signals recorded on different days [130]. In the work of [5], the possibility of using ECG as a biometric on individuals with two specific heart diseases, premature ventricular contraction (PVC) and atrial premature contraction (APC), was studied. In the work of [135], a large data set comprising 502 individuals was used, with training and testing heartbeats taken consecutively on the same day.

Here, we propose a novel robust time-frequency method for ECG biometrics, applying a large data set of one channel ECG signals from 269 individuals varying in their demographic and health status. We focus on contributing to the development of ECG as a biometric with respect to:

- *Analyses on a large representative database.* Most previous results are based on ECG signals selected from only healthy individuals or only unhealthy individuals with specific diseases from three major databases, the MIT-BIH Normal Sinus Rhythm, MIT-BIH Arrhythmia, and PTB databases. In studies that have used these databases, strict individual selection criteria were applied. The present report is based on a large sample ($n = 269$) drawn from the general population.
- *A novel robust method for signal variability and information fusion.* Using fusion over multiple training sessions, we propose an information theoretic method to select features based on their distinguishability and stability. We have previously shown this approach to be useful in another cardiovascular based biometric using Laser Doppler Vibrometry measurements [69, 59, 21, 19, 20].

- *More acceptable electrode placement.* We use a simple recording montage, with electrodes placed bilaterally on the lower rib cage, so that individuals are not required to undress. Despite the use of these non-standard electrode placements, we show that our performance is still comparable with performance reported in other studies.

- *The effects of training and testing sizes.* To understand the utility of ECG as a biometric, we examine the ECG biometric performances using different number of training and testing heartbeats, as well as number of individuals in the database.

We focus on ECG biometric performance under two scenarios: acquiring training and testing heartbeats on the same day, and acquiring testing heartbeats from one week to several months after the training heartbeats have been obtained. For both scenarios, we study the effects of the number of training and testing heartbeats used on the ECG biometric performance. Furthermore, we show how to improve performance through robust feature selection and the fusion of training information from different days. Both authentication and recognition performances are studied. For within-session analysis, the proposed method achieves 0.37% equal error rate (EER) based on a single heartbeat, and 0.02% based on multiple test heartbeats, which is comparable to the best existing methods. Moreover, based on multiple test heartbeats, the recognition accuracy is 99.2%. In particular, for healthy individuals, we achieve a recognition accuracy of 99.67%, which is one of the highest reported accuracies in the existing literature.

This chapter addresses a method of scientific analysis of ECG signals as a biometric. It does not suggest ways to implement and deploy the proposed ECG biometric system.

The chapter is organized as follows. Section 2.3 introduces the experimental data set and ECG signal preprocessing methods. In Section 2.4, we discuss the operational scenarios studied in the chapter. The novel short-time frequency method is introduced in Section 2.5, with simulation results given in Section 2.6, and a summary in Section 2.7.

2.3 ECG Data Set and Preprocessing

The ECG signals were obtained from a single channel, with the electrodes placed bilaterally on the lower rib cage. Signals were recorded with a Biopac TEL-100 system, using a 0.5 Hz high pass filter and 500 Hz low pass filter. The ECG signal acquired at this location has a strong R wave and is less affected by movement artifact than some of the conventional ECG

leads. The individuals were asked to sit for five minutes as the recording took place. Each of the 269 individuals had ECG recorded on three different days, which we call sessions 1, 2, and 3, where the separation between recording days ranged from two weeks to six months. The ages of the individuals varied from 18 to 66 years, with mean and standard deviation of 38.8 and 14.1 years respectively. 40.15% of the individuals had some self-reported cardiovascular disorder, including hypertension, although individuals were excluded if they reported frank heart disease including history of infarction, cardiomyopathy, or valve disorders. 46.84% of the individuals used medicines or other substances that may affect the ECG signal. 27.88% of the individuals were healthy and did not use substances that may affect the ECG signal. 53.9% of the individuals were female, and 72.12%, Caucasian.

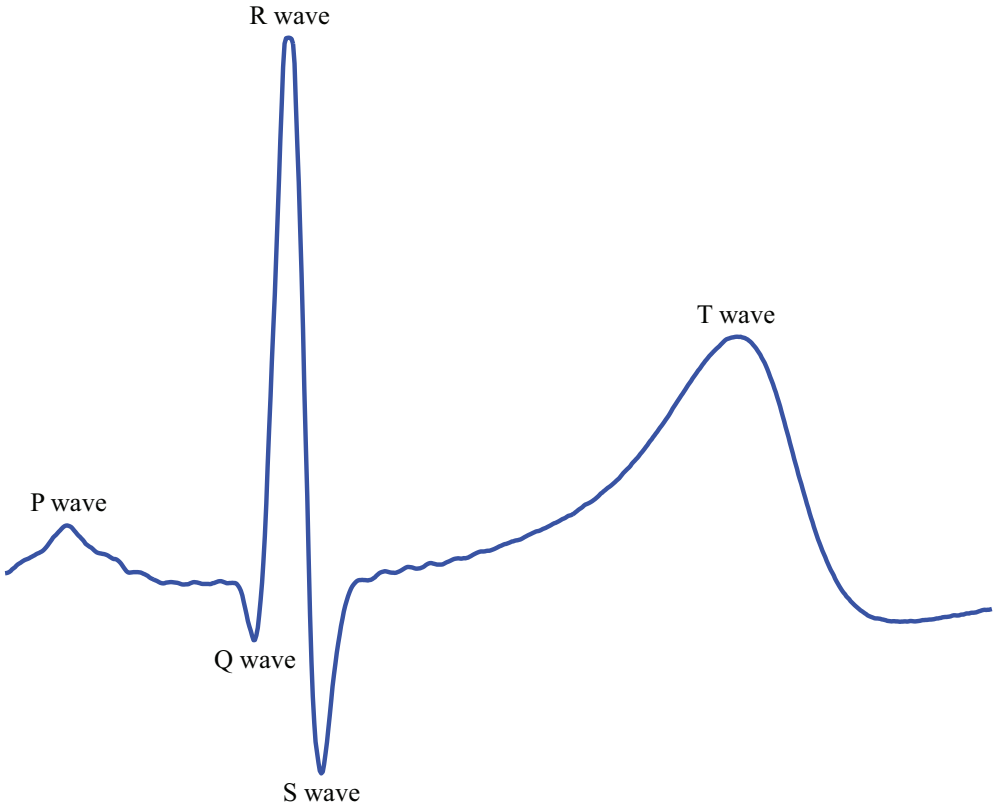


Figure 2.1: A typical ECG pulse with marked P, Q, R, S, T waves.

The ECG signals, which were originally sampled at 10 kHz, were subsequently down-sampled to 1 kHz and digitally notch filtered at 60 Hz to remove power line interference. The resulting ECG signal was reduced to individual 700 msec segments aligned to the respective R wave peak, beginning 200 msec prior to the peak. The segment (heart pulse) duration was chosen to ensure that all of the major components (P, Q, R, S, and T waves of a single heartbeat) were included while minimizing the possibility of including portions of adjacent beats. Each ECG heart pulse, for a given individual, was normalized by subtracting the sample mean of

the pulse, and dividing by the sample standard deviation. Figure 2.1 shows a typical ECG pulse signal with the waves delineated.

In order to align the ECG segments to the peak of the R wave, the peak has to be detected. For each study participant, the data were first high-pass filtered using an IIR elliptic filter (5 Hz cutoff, 90 dB attenuation). An artifact-free calibration epoch of 15 sec was selected. All positive inflection points (peaks) within this epoch were then determined. Only peaks whose amplitude was at least 75% of the amplitude of the maximum detected peak were retained for further analysis. An initial detection template, individualized for each study participant, was created by averaging periods of ECG activity (epoch duration = 750 msec) time-locked to each peak (epoch onset = 250 msec prior to detected peaks). This initial detection template generally represented an accurate model of an individual's typical heart beat; however, such averages can be inappropriately influenced by factors such as noise spikes in the data, or abnormally large amplitude T waves. In order to reduce the impact of such influences, a final detection template was constructed according to the following procedure. The initial detection template was cross-correlated with the 15 sec calibration epoch. The resulting cross-correlation function was then examined for inflection points whose values represented correlations greater than or equal to 0.35. Each resulting peak in the cross-correlation function was assumed to indicate a detected heart beat. The peak of the R wave associated with each detected beat was determined by finding the maximum value in the filtered ECG data in a 100 msec window beginning 50 msec prior to the peak in the associated cross-correlation function. A final detection template was then created by averaging 750 msec ECG epochs beginning 250 msec prior to the peak of each detected R wave in the calibration epoch. R waves in the entire ECG recording were then detected by cross-correlating the final detection template with the complete ECG data vector. In a manner analogous to that used to create the final detection template, detected heart beats were defined as peaks in the cross-correlation function whose value exceeded 0.35. R wave peaks were estimated by determining the time of the maximum value in the ECG signal within a 100 msec window centered on the time at which the heart beat detection peak occurred in the cross-correlation function.

2.4 Biometric Performance Evaluation Scenarios

In addition to having a database of diverse individuals, we are interested in evaluating ECG biometric performance under the effects of temporal variability due to different data

acquisition dates, differences in the number of heartbeats used in the testing and training sets, and varying the number of individuals in the database. To study the temporal effects, there were three testing scenarios, one of which involved using training and testing heartbeats from the same day, and two of which involve training and testing heartbeats from different days:

- *Within session tests.* For a given session, ECG recordings from each individual are separated into two parts. The training heartbeats comprise subsets of the first half of the recordings, while the testing heartbeats comprise subsets of the second half of the recordings.
- *Across sessions tests: single training session.* In this case, training heartbeats consist of subsets of heartbeats from a session, whereas the testing heartbeats come from another session. This study focuses on the effects of ECG time variability.
- *Across sessions tests: multiple training sessions.* In this case, the training heartbeats come from two sessions while the testing heartbeats come from the third session. This study focuses on how to train a classifier using heartbeats from multiple days, and how much performance increases by doing such.

We evaluate performance based on verification (authentication) and recognition (identification). In verification, the system is designed to verify a claimed identity using the biometric trait. In recognition, the system is required to infer the individual's identity based on the biometric trait. For each test and task, different numbers of heartbeats ranging from 1 to 256 are used for training and testing. This helps to understand the sample size effect on ECG biometric performance, and for operational purposes, to optimize data acquisition time. For recognition, we further investigate the performance when the database includes different numbers of individuals. To do so, we simulate databases consisting of 20 to 260 individuals, in steps of 20, by selecting contiguous subsets of individuals from our original data set.

2.5 The Time-frequency Method and Robust Feature Selection

From each ECG pulse signal, we compute a spectrogram which is the logarithm of the square of the magnitude of the short-time Fourier transform of a normalized ECG heart pulse. In

computing the short-time Fourier transform (STFT), we use a Hamming window of size 64ms, with a step size (the distance between the beginnings of two consecutive windows) of 10ms. Thus, there is an overlap of size 54ms between consecutive time frames. This window size was chosen empirically so that it yields robust and good single-heartbeat authentication performance in terms of equal error rate (EER). After computing the STFT, the frequency content was truncated at 250Hz to reduce boundary effects. The spectrogram is then computed as the logarithm of the squared-magnitude of the truncated STFT. We refer to the index of each point of the spectrogram as a time-frequency bin. Thus each ECG heart pulse can be represented by $L = 2048$ time-frequency components denoted as $Y(l)$. To build a generative classifier, we use independent normal distributions to model the time-frequency bins of each individual. During training, only the means and variances have to be estimated. For each bin l of individual i, we use the maximum likelihood (ML) estimates which are the sample means and variances denoted as $\hat{\theta}_i(l) = (\mu_{il}, \sigma_{il}^2)$.

It is known that feature selection or dimensionality reduction often improves performance of classification problems [104] (and many other papers) and cardiovascular based biometrics [69, 59, 21, 20]. We use a robust informative feature selection method to select informative time-frequency bins for verification and recognition. The two key elements considered in our feature selection method are distinguishability and stability. The feature should help distinguish the individual from a reasonably large subset of other individuals, and it should be stable across sessions. The l-th feature of the i-th individual is selected if the symmetric relative entropy, i.e. the symmetric Kullback-Leibler divergence, between $\mathcal{N}(\mu_{il}, \sigma_{il}^2)$ and the nominal distribution $\mathcal{N}(\mu_{0l}, \sigma_{0l}^2)$ is larger than a threshold $\kappa > 0$. The relative entropy between two densities p and q is defined by

$$D(p\|q) = \int p \log \frac{p}{q}, \tag{2.1}$$

where the integral is taken over the support set of p. The symmetric relative entropy between the two densities is defined as

$$d(p, q) = D(p\|q) + D(q\|p). \tag{2.2}$$

For the Gaussian distributions used in our model, the symmetric relative entropy between $\mathcal{N}(\mu_{il}, \sigma_{il}^2)$ and $\mathcal{N}(\mu_{0l}, \sigma_{0l}^2)$ is

$$d(\hat{\theta}_i(l), \hat{\theta}_0(l)) = \frac{\sigma_{il}^2 + (\mu_{il} - \mu_{0l})^2}{2\sigma_{0l}^2} + \frac{\sigma_{0l}^2 + (\mu_{il} - \mu_{0l})^2}{2\sigma_{il}^2} - 1, \tag{2.3}$$

where the nominal model is obtained by using the spectrograms of all the individuals in the database. Using the symmetric relative entropy for feature selection ensures that only those bins whose distributions are far from the nominal are selected for each individual, thereby ensuring distinguishability. Moreover, stability of features is enforced by the variance of the individual's bin σ_{il}^2. It follows from the construction of the nominal model that for the most part, $\sigma_{il}^2 < \sigma_{0l}^2$, so that the second term in equation (2.3), with σ_{il}^2 in the denominator, increases as σ_{il}^2 decreases; for individual bins with small variances, the symmetric relative entropy tends to be large.

The score of a test heartbeat using the i-th individual's model is given by the log-likelihood ratio (LLR)

$$\Lambda_i = \sum_{l=1}^{L} \log \left[\frac{p_i(Y(l)|\hat{\theta}_i(l))}{p_0(Y(l)|\hat{\theta}_0(l))} \right] I_{d(\hat{\theta}_i(l),\hat{\theta}_0(l))>\kappa}, \tag{2.4}$$

where $I_{\{\cdot\}}$ is the truth function indicating which time-frequency bins are selected; l is the index of the bins. For verification, the LLR given in equation (2.4) is compared with a threshold τ, so that if $\Lambda_i > \tau$, the heartbeat with the claimed identity is accepted, otherwise the heartbeat is rejected. For recognition, the LLR is computed for every individual model, and the individual whose model gives the largest score is declared. For rank-k recognition, individuals with models yielding the top k scores are declared. For across-session verification, the score function was modified so as to disregard the role of the variances of the time-frequency bins. That is, we set $\hat{\theta}_i(l)$ and $\hat{\theta}_0(l)$ to a constant θ.

In recognition, to ensure that a variable number of time-frequency bins can be selected for each individual's model, the score obtained from comparing a test heartbeat to an individual's model is normalized by a score obtained from comparing the heartbeat to the nominal model. This normalization ensures that there is no direct relationship between the number of bins used in an individual's model, and the value of the computed score [104]; more selected bins does not mean higher scores.

2.6 Results

2.6.1 Within Session Analysis

In this analysis, ECG signals of each individual collected on the same day are separated into various sizes of training and testing sets. The number of ECG signals used in training ranges

from 4 to 64, in powers of 2. After training each model, which includes feature selection, we tested on 1, 2, 4, 8, 16, 32, and 64 pulses. For each trained model and a given number of testing pulses t, we performed n-independent tests, where n is the number of non-overlapping pulse blocks of size t there are in 64 pulses; $n = \lfloor \frac{64}{t} \rfloor$. For recognition, the number of individuals in the database was varied from 20 to 260, in steps of 20, and the analysis was repeated for contiguous blocks of size m, where $m = 20 \times i, i = 1, 2, 3, \ldots, 13$. Figure 2.2 shows the authentication performance, represented by the equal error rates (EERs), for all the training and testing sets. The solid lines represent the cases of using feature selection. We see that an EER as low as 0.37% is attained when decisions are based on a single test heartbeat. As the number of test heartbeats increases, the EERs tend to decrease, reaching a nadir of 0.02%. Figure 2.3 shows the k-Rank recognition accuracy for training on 4, 8, 16, 32, 64 heartbeats and testing on a single heartbeat. A rank-1 recognition accuracy of 99% can be obtained, when decisions are based on a single heartbeat. As expected, the recognition accuracy increases as the rank increases. In general, recognition accuracy increases as the number of individuals in the database decreases. This is in agreement with our expectation; fewer individuals to choose from yield fewer opportunities for misclassification.

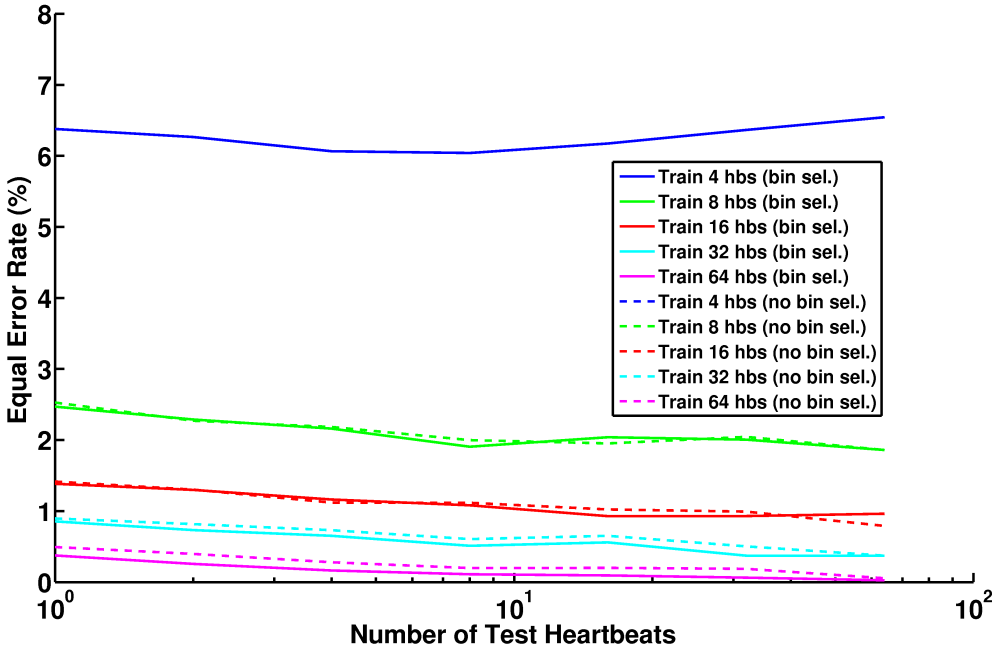


Figure 2.2: EER curves for training on 4, 8, 16, 32, and 64 heartbeats and testing on 1, 2, 4, 8, 16, 32, and 64 heartbeats, with and without feature selection.

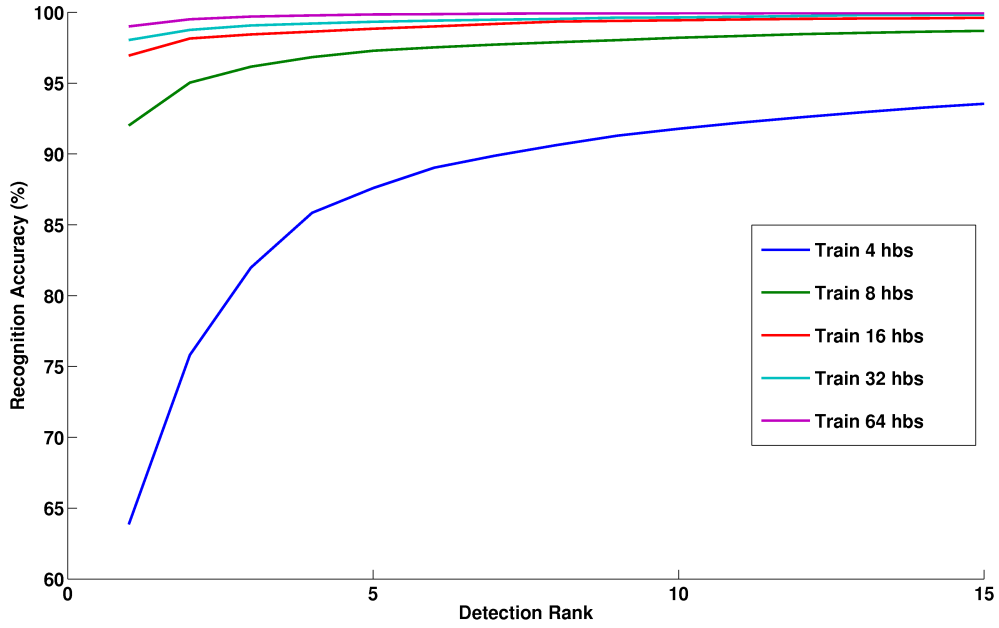


Figure 2.3: Rank-k recognition accuracy for training on 4, 8, 16, 32, and 64 heartbeats and testing on 1 heartbeat, with feature selection and 260 individuals in the database.

2.6.2 Cross Session Analysis I: One Training Session

In this analysis, training heartbeats are obtained from one session, and testing from another. The number of ECG signals used in training ranges from 4 to 128, in powers of 2. After training each model, which includes feature selection, we tested on 1, 2, 4, 8, 16, 32, 64, and 128 pulses. For each trained model and a given number of testing pulses t, we performed n-independent tests, where n is the number of non-overlapping pulse blocks of size t there are in 128 pulses; $n = \lfloor \frac{128}{t} \rfloor$. The biometric performances obtained by training on the first session, and testing on either the second or the third session are shown in Figure 2.4. We see that the performance based on testing on heartbeats from the second session tends to be better; ECG performance is negatively affected by the amount of time between training and testing heartbeats. This difference in ECG performance between testing on heartbeats from the second and third sessions might also be due to minor differences in electrode placement on the different days. To isolate time variability as the principal cause of the ECG performance differences, a study, where the electrode placement is highly standardized in all sessions, would have to be performed.

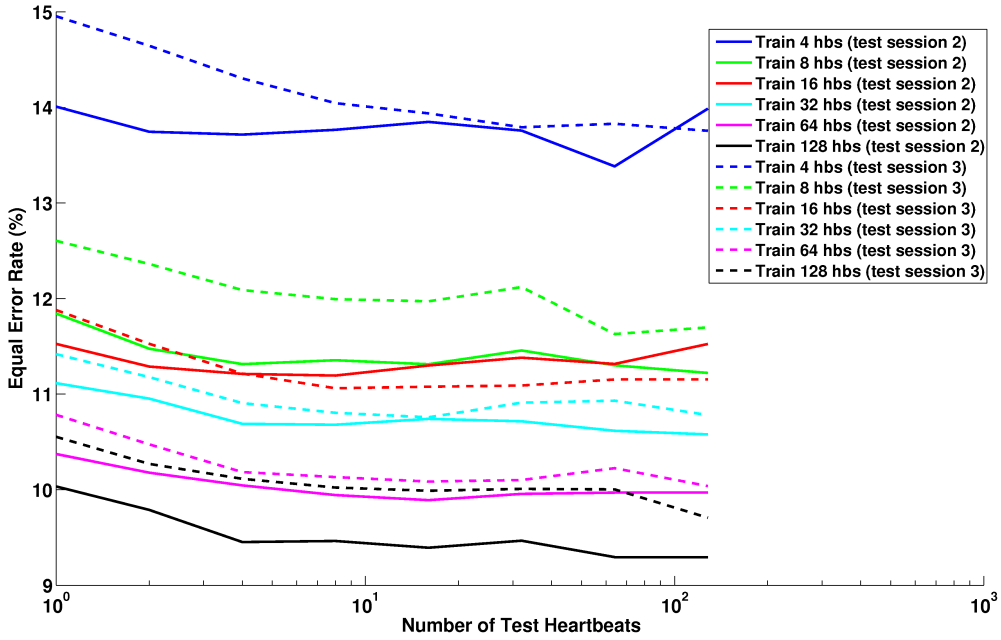


Figure 2.4: EER curves for training on 4, 8, 16, 32, 64, and 128 heartbeats and testing on 1, 2, 4, 8, 16, 32, 64, and 128 heartbeats, using feature selection, when testing on session 2 (solid lines) and session 3 data.

2.6.3 Cross Session Analysis II: Two Training Sessions

In this analysis, training heartbeats are obtained from two days, and testing, from the third. Simple data fusion that trains a combined model for each individual's bin is used. The number of ECG heartbeats used in training ranges from 8 to 256, in powers of 2, where half of the heartbeats come from one day, and the remaining from another. Figure 2.5 shows the EERs for all the training and testing sets, with (solid lines) and without feature selection. It is evident that feature selection improves the authentication performance in all cases. We see that an EER as low as 5.58% is attained when decisions are based on multiple test heartbeats. Figure 2.6 shows the EERs for testing on a session, when training is performed on heartbeats from either two sessions or a single previous session. We observe a two-fold improvement in authentication performance due to data fusion. Figure 2.7 shows the k-Rank recognition accuracy for training on 8, 16, 32, 64, 128, 256 heartbeats from two training sessions and testing on half the number of training heartbeats, respectively. A rank-1 recognition accuracy of 76.9% can be obtained, when decisions are based on multiple heartbeats. As expected, the recognition accuracy increases, as the rank increases, reaching a rank-15 value of 93.5%. When a verification analysis is performed on a database including only people of good health (those who did not report any heart disease or use of substances

that may affect the ECG signal), a 4.75% EER is achieved, when training and testing on 256 and 32 heartbeats respectively. Similarly, when the analysis is performed on a database including only unhealthy people, a 5.7% EER is achieved. The performance based on a database including only people of good health or only unhealthy people is comparable to the performance using the database including both health statuses. This shows that our proposed method is robust to different health statuses.

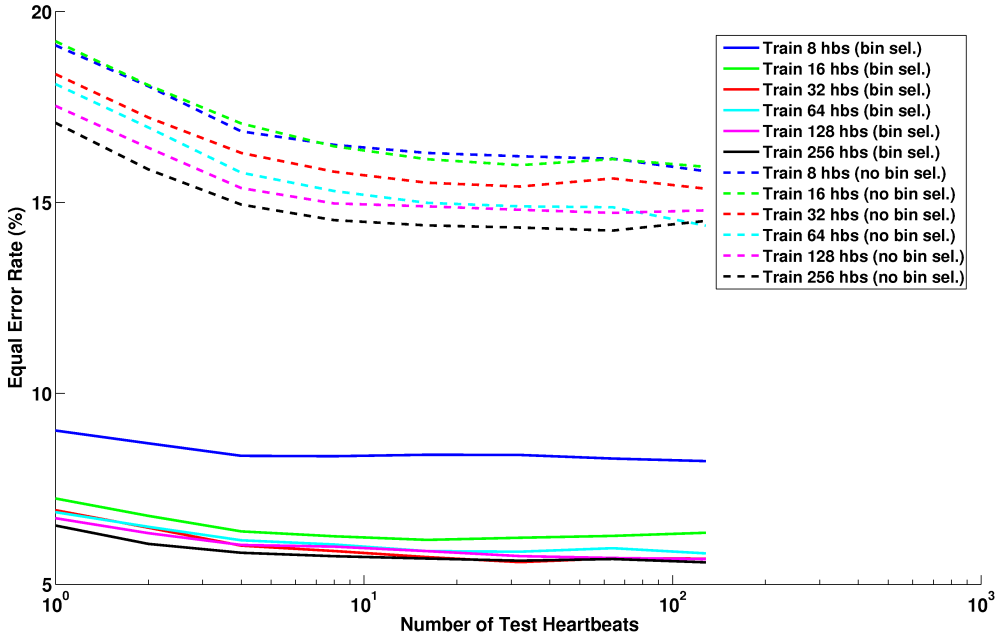


Figure 2.5: EER curves for training on 8, 16, 32, 64, 128, and 256 heartbeats and testing on 1, 2, 4, 8, 16, 32, 64, and 128 heartbeats, with and without feature selection.

2.7 Summary

In this chapter, a novel short-time frequency method with robust feature selection is proposed for use as an ECG biometric. This approach does not require extraction of fiducial points beyond the peak of the R wave. Rather, a systematic approach to feature generation and selection of informative features is presented.

Since the data comprises single lead ECG signals from 269 individuals that represent different demographic groups and health statuses, and were collected on three separate occasions over a seven-month period, we are able to study the impact of long-term variability, health status, data fusion, the number of training and testing heartbeats, and database size on ECG biometric performance. The method uses informative feature selection via symmetric

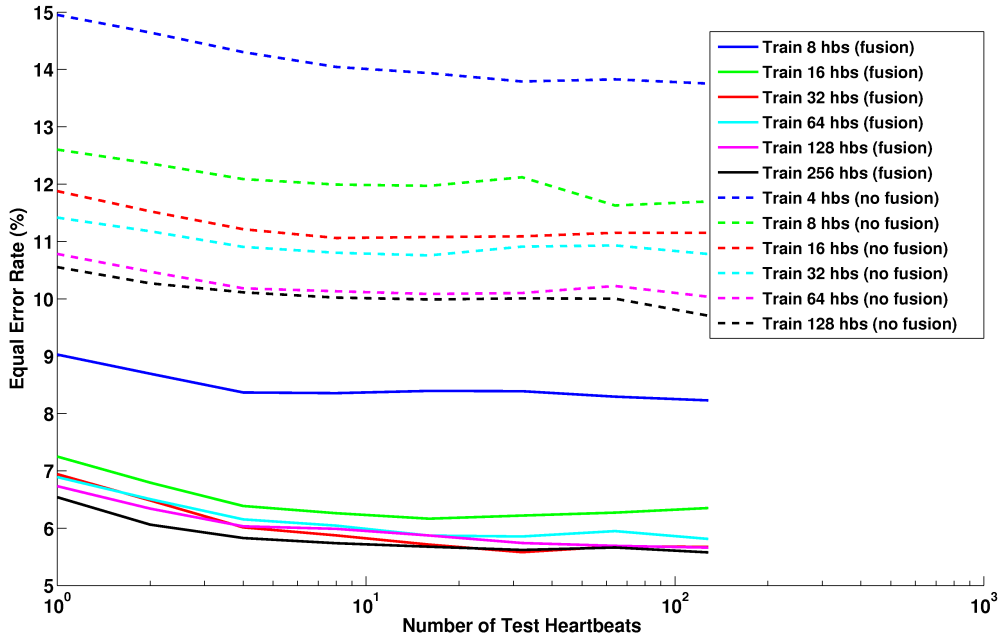


Figure 2.6: EER curves for testing on 1, 2, 4, 8, 16, 32, 64, and 128 heartbeats, using feature selection, when training on two sessions (solid lines) and a single session (dashed lines).

relative entropy, and fuses heartbeats from two training occasions to achieve 5.58% EER in verification, 76.9% accuracy in rank-1 recognition, and 93.5% accuracy in rank-15 recognition when testing heartbeats are measured on another day. When the heartbeats used for training and testing are recorded on the same day, an EER of 0.37% and a recognition accuracy of 99% is attained, for decisions based on a single heartbeat. When a similar analysis is performed on a database including only people of good health, or only unhealthy people, the authentication performance and recognition accuracy are comparable to what we obtain using the database including both health statuses. This shows that our proposed method is robust to different health statuses.

The major contributions of this work are by way of analyzing a large realistic data set comprising several races, genders, age groups, and health statuses, proposing a novel robust method for capturing signal variability and information fusion, proposing a method that performs well using non-standard electrode placements, and demonstrating the effects of training and testing sample sizes on the authentication performance and recognition accuracy of ECG biometrics.

The MIT databases include ECG recordings from a single day for each individual, permitting only within session analysis. On the other hand, the PTB database has ECG recordings from multiple days for some, but not all individuals. Since the ECG recordings from the MIT and

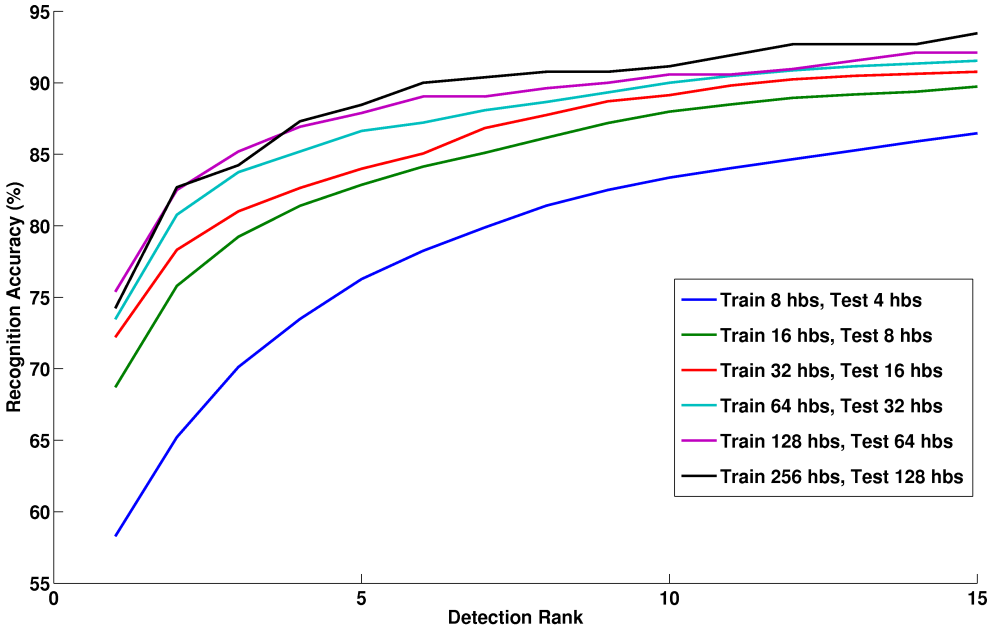


Figure 2.7: Rank-k recognition accuracy for training on 8, 16, 32, 64, 128, and 256 heartbeats from two training sessions, and testing on half the number of training heartbeats, respectively, using feature selection, with 260 individuals in the database.

the PTB databases do not directly lend themselves to our multiple sessions data protocol, it would be easier to apply the methods proposed by authors of previous ECG biometric papers to our database, in order to compare the biometric performance of our proposed method to those of other authors.

2.8 Lessons Learned

The proposed ECG biometric system suffers from performance degradation across repeated observations, mainly due to false rejection errors. That is, the individual looks a lot less like the originally created template.

It is difficult to make generalizable performance claims using a small data set. This renders a direct comparison between the performance of our method with those of others difficult. Moreover, this is compounded by data sets that include individuals from one demographic (e.g. only healthy individuals) and those can contain a mixture of within-session and across-session samples.

Chapter 3

A Comprehensive Review of ECG Biometric Methods

This chapter appeared in a previous IEEE journal publication [88].

3.1 Highlights

The electrocardiogram (ECG) is an emerging biometric modality that has undergone about fifteen years of development, in peer-reviewed literature, and as such deserves a systematic review and discussion of the associated methods and findings. In this dissertation, we review most of the techniques that have been applied to the use of the electrocardiogram for biometric recognition. In particular, we categorize the methodologies based on the features and the classification schemes. Finally, a comparative analysis of the authentication performance of a few of the ECG biometric systems is presented, using our in-house database. The comparative study includes the cases where training and testing data come from the same and different sessions (days). The authentication results show that most of the algorithms that have been proposed for ECG based biometrics perform well, when the training and testing data come from the same session. However, when training and testing data come from different sessions, a performance degradation occurs. Multiple training sessions were incorporated to diminish the loss in performance, but only a few of the proposed ECG recognition algorithms appear to be able to support performance improvement due to multiple training sessions. Only three of these algorithms produced equal error rates (EERs) in the single digits, including an EER of 5.5% using a method proposed by us.

3.2 Introduction

Inherent in the concept of identity is both the permanence and uniqueness of every individual. Biometric systems that perform identity recognition on the basis of informative data collected from an individual are vital for security. Such systems may take many forms, varying with respect to the type and quantity of data collected, specific algorithms used, and operational modes. Biometric identity recognition is attractive because data are measured from the body itself, and are ubiquitous (or nearly so). Various biometrics have been proposed for use in identity recognition, such as fingerprint, iris, face, and speech. These biometrics each have operational tradeoffs in terms of performance, measurability (ease of collecting data), and circumvention (ease of replication), and acceptability [11, 23, 8, 127].

Recently, cardiovascular signals have been studied for use in identity recognition problems, using electrocardiography [10, 67, 51] and carotid laser Doppler vibrometry (LDV) [69, 59, 20]. These signals differ from the signals mentioned above in that they are intrinsically connected to critical biological function. Circumvention is significantly more difficult with these biometrics, and measurability is nearly always guaranteed. Significant challenges remain to incorporate this information into successful recognition systems.

An ECG is a recording of the electrical activity of the heart. Electrodes placed on the surface of the body are used to measure the electrical signals originating from the myocardium, the heart muscle. The ECG consists of three main components: P wave, QRS complex, and T wave. The P wave occurs due to atrial depolarization, the QRS complex, due to ventricular depolarization, and the T wave, due to ventricular repolarization.

Identity recognition based on the ECG dates back to the pioneering work of Biel *et al.* [10], Irvine *et al.* [51], Kyoso and Uchiyama [67]. The premise of these and other studies is that the ECG contains sufficiently detailed information pertaining to the electrical operation of the heart, and that the nature of this activity is highly personalized. The ECG signal will be highly individualized, insofar as it depends on functional and structural properties including conductivity of the heart and other tissue. The main hypothesis shared by these studies is that *the detailed electrical activity of the heart, as captured by the ECG, is of sufficient quality to be used in high performance identity recognition systems.*

In this chapter, we summarize existing methods from the literature on identity recognition systems based on the ECG. The use of the ECG in this setting has three key properties [87]:

- ECG signals are difficult to counterfeit, in supervised conditions.
- The ECG signal is present in all living individuals.
- ECG signals provide additional information related to psychological states, and physiological and clinical status, which may be of interest.

An identity recognition system typically operates in either the identification or authentication mode. In the identification mode, the system outputs the identity of an individual using the input data. In the authentication mode, the system accepts or rejects a claimed identity associated with the input data. When the claimed identity is wrongly rejected, the system is said to have incurred a false rejection error. When the claimed identity is wrongly accepted, a false acceptance error occurs.

The methods that have been proposed for ECG biometrics can be grouped based on the number of ECG data channels used, the operational setting, the method for generating desirable features, and the type of classifier adopted.

All the studies on ECG biometrics are based on one-channel, two-channel, three-channel, or 12-lead ECG signals. These are by far the most common lead configurations employed in clinical practice. Of these, ECG as a biometric based on a single channel is the most studied; simplicity plays a major role in this, since one does not have to worry about effective channel combination schemes. However, in this context, it is not clear that simpler is necessarily better. Many of the single-channel studies use the data from one of the channels of a standard 12-lead ECG recording. However, there are studies that use non-standard electrode placement techniques; Chan *et al.* [17] studied the biometric performance of one-channel ECG signals recorded from the pads of individuals' thumbs, Shen *et al.* [107] used single-channel ECG signals obtained from the palms, while Odinaka *et al.* [87] used single-channel ECG signals obtained from electrodes placed bilaterally across the lower rib cage. Wübbeler *et al.* [130] used data from three channels, while Agrafioti *et al.* [6] fused data from all 12 ECG channels for recognition purposes.

Most studies of ECG biometrics require the segmentation of an ECG recording into single heartbeat signals. One of the reasons for this is the ease of signal alignment, which leads to coherent feature extraction. Another reason for segmenting ECG recordings into single heartbeat signals is that the variations across individuals within one cardiac cycle is thought to be sufficient in discriminating among them. Exceptions include the works of Plataniotis *et al.* [96], Agrafioti *et al.* [4], Li and Narayanan [73], and Loong *et al.* [74]. Plataniotis *et al.*

and Agrafioti *et al.* used features based on the autocorrelation of non-overlapping segments of the ECG recording, while Li and Narayanan, and Loong *et al.* used linear frequency cepstral coefficients (LFCC) and linear predictive coding (LPC) spectral coefficients of overlapping segments of the ECG recording, respectively.

ECG signals obtained during normal resting conditions have been investigated by most studies [10, 106, 87]. In addition, there have been studies designed to test the feasibility of ECG biometrics during changes in emotional and mental states [53], exercise [63], and benign cardiovascular conditions [5].

Based on the features that are extracted from ECG signals, we can classify ECG biometric methods as either fiducial-based, non fiducial-based, or a hybrid. The fiducial-based methods extract temporal, amplitude, area, angle, or dynamic (across heartbeats) features from characteristic points on the ECG signal. The features include but are not limited to the amplitudes of the P, R, and T waves, the temporal distance between wave boundaries (onset and offset of the P, Q, R, S, and T waves), the area of the waves, and slope information [10, 67, 106, 53, 48]. The non fiducial-based methods do not use the characteristic points as features. Instead, features like wavelet coefficients [17] and autocorrelation coefficients [96] are utilized. The hybrid methods combine both fiducial-based features and non fiducial-based features [126, 109].

There have been several choices of classifiers in the literature, ranging from simple nearest neighbor/center classifiers [4] to neural networks [124] and support vector machines [73]. Previous literature reviews of some of the methods that have been applied to ECG based identification can be found in the reports of Nasri *et al.* [86], Sufi *et al.* [119], Chauhan *et al.* [18], and Israel and Irvine [52].

The rest of this chapter is organized as follows. In Section 3.3, we review the ECG biometric methodologies; in Section 3.4, we provide a comparative analysis of a few of the methods using our in-house multi-session database. Section 3.5 is devoted to open issues in ECG recognition, while Section 3.6 provides a summary. Appendix A describes some of the ECG databases that have been utilized in the literature.

3.3 Survey of ECG Recognition Methods

Most studies in ECG biometrics have employed single-channel ECG signals following the work of Biel *et al.* [10], who showed that a single channel contains sufficient information to support biometric recognition. However, some studies have adopted multiple channels in an effort to improve performance. Such studies include those that utilized two channels [134], three channels [130, 33], and 12 channels [10, 6].

A host of feature extraction and reduction techniques have been proposed for use in ECG recognition. Moreover, different types of classifiers have been utilized for placing test feature vectors into predefined classes. Table 3.1 shows most of the methodologies that have been proposed for ECG recognition, in chronological order of publication. Within a year, the studies are arranged in lexicographical order based on the first author's last name. Similarly, the abbreviation descriptors are ordered in lexicographical order.

3.3.1 Categorization Based on Features

Algorithms Based on Fiducial Features

Algorithms based on fiducial features use the characteristic points — wave onset, peak (minima or maxima), and offset, extracted from an ECG trace to generate the feature set. We define characteristic points to be the actual points located on an ECG trace and fiducial features to be the features that are derived from these characteristic points. For example, the peak of the R wave is a characteristic point, while the time difference between the peaks of the R and T waves, the RT interval, is a fiducial feature. There are four types of fiducial features [113] that have been used for ECG based recognition: temporal, amplitude, angle, and dynamic (R-R intervals). Several subsets of these fiducial features have been used in the literature [10, 67, 66, 51, 50, 53, 48, 94, 63, 135, 111, 112, 38, 44, 123].

Algorithms Based on Non-fiducial Features

Algorithms based on non-fiducial features do not use the characteristic points for generating the feature set. Instead, some of the algorithms use one or more of the characteristic points for heartbeat segmentation [17, 35], while others do not use the characteristic points at all, but

Table 3.1: Summary of existing literature on ECG recognition
FT = Feature Type, CT = Classifier Type, AP = Authentication Performance, IP = Identification Performance

Study	Sample Size n	Technique	FT	CT	Session Type	AP (%)	IP (%)
Biel *et al.* [10]	20	10 fiducials + PCA + GMC	F	GMC	Multiple days	-	100
Irvine *et al.* [51]	5	3 temporal fiducials + MD	F	LDAC	Multiple states	-	-
Kyoso and Uchiyama [67]	9	4 temporal fiducials + GMD	F	LDAC	Single day	-	>90
Shen *et al.* [106]	20	7 fiducials and HB + TMCC + DBNN (HC)	H	NN	Single day	-	100
Irvine *et al.* [50]	104 (WS), 95(AS)	15 temporal fiducials + WLFS + LDAC	F	LDAC	Multiple states	-	91(WS), 88(AS)
Kyoso[66]	21	9 fiducials + GMD	F	LDAC	Single day	-	-
Palaniappan and Krishnan [94]	10	6 fiducials + MLP-BP or SFA	F	NN	Single day	-	97.6(MLP-BP)
Israel *et al.* [53]	29	15 temporal fiducials + WLFS + LDAC	F	LDAC	Multiple states	-	97-98(AST), 100(AEL)

Table 3.1: Summary of existing literature on ECG recognition (Continued)
FT = Feature Type, CT = Classifier Type, AP = Authentication Performance, IP = Identification Performance

Study	Sample Size n	Technique	FT	CT	Session Type	AP (%)	IP (%)
Kim *et al.* [63]	10	4 RHBTF + GMD	F	LDAC	Multiple states	-	-
Saechia *et al.* [102]	35	FTF + MLP-BP	NF	NN	-	-	97.15
Plataniotis *et al.* [96]	14	AC/DCT + NNC,ED or GLLC	NF	*k*NN	Single day	-	92.9(NNC), 100(GLLC)
Zhang and Wei [135]	520	14 fiducials + PCA + BC	F	GMC	Single day	-	97.4
Molina *et al.* [84]	10	MSRRS + ED	NF	NC	Multiple days	2(EER)	-
Silva *et al.* [109]	168	8 fiducials and SSMW + FSC + NNC,ED	H	*k*NN	Multiple states	-	99.63, 99.97(FSC)
Wübbeler *et al.* [130]	74	WHVD + NC	NF	NC	Multiple days	2.8(EER, 3 channels)	98.1(3 channels)
Agrafioti and Hatzinakos [6]	14	AC/LDA + NNC,ED	NF	*k*NN	Single day	-	100(12 leads)

Table 3.1: Summary of existing literature on ECG recognition (Continued)
FT = Feature Type, CT = Classifier Type, AP = Authentication Performance, IP = Identification Performance

Study	Sample Size n	Technique	FT	CT	Session Type	AP (%)	IP (%)
Agrafioti and Hatzinakos [4]	27	TMAC + AC/LDA or AC/DCT + NNC,ED (HC)	NF	kNN	Single day	<1(EER)	96.3(AC/DCT), 100(AC/LDA)
Agrafioti and Hatzinakos [5]	56	APC and PVC Screening + AC/LDA + NNC,ED	NF	kNN	Single day	∼5(EER)	96.42
Chan *et al.* [17]	60	SAECG + WDIST and CCORR	NF	NC	Multiple days	-	90.8
Chiu *et al.* [24]	45	DWTMSS + NC,ED	NF	NC	Single day	12.50(FAR), 5.11(FRR)	95.71(PHBIA)
Fatemian and Hatzinakos [35]	27	MRHB + CC	NF	HMS	Single day	-	99.63
Gahi *et al.* [38]	16	24 fiducials + IGRFS + NC,MD	F	LDAC	Single day	-	100
Irvine *et al.* [49]	39	EigenPulse + NC,ED	NF	NC	Single day	-	100
Khalil and Sufi [62]	15	LPE of QRS wave	NF	LMS	Single day	-	-

Table 3.1: Summary of existing literature on ECG recognition (Continued)
FT = Feature Type, CT = Classifier Type, AP = Authentication Performance, IP = Identification Performance

Study	Sample Size n	Technique	FT	CT	Session Type	AP (%)	IP (%)
Singh and Gupta [111, 112]	25, 50	19 fiducials	F	LMS, HMS	Single day	-	99
Sufi *et al.* [118]	15	PDM of P, QRS, T waves, NC	NF	NC	Single day	0(EER)	100
Wan and Yao [124]	23	DWTMRRS + MLP-BP	NF	NN	Multiple days	-	100 (n = 15)
Wang *et al.* [126]	13, 13	21 fiducials and RDHB + WLFS + LDAC or kNN (HC)	H	LDAC or kNN	Single day	-	100
Yao and Wan [132]	20	DWTMRRS + PCA	NF	LMS	Mixed	-	91.5
Agrafioti *et al.* [3]	10	Template Updating	NF	kNN	Multiple states	3.4(EER, HFU), 6.3(EER, MFU), 14.7(EER, LFU)	-
Boumbarov *et al.* [12]	9	HMMHS + PCA or LDA + RBFNN	NF	NN	-	-	∼86
Fang and Chan [33]	100	ECGPSR + SC or MNPD	NF	HMS or LMS	Single day	-	93(1 channel, MNPD), 99(3 channels, MNPD)

Table 3.1: Summary of existing literature on ECG recognition (Continued)
FT = Feature Type, CT = Classifier Type, AP = Authentication Performance, IP = Identification Performance

Study	Sample Size n	Technique	FT	CT	Session Type	AP (%)	IP (%)
Homer *et al.* [47]	12	GF and RARMA + NNC,ED	NF	*k*NN	Multiple states	-	∼85.2
Irvine and Israel [48]	29, 75	WSPR	F	GMC	Multiple states	∼0.01(EER, n = 29), >0.05(EER, n = 75)	-
Agrafioti and Hatzinakos [7]	52	AC/LDA, PT + NNC,ED	NF	*k*NN	Mixed	>10	92.3
Coutinho *et al.* [26]	26	QHB + ZMCP + MDL	NF	NC	Multiple states	-	100
Ghofrani and Bostani [40]	12	ARC, MPSD, HD, LE, AE + *k*NN or MLP-BP or PNN	NF	*k*NN or NN	Mixed	-	100
Jang *et al.* [57]	65	EigenPulse + Heartbeat Screening	NF	NC	Multiple states	-	>96.92
Li and Narayanan [73]	18	HPE + SVMLK	NF	SVM	Single day	0.55(EER)	98.11
Li and Narayanan [73]	18	LFCC + HLDA + GMM + LLRC or SVMGSV	NF	GMC or SVM	Single day	4.05(EER, LLRC), 2.5(EER, SVMGSV)	94.78(LLRC), 95.9(SVMGSV)

Table 3.1: Summary of existing literature on ECG recognition (Continued)
FT = Feature Type, CT = Classifier Type, AP = Authentication Performance, IP = Identification Performance

Study	Sample Size n	Technique	FT	CT	Session Type	AP (%)	IP (%)
Li and Narayanan [73]	18	SLF [SVMGSV + (HPE + SVMLK)]	NF	SVM	Single day	0.5(EER)	98.26
Loong *et al.* [74]	15	LPCS + MLP-BP	NF	NN	Single day	-	100
Odinaka *et al.* [87]	269	log-Normal spectrogram	NF	GMC	Multiple days	0.37(EER, WS), 5.58(EER, ASWF)	99(WS), 76.9(ASWF)
Venkatesh and Jayaraman [123]	15	6 fiducials + FLDA,kNN + NNC,DTW	F	kNN	Single day	-	100
Yao and Wan [133]	30	DWTMRRS + BMS + WDIST	NF	NC	Mixed	-	>80
Ye *et al.* [134]	18, 18, 47, 65	DWT and ICA + SVMRBK	NF	SVM	Single day	-	99.6(2 channels, PHBIA, n = 47)
Lourenço *et al.* [75]	16	MANRHB + NC,ED	NF	NC	Single day	13(EER)	94.3
Safie *et al.* [103]	112	PAR + NC,ED	NF	NC	Mixed	9.89	-

Table 3.1: Summary of existing literature on ECG recognition (Continued)
FT = Feature Type, CT = Classifier Type, AP = Authentication Performance, IP = Identification Performance

Study	Sample Size n	Technique	FT	CT	Session Type	AP (%)	IP (%)
Shen [107]	168	17 fiducials and HB + NNC,WED (HC)	H	kNN	Single day	-	95.3
Tawfik *et al.* [120]	22	DCT of QRS + MLP-BP	NF	NN	Multiple days	-	99.09

AC/DCT = Auto-Correlation/Discrete Cosine Transform; AC/LDA = Auto-Correlation/Linear Discriminant Analysis; AE = Approximation Entropy; AEL = Across Electrode Locations; APC = Atrial Premature Contraction; ARC = AutoRegression Coefficients; ARTMAP = Supervised (MAP) Adaptive Resonance Theory; AS = Across Sessions; AST = Across Stress Tasks; ASWF = Across Sessions With Fusion; BC = Bayes' Classifier; BMS = Birge-Massart Strategy; CC = Cross Correlation; CCORR = Cross CORRelation measure; DBNN = Decision-Based Neural Network; DCT = Discrete Cosine Transform; DTW = Dynamic Time Warping; DWT = Discrete Wavelet Transform; DWTMRRS = DWT of Mean R-R Segments; DWTMSS = DWT of Mean Synthetic Signal; ECGPSR = ECG Phase Space Reconstruction; ED = Euclidean Distance; EER = Equal Error Rate; F = Fiducial; FAR = False Accept Rate; FLDA = Fisher's LDA; FRR = False Reject Rate; FSC = Feature Selection Context; FTF = Fourier Transform Features; GF = Gaussian Fit; GLLC = Gaussian Log-Likelihood Classifier; GMC = Generative Model Classifier; GMD = Generalized Mahalanobis Distance; GMM = Gaussian Mixture Model; H = Hybrid; HB = Heart Beat; HC = Hierarchical Classification; HD = Higuchi Dimension; HFU = High Frequency Updating; HLDA = Heteroscedastic LDA; HMMHS = Hidden Markov Model Heartbeat Segmentation; HMS = High Match Score; HPE = Hermite Polynomial Expansion; ICA = Independent Component Analysis; IGRFS = Information Gain Ratio Feature Selection; kNN = k Nearest Neighbors

Table 3.1: Summary of existing literature on ECG recognition (Continued)
FT = Feature Type, CT = Classifier Type, AP = Authentication Performance, IP = Identification Performance

Study	Sample Size n	Technique	FT	CT	Session Type	AP (%)	IP (%)

LDA = Linear Discriminant Analysis; LDAC = LDA Classifier; LE = Lyapunov Exponent; LFCC = Linear Frequency Cepstral Coefficient; LFU = Low Frequency Updating; LLRC = Log-Likelihood Ratio Classifier; LMS = Low Match Score; LPCS = Linear Predictive Coding Spectrum; LPE = Legendre Polynomial Expansion; MANRHB = Mean of Amplitude-Normalized Resampled Heart Beats; MD = Mahalanobis Distance; MDL = Minimum Description Length; MFU = Medium Frequency Updating; MLP-BP = Multi-Layer Perceptron Back-Propagation neural network; MNPD = Mutual Nearest-Point Distance; MPSD = Mean Power Spectral Density; MRHB = Median of Resampled Heart Beats; MSRRS = Morphological Synthesis of R-R Segments; NC = Nearest Center; NF = Non-Fiducial; NN = Neural Network; NNC = Nearest Neighbor Classifier; PAR = Pulse Active Ratio; PCA = Principal Component Analysis; PDM = Polynomial Distance Measurement; PHBIA = Per Heart Beat Identification Accuracy; PNN = Probabilistic Neural Network; PT = Periodicity Transform; PVC = Premature Ventricular Contraction; QHB = Quantization of Heart Beat; RARMA = Residual Auto-Regressive Moving Average; RBFNN = Radial Basis Function Neural Network; RDHB = Reduced-Dimension Heart Beat(via PCA or LDA); RHBTF = Resampled Heart Beats; SAECG = Signal-Averaged ECG; SC = Spatial Correlation; SF = Score Function; SFA = Simplified Fuzzy ARTMAP; SLF = Score Level Fusion; SSMW = Sub-Sampled Mean Wave; SVM = Support Vector Machine; SVMGSV = SVM GMM Super Vector; SVMLK = SVM Linear Kernel; SVMRBK = SVM Radial Basis Kernel; TMAC = Template Matching based on Auto-Correlation coefficients; TMCC = Template Matching based on Cross Correlation; WDIST = Wavelet DISTance; WED = Weighted Euclidean Distance; WHVD = Wübbeler's Heart Vector Distance; WLFS = Wilks' Lambda Feature Selection; WS = Within Session; WSPR = Wald's Sequential Probability Ratio; ZMCP = Ziv-Merhav Cross Parsing

segment the ECG recording into segments that may be overlapping [73] or non-overlapping [96, 6].

Most of the methods based on non-fiducial features require the detection of the R peaks for heartbeat segmentation and alignment [24, 102, 17, 84, 130, 49, 57, 132, 124, 133, 47, 33, 12, 26, 134, 73, 87].

There are a few methods that not only require the detection of the R peaks, but also some other characteristic points such as the onset and peak of the P wave, the onset and end of the QRS complex, and the peak and end of the T wave [35, 75, 103, 120]. Some methods require the detection of all or a subset of the three major components of each heartbeat (P wave, QRS complex, and T wave) for feature extraction [102, 118, 62].

However, there are other methodologies that do not extract any characteristic points, but rather segment the entire ECG trace into non-overlapping or overlapping windows, and extract features from those windows [96, 6, 4, 5, 3, 117, 73, 74, 7, 40, 39].

Algorithms Based on Hybrid Features

There are a few algorithms that use non-fiducial features for pruning the match space and fiducial features for the final classification [106, 107], while there are others that combine fiducial features with non-fiducial features to create the feature set [126, 109, 116].

3.3.2 Categorization Based on Classifier

Based on the type of classifier utilized, one can divide ECG recognition methodologies into seven groups: k nearest neighbors, nearest center, LDA, neural networks (NNs), generative model classifiers (GMCs), support vector machines (SVMs), and others. Of these groups, the most frequently used in the ECG recognition literature are nearest neighbor, nearest center, and LDA.

kNN Classifiers

k nearest neighbors (kNN) classifiers, which include the nearest neighbor classifier (NNC) as a special case ($k = 1$), is the most frequently used classifier type in the ECG recognition

literature [96, 109, 6, 4, 5, 126, 47, 40, 123, 107]. It involves comparing a feature vector to a collection of feature vectors, and selecting the top k vectors that produce the best match.

Nearest Center Classifiers

A nearest center classifier can be seen as a special kind of nearest neighbor classifier, where a representative training feature vector is created during training, as opposed to using the entire training feature vector set [130, 84, 17, 24, 49, 118, 26, 57, 133, 75, 103].

LDA Classifiers

Classification based on linear discriminant analysis, a special case of generative model classifiers (GMCs), has been used by several studies in the ECG biometrics literature [51, 67, 50, 66, 53, 63, 38, 126].

Neural Network Classifiers

Neural network classifiers have been used extensively for classification, because of their ability to learn complex relationships between the feature vectors in the training set. The most commonly used neural network for ECG biometric recognition is the multi-layer feedforward (perceptron) neural network [94, 102, 124, 40, 74, 120]. Other neural networks that have been used in the literature include decision-based neural network (DBNN) [106] and radial basis function neural network (RBFNN) [12].

Generative Model Classifiers

Generative model classifiers depend on modeling the distribution of the feature vectors. The estimated models are later used for classification. These classifiers include the log-likelihood ratio (LLR) [96, 48, 73, 87], Bayes' classifier [135], and SIMCA [10].

SVM Classifiers

Support vector machines have also been used in a few studies to find the linear boundaries between classes, after projecting the feature vectors in the training set to a high (possibly infinite) dimensional space. Ye *et al.* [134] used an SVM based on a Gaussian radial basis kernel. Also, Li and Narayanan [73] used SVMs based on a linear kernel for classification.

Match Score Classifiers

This category of algorithms include those that cannot be strictly put into any of the six groups described above. Most of the algorithms in this category depend on the computation of match scores based on the similarity (cross correlation [35], spatial correlation [33]) or the dissimilarity (MNPD [33]) between a feature vector and a stored template/model.

When similarity is sought, during identification, the template that gives the highest match score is associated with the test (probe) signal; during authentication, the score is compared to a threshold and the claimed identity is accepted, if the score is greater. On the other hand, when dissimilarity is desired, during identification, the template that gives the lowest match score is associated with the test (probe) signal; during authentication, the score is compared to a threshold and the claimed identity is accepted, if the score is smaller.

Examples of classifiers that seek a similarity between the feature vector and the stored template include template matching algorithms [106, 4, 35, 33, 112]. Some classifiers find the dissimilarity between the feature vector and the stored template [62, 111, 132, 33].

3.3.3 ECG in a Multimodal Framework

ECG has been used in combination with other modalities (biometric or non-biometric) in multimodal systems, either as a means of liveness detection, to prevent replacement attacks in a continuous monitoring setting, or to improve overall biometric performance.

ECG has been used in combination with other modalities for human recognition. Israel et al. [54] combined ECG features with those from the face to enhance identification performance. Fatemian *et al.* [34] combined the ECG and phonocardiogram (PCG) at the decision level, to obtain an improvement in recognition performance. Moreover, ECG has been combined

with electroencephalogram (EEG) [99] for human recognition purposes. ECG has also been combined with accelerometer readings for continuous authentication in a remote health monitoring setting [116]. Damousis et al. [29] looked at ECG as part of a larger framework of multiple biometrics including face and voice.

Agrafioti *et al.* [2] proposed an identity management system based on the ECG signals recorded as part of a body area sensor network (BAN). To avoid misclassifications, which can have drastic impacts on the routing and storing of health information, a two-stage identification scheme was implemented, where the AC/LDA system [4] was used in the first stage to obtain a ranked list of the top matching individuals to the probe ECG signal. In the second stage, a fuzzy commitment scheme (a key-binding authentication method) was used for validation — to select the best matching individual from the ranked list of individuals.

3.4 Comparative Analyses and Results

Using our in-house database [87], we studied the authentication performance of some of the methodologies that have appeared in the ECG recognition literature. The criteria for selecting the algorithms include: uses non-fiducial features, and requires a reasonable amount of training and testing times. There are several reasons for not implementing methodologies that use fiducial features including [49]:

- Variability among standards for detection of some characteristic points, such as the onset and offset of the component waves of the ECG.
- Location of some characteristic points are disproportionately affected by the presence of noise, even using a fixed fiducial detector.
- Difficulty in defining the boundaries and peaks of atypical heartbeats usually leads to an increased failure to enroll.
- Problems with generalizability to larger databases, when the number of features are limited.

We studied the effect of variability across different measurement times on the performance of the biometric system, and how to improve the performance by fusing information from

multiple sessions. Although considerable variation can occur in an ECG recording over the course of a day, nonetheless, we define multiple sessions as data collected on different days.

The authentication performance of the ECG biometric systems was evaluated using equal error rate (EER) and the detection error trade-off (DET) curve. A DET curve is a plot of the error rates on each axis [82]; EER is a point on the DET curve where the false acceptance (match) rate (FAR) equals the false rejection (non-match) rate (FRR) [80]. We consider three operational scenarios using ECG recordings obtained on three different days [87]:

- Training and testing on session 1 (within-session analysis).
- Training on session 1 and testing on session 3 (across-session analysis without fusion).
- Training on sessions 1 and 2, and testing on session 3 (across-session analysis with fusion). In this scenario, the training data from both sessions are simply pooled.

For each of the scenarios, we consider using a varying number of heartbeats (or chronological time equivalent) for training and testing. For the top performing (in terms of EERs) algorithms, we also present the DET curves, for across-session (with fusion) analysis.

All the methodologies were implemented to follow the descriptions provided by the authors. Exceptions are the works of Fatemian and Hatzinakos [35], and Yao and Wan [132, 124]. Fatemian and Hatzinakos used a stationary wavelet transform (SWT) to reconstruct the signal part of the raw ECG recording after which a moving window was used to smooth the signal. As implemented by us, a 1-40 Hz band-pass filter was used because there was significant baseline wander uncorrected by the SWT preprocessing. Yao and Wan used a wavelet-based de-noising approach, using hard thresholding, for preprocessing. We employed a 1-40 Hz band-pass filter for preprocessing, in the place of wavelet-based de-noising, because the de-noising approach led to a noticeable distortion of the ECG recording. Also, in addition to implementing the algorithm proposed by Molina *et al.* [84], we also implemented a slight modification of the algorithm, where band-pass filtering, using a 1-40 Hz frequency band, supplanted the morphological baseline wander removal technique. In the tables that follow, "Molina (M)" will be used to represent this modified version. Fang and Chan [33] implemented an algorithm that creates a three-dimensional ECG portrait, and finds the similarity or dissimilarity between the ECG portraits during authentication. In particular, we implemented the algorithm that finds the similarity between portraits via spatial correlation (SC), which is suboptimal to the one which finds the dissimilarity between portraits through

the mutual nearest point distance (MNPD). MNPD was not utilized because it does not satisfy the timing constraint.

Moreover, the algorithm proposed by Wübbeler *et al.* [130] uses three channels, but is adapted here to use a single channel.

3.4.1 ECG Data Set

We used the same database from previous work [87], with the exclusion of data from four individuals for whom the ECG signal was technically flawed. The current database has 265 individuals (121 males and 144 females). The ages of the individuals varied from 18 to 66 years, with mean and standard deviation of 38.9 and 14.1 years, respectively. As is typical in a community sample of this sort, 40.4% of the individuals self-reported some heart-related disease, including hypertension. 46.8% of the individuals reported using medicines or other substances that may affect the ECG signal. 28.3% of the individuals were healthy and did not use substances that may affect the ECG signal. 72.1% of the individuals were Caucasians.

The ECG signals were obtained from a single channel, with the electrodes placed bilaterally on the lower rib cage. This differs from that of the X channel, in the standard orthogonal (Frank) lead configuration, where the electrodes are located at the mid axillary lines in the fourth intercostal space [25]. The ECG signal acquired at the lower rib cage location has strong R and P wave components and is less affected by movement artifact than some of the conventional ECG channels. Additionally, these electrode sites have an advantage over standard electrode sites in terms of test subject acceptability; the individual does not have to disrobe before the signal can be acquired.

The signals were recorded with a Biopac TEL-100 system, using a 0.5 Hz high-pass filter and 500 Hz low-pass filter. The signals were further filtered to match the frequency range stipulated by each author whose method(s) we implemented. The individuals were asked to sit for five minutes as the recording took place. Each individual had three ECG recordings, each taken on a different day. We will call these three recordings, sessions 1, 2, and 3, where the separation between two consecutive sessions ranged from one week to six months. The mean and standard deviation of the time interval between sessions 1 and 2 were 15.4 and 14.4 days, respectively; for sessions 2 and 3, they were 47.2 and 79.4 days, respectively.

3.4.2 Results

The results presented here were obtained based on implementing the methodologies as described in the literature. The within-session analysis results are given in Table 3.2, which shows each algorithm, the authentication performance reported in the cited paper (if any), and its performance using our database. In the table, FS and NFS stand for "feature selection" and "no feature selection" respectively [87]; FS and NFS correspond to the cases where relative entropy based feature selection is or is not used, respectively. Moreover, "train 8, test 8" represents using 8 heartbeats (or 8 s, for the cases of Agrafioti *et al.* and Wang *et al.*) for training and the same number for testing. From the table, we can see that most algorithms do a decent job in modeling the ties within a class (individual) and discriminating between individuals. However, for some algorithms there are noticeable differences between the authentication performance reported in the literature and what we obtained using our database.

The original algorithm proposed by Molina *et al.* [84] uses a morphological baseline wander removal technique during preprocessing, which introduces distortions in the ECG recording; when band-pass filtering was used for preprocessing instead, the authentication performance improved. Also, the polynomial-based algorithm proposed by Sufi *et al.* [118] suffers from performance deficiencies compared to what was reported in the literature. This is likely due to the large sample size we used for this study; only 15 individuals were used in the original study performed by the authors. When the first 15 individuals from our database were used for the biometric study, an equal error rate of 0.95% was obtained.

The same phenomenon holds true for the algorithm proposed by Coutinho *et al.* [26]. The original study performed by the authors used ECG data obtained from 26 individuals. When the first 26 individuals from our database were used for the biometric study, an equal error rate of 0% was obtained, in comparison to the much higher rates (in the range of 35%) observed when applied to our full database of 265 individuals. The algorithm proposed by Yao and Wan [132] did not perform as well as some of the other methodologies. One possible reason for this is that only a single principal component was used for classification. The principal component may not be adequate to completely separate overlapping classes in the feature space. In general, when training and testing data come from the same session, most algorithms are good at accepting a true identity and rejecting a false one, as evidenced by their within-session authentication performance.

However, when training and testing are on different days, all the algorithms suffer deterioration in performance, as is reflected in Table 3.3. In the table, "train 32, test 16" represents using 32 heartbeats (or 32 s) from session 1 for training and using 16 heartbeats (or 16 s) from session 3 for testing.

The results for across-session testing, when the training data are obtained from two different days is given in Table 3.4. In the table, "train (8+8), test 16" represents using 8 heartbeats (or 8 s) each from sessions 1 and 2 for training and using 16 heartbeats (or 16 s) from session 3 for testing. Cross-session training is vital to the improvement of biometric performance as it accommodates variability across different measurement times in the model.

Comparing the last columns in Table 3.3 and Table 3.4, where a total of 32 heartbeats (or 32 s) are used for training, and 16 heartbeats (or 16 s) are used for testing, we can see the effect of fusing data from more than one session during training, on the authentication performance. With the exception of the algorithm by Fang and Chan [33], all the algorithms show a varied degree of improvement in performance, which can be attributed to data fusion. The most remarkable improvement in performance can be seen in the algorithm by Odinaka *et al.* [87] and Wan and Yao [124], where data fusion accounts for about a 52% and 67% change in EER, respectively.

Based on the across-session performance when fusion is used, we can see that a few of the methodologies provide the framework to capture the variability across time during training and provide for a significant improvement in authentication performance. It seems that most of the methods under review do not directly extend to across-session scenarios. However, with further research on how to extend them to include across-session variability, many of the methods may perform well.

A precise explanation as to why some of the algorithms benefit greatly from fusion approaches, while others do not, would constitute an important contribution to the field. Unfortunately, such explanations are not always readily apparent. For example, the Wan and Yao [124] neural networking approach involves self-organizing networks. It is extremely difficult to deconstruct the inner workings of this type of algorithm. However, we suspect that the way the feature vectors are processed prior to passing them through the network may explain the improved performance of the neural network algorithm; Wan and Yao [124] concatenated feature vectors obtained from different individuals, alongside the concatenated feature vectors from the same individual, in training a neural network to learn the contrast between the concatenated feature vectors. Providing the neural network with examples from

multiple training sessions makes it better able to discriminate future test data. Moreover, the algorithm by Odinaka *et al.* [87] uses a generative model which was selected in large part because of its ability to capture the variability across sessions. That is, the model is inherently robust. As a result of the robustness, multiple training sessions lead to better estimates of the model parameters.

Figure 3.1 shows the detection error trade-off curves, plotted using a normal deviate scale as prescribed by Martin *et al.* [82], for the top performing (based on EERs) methodologies in the across-session (with fusion) analysis, when 16 heartbeats each from sessions 1 and 2 are used for training, and 32 heartbeats from session 3 are used for testing. The DET curves are plotted this way to spread out the curves that would otherwise be bunched up by receiver operating characteristic (ROC) curves [82].

3.5 Open Issues in ECG Recognition

Irvine and Israel [48] raised several issues that need to be resolved before biometric systems based on the ECG can be used in practice. These include heart rate variability due to mental, emotional, and physical changes, issues relating to sensor placement, scalability to larger populations, and the time-varying nature of the ECG signal.

The results we presented above are for ECG recordings obtained during a calm, seated resting condition, where heart rate changes during the signal measurement can be expected to be minimal. We should add that the database, in an albeit uncontrolled fashion, allows for substantial variability in factors which could significantly impact the state of the individual. For instance, we do not restrict the individuals' daily activities, food or substance consumption and medicinal use. Indeed, heart rate varied considerably across sessions for some individuals. In a real world setting, individuals will usually not be in a normal resting condition. As such, if ECG recognition is to become a reality, extensive studies have to be designed to assess the effects of heart diseases, mental and physical stressors, exercise [63, 19], and other factors including common drugs, medications, and diet. To our knowledge, the work by Agrafioti *et al.* [5], is the only one that examines the use of the ECG for biometric recognition in a cardiac irregularity condition.

Heart rate variability due to physical, mental or other stressful activities can have the effect of changing the morphology of the ECG. A few methods have been proposed to compensate for such changes, such as re-sampling the entire heartbeat [63] or ST segment [35, 120] and

Table 3.2: Authentication performance for within session analysis

Researchers	Equal Error Rates (%)				
	Literature	Train 8, Test 8	Train 16, Test 16	Train 32, Test 32	Train 64, Test 64
Agrafioti *et al.* [4]	0.6	3.88	0.85	0.57	0.38
Chan *et al.* [17]	-	5.82	3.84	3.02	2.26
Chiu *et al.* [24]	0.83 - 0.86	4.15	2.64	1.76	1.01
Coutinho *et al.* [26]	-	42.54	38.92	35.14	33.34
Fang and Chan (SC) [33]	-	19.81	19.10	19.03	18.82
Fatemian and Hatzinakos [35]	-	8.69	5.99	4.37	2.26
Irvine *et al.* [49]	-	2.25	1.74	1.26	0.69
Khalil and Sufi [62]	-	5.28	2.64	1.56	1.13
Li and Narayanan (HPE+SVM) [73]	0.55	2.19	1.24	1.17	0.96
Lourenço *et al.* [75]	13	12.01	9.30	6.56	5.25
Molina *et al.* [84]	2	19.99	16.31	16.27	15.98
Molina (M) [84]	-	13.71	7.53	6.16	5.57
Odinaka *et al.* (FS) [87]	0.02	1.89	0.93	0.38	0.03
Odinaka *et al.* (NFS)[87]	-	1.93	1.04	0.51	0.06
Sufi *et al.* [118]	-	27.39	21.97	17.14	13.42
Wan and Yao *et al.* [124]	-	7.98	2.15	0.75	0.27
Wang *et al.* (DCT) [126]	-	3.90	2.22	1.74	1.36
Wübbeler *et al.* [130]	-	1.08	0.57	0.57	0.38
Yao and Wan [132]	-	24.46	22.32	20.63	18.49
Ye *et al.* [134]	-	5.11	2.84	1.64	1.13

"Train 8, Test 8" represents training on 8 heartbeats (or 8 s) and testing on 8 heartbeats (or 8 s) from the same session. "Molina (M)" represents a modified Molina algorithm. DCT = Discrete Cosine Transform; FS = Feature Selection; HPE = Hermite Polynomial Expansion; NFS = No Feature Selection; SC = Spatial Correlation; SVM = Support Vector Machine

Table 3.3: Authentication performance for across session (without fusion) analysis

Researchers	Equal Error Rates (%)				
	Train 8, Test 8	Train 16, Test 16	Train 16, Test 32	Train 32, Test 16	Train 32, Test 32
Agrafioti *et al.* [4]	17.95	11.73	11.64	10.48	10.36
Chan *et al.* [17]	16.83	15.37	14.99	14.91	14.64
Chiu *et al.* [24]	26.56	26.28	26.20	26.38	26.36
Coutinho *et al.* [26]	47.14	45.77	44.67	44.59	43.93
Fang and Chan (SC) [33]	29.53	29.54	29.59	29.57	29.71
Fatemian and Hatzinakos [35]	20.40	20.30	20.18	19.42	19.31
Irvine *et al.* [49]	22.25	21.93	21.87	21.66	21.57
Khalil and Sufi [62]	24.16	22.89	22.13	21.94	21.13
Li and Narayanan (HPE+SVM) [73]	19.94	19.20	19.16	18.33	18.16
Lourenço *et al.* [75]	26.11	25.29	25.33	24.85	24.58
Molina *et al.* [84]	37.07	31.75	30.77	30.20	29.42
Molina (M) [84]	34.61	28.65	27.13	25.92	24.67
Odinaka *et al.* (FS) [87]	12.30	11.29	11.13	11.11	11.30
Odinaka *et al.* (NFS)[87]	21.46	20.66	20.37	19.80	20.00
Sufi *et al.* [118]	35.34	33.17	31.98	32.23	31.64
Wan and Yao [124]	16.93	18.44	16.82	21.65	19.22
Wang *et al.* (DCT) [126]	17.94	17.69	17.63	17.72	17.61
Wübbeler *et al.* [130]	16.00	15.60	15.79	15.59	15.77
Yao and Wan [132]	33.33	31.79	31.38	30.92	30.13
Ye *et al.* [134]	22.98	19.19	20.17	20.17	18.55

"Train 32, Test 16" represents training on 32 heartbeats (or 32 s) from session 1 and testing on 16 heartbeats (or 16 s) from session 3. "Molina (M)" represents a modified Molina algorithm. DCT = Discrete Cosine Transform; FS = Feature Selection; HPE = Hermite Polynomial Expansion; NFS = No Feature Selection; SC = Spatial Correlation; SVM = Support Vector Machine

Table 3.4: Authentication performance for across session (with fusion) analysis

Researchers	Equal Error Rates (%)			
	Train(8+8), Test 8	Train(8+8), Test 16	Train(16+16), Test 16	Train(16+16), Test 32
Agrafioti *et al.* [4]	10.68	10.53	9.56	9.51
Chan *et al.* [17]	12.30	11.91	11.57	11.22
Chiu *et al.* [24]	21.34	21.18	21.00	20.97
Coutinho *et al.* [26]	46.17	45.64	44.41	43.77
Fang and Chan (SC) [33]	30.18	29.85	30.22	30.18
Fatemian and Hatzinakos [35]	17.13	16.66	16.92	16.35
Irvine *et al.* [49]	19.65	19.43	19.35	19.22
Khalil and Sufi [62]	18.91	18.58	18.87	18.53
Li and Narayanan (HPE+SVM) [73]	17.40	17.38	17.06	17.09
Lourenço *et al.* [75]	23.22	22.48	22.46	21.97
Molina *et al.* [84]	27.62	27.21	26.12	26.19
Molina (M) [84]	22.24	21.46	20.89	20.48
Odinaka *et al.* (FS) [87]	6.12	6.04	5.64	5.47
Odinaka *et al.* (NFS)[87]	16.08	15.85	14.91	14.73
Sufi *et al.* [118]	33.49	31.03	31.35	29.95
Wan and Yao [124]	9.31	9.45	6.23	6.28
Wang *et al.* (DCT) [126]	16.16	15.92	15.85	15.93
Wübbeler *et al.* [130]	14.62	14.29	14.11	13.98
Yao and Wan [132]	30.99	30.69	30.15	29.84
Ye *et al.* [134]	16.74	17.06	14.32	13.67

"Train (8+8), Test 16" represents training on 8 heartbeats (or 8 s) each from sessions 1 and 2, and testing on 16 heartbeats (or 16 s) from session 3. "Molina (M)" represents a modified Molina algorithm. DCT = Discrete Cosine Transform; FS = Feature Selection; HPE = Hermite Polynomial Expansion; NFS = No Feature Selection; SC = Spatial Correlation; SVM = Support Vector Machine

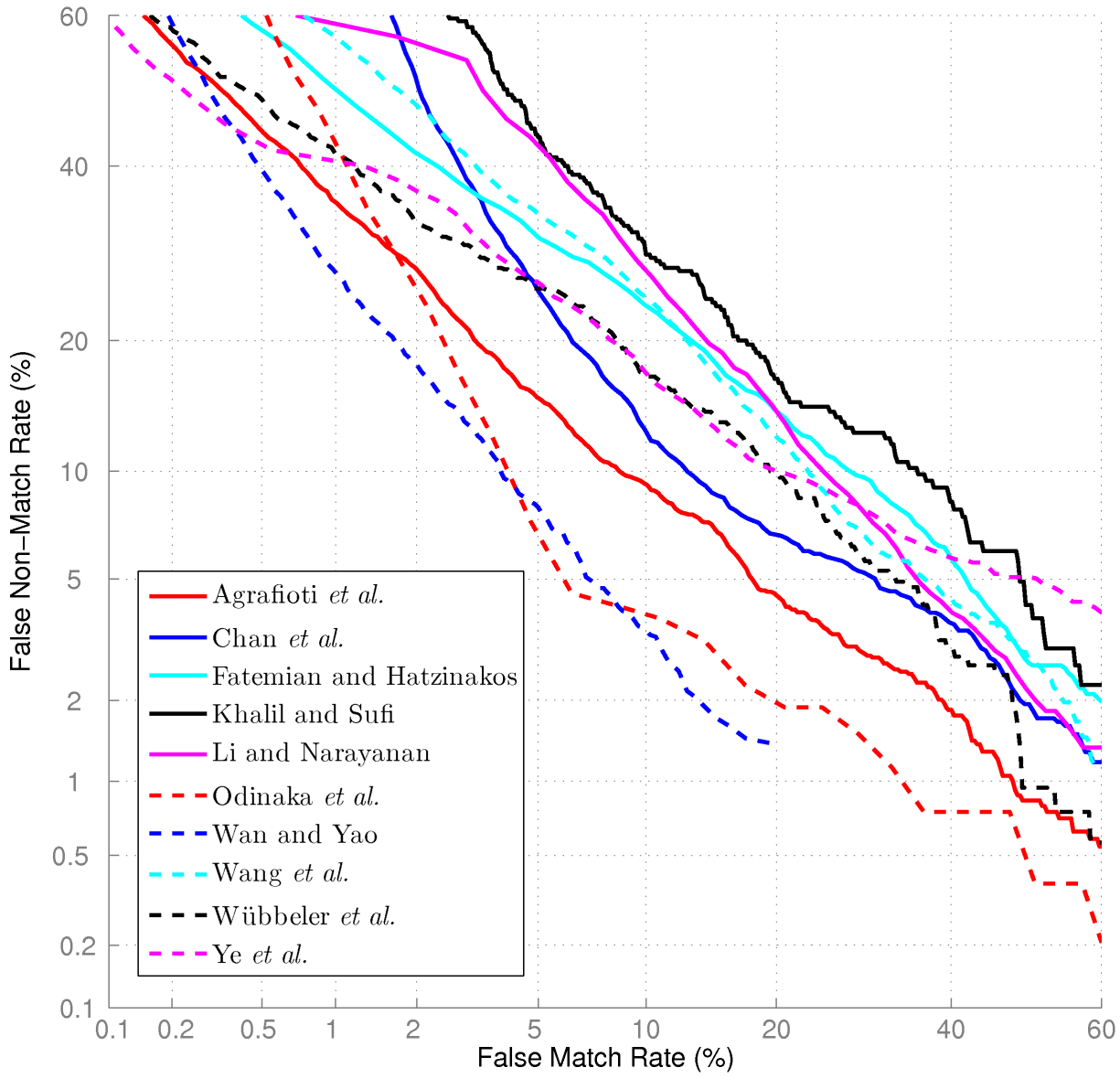


Figure 3.1: Detection error trade-off (DET) curves for the top performing methodologies in the across-session (with fusion) analysis. Training was performed using 16 heartbeats each from sessions 1 and 2, while testing was based on 32 heartbeats from session 3.

normalizing fiducial features using the length of the heartbeat [50, 53]. Agrafioti *et al.* [4] proposed that the autocorrelation sequence obtained in the AC/LDA methodology has the potential of reducing the effect of heart rate variability on the recognition process. Moreover, Safie *et al.* [103] stated that pulse active ratio features can adapt to changes in heart rate. That notwithstanding, an extensive study that would put the compensation techniques to the test is yet to be performed. If the techniques are found short of success, other ideas would need to be investigated.

The ECG signal undergoes both short-term and long-term changes. Short-term changes can be attributed to short-lived (impulsive) activities such as physical or mental activities or the consumption of substances like caffeine. However, long-term changes are mainly due to changes in lifestyle, like the use of medication. Based on the across-session results, we can see that the performance of all the biometric methods degrades over time between training and testing. However, some of the methods were able to regain some of the performance loss by pooling data from multiple sessions.

Perhaps, increased emphasis should be placed upon across-session effects when developing methodologies for ECG biometric recognition, since simply pooling the data from multiple days for the purpose of training does not seem to considerably improve the across-session performance of most of the methods. More data collection efforts will be required to show definitively that such fusion of data from multiple sessions will improve system performance as the number of sessions increases.

Some reports in the literature [130, 120] have proposed the use of only the QRS complex for biometric recognition because it is the most invariant among the other wave components of an ECG heartbeat to heart rate changes and changes over time. However, it is uncertain whether its sole use can support ECG biometric recognition over a large database.

An equally important issue is the database size. Most studies have used a database size below 50 individuals. Based on previous results [87] it is known that ECG identification performance tends to decrease as the number of individuals in the database increases. As such, it is necessary to test the methodologies that have been proposed in the literature, to see if they can carry over to larger databases.

In addition to the performance of the ECG system, other issues have raised some concerns, including privacy. ECG signals have health information embedded in them. Such critical information in the wrong hands can be very devastating. As such there is an interest in

methods for template protection or obfuscating such information prior to the use of the ECG for biometric recognition.

3.6 Summary

The inability to fool ECG sensors, in a supervised setting, is one of the reasons why an ECG based biometric offers an attractive alternative to other traditional biometrics. Although Tsao *et al.* [121] showed that, in theory, it is possible to fool an ECG biometric system by synthesizing an ECG recording using measured features, in practice it may be difficult to replicate an ECG signal at the sensor level.

Despite the considerable effort aimed at developing the ECG as a biometric modality, several important issues remain. These include factors associated with heart rate variability, ECG signal changes over time, and privacy concerns.

We took a closer look at some of the methodologies that have been proposed in the literature using our in-house multi-session database. From that study, we observed that time between training and testing has a drastic impact on system performance, and that some systems have the ability to capitalize on multiple training sessions to achieve improved performance. However, there clearly remains a need for further research involving multi-session protocols. We suspect that robust generative models perform better with multiple training sessions because of their ability to capture the variability across sessions and lead to better model parameter estimates.

3.7 Lessons Learned

Biometric methods based on appearance features, as opposed to fiducial features, are more robust to increase in the number of individuals in the data set.

Although the Neural Network based method works well, it requires a lot of time to test the heartbeat features.

When applied to LDV biometrics, the top three performing methods highlighted in this chapter also achieve single-digit authentication performance. Perhaps, an algorithmic fusion

of the features, scores, or decisions of the methods may improve the biometric performance of the LDV or ECG biometric system.

The AC/LDA method by Agrafioti *et al.* [6] requires the inversion of the within-class covariance matrix for LDA classification. This matrix is rank deficient and can be made invertible by adding an ϵ-scaled identity matrix to the matrix.

Chapter 4

Cardiovascular Biometrics: Combining Mechanical and Electrical Signals

This chapter will appear in an IEEE journal publication. It is currently under IEEE review [91].

4.1 Highlights

The electrical signal originating from the heart, the electrocardiogram, has been examined for its potential use as a biometric. Recent ECG studies, reported in Table 3.4, have shown that an inter-session authentication performance below 6% equal error rate (EER) can be achieved using training data from two days while testing with data from a third day. More recently, a mechanical measurement of cardiovascular activity, the laser Doppler vibrometry (LDV) signal, was proposed by our group as a biometric trait. The inter-session authentication performance of the LDV biometric system is comparable to that of the ECG biometric system. By combining both the electrical and mechanical aspects of the cardiovascular system, an overall improvement in authentication performance can be attained. In particular, the multibiometric system achieves about 2% EER. Moreover, in the identification mode, with a testing database containing 200 individuals, the rank-1 accuracy improves from about 80% for each individual biometric system, to about 92% for the multibiometric system. Although there are implementation issues that would need to be resolved before this combined method could be applied in the field, this chapter establishes the basis and utility of the method in principle, and it identifies effective signal analysis approaches.

4.2 Introduction

Two emerging biometric traits that have been investigated for their potential use for human recognition are the electrocardiogram (ECG) and the laser Doppler vibrometry (LDV) signal. Although the ECG and LDV signals are both measures of cardiovascular activity, the differences between them are substantial and important, and reflect fundamentally different aspects of cardiovascular performance. Whereas the ECG is an electrical signal, which is generated in the myocardium and is conducted by volume to the surface electrodes, the LDV signal (recorded here from the neck, overlying the carotid artery) derives from mechanical displacements of the skin overlying the carotid artery. These displacements in turn are thought to originate in movements of the carotid vessel wall over the blood pressure pulse cycle. The LDV method is but one of several methods that have been developed for assessing this and related mechanophysiological signals.

Previous studies have shown that each biometric system can achieve single digit authentication performance, in terms of equal error rate (EER). In particular, EERs between 5 and 7% have been reported, when training and testing data are obtained from different sessions [20, 88]. Although the recognition performance from each of the individual systems is promising, there is room for improvement. In this chapter, we show that ECG and LDV signals provide different information about an individual's cardiovascular activity. This is a continuation of our previously published works on LDV [20] and ECG [88] biometrics.

To our knowledge, this is the first work that proposes to combine ECG and LDV signals for human recognition. In addition, we employ standard fusion techniques for aggregating the information provided by each biometric system, to achieve a substantial increase in recognition performance. We show that a reduction by at least half in the EER can be obtained from this fusion. Moreover, we report on the methods we have found to work best for authentication and identification. The principal goal of this effort is thus to assess authentication and identification performance using combined ECG and LDV signals. We have pursued this objective using a number of well-established and accepted fusion methods.

When individual biometric systems are shy of the desired level of performance, one of the natural avenues for improvement of biometric performance involves combining the systems in a multiple biometric setting. A multibiometric system is one that uses more than one biometric trait (multimodal), biometric classifier (multi-algorithm), instance of the same biometric trait (multi-instance), sensor for extracting the same biometric trait (multi-sensor), or sample of the same biometric trait (multi-sample) for biometric recognition [56, 100]. A

multi-instance system applies to biometric traits that do not occur singly; e.g., ten fingers, two irises, and multiple ECG channels. A multi-sample system combines multiple samples or different portions of the same biometric trait; e.g. the front and side of a face, segments of the same finger, multiple ECG heartbeats. For physiologically related signals, which change over time, we can combine signals from more than one session, the so called multi-session system [20]. Schmid and O'Sullivan [105] provide bounds on the recognition performance of multibiometric systems. In this chapter, we present the performance gains that result from a multimodal, multi-session, multi-sample system involving the ECG and LDV signals.

The individual systems that constitute a multibiometric system can be combined at multiple levels, ranging from the raw signal to decision processes. Previous work involving combining multiple ECG channels or an ECG channel with other modalities include those by Israel *et al.* [54], Wübbeler *et al.* [130], Agrafioti and Hatzinakos [6], Riera *et al.* [99], Fang and Chan [33], Ye *et al.* [134] and Fatemian *et al.* [34]. Both Wübbeler *et al.* [130] and Fang and Chan [33] combined three ECG channels at the raw level. Israel *et al.* [54] combined ECG and face at the feature level. Riera *et al.* [99] combined ECG and electroencephalogram (EEG) at the score level. Agrafioti and Hatzinakos [6] combined 12 ECG channels at the decision level. Moreover, Ye *et al.* [134] combined two ECG channels while Fatemian *et al.* [34] combined ECG and phonocardiogram (PCG), at the decision level.

In this chapter, we compare the authentication and identification performance of the cardiovascular multibiometric system, when fusion is performed at the feature, score, rank, and decision levels of the biometric pipeline. In addition, we analyze the performance gains of the cardiovascular multibiometric system over the individual ECG and LDV biometric systems. For the purpose of showing the improvement in performance of the multibiometric system, we could have considered only feature- and score-level fusion, but for completeness, we have included our results for rank- and decision-level fusion as well.

The subtle differences between ECG and LDV signals may not be readily appreciated, and it may not be directly obvious that combining the ECG and LDV signals will lead to improved performance. For example, based on Figure 4.1, the use of temporal fiducial features may decrease the performance of the joint biometric system, since the timing information of the two signals are highly correlated. However, the waveform features of the two signals are very different, suggesting that amplitude-based, angle-based, area-based fiducial features, and non-fiducial features are better able to capture uncorrelated information from the ECG and LDV signals.

We have adopted the short-time Fourier transform approach to feature extraction. This is by no means the only way the features can be gotten from each modality for later fusion. See Chapter 3 and the paper by Odinaka *et al.* [88] for a review of the feature extraction methods that have been proposed for ECG biometrics.

The main contributions of this chapter are as follows:

- We performed a large scale cardiovascular biometric experiment on 258 individuals from three sessions. This represents a significant extension of the results of uni-variate modeling [20, 88] to a multidimensional case.
- We showed that ECG and LDV signals provide uncorrelated information for human recognition. Although the feature sets for the ECG and LDV biometric systems were generated in the same way, the scores generated from the two modalities are uncorrelated.
- We performed a comprehensive study of the performance of the multibiometric system at different levels of the biometric pipeline, using standard fusion techniques.
- In one measure of performance, namely equal error rates, we showed that performance improves from about 5.3% to about 2.1%.
- Based on our analysis, we make recommendations on the methods that we found to work best in the authentication and identification modes of operation.

The remainder of this chapter is organized as follows: In Sections 4.3 and 4.4, we provide a brief description of the LDV and ECG biometric systems, respectively. In Section 4.5, we describe the cardiovascular multibiometric system. The results of the different levels of fusion of the LDV and ECG biometric systems are given in Section 4.6, while Section 4.7 provides a summary. Appendix B provides detailed description of methods for extracting the LDV and ECG features.

4.3 LDV Biometric System

4.3.1 LDV as a Physiological Recording Method

Laser Doppler vibrometry is widely used as a non-contact and non-destructive method for assessing mechanical vibrations. The application to the measurement of physiological activity is based on the premise that physiological activity at the system level, particularly in the cardiorespiratory system, has a mechanical component which is substantial and which is transmitted to the skin in the form of sounds and pulses. The associated movements are widely distributed over the body and reflect a combination of gross ballistic forces as well as more focal activities. In the case of measures from the carotid (neck) site, the principal source appears to be associated with movements of the carotid vessel wall produced by the blood pressure pulse. The effectiveness of the LDV method as a physiological recording technique has been demonstrated in a number of settings, for example, assessment of heart and respiration rates in the neonatal ICU [81].

4.3.2 Data Acquisition

Carotid LDV signals were recorded using a Polytec PSV-400 scanning laser Doppler vibrometer [97], with a sensitivity of 10 mm/sec/Volt, and a bandwidth of 0 to 5 kHz. The laser was targeted at the neck region overlying the carotid artery, thereby extracting the component of the velocity of the carotid site that is parallel to the direction of the laser beam. The target site (at the level of the carotid sinus) was established with respect to observable anatomic features, and was standardized over multiple sessions by referring to a photographic record. The data were sampled at a rate of 10 kHz using a Biopac MP150 recording system. The displacement signal (cumulative sum of the velocity signal) can be considered a surrogate for the carotid pressure waveform [60].

4.3.3 LDV Signal

Figure 4.1 shows a 5-s segment of the simultaneously recorded ECG and LDV velocity waveforms of an individual. The maximum velocity peak in the LDV signal, which follows the peak of the R wave in the ECG, is produced by the initial systolic thrust during left ventricular ejection, associated with the upstroke of the blood pressure waveform. The later

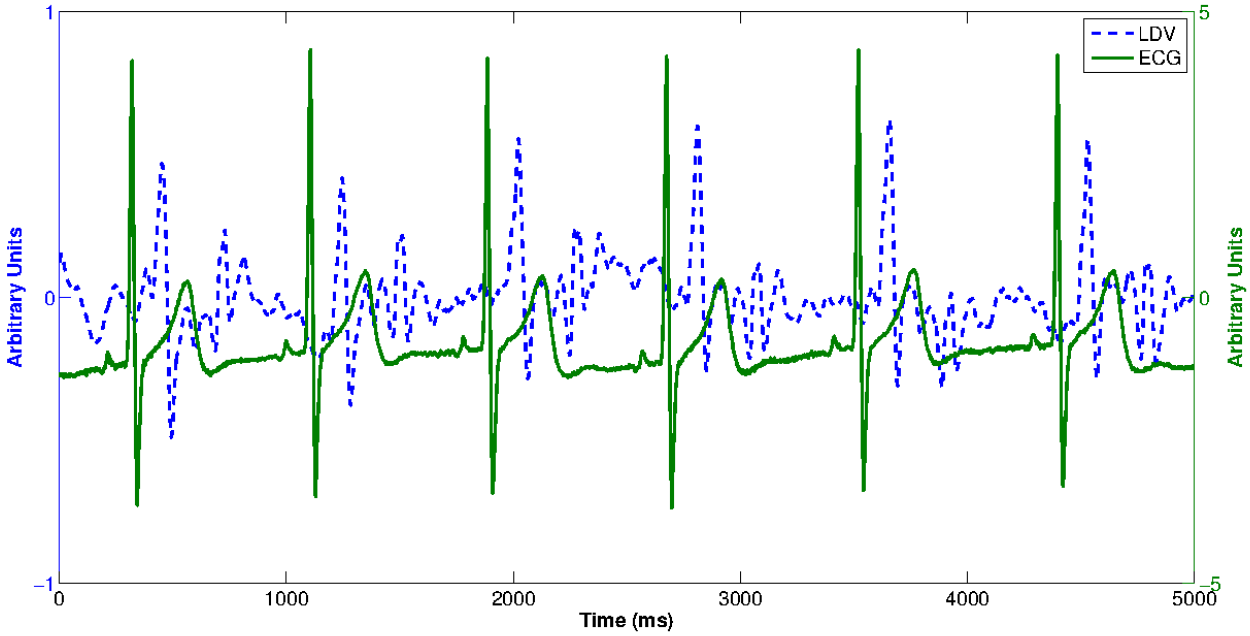


Figure 4.1: A five-second segment of an individual's simultaneously recorded ECG and LDV velocity waveforms.

peaks in the LDV velocity signal are typically of smaller amplitude and are produced by subsequent reflection waves and pressure waves associated with the reservoir function in the arterial system [20, 60].

An LDV pulse signal is a segment of the LDV waveform that corresponds to a single heartbeat. Within each LDV pulse signal, there are two important landmarks: the maximum velocity peak (main peak) early in systole, and the location of the incisura — the incisura can be identified as a distinct peak that appears in the differentiated velocity (acceleration) signal. The incisura represents a major event that marks the end of systole [20].

The LDV waveform is first downsampled to 1 kHz and high-frequency artifacts associated with speckle dropout are suppressed. This procedure is based on the detection of physiologically improbable high frequency events (in the second derivative of the original velocity signal), then removing and resampling through the affected points. Following this, the maximum velocity peak locations are obtained by template matching. Using the detected peak locations, the incisura locations are obtained [20].

After segmenting the LDV waveform into pulses, features are extracted from each LDV pulse signal, as described in Appendix B.

4.4 ECG Biometric System

4.4.1 Data Acquisition

The single channel ECG signals were obtained from electrodes placed bilaterally on the lower rib cage. The ECG signal acquired at this location has a strong R wave, as seen in Figure 4.1, and is less affected by movement artifact than some of the conventional ECG channels. Additionally, these electrode sites have an advantage in terms of test subject acceptability; the individuals do not have to disrobe prior to data acquisition. The signals were recorded with a Biopac TEL-100 system at a sampling rate of 10 kHz [88]. Placement of electrodes was guided with respect to observable and palpable features of the lower ribcage and torso, although the precise locations were not permanently marked on the skin to guarantee exact replacement in subsequent recording sessions.

4.4.2 ECG Signal

An ECG pulse signal is a segment of the ECG waveform that corresponds to a single heartbeat. Within each typical ECG pulse signal, there are three important landmarks: the P wave, QRS complex, and T wave.

The ECG signal is first downsampled to 1 kHz, and the power line frequency removed with a notch filter. Following this, the peak locations of the R waves are obtained by cross-correlation [87].

Following the segmentation of the ECG waveform into pulses, features are extracted from each ECG pulse signal, as described in Appendix B.

4.5 Cardiovascular Multibiometric System

4.5.1 Database Description

Our database consists of recordings from 258 individuals (117 males and 141 females) whose ECG and LDV recordings were obtained concurrently during a continuous 5 min period.

The recordings took place with the individuals in a normal sitting and resting condition [20, 87]. The ages of the individuals varied from 18 to 66 years, with mean and standard deviation of 39 and 14 years, respectively. As is typical in a community sample of this sort, 41% of the individuals self-reported some heart-related disease, including hypertension. 46% of the individuals reported using medications that may affect the ECG/LDV signal. 28% of the individuals did not report any significant health issues or use of medications that may affect the ECG/LDV signal. 73% of the individuals were Caucasians. The individuals were asked to sit for five minutes as the recording took place. Each individual provided three ECG/LDV recordings, each taken on a different day. We call these three recordings, sessions 1, 2, and 3, where the separation between two consecutive sessions ranged from one week to six months. The average time interval between sessions 1 and 2 was 15.4 days and 40.6 days for the time interval between sesssions 2 and 3. The time interval from session 1 to 2 lies between 4 and 92 days, while that from session 2 to 3 is between 5 and 156 days.

4.5.2 Information Fusion

The ECG and LDV biometric systems can be combined at different levels of a biometric pipeline. Ideally, multiple biometrics can be fused at the raw data, feature, score, rank, or decision levels. Since the LDV signal is a mechanical signal and the ECG signal is an electrical one, and their morphologies are distinct, it may be difficult to achieve fusion at the raw data level. Consequently, we consider fusing the two biometric systems at the feature, score, rank, and decision levels. We are currently working on methods for fusing the raw data, and will present our results in future reports.

According to the data processing inequality [27, 100], the amount of information about an individual contained in the raw signal is at least as great as that in the feature vector. Similarly, the amount of information about the individual contained in the feature vector is not smaller than that in the test scores. The system decisions usually contain the least amount of information about the individual. As such, it can be anticipated that when the information from earlier stages of the biometric pipeline are optimally combined, the multi-biometric system will achieve better performance than one which relies on further processed information. Note that this statement is true only if the combination of the modalities is optimal, accounting for signal noise and the relationship between the modalities, if any.

It is known [100] that factors such as correlation, and differences in classifier accuracy among the biometric systems to be combined, affect the resulting performance of the multibiometric

system. There is no guarantee of performance gain when the biometric systems are highly correlated. Moreover, when the biometric systems have very different performance, the weakest (least performing) system tends to overwhelm the fusion.

Feature-level Fusion

Feature-level fusion is one of the two ways (along with raw-level fusion) to consolidate the information from multiple biometric systems, prior to the actual computation of match scores. After creating the fused feature vectors, a single classifier is employed.

The simplest way to perform feature-level fusion is to concatenate the features vectors from the different modalities. For the methods we explore, the concatenation can be performed before feature selection (BFS) or after feature selection (AFS). For our multibiometric system, when a single threshold is used for both the LDV and ECG feature vectors, we say that feature-level fusion occurs prior to feature selection. On the other hand, when two different thresholds are adopted by the individual systems, the fusion occurs after feature selection. See Appendix B for more information about the feature selection methodology.

The LDV [20] and ECG [88] biometric systems, as described in Appendix B, make use of different window parameters for generating the time-frequency feature vector. As a result of this, they may not be directly compatible for feature-level concatenation. Moreover, due to the inherently different nature of the signals, it may not be optimal to adopt a single threshold for selecting features from both systems; ECG and LDV signals have differing frequency content, for each time frame.

To solve the problem of different parameters used in feature generation, for the purpose of feature-level fusion, we used the LDV feature generation parameters for both the ECG and LDV signals. Feature-level fusion BFS was performed by using a single threshold for the entire concatenated vector.

Based on prior work on ECG and LDV biometrics [20, 88], we know that both systems have comparable recognition performance. Informed by this, during training for feature-level fusion AFS, we found the feature selection thresholds that correspond to a given fraction of selected time-frequency bins of the LDV and the ECG, for the entire training population. For example, if 23% of the LDV time-frequency bins are selected on average across the training population, approximately the same proportion of the ECG time-frequency bins will be selected, and vice versa. An alternative way to obtain the feature selection thresholds is to

perform a cross-validation study by sweeping the thresholds used for the two modalities, and finding the pair that gives the best recognition performance on the training population.

We expect the recognition performance of the AFS feature-level fusion scheme to exceed that of the BFS since the AFS scheme utilizes different bin-selection thresholds for the ECG and LDV signals.

To determine the bin-selection thresholds for feature-level fusion, a subset of 58 individuals (training subjects) in the database were utilized. These individuals were subsequently excluded from the test set.

A detailed description of the score functions used in authentication and identification is given in Appendix B.

Score-level Fusion

Score-level fusion is the richest post-matcher fusion technique, in terms of the amount of information about the individual to be classified. There are three major approaches to score-level fusion: score normalization, density estimation, and classifier-based fusion [56, 100]. Score normalization involves heuristically normalizing the scores from each biometric system, so that the scores become compatible. Examples of score normalization techniques include: Min-Max, Z-Score (Mean-Std), Median-MAD (median absolute deviation), Double Sigmoid [15], and Tanh estimators [56, 100]. For all score normalization techniques, during training, some parameters are estimated from the training scores. The parameters are later used during the actual authentication or identification process.

Min-Max score normalization computes the minimum and maximum of the scores obtained during training, for each biometric system, and then uses them to normalize the scores obtained during testing. The minimum score is subtracted from the test score, and the result is divided by the difference between the maximum and minimum scores. Z-Score normalization involves computing the sample mean and standard deviation of the scores. During testing, the test scores are normalized by subtracting the mean and dividing by the standard deviation. Median-MAD normalization replaces the mean and standard deviation in the Z-Score normalization procedure with the median and standard median absolute deviation, respectively.

The Double Sigmoid normalization technique pieces two sigmoidal functions together in their linear regime, in order to normalize test scores. See Appendix B for more details on how the Double Sigmoid parameters are estimated.

Tanh estimator score normalization is based on robust estimators of location and scale of a distribution. Using the robust location and scale estimates of the genuine training scores as parameters to a hyperbolic tangent function, test scores can be normalized [56, 100]. In particular, we used the median and standard median absolute deviation of the genuine training scores to normalize the scores before using the tanh normalization function. See the above-referenced publications for details on each score normalization technique.

Following the normalization of each biometric system's score, the scores can be combined to obtain a representative score for the multibiometric system. We consider three approaches for combining the normalized scores: mean, maximum, and minimum [100]. We did not consider a weighted average method for combining the normalized scores, because the normalization should put the scores on the same scale. The score module, in the biometric pipeline, takes care of both score normalization and combination. In this chapter, we examine the multibiometric system performance when each of the five score normalization and three normalized-score combination techniques are applied.

Density estimation score fusion requires that the distribution of the genuine and impostor scores be estimated from the training scores. The estimated distributions can then be used in a likelihood ratio test. We estimated the joint distributions of the genuine and impostor scores from the ECG and LDV biometric systems via the Fukunaga method of Gaussian kernel density estimation [110].

Classifier-based score fusion uses a classifier to separate or cluster the training scores into classes. The scores from the different biometric systems are concatenated into a single vector. During the training phase, a classifier is trained by using the score vectors as feature vectors. During testing, the test scores can be assigned to the best matching class [56, 100]. Support vector machines (SVMs) and neural networks have been used to classify test scores [100]. Since we consider only two biometric systems, and there are only two training classes, namely, genuine and impostor, SVM based on a linear kernel is akin to fitting a straight line between the genuine and impostor scores. The line should separate the two classes and be as far away as possible from either class. For neural network score fusion, we used a multi-layer perceptron neural network (MLP-NN) with one hidden layer of 4 neurons to fuse the genuine and impostor scores.

Due to relatively larger instances of the impostor scores compared to the genuine scores, it becomes necessary to balance the number of training examples from each class. We employed random downsampling (without replacement) of the impostor examples to match the number of genuine scores.

Rank-level Fusion

In addition to combining biometric information at the score level, in the identification mode, information can be fused at the rank level. After computing the match scores of each test signal, each biometric matcher outputs a rank for each of the classes (enrolled individuals), with a rank of 1 signaling the best match. There are several techniques that have been proposed in the literature for rank-level fusion. These include highest rank, Borda count, weighted Borda count (logistic regression) [46], modified highest rank [1], and nonlinear weighted ranks [64]. Appendix B provides useful information on the parameter estimation for the weighted Borda count and nonlinear weighted ranks. For a more detailed description of each rank-level fusion technique, see the above-referenced literature and Appendix B.

In a rank-m identification task, after the fused ranks are reordered from smallest to largest, if the true identity of the test signal is among the identities associated with the top m ranks, then the identification is a success.

Decision-level Fusion

The last point along the biometric pipeline where information can be consolidated between multiple biometric systems is the decision level. The decision made by a biometric matcher can be an acceptance or rejection, in authentication mode, or an identity verification or rejection, in identification mode. This should come as a last resort when all other forms of biometric fusion are infeasible, because the least amount of information about the individual is available at this level; a binary (at best multinomial, in identification mode) response from the individual biometric matchers.

Several techniques have been proposed to achieve this level of fusion, including, "AND" and "OR" rules [95], majority voting [70], weighted majority voting [65], Bayesian decision fusion and Dempster-Shafer fusion [131]. The majority voting and weighted majority voting techniques do not directly apply to a multibiometric system with only two biometric matchers;

majority voting with test sample rejection or random tie breaks can be used in their stead. That is, if the two biometric systems give different decisions, either a combined decision is made by choosing one of the two decisions at random or the sample is rejected. Appendix B briefly describes the "AND" and "OR" rules, Bayesian decision fusion, and Dempster-Shafer fusion. See Appendix B and the references above for more details on each of the decision-level fusion schemes.

In this chapter, we analyze the authentication and identification performance of the cardiovascular multibiometric system using feature-level, score-level, rank-level and decision-level fusion. At the score level, we consider the score normalization, density estimation, and score classifier based approaches. At the rank level, we consider the highest and modified highest rank, Borda and weighted Borda count, and nonlinear rank combination techniques. At the decision level, since we are combining only two biometric modalities, for authentication, we consider using the "AND" and "OR" rules. In addition, for identification, we use the naïve Bayes', Dempster-Shafer, majority voting with test sample rejection, and majority voting with random tie breaks, decision fusion methodologies.

4.6 Results

Combining multiple biometric systems can provide an improvement in recognition performance, if the individual biometrics are uncorrelated [100]. We examined the authentication and identification performance of the ECG and LDV biometric systems, and the cardiovascular multibiometric system in the inter-session with data fusion setting. In particular, we considered the scenario where the training heartbeats were obtained from sessions 1 and 2, while the testing heartbeats came from the third session.

Chen *et al.* [20] and Odinaka *et al.* [88] consider the cases of within-session and multi-session without data fusion, in addition to the case of multi-session with data fusion. We have omitted the treatment of the first two cases here in the interest of length considerations. However, the performance gains due to multibiometric fusion carry over to them as well.

To permit a fair comparison between the cardiovascular (ECG + LDV) multibiometric systems and the individual ECG and LDV biometric systems, if H heartbeats are used for training the individual biometric systems and G heartbeats used in testing, $H/2$ and $G/2$ heartbeats, each from the LDV and ECG datasets, will be used for training and testing the cardiovascular multibiometric systems, respectively. When G heartbeats are used for testing,

for authentication, the scores from each of the heartbeats are combined by averaging; for identification, the scores are combined by computing the maximum score. The maximum function was chosen over others (mean, median, and min) because it gave the best performance during training [87]. This chapter represents the results not for single, but multiple testing heartbeats.

4.6.1 Authentication

We examined the authentication performance of the cardiovascular multibiometric system under different fusion schemes — feature, score, and decision levels. To assess the performance of the multibiometric systems, 64 ECG pulse signals and 64 LDV pulse signals, each from sessions 1 and 2, were obtained for training; for testing, 8 ECG pulse signals and 8 LDV pulse signals were obtained from the third session. In order to compare the performance of the individual biometric systems — ECG and LDV, to those of the cardiovascular multibiometric systems, each of the biometric systems used 128 and 16 pulse signals for training and testing, respectively.

For training and cross validation, the data obtained from 58 individuals were used; the remaining 200 individuals were used in testing. This approach of splitting the database into training/validation and testing individuals allows for generalizability, since the individuals involved in the training process are not utilized during testing.

The authentication performance is reported by means of equal error rates (EERs) and detection error trade-off (DET) curves. For the EERs, in addition, we report a 95% confidence interval, computed by assuming a certain correlation structure between the individual tests [58]. The format for reporting the EERs and confidence intervals will be $b, [a, c]\%$, where b is the average EER, and a and c are the lower and upper bounds on the EER, respectively.

When fusion was performed at the feature level before feature selection, EERs of 2.6, [1.5, 3.6]% were obtained; EERs of 2.3, [1.2, 3.4]% were obtained for feature-level fusion after feature selection. Based on the confidence intervals, the performance of the two schemes are similar. However, on average, as anticipated, feature-level fusion after feature selection is preferable.

For score-level fusion, the training scores were obtained from the training individuals. The training scores were then used to estimate the necessary parameters for each score normalization technique, the genuine and impostor score probability distribution via Gaussian kernel density estimation, and the parameters of the linear SVM and Neural Network classifiers. Table 4.1 shows the authentication performance (in EER) for the different score-level fusion methodologies.

From the table, we can see that by combining the ECG and LDV biometric systems at the score level, we can achieve a lowest EER of about 2.1%, on average. For score-level fusion based on score normalization, the best performance is obtained when the normalized scores are averaged across the biometric matchers. On average, the Median-MAD and Tanh estimator approaches provide the best techniques for score normalization, during authentication. The overall best performance, on average, for all the score-level fusion techniques was obtained using the Neural Network score classifier technique, with comparable performance obtained by using the Median-MAD and Tanh estimator score normalization techniques, linear SVM, and Gaussian kernel density estimation. Due to the overlap of the estimated confidence intervals, we cannot claim that one method is always better than the other, rather, we can compare performance on an average basis.

The ECG and LDV biometric systems were also combined at the decision level using "AND" and "OR" rules, yielding EERs of 3.3, [1.1, 5.5]% and 4.0, [1.8, 6.3]% respectively. As expected, the best decision-level performance is lower than the best score-level performance.

The authentication performance of the cardiovascular multibiometric systems can be contrasted with those of the individual ECG and LDV biometric systems. The authentication performance (%EER, [95% confidence interval]) of the individual biometric systems were 5.3, [3.5, 6.9]% and 5.4, [3.8, 7.3]%, respectively, for the ECG and LDV signals. Feature-level (AFS) fusion offers an improvement in average EER of about 57% over the single ECG and LDV biometric systems. Score-level fusion (Median-MAD, Tanh estimator, and Neural Network classifier) provides an improvement in average EER of about 60% and 61% over the single ECG and LDV biometric systems, respectively. Decision-level fusion, using "AND" rules, improves the average EER by 38% and 39% respectively, over the single ECG and LDV biometric systems.

For clarity, we have separated the authentication performance curves of the score-level fusion techniques into two figures. Figure 4.2a shows the detection error trade-off (DET) curves,

Table 4.1: Authentication performance (% EER, [95% confidence interval]) for the score-level multibiometric system. 58 individuals were utilized for obtaining training scores and the remaining 200 individuals for testing. Training was performed using 64 heartbeats each from sessions 1 and 2, each from LDV and ECG, while testing was performed on 8 heartbeats each from the ECG and LDV testing datasets (session 3).

Score Combination	Score Normalization				
	Min-Max	Z-Score	Median-MAD	Double Sigmoid	Tanh
Max	3.5, [1.8, 5.0]	3.6, [2.2, 5.2]	3.3, [1.7, 5.0]	3.3, [1.8, 4.9]	3.4, [1.9, 5.0]
Mean	3.2, [1.9, 4.4]	2.6, [1.4, 3.8]	2.1, [1.1, 3.2]	2.6, [1.5, 3.8]	2.1, [1.1, 3.3]
Min	5.8, [3.7, 7.4]	5.8, [3.9, 7.7]	4.1, [2.8, 5.5]	5.8, [3.9, 7.6]	4.2, [2.9, 5.3]
	Classifier-Based Score Fusion				
	Linear SVM			Neural Network	
	2.2, [1.1, 3.4]			2.1, [1.1, 3.2]	
	Gaussian Kernel Density Estimation-Based Score Fusion				
	2.3, [1.2, 3.4]				

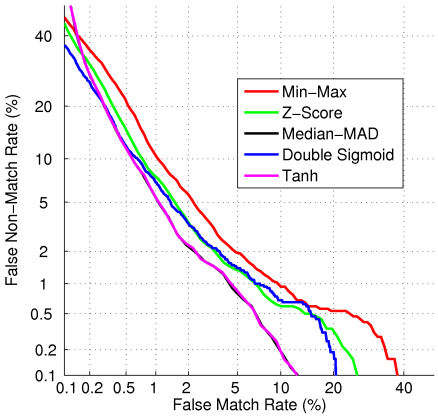
plotted using a normal deviate scale as prescribed by Martin *et al.* [82], for the cardiovascular multibiometric systems based on score normalization score-level fusion, while Figure 4.2b shows the DET curves for all the score-level fusion methodologies, including only one curve from the score normalization methods. The DET curves for the score normalization techniques are for the case where the normalized classifier scores are combined by averaging. From the figures, we can see that, on average, the best score normalization techniques are the Median-MAD and Tanh estimator. Among the score-fusion techniques, the Median-MAD and Tanh estimator score normalization methods, linear SVM and Neural Network score classifiers have fairly similar performance, on average.

We examined the overall biometric performance of the individual unimodal systems in comparison to those of the multimodal systems. The DET curves for the ECG, LDV, and cardiovascular multibiometric systems based on feature-level fusion AFS and one of the best score-level fusion techniques (Median-MAD) are shown in Figure 4.2c. We can observe from the figure that the multibiometric systems exceed each of the single biometric systems at each operation point. Moreover, the score-level Median-MAD score normalization scheme outperforms the feature-level fusion scheme, on average. This clearly means that the features from both modalities are not optimally combined. The design of better feature-level fusion schemes is one of our future goals.

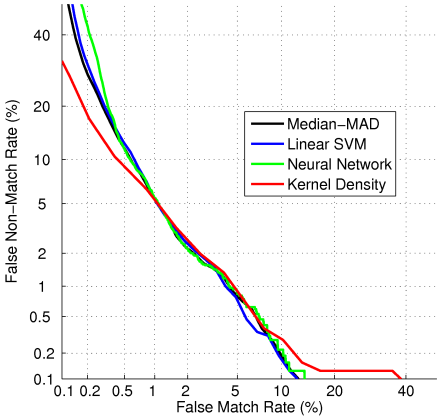
Figure 4.3 shows the scatter plot between the training scores obtained from the ECG biometric system and those obtained from the LDV biometric system. The genuine scores are given by the blue circular points, while the red circular points correspond to the imposter scores. The correlation coefficient between the scores is about 0.095, meaning that the scores from the ECG and LDV biometric systems are mostly uncorrelated. Thus, it is not surprising that a proper combination of the two systems, at the score level, offers an improvement in biometric performance. This provides an empirical validation of the physiological differences between the two signals as stated earlier in the introduction.

4.6.2 Identification

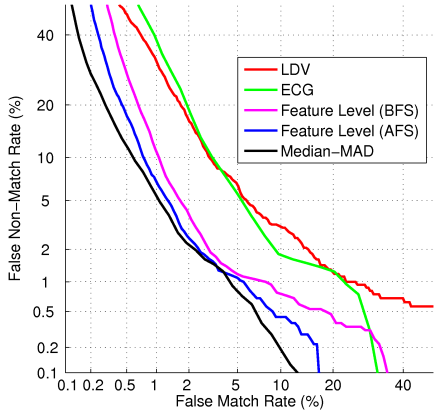
We evaluated the performance of the cardiovascular multibiometric system in the identification mode, under different fusion schemes — feature, score, rank, and decision levels. The number of heartbeats used for training, cross validation, and testing were identical to those used in the authentication mode.



(a)



(b)



(c)

Figure 4.2: Detection error trade-off (DET) curves for cardiovascular multibiometric and uni-biometric systems. (a) Score normalization fusion. Averaging to combine normalized scores. (b) Score-level fusion. (c) LDV, ECG, feature-, score-level fusion.

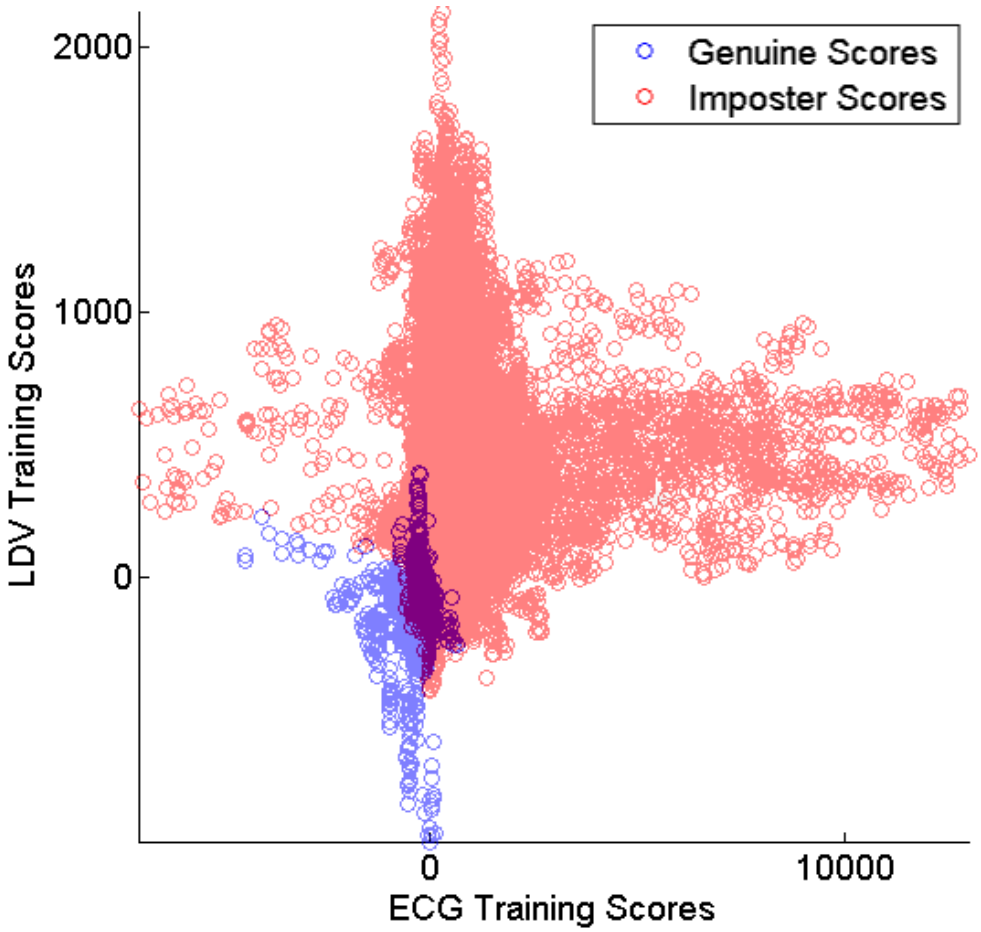


Figure 4.3: Scatter plot of training LDV scores versus training ECG scores. The blue and red circular points correspond to the genuine and imposter scores, respectively.

Exceptions to this rule are the naïve Bayes' and Dempster-Shafer decision fusion techniques, which require training based on the individuals to be tested on. To achieve this training, half of the training heartbeats from sessions 1 and 2 were used for training the log-normal feature models, while the remaining half were used to obtain training scores, which were utilized in estimating the probabilities needed by the two decision fusion methods.

When the multimodal fusion was performed at the decision level, using naïve Bayes' and Dempster-Shafer's fusion rules, identification accuracies of 68.8% and 79.3% were obtained, respectively. The performance of the multibiometric system, based on naïve Bayes' and Dempster-Shafer's decision-level fusion, degraded in comparison to that of the individual unimodal biometric systems; 80.4% and 80.3% for the ECG and LDV biometric systems, respectively. The performance degradation may be due to the limited number of training decision labels used in estimating the conditional probabilities in naïve Bayes' and the degrees

of belief in Dempster-Shafer's fusion rules. In addition, the assumption of independence between the ECG and LDV decisions, made by naïve Bayes', may not be appropriate.

When majority voting with random tie breaks was used to combine the ECG and LDV decisions, an identification performance of 78.6% was obtained. Moreover, when majority voting with test sample rejection was used for decision fusion, an identification performance of 99.9% was obtained, with 34.2% of the total test heartbeats and 35 out of 200 subjects being rejected. Since decision-level fusion does not offer a significant performance gain, we look to other levels along the biometric pipeline.

The identification performance of a biometric system can be summarized by using a cumulative match characteristic (CMC) curve. It plots the identification accuracy against the cumulative rank of the biometric system. For clarity, we have separated the CMC curves for score-level fusion into two figures. Figure 4.4a shows the CMC curves for the score-level fusion techniques based on score normalization, using averaging to combine the normalized scores, while Figure 4.4b shows the CMC curves for the best score normalization technique, and the other score-level fusion methods, including: SVMs, NN, and kernel density estimation. It is evident from both figures that score-level fusion based on Double Sigmoid score normalization outperforms the other score-level fusion techniques for all ranks; the other techniques have comparable identification accuracy. The Double Sigmoid score normalization technique achieves a rank-1 identification accuracy of about 92%. The Double Sigmoid technique is the best method for score-level fusion in the identification mode. At present, the reasons remain unknown.

The CMC curves for the rank-level fusion techniques are shown in Figure 4.4c. From the figure, except for the highest rank fusion scheme, which has an accuracy of about 79%, each rank fusion scheme achieves an accuracy of about 89%, for rank-1 identification. The highest rank and modified highest rank fusion schemes achieve the best identification performance, beyond a cumulative rank of 5. A similar trend in the CMC curves of the highest and modified highest rank fusion methods can be observed in the paper by Abaza and Ross [1]. A possible explanation for why the highest rank method under-performs for rank-1, is that it encounters frequent ties for best rank and resolves them randomly, selecting the wrong identity half the time. The modified highest rank method first uses the Borda count, before resorting to random tie breaks, resulting in fewer random tie breaks; the Borda and weighted Borda count methods are more conservative in fusing the ranks, resulting in less ties for best rank. However, for rank-m identification, as m increases, when ties are broken randomly, the assigned rank of the true identity is still within the top m positions, resulting in much

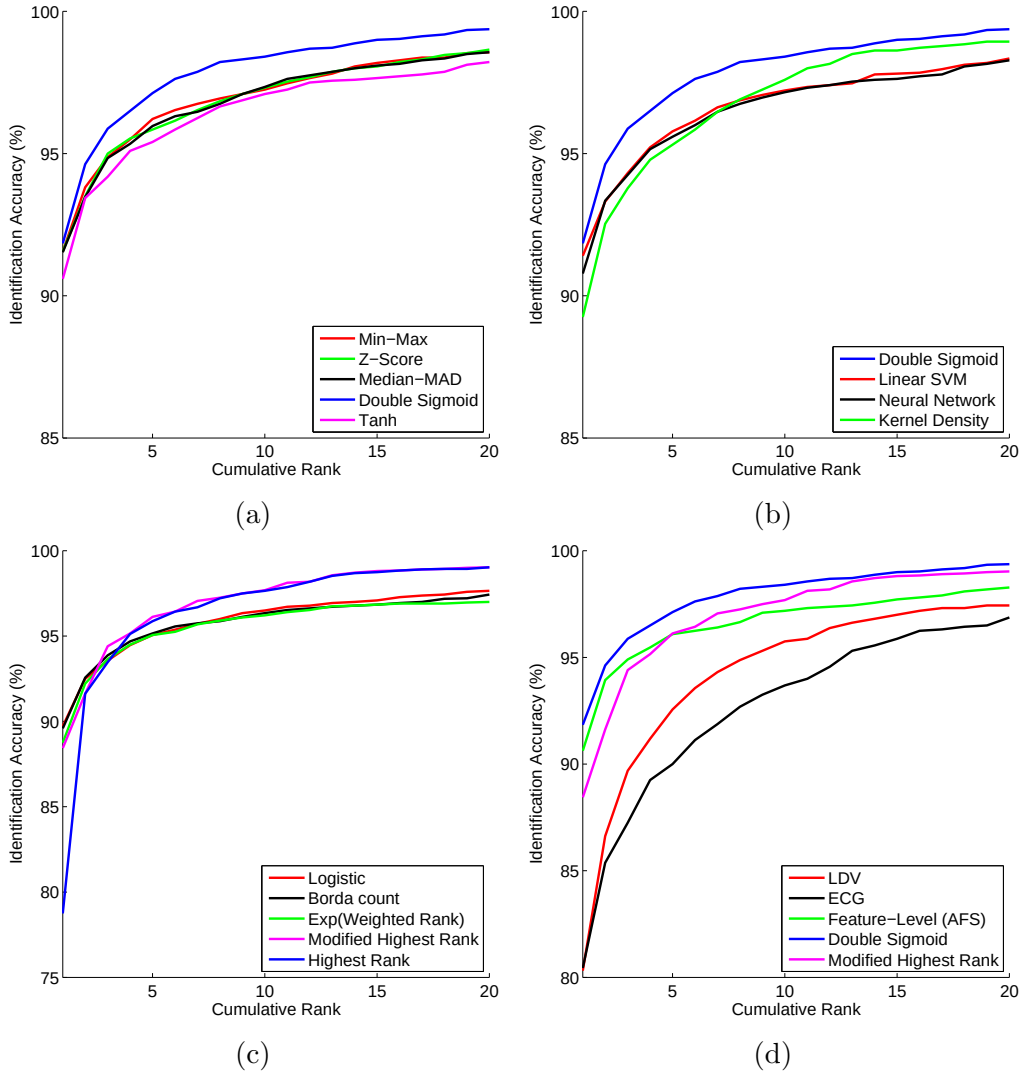


Figure 4.4: Cumulative Match Characteristic (CMC) curves for the cardiovascular multi-biometric and uni-biometric systems. (a) Score normalization fusion. Averaging to combine normalized scores. (b) Score-level fusion. (c) Rank-level fusion. (d) LDV, ECG, feature-, score-, rank-level fusion.

better performance than Borda count, which may place the true identity outside the top m spots, since it adds the ranks from the two modalities.

We examined the overall identification performance of the individual unimodal systems in contrast to those of the multibiometric systems. The CMC curves for the LDV, ECG, and cardiovascular multibiometric systems based on feature-level fusion, the best score-level fusion technique (Double Sigmoid), and the best rank-level fusion scheme (modified highest rank) are shown in Figure 4.4d. We can observe from the figure that the multibiometric systems perform better than each of the single biometric systems. The score-level fusion technique performs better than all the other fusion schemes, including feature-level fusion. This is similar to what we observed in the authentication mode, providing further evidence that the feature-level fusion scheme can be improved.

4.7 Summary

Previous studies have shown that the ECG and LDV biometric systems can attain authentication performance below 6% EER, and identification accuracy of about 80%, when training on heartbeats from multiple sessions and testing on multiple heartbeats from a different session [20, 88]. Although these levels of performance may be useful in some scenarios, there is opportunity for improvement. This has motivated our approach based on combining the two biometric systems in a multimodal fashion.

We fused the two biometric systems at the feature, score, rank, and decision levels. Based on our analysis, score-level fusion gave consistently the best performance for both authentication and identification. We demonstrated the gains in authentication performance, when single biometric modalities, ECG and LDV, are combined into a cardiovascular multibiometric system. In particular, about a 60% decrease in equal error rates was obtained. Moreover, the best rank-1 identification accuracy of the multibiometric system of about 92% was achieved when score normalization based on a Double Sigmoid function was utilized for fusion.

We investigated using signals acquired concurrently from both modalities. We have also investigated the effects of fusing ECG and LDV signals that are not obtained at the same instant. Similar performance gains were obtained for that setting.

The LDV signals used in this study came from measurements at the carotid site, while the ECG signals came from bilateral chest electrodes. Such spatially separated measurements

could be cumbersome for a multimodal biometric system, in practice. However, Chan *et al.* [17] demonstrated that ECG signals can be acquired from the fingers by using a mouse. Moreover, mechanical cardiovascular signals such as blood pressure can also be extracted from the finger tips. We envision a cardiovascular multibiometric sensor that is housed within a single unit, making signal acquisition less cumbersome.

4.8 Lessons Learned

The feature-level fusion methods did not perform as well as the best performing score-level fusion method, suggesting that the feature-level fusion methods need to be further improved.

Chapter 5

Estimating Biometric Capacity

This chapter appeared in a previous IEEE conference publication [89].

5.1 Highlights

It is known that when a pattern recognition system is not subject to memory constraints on pattern representations and sensory information is not compressed, the recognition rate is bounded by the mutual information between the memory and sensory representations [128]. In this dissertation, we investigate the recognition rates of Laser Doppler Vibrometry (LDV) signals obtained from 285 individuals. We consider four cases corresponding to four different assumptions as to the structure of the data source and noisy measurements, and present the results of the mutual information bounds. In particular, we show the bounds of the recognition rates for each feature of the LDV signal.

5.2 Introduction

In an earlier work [128], Westover and O'Sullivan proposed an information theoretic framework for understanding pattern recognition capacity and the trade-off among the pattern complexity, memory and sensory resources, and noisy measurements. The patterns to be recognized are denoted as a length n vector x drawn from a distribution p_x, and the sensor input is a noisy version y, modeled as x passing through a channel with transition probability $p_{y|x}$. In general, there may be memory compression and sensory compression due to system limitations as discussed in [128]. However, we focus on the case where no compression takes place. In this special case, the maximum number of patterns that a system can

recognize is $M_c = 2^{nR_c}$, with the probability of error asymptotically approaching zero as n gets large. The number of recognizable patterns is bounded by $2^{nI(X;Y)}$, where $I(\cdot,\cdot)$ is the mutual information. This implies that $R_c \leq I(X;Y)$.

In this work, we estimate the recognition rate of real signals under four scenarios. This chapter is organized as follows: Section 5.3 introduces the LDV experimental dataset and the generation of features. In Section 5.4, we discuss the scenarios in which the recognition rates of LDV signals rate are computed. In Section 5.5, we discuss the cases presented in Section 5.4 in more detail. The simulation results are given in Section 5.6, and a summary in Section 5.7.

5.3 LDV Signal and Short-time Fourier Transform

A Polytec PSV-400 scanning laser Doppler vibrometer, with a sensitivity of 10 mm/sec/Volt, and a bandwidth of 5 kHz, starting from DC, was used to measure vibrations of the surface of the skin overlying the carotid artery [20]. The signals, which we call LDV signals, were obtained for a period of 5 minutes, using a Biopac MP150 recording system, while the individuals were in a normal resting condition. The LDV signals, which were originally sampled at 10 kHz, were subsequently down-sampled to 1 kHz. The resulting LDV signal was reduced to individual 900 ms segments aligned to the respective peaks of the velocity waveform, beginning 400 ms prior to each peak. Each LDV heart pulse was normalized by subtracting the sample mean of the pulse, and dividing by the sample standard deviation.

From each LDV pulse signal, we computed a spectrogram which is the logarithm of the square of the magnitude of the short-time Fourier transform of a normalized LDV heartbeat. In computing the short-time Fourier transform (STFT), we used a Nuttall window of size 96 ms, with a step size (the distance between the beginning of two consecutive windows) of 16 ms. Thus, there was an overlap of 80 ms between consecutive time frames [20]. After computing the STFT, the frequency content was truncated at 150 Hz to reduce boundary effects. The spectrogram was then computed as the logarithm of the squared-magnitude of the truncated STFT. We refer to the index of each point of the spectrogram as a time-frequency bin or feature. Thus each LDV heartbeat can be represented by $L = 1479$ time-frequency components (corresponding to 51 time frames and 29 frequency bins), with the value of the i-th time-frequency bin denoted as Y_i.

5.4 Bound Evaluation Settings

We assume that each time-frequency bin of an individual's signal is Gaussian distributed with mean μ and variance σ^2, where the mean and the variance are drawn from certain distributions. The parameters μ and σ^2 correspond to the vector x of length n. We also assume that the parameters μ and σ^2 are independent of each other. For each time-frequency bin of each individual, we compute the maximum likelihood (ML) estimates of the mean and variance.

We consider four settings, corresponding to the distributions of μ and σ^2. For the first three settings, we consider two cases. In the first case, the mean and variance (parameters) of all the time-frequency bins are drawn from the same distribution. In the second case, the parameters of each time-frequency bin are independent. That is, the parameters of a given time-frequency bin are independent of those of other time-frequency bins. Moreover, for the fourth setting, we consider the case where the parameters of each time-frequency bin are independent. The four settings are as follows:

- *Gaussian distributed μ, Constant σ^2.* In the first case, we assume the variance of all time-frequency bins is the same constant, while the mean is Gaussian distributed. For the second case, we assume a constant variance for each time-frequency bin, while the mean is drawn from a Gaussian distribution.

- *Constant μ, Gamma distributed $\frac{1}{\sigma^2}$.* In the first case, we assume the mean of all time-frequency bins is the same constant, while the reciprocal of the variance is Gamma distributed. For the second case, we assume a constant mean for each time-frequency bin, while the reciprocal of the variance is Gamma distributed.

- *Gaussian distributed μ, Gamma distributed $\frac{1}{\sigma^2}$.* In the first case, we assume the mean of all time-frequency bins is drawn from a single Gaussian distribution, while the reciprocal of the variance is drawn from a single Gamma distribution. For the second case, we assume the mean of each time-frequency bin is Gaussian distributed, while the reciprocal of the variance is Gamma distributed.

- *Nonparametric kernel density estimation.* Here, we consider the case where the parameters of each time-frequency bin are independent. We estimate the joint distribution of the parameters μ and σ^2, and use this in the computation of the mutual information bound on the LDV recognition rate.

The parameters μ and σ^2 are representative of an individual. Once the parameters have been generated by their respective distributions, the i-th time-frequency bin Y_i is assumed to be Gaussian distributed with the generated parameters.

After computing the mutual information between the parameters and the measurement for each time-frequency bin, we also estimate the number of independent time-frequency bins (intrinsic dimension) $\hat{n}$.

We normalize the test spectrogram (measurement) of each individual, by subtracting the mean and dividing by the standard deviation of each bin. For every pair of individuals, with normalized spectrograms Y^k and Y^l, we compute the sample correlation coefficient r_{kl} between the spectrograms. We use the r to $\dot{z}$ transformation attributed to R.A. Fisher [101, 125],

$$\dot{z}_{kl} = \frac{1}{2} \ln \left[\frac{1 + r_{kl}}{1 - r_{kl}} \right]. \tag{5.1}$$

The resulting random variable $\dot{z}$, with samples $\dot{z}_{kl}, k = 1, 2, \ldots, 285$ and $l = k + 1, k + 2, \ldots, 285$, is approximately Gaussian, with variance given by $\sigma^2 = \frac{1}{\hat{n}-3}$ [101, 125].

5.5 Estimation of Biometric Capacities

We consider four settings where we estimate the recognition capacity of the LDV biometric system.

5.5.1 Gaussian Distributed Mean, Constant Variance

In this setting, the distribution of the mean of a time-frequency bin is assumed to be Gaussian, while the variance of the bin is constant. We consider two cases. In the first case, we assume the variance σ_0^2 of all time-frequency bins is a constant σ_{c0}^2, while the mean μ_0 is Gaussian distributed with zero mean and variance $\sigma_{\mu 0}^2$. For the second case, we assume the variance σ_i^2 of the i-th time-frequency bin is a constant σ_{ci}^2, while the mean μ_i is drawn from a Gaussian distribution with zero mean and variance $\sigma_{\mu i}^2$.

For the case where the parameters of all the time-frequency bins are drawn from the same distribution, the mutual information between the parameters $X_0 = (\mu_0, \frac{1}{\sigma_0^2})$ and the measurement Y_0 is

$$I(X_0; Y_0) = \frac{1}{2} \log \left[1 + \frac{\sigma_{\mu 0}^2}{\sigma_0^2} \right], \tag{5.2}$$

where it is understood that $\log(.)$ represents $\log_2(.)$.

For the case of independent parameters for each time-frequency bin, the mutual information between the parameters $X_i = (\mu_i, \frac{1}{\sigma_i^2})$ of the i-th time-frequency bin, and the measurement Y_i is given as

$$I(X_i; Y_i) = \frac{1}{2} \log \left[1 + \frac{\sigma_{\mu i}^2}{\sigma_i^2} \right]. \tag{5.3}$$

5.5.2 Constant Mean, Gamma Distributed Inverse Variance

In this setting, the reciprocal of the variance of a time-frequency bin is assumed to be Gamma distributed, while the mean of the bin is constant. We consider two cases. In the first case, we assume the mean μ_0 of all time-frequency bins is a constant μ_{c0}, while the reciprocal of the variance $\frac{1}{\sigma_0^2}$ is Gamma distributed with shape and scale parameters α_0 and β_0 respectively. We chose a Gamma distribution for the reciprocal of the variance, because it is a conjugate prior (reproducing) density [122]. For the second case, we assume the mean μ_i of the i-th time-frequency bin is a constant μ_{ci}, while the reciprocal of the variance $\frac{1}{\sigma_i^2}$ is Gamma distributed with shape and scale parameters α_i and β_i respectively.

The mutual information between the parameters $X_0 = (\mu_0, \frac{1}{\sigma_0^2})$ and the measurement Y_0 is

$$I(X_0; Y_0) = h(X_0) + h(Y_0) - h(X_0, Y_0), \tag{5.4}$$

where

$$
\begin{aligned}
h(X_0) &= \Bigg[\alpha_0 + \ln(\beta_0) + \ln(\Gamma(\alpha_0)) \\
&+ (1-\alpha_0)\psi(\alpha_0)\Bigg] \log e, \\
h(Y_0) &= \log e \Bigg[\ln\left(\frac{\pi^{\frac{1}{2}}\Gamma(\alpha_0)}{2^{\alpha_0}\beta_0^{\frac{1}{2}}\Gamma(\alpha_0+\frac{1}{2})}\right) \\
&+ \left(\alpha_0+\frac{1}{2}\right)\int_{-\infty}^{\infty} \ln(2+\beta_0(y-\mu_0)^2) \\
&\quad \left(\frac{2^{\alpha_0}\beta_0^{\frac{1}{2}}(2+\beta_0(y-\mu_0)^2)^{-\frac{1}{2}-\alpha_0}}{\pi^{\frac{1}{2}}}\right) \\
&\quad \left(\frac{\Gamma(\alpha_0+\frac{1}{2})}{\Gamma(\alpha_0)}\right)dy\Bigg], \\
h(X_0,Y_0) &= \Bigg[\ln\left((2\pi)^{\frac{1}{2}}\beta_0^{\alpha_0}\Gamma(\alpha_0)\right) + \frac{1}{2} + \alpha_0 \\
&+ \left(\frac{1}{2}-\alpha_0\right)\left(\psi(\alpha_0)+\ln(\beta_0)\right)\Bigg] \log e,
\end{aligned}
$$

and $\psi(.)$ is the digamma function, the derivative of the natural logarithm of the Gamma function.

For the case of independent parameters for each time-frequency bin, the mutual information between the parameters $X_i = (\mu_i, \frac{1}{\sigma_i^2})$ and the measurement Y_i is given as

$$I(X_i;Y_i) = h(X_i) + h(Y_i) - h(X_i,Y_i), \tag{5.5}$$

where

$$
\begin{aligned}
h(X_i) &= \Bigg[\alpha_i + \ln(\beta_i) + \ln(\Gamma(\alpha_i)) \\
&+ (1-\alpha_i)\psi(\alpha_i)\Bigg] \log e, \\
h(Y_i) &= \log e \Bigg[\ln\left(\frac{\pi^{\frac{1}{2}}\Gamma(\alpha_i)}{2^{\alpha_i}\beta_i^{\frac{1}{2}}\Gamma(\alpha_i+\frac{1}{2})}\right) \\
&+ \left(\alpha_i + \frac{1}{2}\right)\int_{-\infty}^{\infty} \ln(2+\beta_i(y-\mu_i)^2) \\
&\quad \left(\frac{2^{\alpha_i}\beta_i^{\frac{1}{2}}(2+\beta_i(y-\mu_i)^2)^{-\frac{1}{2}-\alpha_i}}{\pi^{\frac{1}{2}}}\right) \\
&\quad \left(\frac{\Gamma(\alpha_i+\frac{1}{2})}{\Gamma(\alpha_i)}\right) dy\Bigg], \text{ and} \\
h(X_i, Y_i) &= \Bigg[\ln\left((2\pi)^{\frac{1}{2}}\beta_i^{\alpha_i}\Gamma(\alpha_i)\right) + \frac{1}{2} + \alpha_i \\
&+ \left(\frac{1}{2}-\alpha_i\right)\left(\psi(\alpha_i) + \ln(\beta_i)\right)\Bigg] \log e.
\end{aligned}
$$

5.5.3 Gaussian Distributed Mean, Gamma Distributed Inverse Variance

In this setting, the reciprocal of the variance of a time-frequency bin is assumed to be Gamma distributed, while the mean of the bin is assumed to be Gaussian. Similarly, we consider two cases. In the first case, we assume the mean μ_0 of all time-frequency bins is Gaussian distributed with zero mean and variance $\sigma_{\mu 0}^2$, while the reciprocal of the variance $\frac{1}{\sigma_0^2}$ is Gamma distributed with shape and scale parameters α_0 and β_0 respectively. For the second case, we assume the mean μ_i of the i-th time-frequency bin is Gaussian distributed with zero mean and variance $\sigma_{\mu i}^2$, while the reciprocal of the variance $\frac{1}{\sigma_i^2}$ is Gamma distributed with shape and scale parameters α_i and β_i respectively.

For the case of independent parameters for each time-frequency bin, the conditional differential entropy of Y_i given X_i is

$$h(Y_i|X_i) = E_{X_i}\left[\frac{1}{2}\log(2\pi e\sigma_i^2)\right], \tag{5.6}$$

$$= \frac{\log e}{2}[\ln(2\pi e) - (\psi(\alpha_i) + \ln(\beta_i))]. \tag{5.7}$$

There is no closed-form expression for the differential entropy of Y_i, so it is given in integral form as

$$h(Y_i) = -\int_{-\infty}^{\infty} p_{Y_i}(y)\log p_{Y_i}(y)dy, \tag{5.8}$$

where $p_{Y_i}(y) = \int_0^\infty \frac{(\frac{\lambda_i}{\beta_i})^{\alpha_i}\exp(-\lambda_i(\frac{1}{\beta_i}+\frac{y^2}{2+2\lambda_i\sigma_{\mu i}^2}))}{\sqrt{2\pi}\sqrt{\lambda_i(1+\lambda_i\sigma_{\mu i}^2)}\Gamma(\alpha_i)}d\lambda_i$ and $\lambda_i = \frac{1}{\sigma_i^2}$.

The mutual information between X_i and Y_i is given as

$$I(X_i;Y_i) = h(Y_i) - h(Y_i|X_i). \tag{5.9}$$

5.5.4 Nonparametric Kernel Density Estimation

In this setting, the probability density function of the parameters of each time-frequency bin is estimated using a nonparametric kernel density estimation method. This means that the density is not assumed to be in any particular family of distributions. For the spectrograms of each individual, we assume that the i-th time-frequency bin is Gaussian distributed with parameters given by mean μ_i and variance σ_i^2. Also, we assume that the time-frequency bins are independent. However, for bin i the joint density of (μ_i, σ_i^2) is not assumed to be in any particular family of distributions; it is to be estimated nonparametrically. Here, we use the kernel density estimation method, where given a set of data $(y_i, \cdots, y_n)$, the density is estimated as

$$p(y) = C\sum_{i=1}^{n} f(h^{-1}(y - y_i)), \tag{5.10}$$

where f is a kernel function, usually smooth and unimodal, whose integral is finite, h is called the width parameter, and C is a normalizing constant such that $p(y)$ is a probability density. The idea is that the closer y is to an available data point, the higher the density should be, and if y is close to many data points, the density should be even higher. Putting a unimodal function around each data point and summing over all the effects yields the kernel density estimate. We use the two dimensional Gaussian kernel proposed by Fukunaga, and Silverman [37, 110] to estimate the joint density. The Gaussian kernel is defined as

$$f(y, y_i) = e^{-\frac{h^{-2}(y-y_i)^{\mathrm{T}}\hat{S}^{-1}(y-y_i)}{2}}, \tag{5.11}$$

where $\hat{S}$ is the empirical covariance matrix of $(y_1, \cdots, y_n)$. A reasonable choice of h for a Gaussian kernel, discussed in [110], is

$$h = (\frac{4}{n(2d+1)})^{\frac{1}{d+4}}, \tag{5.12}$$

where d is the dimension of the data, which in our case equals 2. Figure 5.1 shows the estimated density using the described nonparametric density estimator for an example of 400 data points drawn from a mixture of two uniform distributions with means at $(0.5, 0.5)$ and $(5.5, 0.5)$, and the width of both axes of both uniform distributions is 1. Note that the density estimate is smoother than the underlying true distribution in this case, which is particularly due to the uniform distribution having discontinuous regions, while the kernel used is smooth.

5.6 Results

After computing the mutual information for each time-frequency bin, we represent the result as an image by reshaping the mutual information vector into a mutual information matrix, with 51 time and 29 frequency indices.

The mutual information for the case of constant variance and Gaussian distributed mean, computed by using equation (5.3), is given in Figure 5.2, when the parameters for each time-frequency bin are independent. The average mutual information over all the time-frequency bins is 0.1911 bits. The estimated intrinsic dimension, for this case of each time-frequency bin having a constant variance and Gaussian distributed mean, is $\hat{n} = 72$, to the nearest whole number. Based on the average mutual information and the estimated intrinsic dimension,

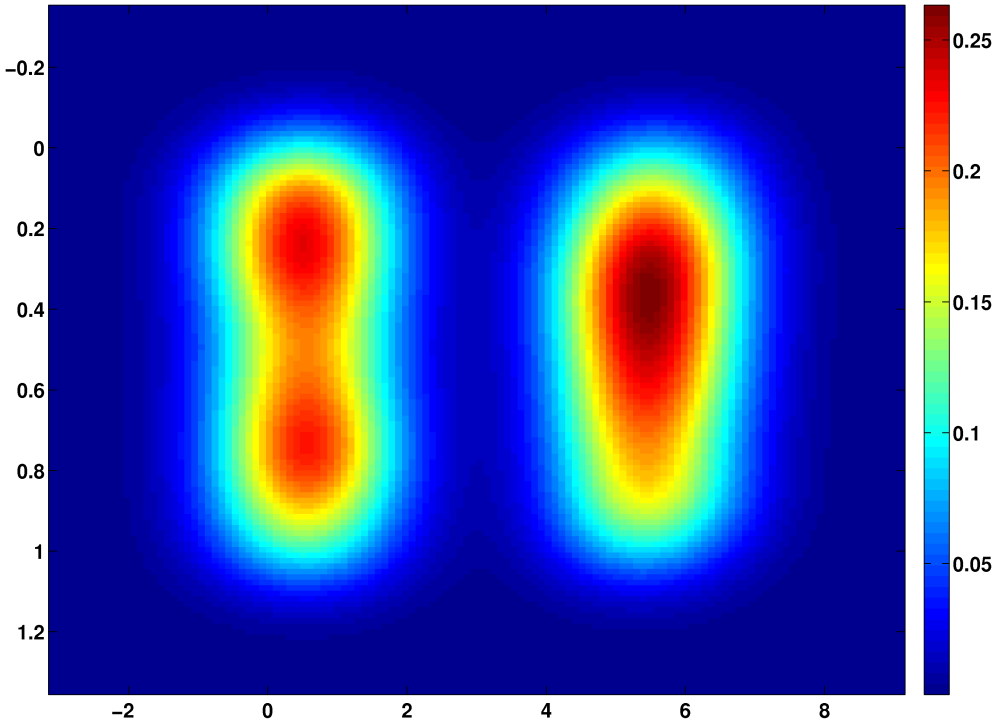


Figure 5.1: Density estimated using the proposed nonparametric method.

the biometric capacity of the LDV system is 13.76 bits. In particular, this result implies that variations between individuals that are based on the mean of each time-frequency bin can be captured using approximately 14 bits. When the parameters of all the time-frequency bins are drawn from a single distribution, the computed mutual information, given by equation (5.2), is 0.4647 bits.

The mutual information for the case of constant mean and Gamma distributed reciprocal-variance, computed by using equation (5.5), is given in Figure 5.3, when the parameters for each time-frequency bin are independent. The average mutual information over all the time-frequency bins is 0.093 bits. The estimated intrinsic dimension, for this case of each time-frequency bin having a constant mean and Gamma distributed reciprocal-variance, is $\hat{n} = 40$, to the nearest whole number. Based on the average mutual information and the estimated intrinsic dimension, the biometric capacity of the LDV system is 3.72 bits. In particular, this result implies that variations between individuals that are based on the variance of each time-frequency bin can be captured using approximately 4 bits. When the parameters of all the time-frequency bins are drawn from a single distribution, the computed mutual information, given by equation (5.4), is 0.4448 bits.

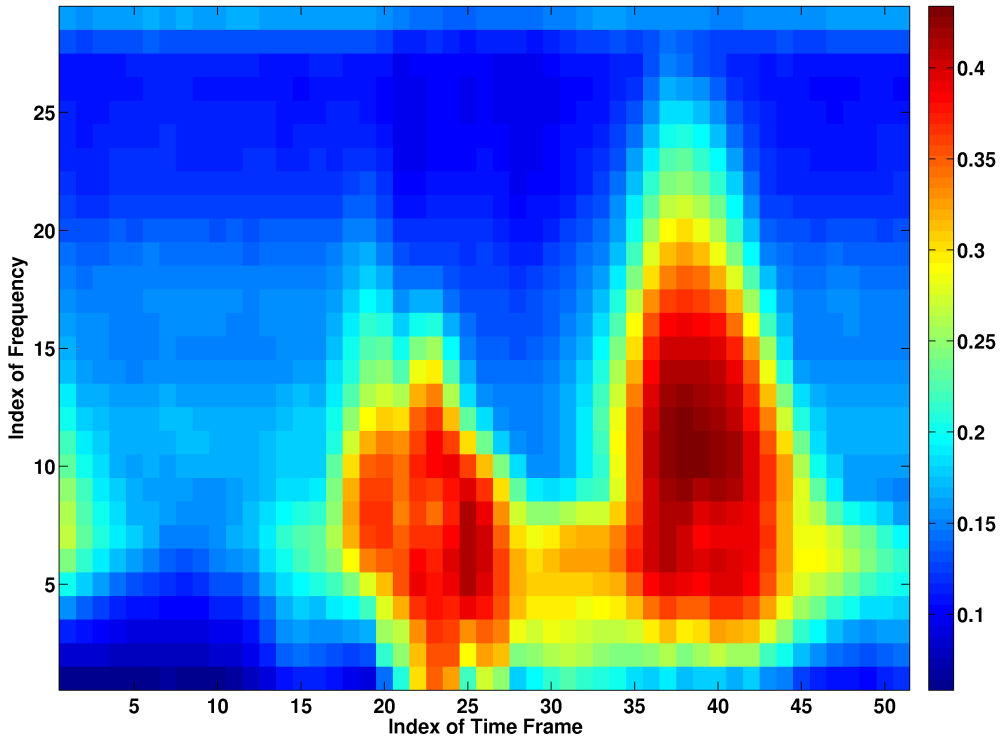


Figure 5.2: The mutual information between the mean (Gaussian distributed) and reciprocal-variance (constant) of each time-frequency bin and the value of each time-frequency bin of each spectrogram. The value of each time-frequency bin of each spectrogram is assumed to be Gaussian distributed.

The mutual information for the case of Gaussian distributed mean and Gamma distributed reciprocal-variance, computed by using equations (5.6) through (5.9), is given in Figure 5.4, when the parameters for each time-frequency bin are independent. The average mutual information over all the time-frequency bins is 0.406 bits. The estimated intrinsic dimension, for this case of each time-frequency bin having a Gaussian distributed mean and Gamma distributed reciprocal-variance, is $\hat{n} = 77$, to the nearest whole number. Based on the average mutual information and the estimated intrinsic dimension, the biometric capacity of the LDV system is 31.26 bits. In particular, this result implies that variations between individuals that are based on both the mean and variance of each time-frequency bin can be captured using approximately 32 bits. When the parameters of all the time-frequency bins are drawn from a single distribution, the computed mutual information is 1.9571 bits.

The mutual information computed by using the proposed nonparametric density estimate for the joint density of the mean and variance is shown in Figure 5.5. The average mutual information over all the time-frequency bins is 0.4114 bits. The estimated intrinsic dimension, for this case of each time-frequency bin having a mean and variance, whose joint distribution

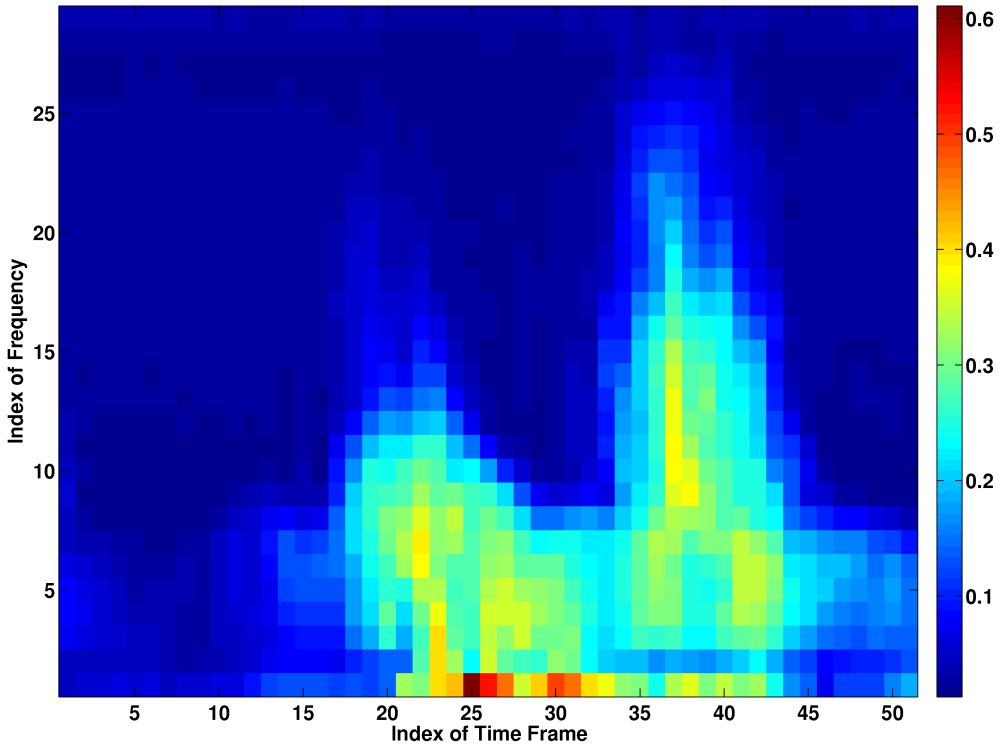


Figure 5.3: The mutual information between the mean (constant) and reciprocal-variance (Gamma distributed) of each time-frequency bin and the value of each time-frequency bin of each spectrogram. The value of each time-frequency bin of each spectrogram is assumed to be Gaussian distributed.

is estimated nonparametrically, is $\hat{n} = 77$, to the nearest whole number. Based on the average mutual information and the estimated intrinsic dimension, the biometric capacity of the LDV system is 31.68 bits. In particular, this result implies that variations between individuals that are based on both the mean and variance of each time-frequency bin can be captured using approximately 32 bits, when we do not assume a particular parameterized family of distributions for the mean and variance.

Table 5.1 summarizes the results obtained from the seven models discussed above. In the table, a large value of mutual information, intrinsic dimension, and biometric capacity is desirable. The larger the biometric capacity, the greater the number of recognizable patterns. From the table, we can conclude that the models that assume that all the time-frequency bins are drawn from the same distributions are worse than those that assume independent bins. In addition, the best models are the Gaussian mean, Gamma reciprocal variance model and the non-parametric mean and variance model.

Table 5.1: Summary of biometric capacity for mean and variance models.

Model	Average Mutual Information (bits)	$\hat{n}$	Biometric Capacity (bits)
Gaussian Mean, Constant Variance, Independent	0.1911	72	13.76
Gaussian Mean, Constant Variance, Same	0.4647	1	0.4647
Constant Mean, Gamma Reciprocal Variance, Independent	0.0930	40	3.72
Constant Mean, Gamma Reciprocal Variance, Same	0.4448	1	0.4448
Gaussian Mean, Gamma Reciprocal Variance, Independent	0.4060	77	31.26
Gaussian Mean, Gamma Reciprocal Variance, Same	1.9571	1	1.9571
Joint Non-parametric Mean and Variance	0.4114	77	31.68

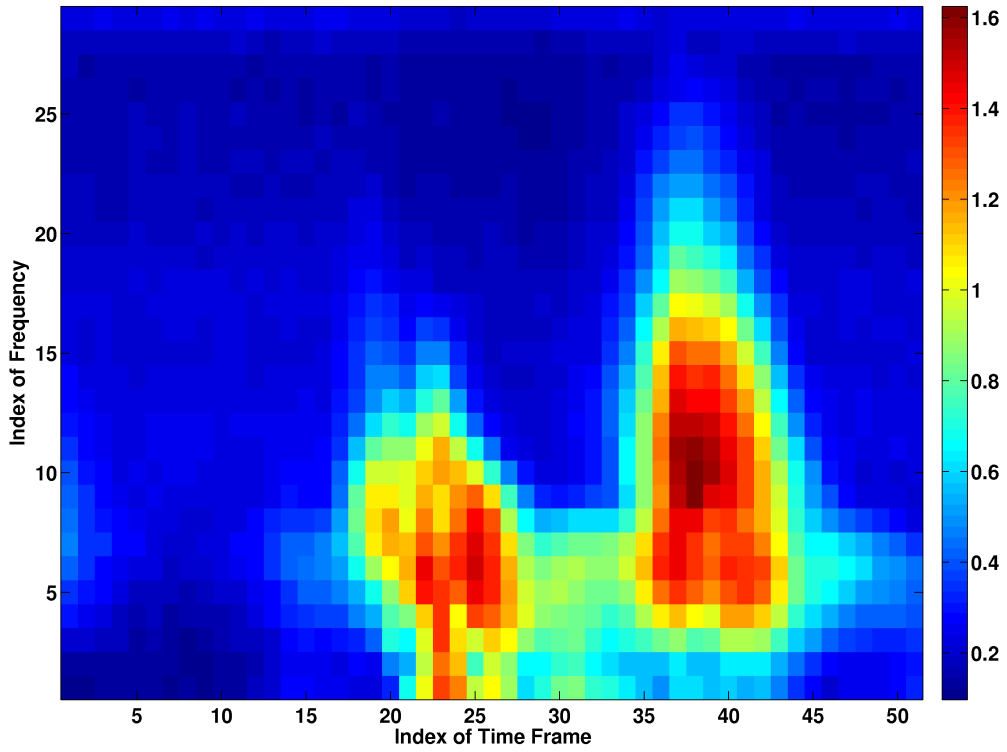


Figure 5.4: The mutual information between the mean (Gaussian distributed) and reciprocal-variance (Gamma distributed) of each time-frequency bin and the value of each time-frequency bin of each spectrogram. The mean and variance are assumed to be independent, and the value of each time-frequency bin of each spectrogram is assumed to be Gaussian distributed.

From the results we obtained, we observe that the biometric recognition capacity of the LDV system may be over-estimated. This can be attributed to errors in estimating the source (parameter) p_x and channel $p_{y|x}$ distributions, difficulty in estimating the intrinsic dimension of the time-frequency bins, and the assumption of independence between every pair of time-frequency bins. The over-estimation of the recognition capacity is less likely due to errors in estimating $p_{y|x}$. From empirical data, the short-time frequency decomposition of the LDV pulse signal results in coefficients, whose logarithm of the squared magnitude is approximately Gaussian.

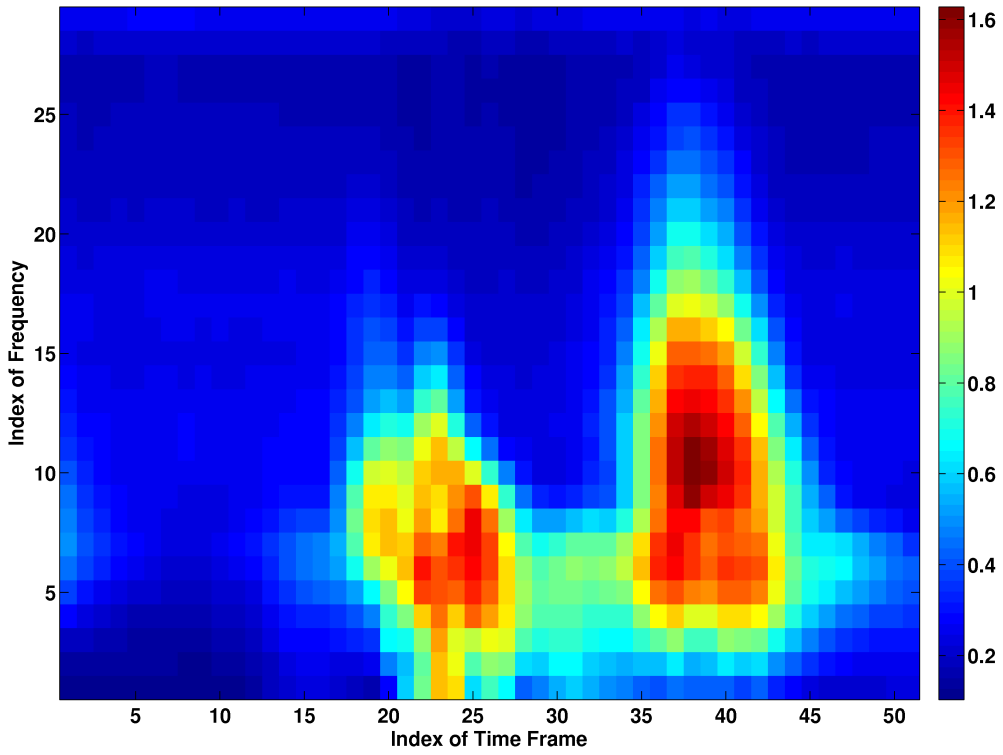


Figure 5.5: The mutual information between the mean and variance of distributions of each time-frequency bin and the value of each time-frequency bin of each spectrogram. The joint density of the mean and variance of each bin is computed using the nonparametric density estimate, and the value of each time-frequency bin of each spectrogram is assumed to be Gaussian distributed.

5.7 Summary

In this chapter, we illustrate how the recognition bounds may be estimated using four different settings, from a simple parametric model to a nonparametric method. The chapter connects the information theoretic pattern recognition results to a practical problem involving human recognition. This work is potentially useful in estimating the number of individuals that can possibly be recognized, using a given biometric system, and it also provides a way to investigate the usefulness of each feature in recognition.

Moreover, it is of interest and an on-going research focus to incorporate dependency of features in order to obtain a better recognition capacity estimate, and to use the results as a means of feature selection in recognition problems with large alternatives.

Chapter 6

Biometric Recognition System Design Issues in the Presence of Hidden States

6.1 Highlights

In biometric recognition systems, one fundamental problem related to the use of model-based approaches is the derivation of the appropriate model. Models are typically trained using data that is representative of the problem of interest. While some biometrics do not vary with time or the state of the individual, a second fundamental problem is that some biometrics do change either with aging or with state. In the latter case, effective models are robust enough to perform well in the presence of conditions that are not represented in the training data. We motivate the need for a theoretical basis using the results of a large study we have recently performed on recognition using electrocardiogram (ECG) signals. We propose an analytical framework that captures random hidden states. We propose an analytical basis for the use of more robust models for biometrics. We speculate on the impact of these observations for biometric system design.

6.2 Introduction

The use of biometric recognition systems has grown tremendously over the past few years. As the uses have grown, new biometrics have been proposed, including recognition based on

the shape of the heartbeat, either electrically through the electrocardiogram (ECG) or mechanically through the use of laser Doppler vibrometry (LDV) signals or phonocardiograms (heart sounds). ECG and LDV signals depend on the state of the individual. Gross changes in the signals occur due to changes in heart rate. Other changes are more subtle and may be based upon breathing phase (short term) and consumption of stimulants or depressants (longer term). Similarly, in face recognition, there are changes in the state of the face, either through pose or expression (short term) or through facial hair, make-up, and aging (long term). Effective models for biometrics with hidden states must be robust with respect to the set of possible states.

Robustness to hidden states can be accomplished in a number of ways. We focus on two fundamental, intertwined approaches to robustness. One approach is through a set of measurements that span the set of hidden states, thereby training the model on a sufficient set of states to be able to perform well when new states are observed. A second approach is to train a model on a small number of states, but use a distribution that can extrapolate from these few states to the many possible states.

Note that models trained on only one state have limited ability to extrapolate to unknown states. We have shown this through our study of performance of ECG recognition systems [88] in Chapter 3. As shown in Table 6.1, nearly all algorithms performed well within session. That is, when trained and tested on data collected within a single data recording session, nearly all algorithms performed well. However, as depicted in Table 6.2, when the state is allowed to vary, in this case simply by collecting data a couple of weeks separated in time, then all algorithms performed less well (all algorithms performed worse across sessions than within session).

What is most intriguing is the performance when biometric recognition algorithms are allowed to be trained on multiple sessions. The goal of using multiple sessions is to allow the algorithms to see a greater variety of states, thereby increasing performance when an unknown state is presented to the system. In our study, most algorithms saw a negligible increase in performance, as portrayed in Table 6.3. One algorithm that performed well was designed by our group [87] with the goal of achieving this robustness. Another algorithm that performed well was a neural network model by Wan and Yao [124].

The goal of this chapter is to present an analytical framework that captures two important aspects of the performance. First, the framework should demonstrate that algorithms can perform very well within session, yet suffer in performance dramatically across sessions.

Table 6.1: Within-session authentication performance. Each individual's model was trained by using 32 heartbeats, while 32 other heartbeats, per individual were used in testing. The equal error rate is a point on the receiver operating characteristic (ROC) curve, where the false acceptance rate equals the false rejection rate of a biometric recognition system. This is an excerpt from a table given in [88], where we considered a few more algorithms and training and testing scenarios.

Researchers	Equal Error Rates (%)
Agrafioti *et al.* [4]	0.57
Chiu *et al.* [24]	1.76
Irvine *et al.* [49]	1.26
Khalil and Sufi [62]	1.56
Li and Narayanan [73]	1.17
Odinaka *et al.* [87]	0.38
Wan and Yao *et al.* [124]	0.75
Wang *et al.*[126]	1.74
Wübbeler *et al.* [130]	0.57
Ye *et al.* [134]	1.64

Table 6.2: Across-session authentication performance. Each individual's model was trained by using 32 heartbeats from one session, while 32 heartbeats, per individual, from another session, were used in testing. This is an excerpt from a table given in [88], where we considered a few more algorithms and training and testing scenarios.

Researchers	Equal Error Rates (%)
Agrafioti *et al.* [4]	10.36
Chiu *et al.* [24]	26.36
Irvine *et al.* [49]	21.57
Khalil and Sufi [62]	21.13
Li and Narayanan [73]	18.16
Odinaka *et al.* [87]	11.30
Wan and Yao [124]	19.22
Wang *et al.* [126]	17.61
Wübbeler *et al.* [130]	15.77
Ye *et al.* [134]	18.55

Table 6.3: Across-session (with multiple training sessions) authentication performance. Each individual's model was trained by using 32 heartbeats from two sessions (16 heartbeats from each session), while 32 heartbeats, per individual, from another session were used in testing. This is an excerpt from a table given in [88], where we considered a few more algorithms and training and testing scenarios.

Researchers	Equal Error Rates (%)
Agrafioti *et al.* [4]	9.51
Chiu *et al.* [24]	20.97
Irvine *et al.* [49]	19.22
Khalil and Sufi [62]	18.53
Li and Narayanan [73]	17.09
Odinaka *et al.* [87]	5.47
Wan and Yao [124]	6.28
Wang *et al.* [126]	15.93
Wübbeler *et al.* [130]	13.98
Ye *et al.* [134]	13.67

Second, it should demonstrate the increase in performance when trained on multiple sessions that contain multiple states. This idea in itself is not new, but is often neglected in derivation of biometric recognition algorithms. We describe a simple framework for analyzing such systems and performance gains in idealized cases.

In Section 6.3, we describe the analytical framework for hidden states. The parameters of a prior on the state become the biometric parameters of interest. In Section 6.4, we describe performance gains based on multiple training sessions, in particular, two training sessions. A summary of the chapter is provided in Section 6.5.

6.3 Analytical Framework for Biometric Recognition with Hidden States

Our previous work [89] on estimating the biometric capacity of a system gave estimates for the number of individuals that the system can handle and still guarantee exponential recognition rates, assuming known models for the data. However, based on experimental results, the recognition rates are not exponential, especially when the training and testing data are from different sessions. Within a single session, the model can be improved by using several realizations of the biometric signature, which improves the recognition system performance. However, when the testing data come from a different session, the model

improvement due to several realizations of the biometric signature is not commensurate with the performance improvement. One possible explanation for this is that the underlying state of the individual changes from session to session.

6.3.1 Analytical Framework with Hidden States

Suppose that individuals are indexed by $i \in \{1, 2, \cdots, I\}$. Given individual i and state s, there is a probability density function on the data given by $p(x|s, i)$, where x should be interpreted as the biometric signature (for example, a facial image or a cardiac signal). We concentrate our analysis on the case where given the state s, the data are independent of the individual i, that is $p(x|s) = p(x|s, i)$. This model is simplistic in that the transformation from the hidden state to the data is independent of the individual.

Let the session number be indexed by $m \in \{1, 2, \cdots, M\}$, the biometric signature number (within a session) be indexed by $j \in \{1, 2, \cdots, J\}$, and the signal points within each signature be indexed by $n \in \{1, 2, \cdots, N\}$. We assume that the sessions are independent, and given the state s, the biometric signature realizations are independent within a session. Moreover, we assume that the signal points within a signature are independent. Furthermore, we assume that data recorded in a single session has a single realization of the hidden state. Figure 6.1 is a rendition of the framework for biometric recognition with hidden states. This framework is an improvement over our previous one [89], which only considers what happens within a single session.

For individual i, the joint distribution across sessions and biometric signature realizations is given as

$$P(i) \prod_{m=1}^{M} \left(\prod_{n=1}^{N} \left\{ \left[\prod_{j=1}^{J} p(x_{njm}|s_{nm}) \right] p(s_{nm}|i, \theta_n) \right\} \right), \tag{6.1}$$

where θ is the individual's parameter vector, $s_{nm} \in \mathbb{R}$, $\mathbb{R}^2$, or $\mathbb{R} \times \mathbb{R}_{\geq 0}$, $x_{njm} \in \mathbb{R}$.

The problem we consider is that of biometric recognition where the goal is to determine the posterior probability of each individual given testing data $P(i|x)$. For detection (classification) problems that employ the plug-in rule, during a training phase, the individual's parameter vector is estimated. During the test phase, the test biometric signature and estimated parameter vector are plugged into the data likelihood function.

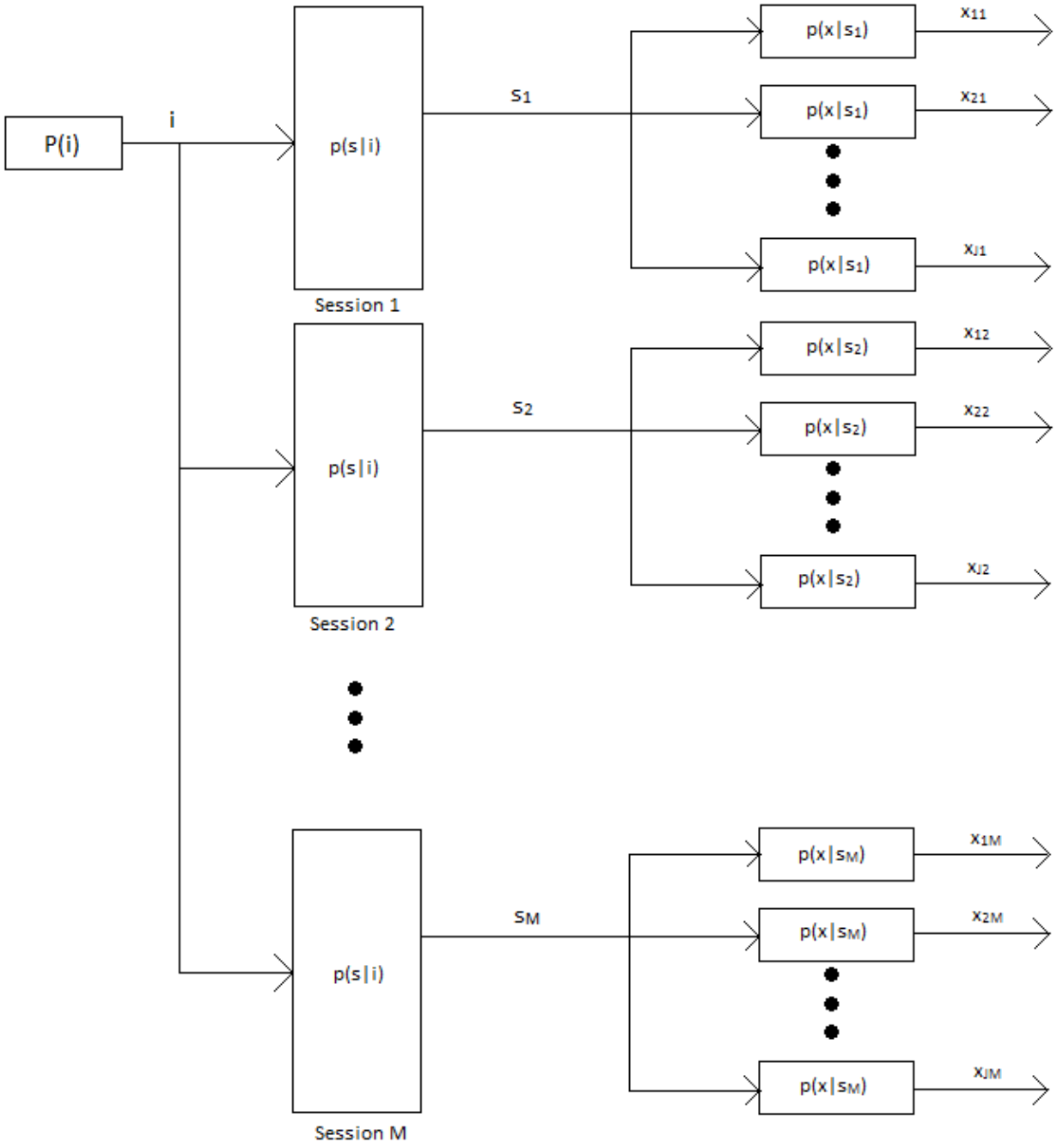


Figure 6.1: Framework for biometric recognition with hidden states.

There are several strategies to training a biometric recognition system, including:

1. For each session m, estimate the realization of the hidden state $\{s_{nm}\}$, ignoring the prior distribution of the states given the individual. Given the state estimates $\{\hat{s}_{nm}\}$, the parameters $\{\theta_n\}$ can be estimated.

2. Directly estimate $\{\theta_n\}$ from the marginal distributions by summing the joint distribution over the states, or jointly estimate $\{s_{nm}\}$ and $\{\theta_n\}$ using an expectation-maximization (EM) algorithm.

3. Use a Bayesian approach of assigning a universal prior on θ_n, and then jointly estimate $\{s_{nm}\}$ and $\{\theta_n\}$. An EM algorithm can be used to this end.

Comments on these approaches:

1. In the estimation for each s_{nm} using J independent and identically distributed random variables $\{x_{njm}, j = 1, 2, \cdots, J\}$, the variance in the error is proportional to $1/J$. For small J, this error can be significant. Even for moderate values of J (on the order of 100), the standard deviation may be a significant fraction of the value of the estimate $\hat{s}_{nm}$. Given typical cases for problems of interest to us where the number of training sessions is small, estimates of θ_n have even larger errors. For only one training session (as is the case in much of the literature), there is only one state ($M = 1$), the worst possible case. In this sense, using a correct model and applying any methodology similar to this is doomed to have low performance for data drawn from an unobserved state.

   We note that this approach is also consistent with the use of mixture models for the data. If we know that the data come from different states and that each session corresponds to a different state, then a reasonable approach is to estimate one probability density function for each session. These probability density functions are then mixed, so that the resulting density function is a sum of components, one from each session. We have studied this approach extensively for ECG and LDV recognition and have found that it does not work well. In particular, these mixture models still suffer from an inability to extrapolate to data from unobserved states.

2. Taking into account the complete probability density function and estimating (perhaps using maximum likelihood estimation) the variables $\{\theta_n\}$ should yield better performance than the first approach. However, in the presence of only one observed state

in the training data, the estimation performance is lower bounded by the performance of a genie that directly observes $\{s_{nm}\}$ (this performance is achieved under the prior approach or this approach as J gets large). That is, one nonrandom parameter is estimated using one random variable.

3. Using a Bayesian approach with a universal prior on θ_n can provide better performance when extrapolating to unobserved states. There is a tradeoff between a prior on the parameters $\{\theta_n\}$ and the data that is necessarily limited due to a small number of observed states.

6.3.2 Approximate Framework Integrating Out Hidden States

One strategy to training a biometric recognition system that we have found to be quite effective in both ECG-based and LDV-based biometrics involves using the following joint distribution, and estimating the parameters θ_n:

$$P(i) \prod_{m=1}^{M} \prod_{j=1}^{J} \prod_{n=1}^{N} \tilde{p}(x_{njm}|i, \theta_n), \tag{6.2}$$

where θ is the individual's parameter vector. This model can be interpreted as using the conditional distributions on the data that result from integrating out the hidden states $\{s_{nm}\}$ in the joint conditional density function

$$p(x_{njm}|s_{nm})p(s_{nm}|i, \theta_n). \tag{6.3}$$

This is clearly different from the model in (6.1). Thus, we do not interpret the conditional density $\tilde{p}(x_{njm}|i, \theta_n)$ as directly corresponding to the result of this integral.

For such an approach to be successful, $\tilde{p}$ should be a robust distribution, so that it can generalize to unobserved (not in the training set) states. That is, if the unobserved state is close enough to the training set states, the trained model must perform well. Moreover, the model has to be chosen so that it does not compromise the discriminability between individuals. In order words, the model needs to preserve discriminability while maintaining robustness.

The following are examples of models for the state and biometric signature that we have used for our idealized analytical models:

1. s is a Gaussian vector, with zero mean and independent entries; x is a Gaussian vector, with mean s and conditionally independent entries with known variances; given individual i, the parameter vector may be the variances of the entries of s.

2. s is a vector of independent Gamma-distributed random variables; x is a Gaussian vector, with zero mean, independent entries, and variances $\{1/s_{nm}\}$; given individual i, the parameters of the Gamma distributions comprise the parameter vector.

3. s is a pair of vectors $(\mu_{nm}, 1/\sigma_{nm}^2)$, where μ_{nm} are independent Gaussian random variables and $1/\sigma_{nm}^2$ are independent Gamma distributed random variables, independent of $\{\mu_{nm}\}$; x is a Gaussian vector with independent entries, with means μ_{nm} and variances σ_{nm}^2; given individual i, the parameters are those for the Gaussian distributions for $\{\mu_{nm}\}$ and the Gamma distributions for $\{1/\sigma_{nm}^2\}$.

The first example assumes that the biometric signature is Gaussian, and that the mean vector is a realization of a zero-mean Gaussian random state s. The second example assumes that the biometric signature is a zero-mean Gaussian, whose variance is a realization of an inverse Gamma distribution. The third example assumes the Gaussian biometric signature has mean and variance vectors that are realizations of a zero-mean Gaussian and an inverse Gamma distribution, respectively. These model examples are reminiscent of the cases we considered in a previous study [89].

6.4 Multiple Training Sessions

Barring the choice of model, with an adequate amount of training data, the realization of the hidden state s_{nm} can be estimated. When data from the same session are used during the testing phase, since the underlying state is the same, biometric recognition performance is expected to be excellent. This is reflected in Table 6.1, where some of the algorithms achieve an equal error rate below 1%.

However, when data from one session are used in training and data from another session are used in testing, since the two different sessions may potentially possess two different realizations of the hidden state, it is not surprising that the biometric recognition performance

degrades, as shown in Table 6.2. To improve this performance loss due to differences in hidden state realizations between training and testing, one needs multiple training sessions and a model that is able to learn from the known realizations of the hidden state and extrapolate to other unknown realizations. Such a model is said to be robust to hidden states.

One approach to achieving robustness to hidden states is through a set of measurements that span the set of hidden states, thereby training the model on a sufficient set of states to be able to perform well when new states are present. Another approach is to train a model on a small number of states, but use a distribution that can extrapolate from these few states to the many possible states. Consider the model examples given in the previous section. If the biometric recognition system is presented with a very large number of training sessions, and a very large number of biometric signature realizations per session, it can estimate both the hidden state realizations for each session, and then the distribution of the hidden state using all the hidden state realizations. This is the idea behind the first approach. It is a rather expensive approach to recognition with dynamic biometric signatures.

The first approach requires a large number of sessions in order to obtain as many realizations of the hidden state as possible, so that when the system is presented with a new state realization, one of the observed states will be close enough to it, and result in a successful recognition. On the other hand, the success of the second approach relies more on the nature of the distribution than on the number of sessions. By using a robust distribution (one with heavy tail(s)), the model can be trained on a small number of sessions (hidden state realizations). The robust distribution allows the model to be able to generalize to new hidden state realizations. Rather than trying to fit a distribution to the data from each training session separately, the second approach pools the data together, and utilizes a robust distribution to fit the data. In essence, instead of modeling each individual realization of the hidden state, the average hidden state realization is modeled.

The second approach to attaining robustness to hidden states was utilized by our group in previous studies [87, 88], where we showed that the performance degradation that occurs when training and testing across sessions, can be reduced by using a log-normal (robust) distribution to train on two sessions. This performance improvement can be seen in Table 6.3, where an equal error rate below 6% (in contrast to the 11.3% EER from Table 6.2) was achieved.

In designing a biometric system, due to the costs associated with obtaining measurements of the biometric signature for training, one needs to determine how many sessions is adequate in

training the models and how long each session should last. In a previous study [87], we showed that as the number of realizations of the biometric signature increases, so does the recognition performance. We also showed that using multiple sessions for training drastically improves the recognition performance. The improvement in performance due to using multiple training sessions was more significant than that attributed to increasing the number of realizations of the biometric signature within a session. We speculate that in designing a biometric system where the biometric signature changes with state or aging, more training sessions will ultimately lead to improvement in performance. As such, during training, one could elicit different states from the individual, or collect data from different days to allow for aging.

6.5 Summary

Biometric recognition systems that rely on model-based approaches are faced with the fundamental problem of choosing the appropriate model. The choice of model depends on the particulars of the problem. A second fundamental problem that plagues some biometric recognition systems is that the biometric signature varies with time or the state of the individual. Effective models are robust enough to state or time changes. That is, they perform well when presented with conditions that are not captured in the training data.

We propose analytical and approximate frameworks that incorporate the random hidden states. These frameworks serve as a basis for the use of more robust models for biometrics.

The recognition rates that are observed for a system that uses biometric signatures that change with state and aging are far from exponential as implied by our previous work on estimating the biometric capacity of a system [89] given in Chapter 5, especially when the training and testing data are from different sessions. To improve the system performance, biometric systems often rely on training data from multiple sessions.

A biometric system can be trained on a large number of representative sessions which will allow it to perform well when it encounters biometric signatures with a new underlying state. However, when the number of training sessions are limited, one has to rely on robust distributions that can extrapolate to unknown states. This approach has been applied by our group to the case where there are only two training sessions available to an ECG biometric recognition system. By training on two sessions, using a log-normal distribution, an improvement in recognition performance of about 50% was obtained, in comparison to training

on biometric signatures from a single session. This has important consequences for biometric system design; more efforts should be channeled towards choosing a robust distribution and obtaining training data across sessions, than within a single session.

Chapter 7

Computational Imaging Approach for Fan Beam X-ray Coherent Scatter Imaging

This chapter will appear in a journal publication. It is currently under internal review [90].

7.1 Highlights

In X-ray scatter imaging, tomographic measurements of the forward scatter distribution are used to infer scatter densities within a volume. A mask placed between the object and the detector array provides information about scatter angles. An efficient computational implementation of the forward and backward model facilitate iterative algorithms based upon a Poisson log-likelihood. Results, validating the forward model and the reconstruction algorithm, are presented on simulated and Monte Carlo data. The design of the scatter imaging system influences the algorithmic choices. Moreover, the system design is guided by the need for efficient algorithms.

7.2 Introduction

X-ray scatter imaging reconstructs a volumetric scatter density from tomographic measurements of the scattered X-ray data. For objects comprised of particles, the forward scatter distribution is modeled as a superposition over all source angles and source intensities of the scatter from all object points at those source angles, at the corresponding scatter angles.

The detailed model is given below, but it is immediately apparent that the central challenge in X-ray scatter imaging is the separation of the scatter back into a volumetric scatter density. Two important design elements are a primary mask (for beam shaping) between the source and the object, and a secondary mask (to block direct X-rays and disambiguate scatter angles) between the object and the detector array.

In its simplest form, the primary mask has a single small hole, resulting in a pencil beam distribution for the source. An array of pencil beams may be selected. Alternatively, a continuous set of angles may be selected as in a fan beam source. This chapter describes the use of a fan beam source, but the techniques for modeling the measured data are applicable to a wide variety of sources distributions. A discussion of the system based on a pencil beam source can be found in the paper by MacCabe *et al.* [79].

The simplest secondary mask that restricts the measurements to scattered X-rays is complementary to the primary mask (a "complementary mask"). The array of holes or the aperture of the primary mask is magnified in the direction of the secondary mask to yield the secondary mask design. This complementary mask design does not yield a good design for any reasonable source design including a pencil beam, an array of pencil beams, and a fan beam. That the complementary mask fails is easily seen in the pencil beam case where there is a scatter angle at every point along the pencil beam for every ring on the detector around the beam-stop. Thus, for this design it is not possible to determine a unique scatter angle distribution for every point along the pencil beam.

More complicated secondary masks selectively block scatter angles. Consider a collection of holes or an aperture placed between the object and the detector array, in addition to the complementary mask. Scatter from a given object location through a given location on the aperture hits the detector at a known location. In inference for fan beam applications, given a measured detector location and an aperture location, the object location and its scatter angle are determined.

Mask design is based upon a combination of objectives. As many photons as possible should be allowed to hit the detector to increase the signal-to-noise ratio. However, a smaller number of holes in an aperture allows more accurate determination of object location and scatter angle. For every object location, some photons that get scattered at every angle should be allowed to hit the detector. Designs of masks range from intuitive to analytical. On the analytical side, there is a desire for a uniform sensitivity to a wide range of scatter locations and scatter angles. Such analytical designs are described elsewhere [13].

There is also a computational imaging aspect of the mask design. The mask and the inference algorithm may be designed at the same time as a pair, with a goal of optimizing a measure of the reconstruction performance. In this view, the ad hoc intuitive arguments for mask selection play a subsidiary role to quantitative measures of performance. While this is our ultimate goal, this chapter focuses on inference algorithm development from a penalized maximum likelihood estimation point of view.

The algorithm relies critically on a computational representation of the forward model for the data. Our forward model has been derived analytically, tested using Monte Carlo simulations, and compared against experimental measurements. The model is described in detail below, as is an efficient representation of the forward model based upon symmetry assumptions that are readily verifiable in analytical and Monte Carlo models, but less so in experiments.

The ultimate benefit of the model and approach described here will be determined in part by the calibration process for verifying alignment and symmetry. Even if the symmetry assumptions break down in practice, the remaining aspects of the forward model are valid. The computational implementation involves a tradeoff between on-line and off-line computations which has been optimized for our single processor, six core computer.

There is a connection between the computational time of the iterative algorithm and judicious choices for the sampling in the image domain. If the voxel size is chosen to be an integer multiple of the spacing of pixels on the detector, many computations can be reused. Furthermore, the symmetries available in the ideal case motivate a particular choice of subsets in an ordered subset expectation-maximization (EM) algorithm. This is described in detail below.

Analysis of the recovered resolution from simulated point objects with a single momentum transfer peak reveals that spectral resolution degrades as momentum transfer increases. However, statistical correlation analysis shows that each spatio-spectral voxel has a limited effect on voxels that are far away.

Finally, reconstruction results for simulated and Monte Carlo data with Poisson noise demonstrate the performance for particular choices of the secondary mask. The scatter spatial distribution and the momentum transfer profile are recovered accurately and quickly, with only a few ordered-subset EM iterations required.

The remainder of this chapter is organized as follows: In the next section, we describe the forward operator. Section 7.4, considers ways of efficiently computing the forward operator.

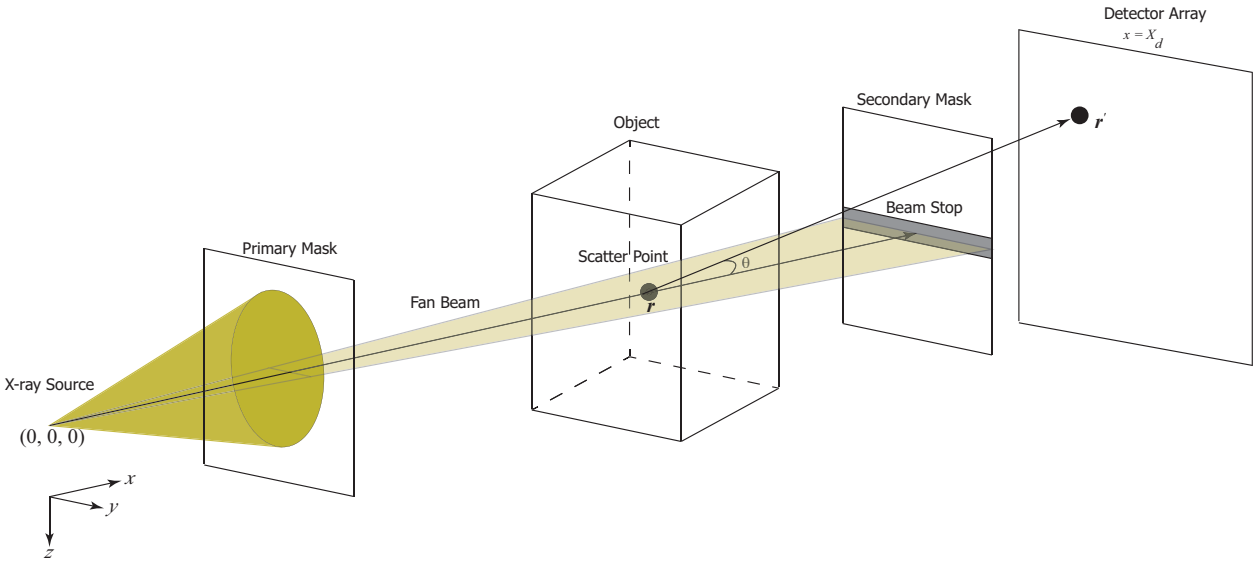


Figure 7.1: Schematic of the fan beam geometry.

Section 7.5 describes the backward operator. The ordered subset EM algorithm is developed in Section 7.6. In Section 7.7, we analyze the resolution of reconstructed momentum transfer values and inter-voxel correlation by using simulated data. The results of applying the ordered subset EM algorithm to simulated and Monte Carlo data are presented in Section 7.8. Section 7.9 summarizes the chapter.

7.3 Forward Operator

In the fan beam model, the X-ray source transmits X-rays in a fan within the $z = 0$ plane. The source is at the origin of the coordinate system. Object scatter locations are indexed by (x, y) coordinates and the corresponding object point $\mathbf{r} = [x, y, 0]$, with the positive x coordinate pointing from the source through the object to the secondary mask and the detector. The detector plane is perpendicular to the central ray of the X-ray fan which is oriented along the x axis. A point on the detector is $\mathbf{r}' = [X_d, y', z']$. The object point and the detector point determine the scatter vector $\mathbf{s} = \mathbf{r}' - \mathbf{r}$. The scatter vector and the object point determine the scatter angle θ. Figure 7.1 provides a schematic of the fan beam geometry.

The scattering follows a radiance model using ray propagation. The scatter density is a function of the object point and momentum transfer, $f(\mathbf{r}, q)$, where q is given in rad nm^{-1}.

Bragg's Law determines the relationship between the energy of the incident X-rays k, the scatter angle θ, and the momentum transfer q, by $q = 2k\sin(\theta/2)$.

For a point on the detector with incoming scatter vector, there are geometric factors due to the cosine of the incident angle and due to the squared distance to the object point, $\frac{|\mathbf{n}\cdot\hat{\mathbf{s}}|}{s^2}$, where $\mathbf{n}$ is a unit normal vector on the detector, $\hat{\mathbf{s}}$ is the unit vector in the direction of the scatter vector $\mathbf{s}$, and $s = |\mathbf{s}|$.

The secondary mask (coded aperture) is between the object and the detector array. It is assumed to be on a plane parallel to the detector array, centered on the middle of the fan beam array. There is a beam stop in front of the fan beam [78]. This secondary mask modulates the X-ray flux. Given the object point and detector point, the intersection of the scatter vector and the aperture is defined. Equivalently, given the detector point and a unit vector proportional to the scatter vector, the intersection with the aperture is defined. Denote the resulting modulation of the X-ray flux by $T(\mathbf{r}, \hat{\mathbf{s}})$. In practice, we estimate the values of the mask modulation using a calibration scan with a flood illumination. The resulting measured image is rescaled from the detector plane back to the secondary mask plane.

The source is assumed to be polychromatic. Whatever source filtering that is present is included in the model whose energy-dependent flux is $W(k)$.

Bringing together these effects yields the model for the flux at detector location $\mathbf{r}'$ as [76]

$$g(\mathbf{r}') = \int d\mathbf{r} \int dq\, H(\mathbf{r}', \mathbf{r}, q) f(\mathbf{r}, q), \tag{7.1}$$

where

$$H(\mathbf{r}', \mathbf{r}, q) = C \left| \frac{\mathbf{n}\cdot\hat{\mathbf{s}}}{s^2} \right| T(\mathbf{r}, \hat{\mathbf{s}}) \left(\frac{1}{2q\sin\frac{\theta}{2}} \right) W\left(\frac{q}{2\sin\frac{\theta}{2}} \right),$$

and C is a normalization constant.

7.4 Computation of the Forward Operator

In anticipation of our iterative algorithm for estimating the scatter density, efficient computational implementations of the forward and backward operators are required. Let us define

the source factor as $S(\theta, q) = \left(\frac{1}{2q \sin\frac{\theta}{2}}\right) W\left(\frac{q}{2\sin\frac{\theta}{2}}\right)$, and the geometry factor as $G(\mathbf{n}, \mathbf{s}) = \frac{|\mathbf{n}\cdot\hat{\mathbf{s}}|}{s^2}$. A straightforward implementation of the forward operator yields a computational structure given in Pseudocode 7.1. These nested loops indicate several opportunities for increasing the efficiency of an online implementation of the code. For large problems, some of the computations need to be performed efficiently and online, with as many computations reused in the code as possible.

Pseudocode 7.1 Raw computational structure for forward operator

```
Given object scattering density f(r, q)
for each object point r do
  for each detector point r′ do
    compute geometry factor G(n, s)
    compute scatter angle θ(r, s)
    compute mask modulation factor T(r, ŝ)
    for each momentum transfer q do
      compute source factor S(θ, q)
      accumulate source factor × object scattering density into result
    end for
    accumulate geometry factor × mask modulation factor × result
  end for
end for
```

Three ways to improve the efficiency of the code are highlighted here, namely: efficient implementation of the source factor times the scattering density, accompanied by interpolation; symmetries in the geometry factors; and the need to balance online and offline computations. Each subsequent pseudocode will introduce an element of efficiency to the raw pseudocode. We present the incremental pseudocode as we discuss each aspect of forward operator optimization.

7.4.1 Source Factor Implementation and Interpolation

The integral over the momentum transfer values is given by

$$\tilde{f}(\mathbf{r}, \theta) = \int dq\, S(\theta, q) f(\mathbf{r}, q). \tag{7.2}$$

This integral takes scattering density at a given object point as a function of momentum transfer and produces scattering density at that same point as a function of scatter angle.

Given a vector of momentum transfer values (in our simulations, we use around 250 samples of momentum transfer), we compute this integral using a matrix-vector multiply to get samples of scattering densities at a set of predefined scatter angles. In our simulations, we used around 250 scatter angles. When called afterward in the code, these samples are interpolated to the actual scatter angles.

The source matrix $S(\theta, q)$ is precomputed by using the predefined samples of momentum transfer and scatter angles. An example of the matrix used is shown in Figure 7.2. This corresponds to a filtered 125 kVp source. Note that such practical sources introduce blurring in the resulting scatter angles. That is, a monochromatic source would yield a single scatter angle for each momentum transfer. However, for a polychromatic source, each momentum transfer has a range of scatter angles that result, of known intensity. From the figure, we can see that the source matrix is smooth and slowly varying with θ and q. This suggests that interpolation will lead to fairly accurate values of the source factor, for a standard X-ray source (polyenergetic). Note that for a monoenergetic source, the source matrix is no longer slowly varying, so that interpolation is no longer a valid strategy. However, with such a source, each value of q will have a corresponding unique value of θ associated with it, making acceleration techniques such as interpolation, unnecessary.

The updated pseudocode for the computational structure that uses the precomputed source matrix and scatter angle interpolation is given in Pseudocode 7.2.

Pseudocode 7.2 Computational structure for forward operator using precomputed source matrix and scatter angle interpolation

Given object scattering density $f(\mathbf{r}, q)$
compute predefined momentum transfer q samples, scatter angle θ samples, and source factor matrix $S(\theta, q)$ for predefined q and θ samples
initialize detector image to zero
for each object point **r do**
  $Sf(\theta) = S(\theta, :) \times f(\mathbf{r}, :)$
  compute geometry factor image
  compute and interpolate scatter angle image
  interpolate $Sf(\theta)$ at scatter angle image
  compute mask modulation factor image
  detector image += geometry factor image × mask modulation factor image × interpolated $Sf(\theta)$
end for

From Pseudocode 7.2, we can see that the innermost momentum transfer loop is implemented as a matrix-vector product. The precomputation of the source factor matrix and

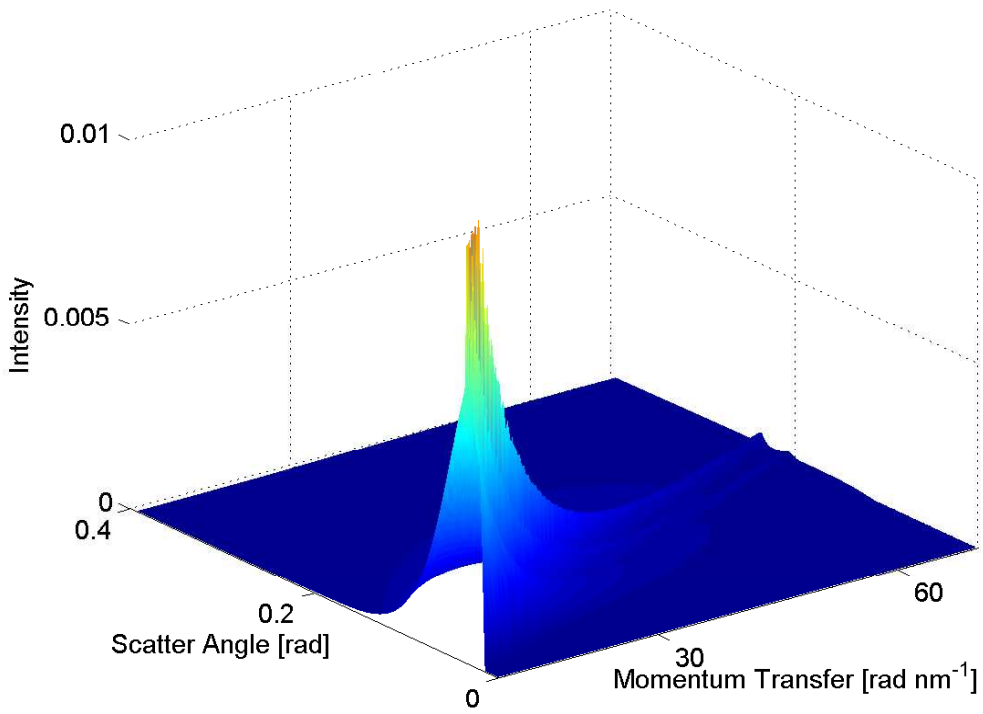


Figure 7.2: Source matrix that transforms scattering density as a function of momentum transfer to scattering density as a function of scatter angle. This is for a filtered 125 kVp source.

the interpolation of scatter angles avoids the computation of the source factor for each (θ, q) pair.

7.4.2 Geometry Factor Symmetries

The geometry factor relies on computation of the scatter vector $\mathbf{s} = \mathbf{r}' - \mathbf{r}$, which has length $s = |\mathbf{s}|$ and the cosine factor at the detector surface $|\mathbf{n} \cdot \hat{\mathbf{s}}|$. For our geometry, the cosine factor equals the magnitude of the x component of the corresponding unit vector $\hat{\mathbf{s}} = \frac{\mathbf{s}}{s}$.

Translation Symmetry

For the system envisioned in this chapter, there is a large rectangular array of small detector pixels. A typical detector may have an array of approximately 1500 by 2000 pixels with pixel widths of 0.19 mm. One choice in reconstruction is the reconstructed pixel widths in the object domain. These pixel widths should be motivated from a first principles analysis of achievable resolution of the final system. Even then, some flexibility in the exact pixel width remains. Suppose that the displacement between pixel centers in the y direction (across the source fan) is an integer multiple of the detector pixel widths. For our simulations, we use a factor of 16 to get 16*0.19 = 3.04 mm width in the y direction. Let $\mathbf{v}$ be a vector corresponding to the translation of a pixel in the object by one pixel width along the y direction. Then, the trivial equality

$$(\mathbf{r}' + \mathbf{v}) - (\mathbf{r} + \mathbf{v}) = (\mathbf{r}' - \mathbf{r}) = \mathbf{s} \tag{7.3}$$

implies that for each scatter vector from object location $\mathbf{r}$ to detector location $\mathbf{r}'$, there is a scatter vector from object location $\mathbf{r} + \mathbf{v}$ to detector location $\mathbf{r}' + \mathbf{v}$. Thus, given scatter vectors computed for one value of y in the object, the scatter vectors for adjacent object locations are nearly all determined, with only scatter vectors corresponding to detector pixels near the edge of the detector array needing to be computed. This gives a very efficient update for their computations.

Let us assume there are M detector rows (z-direction), N detector columns (y-direction), and TSF_y represents the number of y-directional detector pixels that correspond to 1 y-directional object pixel. Pseudocode 7.3 shows the effect translation symmetry has on the computational structure for the forward operator. From the pseudocode, we see that previous geometry

factor computations are reused, and only a miniscule fraction of the geometry factor image needs to be recomputed. For a 1500 by 2000 detector array and a translation symmetry factor of 16, the reduction in geometry factor computation is about 98.9%.

Pseudocode 7.3 Computational structure for forward operator incorporating translation symmetry

Given object scattering density $f(\mathbf{r}, q)$
compute predefined momentum transfer q samples, scatter angle θ samples, and source factor matrix $S(\theta, q)$ for predefined q and θ samples
initialize detector image to zero
for each object pixel x **do**
  for each object pixel y **do**
    $Sf(\theta) = S(\theta, :) \times f(\mathbf{r}, :)$
    if $\mathbf{r}_y$ is first pixel along y direction **then**
      compute geometry factor image $G(1 : M, 1 : N)$
    else
      update geometry factor image $G(:, TSF_y + 1 : N) = G(:, 1 : N - TSF_y)$
      recompute $G(:, 1 : TSF_y)$
    end if
    compute and interpolate scatter angle image
    interpolate $Sf(\theta)$ at scatter angle image
    compute mask modulation factor image
    detector image += geometry factor image $\times$ mask modulation factor image $\times$ interpolated $Sf(\theta)$
  end for
end for

Left-right Mirror Symmetry

Another form of symmetry that we can exploit in the computation of the forward model is a left-right mirror symmetry. Let us assume that the fan beam is perpendicular to the detector plane. If the y-coordinates of the source and the center of the detector array are equal, then we can consider an object reconstruction region whose center's y-coordinate is aligned with the source and detector. For such an object region, we can select an even number of pixels along the y direction. With this choice, the scatter vectors from the left half of the object region are a mirror reflection of those on the right half of the region. The scatter angles and magnitude of the scatter vectors are equal for both halves of the object region.

As a result of the left-right mirror symmetry, we only need to compute the geometry factor and scatter angles for one half of the object region, a savings of 50% in geometry computation.

Pseudocode 7.4 shows the the modification of the forward operator computation due to left-right mirror symmetry. The pseudocode assumes that the mask modulation factor is left-right symmetric.

Pseudocode 7.4 Computational structure for forward operator incorporating left-right mirror symmetry. "fliplr"flips the image from left to right.

Given object scattering density $f(\mathbf{r}, q)$
compute predefined momentum transfer q samples, scatter angle θ samples, and source factor matrix $S(\theta, q)$ for predefined q and θ samples
initialize **two symmetric** detector images to zero
for each object pixel x **do**
  for each object pixel y in the left half **do**
    $Sf_L(\theta) = S(\theta, :) \times f([x, y, 0], :)$
    $Sf_R(\theta) = S(\theta, :) \times f([x, -y, 0], :)$
    if y is first pixel along y direction **then**
      compute geometry factor image $G(1 : M, 1 : N)$
    else
      update geometry factor image $G(:, TSF_y + 1 : N) = G(:, 1 : N - TSF_y)$
      recompute $G(:, 1 : TSF_y)$
    end if
    compute and interpolate scatter angle image
    interpolate $Sf_L(\theta)$ at scatter angle image
    interpolate $Sf_R(\theta)$ at scatter angle image
    compute mask modulation factor image
    detector image left += geometry factor image × mask modulation factor image × interpolated $Sf_L(\theta)$
    detector image right += geometry factor image × mask modulation factor image × interpolated $Sf_R(\theta)$
  end for
end for
detector image = detector image left + fliplr(detector image right)

One of the difficulties with this symmetry is that it is easily broken in reality. The source and the detector array's center may not share the same y-coordinate. In that case, we can transform the measurements, with a potential data loss near the left and right edges, by shifting the measurement window to align the source and detector center.

In addition, the mask modulation may not be symmetric about $y = 0$. This completely breaks the left-right mirror symmetry. A method that alleviates this break in left-right mirror symmetry is discussed later on in the chapter.

Up-down Mirror Symmetry

In addition to left-right mirror symmetry, further improvements in the efficiency of the forward operator computation are possible, with additional constraints on the geometry. If we assume that the fan beam is perpendicular to the detector plane and that the z-coordinate of the source equals that of the detector array center, then the scatter vectors from the object to the top portion of the detector array are reflections of those to the bottom half.

Pseudocode 7.5 depicts the incremental change of the forward operator computational structure attributed to up-down mirror symmetry. The pseudocode assumes that the mask modulation factor is left-right and up-down symmetric. Since only a half of the detector arrays are used in geometry factor and scatter angle computation, a further savings of 50% is achieved. Using both left-right and up-down mirror symmetries result in a 75% decrease in geometry and scatter angle computations, since we only need to compute the scatter vector from one half of the object to a half of the detector array.

In a real system, it may be difficult to align the fan beam and detector array well enough to achieve this setting. However, this vertical centering can be mimicked at a loss of a few measurements near the top and bottom of the detector array; the measured data can be cropped so that the fan beam is centered.

As was the case with left-right mirror symmetry, if the mask modulation factor is not symmetric about $z = 0$, then the up-down symmetry is broken. In essence, the choice of the computational algorithm affects the design of the mask and the geometry as a whole. If we anticipate an efficient algorithm with includes online computation of the mask modulation factor, then the secondary mask should be designed to allow for the mirror symmetries. On the other hand, given a fixed secondary mask, the algorithm should be designed to accommodate the potential absence of mirror symmetries. As we will see later, the left-right and up-down mirror symmetries can still be enforced even when the mask is not mirror symmetric. This algorithmic choice is directly influenced by the system design.

7.4.3 Online-offline Computations

Part of the choice in online computations involves the mask modulation factor. If this factor has desirable symmetry properties that are consistent with the symmetry properties of the geometry factor, then many aspects of its use may be computed online. When the mask

Pseudocode 7.5 Computational structure for forward operator incorporating up-down mirror symmetry. "fliplr"flips the image from left to right, "flipud"flips the image from top to bottom, and "vertcat"stitches two images together vertically.

Given object scattering density $f(\mathbf{r}, q)$
compute predefined momentum transfer q samples, scatter angle θ samples, and source factor matrix $S(\theta, q)$ for predefined q and θ samples
initialize **two symmetric half**-detector images to zero
for each object pixel x **do**
  for each object pixel y in the left half **do**
    $Sf_L(\theta) = S(\theta, :) \times f([x, y, 0], :)$
    $Sf_R(\theta) = S(\theta, :) \times f([x, -y, 0], :)$
    if y is first pixel along y direction **then**
      compute geometry factor image $G(1 : M/2, 1 : N)$
    else
      update geometry factor image $G(:, TSF_y + 1 : N) = G(:, 1 : N - TSF_y)$
      recompute $G(:, 1 : TSF_y)$
    end if
    compute and interpolate scatter angle image
    interpolate $Sf_L(\theta)$ at scatter angle image
    interpolate $Sf_R(\theta)$ at scatter angle image
    compute mask modulation factor image
    detector image up left += geometry factor image $\times$ mask modulation factor image $\times$ interpolated $Sf_L(\theta)$
    detector image up right += geometry factor image $\times$ mask modulation factor image $\times$ interpolated $Sf_R(\theta)$
  end for
end for
detector image = detector image up left + fliplr(detector image up right)
detector image = vertcat(detector image, flipud(detector image))

---

is completely determined through experimental measurements, then a lot of the potential symmetries break down. In addition, the types of masks that we have considered for the fan beam Monte Carlo study do not have these desirable symmetries. Taking these observations into account, we have chosen to precompute the mask factors offline and load them into our code at run-time. Other choices here include interpolation between a subset of precomputed geometry factors, but we do not pursue that avenue in this chapter.

Pseudocode 7.6 shows the final forward operator computational structure, incorporating offline computation of source factor matrix and mask modulation factor, interpolation of smooth scatter angles, translation symmetry, and left-right and up-down mirror symmetries. In addition, a few MATLAB-style [83] comments are given to illustrate what aspects of the geometry are exploited or improved.

7.5 Backward Operator

A raw computational structure of the backward operator is given in Pseudocode 7.7.

An efficient implementation of the backward operator is paramount for an overall efficient iterative algorithm for scatter density estimation. As was the case for the forward operator, the geometry offers several opportunities for efficient computation. The interpolation and symmetry classes identified in the computation of the forward operator can be easily incorporated into an efficient computation of the backward operator. Pseudocode 7.8 gives such an implementation. Note that there is an inner loop, of the dimension of the sampled scatter angles, which is so much smaller than the number of detector pixels, as used by the inner loop in Pseudocode 7.7.

7.6 Reconstruction Algorithm

The forward model given in equation 7.1 predicts the mean detector values. Assuming a discrete representation of the forward operator $\mathbf{H}$ and the object scattering density $\mathbf{f}$, the

Pseudocode 7.6 Computational structure for forward operator incorporating all the code optimizations, including offline computation of mask modulation factors. "fliplr" flips the image from left to right, "flipud" flips the image from top to bottom, and "vertcat" stitches two images together vertically.

Given object scattering density $f(\mathbf{r}, q)$
compute predefined momentum transfer q samples, scatter angle θ samples, and source factor matrix $S(\theta, q)$ for predefined q and θ samples
compute **four symmetric half**-mask modulation factor images: mask up left, mask down left, mask up right, mask down right % design issue
initialize **four symmetric half**-detector images to zero
for each object pixel x **do**
  for each object pixel y in the left half **do**
    $Sf_L(\theta) = S(\theta, :) \times f([x, y, 0], :)$ % reduce number of multiplies
    $Sf_R(\theta) = S(\theta, :) \times f([x, -y, 0], :)$ % left-right symmetry class
    if y is first pixel along y direction **then**
      compute geometry factor image $G(1 : M/2, 1 : N)$ % exploit up-down symmetry
    else
      update geometry factor image $G(:, TSF_y + 1 : N) = G(:, 1 : N - TSF_y)$ % exploit translation symmetry
      recompute $G(:, 1 : TSF_y)$
    end if
    compute and **interpolate** scatter angle image % exploit smoothness of angles
    interpolate $Sf_L(\theta)$ at scatter angle image
    interpolate $Sf_R(\theta)$ at scatter angle image
    lookup mask modulation factors
    detector image up left += geometry factor image × mask up left × interpolated $Sf_L(\theta)$
    detector image down left += geometry factor image × mask down left × interpolated $Sf_L(\theta)$
    detector image up right += geometry factor image × mask up right × interpolated $Sf_R(\theta)$
    detector image down right += geometry factor image × mask down right × interpolated $Sf_R(\theta)$
  end for
end for
detector image up = detector image up left + fliplr(detector image up right)
detector image down = detector image down left + fliplr(detector image down right)
detector image = vertcat(detector image up, flipud(detector image down))

Pseudocode 7.7 Raw computational structure for backward operator

```
Given detector image
for each object point r do
  for each momentum transfer q do
    for each detector point r' do
      compute geometry factor G(n, s)
      compute scatter angle θ(r, s)
      compute source factor S(θ, q)
      compute mask modulation factor T(r, ŝ)
      accumulate geometry factor times mask modulation factor times detector image
      into result
    end for
    object image f(r, q) = result
  end for
end for
```

mean detector values are given by [79]

$$\begin{aligned} \mathbf{g} &= \mathbf{H}\mathbf{f}, \\ g(m) &= \sum_{i \in I} H(m,i) f(i), \end{aligned} \tag{7.4}$$

where $\mathbf{g}$ is an $|M| \times 1$ vector, $\mathbf{H}$ is an $|M| \times |I|$ matrix, $\mathbf{f}$ is an $|I| \times 1$ vector, and $m \in M$.

Note that this representation implies the forward operator matrix $\mathbf{H}$ is precomputed. However, for the results we present, the forward operator and its adjoint, the backward operator, are computed online using an implementation of Pseudocodes 7.6 and 7.8, respectively.

A Poisson model for the measured data is appropriate in many applications. Let us denote the random data by

$$\begin{aligned} \mathbf{y} &\sim \text{Poisson}\left(\mathbf{H}\mathbf{f} + \boldsymbol{\mu}_b\right), \\ y(m) &\sim \text{Poisson}\left(\sum_{i \in I} H(m,i) f(i) + \mu_b(m)\right), \end{aligned}$$

where $\boldsymbol{\mu}_b$ is the mean number of background counts, assumed to be known.

The unpenalized log-likelihood function for the data is given as

Pseudocode 7.8 Efficient computational structure for backward operator

Given detector image
extract **four symmetric half**-detector images: detector up left, detector up right, detector down left, detector down right
compute predefined momentum transfer q samples, scatter angle θ samples, and source factor matrix $S(\theta, q)$ for predefined q and θ samples
compute **four symmetric half**-mask modulation factor images: mask up left, mask down left, mask up right, mask down right
for each object pixel x **do**
  for each object pixel y in the left half **do**
    if y is first pixel along y direction **then**
      compute geometry factor image $G(1 : M/2, 1 : N)$
    else
      update geometry factor image $G(:, TSF_y + 1 : N) = G(:, 1 : N - TSF_y)$
      recompute $G(:, 1 : TSF_y)$
    end if
    lookup mask modulation factors
    temp image left = geometry factor image $\times$ (mask up left $\times$ detector up left + mask down left $\times$ detector down left)
    temp image right = geometry factor image $\times$ (mask up right $\times$ detector up right + mask down right $\times$ detector down right)
    compute and **interpolate** scatter angle image
    for each sampled scatter angle θ **do**
      temp object left(θ) = **accumulate** temp image left locations at angle θ
      temp object right(θ) = **accumulate** temp image right locations at angle θ
    end for
    object image $f([x, y, 0], q)$ = temp object left(:) $\times$ $S(:, q)$
    object image $f([x, -y, 0], q)$ = temp object right(:) $\times$ $S(:, q)$
  end for
end for

$$l(\mathbf{y}|\mathbf{f}) = \left[\sum_{m\in M} y(m)\ln\left(\sum_{i\in I} H(m,i)f(i)+\mu_b(m)\right)\right. \\ \left. - \left(\sum_{i\in I} H(m,i)f(i)+\mu_b(m)\right)\right]. \tag{7.5}$$

In this chapter, since we consider an unregularized version of the reconstruction algorithm, the goal is to find the object scattering density $\mathbf{f}$ which maximizes $l(\mathbf{y}|\mathbf{f})$. An iterative algorithm for this optimization problem is described by Snyder *et al.* [114], MacCabe *et al.* [79] and is shown in Pseudocode 7.9.

Pseudocode 7.9 Expectation-maximization (EM) reconstruction algorithm

initialize $\hat{f}^{k=0}(i)$ to positive values and $w(i) = \sum_{m\in M} H(m,i)$

for $k = 0$ **to** $K-1$ **do**

$$\hat{f}^{k+1}(i) = \frac{\hat{f}^k(i)}{w(i)} \sum_{m\in M} H(m,i) \times \left(\frac{y(m)}{\sum_{i'\in I} H(m,i')\hat{f}^k(i') + \mu_b(m)}\right)$$

end for

To increase the convergence rate of the EM algorithm, we included a range decomposition method a.k.a. ordered subsets. Pseudocode 7.10 shows the structure of the ordered subset EM algorithm. The measurements are partitioned into J subsets.

The outer sum corresponds to a backward projection, while the denominator is a forward projection. The weights $\mathbf{w}$ are computed as a backward projection of a detector image of all ones. Let us denote the ratio of the measured data $\mathbf{y}$ to the estimated data (mean detector values + background values) as an error ratio. The ordered subset expectation-maximization algorithm can be summarized in terms of the forward and backward operators as shown in Pseudocode 7.11, where ./ and .$*$ represent element-wise division and multiplication respectively.

Pseudocode 7.10 Ordered subset expectation-maximization (OSEM) reconstruction algorithm

define J subsets $\{M_j\}_{j=1}^{J}$ such that $M = \cup_{j=1}^{J} M_j$
initialize $\hat{f}^{k=0,j=0}(i)$ to positive values and $w_j(i) = \sum_{m \in M_j} H(m,i)$
for $k = 0$ **to** $K-1$ **do**
  for $j = 1$ **to** J **do**

$$\hat{f}^{k,j}(i) = \frac{\hat{f}^{k,j-1}(i)}{w_j(i)} \sum_{m \in M_j} H(m,i) \times \left(\frac{y(m)}{\sum_{i' \in I} H(m,i') \hat{f}^{k,j-1}(i') + \mu_b(m)} \right)$$

  end for
  $\hat{f}^{k+1,0}(i) = \hat{f}^{k,J}(i)$
end for

Pseudocode 7.11 Ordered subset expectation-maximization reconstruction algorithm based on forward and backward operators

define J subsets $\{M_j\}_{j=1}^{J}$ such that $M = \cup_{j=1}^{J} M_j$
initialize $\hat{\mathbf{f}}^{k=0,j=0}$ to positive values and $\mathbf{w}_j$ = backward(detector image of all ones)
for $k = 0$ **to** $K-1$ **do**
  for $j = 1$ **to** J **do**
    error ratio = $\mathbf{y}$ ./ (forward($\hat{\mathbf{f}}^{k,j-1}$) + $\boldsymbol{\mu}_b$)
    $\hat{\mathbf{f}}^{k,j} = \left(\hat{\mathbf{f}}^{k,j-1} ./ \mathbf{w}_j \right) .*$ backward(error ratio)
  end for
  $\hat{\mathbf{f}}^{k+1,0} = \hat{\mathbf{f}}^{k,J}$
end for

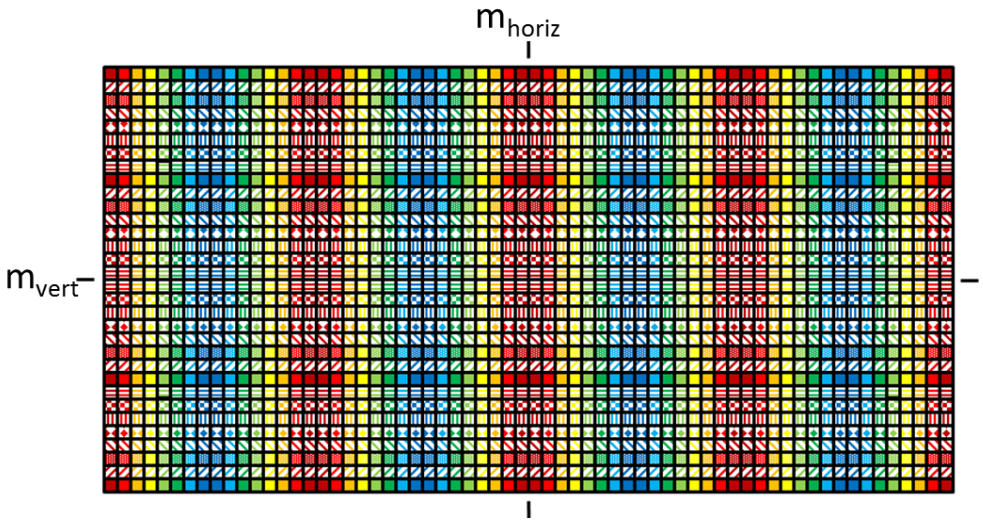


Figure 7.3: Layout of the choice of 64 subsets using a mini detector array. $\mathrm{m}_{\mathrm{horiz}}$ and $\mathrm{m}_{\mathrm{vert}}$ are the midpoints of the detector array. This illustration was created by Dr. David Politte.

There are several choices for $M_1, \cdots, M_J$. In keeping with the left-right and up-down mirror symmetries discussed above, if one detector pixel is in a subset, then its three mirror reflections must also be included in that subset. In addition, if we want to utilize the translation symmetry in saving computations, then the pixels that are a given horizontal (y) translation away from each of the four pixels above must also be included in the subset.

We have found that the choice in Figure 7.3 satisfies this requirement, for a scaled-down version of the detector array with 32 rows and 64 columns, and a translation symmetry factor of 16. Pixels with each distinct combination of color and pattern represent a subset. For example, the horizontal pixels in the first subset are obtained from horizontal detector pixels $1 : 16 : 64$ and $64 : -16 : 1$, where we have used MATLAB's [83] listing notation $a : b : c$ to mean numbers from a to c in steps of b, and only a is guaranteed to be in the list. The choice of 16 as the translation symmetry factor gives a total of 8 ($= 16/2$) horizontal subsets.

For the vertical pixels, due to up-down mirror symmetry, we can consider partitioning only one half of the detector array. For a detector array with 32 rows, this means we can choose any factor of 16 as the step size in the vertical direction, in selecting pixels that belong to the same subset. In Figure 7.3, 8 was chosen as the step size, resulting in 8 vertical subsets. This gives a total of 64 subsets for the ordered subset EM algorithm.

7.7 Resolution Recovery and Inter-voxel Correlation

In this section, we empirically characterize the fan beam flat-panel energy-integrating detection system. In particular, we want to understand whether and how the momentum transfer (spectral) resolution degrades with increasing spectral values. In addition, we want to know if the system is well-conditioned by studying the correlation extent of spatio-spectral voxels. That is, how much of the signal in a voxel is misrepresented in another voxel by the system.

For the resolution recovery analysis, we simulated a point object at the spatial center of the object region. At this location, we utilized a discrete impulse function centered at a momentum transfer bin. This constitutes a momentum transfer profile with a single peak. Several momentum transfer profiles were simulated by centering the discrete impulse function on different momentum transfer bins. For each of the simulated objects, the forward model was used to create scatter data. The data was then reconstructed using an ordered subset EM algorithm.

The full width half maximum (FWHM) of the simulated and reconstructed momentum transfer profiles was used as the measure of spectral resolution. Since the momentum transfer bins have a finite size, the FWHM of the simulated momentum transfer profile is given by the bin width. Let the resolution recovery ratio be defined as

$$\text{Resolution Recovery Ratio} = \frac{\text{FWHM}_{\text{simulated}}}{\text{FWHM}_{\text{reconstructed}}}. \tag{7.6}$$

That is, the ratio of the FWHM of the simulated momentum transfer profile to the FWHM of the reconstructed momentum transfer profile.

For the simulation, we used a smaller detector array with 384 rows and 512 columns, with a pixel pitch of 0.76 mm in both z and y directions, to save computational time. Also, a smaller object volume was reconstructed; approximately 15 mm by 15 mm in the xy-plane, with a pixel pitch of 2.5 mm in both directions. 277 momentum transfer bins evenly spaced between 2 and 140 rad nm^{-1} were used in the simulations and reconstructions. The translation symmetry factor as described in previous sections was not utilized here for generating the subsets. Rather, 64 subsets of the detector data were constructed by using a row and column spacing of 8 pixels. The other details about the geometry are the same as those in Section 7.8 below.

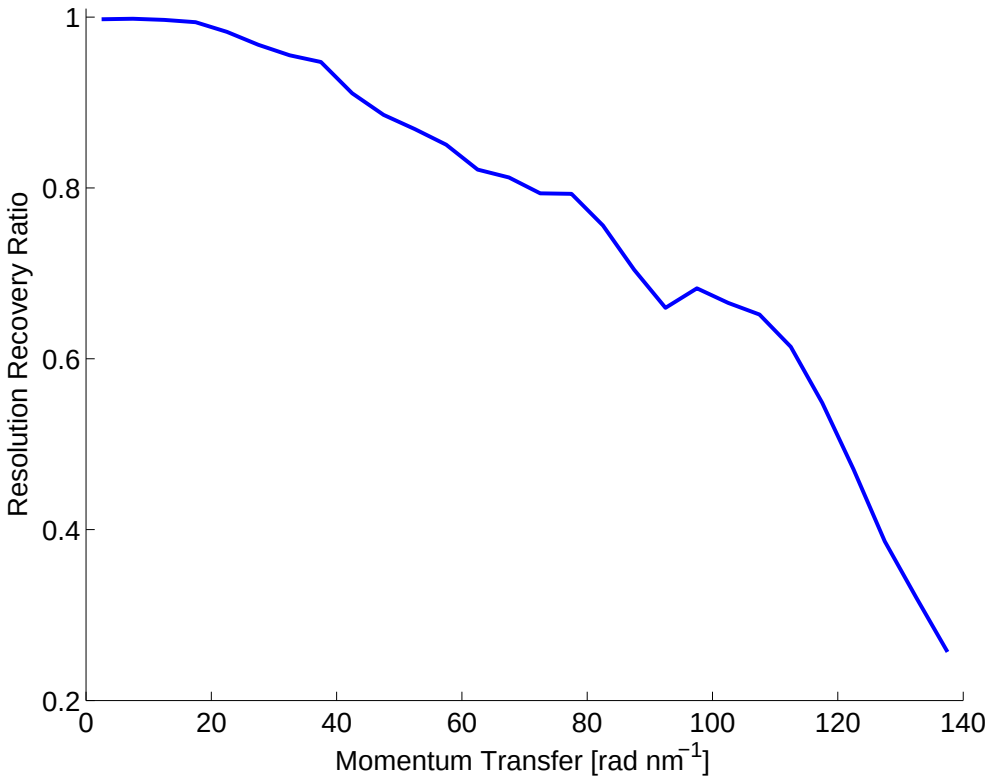


Figure 7.4: Resolution recovery plot. A plot of the ratio of the resolution of the simulated momentum transfer profile to that of the reconstructed momentum transfer profile against the value of the momentum transfer bin.

Figure 7.4 shows a plot of the resolution recovery for each momentum transfer bin value, after 100 ordered-subset EM iterations. From the figure, we see that the ratio starts close to 1, suggesting nearly perfect recovery, but degrades as we step along the momentum transfer direction.

The result presented in Figure 7.4 suggests that the scatter system is better able to resolve lower momentum transfer values. The results also suggest using non-uniform momentum transfer bins, with increasing bin widths as momentum transfer increases, since the system is better able to reconstruct lower values of momentum transfer. In this work, we consider uniformly spaced momentum transfer bins.

For the correlation analysis, we simulated a uniform random object. The forward model was used to create the mean scatter data. 1000 realizations of Poisson data based on the mean scatter data, were created. The data was then reconstructed using an ordered subset EM algorithm. The same 64 subsets used in the resolution recovery analysis were utilized here. The sample correlation coefficient matrix for the spatio-spectral voxels is given as

$$CM_{ij} = \frac{CV_{ij}}{\sqrt{CV_{ii}CV_{jj}}}, \tag{7.7}$$

where $i, j = 1, \cdots, V$, V is the number of voxels, and CV is the sample covariance matrix.

The sample covariance matrix is defined as

$$CV_{ij} = \frac{1}{N-1}\sum_{n=1}^{N}(f_n(i) - \mu_i)(f_n(j) - \mu_j), \tag{7.8}$$

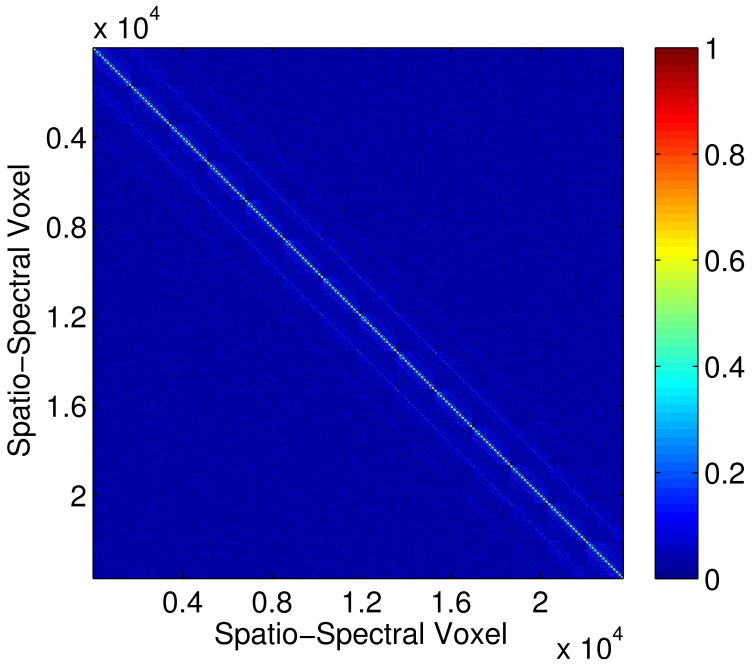
where N is the number of realizations, $f_n(i)$ is the reconstructed scattering density of the i^{th} voxel of the n^{th} realization, and $\mu_i = \frac{1}{N}\sum_{n=1}^{N} f_n(i)$ is the sample mean of the i^{th} voxel.

Since so many realizations are considered, to save computational time, as was the case for the resolution recovery analysis, we used a smaller detector array with 384 rows and 512 columns, with a pixel pitch of 0.76 mm in both the z and y directions. A small object volume was reconstructed; approximately 35 mm by 35 mm in the xy-plane, with a pixel pitch of 2.5 mm in both directions. 121 momentum transfer bins evenly spaced between 10 and 70 rad nm^{-1} were used in the simulations and reconstructions. The other details about the geometry are the same as those in Section 7.8 below.

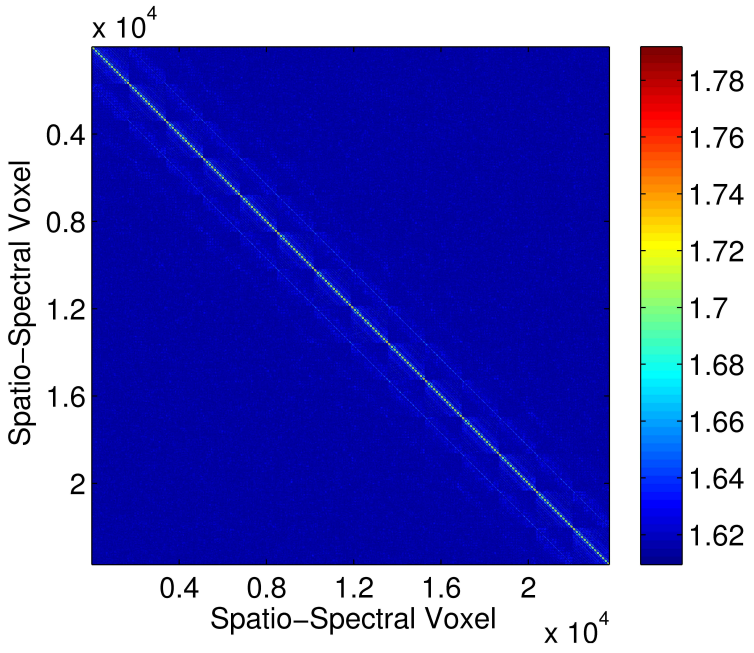
Figure 7.5a shows the absolute value of the sample spatio-spectral correlation coefficient matrix based on the reconstructed object voxels from the Poisson scatter data realizations, obtained using 100 ordered-subset EM iterations. Figure 7.5b shows the same information in Figure 7.5a, but represented on a non-linearly zoomed in scale. The spatio-spectral voxels are arranged such that voxels belonging to the same spatial location are closest. The next closest voxels are those with the same y values (different x values; along the direction of the source). The structure revealed by Figure 7.5 suggests that the extent of influence (cross-talk) across spatio-spectral voxels is minimal.

7.8 Results

The forward and backward models, integral parts of the EM algorithm for estimating object scattering density, were validated on simulated and Monte Carlo data. For both cases, the source was located at the origin, and the center of the flat-panel detector array was 1546.5



(a)



(b)

Figure 7.5: Absolute spatio-spectral sample correlation coefficient matrix. A plot of the absolute value of the sample correlation coefficient matrix for the reconstructed object. (a) Absolute correlation coefficient matrix. (b) Logarithm of the sum of the absolute correlation coefficient matrix and 5.

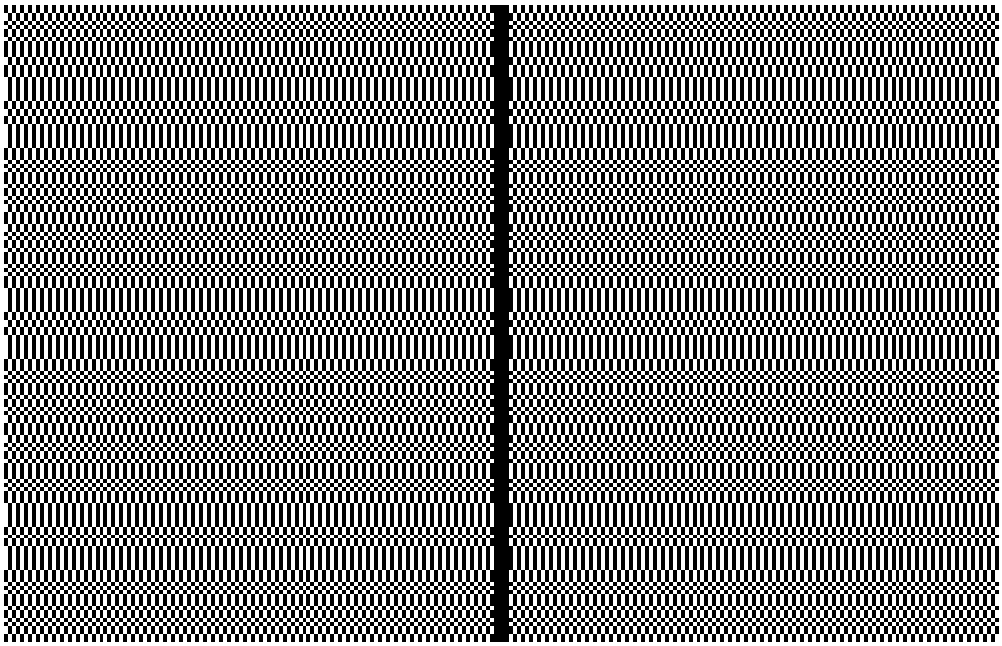


Figure 7.6: Secondary mask with fan beam stop.

mm away along the positive x-axis. The source was operated at 125 kVp, and the spectrum was filtered before being shaped into a fan by the primary aperture (slit). The secondary mask was placed 100 mm in front of and parallel to the plane of the detector array. The center of the object was located 1035 mm from the source, along the positive x-axis. The detector array had 1536 rows and 2048 columns, with a pixel pitch of 0.19 mm in both the z and y directions.

For the reconstructions, a translation symmetry factor of 16 was used, with 8 vertical subsets, giving a total of 64 subsets of the measurements. A region of approximately 70 mm by 85 mm, in the xy-plane, was reconstructed, with a pixel pitch of 2.5 mm and 3.04=(16×0.19) mm along the x and y directions, respectively. 261 evenly spaced momentum transfer bins from 10 to 140 rad nm^{-1} were used. These predefined momentum transfer samples were later converted into inverse inter-atomic distances $1/(2\mathrm{D})$, measured in nm^{-1} by dividing them by 4π. There were 250 scatter angle samples chosen uniformly from 0 to $\frac{\pi}{6}$ radians, excluding 0.

Figure 7.6 shows the secondary mask used in both the simulated and Monte Carlo data. This mask has been described in detail by MacCabe *et al.* [78]. Since the mask does not satisfy the mirror symmetry requirements, we precomputed the four mask modulation factors as hinted to in Pseudocode 7.6.

7.8.1 Simulated Data

We considered two examples for the simulated data. The first example models a scenario where there are multiple strong point scatterers in close proximity. For this example, illustrated in Figure 7.7b, point scatterers of sodium chloride (NaCl) and aluminum (Al) powder were placed in a plus sign configuration. The middle (12.5 mm span) column of the object spatial distribution is occupied by NaCl, while Al is in the middle row (except for the center NaCl pixel). Al occupies 6.08 mm on each side of the NaCl column.

For the second example, we modeled the object configuration used in generating the Monte Carlo data — a 5 mm by 50 mm by 50 mm rectangular block of graphite powder. Since our model considers a single fan beam in the $z = 0$ plane, we simulated data for a 5 mm by 48.64 mm slice of the graphite slab. The simulated data for this example is given in Figure 7.8a.

For both examples, the forward operator was applied to the object scattering density in creating the noiseless scatter data. Poisson noise was later introduced with a maximum photon count of 50. The simulated noisy data for both examples are given in Figures 7.7a and 7.8a.

Figures 7.7 and 7.8 show the results of estimating the scattering profiles of the plus sign configuration and the graphite slice configuration, using 10 iterations of the ordered subset EM algorithm, respectively. From the figures we can see that the spatial distributions ($\tilde{f}(x, y) = \int dq\, f(x, y, q)$) and the momentum transfer profiles are recovered accurately in a few iterations.

The spatial distributions and momentum transfer profiles closely match those of the reference configurations and materials. These results are in agreement with those presented in the paper by MacCabe *et al.* [79] using real data from a pencil-beam system. However, from the momentum transfer profile of Al shown in Figure 7.7f, we can see that the estimation struggles with momentum transfer resolution at higher momentum transfer values. The two Al peaks between 5 and 6 nm^{-1} are merged into a single broad peak.

Ordered subsets EM promises an acceleration of the convergence rate of a regular EM algorithm that is comparable to the number of subset [31]. Figure 7.9 shows the value of the log-likelihood objective as a function of the iteration number. We can see that an acceleration factor of about 54 is obtained by using the 64 subsets described above.

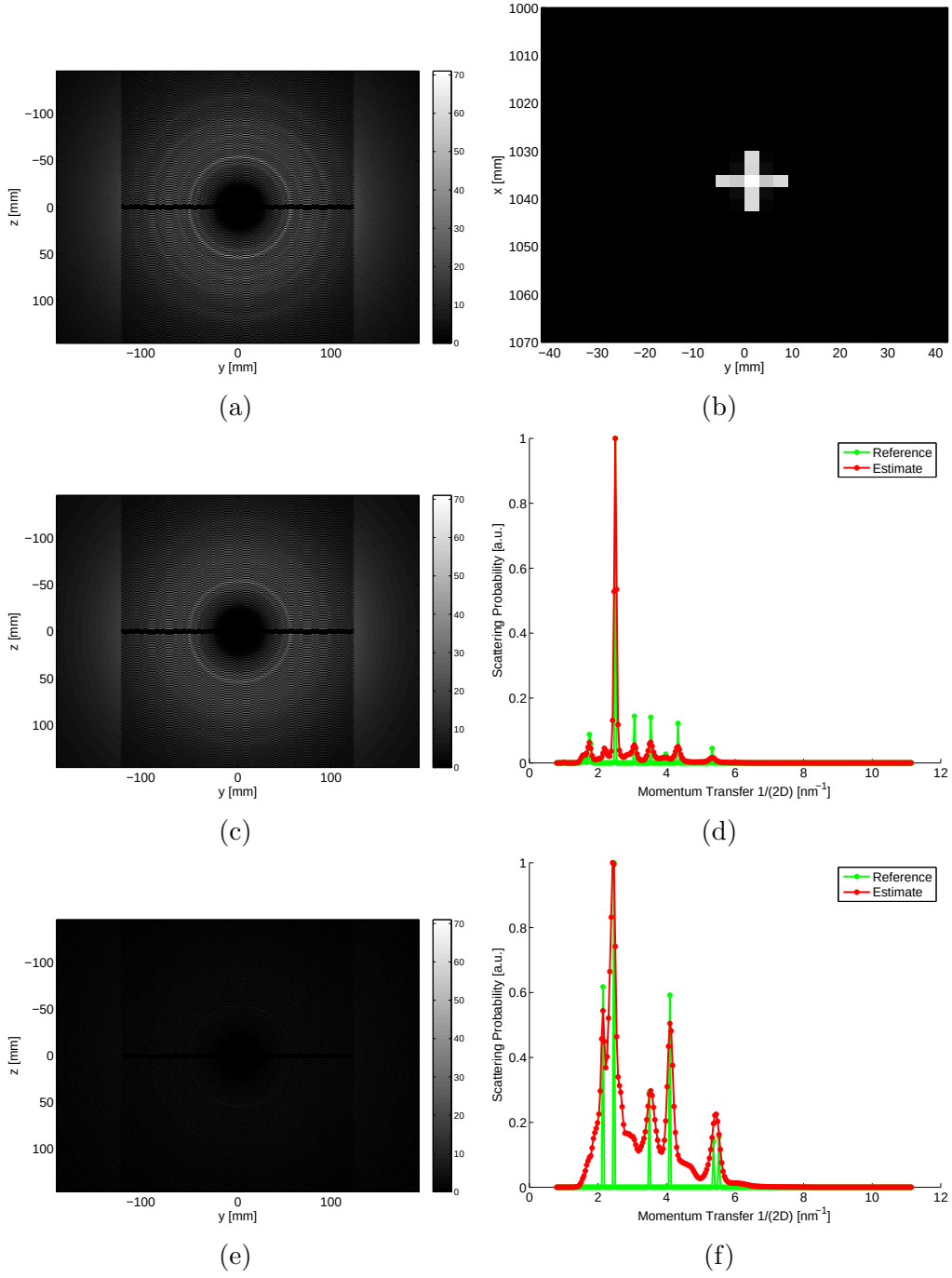


Figure 7.7: Estimation of scattering density from a plus sign configuration involving NaCl and Al. (a) Simulated noisy data. (b) Estimated spatial distribution. (c) Estimated mean data. (d) Momentum transfer profile for NaCl. (e) Absolute difference between simulated and estimated data. (f) Momentum transfer profile for Al.

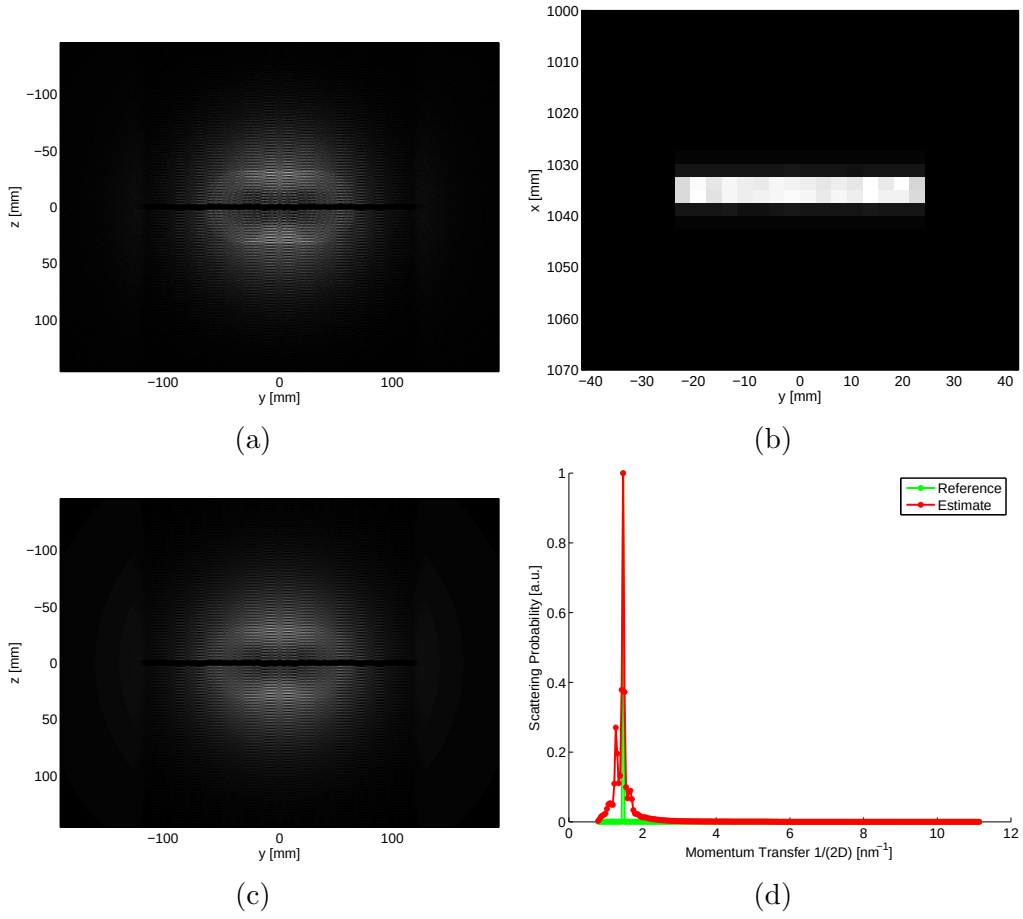


Figure 7.8: Estimation of scattering density from a graphite slice configuration. (a) Simulated noisy data. (b) Estimated spatial distribution. (c) Estimated mean data. (d) Momentum transfer profile for graphite.

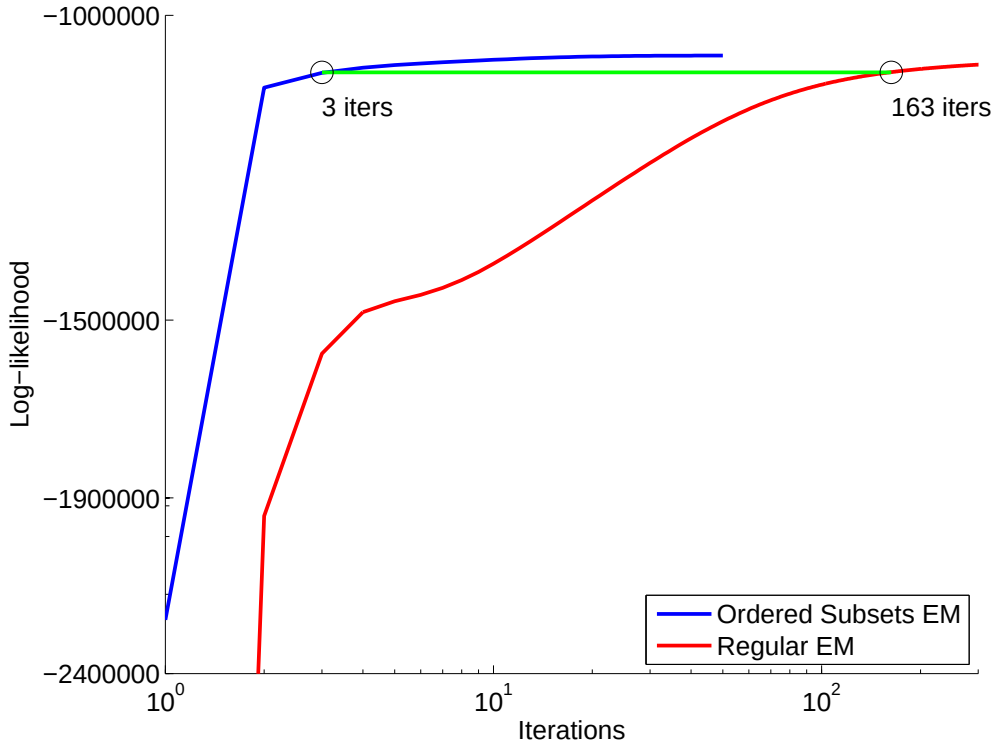


Figure 7.9: Log-likelihood as a function of iteration number for the regular and ordered subsets EM algorithms.

7.8.2 Monte Carlo Data

The Monte Carlo data used in this chapter was generated by Manu Lakshmanan at Duke University. The object used in the Monte Carlo simulation was a 5 mm by 50 mm by 50 mm right rectangular prism of graphite powder whose form factor is shown as the reference in Figures 7.8d and 7.10d. The Monte Carlo data includes single scattering events from the graphite crystalline powder, multiple scattering events, and scatter from the secondary aperture.

Since the forward model does not account for X-ray penetration of the beam-stop on the secondary mask, the magnified region occupied by the beam-stop on the detector array has all zeros, while the Monte Carlo data does not. To account for this model mismatch, a detector (0-1) mask was applied to the Monte Carlo data, effectively ignoring the photons that were measured in the magnified beam-stop region.

Figure 7.10 shows the results of estimating the mean detector photon counts and the object scattering density, using 10 iterations of the ordered subset EM algorithm described above.

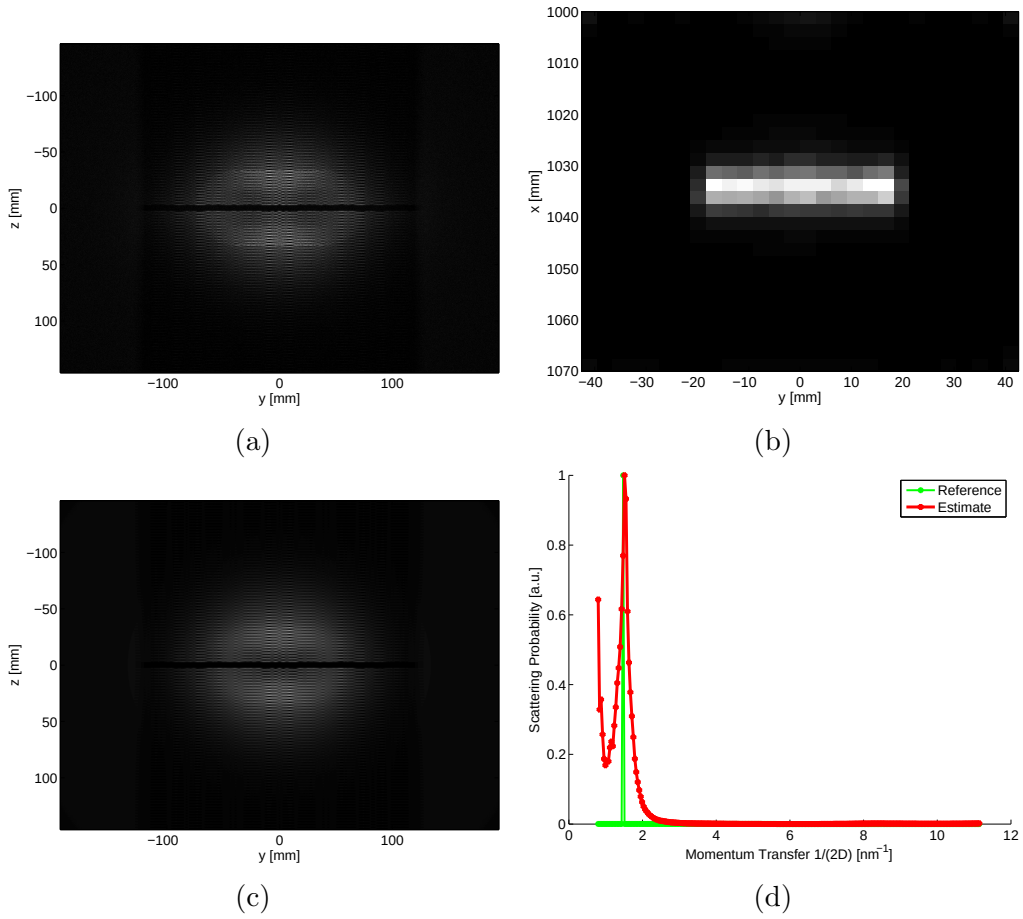


Figure 7.10: Estimation of scattering density from Monte Carlo data of a graphite rectangular prism. (a) Monte Carlo data. (b) Estimated spatial distribution. (c) Estimated mean data. (d) Momentum transfer profile.

Again, we see that the spatial distribution is recovered quickly and accurately. The correct momentum transfer peak was recovered with spurious contributions to the lower edge of the momentum transfer range. This edge effect diminishes with further iterations.

These results demonstrate that without prior knowledge of the location of the scatterers, the fan beam system can be used to efficiently estimate their scattering densities $f(x, y, q)$.

7.9 Summary

This chapter described the use of tomographic scatter measurements in estimating volumetric scatter density, using a fan beam source distribution and a secondary mask. However the

methods we described for modeling the data can be easily carried over to other source distributions.

The design of the scatter imaging system, including the choice of the detector array, its pixel pitch, and placement relative to the source, and the secondary mask and its placement, influences the choices made in implementing the forward and backward operators and the reconstruction algorithm. For example, the mask modulation factor of a secondary mask that fails to satisfy the left-right and up-down mirror symmetries is better computed offline, since its online computation may become the bottleneck in the computation of the forward operator.

On the other hand, the need for efficient algorithms also affects several elements of the system design. For example, placing the center of the detector array so that the central ray from the source strikes it and using a mask that satisfies the mirror symmetry requirements, permit efficient online implementations of the forward model and reconstruction algorithms. Also, having a detector with a smaller pixel pitch permits the potential use of a larger translation symmetry factor, without significant deterioration of the recoverable object spatial resolution in the y direction.

The algorithm we described depends heavily on a computational representation of the forward model for the data. We identified and described three ways of significantly reducing the computational burden of the forward model and the overall reconstruction algorithm, namely: efficient implementation of the source factor times the scattering density, accompanied by interpolation; symmetries in the geometry factors; and the need to balance online and offline computations.

The recovered resolution from simulated point objects show that spectral resolution degrades with increasing momentum transfer values. Correlation of spatio-spectral voxel values over many realizations of reconstructed Poisson data reveal that each spatio-spectral voxel has a limited effect on voxels that are far away.

The forward model and reconstruction algorithm were validated using simulated and Monte Carlo data. The scatter spatial distribution and the momentum transfer profiles were recovered accurately using a few iterations of the ordered subset EM reconstruction algorithm.

7.10 Lessons Learned

To permit efficient computation of a forward model, some algorithmic needs must be satisfied by the scanner geometry.

Interpolation is an invaluable tool for speeding up computations, when applicable. Interpolation is useful when the a function is smooth. In addition, if this function needs to be computed over and over, using different input parameters, that is a good indication for interpolation.

Exploiting symmetry in the scanner geometry leads to a speed up in online geometry factor computations. Even when this symmetry is initially broken, certain system design decisions can be made to approximately mend it.

With the source, object, and detector fixed, the structure of the spatio-spectral correlation matrix can be a measure of the utility of a secondary mask.

The scatter system estimates low spectral values with better resolution than it does higher spectral values.

Chapter 8

Special Case and Extensions of Fan Beam Model for X-ray Coherent Scatter Imaging

8.1 Introduction

The forward model described in Chapter 7 for a fan beam source with energy-integrating detectors can be adapted to other source distributions or detection systems. In this chapter, we consider two such adaptations of the fan beam source distribution, energy-integrating detection model: a pencil beam source distribution model and an energy-sensitive detection model.

A momentum transfer basis was used to represent the object's scattering density in the previous chapter. Due to the large number of degrees of freedom, this basis may introduce unrealistic noise into the reconstruction. Other bases that could be used within the framework we developed in the previous chapter are introduced in this chapter. We consider one such basis, namely, a material basis, for the energy-sensitive detector setting.

In addition to choosing a different representational basis, noisy reconstructions that result from using the more natural momentum transfer basis can be improved by regularization. This chapter considers the use of an edge-preserving neighborhood-based regularization function [71] to reduce noise in the reconstructions. Note that regularization can also be employed in other bases. We demonstrate the utility of regularization in the energy-sensitive detection setting.

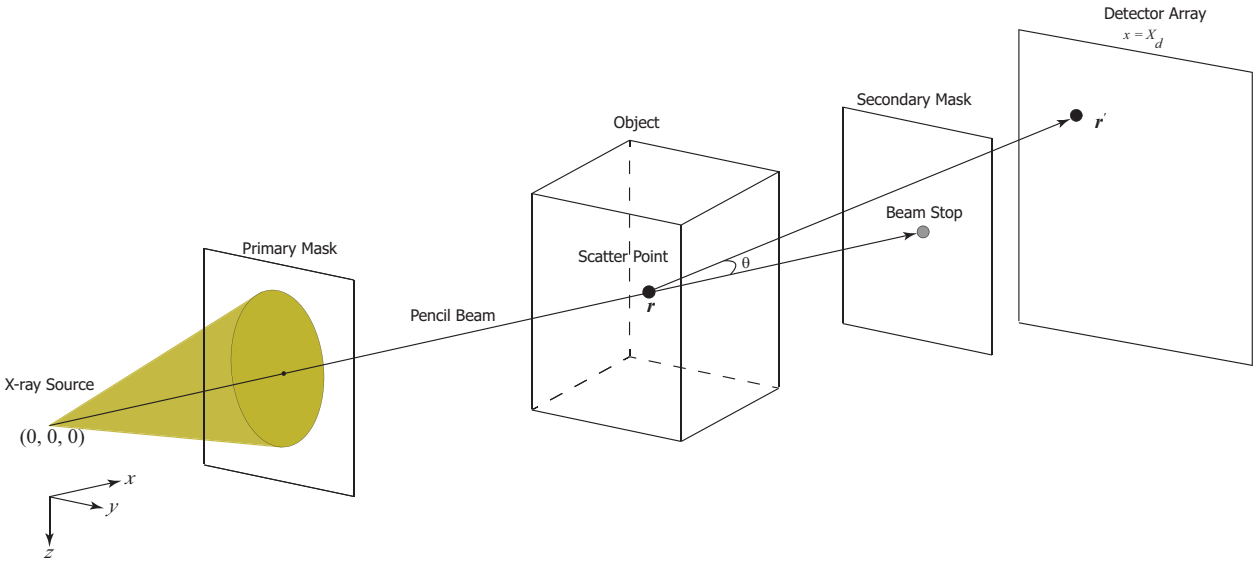


Figure 8.1: Schematic of the pencil beam geometry.

We observed that the slowness of the EM algorithm in recovering the object scattering density, when higher momentum transfer values are utilized, may be improved by effective regularization either using an expressive material basis or an edge-preserving neighborhood-based regularizer.

8.2 Pencil Beam X-ray Coherent Scatter System

A cone beam emanating from a point source can be shaped into a pencil beam of a given width by using a primary mask (collimator) with a single small hole. The pencil beam source illuminates an object in its path, which scatters the lights. The scattered photons pass through a secondary mask on their way to the flat-panel, energy-integrating detector. The secondary mask helps in delineating the scatter angles. Figure 8.1 provides a schematic of the pencil beam geometry.

8.2.1 Forward and Backward Models

In the pencil beam model, the X-ray source transmits X-rays along a line, the x axis. The source is at the origin of the coordinate system. Object scatter locations are indexed by x and corresponding object point $\mathbf{r} = [x, 0, 0]$, with the positive x coordinate pointing from

the source through the object to the secondary mask and the detector. The detector plane is perpendicular to the pencil beam. A point on the detector is $\mathbf{r}' = [X_d, y', z']$. The object point and the detector point determine the scatter vector $\mathbf{s} = \mathbf{r}' - \mathbf{r}$. The scatter vector and the object point determine the scatter angle θ.

The forward model for the pencil beam geometry is identical to that of the fan beam geometry 7.1, and is restated here as

$$g(\mathbf{r}') = \int d\mathbf{r} \int dq\, H(\mathbf{r}', \mathbf{r}, q) f(\mathbf{r}, q), \tag{8.1}$$

where

$$H(\mathbf{r}', \mathbf{r}, q) = C \left| \frac{\mathbf{n} \cdot \hat{\mathbf{s}}}{s^2} \right| T(\mathbf{r}, \hat{\mathbf{s}}) \left(\frac{1}{2q \sin \frac{\theta}{2}} \right) W \left(\frac{q}{2 \sin \frac{\theta}{2}} \right),$$

$f(\mathbf{r}, q)$ is the scattering density at the object location $\mathbf{r}$ and momentum transfer q, C is a normalization constant, $\mathbf{n}$ is a normal to the detector array plane, k is the energy of the incident X-rays, $T(\cdot)$ is the mask modulation factor, and $W(\cdot)$ is the energy-dependent filtered source flux.

The original fan beam model described in Chapter 7 can be adapted to the pencil beam model by setting $y = 0$. If we assume a discrete representation of the object volume, there is a single pixel in the y direction. Since there is only one pixel in that direction, the utility of the translation symmetry in accelerating the computation of the forward operator is diminished.

This leaves us with the source computation coupled with scatter angle interpolation, left-right and up-down mirror symmetry, and balancing online/offline computations, for speeding up the computation of the forward operator. Pseudocode 8.1 shows the modified version of the fan beam forward operator given in Pseudocode 7.6.

Similarly, the backward operator for the pencil beam geometry is given in Pseudocode 8.2. This pseudocode is a direct modification of Pseudocode 7.8.

The structure of the reconstruction algorithm is identical to that given in Pseudocode 7.11.

Pseudocode 8.1 Computational structure for forward operator with pencil beam source and energy-integrating flat-panel detectors. The detector array has M by N pixels. "fliplr" flips the image from left to right, "flipud" flips the image from top to bottom, and "vertcat" stitches two images together vertically.

compute predefined momentum transfer q samples, scatter angle θ samples, and source factor matrix $S(\theta, q)$ for predefined q and θ samples
compute **four symmetric quarter**-mask modulation factor images: mask up left, mask up right, mask down left, mask down right
initialize **four symmetric quarter**-detector images to zero
Given object scattering density $f(\mathbf{r}, q)$
for each object pixel x **do**
  $Sf(\theta) = S(\theta, :) \times f([x, 0, 0], :)$
  compute geometry factor image $G(1 : M/2, 1 : N/2)$
  compute and **interpolate** scatter angle image
  interpolate $Sf(\theta)$ at scatter angle image
  lookup mask modulation factors
  detector image up left += geometry factor image × mask up left × interpolated $Sf(\theta)$
  detector image up right += geometry factor image × mask up right × interpolated $Sf(\theta)$
  detector image down left += geometry factor image × mask down left × interpolated $Sf(\theta)$
  detector image down right += geometry factor image × mask down right × interpolated $Sf(\theta)$
end for
detector image up = horcat(detector image up left, fliplr(detector image up right))
detector image down = horcat(detector image down left, fliplr(detector image down right))
detector image = vertcat(detector image up, flipud(detector image down))

Pseudocode 8.2 Computational structure for backward operator with pencil beam source and energy-integrating flat-panel detectors. The detector array has M by N pixels.

```
compute predefined momentum transfer q samples, scatter angle θ samples, and source
factor matrix S(θ, q) for predefined q and θ samples
compute four symmetric quarter-mask modulation factor images: mask up left, mask
up right, mask down left, mask down right
Given detector image
extract four symmetric quarter-detector images: detector up left, detector up right,
detector down left, detector down right
for each object pixel x do
  compute geometry factor image G(1 : M/2, 1 : N/2)
  lookup mask modulation factors
  temp image = geometry factor image × (mask up left × detector up left + mask up
  right × detector up right + mask down left × detector down left + mask down right ×
  detector down right)
  compute and interpolate scatter angle image
  for each sampled scatter angle θ do
    temp object(θ) = accumulate temp image locations at angle θ
  end for
  object image f([x, 0, 0], q) = temp object(:) × S(:, q)
end for
```

8.2.2 Results

The pencil beam model was validated using Monte Carlo data from a suitcase. This data was generated by Pooyan Sahbaee at Duke University. The components of the suitcase are given in Figure 8.2. The suitcase is made from polypropylene and filled with foam, between its interior and exterior surfaces. The suitcase phantom also includes 3 layers of different materials. The first layer (closest to the source) and the third layer (closest to the detector) are made of acrylic. The middle layer is made of the target material, sodium chloride (NaCl) crystalline powder. Each of the layers is 3.175 mm deep along the x direction.

The source is located at the origin, and the center of the flat-panel detector array is 1079.5 mm away along the x direction. The source was operated at 150 kVp with 0.1 mm Tungsten filtration. The X-ray exposure was 5 mAs. The secondary mask was placed 100 mm in front of and parallel to the plane of the detector array. The plane of the middle layer is 584.2 mm from the source. The detector array had 1536/4 rows and 2048/4 columns, with a pixel pitch of 0.19*4 mm in both the z and y directions.

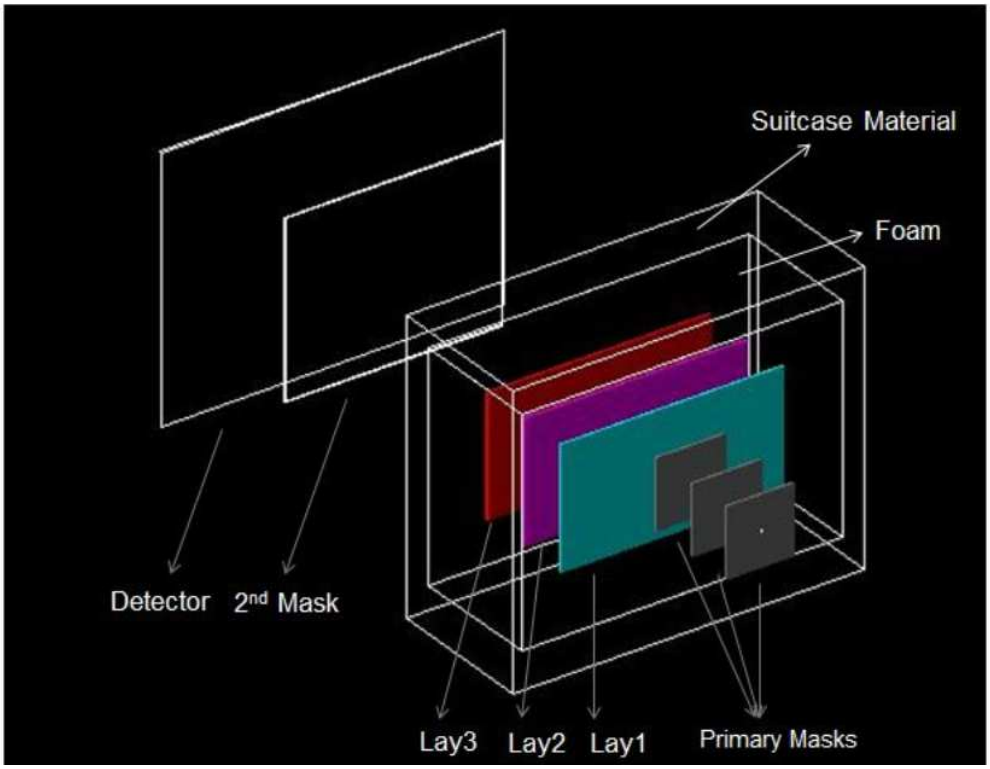


Figure 8.2: Schematic of suitcase phantom with three layers. The middle layer is NaCl, sandwiched between two layers of acrylic. The suitcase exterior and interior are made from polypropylene, and the space between them is filled with foam. This illustration was provided by Pooyan Sahbaee from Duke University.

For the reconstructions, a total of 64 subsets of the measurements were chosen in the same way as described in Section 7.8 and Figure 7.3 of Chapter 7. A region of approximately 448.50 mm along the pencil beam was reconstructed, with a pixel pitch of 1.5 mm. 261 evenly spaced momentum transfer bins from 10 to 140 rad nm^{-1} were used. These predefined momentum transfer samples were later converted into inverse inter-atomic distances $1/(2D)$, measured in nm^{-1} by dividing them by 4π. There were 250 scatter angle samples chosen uniformly from 0 to $\frac{\pi}{3}$ radians, excluding 0.

Since the forward model does not account for X-ray penetration of the beam-stop on the secondary mask, the magnified region occupied by the beam-stop on the detector array has all zeros, while the Monte Carlo data does not. To account for this model mismatch, a detector (0-1) mask was applied to the Monte Carlo data, effectively ignoring the photons that were measured in the magnified beam-stop region.

Figure 8.3 shows the results of estimating the mean detector photon counts and the object scattering density of the NaCl layer amidst the clutter of the suitcase material, foam, and acrylic, using 20 iterations of the ordered subset EM algorithm described in Chapter 7, with 64 subsets. We see that the spatial distribution and the momentum transfer profile are recovered within a few iterations. The three bright bands in Figure 8.3d correspond to the locations of the acrylic, NaCl, and acrylic layers, respectively.

8.3 Energy-sensitive Detectors

In this section, we consider a fan beam source distribution with a linear array of energy-sensitive detectors as shown in Figure 8.4. From the figure, we can see that a fan beam stop is not needed to prevent the detectors from being flooded by photons because the line of detectors are displaced vertically from the fan beam's reach. However, the linear detector array is illuminated by fewer photons compared to the flat-panel detectors, thus requiring longer integration times to acquire a useable signal.

The difference between this geometry and that described in Chapter 7 is the use of energy-sensitive detectors. This substitution results in a modification of the forward model by the introduction of a detector energy response function and an integral over source energies.

Energy-sensitive detectors have been used by Greenberg *et al.* [43] in a pencil beam setting.

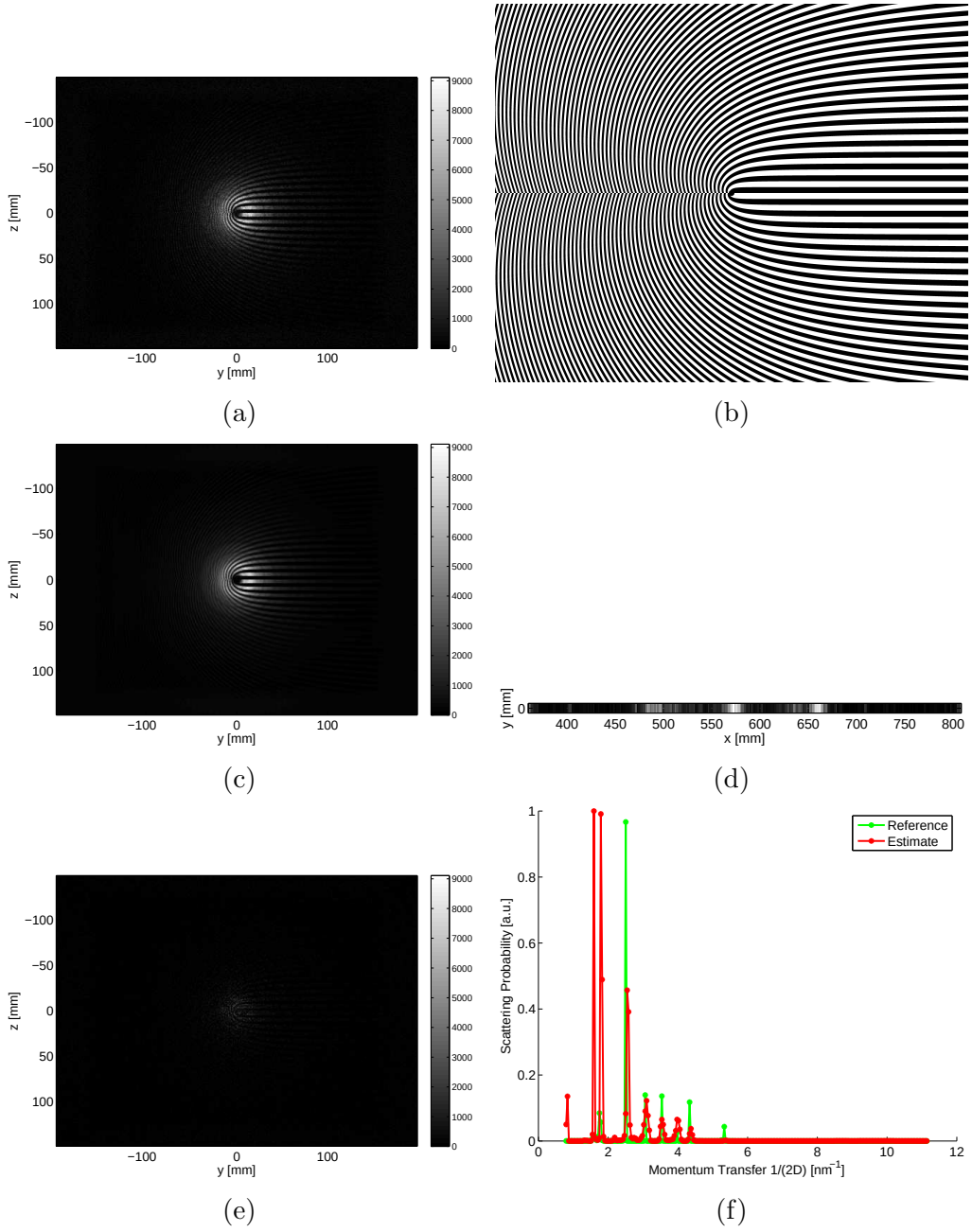


Figure 8.3: Estimation of scattering density from Monte Carlo data of a suitcase with NaCl concealed within. (a) Monte Carlo data. (b) Secondary mask. (c) Estimated mean data. (d) Estimated spatial distribution. (e) Absolute difference between Monte Carlo data and estimated data. (f) Momentum transfer profile of NaCl layer.

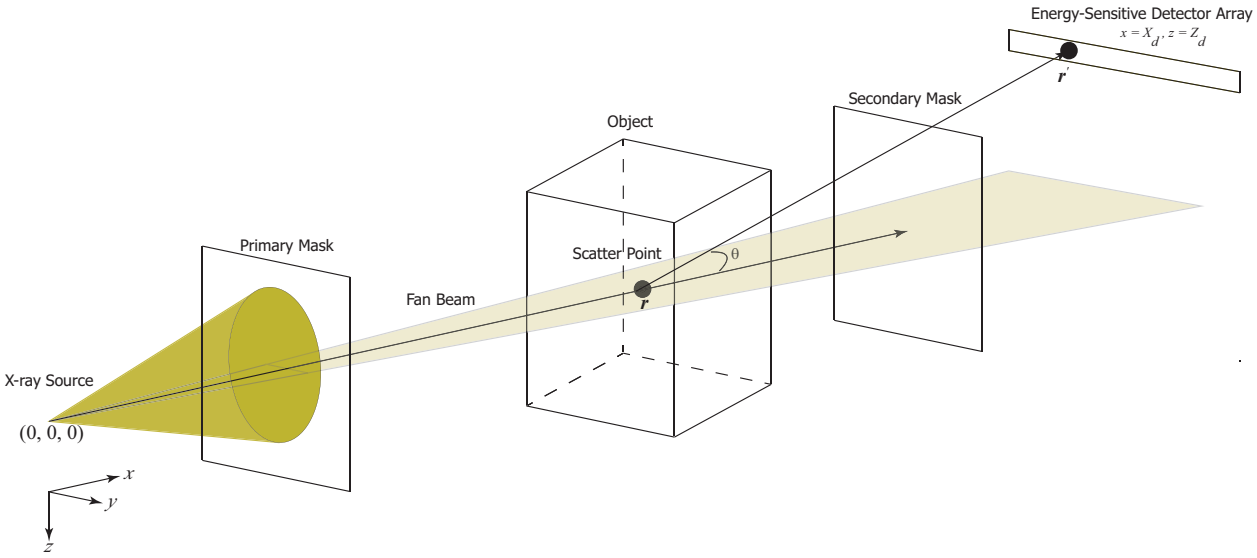


Figure 8.4: Schematic of the fan beam geometry with a line of energy sensitive detectors.

8.3.1 Forward Operator

In this fan beam model, the X-ray source transmits X-rays along the $z = 0$ plane, with the central ray along the x axis. The source is at the origin of the coordinate system. Object scatter locations are indexed by (x, y) and corresponding object point $\mathbf{r} = [x, y, 0]$, with the positive x coordinate pointing from the source through the object to the secondary mask. The linear detector array is aligned parallel to the y-axis. A point on the detector is $\mathbf{r}' = [X_d, y', Z_d]$. The object point and the detector point determine the scatter vector $\mathbf{s} = \mathbf{r}' - \mathbf{r}$. The scatter vector and the object point determine the scatter angle θ.

The scattering follows a radiance model based on ray propagation. The scatter density is a function of the object point and momentum transfer, $f(\mathbf{r}, q)$, where q is given in rad nm^{-1}. Bragg's Law determines the relationship between the energy of the incident X-rays k, the scatter angle θ, and the momentum transfer q by $q = 2k \sin(\theta/2)$.

For a point on the detector with incoming scatter vector, there are geometric factors due to the cosine of the incident angle and due to the squared distance to the object point, $\frac{|\mathbf{n} \cdot \hat{\mathbf{s}}|}{s^2}$, where $\mathbf{n}$ is a unit normal vector on the detector, $\hat{\mathbf{s}}$ is the unit vector in the direction of the scatter vector $\mathbf{s}$, and $s = |\mathbf{s}|$.

The secondary mask between the object and the detector determines which scattered rays are allowed to reach the linear detector array.

The source is polychromatic and filtered before illuminating the object. The flux of the filtered source is energy dependent and given by $W(k)$.

Equation 8.2 shows the model for the flux at detector location $\mathbf{r}'$ and energy k'

$$g(\mathbf{r}', k') = \int d\mathbf{r} \int dq \int dk\, H(\mathbf{r}', k', \mathbf{r}, k, q) f(\mathbf{r}, q), \tag{8.2}$$

where

$$H(\mathbf{r}', k', \mathbf{r}, k, q) = C \left| \frac{\mathbf{n} \cdot \hat{\mathbf{s}}}{s^2} \right| T(\mathbf{r}, \hat{\mathbf{s}}) R(k', k) \left(\frac{1}{2q \sin \frac{\theta}{2}} \right) W \left(\frac{q}{2 \sin \frac{\theta}{2}} \right),$$

C is a normalization constant, and $R(\cdot, \cdot)$ is the detector energy sensitivity (response) function.

In Chapter 7, we gave a detailed description of the forward and backward models for a fan beam source distribution, with an energy-integrating flat-panel detector array. There we described three accelerating techniques, namely: source matrix computation with interpolation of scatter angles, symmetry classes, balancing online/offline computations. The symmetry classes include translation symmetry, left-right symmetry, and up-down symmetry.

For a linear detector array along the y direction, the up-down symmetry is not applicable. Moreover, if the detector is not centered at $y = 0$, the left-right symmetry is also not of consequence. However, the translation symmetry can be used to efficiently compute the forward operator.

The inability to use the mirror symmetry acceleration techniques is placated by the small number of detector pixels, usually on the order of a few hundreds, on the linear detector array, and a small object volume. In such a case, the entire forward operator can be precomputed, saved as a forward matrix, and used offline. However, when the object volume is big enough that the entire forward matrix cannot be stored in random-access memory (RAM), an online computation of the forward operator is needed. In this section, we describe such an online computation of the forward model that makes reconstructing large object volumes possible.

One of the methods we described for reducing the computational burden of the forward operator was the computation of the source matrix along with the interpolation of scatter angles. Predefined samples of the source energy k, scatter angles θ, and momentum transfer q are used in computing the source matrix. This matrix is sparse since it has nonzero elements

only when $(q, \theta, k_d \in (k - \Delta k, k + \Delta k])$ satisfy Bragg's equation, where $2\Delta k$ is the distance between consecutive elements of the uniformly-spaced source energy samples.

Multiplying the source matrix and the scatter density $f(\mathbf{r}, q)$, eliminates the momentum transfer, leaving $Sf(k, \theta)$, a matrix containing values for the samples of the scatter angle θ and source energy k

Pseudocode 8.3 outlines the steps in computing the forward operator and is an extension to energy-sensitives detectors of Pseudocode 7.6, given in Chapter 7.

Pseudocode 8.3 Computational structure for forward operator with energy-sensitive detectors. The linear detector array has N pixels.

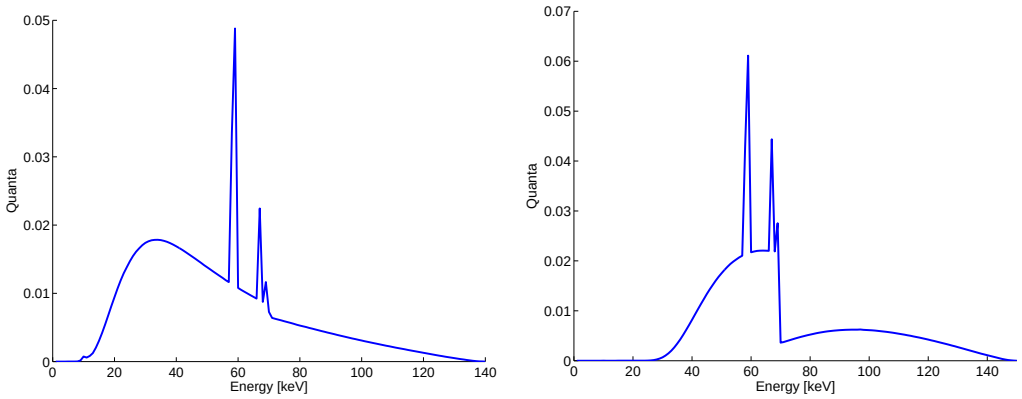
```
compute predefined momentum transfer q samples, scatter angle θ samples, source energy
samples k, and source factor matrix S(k, θ, q) for predefined q, θ, and k samples
compute detector energy response matrix R(k', k)
compute mask modulation factor image
initialize detector image to zero
Given object scattering density f(r, q)
for each object pixel x do
  for each object pixel y do
    Sf(k, θ) = S(k, θ, :) × f([x, y, 0], :)
    if y is first pixel along y direction then
      compute geometry factor image G(1 : N)
    else
      update geometry factor image G(TSF_y + 1 : N) = G(1 : N − TSF_y)
      recompute G(1 : TSF_y)
    end if
    compute and interpolate scatter angle image
    interpolate Sf(k, θ) at scatter angle image
    lookup mask modulation factors
    detector image(k, r') += geometry factor image × mask × interpolated Sf(k, θ)
  end for
end for
detector image = R(k', :) × detector image(:, r')
```

8.3.2 Backward Operator

The procedure for computing the backward operator for a fan beam model with an energy-sensitive detector array stems from that for a fan beam energy-integrating model discussed in Chapter 7 and is given in Pseudocode 8.4.

Pseudocode 8.4 Computational structure for backward operator with energy-sensitive detectors. The linear detector array has N pixels.

compute predefined momentum transfer q samples, scatter angle θ samples, source energy samples k, and source factor matrix $S(k, \theta, q)$ for predefined q, θ, and k samples
compute detector energy response matrix $R(k', k)$
compute mask modulation factor image
Given detector image
detector image($\mathbf{r}'$, k) = $R(:, k) \times$ detector image($\mathbf{r}'$, :)
for each object pixel x **do**
  for each object pixel y **do**
    if y is first pixel along y direction **then**
      compute geometry factor image $G(1 : N)$
    else
      update geometry factor image $G(TSF_y + 1 : N) = G(1 : N - TSF_y)$
      recompute $G(1 : TSF_y)$
    end if
    lookup mask modulation factors
    temp image = geometry factor image $\times$ mask $\times$ detector image
    compute and **interpolate** scatter angle image
    for each sampled scatter angle θ **do**
      temp object (k, θ) = **accumulate** temp image locations at angle θ
    end for
    object image $f([x, y, 0], q)$ = temp object(:) $\times$ $S(:, :, q)$
  end for
end for



(a) (b)

Figure 8.5: Energy spectra for 140 kVp and 150 kVp sources. (a) 140 kVp source. (b) 150 kVp source.

8.3.3 Results

The real data used in validating the fan beam energy-sensitive model was measured by Dr. Joel Greenberg at Duke University using a Smiths HI-SCAN 6040i X-ray attenuation machine [30], which he modified to obtain scatter measurements.

The modified machine has two linear detector arrays, each with 128 pixels and 64 energy channels, that were chained together to form a 256 pixel detector array, with a spacing of 3 mm between the two linear detector arrays. Both of the detectors are Multix energy-sensitive detectors [85], with a full width half maximum (FWHM) of about 6.2 keV. A Gaussian probability density function with the detector energy bin centers, as its mean, and a full width half max (FWHM) of 6.2 keV was used to model the detector energy response for each detector pixel.

The chained detector arrays are laid out along the y direction. With the source at the origin, the center of the chained detector array is located at (967.8, 285.86, 25.4) mm. The pixel pitch along the y direction is 0.8 mm. For the detector energy bins, there are 64 bin centers evenly spaced between 22 and 177 keV. The source was operated at 140 kVp, and the spectrum was filtered before being shaped into a fan by the primary aperture. With the detector energy bin centers given above, and a 140 kVp source, only the first 48 detector energy bins are used. Figure 8.5 shows the spectrum of the 140 kVp and 150 kVp sources used in the pencil beam and energy-sensitive detection settings, respectively.

A 134 mm by 120.7 mm periodic secondary mask was placed 159.5 mm in front of and parallel to the plane of the linear detector array (The linear detector array has a finite size in the direction perpendicular to the direction the pixels are laid out). The center of the mask with respect to the source is (808.3, 260.26, 25.4) mm. The center of the reconstructed object was located at (534.25, 165.96, 0) mm relative to the source. A region of approximately 112 mm by 80 mm, in the xy-plane, was reconstructed, with a pixel pitch of 2.5 mm along the x and y directions. 89 evenly spaced momentum transfer bins from 6 to 50 rad nm^{-1} were used. There were 250 scatter angle samples chosen uniformly from 0 to $\frac{\pi}{6}$ radians, excluding 0.

The results presented here are for a cluttered bag with aluminum crystalline powder inside it. The background data consists of the bag without the aluminum target. For the reconstruction, 16 subsets of the measured data were obtained by using a spacing of 4 in the energy and pixel directions. The results shown in Figure 8.6 were obtained after 10 iterations of the EM algorithm. Based on a photo of the secondary mask in Figure 8.6c, the processed binary secondary mask in Figure 8.6d was obtained by cropping, thresholding, and smoothing using a median filter. The real data displayed in Figure 8.6a is actually the clipped difference between the raw data with target and the background data without the target. Negative values are replaced by 0. This is by no means the best way to deal with the background data in an EM algorithm. This clipped difference is shown here for easy comparison with the estimated data, without the background.

From the results, we can see that the spatial distribution of the target aluminum powder material is accurately estimated, further validating the forward model in an energy-sensitive detection setting. However, the reconstructed momentum transfer profile has a single broad peak, which is flanked by two narrow peaks from the true momentum transfer profile. Also, the other Al powder peak is not visible in the reconstruction. It is not clear why such broadening occurs or why the other peak is attenuated, but a possible explanation could be due to the attenuation of the incident and scattered photons, for which the model does not account.

Figure 8.7 shows the estimation result when a larger range of momentum transfer values are used in the reconstruction; 269 evenly spaced momentum transfer bins from 6 to 140 rad nm^{-1} were used. The momentum transfer profile for the brightest region of the spatial distribution is also plotted alongside that of a reference Al powder.

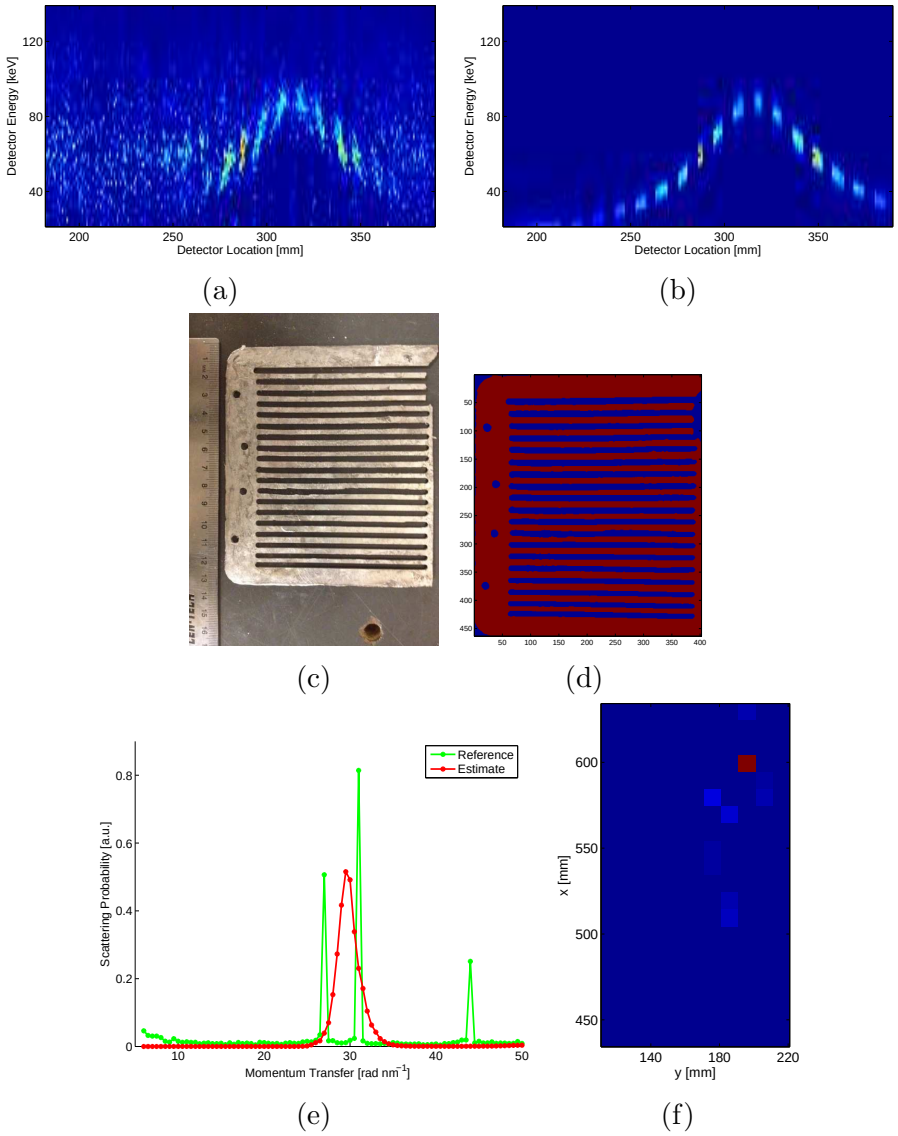


Figure 8.6: Estimation of scattering density from real data of Al in a cluttered bag. The momentum transfer profile of the location with the largest spatial density is plotted alongside the actual profile for Al powder. (a) Real data. (b) Estimated mean data. (c) Raw secondary mask. (d) Processed secondary mask. (e) Momentum transfer profile. (f) Estimated spatial distribution.

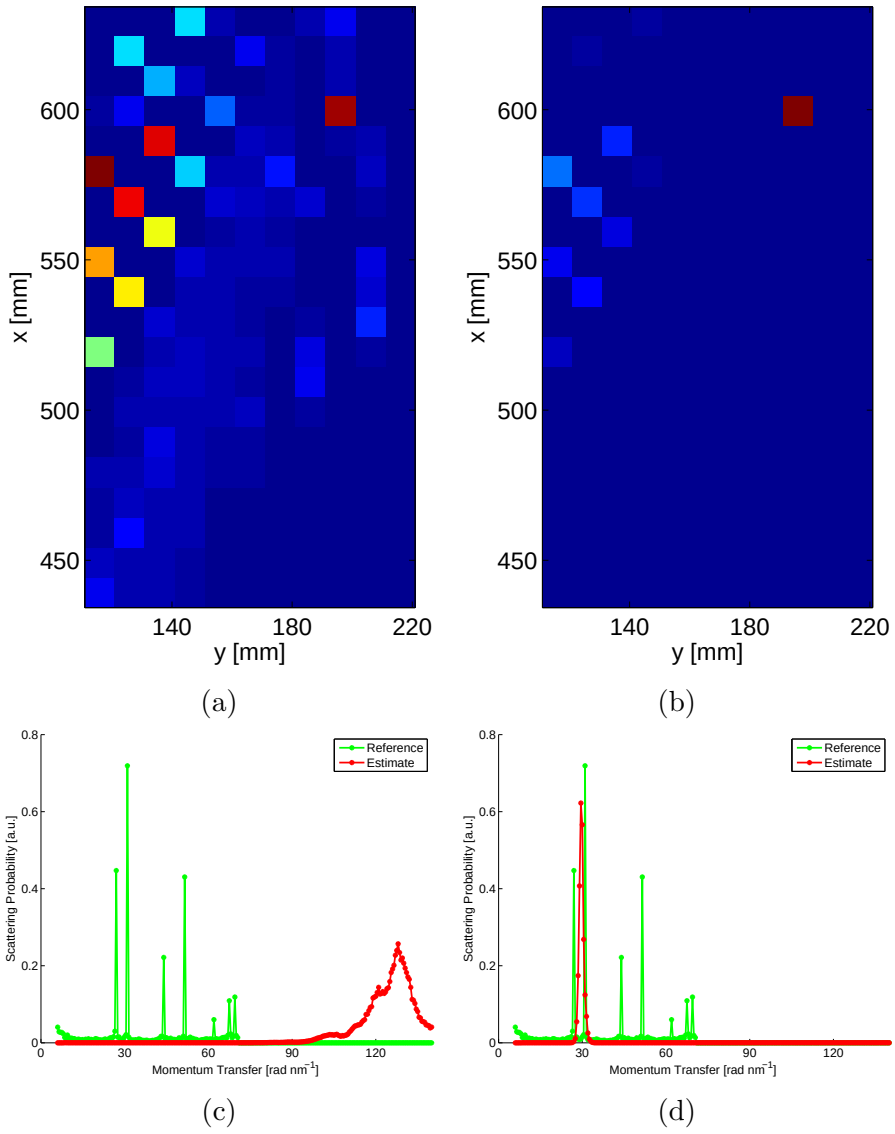


Figure 8.7: Estimation of scattering density from real data of Al in a cluttered bag with extended momentum transfer range. The momentum transfer profile of the location with the largest spatial density is plotted alongside the actual profile for Al powder. The results are shown for 10 and 20 iterations of the ordered subset EM algorithm. (a) Estimated spatial distribution, 10 iterations. (b) Estimated spatial distribution, 20 iterations. (c) Momentum transfer profile, 10 iterations. (d) Momentum transfer profile, 20 iterations.

From the results, we see that after 10 iterations, the scattering density is not yet correctly estimated. It takes about 10 further iterations for a much improved estimate of the scattering density to be obtained. Since the only difference between this setup and the previous one is the momentum transfer range, a possible explanation for this is given by the resolution recovery analysis of Section 7.7, where we showed that the unregularized EM algorithm has poor spectral resolution for larger momentum transfer values, which effectively leads to slower convergence toward the true scattering density. So, it would seem that using a smaller momentum transfer range as in Figure 8.6 is a better reconstruction choice than a larger range as in Figure 8.7.

The results shown here are for the case when the bag is scanned twice, once without the target and the second time, with the target. There is ongoing work in reconstructing an entire bag using only an air scan as the background data. This will require some knowledge about the attenuation map of the bag and may require modeling of multiple scattering phenomena.

8.4 Material Basis

The forward models given in equations 7.1, 8.1, and 8.2 utilize one particular choice of basis for representing the scattering density. The momentum transfer basis seems like the most natural choice since our goal is to use the momentum transfer profile to disambiguate materials. However, due to its high degree of freedom, using the momentum transfer basis could lead to noisy reconstructions, especially from noisy scatter data.

Two ways to compensate for this noisy reconstructions include using a more rigid representational basis or a regularizer. The next section considers the use of an edge-preserving potential function in a neighborhood-based (local) regularizer. The use of a material database as a form of a more rigid (fewer degrees of freedom) representational basis is an attractive choice because it can be easily incorporated into the frameworks described in this chapter and Chapter 7. This ease is contrasted with the method of regularization which prevents a closed-form update rule for the image iterates.

In order to utilize a material basis, we need a material database of momentum transfer profiles. This database of measured momentum transfer profiles was created by our colleague, Dr. Scott Wolter at Duke University. For the predefined momentum transfer values, we interpolate each material's momentum transfer profile, creating a matrix of interpolated momentum transfer profiles.

Pseudocodes 8.5 and 8.6 show the computation of the forward and backward operators, respectively, for the fan beam energy-sensitive model described in Section 8.3 above, when a material basis is used. There are $\mathcal{M}$ materials in the basis. The material matrix denoted by $M(q, m)$ is created by interpolating the m-th material's momentum transfer profile at the predefined momentum transfer value q.

Pseudocode 8.5 Computational structure for forward operator with energy-sensitive detectors and a material basis. The linear detector array has N pixels. There are $\mathcal{M}$ materials in the basis.

```
compute predefined momentum transfer q samples, scatter angle θ samples, source energy
samples k, and source factor matrix S(k, θ, q) for predefined q, θ, and k samples
compute material matrix M(q, m), m = 1 ··· M
compute source factor material matrix SM(k, θ, m) = S(k, θ, :) × M(:, m)
compute detector energy response matrix R(k', k)
compute mask modulation factor image
initialize detector image to zero
Given object scattering density material concentration f̃(r, m)
for each object pixel x do
  for each object pixel y do
    SMf(k, θ) = SM(k, θ, :) × f̃([x, y, 0], :)
    if y is first pixel along y direction then
      compute geometry factor image G(1 : N)
    else
      update geometry factor image G(TSF_y + 1 : N) = G(1 : N − TSF_y)
      recompute G(1 : TSF_y)
    end if
    compute and interpolate scatter angle image
    interpolate SMf(k, θ) at scatter angle image
    lookup mask modulation factors
    detector image(k, r') += geometry factor image × mask × interpolated SMf(k, θ)
  end for
end for
detector image = R(k', :) × detector image(:, r')
```

As seen in Pseudocodes 8.5 and 8.6, the implementation of a material basis simply involves linearly decomposing the scattering density $f(\mathbf{r}, q)$ in terms of the materials' momentum transfer profiles according to

$$f(\mathbf{r}, q) = \sum_{m=1}^{\mathcal{M}} M(q, m)\tilde{f}(\mathbf{r}, m), \tag{8.3}$$

Pseudocode 8.6 Computational structure for backward operator with energy-sensitive detectors and a material basis. The linear detector array has N pixels. There are $\mathcal{M}$ materials in the basis.

compute predefined momentum transfer q samples, scatter angle θ samples, source energy samples k, and source factor matrix $S(k, \theta, q)$ for predefined q, θ, and k samples
compute material matrix $M(q, m), m = 1 \cdots \mathcal{M}$
compute source factor material matrix $SM(k, \theta, m) = S(k, \theta, :) \times M(:, m)$
compute detector energy response matrix $R(k', k)$
compute mask modulation factor image
Given detector image
detector image($\mathbf{r}'$, k) = $R(:, k)$ $\times$ detector image($\mathbf{r}'$, :)
for each object pixel x **do**
  for each object pixel y **do**
    if y is first pixel along y direction **then**
      compute geometry factor image $G(1 : N)$
    else
      update geometry factor image $G(TSF_y + 1 : N) = G(1 : N - TSF_y)$
      recompute $G(1 : TSF_y)$
    end if
    lookup mask modulation factors
    temp image = geometry factor image $\times$ mask $\times$ detector image
    compute and **interpolate** scatter angle image
    for each sampled scatter angle θ **do**
      temp object (k, θ) = **accumulate** temp image locations at angle θ
    end for
    object image material concentration $\tilde{f}([x, y, 0], m)$ = temp object(:) $\times$ $SM(:, :, m)$
  end for
end for

where M is the material matrix and $\tilde{f}(\mathbf{r}, m)$ is the concentration of the m-th material in object location $\mathbf{r}$.

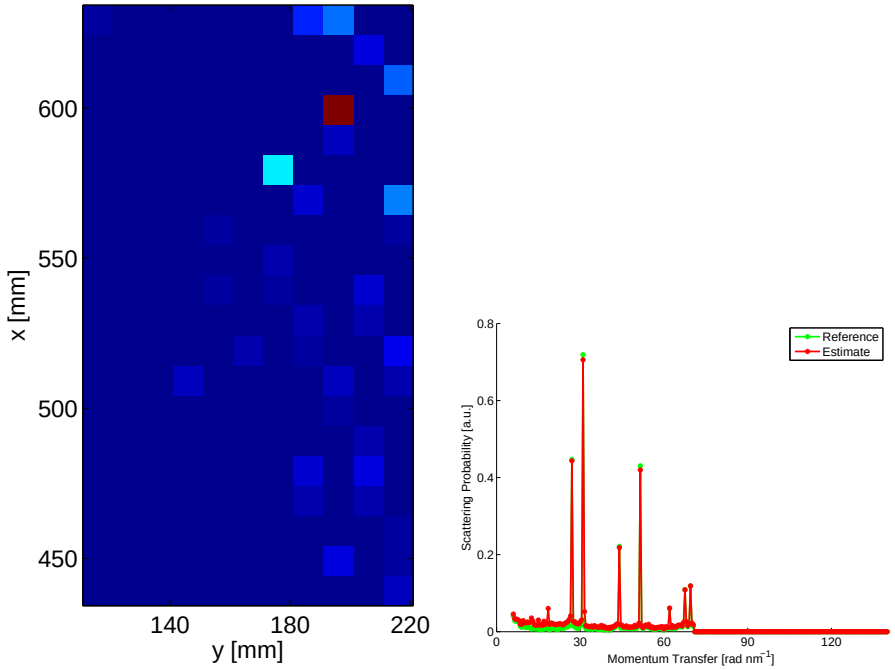
Note that this material basis method can be utilized in any of the other scatter models we have described. We have implemented the material basis method for all the models but only show results for the fan beam energy-sensitive model. To permit accurate reconstructions, special care needs to be taken in designing the material database. For example, a complete basis could be one desirable choice. For the purpose of detecting explosive materials, using potentially dangerous materials as atoms (columns) of the material matrix may be a good design choice. However, as we would show shortly, using a set of materials with limited representational power in the reconstruction problem leads to inaccurate reconstructions.

8.4.1 Results

The same setup that was used in Section 8.3.3 is used here, with the added element of a material basis. The extended momentum transfer range from 6 to 140 rad nm^{-1}, with 269 evenly spaced values, was used to test the ability of the method based on a material basis to quickly recover the object scattering density.

We considered two sets of material bases. In the first case, Al is part of the database, in the second, Al is not. These two choices simulate material bases with strong and weak representational power, respectively. The material database for the first scenario consists of 8 materials, namely: water, teflon, aluminum (Al) powder, high-density polyethylene (HDPE) sheet, 50% hydrogen peroxide, sodium chloride powder, methanol, and graphite powder. Ammonium nitrate powder takes the place of Al powder in the second scenario.

Figure 8.8 shows the results of estimating the scattering density of the Al powder in the cluttered bag using only 10 iterations, with the first set of materials, while Figure 8.9 shows the results when using the second set, using 10 and 20 iterations of the ordered subset EM algorithm. From the figures we can see that the spatial distribution and the momentum transfer profile (at the spatial location with the largest value) are recovered quickly and accurately with Al powder in the material basis. However, when Al is replaced with ammonium nitrate, the basis is no longer expressive for the problem of estimating the location and momentum transfer profile of Al in the bag, and thus suffers from severe estimation errors.



(a) (b)

Figure 8.8: Estimation of scattering density from real data of Al in a cluttered bag using material basis with Al. The momentum transfer profile at the spatial location with the largest value is displayed alongside the momentum transfer profile of Al powder. (a) Estimated spatial distribution. (b) Momentum transfer profile.

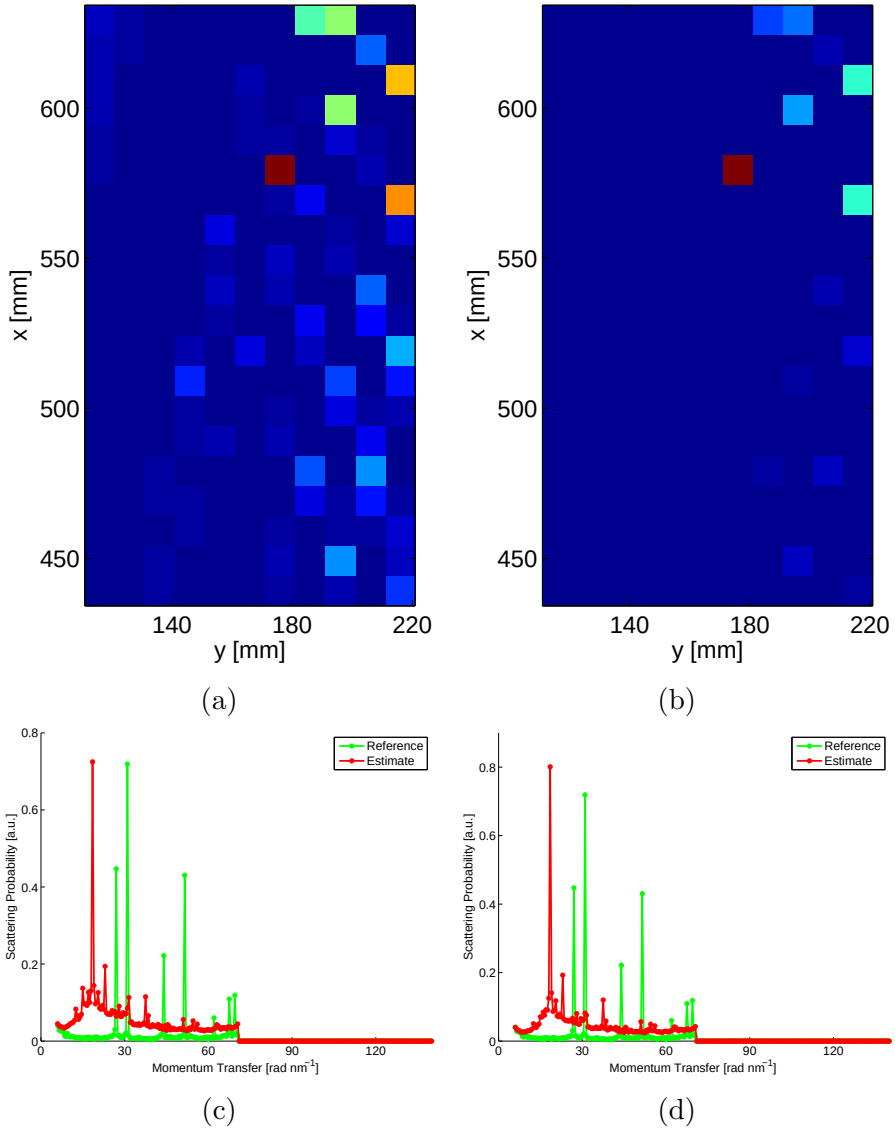


Figure 8.9: Estimation of scattering density from real data of Al in a cluttered bag using material basis without Al. The momentum transfer profile at the spatial location with the largest value is displayed alongside the momentum transfer profile of Al powder. (a) Estimated spatial distribution, 10 iterations. (b) Estimated spatial distribution, 20 iterations. (c) Momentum transfer profile, 10 iterations. (d) Momentum transfer profile, 20 iterations.

8.5 Regularization

Introducing a regularization term in the original scatter objective given in equation 7.5 allows us to use the native momentum transfer basis which may include large momentum transfer values. However, the presence of a regularization term prevents a closed-form expression for the image update.

The penalized (regularized) log-likelihood function for the data is given as

$$\begin{aligned} l(\mathbf{y}|\mathbf{f}) &= \sum_{m\in M}\left[y(m)\ln\left(\sum_{i\in I}H(m,i)f(i)+\mu_b(m)\right)\right. \\ &- \left.\left(\sum_{i\in I}H(m,i)f(i)+\mu_b(m)\right)\right]-\lambda P(\mathbf{f}), \end{aligned} \tag{8.4}$$

where $\mathbf{y}$ is the measured data, $\mathbf{f}$ is the object image, $H(\cdot,\cdot)$ is the system matrix, $\boldsymbol{\mu}_b$ is the mean number of background counts (assumed known), λ is the regularization coefficient that controls the strength of the penalty, and $P(\cdot)$ is the regularization term.

Pseudocode 8.7 shows a modification of Pseudocode 7.9 described in Section 7.6, when regularization is used. We consider an edge-preserving [71] neighborhood-based regularizer, $P(\mathbf{f})$, with decoupled form, $\hat{P}(\mathbf{f})$. The image update is obtained by solving the gradient equation given in equation 8.5.

A one step late (OSL) approach [42] can be used to regain the closed-form solution, but there is no convergence guarantee for large values of the regularization coefficient. In this section, we consider using a trust-region modified Newton's method (see Section 9.5.1) for solving the gradient equation. Chapter 9 provides a more detailed discussion on the regularization issues, including the decoupling of the regularization term and the derivative of the decoupled regularization term.

Since regularization only manifests itself in the EM algorithm and not in the forward and backward operator, regularization can be used with any representational basis for any of the scatter models discussed so far. Although regularization has been implemented for all the models, we only present the results of applying regularization to the fan beam energy-sensitive detection model, in this chapter.

Pseudocode 8.7 Regularized expectation-maximization reconstruction algorithm. λ is the regularization coefficient and $\hat{P(\cdot)}$ is the decoupled regularization term.

initialize $\hat{f}^{k=0}(i)$ to non-negative values
backward project a vector of ones $w_o(i) = \sum_{m \in M} H(m,i)$
for $k = 0$ **to** $K-1$ **do**
  forward project the current estimate $\hat{\mathbf{f}}^k$ to get $\mathbf{g}$
  compute data error ratio $e(m) = y(m)/g(m)$
  backward project data error ratio $w_d(i) = \sum_{m \in M} H(m,i)e(m)$
  update image estimate by solving the equation

$$\frac{\hat{f}^k(i)}{f(i)} w_d(i) - w_o(i) - \lambda \frac{\partial \hat{P}(\mathbf{f})}{\partial f(i)} = 0. \tag{8.5}$$

  set $\hat{\mathbf{f}}^{k+1}$ to the solution of the equation
end for

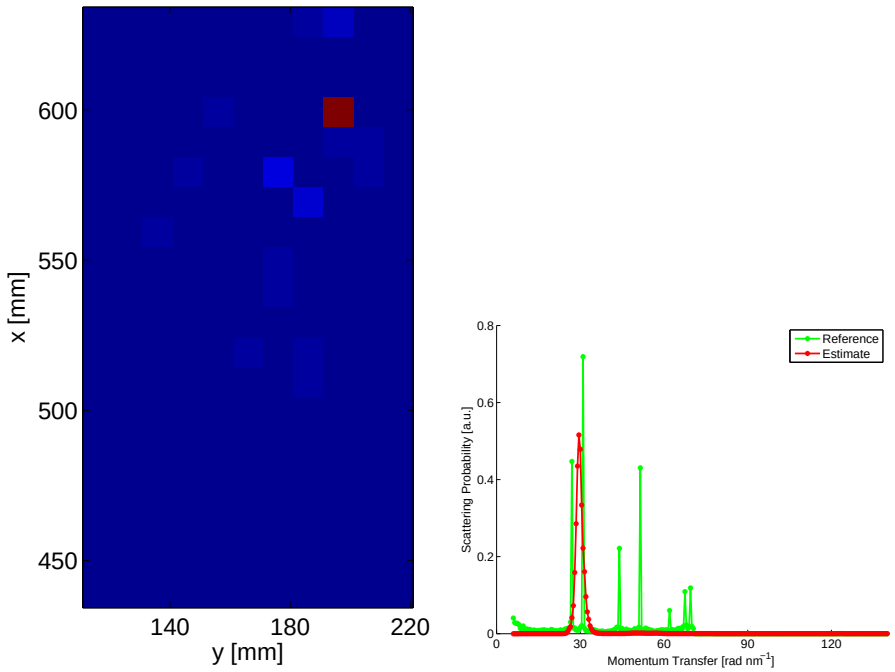
8.5.1 Results

The same setup that was used in Section 8.3.3 is used here, with the added element of a Lange [71] edge-preserving regularization function (see Section 9.5). The extended momentum transfer range from 6 to 140 rad nm^{-1}, with 269 evenly spaced values, was used to reconstruct the object scatter density from the scatter data. The regularization coefficient was $\lambda = 0.00001$ while the regularization scale parameter was set at $\delta = 0.00001$ (See equation 9.49).

Figure 8.10 shows the results of estimating the scattering density of Al powder in the cluttered bag, with the regularizer, using 10 iterations of the ordered subset EM algorithm. The figure suggests that the spatial distribution is recovered quickly and accurately. Again, the momentum transfer profile is split between two major peaks of the true Al powder momentum transfer profile. Also, several peaks with momentum transfer greater than 40 rad nm^{-1} were missed.

8.6 Summary

This chapter presented two modifications to the fan beam energy-integrating detection framework developed in Chapter 7 to accommodate a pencil beam source distribution and energy-sensitive detectors.



(a) (b)

Figure 8.10: Estimation of scattering density from real data of Al in a cluttered bag using regularization. The momentum transfer profile at the spatial location with the largest value is displayed alongside the momentum transfer profile of Al powder. (a) Estimated spatial distribution. (b) Momentum transfer profile at the brightest spatial location.

In addition, two techniques for reducing the sensitivity of the EM algorithm to noise in the measured scatter data were discussed, namely: material basis representation and regularization. It was observed that these two noise-reduction techniques helped the EM algorithm cope with higher momentum transfer values; the EM algorithm quickly converged to the true spatial distribution.

The forward and backward operators and the reconstruction algorithms for the pencil-beam and energy-sensitive geometries were validated using Monte Carlo and real data, respectively.

The results from both geometries and noise reduction techniques are promising, further validating the forward model and its implementation.

8.7 Lessons Learned

Although not shown here, the objective function (log-likelihood) achieved by the EM algorithm increases less rapidly from iteration to iteration when a scatter background, with magnitudes on the order of the real data, is provided.

Knowledge of the clutter, in the form of a background measurement, may make it possible to ignore the effects due to self-attenuation, Compton scattering, and coherent scatter from other materials, in estimating the object spatial distribution. However, ignoring these effects may lead to inaccurate estimates of the momentum transfer profiles, as seen in the reconstructions (the widening of the Al momentum transfer peaks to span multiple reference peaks, in the real data reconstruction).

As observed in the previous chapter, the scatter system better resolves lower momentum transfer values than it does higher ones. Using a smaller momentum transfer range can speed up the convergence of the EM algorithm, provided material identification is still possible within the limited range. The use of regularization or a material basis, which includes the target material, may help in speeding up convergence when using a wide range of momentum transfer values.

Chapter 9

X-ray Attenuation System

9.1 Introduction

Several X-ray computed tomography (CT) machines have been deployed for use in different settings. Most of them rely on a rotating gantry along with a translating stage. In this chapter, we consider a prototype X-ray CT machine that does not utilize a rotating gantry, but rather has a tunnel along which linear detector arrays are laid out, with a few X-ray sources strategically placed. The object of interest is placed on a translational belt which moves it along within the tunnel.

The choice of the source and detector arrays determine what algorithms are appropriate for reconstructing the object from the measurements. If the source has a single peak, then a monoenergetic algorithm is appropriate. On the other hand, a polyenergetic source requires a polyenergetic reconstruction algorithm. In certain cases, a monoenergetic approximation of the polyenergetic model is possible, when certain corrections including beam-hardening have been accounted for.

Measurements made on a photon counting detector, as used in a lot of scanners, can be modeled by a simple Poisson distribution and in the case of large photon flux, a Gaussian approximation can be made. On the other hand, energy integrating (energy fluence) detectors are not Poisson, but compound Poisson distributed [129]. Alternating minimization (AM) Algorithms based on the correct probability distribution have been developed by our group in the past [93, 92].

The algorithm based on a compound Poisson distribution is more accurate, but more computationally intensive than that based on a simple Poisson distribution for energy fluence detection. However, Lasio *et al.* [72] showed that a polyenergetic AM algorithm based on

a simple Poisson distribution is sufficient for reconstructing compound Poisson data, when the mean data is matched by using an energy-scaled source spectrum.

In conjunction with our colleagues at Duke University, we have developed an X-ray software reconstruction package capable of modeling any arbitrary geometry and photon detection method, including photon-counting, energy-sensitive, and energy-integrating detectors. This package allows for a feasible cycle of system design guided by algorithm. Questions like: "Where should the sources be placed?" and "Given a set of detector layouts, which is optimal?", can be easily answered empirically. The prototype X-ray machine uses multiple sources and energy-integrating linear detector boards laid out on the tunnel walls.

This chapter considers a polyenergetic forward model for the system, an optimization problem for estimating the position and orientation of each line of detector, and the steps in reconstructing an object from its attenuation measurements. In addition, we present a trust region modification of the Newton method used in estimating the attenuation coefficients in the presence of a regularization term.

9.2 Prototype X-ray System

An X-ray CT machine was built by our colleagues at Duke University, using the design specifications provided by the Software and Algorithm group, which Washington University researchers are an integral part of. This machine was built as part of our project sponsored by the Department of Homeland Security (DHS). It is a rectangular tunnel with detectors laid out along the walls of the tunnel in ∩s. A translation stage conveys the object of interest across the tunnel as it is illuminated by two sources at opposite ends of the tunnel. Figure 9.1 shows the layout of the detector, sources, and the translation stage. From the figure, we can see that there are some misalignments of the detector boards with the tunnel wall. Section 9.4 considers an approach for estimating the position and orientation of the misaligned detector boards.

The scanner has 10 ∩s (columns) of detectors, each having 12 detector boards, giving a total of 120 detector boards. Each detector board has 128 energy-integrating detector pixels. This gives a total of 15360 detector pixels (measurements per frame). Let us denote the location of each of the 12 detector boards in a ∩ as a row. That is, there are 12 rows. X-rays emanate from two 170 kVp sources with cone angles $(70^\circ, 30^\circ)$. Let the direction along the conveyor

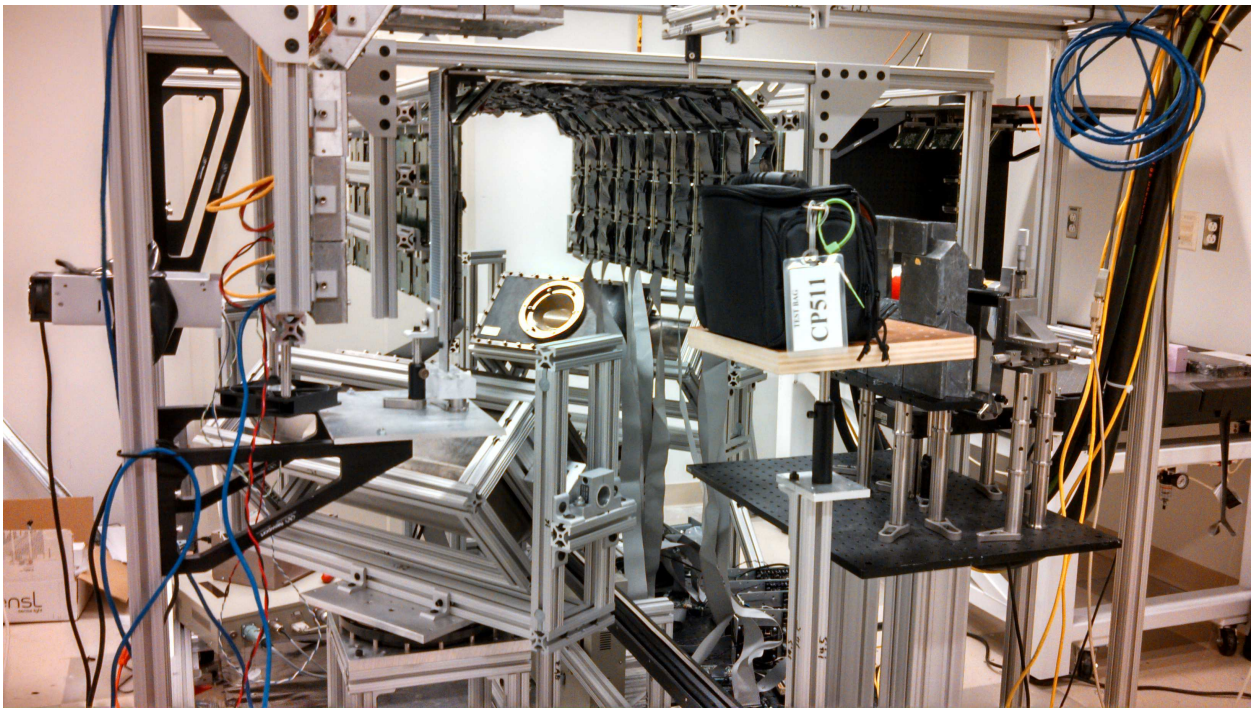


Figure 9.1: Prototype X-ray attenuation system. This photo was taken by Prof. Joseph O'Sullivan at Duke University, in January 2014.

belt be the x direction, while the horizontal and vertical directions are the y and z directions, respectively.

Based on the layout, we can see that the system is view-starved. Not all paths through an extended object is covered as it translates through the tunnel. The scanner was designed to inform material identification using scatter data, by accounting for the attenuation that occurs in the path of the incident and scattered photons.

9.3 Attenuation Data Model

Following the notation of O'Sullivan and Benac [92], let us denote the transmission data by $d(\gamma)$, where γ is a source-detector pair index. We assume the detector measurements are independent Poisson random variables with means $g(\gamma : \mu)$, where the colon is used to indicate the dependence of g on μ. Let $I_0(\gamma, E)$ represent the mean data for source-detector pair γ at energy E, in the absence of background events and the attenuating object of interest. This energy-dependent mean data incorporates whether the sources are on or off,

the X-ray geometric loss due the separation of the source and detector, detector size and shape, a finite source spot, and the detector energy response.

The data means are given as [92]

$$g(\gamma : \mu) = \sum_{E} I_0(\gamma, E) \exp\left(- \sum_{\chi \in \mathrm{X}} h(\gamma|\chi) \mu(\chi, E) \right) + \beta(\gamma), \tag{9.1}$$

where $h(\gamma|\chi)$ is the point spread function, $\mu(\chi, E)$ is the energy-dependent attenuation map, χ is the object's spatial coordinate, and $\beta(\gamma)$ is the mean background data.

Since $\mu(\chi, E)$ has units of mm^{-1}, $h(\gamma|\chi)$ has units of mm and corresponds to the intersection length of the ray along the source-detector pair γ with the object location represented by χ. It accounts for the different ways of discretizing the extended object into voxels. $\beta(\gamma)$ is assumed to be non-negative and known. It includes data from the translation stage and scattered photons from the X-ray scatter system, if the X-ray sources in the scatter system are turned on at the same time the attenuation system is active.

Assuming the separability of the attenuation map $\mu(\chi, E)$ into a linear combination of I basis materials with known energy-dependent attenuation, $\mu_i(E)$, we get [92]

$$\mu(\chi, E) = \sum_{i=1}^{I} \mu_i(E) c_i(\chi), \tag{9.2}$$

where $c_i(\chi)$ is the concentration of material i in voxel χ.

Incorporating the separable attenuation map, the data means are given as

$$g(\gamma : \mu) = \sum_{E} I_0(\gamma, E) \exp\left[- \sum_{\chi \in \mathrm{X}} \sum_{i=1}^{I} h(\gamma|\chi) \mu_i(E) c_i(\chi) \right] + \beta(\gamma). \tag{9.3}$$

9.4 Detector Calibration

We propose to estimate the position and orientation of each line of detector pixels (detector board) laid out along the walls of the tunnel.

In a previous attempt at finding these estimates, we assumed the orientation of each detector board was known, and estimated the position of the board. In addition, our previous approach assumed a single source was illuminating each detector board through an object phantom. When there were multiple phantom placements coupled with multiple sources, the positions from each source-phantom placement pair were estimated separately, and later averaged together. These assumptions allowed the detector position estimation to be an unconstrained least squares problem, whose solution was readily obtained by means of pseudo-inverses.

The new framework allows data from multiple sources and phantom placements to be used in estimating the position and orientation of each detector board. We consider two different objective functions, and a set of constraints which force the solutions to be physically realizable.

In the next subsection, we describe the problem of estimating the location and orientation of a detector board.

9.4.1 Problem Description

We are given a set of sources indexed by $s = 1, \cdots, S$, whose locations are known as $[x_s, y_s, z_s]^\top$, a 2-dimensional phantom grid with G beads, whose relative locations on the grid are known, and a set of D, 1-dimensional detector boards, each with K pixels of uniform pitch p.

Assume the 2-dimensional phantom lies along the yz-plane, centered at $y = 0$, $z = 0$, and is later rotated and translated into the scanner coordinate system, prior to being conveyed along the x-axis of the X-ray tunnel. The phantom can be placed and oriented in different ways prior to being run through the scanner. Let $q = 1, \cdots, Q$ index the different phantom placements. The vector from the location of source s to the g-th grid point (bead) of the q-th phantom placement is given as

$$\begin{bmatrix} x_{sqg} \\ y_{sqg} \\ z_{sqg} \end{bmatrix} = \begin{bmatrix} 1 & R_{q12} & R_{q13} \\ 0 & R_{q22} & R_{q23} \\ 0 & R_{q32} & R_{q33} \end{bmatrix} \begin{bmatrix} 0 \\ y_g \\ z_g \end{bmatrix} + \begin{bmatrix} x(t) \\ Y_q \\ Z_q \end{bmatrix} - \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}, \tag{9.4}$$

where we do not care about the first column, since the first element of the initial position is 0.

Since only the x-coordinate of the phantom changes for a given phantom placement, as it is run through the scanner, we can express $x(t)$ as $x(t) = X_q + \Delta x_g$. As a result, equation 9.4 above becomes

$$\begin{bmatrix} x_{sqg} \\ y_{sqg} \\ z_{sqg} \end{bmatrix} = \begin{bmatrix} 1 & R_{q12} & R_{q13} \\ 0 & R_{q22} & R_{q23} \\ 0 & R_{q32} & R_{q33} \end{bmatrix} \begin{bmatrix} 0 \\ y_g \\ z_g \end{bmatrix} + \begin{bmatrix} X_q \\ Y_q \\ Z_q \end{bmatrix} + \begin{bmatrix} \Delta x_g \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}. \tag{9.5}$$

Overall, the source to phantom vector is given by

$$\begin{bmatrix} x_{sqg} \\ y_{sqg} \\ z_{sqg} \end{bmatrix} = \begin{bmatrix} X_q + \Delta x_g + R_{q12} y_g + R_{q13} z_g - x_s \\ Y_q + R_{q22} y_g + R_{q23} z_g - y_s \\ Z_q + R_{q32} y_g + R_{q33} z_g - z_s \end{bmatrix}. \tag{9.6}$$

Assume the 1-dimensional detector boards are originally laid out along the z-axis, with the first pixel's lowest edge at $z = 0$, and later rotated and translated into the scanner coordinate system. The vector from the location of source s to the k-th pixel of detector board d is given as:

$$\begin{bmatrix} \nu_{sdk} \\ \eta_{sdk} \\ \phi_{sdk} \end{bmatrix} = \begin{bmatrix} 1 & 0 & R_{d13} \\ 0 & 1 & R_{d23} \\ 0 & 0 & R_{d33} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ kp \end{bmatrix} + \begin{bmatrix} X_d \\ Y_d \\ Z_d \end{bmatrix} - \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}, \tag{9.7}$$

where we do not care about the first and second columns, since the first and second elements of the initial pixel position are 0.

Overall, the source to detector pixel vector is given by

$$\begin{bmatrix} \nu_{sdk} \\ \eta_{sdk} \\ \phi_{sdk} \end{bmatrix} = \begin{bmatrix} kpR_{d13} + X_d - x_s \\ kpR_{d23} + Y_d - y_s \\ kpR_{d33} + Z_d - z_s \end{bmatrix}. \tag{9.8}$$

The goal of finding the location and orientation of the detector pixels can be turned into an optimization problem, where the objective is related to how well the vector from a source through a phantom grid point aligns with the vector from the source to the detector pixel — which captures the shadow of the phantom grid point. Finding the central location of the shadow of the grid point requires a centroid estimation algorithm. We assume that the locations $k, 1 \leq k \leq K$ have already been estimated via centroiding.

9.4.2 Detector Pose Estimation

The problem of estimating the location and orientation of the detector pixels or the location and orientation of the phantom can be cast as a constrained optimization problem. For each of the two problems of estimating the detector pixels' pose and those of the phantoms, we consider using either of two objectives: a method based on the magnitude of the normalized cross product, and another based on the negative normalized dot product of the source to phantom and source to detector pixel vectors.

Assume we are given the location of the sources, those of the phantom grid points which cast their shadows on a given detector board d, where $d = 1, \cdots, D$, and where in the detector board the center of the shadow occurs, k, where $1 \leq k \leq K$. That is, x_s, y_s, z_s, x_{sqg}, y_{sqg}, z_{sqg}, k, and d are known. We would like to estimate the pose parameters $\mathbf{w}_d$ of the board,

$$\mathbf{w}_d = [R_{d13}, R_{d23}, R_{d33}, X_d, Y_d, Z_d]^\top .$$

Normalized Cross Product

The magnitude of the normalized cross product between the source to phantom grid point vector and the source to detector pixel vector is given by

$$f^1_{nd} = \frac{\sqrt{(y_n\phi_{nd} - z_n\eta_{nd})^2 + (z_n\nu_{nd} - x_n\phi_{nd})^2 + (x_n\eta_{nd} - y_n\nu_{nd})^2}}{\sqrt{x_n^2 + y_n^2 + z_n^2}\sqrt{\nu_{nd}^2 + \eta_{nd}^2 + \phi_{nd}^2}}, \tag{9.9}$$

where n in x_n, y_n, z_n indexes the source, phantom-placement, grid point triple (s, q, g) and n in $\nu_{nd}, \eta_{nd}, \phi_{nd}$ indicates that the detector pixel (pixel in the shadow) k depends on the triple.

The overall objective function based on all the phantom grid points and sources is given by the sum

$$f_d^1 = \sum_{n=1}^{N_d} f_{nd}^1, \tag{9.10}$$

where N_d is the number of phantom grid points that cast their shadows on detector board d.

The gradient of f_d^1 with respect to the optimization vector $\mathbf{w}_d$ is given by

$$\begin{aligned}\nabla_{\mathbf{w}_d} f_d^1 = & \sum_{n=1}^{N_d} \frac{1}{\sqrt{x_n^2+y_n^2+z_n^2}\sqrt{\nu_{nd}^2+\eta_{nd}^2+\phi_{nd}^2}} \times \\ & \begin{bmatrix} k_n p\left(\alpha_1\left(z_n\left(z_n\nu_{nd}-x_n\phi_{nd}\right)-y_n\left(x_n\eta_{nd}-y_n\nu_{nd}\right)\right)-\alpha_2\nu_{nd}\right) \\ k_n p\left(\alpha_1\left(x_n\left(x_n\eta_{nd}-y_n\nu_{nd}\right)-z_n\left(y_n\phi_{nd}-z_n\eta_{nd}\right)\right)-\alpha_2\eta_{nd}\right) \\ k_n p\left(\alpha_1\left(y_n\left(y_n\phi_{nd}-z_n\eta_{nd}\right)-x_n\left(z_n\nu_{nd}-x_n\phi_{nd}\right)\right)-\alpha_2\phi_{nd}\right) \\ \alpha_1\left(z_n\left(z_n\nu_{nd}-x_n\phi_{nd}\right)-y_n\left(x_n\eta_{nd}-y_n\nu_{nd}\right)\right)-\alpha_2\nu_{nd} \\ \alpha_1\left(x_n\left(x_n\eta_{nd}-y_n\nu_{nd}\right)-z_n\left(y_n\phi_{nd}-z_n\eta_{nd}\right)\right)-\alpha_2\eta_{nd} \\ \alpha_1\left(y_n\left(y_n\phi nd-z_n\eta_{nd}\right)-x_n\left(z_n\nu_{nd}-x_n\phi_{nd}\right)\right)-\alpha_2\phi_{nd} \end{bmatrix},\end{aligned} \tag{9.11}$$

where we use the subscript n in k_n to explicitly show the dependence of the detector pixel on the (s,q,g) triple, and the constants α_1, α_2 are given as

$$\alpha_1 = \frac{1}{\sqrt{(y_n\phi_{nd}-z_n\eta_{nd})^2+(z_n\nu_{nd}-x_n\phi_{nd})^2+(x_n\eta_{nd}-y_n\nu_{nd})^2}},$$

$$\alpha_2 = \frac{\sqrt{(y_n\phi_{nd}-z_n\eta_{nd})^2+(z_n\nu_{nd}-x_n\phi_{nd})^2+(x_n\eta_{nd}-y_n\nu_{nd})^2}}{\nu_{nd}^2+\eta_{nd}^2+\phi_{nd}^2}.$$

Normalized Dot Product

The normalized dot product between the source to phantom grid point vector and the source to detector pixel vector is given by

$$f_{nd}^2 = \frac{x_n\nu_{nd} + y_n\eta_{nd} + z_n\phi_{nd}}{\sqrt{x_n^2 + y_n^2 + z_n^2}\sqrt{\nu_{nd}^2 + \eta_{nd}^2 + \phi_{nd}^2}}. \tag{9.12}$$

The overall objective function based on all the phantom grid points and sources is given by the sum

$$f_d^2 = -\sum_{n=1}^{N_d} f_{nd}^2. \tag{9.13}$$

The gradient of f_d^2 with respect to the optimization vector $\mathbf{w}_d$ is given by

$$\nabla_{\mathbf{w}_d} f_d^2 = -\sum_{n=1}^{N_d} \frac{1}{\sqrt{x_n^2 + y_n^2 + z_n^2}\sqrt{\nu_{nd}^2 + \eta_{nd}^2 + \phi_{nd}^2}} \times \begin{bmatrix} k_n p\left(x_n - \frac{x_n\nu_{nd}+y_n\eta_{nd}+z_n\phi_{nd}}{\nu_{nd}^2+\eta_{nd}^2+\phi_{nd}^2}\nu_{nd}\right) \\ k_n p\left(y_n - \frac{x_n\nu_{nd}+y_n\eta_{nd}+z_n\phi_{nd}}{\nu_{nd}^2+\eta_{nd}^2+\phi_{nd}^2}\eta_{nd}\right) \\ k_n p\left(z_n - \frac{x_n\nu_{nd}+y_n\eta_{nd}+z_n\phi_{nd}}{\nu_{nd}^2+\eta_{nd}^2+\phi_{nd}^2}\phi_{nd}\right) \\ x_n - \frac{x_n\nu_{nd}+y_n\eta_{nd}+z_n\phi_{nd}}{\nu_{nd}^2+\eta_{nd}^2+\phi_{nd}^2}\nu_{nd} \\ y_n - \frac{x_n\nu_{nd}+y_n\eta_{nd}+z_n\phi_{nd}}{\nu_{nd}^2+\eta_{nd}^2+\phi_{nd}^2}\eta_{nd} \\ z_n - \frac{x_n\nu_{nd}+y_n\eta_{nd}+z_n\phi_{nd}}{\nu_{nd}^2+\eta_{nd}^2+\phi_{nd}^2}\phi_{nd} \end{bmatrix}. \tag{9.14}$$

Constraints

Since the third column of the detector rotation matrix is $[R_{d13}, R_{d23}, R_{d33}]^\top$, the vector must have unit norm. This gives us an equality constraint,

$$c_d^1 = \mathbf{w}_d^\top \mathbf{A} \mathbf{w}_d - 1 = 0, \tag{9.15}$$

where $\mathbf{A}$ is a 6×6 matrix given as

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The gradient of this equality constraint with respect to the optimization vector $\mathbf{w}_d$ is $\nabla_{\mathbf{w}_d} c_d^1 = 2\mathbf{A}\mathbf{w}_d$.

Also, since the magnitude of the normalized cross product does not account for the difference between parallel and anti-parallel vectors, in using the cross product objective, we must enforce a strict orientation constraint: both the source to phantom and source to detector pixel vectors must be facing the same way. That is, their dot product must be positive (relaxed to be non-negative), or

$$c_{nd}^2 = -\left(x_n \nu_{nd} + y_n \eta_{nd} + z_n \phi_{nd}\right) \leq 0. \tag{9.16}$$

The gradient of this inequality constraint with respect to the optimization vector $\mathbf{w}_d$ is

$$\nabla_{\mathbf{w}_d} c_{nd}^2 = -\begin{bmatrix} x_n k_n p \\ y_n k_n p \\ z_n k_n p \\ x_n \\ y_n \\ z_n \end{bmatrix}. \tag{9.17}$$

This constraint is only used by the optimization problem based on the normalized cross product objective, and is applicable for each $n = 1, \cdots, N_d$.

Moreover, we placed magnification constraints on the source to detector pixel vector, relative to the source to phantom grid point vector. The lowest and highest magnifications are given as m and M, respectively. The magnification constraints are given as

$$c_{nd}^3 = \left(\nu_{nd}^2 + \eta_{nd}^2 + \phi_{nd}^2\right) - M^2\left(x_n^2 + y_n^2 + z_n^2\right) \leq 0, \tag{9.18}$$

$$c_{nd}^4 = m^2\left(x_n^2 + y_n^2 + z_n^2\right) - \left(\nu_{nd}^2 + \eta_{nd}^2 + \phi_{nd}^2\right) \leq 0. \tag{9.19}$$

The gradient of the inequality constraint 9.18 with respect to the optimization vector $\mathbf{w}_d$ is

$$\nabla_{\mathbf{w}_d} c_{nd}^3 = 2\begin{bmatrix} \nu_{nd}k_n p \\ \eta_{nd}k_n p \\ \phi_{nd}k_n p \\ \nu_{nd} \\ \eta_{nd} \\ \phi_{nd} \end{bmatrix}. \tag{9.20}$$

The gradient of the inequality constraint 9.19 with respect to the optimization vector $\mathbf{w}_d$ is $-\nabla_{\mathbf{w}_d} c_{nd}^3$.

In addition, we placed boundary constraints on the position of the detector pixels, by using the approximate dimensions of the tunnel, relative to the scanner coordinate system. Given a tunnel with boundary specified by $[x_{\min}, x_{\max}, y_{\min}, y_{\max}, z_{\min}, z_{\max}]$ the boundary constraints are

$$c_{nd}^5 = x_{\min} - (\nu_{nd} + x_s) \leq 0, \tag{9.21}$$

$$c_{nd}^6 = (\nu_{nd} + x_s) - x_{\max} \leq 0, \tag{9.22}$$

$$c_{nd}^7 = y_{\min} - (\eta_{nd} + y_s) \leq 0, \tag{9.23}$$

$$c_{nd}^8 = (\eta_{nd} + y_s) - y_{\max} \leq 0, \tag{9.24}$$

$$c_{nd}^{9} = z_{\min} - (\phi_{nd} + z_s) \leq 0, \tag{9.25}$$

$$c_{nd}^{10} = (\phi_{nd} + z_s) - z_{\max} \leq 0. \tag{9.26}$$

The gradients of inequality constraints 9.21, 9.22, 9.23, 9.24, 9.25, and 9.26 with respect to the optimization vector $\mathbf{w}_d$ are:

$$\begin{bmatrix} -k_n p \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} k_n p \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -k_n p \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ k_n p \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ -k_n p \\ 0 \\ 0 \\ -1 \end{bmatrix}, \text{ and } \begin{bmatrix} 0 \\ 0 \\ k_n p \\ 0 \\ 0 \\ 1 \end{bmatrix}, \tag{9.27}$$

respectively.

The last constraint placed on the detector boards involves their orientation. If we know a priori that each detector board is closely aligned with a given direction (tethering direction), then we can constrain the optimization variables according to

$$c_d^{11} = \cos(\theta_t) - (\mathbf{v}_t(1)\mathbf{w}_d(1) + \mathbf{v}_t(2)\mathbf{w}_d(2) + \mathbf{v}_t(3)\mathbf{w}_d(3)) \leq 0, \tag{9.28}$$

where θ_t is the tethering angle and $\mathbf{v}_t$ is the tethering direction.

The gradient of constraint 9.28 is given as $\left[-\mathbf{v}_t^\top, 0, 0, 0\right]^\top$.

9.4.3 Phantom Pose Estimation

Assume we are given the location of the sources, those of the detector board pixels, and the phantom grid point indexes corresponding to the given detector board pixel locations. To estimate the pose of phantom placement q, where $q = 1, \cdots, Q$, we utilize only the detector pixels that the phantom casts a shadow on, in pose q. That is, for pose q, we know the values of $x_s, y_s, z_s, \nu_{sdk}, \eta_{sdk}, \phi_{sdk}, y_g, z_g,$ and Δx_g.

Let n denote the triple (s, d, k), where $n = 1, \cdots, N_q$, and N_q is the number of known detector board pixel locations for pose q. For each n, there is a corresponding phantom grid point index g. As such, we write g as $g(n)$. We would like to estimate the pose parameters $\mathbf{w}_q$ of the phantom, $\mathbf{w}_q = [R_{q12}, R_{q22}, R_{q32}, R_{q13}, R_{q23}, R_{q33}, X_q, Y_q, Z_q]^\top$.

Normalized Cross Product

The magnitude of the normalized cross product between the source to phantom grid point vector and the source to detector pixel vector has a form similar to equation 9.9, and is given as

$$f_{nq}^1 = \frac{\sqrt{(y_{nq}\phi_n - z_{nq}\eta_n)^2 + (z_{nq}\nu_n - x_{nq}\phi_n)^2 + (x_{nq}\eta_n - y_{nq}\nu_n)^2}}{\sqrt{x_{nq}^2 + y_{nq}^2 + z_{nq}^2}\sqrt{\nu_n^2 + \eta_n^2 + \phi_n^2}}, \tag{9.29}$$

where n in ν_n, η_n, ϕ_n indexes the source, detector board, pixel (s, d, k) and n in x_{nq}, y_{nq}, z_{nq} indicates that the triple determines the location of the phantom point.

The overall objective function based on all the known detector board pixel locations and sources is given by the sum

$$f_q^1 = \sum_{n=1}^{N_q} f_{nq}^1. \tag{9.30}$$

The gradient of f_q^1 with respect to the optimization vector $\mathbf{w}_q$ is given by

$$\nabla_{\mathbf{w}_q} f_q^1 = \sum_{n=1}^{N_q} \frac{1}{\sqrt{x_{nq}^2 + y_{nq}^2 + z_{nq}^2}\sqrt{\nu_n^2 + \eta_n^2 + \phi_n^2}} \times$$

$$\begin{bmatrix} y_{g(n)} \left(\beta_1 \left(\eta_n \left(x_{nq}\eta_n - y_{nq}\nu_n\right) - \phi_n \left(z_{nq}\nu_n - x_{nq}\phi_n\right)\right) - \beta_2 x_{nq}\right) \\ y_{g(n)} \left(\beta_1 \left(\phi_n \left(y_{nq}\phi_n - z_{nq}\eta_n\right) - \nu_n \left(x_{nq}\eta_n - y_{nq}\nu_n\right)\right) - \beta_2 y_{nq}\right) \\ y_{g(n)} \left(\beta_1 \left(\nu_n \left(z_{nq}\nu_n - x_{nq}\phi_n\right) - \eta_n \left(y_{nq}\phi_n - z_{nq}\eta_n\right)\right) - \beta_2 z_{nq}\right) \\ z_{g(n)} \left(\beta_1 \left(\eta_n \left(x_{nq}\eta_n - y_{nq}\nu_n\right) - \phi_n \left(z_{nq}\nu_n - x_{nq}\phi_n\right)\right) - \beta_2 x_{nq}\right) \\ z_{g(n)} \left(\beta_1 \left(\phi_n \left(y_{nq}\phi_n - z_{nq}\eta_n\right) - \nu_n \left(x_{nq}\eta_n - y_{nq}\nu_n\right)\right) - \beta_2 y_{nq}\right) \\ z_{g(n)} \left(\beta_1 \left(\nu_n \left(z_{nq}\nu_n - x_{nq}\phi_n\right) - \eta_n \left(y_{nq}\phi_n - z_{nq}\eta_n\right)\right) - \beta_2 z_{nq}\right) \\ \beta_1 \left(\eta_n \left(x_{nq}\eta_n - y_{nq}\nu_n\right) - \phi_n \left(z_{nq}\nu_n - x_{nq}\phi_n\right)\right) - \beta_2 x_{nq} \\ \beta_1 \left(\phi_n \left(y_{nq}\phi_n - z_{nq}\eta_n\right) - \nu_n \left(x_{nq}\eta_n - y_{nq}\nu_n\right)\right) - \beta_2 y_{nq} \\ \beta_1 \left(\nu_n \left(z_{nq}\nu_n - x_{nq}\phi_n\right) - \eta_n \left(y_{nq}\phi_n - z_{nq}\eta_n\right)\right) - \beta_2 z_{nq} \end{bmatrix}, \tag{9.31}$$

where the constants β_1 and β_2 are given as

$$\beta_1 = \frac{1}{\sqrt{(y_{nq}\phi_n - z_{nq}\eta_n)^2 + (z_{nq}\nu_n - x_{nq}\phi_n)^2 + (x_{nq}\eta_n - y_{nq}\nu_n)^2}},$$

$$\beta_2 = \frac{\sqrt{(y_{nq}\phi_n - z_{nq}\eta_n)^2 + (z_{nq}\nu_n - x_{nq}\phi_n)^2 + (x_{nq}\eta_n - y_{nq}\nu_n)^2}}{x_{nq}^2 + y_{nq}^2 + z_{nq}^2}.$$

Normalized Dot Product

Similar to equation 9.12, the normalized dot product between the source to phantom grid point vector and the source to detector pixel vector is given by

$$f_{nq}^2 = \frac{x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n}{\sqrt{x_{nq}^2 + y_{nq}^2 + z_{nq}^2}\sqrt{\nu_n^2 + \eta_n^2 + \phi_n^2}}. \tag{9.32}$$

The overall objective function based on all the known detector board pixel locations and sources is given by the sum

$$f_q^2 = -\sum_{n=1}^{N_q} f_{nq}^2. \tag{9.33}$$

The gradient of f_q^2 with respect to the optimization vector $\mathbf{w}_q$ is given by

$$\nabla_{\mathbf{w}_q} f_q^2 = -\sum_{n=1}^{N_q} \frac{1}{\sqrt{x_{nq}^2 + y_{nq}^2 + z_{nq}^2}\sqrt{\nu_n^2 + \eta_n^2 + \phi_n^2}} \times \begin{bmatrix} y_{g(n)}\left(\nu_n - \frac{x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n}{x_{nq}^2 + y_{nq}^2 + z_{nq}^2} x_{nq}\right) \\ y_{g(n)}\left(\eta_n - \frac{x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n}{x_{nq}^2 + y_{nq}^2 + z_{nq}^2} y_{nq}\right) \\ y_{g(n)}\left(\phi_n - \frac{x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n}{x_{nq}^2 + y_{nq}^2 + z_{nq}^2} z_{nq}\right) \\ z_{g(n)}\left(\nu_n - \frac{x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n}{x_{nq}^2 + y_{nq}^2 + z_{nq}^2} x_{nq}\right) \\ z_{g(n)}\left(\eta_n - \frac{x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n}{x_{nq}^2 + y_{nq}^2 + z_{nq}^2} y_{nq}\right) \\ z_{g(n)}\left(\phi_n - \frac{x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n}{x_{nq}^2 + y_{nq}^2 + z_{nq}^2} z_{nq}\right) \\ \nu_n - \frac{x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n}{x_{nq}^2 + y_{nq}^2 + z_{nq}^2} x_{nq} \\ \eta_n - \frac{x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n}{x_{nq}^2 + y_{nq}^2 + z_{nq}^2} y_{nq} \\ \phi_n - \frac{x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n}{x_{nq}^2 + y_{nq}^2 + z_{nq}^2} z_{nq} \end{bmatrix}. \tag{9.34}$$

Constraints

The second and third columns of the phantom rotation matrix are

$$[R_{q12}, R_{q22}, R_{q32}]^\top \text{ and } [R_{q13}, R_{q23}, R_{q33}]^\top,$$

respectively.

Both vectors must form an orthonormal pair. That is, they must each have unit norm, and their dot product must be 0. This gives us three equality constraints, namely

$$c_q^1 = \mathbf{w}_q^\top \mathbf{A}_1 \mathbf{w}_q - 1 = 0, \tag{9.35}$$

$$c_q^2 = \mathbf{w}_q^\top \mathbf{A}_2 \mathbf{w}_q - 1 = 0, \tag{9.36}$$

$$c_q^3 = \mathbf{w}_q^\top \mathbf{A}_3 \mathbf{w}_q = 0, \tag{9.37}$$

where $\mathbf{A}_1, \mathbf{A}_2,$ and $\mathbf{A}_3$ are 9×9 matrices given by

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{A}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{A}_3 = \begin{bmatrix} 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0 & 0 \\ 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The gradient of the equality constraints with respect to the optimization vector $\mathbf{w}_q$ are $\nabla_{\mathbf{w}_q} c_q^1 = 2\mathbf{A}_1\mathbf{w}_q, \nabla_{\mathbf{w}_q} c_q^2 = 2\mathbf{A}_2\mathbf{w}_q,$ and $\nabla_{\mathbf{w}_q} c_q^3 = 2\mathbf{A}_3\mathbf{w}_q$.

As was the case for detector board pose estimation, since the magnitude of the normalized cross product does not account for the difference between parallel and anti-parallel vectors, in using the cross product objective, we enforce an orientation constraint,

$$c_{nq}^4 = -\left(x_{nq}\nu_n + y_{nq}\eta_n + z_{nq}\phi_n\right) \leq 0. \tag{9.38}$$

The gradient of this inequality constraint with respect to the optimization vector $\mathbf{w}_q$ is

$$\nabla_{\mathbf{w}_q} c_{nq}^4 = -\begin{bmatrix} \nu_n y_{g(n)} \\ \eta_n y_{g(n)} \\ \phi_n y_{g(n)} \\ \nu_n z_{g(n)} \\ \eta_n z_{g(n)} \\ \phi_n z_{g(n)} \\ \nu_n \\ \eta_n \\ \phi_n \end{bmatrix}. \tag{9.39}$$

This constraint is only used by the optimization problem based on the normalized cross product objective, and is applicable for each $n = 1, \cdots, N_q$.

In addition, similar to what was done for detector pose estimation, we can place magnification constraints on the source to detector pixel vector, relative to the source to phantom grid point vector. The lowest and highest magnifications are given as m and M, respectively. The magnification constraints are

$$c_{nq}^5 = \left(\nu_n^2 + \eta_n^2 + \phi_n^2\right) - M^2\left(x_{nq}^2 + y_{nq}^2 + z_{nq}^2\right) \leq 0, \tag{9.40}$$

$$c_{nq}^6 = m^2\left(x_{nq}^2 + y_{nq}^2 + z_{nq}^2\right) - \left(\nu_n^2 + \eta_n^2 + \phi_n^2\right) \leq 0. \tag{9.41}$$

The gradient of the inequality constraint 9.40 with respect to the optimization vector $\mathbf{w}_q$ is

$$\nabla_{\mathbf{w}_q} c_{nq}^5 = -2M^2 \begin{bmatrix} x_{nq}y_{g(n)} \\ y_{nq}y_{g(n)} \\ z_{nq}y_{g(n)} \\ x_{nq}z_{g(n)} \\ y_{nq}z_{g(n)} \\ z_{nq}z_{g(n)} \\ x_{nq} \\ y_{nq} \\ z_{nq} \end{bmatrix}. \tag{9.42}$$

The gradient of the inequality constraint 9.41 with respect to the optimization vector $\mathbf{w}_q$ is $-\frac{m^2}{M^2}\nabla_{\mathbf{w}_q} c_{nq}^5$.

The boundary constraints are not necessary for phantom pose estimation, because the magnification constraints will enforce them automatically.

9.4.4 Implementation Details

The pose estimation algorithm was implemented in MATLAB [83] using "fmincon", a function in the Optimization Toolbox. For each estimation problem, two algorithms were used: interior-points and sequential quadratic programming (sqp) methods, and the results obtained from the algorithm with the smaller objective value was used.

The tunnel we considered has 10 $\cap$ columns. Each $\cap$ has 12 rows. The bottom three rows on each side of the $\cap$ are nearly vertically oriented, parallel to the z-axis. The 4th and 9th rows are oriented nearly at 45 degrees to the vertical line, while the 5th, 6th, 7th and 8th rows are nearly horizontally oriented, parallel to the y-axis. Each $\cap$ lies almost completely on a yz-plane (fixed x-coordinate). The $\cap$s are almost evenly laid out along the x-axis, except for the first two, whose separation is almost twice that of the other consecutive $\cap$s.

The detector boards' poses were estimated in multiple waves. First, assuming the phantom grid point locations are known, the detector poses are estimated. We call this "run 1."Next, the detector pose estimates are used to estimate the pose of the phantom placements. Following this, the detector boards' poses were re-estimated using the estimated phantom poses. This is called "run 2."The process of alternating between the estimation of the pose of the phantom placements and those of the detector boards can be repeated multiple times resulting in further runs, "run r,"where $r = 1, \cdots, R$.

The optimization problem is a nonlinear optimization and needs an initial parameter estimate. To initialize the first wave of detector pose estimation, we assumed the boards were aligned perfectly along the sides of the tunnel. To be specific, for a board d, its row r_d and column ($\cap$) indexes c_d are

$$r_d = \mod (d-1, \text{boards per column}) + 1, \text{ and } c_d = \left\lfloor \frac{d-1}{\text{boards per column}} \right\rfloor + 1.$$

The initial position (the location of the first pixel on the board) of detector board d is given as

$$\mathbf{P}(:, r_d) + \begin{bmatrix} x_{\text{offset}}(c_d) \\ 0 \\ 0 \end{bmatrix},$$

where $\mathbf{P}$ is the initial position matrix for the first $\cap$ and $\mathbf{P}(:, r_d)$ uses MATLAB's [83] notation for the r_d-th column of the matrix $\mathbf{P}$. x_{offset} gives the offset of each $\cap$ from the first $\cap$. One of our collaborators at Duke University, Andrew Holmgren, measured the first and last $\cap$s to be at approximately $x = 714$mm and $x = 1859$mm, respectively. Using these measurements, the matrix $\mathbf{P}$ is given as (in mm)

$$\begin{bmatrix} 714 & 714 & 714 & 714 & 714 & 714 & 714 & 714 & 714 & 714 & 714 & 714 \\ -278 & -278 & -278 & -278 & -205 & -103 & 0 & 103 & 205 & 278 & 278 & 278 \\ 500 & 603 & 705 & 808 & 880 & 880 & 880 & 880 & 880 & 808 & 705 & 603 \end{bmatrix}.$$

The offset vector in the x-direction is approximately given as (in mm)

$$[0, 200, 318, 436, 554, 673, 791, 909, 1027, 1145].$$

The initial orientation of board d is given as $\mathbf{O}(:, r_d)$, where $\mathbf{O}$ is given as

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} & 1 & 1 & 1 & 1 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 1 & 1 & 1 & \frac{1}{\sqrt{2}} & 0 & 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & -1 & -1 & -1 \end{bmatrix}.$$

To initialize the phantom pose estimation, we used the originally measured position and orientation of the phantom. For the second wave of detector pose estimation, the parameters that were estimated from the first wave were used to initialize the optimization algorithms.

For the optimization constraints, we chose a different set of magnification constants and tunnel bounds for each detector board. The tunnel bounds were based on the initial position of each detector board, with a window of 50 mm in each direction. The magnification constants m and M were based on the current estimate of the detector position and orientation, with a tolerance of 1% above and below the maximum and minimum magnifications, respectively. The minimum magnification was never less than 1, to prevent the detector board's position from lying between the source and the phantom. For each run (wave), the tethering direction was chosen to match the initial orientation of the detector boards, with tether angle $\theta_t = 2°$.

For the phantom placement parameter estimation, the magnification constraints can either be enforced directly or indirectly. If the constraints are included in the optimization problem, then they are directly enforced. We have found that including them increases the estimation times significantly, since the constraint is checked for every measured phantom point. Moreover, because a lot of detector boards surrounding the phantom are used in estimating its parameters, the magnification constraints are indirectly enforced, due to the different perspectives.

If the location of the phantom grid points are accurately measured, then the estimation of the phantom parameters can be omitted. Moreover, if the number of detector boards surrounding each phantom placement is not sufficient to estimate the phantom parameters, then re-estimating the phantom parameters in each run could lead to incorrect estimates of the detector pose. Indeed, this is the case for our calibration setup.

9.4.5 Results

In the remainder of this section, we consider the results of using the dot product objective to estimate the pose of the detector boards, with the location of the phantom grid points fixed at their measured values. Dr. Kenneth MacCabe performed the calibration experiments, measuring the location of the phantom grid points and estimating the centroids of the images of the grid points. The initial positions and orientations of the detector boards are given in Figure 9.2 while the results from the 1^{st} and 20^{th} waves of detector pose estimation, using the dot product objective function are given in Figures 9.3 and 9.4. Figure 9.5 shows the average value of the normalized dot product (averaged over all the phantom-detector pixel pairs across all detector boards) for each run of the detector pose estimation. From the figure, we see that the objective function increases with each run, suggesting that the source to phantom grid point vectors and the source to detector pixel vectors are better aligned with each run. The reason for the change in the objective with each run in this setting is the magnification parameters m and M, which vary with the current estimate of the detector pose. The objective function effectively saturates after about 20 runs.

9.5 AM Algorithm

The derivation of the AM algorithm and a pseudocode for the algorithm are given in detail in the paper by O'Sullivan and Benac [92]. The pseudocode is repeated here in Pseudocode 9.1, for readability. In the pseudocode, $\hat{q}(\gamma, E)$ (an estimate of $q(\gamma, E)$) is the summand of $g(\gamma : \mu)$ as shown in equation 9.3 above. $\hat{p}(\gamma, E)$ (an estimate of $p(\gamma, E)$) is a hidden-variable and summand of $d(\gamma)$, and is interpreted as the mean data for source-detector pair γ that has energy E. A hat is used to denote estimates of a parameter. $\mathbf{p}$ and $\mathbf{q}$ are the collection of the values of $p(\gamma, E)$ and $q(\gamma, E)$ for all (γ, E) pairs, respectively.

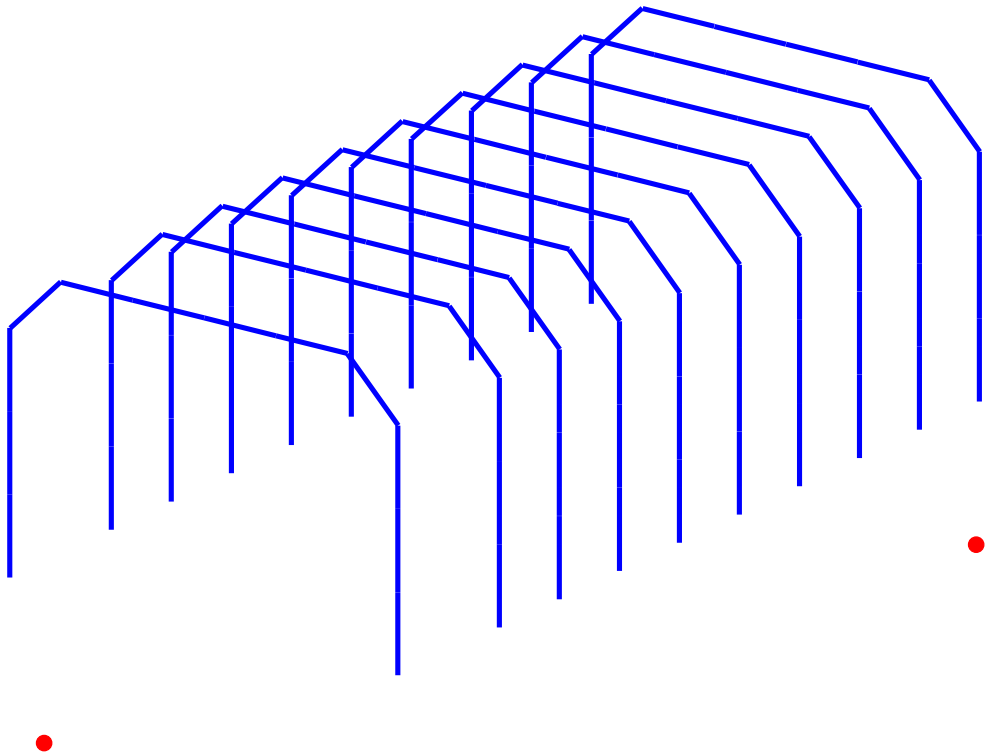


Figure 9.2: Initial detector layout.

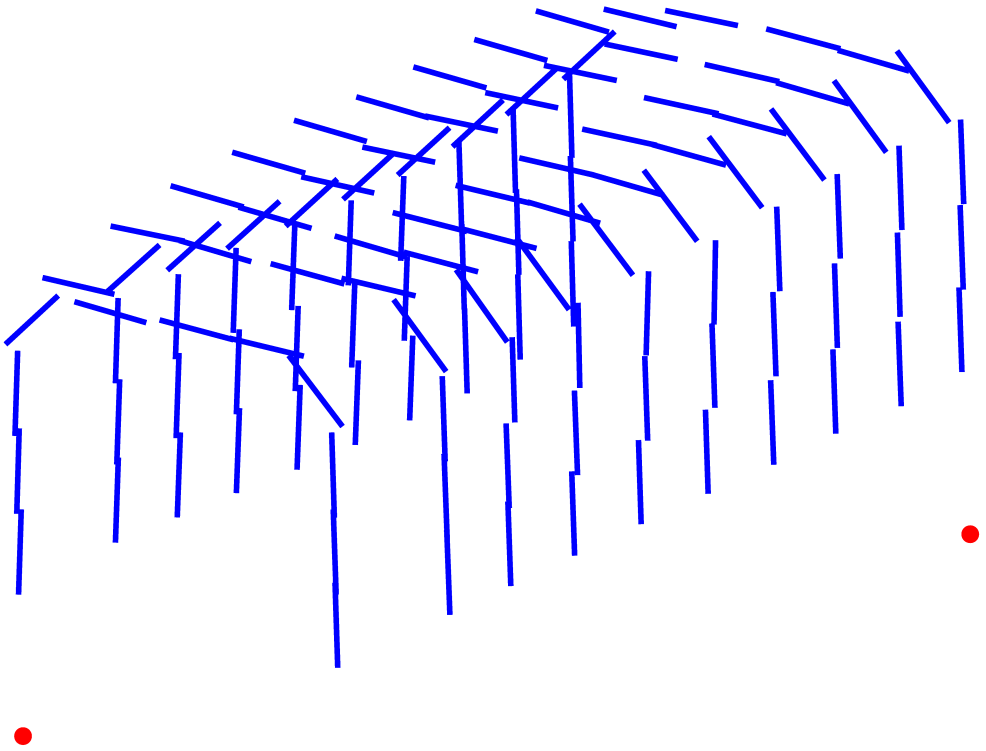


Figure 9.3: First wave detector layout.

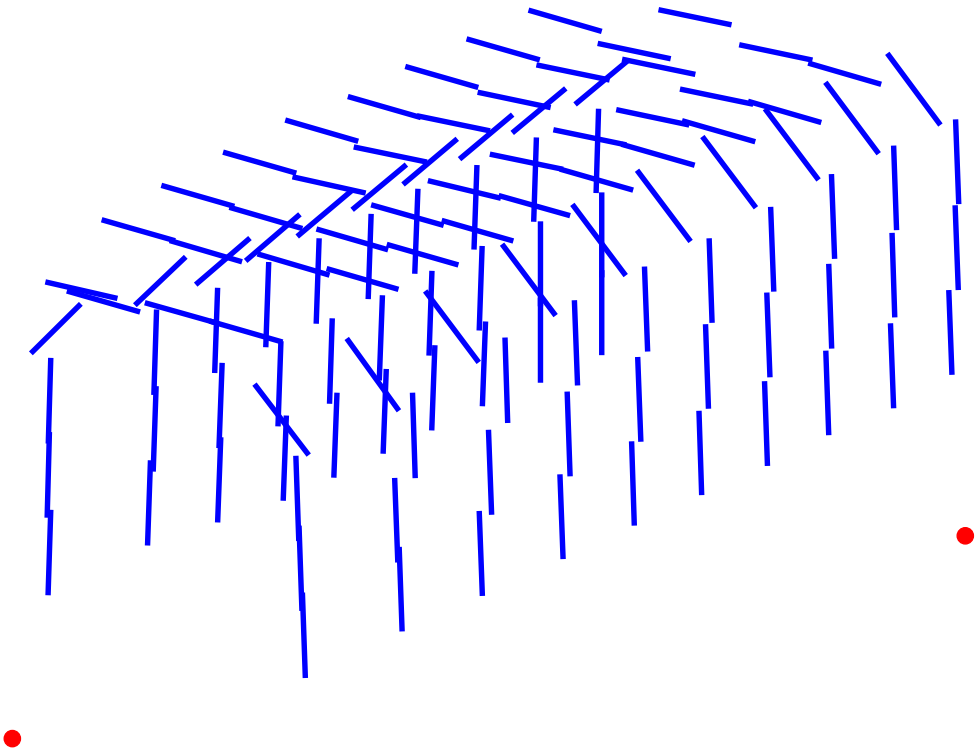


Figure 9.4: Twentieth wave detector layout.

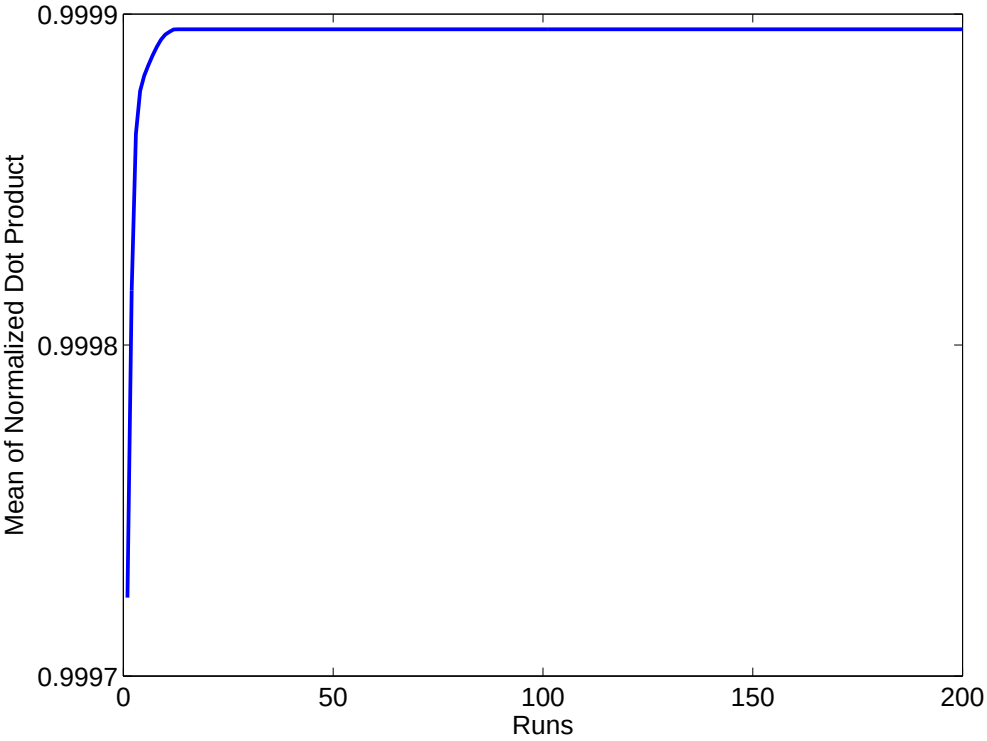


Figure 9.5: Mean of normalized dot product figure of merit.

Pseudocode 9.1 Unregularized polyenergetic AM algorithm. The hat represents estimates of the parameter. $q(\gamma, E)$ is the summand of $g(\gamma : \mu)$. $p(\gamma, E)$ is the hidden-variable summand of $d(\gamma)$. $\mathbf{p}$ and $\mathbf{q}$ are the collection of the values of $p(\gamma, E)$ and $q(\gamma, E)$ for all (γ, E), respectively.

Precompute the normalization factor $Z_i(\chi)$

$$Z_i(\chi) = \max_{(\gamma,E)} \sum_{\chi \in \mathrm{X}} \sum_{i=1}^{I} \mu_i(E) h(\gamma|\chi). \tag{9.43}$$

Set $k = 0$. Set initial condition for each $c_i^{(0)}(\chi)$.
Compute the current estimate of $q(\gamma, E)$:

$$\hat{q}^{(k)}(\gamma, E) = I_0(\gamma, E) \exp\left[- \sum_{\chi \in \mathrm{X}} \sum_{i=1}^{I} \mu_i(E) h(\gamma|\chi) \hat{c}_i^{(k)}(\chi) \right]. \tag{9.44}$$

Compute the current estimate of $p(\gamma, E)$:

$$\hat{p}^{(k)}(\gamma, E) = \hat{q}^{(k)}(\gamma, E) \frac{d(\gamma)}{\sum_{E'} \hat{q}^{(k)}(\gamma, E') + \beta(\gamma)}. \tag{9.45}$$

Compute backprojection of the current estimates of $\mathbf{p}$ and $\mathbf{q}$:

$$\tilde{b}_i^{(k)}(\chi) = \sum_{\gamma \in \mathrm{Y}} \sum_{E} \mu_i(E) h(\gamma|\chi) \hat{p}^{(k)}(\gamma, E), \tag{9.46}$$

$$\hat{b}_i^{(k)}(\chi) = \sum_{\gamma \in \mathrm{Y}} \sum_{E} \mu_i(E) h(\gamma|\chi) \hat{q}^{(k)}(\gamma, E). \tag{9.47}$$

Update the estimate of the concentrations:

$$\hat{c}_i^{(k+1)}(\chi) = \hat{c}_i^{(k)}(\chi) - \frac{1}{Z_i(\chi)} \ln\left(\frac{\tilde{b}_i^{(k)}(\chi)}{\hat{b}_i^{(k)}(\chi)} \right). \tag{9.48}$$

if $\hat{c}_i^{(k+1)}(\chi) < 0$ **then**
  $\hat{c}_i^{(k+1)}(\chi) = 0$.
end if
Set $k = k + 1$.
Check for convergence, iterate if necessary

The introduction of a penalty term for regularizing the estimates of the concentration leads to the loss of the closed-form image update as in equation 9.48. Green [42] proposed a one-step-late (OSL) approach to solve this problem. However, the OSL approach does not guarantee convergence for large values of the regularization coefficient. An alternative to the OSL approach is an iterative approach to solving the gradient equation. This is the approach followed by most researchers [61].

Keesing [61] discusses an extension of the monoenergetic version of the algorithm to a regularized case. Keesing [61] utilized a Lange [71] edge-preserving potential function and derived a surrogate regularization term to allow for independent pixel updates. The Lange [71] edge-preserving potential function is

$$\psi(t) = \delta^2 \left[\left| \frac{t}{\delta} \right| - \ln \left(1 + \left| \frac{t}{\delta} \right| \right) \right], \tag{9.49}$$

where δ is a scaling parameter that determines what level of the image gradient (difference between neighboring voxels) should be used for noise reduction. Roughly speaking, only image gradient values above δ are considered to be from a signal and are preserved.

Erdoğan and Fessler [31] utilize the same potential function but a different decoupling technique to obtain a surrogate regularization term. In this chapter, we follow the approach of Keesing [61] for the regularized polyenergetic AM algorithm.

The regularization term used in penalizing the Poisson log-likelihood (or the I-divergence) is given as

$$R(\mathbf{c}) = \sum_{i=1}^{I} \sum_{\chi \in \mathrm{X}} \sum_{s \in \mathrm{N}_x} w_{\chi s} \psi(c_i(\chi) - c_i(s)), \tag{9.50}$$

where N_x is the set of neighboring voxels to χ and $w_{\chi s}$ is the weight assigned to the neighbor's contribution.

Using the method based on the convexity property of the potential function as described by Keesing [61], the decoupled regularization term is given as

$$\hat{R}(\mathbf{c}) = \sum_{i=1}^{I} \sum_{\chi \in \mathrm{X}} \sum_{s \in \mathrm{N}_x} w_{\chi s} \frac{1}{2} \left[\psi(2c_i(\chi) - \hat{c}_i(\chi) - \hat{c}_i(s)) + \psi(2c_i(s) - \hat{c}_i(\chi) - \hat{c}_i(s)) \right], \tag{9.51}$$

where $\hat{c}_i(\chi)$ and $\hat{c}_i(s)$ are previous estimates of the image at voxel χ and s respectively, for material i.

Pseudocode 9.2 shows the regularized polyenergetic AM algorithm using the decoupled regularization term $\hat{R}(\mathbf{c})$. λ is the regularization coefficient and determines the relative weight assigned to the regularization term in comparison to the data fit term (Poisson log-likelihood or I-divergence).

Newton's method is a popular choice [31, 61] for solving equation 9.53, since it has quadratic convergence to the root, when the initial guess is close to the solution. When δ approaches 0, for a fixed value of λ, the edge-preserving regularization term approaches a total variation penalty (absolute value of the difference between a voxel's and its neighbor's image value). The steepness of the absolute value function is directly proportional to $\frac{\lambda}{\delta}$ [42]. As δ goes to ∞, the edge-preserving regularization term approaches a Huber-type regularizer based on a quadratic potential function [42]. This regularizer does not preserve any edge, but rather smooths all image gradients.

For relatively small values of δ, as the value of λ increases, the steepness of the absolute value function increases. This steepness presents a problem to the standard Newton iterates. For a reasonable value of λ, the Newton iterates do not converge to the solution of the gradient equation, but oscillate ad infinitum. This observation was made by my advisor, Dr. O'Sullivan.

One of the students in our lab, Soysal Degirmenci, proposed a line search approach to solving this convergence problem, while I proposed an alternative solution based on a trust region approach [115]. This chapter considers only the trust region approach, since it is thought to be more robust than line search methods [14].

9.5.1 Trust Region Newton Method

The decoupled objective (data fit term and penalty term) leads to a 1-dimensional (1D) gradient equation given in equation 9.53. The solution of this equation can be obtained by a modification of Newton's method, using a trust region approach. The approach was described by Sorensen [115] for a multidimensional problem and has been adapted and optimized for a 1D setting.

Pseudocode 9.2 Regularized polyenergetic AM algorithm. The hat represents estimates of the parameter. $q(\gamma, E)$ is the summand of $g(\gamma : \mu)$. $p(\gamma, E)$ is the hidden-variable summand of $d(\gamma)$. $\mathbf{p}$ and $\mathbf{q}$ are the collection of the values of $p(\gamma, E)$ and $q(\gamma, E)$ for all (γ, E), respectively. N_x is the set of neighboring voxels to χ and $w_{\chi s}$ is the amount of weight assigned to the neighbor's contribution. λ is the regularization coefficient.

Choose values for λ and δ

Precompute the normalization factor $Z_i(\chi)$

$$Z_i(\chi) = \max_{(\gamma,E)} \sum_{\chi \in \mathrm{X}} \sum_{i=1}^{I} \mu_i(E) h(\gamma|\chi). \tag{9.52}$$

Set $k = 0$. Set initial condition for each $c_i^{(0)}(\chi)$.

Compute the current estimate of $q(\gamma, E)$:

$$\hat{q}^{(k)}(\gamma, E) = I_0(\gamma, E) \exp\left[- \sum_{\chi \in \mathrm{X}} \sum_{i=1}^{I} \mu_i(E) h(\gamma|\chi) \hat{c}_i^{(k)}(\chi) \right].$$

Compute the current estimate of $p(\gamma, E)$:

$$\hat{p}^{(k)}(\gamma, E) = \hat{q}^{(k)}(\gamma, E) \frac{d(\gamma)}{\sum_{E'} \hat{q}^{(k)}(\gamma, E') + \beta(\gamma)}.$$

Compute backprojection of the current estimates of $\mathbf{p}$ and $\mathbf{q}$:

$$\begin{aligned} \tilde{b}_i^{(k)}(\chi) &= \sum_{\gamma \in \mathrm{Y}} \sum_{E} \mu_i(E) h(\gamma|\chi) \hat{p}^{(k)}(\gamma, E), \\ \hat{b}_i^{(k)}(\chi) &= \sum_{\gamma \in \mathrm{Y}} \sum_{E} \mu_i(E) h(\gamma|\chi) \hat{q}^{(k)}(\gamma, E). \end{aligned}$$

Update the estimate of the concentrations by solving the equation:

$$\tilde{b}_i^{(k)}(\chi) - \hat{b}_i^{(k)}(\chi) \exp\left(Z_i(\chi) \left[\hat{c}_i^{(k)}(\chi) - c_i(\chi) \right] \right) + \lambda \sum_{s \in \mathrm{N}_x} w_{\chi s} \left. \frac{\partial \psi(t)}{\partial t} \right|_{t = 2c_i(\chi) - \hat{c}_i^{(k)}(\chi) - \hat{c}_i^{(k)}(s)} = 0. \tag{9.53}$$

if The solution $c_i(\chi) = \hat{c}_i^{(k+1)}(\chi) < 0$ **then**

$\hat{c}_i^{(k+1)}(\chi) = 0$.

end if

Set $k = k + 1$.

Check for convergence, iterate if necessary

The idea of the trust region method is to substitute an unconstrained minimization of a quadratic surrogate function at each Newton iteration with a constrained minimization. If the minimum of the quadratic function falls outside the interval (trust region), then the closest interval endpoint is used. In addition, if the current minimum of the constrained problem does not guarantee a reduction of the overall objective function, the interval is shrunk. That is, small steps will be taken in subsequent steps. On the other hand, if the current minimum leads to a big enough reduction in the objective, the interval will be expanded, approaching a regular Newton iteration; a full Newton step may be taken in the next step.

Pseudocode 9.3 gives the procedure for the trust-region modified Newton method for finding the root of an equation $g(c) = 0$, where $f(c)$ is the integral of $g(c)$ with respect to c, and $h(c)$ is the derivative of $g(c)$ with respect to c.

This 1D trust region modified Newton's method adapts the multidimensional version of the algorithm proposed by Sorensen [115] and further optimizes the step that shrinks the radius of the trust region to avoid wasting iterations. It does this by setting the radius of the trust region to the absolute value of the Newton step, whenever the standard Newton step is within the trust region, but does not provide the desired decrease in the function $f(c)$ or desired increase in proximity to the zero of $g(c)$. This prevents the Newton step from being valid in the next iteration, since it will be outside the shrunk trust region.

9.6 Reconstruction Results

We implemented the polyenergetic and monoenergetic versions of the regularized AM algorithms for reconstructing attenuation objects in a C/C++ software package dubbed "CAXI," for Coded Aperture X-ray Imaging and the results presented in this section were obtained by using the software. The software package can model any arbitrary scanner geometry including the standard clinical CT cylindrical geometry, and any type of detector, including photon-counting, energy-sensitive, and energy-integrating detectors. The results presented here are for the prototype X-ray machine which is fitted with energy-integrating detectors.

We consider reconstructing simulated and real attenuation data using both the monoenergetic [92] and polyenergetic AM algorithms. The results presented here for the monoenergetic AM algorithm do not account for beam hardening. However, the software package has the

Pseudocode 9.3 Trust region modified Newton's method for finding the root of an equation $g(c) = 0$. $f(c)$ is the integral of $g(c)$ with respect to c and $h(c)$ is the derivative of $g(c)$ with respect to c. $\hat{c}$ is an estimate of the root of the equation.

```
Choose values for η1, η2, γ1, and γ2 such that 0 < η1 < η2 < 1 and 0 < γ1 < 1 < γ2
Choose convergence threshold ε
Choose radius of trust region Δ > ε
Set initial value for ĉ.
Compute fun_val = f(ĉ), grad = g(ĉ) , and hess = h(ĉ)
Set newton_update = -grad/hess
Set w = newton_update
while |newton_update| > ε and Δ > ε do
  % Constrain the minimum of the quadratic surrogate
  if w > Δ then
    w = Δ
  else if w < −Δ then
    w = −Δ
  end if
  Compute new_fun_val = f(ĉ + w)
  Compute actual_reduction = fun_val - new_fun_val
  Compute predicted_reduction = −grad × w − 0.5 × hess × w^2
  Compute ratio = actual_reduction / predicted_reduction
  if ratio < η1 then
    % Shrink the radius of the trust region
    if |w| < Δ then % Efficient modification of the original trust region method
      Δ = |w|
    end if
    Δ = γ1Δ
  else
    % Update the estimate of the root
    ĉ = ĉ + w
    Compute fun_val = f(ĉ), grad = g(ĉ) , and hess = h(ĉ)
    Set newton_update = -grad/hess
    Set w = newton_update
    if ratio > η2 then % Expand the radius of the trust region
      Δ = γ2Δ
    end if
  end if
end while
```

methods in place for incorporating beam-hardening corrections and experiments are underway to include it in future reconstructions.

A few of the detector pixels and detector boards may not provide usable data due to excessive noise. These pixels and boards were first detected prior to reconstructing an object from the real data. Using the dark current measurements (with the sources turned off), we can determine which detector pixels or boards are malfunctioning, so that their measurements can be ignored in the reconstruction. Given a large number of frames, we can compute the mean of the dark current for each pixel and the overall mean and standard deviation. Anomaly detection, assuming a Gaussian distribution for the dark current measurements, follows by eliminating the pixels whose means are greater than 2 standard deviations from the overall mean.

The detector layout obtained after the 20^{th} wave of detector position and orientation parameter estimation was used in obtaining the point spread function $h(\gamma|\chi)$ which plays a major role in object reconstruction. Projection data were obtained in a cone-beam geometry with 120 linear energy-integrating detector boards (each with 128 pixels) and the two 170 kVp sources turned on at different times. The detectors had an integration time of 4 ms. The first source was switched on for the first 12 seconds (giving 3000 frames) while the second source was turned on, with the first switched off, for the remaining 12 seconds. Figure 9.6 shows the spectrum of the 170 kVp source.

9.6.1 Simulated Data

A 3D Shepp-Logan [108] phantom with $128 \times 128 \times 128$ voxels and slices along the z direction was simulated by using the parameters of the prototype X-ray attenuation machine. The polyenergetic 170 kVp source was modeled with a discrete energy spectrum with energies from 20 keV to 170 keV in steps of 10 keV. The attenuation of the phantom object is obtained by using a single material, namely, water, with known attenuation coefficients for the different incident photon energies. The values from the 3D Shepp-Logan phantom range between 0 and 1, and represent the concentration of water at each voxel. The 3D images were reconstructed from the mean data (without Poisson noise). For the monoenergetic AM algorithm to match the mean data, the mean spectral energy of 69.68 keV was utilized for the reconstruction. The polyenergetic reconstruction uses the same source energies that were used in simulating the projection data. The images shown here and for the real data are based on the attenuation coefficient of water at 69.68 keV and are displayed in a [0, 0.026]

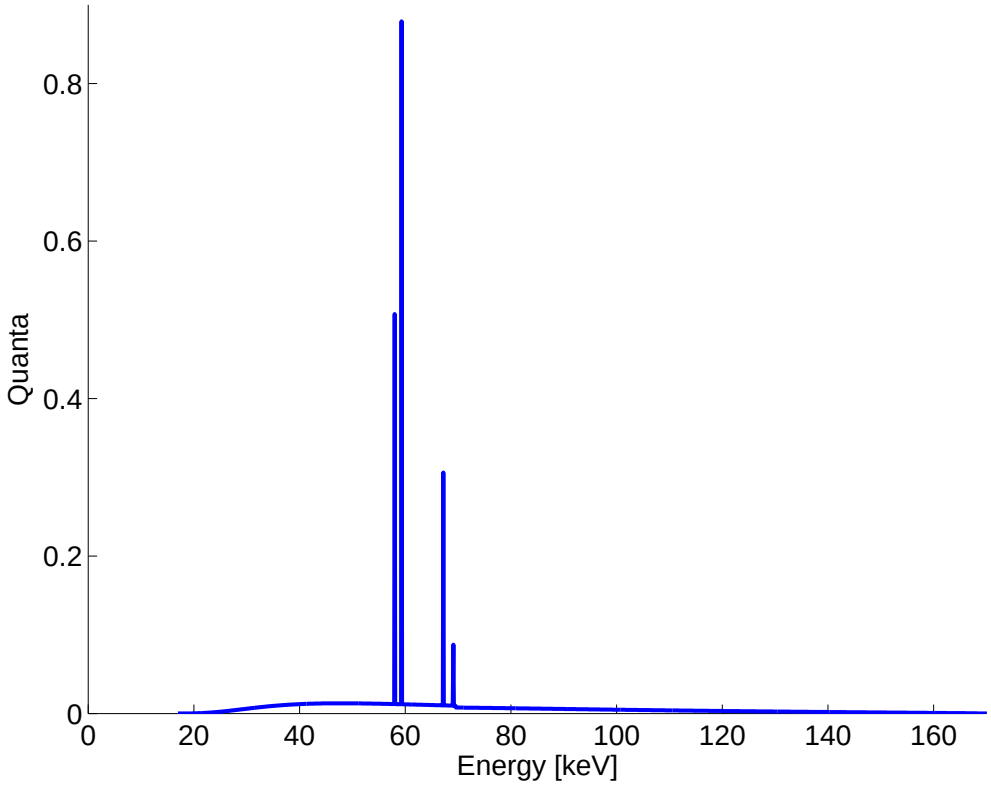


Figure 9.6: Energy spectrum for 170 kVp source.

mm^{-1} viewing window. An object volume of $128 \times 128 \times 128$ voxels with a resolution of 3 mm in each direction was simulated and reconstructed.

The simulated projection data is shown in Figure 9.7. Figures 9.8b and 9.8c show the results after 1000 iterations of the monoenergetic and polyenergetic AM algorithms, respectively. The regularization coefficient was $\lambda = 100$ for both AM algorithms. The regularization scale parameter was set at $\delta = 0.0002\text{mm}^{-1}$ and $\delta = 0.0104$ for the monoenergetic and polyenergetic AM algorithms, respectively. The regularization scale parameter for the polyenergetic AM algorithm was obtained from that of the monoenergetic AM algorithm, by dividing it by the attenuation coefficient of the reference material (water) at the mean spectral energy. The truth and reconstructed images shown in Figure 9.8 are for the 64^{th} slice along the z direction. For this slice, a profile along the x direction at the 64^{th} y pixel is given in Figure 9.9.

We can see that the reconstructions from the monoenergetic and polyenergetic AM algorithms are visually similar and suffer from aliasing artifacts, due to an under-sampling of view angles. However, the profiles in Figures 9.9b and 9.9c show that the monoenergetic and polyenergetic AM algorithms follow the original profile of the phantom, albeit noisily. Given

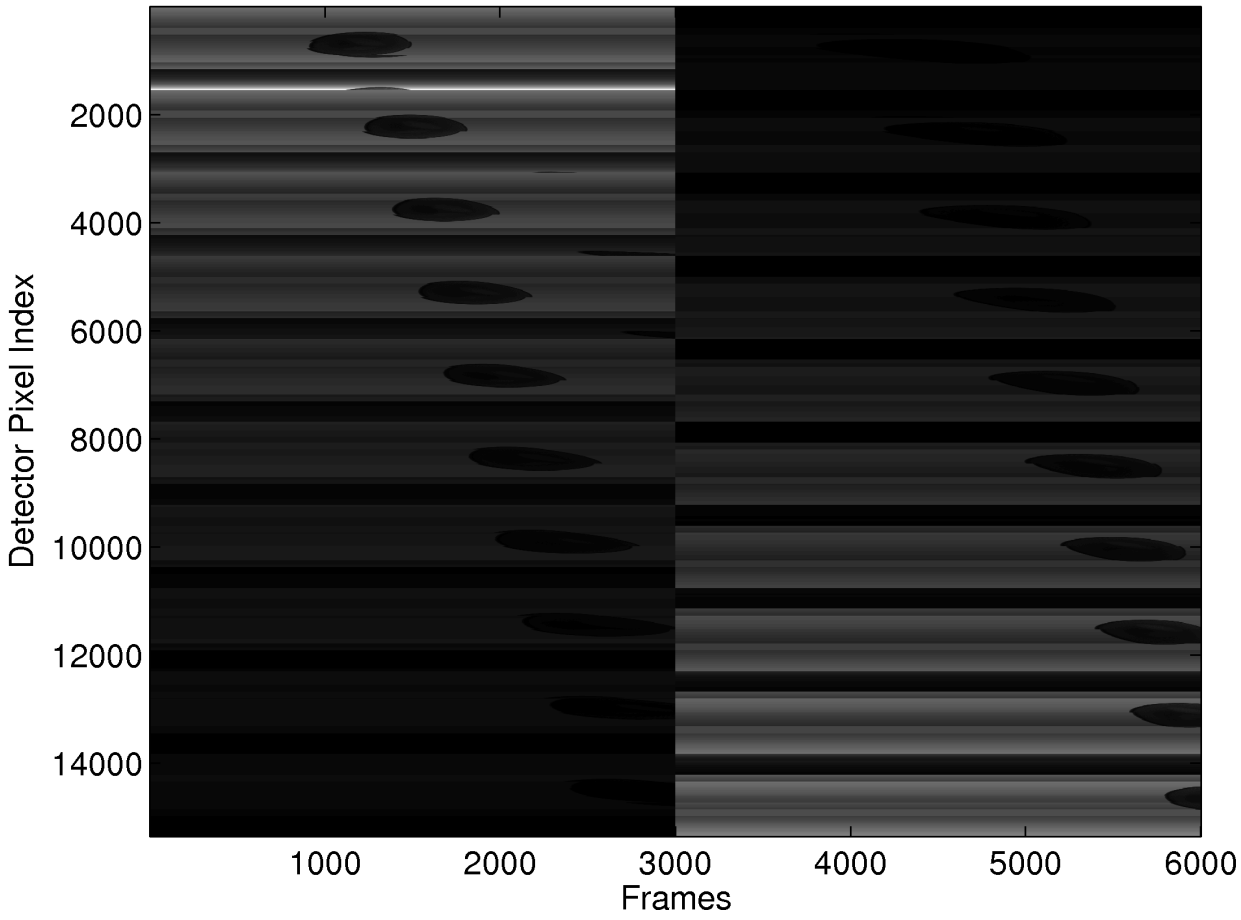
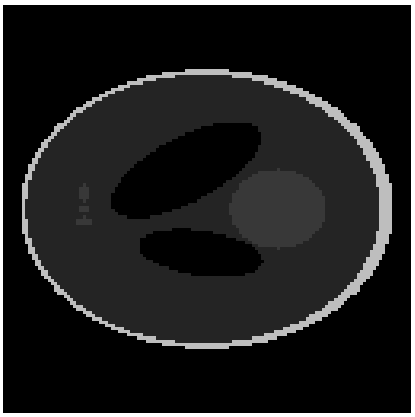
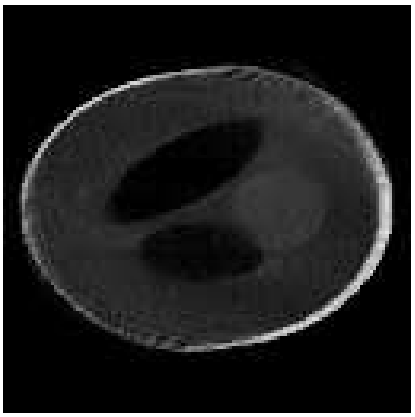


Figure 9.7: Simulated projection data for 3D Shepp-Logan phantom.



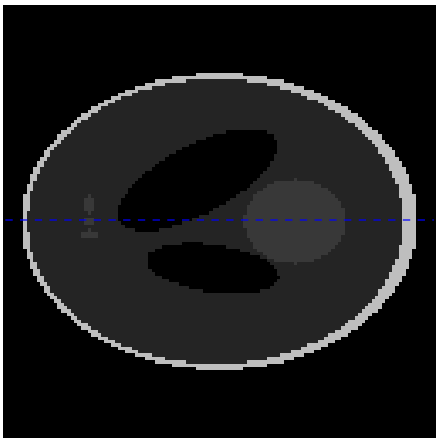
(a)

(b)

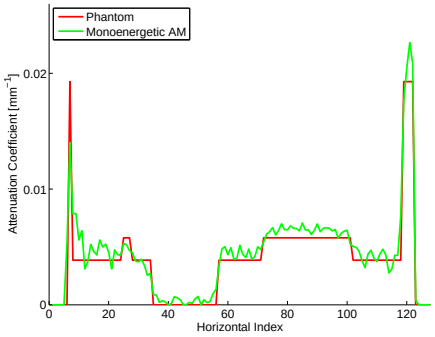


(c)

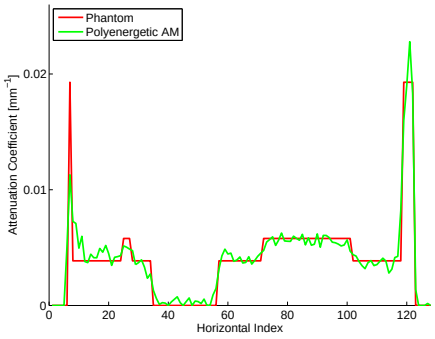
Figure 9.8: Reconstruction of noiseless 3D Shepp-Logan phantom data, 64^{th} z slice. The results for the monoenergetic and polyenergetic AM algorithms were obtained after 1000 iterations. (a) Phantom. (b) Monoenergetic AM. (c) Polyenergetic AM.



(a)



(b)



(c)

Figure 9.9: Reconstruction of noiseless 3D Shepp-Logan phantom data, profile along the 64^{th} y pixel of the 64^{th} z slice. The results for the monoenergetic and polyenergetic AM algorithms were obtained after 1000 iterations. (a) Phantom slice with profile line. (b) Monoenergetic AM profile with phantom profile. (c) Polyenergetic AM profile with phantom profile.

that the monoenergetic AM algorithm is much faster than the polyenergetic AM algorithm, and they provide similar reconstructions, the monoenergetic AM algorithm is favored in this setting.

Based on such simulation studies, we can modify the prototype system by including additional detector boards along the tunnel walls or adding other sources at strategic positions that would allow a more complete measurement of the views through the scanned object, which will lead to better attenuation reconstructions.

9.6.2 Real Data

The data from the prototype X-ray machine used here was acquired by Mehadi Hassan and Andrew Holmgren at Duke University. A briefcase containing a few objects including a cellphone was scanned through the prototype X-ray attenuation machine and the projection data shown in Figure 9.10 were obtained. For the reconstructions, again, the material of choice was water. The same source energies that were used in the reconstruction of the simulated data were also used for reconstructing the real data. The images shown here are based on the attenuation coefficient of water at 69.68 keV and are displayed in a [0, 0.024] mm^{-1} viewing window. A $448mm \times 448mm \times 448mm$ object volume was reconstructed using $128 \times 128 \times 128$ voxels.

The results of reconstructing briefcase data from the prototype X-ray attenuation system are given in Figure 9.11. The reconstructed images shown in the figure are for the 64^{th} slice along the z direction. The results shown in the figure were obtained from 120 iterations of the monoenergetic and polyenergetic AM algorithms.

Aliasing indicative of a shortage of view angles can be observed in Figure 9.11. A few streaks in the image can be attributed to detector board misalignments. These artifacts will be removed when the prototype system is machined and the detector and source locations are accurately characterized.

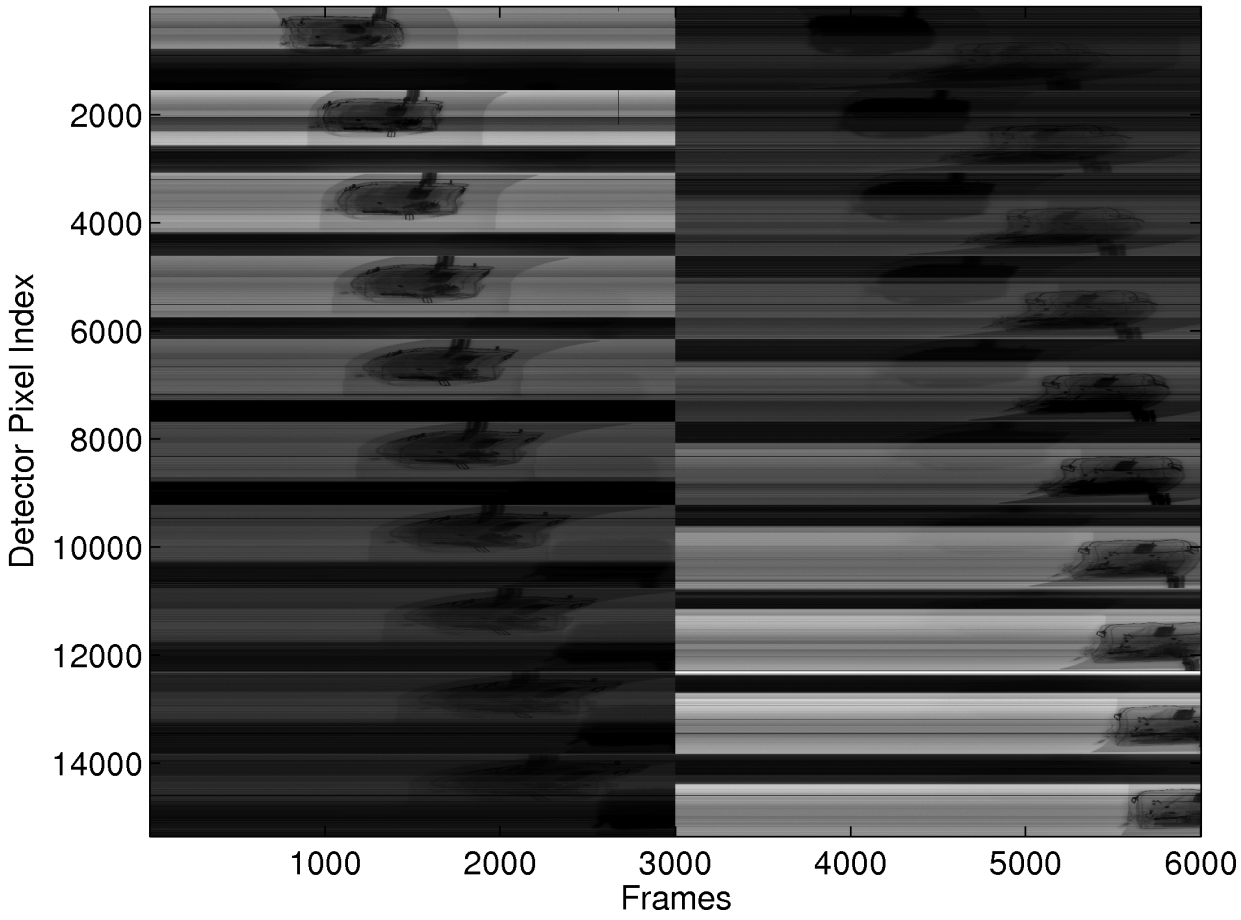
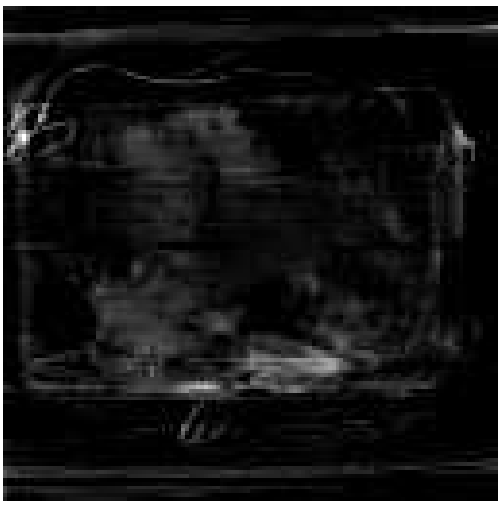
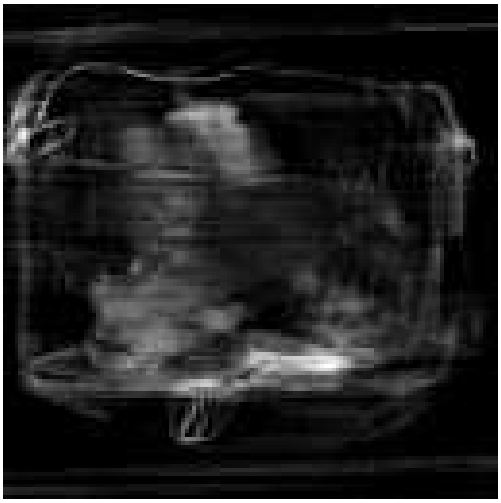


Figure 9.10: Real projection data for briefcase with an embedded cellphone.



(a)



(b)

Figure 9.11: Reconstruction of briefcase data. The results for the monoenergetic and polyenergetic AM algorithms were obtained after 120 iterations. The 2D images slices are taken from the 64^{th} object slice along the z direction. (a) Monoenergetic AM. (b) Polyenergetic AM.

9.7 Summary

In this chapter, we described a prototype X-ray system that was developed to inform material identification, by estimating an attenuation map that can be used by the scatter system to better estimate the scattering density of an object.

We restated the resulting AM algorithm proposed by O'Sullivan and Benac [92], adding a pseudocode for the regularized case.

We discussed a framework for estimating the position and orientation of a set of linear detector boards laid out on the walls of a tunnel, based on a constrained nonlinear optimization problem. This approach allows for side information such as known detector positions, orientations, magnification constraints, or bounds on detector positions to be easily incorporated.

We introduced the use of a trust region method to guarantee the convergence of a 1D Newton iterate for a decoupled regularized alternating minimization algorithm, thus, solving the convergence problem of previous methods that utilize a neighborhood-based edge-preserving penalty [31, 61].

Finally, we presented results on reconstructing simulated data from a 3D Shepp-Logan phantom and real data from a briefcase with an embedded cellphone, among other objects, when the two sources are switched on alternately. Reconstruction results based on monoenergetic and polyenergetic AM algorithms were presented. The current monoenergetic reconstruction results do not utilize a beam-hardening correction. Including this correction may lead to better reconstructed images.

9.8 Lessons Learned

The problem of estimating the position and orientation (pose) of a linear detector array using multiple sources is highly nonlinear. A good initial condition in the form of extra information about the orientation or the position of the detector array helps to accurately estimate the pose of the detector array.

Edge-preserving regularized transmission reconstruction algorithms based on the standard Newton's method do not converge for very small values of the scaling parameter and relatively

large values of the regularization coefficient. Modified Newton's methods via either a trust region modification or a line search are necessary to guarantee convergence.

An AM algorithm derived for simple Poisson data can be utilized for compound Poisson data, if the mean data utilized by the AM algorithm is adjusted to match that of the compound Poisson data.

Chapter 10

Future Work

The biometric studies reported in this dissertation are for individuals sitting in a normal resting condition. As indicated in Chapters 2 and 3, there need to be more studies performed to understand the effects of mental and physical stressors on the ECG signal, and a biometric system based on it. The same applies to the LDV signals and biometric system. Moreover, the effects of body positions, e.g. standing versus sitting, on cardiovascular biometrics need to be fully understood to allow such systems to be deployed in the field.

The work on fusing cardiovascular biometric traits, namely, ECG and LDV signals discussed in Chapter 4, omitted fusion at the raw signal level. In addition, several score-level fusion schemes outperformed the feature-level fusion method developed in that chapter. Since more information about the individuals to be recognized is available at the feature level than at the score level, we expect that feature-level fusion should give better recognition performance than score-level fusion. This tells us that the feature-level fusion methodology we adopted in Chapter 4 is not optimal.

We are currently investigating an information theoretic approach to fusing ECG and LDV signals at the raw and feature levels of the biometric pipeline. The method is similar to that used by Fisher *et al.* [36] for fusing data from multiple sensors. The idea is to learn a weight vector or matrix for ECG raw signals/feature vectors and one for LDV and maximize the mutual information between the resulting ECG and LDV projections. This projection to a lower dimensional space followed by maximizing mutual information effectively learns a subspace where the two signals are maximally aligned. That is, we learn a common representation for the two cardiovascular related signals. The distribution of the resulting projections can be estimated non-parametrically or fit to a parametric family of distributions. Non-parametric density estimation has the advantage of lowering the bias induced by parametric estimation at the cost of an increase in computational complexity when learning the weights.

One of the problems encountered in fusing ECG and LDV signals either at the raw or feature level, which is not as pronounced in the application areas described by Fisher *et al.* [36], is over-fitting. In ECG and LDV biometrics, there is a limited amount of training data, e.g. a few seconds of ECG recording yields a few heartbeats. Learning the projection weights through non-parametric density estimation can lead to weights which maximize mutual information for the training set, but do not generalize well to testing data. Adding a regularization term to the mutual information objective function will help prevent this over-fitting. An alternative to over-fitting is to use parametric density estimation, which introduces a bias preventing the weights from "memorizing" the training data.

There is ongoing work in measuring the attenuation through blocks of reference materials such as polymethyl methacrylate (PMMA) to create a data set for beam-hardening correction. Also, there are plans for collecting more detector calibration data with multiple phantom placements which will further help the nonlinear optimization problem in triangulating the actual detector positions and orientations. With these corrections in place, the streaking artifact observed in the reconstructed data in Chapter 9 should be greatly diminished.

The main task of the X-ray scatter system is to classify materials as dangerous or not. It does this by first estimating the scattering properties of the materials. The models we described in Chapters 7 and 8 ignore the effects of Compton scattering and attenuation. To obtain accurate estimates of the scatter properties for a real bag, our forward model will need to account for these other effects and estimate them jointly. There is ongoing work to create and validate an efficient joint coherent scatter and attenuation forward model for low-angle scatter, where Compton scatter effects are minimal.

References

[1] A. Abaza and A. Ross. Quality based rank-level fusion in multibiometric systems. In *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems (BTAS)*, pages 1 – 6, Sep. 2009.

[2] F. Agrafioti, , F. M. Bui, and D. Hatzinakos. On supporting anonymity in a BAN biometric framework. In *Proceedings of the 16th international conference on Digital Signal Processing, Santorini, Greece*, pages 787 – 792, 2008.

[3] F. Agrafioti, F. M. Bui, and D. Hatzinakos. Medical biometrics: the perils of ignoring time dependency. In *Proceedings of the 3rd IEEE international conference on Biometrics: Theory, applications and systems, Washington, D.C., USA*, pages 358 – 363, 2009.

[4] F. Agrafioti and D. Hatzinakos. ECG Based Recognition Using Second Order Statistics. In *Proceedings of the Communication Networks and Services Research Conference 2008, Nova Scotia, Canada*, 2008.

[5] F. Agrafioti and D. Hatzinakos. ECG biometric analysis in cardiac irregularity conditions. *Signal, Image and Video Processing*, 3(4):1683 – 1706, 2008.

[6] F. Agrafioti and D. Hatzinakos. Fusion of ECG sources for human identification. In *The 3rd International Symposium on Communications, Control and Signal Processing (ISCCSP 2008), St. Julians, Malta*, 2008.

[7] F. Agrafioti and D. Hatzinakos. Signal validation for cardiac biometrics. In *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, pages 1734 – 1737, Mar. 2010.

[8] C. J. Ashbourn. *Biometrics: Advanced Identity Verification*. Springer-Verlag, London, 2000.

[9] L. Biel, O. Pettersson, L. Philipson, and P. Wide. ECG analysis: a new approach in human identification. In *Proceedings of the IEEE Instrumentation and Measurement Technology Conference*, volume 16, 1999.

[10] L. Biel, O. Pettersson, L. Philipson, and P. Wide. ECG analysis: a new approach in human identification. *IEEE Trans. Instrum. Meas.*, 50:808 – 812, 2001.

[11] A. R. M. Bolle, J. H. Connell, S. Pankanti, N. K. Ratha, and A. W. Senior. *Guide to Biometrics*. Springer-Verlag, New York, NY, 2003.

[12] O. Boumbarov, Y. Velchev, and S. Sokolov. ECG personal identification in subspaces using radial basis neural networks. In *IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, pages 446 – 451, Sep. 2009.

[13] D. J. Brady, D. L. Marks, K. P. MacCabe, and J. A. O'Sullivan. Coded apertures for x-ray scatter imaging. *Appl. Opt.*, 52(32):7745 – 7754, Nov. 2013.

[14] M. Braun. trustOptim: An R Package for Trust Region Optimization with Sparse Hessians, 2013.

[15] R. Cappelli, D. Maio, and D. Maltoni. Combining Fingerprint Classifiers. In *Proceedings of the First International Workshop on Multiple Classifier Systems, Cagliari, Italy*, pages 351 – 361, 2000.

[16] A.D.C. Chan, M. M. Hamdy, A. Badre, and V. Badee. Person identification using electrocardiograms. In *IEEE Canadian Conference on Electrical and Computer Engineering, Ottawa*, 2006.

[17] A.D.C. Chan, M. M. Hamdy, A. Badre, and V. Badee. Wavelet Distance Measure for Person identification using electrocardiograms. *IEEE Transactions on Instrumentation and Measurement*, 57(2):248 – 253, 2008.

[18] S. Chauhan, A.S. Arora, and A. Kaul. A survey of emerging biometric modalities. *Procedia Computer Science*, 2:213 – 218, 2010. Proceedings of the International Conference and Exhibition on Biometrics Technology.

[19] M. Chen, J. A. O'Sullivan, A. D. Kaplan, P-.H. Lai, E. J. Sirevaag, and J. W. Rohrbaugh. Biometrics with Physical Exercise Using Laser Doppler Vibrometry Measurements of the Carotid Pulse. In *The First IEEE International Conference on Biometrics, Identity and Security, Tampa, Florida, USA*, 2009.

[20] M. Chen, J. A. O'Sullivan, N. Singla, E. J. Sirevaag, S. D. Kristjansson, P-.H. Lai, A. D. Kaplan, and J. W. Rohrbaugh. Laser Doppler vibrometry measures of physiological function: evaluation of biometric capabilities. *IEEE Trans. on Info. For. Sec.*, 5(3):449 – 460, 2010.

[21] M. Chen, J. A. O'Sullivan, N. Singla, E. J. Sirevaag, and J. W. Rohrbaugh. Laser Doppler vibrometry measures of physiological function: evaluation of biometric capabilities. In *Proceedings of SPIE, Orlando, Florida, USA*, 2009.

[22] T. Chighvinadze and S. Pistorius. WEG21103: Multiple Projection Compton Scatter Tomography. *Medical Physics*, 38(6):3835 – 3835, 2011.

[23] B. J. Chirillo and S. Blaul. *Implementing Biometric Security*. Wiley Publishing Inc., Indianapolis, IN, 2003.

[24] C-.C. Chiu, C-.M. Chuang, and C-.Y. Hsu. A Novel Personal Identity Verification Approach Using a Discrete Wavelet Transform of the ECG Signal. In *The 2nd International Conference on Multimedia and Ubiquitous Engineering, Busan, Korea*, 2008.

[25] M. B. Conover. *Understanding Electrocardiography.* Mosby, St. Louis, MO, 8 edition, 2003.

[26] D. P. Coutinho, A. N. Fred, and M. A. T. Figueiredo. One-Lead ECG-based Personal Identification Using Ziv-Merhav Cross Parsing. In *20th International Conference on Pattern Recognition (ICPR)*, pages 3858 – 3861, Aug. 2010.

[27] T. M. Cover and J. A. Thomas. *Elements of Information Theory.* John Wiley and Sons, Inc., New York, NY., 1991.

[28] C. Cui, S. M. Jorgensen, D. R. Eaker, and E. L. Ritman. Direct three-dimensional coherently scattered x-ray microtomography. *Medical Physics*, 37(12):6317 – 6322, 2010.

[29] I.G. Damousis, D. Tzovaras, and E. Bekiaris. Unobtrusive Multimodal Biometric Authentication: The HUMABIO Project Concept. *EURASIP Journal on Advances in Signal Processing*, 2008.

[30] Smiths Detection. HI-SCAN 6040i, Mar. 2014.

[31] H. Erdoğan and J. A. Fessler. Ordered subsets algorithms for transmission tomography. *Phys. Med. Biol.*, 44(11):2835 – 2851, Nov. 1999.

[32] B.L. Evans, J.B. Martin, L.W. Burggraf, and M.C. Roggemann. Nondestructive inspection using Compton scatter tomography. In *IEEE Nuclear Science Symposium*, pages 386 – 390, Nov. 1997.

[33] S-.C. Fang and H-.L. Chan. Human identification by quantifying similarity and dissimilarity in electrocardiogram phase space. *Pattern Recognition*, 42:1824 – 1831, Sep. 2009.

[34] S. Z. Fatemian, F. Agrafioti, and D. Hatzinakos. HeartID: Cardiac biometric recognition. In *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, pages 1 – 5, Sep. 2010.

[35] S. Z. Fatemian and D. Hatzinakos. A new ECG feature extractor for biometric recognition. In *The 16th International Conference on Digital Signal Processing*, 2008.

[36] J. W. Fisher, M. J. Wainwright, E. B. Sudderth, and A. S. Willsky. Statistical and Information-Theoretic Methods for Self-Organization and Fusion of Multimodal, Networked Sensors. *International Journal of High Performance Computing Applications*, 16(3):337 – 353, 2002.

[37] K. Fukunaga. *Introduction to Statistical Pattern Recognition.* Academic Press, San Diego, CA, 1991.

[38] Y. Gahi, M. Lamrani, A. Zoglat, M. Guennoun, B. Kapralos, and K. El-Khatib. Biometric Identification System Based on Electrocardiogram Data. In *New Technologies, Mobility and Security, Tangier, Marrocco*, 2008.

[39] J. Gao, F. Agrafioti, H. Mohammadzade, and D. Hatzinakos. ECG for blind identity verification in distributed systems. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1916 – 1919, May 2011.

[40] N. Ghofrani and R. Bostani. Reliable features for an ECG-based biometric system. In *17th Iranian Conference on Biomedical Engineering (ICBME)*, pages 1 – 5, Nov. 2010.

[41] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, 101(23):e215 – e220, Jun. 2000.

[42] P. J. Green. Bayesian reconstructions from emission tomography data using a modified EM algorithm. *IEEE Trans. Med. Imag.*, pages 84 – 93, 1990.

[43] J. A. Greenberg, K. Krishnamurthy, M. Lakshmanan, K. P. MacCabe, S. Wolter, A. Kapadia, and D. J. Brady. Coding and sampling for compressive x-ray diffraction tomography. In *Proc. SPIE*, volume 8858, pages 885813–1 – 885813–11, 2013.

[44] M. Guennoun, N. Abbad, J. Talom, S.M.M. Rahman, and K. El-Khatib. Continuous authentication by electrocardiogram data. In *IEEE Toronto International Conference on Science and Technology for Humanity (TIC-STH)*, pages 40 – 42, Sep. 2009.

[45] G. Harding and J. Kosanetzky. Elastic scatter computed tomography. *Physics in Medicine and Biology*, 30(2):183, 1985.

[46] T. K. Ho, J.J. Hull, and S.N. Srihari. Decision combination in multiple classifier systems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(1):66 – 75, Jan. 1994.

[47] M. Homer, J. M. Irvine, and S. Wendelken. A model-based approach to human identification using ECG. In *Proceedings of SPIE, Orlando, Florida, USA*, volume 7306, 2009.

[48] J. M. Irvine and S. A. Israel. A Sequential Procedure for Individual Identity Verification Using ECG. *EURASIP Journal on Advances in Signal Processing*, 2009.

[49] J. M. Irvine, S. A. Israel, W. T. Scruggs, and W. J. Worek. eigenPulse: Robust human identification from cardiovascular function. *Pattern Recognition*, 41:3427 – 3435, 2008.

[50] J.M. Irvine, S.A. Israel, M.D. Wiederhold, and B.K. Wiederhold. A new biometric: human identification from circulatory function. In *Joint Statistical Meetings of the American Statistical Association, San Francisco*, 2003.

[51] J.M. Irvine, B.K. Wiederhold, L.W. Gavshon, S.A. Israel, S.B. McGehee, R. Meyer, and M.D. Wiederhold. Heart rate variability: a new biometric for human identification. In *International Conference on Artificial Intelligence (IC-AI'2001), Las Vegas, Nevada, USA*, 2001.

[52] S. A. Israel and J. M. Irvine. Heartbeat biometrics: a sensing system perspective. *International Journal of Cognitive Biometrics*, 1(1):39 – 65, 2012.

[53] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold. ECG to identify individuals. *Pattern Recognition*, 38:133 – 142, 2005.

[54] S. A. Israel, W. T. Scruggs, W. J. Worek, and J. M. Irvine. Fusing Face and ECG for Personal Identification. In *Proceedings of the 32nd Applied Imagery Pattern Recognition Workshop*, 2003.

[55] F. Jager, A. Taddei, M. Emdin, G. Antolic, R. Dorn, G. B. Moody, B. Glavic, A. Smrdel, M. Varanini, M. Zabukovec, S. Bordigiago, C. Marchesi, and R. G. Mark. The Long-Term ST Database: A Research Resource for Algorithm Development and Physiologic Studies of Transient Myocardial Ischemia. *Computers in Cardiology*, 27, 2000.

[56] Anil K. Jain and Arun Ross. Multibiometric systems. *Commun. ACM*, 47:34 – 40, Jan. 2004.

[57] D. Jang, S. Wendelken, and J. M. Irvine. Robust Human Identification Using ECG: Eigenpulse Revisited. In *Proceedings of SPIE, Orlando, Florida, USA*, volume 7667, 2010.

[58] A. D. Kaplan. Information Processing for Biological Signals: Application to Laser Doppler Vibrometry. In *PhD Dissertation, Department of Electrical and Systems Engineerig, School of Engineering and Applied Science, Washington University in St Louis*, 2011.

[59] A. D. Kaplan, J. A. O'Sullivan, E. J. Sirevaag, and J. W. Rohrbaugh. Laser Doppler vibrometry measurements of the carotid pulse: biometrics using hidden Markov models. In *Proceedings of SPIE, Orlando, Florida, USA*, 2009.

[60] A.D. Kaplan, J.A. O'Sullivan, E.J. Sirevaag, P-.H. Lai, and J.W. Rohrbaugh. Hidden State Models for Noncontact Measurements of the Carotid Pulse Using a Laser Doppler Vibrometer. *IEEE Transactions on Biomedical Engineering*, 59(3):744 – 753, Mar. 2012.

[61] D. B. Keesing. Development and implementation of fully 3D statistical image reconstruction algorithms for helical CT and half-ring PET insert system. In *Ph.D. Dissertation, Washington University in St. Louis*, 2009.

[62] I. Khalil and F. Sufi. Legendre Polynomials Based Biometric Authentication Using QRS Complex of ECG. In *Proceedings of the 4th International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, 2008.

[63] K-.S. Kim, T-.H. Yoon, J-.W. Lee, D-.J. Kim, and H-.S. Koo. A Robust Human Identification by Normalized Time-Domain Features of Electrocardiogram. In *Proceedings of the 2005 IEEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China*, 2005.

[64] A. Kumar and S. Shekhar. Personal Identification Using Multibiometrics Rank-Level Fusion. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 41(5):743 – 752, Sep. 2011.

[65] L. I. Kuncheva. *Combining Pattern Classifiers: Methods and Algorithms*. Wiley-Interscience, New Jersey, 2004.

[66] M. Kyoso. A technique for avoiding false acceptance in ECG identification. In *IEEE EMBS Asian-Pacific Conference on Biomedical Engineering*, 2003.

[67] M. Kyoso and A. Uchiyama. Development of an ECG identification system. In *Proceedings of the 23rd Annual EMBS International Conference, Istanbul, Turkey*, 2001.

[68] P. Laguna, R. G. Mark, A. Goldberger, and G. B. Moody. A Database for Evaluation of Algorithms for Measurement of QT and Other Waveform Intervals in the ECG. *Computers in Cardiology*, 24:673 – 676, 1997.

[69] P-.H. Lai, J. A. O'Sullivan, M. Chen, E. J. Sirevaag, A. D. Kaplan, and J. W. Rohrbaugh. A Robust Feature Selection Method for Noncontact Biometrics Based on Laser Doppler Vibrometry. In *Biometric Symposium (BSYM), Tampa, Florida, USA*, 2008.

[70] L. Lam and S.Y. Suen. Application of majority voting to pattern recognition: an analysis of its behavior and performance. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 27(5):553 – 568, Sep. 1997.

[71] K. Lange. Convergence of EM image reconstruction algorithms with Gibbs smoothing. *IEEE Transactions on Medical Imaging*, 9(4):439 – 446, Dec. 1990.

[72] G. M. Lasio, B. R. Whiting, and J. F. Williamson. Statistical reconstruction for x-ray computed tomography using energy-integrating detectors. *Phys. Med. Biol.*, 52(8):2247 – 2266, Apr. 2007.

[73] M. Li and S. Narayanan. Robust ECG Biometrics by Fusing Temporal and Cepstral Information. In *20th International Conference on Pattern Recognition (ICPR)*, pages 1326 – 1329, Aug. 2010.

[74] J. L. C. Loong, K. S. Subari, R. Besar, and M. K. Abdullah. A New Approach to ECG Biometric Systems: A Comparative Study between LPC and WPD Systems. In *World Academy of Science, Engineering and Technology*, pages 759 – 764, Aug. 2010.

[75] A. Lourenço, H. Silva, and A. Fred. Unveiling the biometric potential of finger-based ECG signals. *Computational Intelligence and Neuroscience*, 2011(720971), 2011.

[76] K. MacCabe and D. Brady. Fan beam CAXSI radiance model. Personal Communication, 2012.

[77] K. P. MacCabe. *X-RAY SCATTER TOMOGRAPHY USING CODED APERTURES*. Ph.D. dissertation, University of North Carolina., Chapel Hill, NC, 2014.

[78] K. P. MacCabe, A. D. Holmgren, M. P. Tornai, and D. J. Brady. Snapshot 2D tomography via coded aperture x-ray scatter imaging. *Appl. Opt.*, 52(19):4582 – 4589, Jul. 2013.

[79] K. P. MacCabe, K. Krishnamurthy, A. Chawla, D. L. Marks, E. Samei, and D. J. Brady. Pencil beam coded aperture x-ray scatter imaging. *Optics Express*, 20(15):16310 – 16320, 2012.

[80] A. J. Mansfield and J. L. Wayman. Best Practices in Testing and Reporting Performance of Biometric Devices, Aug. 2002. Retrieved on Feb. 1, 2012.

[81] P. Marchionni, L. Scalise, I. Ercoli, and E. P. Tomasini. An optical measurement method for the simultaneous assessment of respiration and heart rates in preterm infants. *Review of Scientific Instruments*, 84(12):1 – 9, 2013.

[82] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki. The DET Curve in Assessment of Detection Task Performance. In *Proceedings of Eurospeech*, volume 4, pages 1895 – 1898, 1998.

[83] MATLAB. *version 7.11.0.584 (R2010b)*. The MathWorks Inc., Natick, Massachusetts, 2010.

[84] G. G. Molina, F. Bruekers, C. Presura, M. Damstra, and M. van der Veen. MORPHOLOGICAL SYNTHESIS OF ECG SIGNALS FOR PERSON AUTHENTICATION. In *European Signal Processing Conference, Poznan, Poland*, 2007.

[85] Multix. Multix Xray Spectrometric Imaging, Mar. 2014.

[86] B. Nasri, M. Guennoun, and K. El-Khatib. Using ECG as a measure in biometric identification systems. In *IEEE Toronto International Conference on Science and Technology for Humanity (TIC-STH)*, pages 28 – 33, Sep. 2009.

[87] I. Odinaka, P-.H. Lai, A. D. Kaplan, J. A. O'Sullivan, E. J. Sirevaag, S. D. Kristjansson, A. K. Sheffield, and J. W. Rohrbaugh. ECG biometrics: A robust short-time frequency analysis. In *IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1 – 6, Dec. 2010.

[88] I. Odinaka, P-.H. Lai, A. D. Kaplan, J.A. O'Sullivan, E.J. Sirevaag, and J.W. Rohrbaugh. ECG Biometric Recognition: A Comparative Analysis. *IEEE Transactions on Information Forensics and Security*, 7(6):1812 – 1824, Dec. 2012.

[89] I. Odinaka, P-.H. Lai, A.D. Kaplan, and J.A. O'Sullivan. On estimating biometric capacity: An example based on LDV biometrics. In *48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 146 – 151, Sep. 2010.

[90] I. Odinaka, J.A. O'Sullivan, D.G. Politte, K.P. MacCabe, K. Krishnamurthy, M. Lakshmanan, A. Kapadia, and D.J. Brady. Computational Imaging Approach for Fanbeam X-ray Coherent Scatter Imaging. *Under Internal Review*, Mar. 2014.

[91] I. Odinaka, J.A. O'Sullivan, E.J. Sirevaag, and J.W. Rohrbaugh. Cardiovascular Biometrics: Combining Mechanical and Electrical Signals. *Under Review for IEEE Transactions on Information Forensics and Security*, Mar. 2014.

[92] J.A. O'Sullivan and J. Benac. Alternating minimization algorithms for transmission tomography. *IEEE Transactions on Medical Imaging*, 26(3):283 – 297, 2007.

[93] J.A. O'Sullivan, D. L. Snyder, and B.R. Whiting. Alternating minimization algorithms for transmission tomography using energy detectors. In *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*, volume 1, pages 144 – 147, Nov. 2002.

[94] R. Palaniappan and S. M. Krishnan. Identifying individuals using ECG beats. In *International Conference on Signal Processing and Communications*, 2004.

[95] PhysioNet. Combining Multiple Biometrics., 2000. Retrieved on Feb. 1, 2012.

[96] K. N. Plataniotis, D. Hatzinakos, and J. K. M. Lee. ECG biometric recognition without fiducial detection. In *Proceedings of Biometric Symposium (BSYM), Baltimore, Maryland, USA*, 2006.

[97] Polytec. PSV-400 Scanning Vibrometer, Mar. 2014.

[98] J. L. Prince and J. M. Links. *Medical Imaging Signals And Systems*. Pearson Prentice Hall Bioengineering. Pearson Prentice Hall, 2006.

[99] A. Riera, S. Dunne, I. Cester, and G. Ruffini. STARFAST: a Wireless Wearable EEG/ECG Biometric System based on the ENOBIO Sensor. In *Proceedings of the 5th International Workshop on Wearable Micro and Nanosystems for Personalised Health*, 2008.

[100] A. A. Ross, K. Nandakumar, and A. K. Jain. *Handbook of Multibiometrics (International Series on Biometrics)*. Springer-Verlag New York, Inc., Secaucus, NJ, 2006.

[101] L. Sachs. *Applied Statistics: A Handbook of Techniques*. Springer-Verlag, New York, NY, 2 edition, 1984.

[102] S. Saechia, J. Koseeyaporn, and P. Wardkein. Human Identification System Based ECG Signal. In *TENCON 2005 IEEE Region 10*, pages 1 – 4, Nov. 2005.

[103] S. I. Safie, J. J. Soraghan, and L. Petropoulakis. Electrocardiogram (ECG) Biometric Authentication Using Pulse Active Ratio (PAR). *IEEE Transactions on Information Forensics and Security*, 6(4):1315 – 1322, Dec. 2011.

[104] N. A. Schmid and J. A. O'Sullivan. Thresholding Method for Reduction of Dimensionality. *IEEE Trans. Information Theory*, 47(7):2903 – 2920, 2001.

[105] N. A. Schmid and J. A. O'Sullivan. Performance Prediction Methodology for Multibiometric Systems. In R.I. Hammoud, B.R. Abidi, and M.A. Abidi, editors, *Face Biometrics for Personal Identification. Multi-Sensory Multi- Modal Systems*, pages 213 – 227. Springer-Verlag, Berlin, 2007.

[106] T-.W. Shen, W. J. Tompkins, and Y. H. Hu. One-Lead ECG for identity verification. In *2nd Joint Conference of the IEEE Engineering in Medicine and Biology Society and the Biomedical Engineering Society, Houston, TX, USA*, 2002.

[107] T.-W. Shen, W. J. Tompkins, and Y. H. Hu. Implementation of a one-lead ECG human identification system on a normal population. *Journal of Engineering and Computer Innovations*, 2(1):12 – 21, Jan. 2011.

[108] L.A. Shepp and B.F. Logan. The Fourier reconstruction of a head section. *IEEE Transactions on Nuclear Science*, 21(3):21 – 43, Jun. 1974.

[109] H. Silva, H. Gamboa, and A. Fred. One Lead ECG Based Personal Identification with Feature Subspace Ensembles. In *Proceedings of the 5th International Conference on Machine Learning And Data Mining In Pattern Recognition*, 2007.

[110] B. W. Silverman. *Density estimation for statistics and data analysis*. Chapman and Hall, New York, NY, 1986.

[111] Y. N. Singh and P. Gupta. ECG to individual identification. In *Proceedings of the 2nd IEEE International Conference on Biometrics: Theory, Applications and Systems*, 2008.

[112] Y. N. Singh and P. Gupta. Correlation-based classification of heartbeats for individual identification. *Soft computing - A Fusion of Foundations, Methodologies and Applications*, 2009.

[113] S. K. Singla and A. Sharma. ECG as Biometric in the Automated World. *International Journal of Computer Science and Communication*, 1(2):281 – 283, 2010.

[114] D. L. Snyder, T. J. Schulz, and J. A. O'Sullivan. Deblurring subject to nonnegativity constraints. *IEEE Transactions on Signal Processing*, 40(5):1143 – 1150, May 1992.

[115] D.C. Sorensen. Newton's Method with a Model Trust Region Modification. *SIAM Journal on Numerical Analysis*, 19(2):409 – 426, 1982.

[116] J. Sriram, M. Shin, T. Choudhury, and D. Kotz. Activity-aware ECG-based patient authentication for remote health monitoring. In *Proceedings of the 2009 international conference on Multimodal interfaces*, pages 297 – 304, 2009.

[117] F. Sufi, Q. Fang, I. Khalil, and S. S. Mahmoud. Compressed ECG Biometric: A Fast, Secured and Efficient Method for Identification of CVD Patient. *Journal of Medical Systems*, 2009.

[118] F. Sufi, I. Khalil, and I. Habib. Polynomial distance measurement for ECG based biometric authentication. *Security and Communication Networks Online*, 2008.

[119] F. Sufi, I. Khalil, and J. Hu. ECG-based authentication. In Peter Stavroulakis and Mark Stamp, editors, *Handbook of Information and Communication Security*, pages 309 – 331. Springer, 2010.

[120] M. M. Tawfik and H. S. T. Kamal. Human Identification Using QT Signal and QRS Complex of the ECG. *The Online Journal on Electronics and Electrical Engineering (OJEEE)*, 3(1):383 – 387, 2011.

[121] Y-.T. Tsao, T-.W. Shen, T-.F. Ko, and T-.H. Lin. The Morphology of the Electrocardiogram for Evaluating ECG Biometrics. In *The 9th International Conference on e-Health Networking, Application and Services, Taipei, Taiwan*, 2007.

[122] H. L. Van Trees. *Detection, Estimation, and Modulation Theory*. John Wiley and Sons, Inc., Cambridge, MA, 2001.

[123] N. Venkatesh and S. Jayaraman. Human Electrocardiogram for Biometrics Using DTW and FLDA. In *20th International Conference on Pattern Recognition (ICPR)*, pages 3838 – 3841, Aug. 2010.

[124] Y. Wan and J. Yao. A Neural Network to Identify Human Subjects with Electrocardiogram Signals. In *Proceedings of the World Congress on Engineering and Computer Science*, 2008.

[125] X. Wang and S. S. Shen. Estimation of Spatial Degrees of Freedom of a Climate Field. *Climate*, 12:1280 – 1291, 1999.

[126] Y. Wang, F. Agrafioti, D. Hatzinakos, and K. N. Plataniotis. Analysis of Human Electrocardiogram for Biometric Recognition. *EURASIP Journal on Advances in Signal Processing*, 2008(148658), 2008.

[127] D. J. L. Wayman. Fundamentals of Biometric Authentication Technologies. *International Journal of Image and Graphics*, 1(1):93 – 113, 2001.

[128] M. B. Westover and J. A. O'Sullivan. Achievable Rates for Pattern Recognition. *IEEE Transactions of Information Theory*, 54(1):299 – 320, 2008.

[129] B. R. Whiting, P. Massoumzadeh, O. A. Earl, J. A. O'Sullivan, D. L. Snyder, and J. F. Williamson. Properties of preprocessed sinogram data in x-ray computed tomography. *Med Phys*, 33(9):3290 – 3303, Sep. 2006.

[130] G. Wübbeler, M. Stavridis, D. Kreiseler, R-.D. Bousseljot, and C. Elster. Verification of humans using the electrocardiogram. *Pattern Recognition Letters*, 28:1172 – 1175, 2007.

[131] L. Xu, A. Krzyzak, and C.Y. Suen. Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Transactions on Systems, Man and Cybernetics*, 22(3):418 – 435, May 1992.

[132] J. Yao and Y. Wan. A wavelet method for biometric identification using wearable ECG sensors. In *Proceedings of the 5th International Workshop on Wearable and Implantable Body Sensor Networks*, 2008.

[133] J. Yao and Y. Wan. Improving Computing Efficiency of a Wavelet Method Using ECG as a Biometric Modality. *International Journal of Computer and Network Security (IJCNS)*, 2(1):15 – 20, Jan. 2010.

[134] C. Ye, M.T. Coimbra, and B.V.K.V. Kumar. Investigation of human identification using two-lead Electrocardiogram ECG signals. In *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, pages 1 – 8, Sep. 2010.

[135] Z. Zhang and D. Wei. A new ECG identification method using Bayes' Theorem. In *Proceedings of the IEEE Region 10 Conference*, 2006.

Appendix A

Databases

There are six public ECG databases and many more in-house private databases that have been used for evaluating the recognition performance of an ECG biometric system.

A.1 Public Databases

The public databases that have been used for evaluating the performance of ECG recognition systems include the MIT-BIH Arrhythmia (MITDB), MIT-BIH Normal Sinus Rhythm (NSRDB), PTB Diagnostic (PTBDB), QT (QTDB), European ST-T (EDB), and Long-Term ST (LTSTDB) databases. All six databases are hosted in the Physionet website [41].

MITDB and NSRDB are ECG collections that were obtained in the Arrhythmia Laboratory at Boston's Beth Israel Hospital. MITDB is a collection of 48 fully annotated 30-minute excerpts of two-channel ambulatory ECG recordings, obtained from 47 individuals suffering from some form of arrhythmia.

NSRDB is a collection of 18 long-term (about a day) two-channel ECG recordings. The individuals included in the database were found to have had no significant arrhythmias.

PTBDB (Physikalisch-Technische Bundesanstalt Database) is an ECG collection that was provided by the National Metrology Institute of Germany for teaching and research purposes, and for algorithm evaluation. The database contains 549 recordings from 290 individuals. Each individual has between one and seven recordings. Although the subject number runs from one to 294, there are no individuals numbered 124, 132, 134, or 161. Each recording has 15 data channels; the standard 12 leads and the three Frank leads. Fifty two of the 290 individuals are healthy, while the others suffer from a variety of cardiac disorders.

There appears to be some confusion in the literature about the number of healthy individuals from the PTB database with recordings from multiple days. Seven (patient numbers 174, 180, 198, 233, 245, 251, and 284) of the 52 healthy individuals have recordings from at least two different days, while only one individual (patient number 180) apparently has recordings from three or more different days.

QTDB is a collection of 105 15-minute excerpts of two-channel ECG recordings, selected to avoid significant baseline wander or other artifacts [68]. The recordings were obtained from five MIT-BIH ECG databases (including MITDB and NSRDB), the European ST-T database (EDB), and an in-house database containing Holter recordings of patients who experienced sudden cardiac death during the recordings.

EDB is a collection of 90 two-channel recordings from 79 individuals suffering from myocardial ischemia, intended for use in evaluating algorithms that analyze changes in the ST segment and T wave.

LTSTDB contains 86 long-term two- or three-channel ECG recordings from 80 individuals, chosen to represent a variety of changes in the ST segment [55]. The database is mainly intended for use in evaluating algorithms that can differentiate between ischemic and non-ischemic ST events.

A.2 Private Databases

In addition to the public databases, several researchers have utilized recordings obtained in their laboratories for evaluating the biometric capability of their proposed ECG recognition system.

The database used by Biel *et al.* [10] consisted of four to 10 recordings obtained from 20 individuals. The standard 12-lead recordings were taken over a span of about six weeks, during normal resting conditions.

The database used by Irvine *et al.* [51, 50, 49, 48], Israel *et al.* [53], Homer *et al.* [47], and Jang *et al.* [57] were from data collection protocols designed to elicit different states of anxiety in the individuals. Single-channel recordings were obtained from a total of 126 individuals.

The database used by Shen *et al.* [107] consisted of lead I recordings obtained from the palms of 168 healthy individuals during normal resting conditions.

The database used by Kim *et al.* [63] consisted of single-channel ECG recordings obtained from 10 healthy males, both during normal resting condition and a physical activity (running up and down a flight of stairs).

The database used by Zhang and Wei [135] consisted of 10 s standard 12-lead ECG recordings from 520 individuals. This database represents the largest that has been reported for the development of ECG biometrics algorithms.

The database used by Silva *et al.* [109] and Coutinho *et al.* [26] consisted of 26 individuals. Within a session, each individual performed a series of cognitive tasks.

The database used by Molina *et al.* [84] consisted of ECG recordings obtained from 10 individuals, obtained during normal resting conditions. Each individual participated in five 5-minute sessions over the course of four weeks.

The database used by Wübbeler *et al.* [130] consisted of 234 three-channel (leads I, II, and III) 10-s ECG recordings obtained from 74 healthy individuals. Each individual had between two and 20 recordings, with an average time of about 500 days between each recording.

The database used by Chan *et al.* [17] consisted of 90 s ECG recordings obtained from 60 healthy individuals. There were three recordings for each individual, corresponding to three sessions, with a minimum of one day between sessions. The button electrodes were held between the thumb and index finger.

The database used by Yao and Wan [132, 124, 133] consisted of 121 two-minute ECG recordings taken from 30 individuals, during normal resting condition. Each individual participated in at least two sessions, with the time interval between two sessions ranging from several hours to a few weeks.

The database used by Agrafioti and Hatzinakos [7] and Gao *et al.* [39] consisted of three-minute single-channel wrist recordings from 52 healthy individuals. The recordings were repeated about a month later, for 16 of the individuals [39].

The database used by Lourenço *et al.* [75] consisted of two-minute single-channel finger recordings from 16 individuals within a single session.

Appendix B

Multibiometrics

B.1 LDV Biometric Channel

The LDV biometric system that is employed in this study is the two separate aligned segments log-normal method [20]. This method has been shown to be the best technique, so far, for LDV biometric recognition. Two segments of lengths 370 ms and 400 ms are extracted from each LDV pulse signal, centered at the maximum velocity peak and the incisura, respectively. Each sample in each segment is normalized by subtracting the mean value of all the samples in the segment, and dividing by the standard deviation of the samples in the segment. A possible explanation for why the separate aligned segments work better than using a single segment aligned to either the main velocity peak or the incisura location is that the main velocity peak and the incisura are produced by two different underlying processes, the timing of which is variable across multiple heartbeats, depending on the duration of left ventricular ejection. Aligning to the two segments helps reduce gross misalignments that occur further away from the centering point.

Features are extracted from each of the two segments using a short-time Fourier transform (STFT) with a Hamming window of size 96 ms and a step size of 16 ms; a cutoff frequency of 150 Hz was utilized. This resulted in 522 and 580 time-frequency features for the first and second segments, respectively. A log-normal model was then applied to the absolute value of the STFT coefficients [20]. For our purpose, the logarithm of the magnitude of the STFT coefficients is called a spectrogram. Each of the time-frequency components of the spectrogram is called a time-frequency bin (feature) or simply a bin. The log-normal model assumes that for a given heartbeat, each bin is independent and Gaussian distributed, for a given individual, each heartbeat is independent, and that individuals are independent [20]. The appropriate models for the data are not known. The models were not selected based on

their generative abilities, but on their utility for classification, as demonstrated in the works of Chen *et al.* [20] and Odinaka *et al.* [87, 88]. During training, estimates of the mean and variance of the Gaussian distributions are obtained via maximum likelihood estimation.

Informative features (bins) are obtained by thresholding [104], based on the symmetric relative entropy between each individual's bins and those of the population [69, 87]; the distributions of the population bins are obtained by using the spectrograms from all the enrolled individuals. The distribution of the population bins can be obtained from heartbeats from a different set of individuals than the ones being trained/tested, making the training process for an enrolled individual independent of everybody else. The symmetric relative entropy between two densities p and q is defined as

$$d(p,q) = D(p\|q) + D(q\|p), \tag{B.1}$$

where

$$D(p\|q) = \int p \log \frac{p}{q}$$

and the integral is taken over the support set of p.

For the Gaussian distribution used in our model, the symmetric relative entropy is given as

$$\begin{aligned} RE_{ki} &= \frac{\sigma_{ki}^2(j) + (\mu_{ki}(j) - \mu_{k0}(j))^2}{2\sigma_{k0}^2(j)} \\ &+ \frac{\sigma_{k0}^2(j) + (\mu_{ki}(j) - \mu_{k0}(j))^2}{2\sigma_{ki}^2(j)} - 1, \end{aligned} \tag{B.2}$$

where k indexes the biometric trait, j indexes the time-frequency bins, and μ_{k0} and μ_{ki} are the stored mean feature vector for the population and the i-th individual, respectively. σ_{k0}^2 and σ_{ki}^2 are the stored variance feature vectors for the population and the i-th individual, respectively. RE_{ki} is the symmetric relative entropy vector between the i-th individual's model and that of the population.

The relative entropy-based feature selection thresholds are obtained by cross validation using the biometric data from the training individuals. That is, the threshold is swept, and the one that gives the best authentication performance is used subsequently during the authentication and identification of the test individuals.

The score function used in the authentication and identification modes is based on a log-likelihood ratio test statistic obtained from the simple binary hypothesis

$$Y_k(j) \overset{H_0}{\sim} \mathcal{N}\left(\mu_{k0}, \sigma_{k0}^2\right), \quad Y_k(j) \overset{H_1}{\sim} \mathcal{N}\left(\mu_{ki}, \sigma_{ki}^2\right),$$

where Y_k is the test feature vector for the k-th biometric trait.

If all the time-frequency bins are used, the Gaussian log-likelihood ratio function is given as

$$\begin{aligned} S_{ki} &= \sum_{j=1}^{J_k} \left[\log\left(\frac{\sigma_{k0}^2(j)}{\sigma_{ki}^2(j)}\right) + \frac{(Y_k(j) - \mu_{k0}(j))^2}{\sigma_{k0}^2(j)} \right. \\ &- \left. \frac{(Y_k(j) - \mu_{ki}(j))^2}{\sigma_{ki}^2(j)} \right], \end{aligned} \tag{B.3}$$

where J_k is the total number of time-frequency bins for the k-th modality.

During authentication, a modified log-likelihood ratio test is used for classification. The log-likelihood ratio function was modified by setting the variance of the bins belonging to the claimed identity and those of the population to a constant (homoscedastic assumption). Each of the two segments from a test LDV pulse signal produces a score, so the scores are averaged to obtain a score representative of the test signal [20]. The score is then compared to a threshold, and a decision is made whether to accept or reject the claimed identity. Suppose the claimed identity belongs to the i-th individual in the database, then the match score from each of the two segments is given as (the same score function is used for the ECG biometric system, and the multibiometric system fused at the feature level)

$$\begin{aligned} S_{ki} &= \sum_{j=1}^{J_k} \left[(Y_k(j) - \mu_{k0}(j))^2 \right. \\ &- \left. (Y_k(j) - \mu_{ki}(j))^2 \right] I_{\{RE_{ki}(j) > \tau_k\}}, \end{aligned} \tag{B.4}$$

where τ_k is the bin-selection threshold for the k-th biometric trait, and $I_{\{\}}$ is an indicator function.

During identification, the match scores are computed using the log-likelihood ratio between the individual templates (stored models) and the population model. The enrolled individual whose stored model gives the best score is associated with the test signal. In a rank-m identification setting, a classification error occurs if the true identity is not among the enrolled

individuals whose models achieved the top m scores. The identification match score between the test signal Y_k and the i-th enrolled individual's model is given by (the same score function is used for the ECG biometric system, and the multibiometric system fused at the feature level) equation B.3 above, where the summation is only over the selected informative features enforced by the indicator function, as in the case of authentication.

The score functions above were successfully used in the past by Chen *et al.* [20] and Odinaka *et al.*[88]. Intuitively, the closer the features $Y_k(j)$ are to the population mean $\mu_{k0}(j)$, the smaller the match score. Whereas, the closer features are to the mean $\mu_{ki}(j)$ of the i-th enrolled individual, the greater the match score.

B.2 ECG Biometric Channel

The ECG biometric system employed in this study is the same as that proposed in previous studies [87, 88]. A 700 ms segment is extracted from each ECG pulse signal, aligned at the peak of the R wave. The segment has 200 ms before and 499 ms after the peak. Each sample in the segment is normalized by subtracting the mean value of all the samples in the segment, and dividing by the standard deviation of the samples in the segment.

Features are extracted from each segment using an STFT with a Hamming window of size 64 ms and a step size of 10 ms; a cutoff frequency of 250 Hz was used. This resulted in 2048 time-frequency features. A log-normal model was then applied to the absolute value of the STFT coefficients [87]. Each of the time-frequency components of the spectrogram (logarithm of the magnitude of the STFT coefficients) is called a time-frequency bin (feature). The log-normal model, as was the case with the LDV biometric system, assumes that for a given heartbeat, each bin is independent and Gaussian distributed, for a given individual, each heartbeat is independent, and that individuals are independent. Again, the models were not selected based on their generative abilities, but on their utility for classification [87, 88].

The ECG biometric system uses the same methodology as the LDV biometric system for training the log-normal models and computing match scores.

B.3 Feature-level Fusion

In this work, feature-level fusion involves concatenating the informative features extracted from the two segments of the LDV signal and the ECG signal into a single feature vector. The log-likehood and modified log-likelihood ratio tests, as described in subsection B.1 above, are used for identification and authentication, respectively.

B.4 Score-level Fusion

Double Sigmoid

The parameters of the sigmoid functions can be ascertained from the region of overlap between the impostor and genuine training score distribution [56, 100]; impostor scores are scores obtained when an individual claims a false identity, while genuine scores are those obtained when a true identity is claimed. The two sigmoid functions are piecewise-continuously co-joined at the intersection of the non-parametrically estimated genuine and impostor probability density functions. The extent of the linear regime of each of the sigmoid functions is determined by the amount of overlap between the two probability density functions to the left and right of the intersection point. Given a test score S_{ki}^{l}, obtained using the l-th signal, the i-th individual's template and the k-th biometric system ($k = 1, 2$, corresponding to the ECG and LDV biometric systems respectively), the normalized score is given as [100]

$$NS_{ki}^{l} = \begin{cases} \frac{1}{1+\exp\left(-2\left(\frac{S_{ki}^{l}-\nu}{\alpha_1}\right)\right)}, & \text{if } S_{ki}^{l} < \nu, \\ \frac{1}{1+\exp\left(-2\left(\frac{S_{ki}^{l}-\nu}{\alpha_2}\right)\right)}, & \text{otherwise,} \end{cases} \tag{B.5}$$

where ν is the score at the intersection point of the two probability density functions, α_1 is the left edge of the linear regime of the first sigmoidal function and α_2 is the right edge of the linear regime of the second sigmoidal function.

Tanh Estimator

Given a test score S_{ki}^{l}, obtained using the l-th signal, the i-th individual's template and the k-th biometric system ($k = 1, 2$, corresponding to the ECG and LDV biometric systems respectively), the normalized score is given as [100]

$$NS_{ki}^{l} = \frac{1}{2}\left\{\tanh\left(0.01\left(\frac{S_{ki}^{l} - \text{Med}}{\text{MAD}}\right)\right) + 1\right\}, \tag{B.6}$$

where Med and MAD are the median and standard median absolute deviation of the genuine training scores, respectively.

B.5 Rank-level Fusion

Let $r_{k,i}$ be the rank assigned to the i-th class (enrolled individual) by the k-th matcher, $k = 1, \ldots, K$.

The highest rank method assigns to a class, the smallest rank across the ranks output for that class, by all the biometric matchers. The fused rank is given as

$$R_i = \min_{k=1}^{K} r_{k,i}. \tag{B.7}$$

The fused ranks are then reordered from smallest to largest. If there is a tie, it is broken randomly.

The Borda count method adds the ranks from all the biometric matchers. The fused rank is given as

$$R_i = \sum_{k=1}^{K} r_{k,i}. \tag{B.8}$$

Ties between the reordered fused ranks are also broken randomly.

The weighted Borda count method computes a linearly weighted sum of the ranks output for each class, by all the biometric matchers. The fused rank is given as

$$R_i = \sum_{k=1}^{K} w_k r_{k,i}. \tag{B.9}$$

The weights w_k can be determined by means of logistic regression.

After obtaining the fused rank by using the highest rank method, the modified highest rank method resolves ties in the fused rank by using a small fraction of the Borda count for each class [1]. The fused rank using this modified highest rank method is given by [1]

$$R_i = \min_{k=1}^{K} r_{k,i} + \epsilon_i, \tag{B.10}$$

for some large constant C, where

$$\epsilon_i = \frac{\sum_{k=1}^{K} r_{k,i}}{C}. \tag{B.11}$$

The nonlinear weighted ranks methods combine the ranks by using either a weighted sum of a nonlinear function (exp) of the ranks or the sum of a nonlinear function (tanh or exp) of the weighted ranks. The fused rank using the weighted sum of a nonlinear function of the ranks is given by [64]

$$R_i = \sum_{k=1}^{K} w_k f(r_{k,i}), \tag{B.12}$$

where f is *exp*,

while, the fused rank based on the sum of a nonlinear function (tanh or exp) of the weighted ranks is given by [64]:

$$R_i = \sum_{k=1}^{K} f(w_k r_{k,i}), \tag{B.13}$$

where f is *tanh* or *exp*.

The weighted Borda count and the nonlinear weighted ranks methods require a training phase in order to determine the weights for each biometric matcher.

B.6 Decision-level Fusion

In authentication mode, the ECG and LDV biometric systems were combined using "AND" and "OR" rules. The "AND" rule accepts the claimed identity only when all the biometric matchers accept the identity. Alternatively, the "OR" rule accepts the claimed identity when at least one of the biometric matchers accepts it. Since the false acceptance and rejection rates of the multimodal system depend on the score thresholds adopted by each of the unimodal systems (ECG and LDV), we determined the EER of the multimodal system as the lowest point on the curve of intersection between the false acceptance and rejection surfaces.

Bayesian decision fusion combines the decisions output by the individual biometric matchers into probability values by using Bayes' rule. During testing, given the decision labels output by all the matchers, the probability of each of the classes (accept or reject, in authentication, the enrolled individuals, in identification) is computed, and the class with the highest probability is chosen. Let d_k be the decision label assigned by the k-th matcher, $k = 1, \ldots, K$, and c_i, the i-th class (enrolled individual), $i = 1, \ldots, I$. The probability of the i-th class being assigned to the test signal, given the decision labels, $\mathbf{d} = [d_1 \ldots d_K]$, from the matchers is given as

$$\mathrm{P}(c_i|\mathbf{d}) \propto g_i = \mathrm{P}(\mathbf{d}|c_i)\mathrm{P}(c_i). \tag{B.14}$$

The final decision of the multibiometric system is given as the class with the highest value of g_i.

The Bayesian decision fusion technique requires a training phase, where the joint probability of each set of decision labels given each class, $\mathrm{P}(\mathbf{d}|c_i)$, is computed. When the amount of training data that is available is limited, it may be infeasible to compute this joint probability, in which case a simplifying assumption of independence ("naïve") between the biometric matchers is made; reducing the joint probability of the set of decision labels given a class to a product of the marginal probability of each decision label given the class. The joint probability of a set of decision labels, $\mathbf{d} = [d_1 \ldots d_K]$ given the i-th class, c_i is given by [100]

$$\mathrm{P}(\mathbf{d}|c_i) = \mathrm{P}(d_1, \ldots, d_K|c_i) = \prod_{k=1}^{K} \mathrm{P}(d_k|c_i). \tag{B.15}$$

To estimate the marginal probabilities $\mathrm{P}(d_k|c_i)$ and the prior probability of each class $\mathrm{P}(c_i)$, during training, a confusion matrix C^k is constructed for each biometric matcher. A confusion matrix is a square matrix, whose i-th row and j-th column corresponds to the number of training samples belonging to the i-th class (individual) that were classified as belonging to the j-th class. Normalizing the i-th row by dividing by the total number of examples in the row and dividing the total number of examples in the i-th row by the total number of training samples, provides an estimate for the marginal probabilities $\mathrm{P}(d_k|c_i)$ and the prior probability of the class $\mathrm{P}(c_i)$, respectively.

Dempster-Shafer fusion combines the decisions output by the individual biometric matchers into degrees of belief values. During testing, given the decision labels output by all the matchers, the degree of belief of each of the classes (accept or reject, in authentication, the enrolled individuals, in identification) is computed, and the class with the highest degree of belief is chosen. The decision labels output by the matchers can be summarized in a matrix called a decision profile [100]. Let d_k be the decision label assigned by the k-th matcher, $k = 1, \ldots, K$, and c_i, the i-th class (enrolled individual), $i = 1, \ldots, I$. The decision profile is a $K \times I$ matrix whose ki-th entry is the degree of support of the i-th class, by the k-th biometric matcher. The degree of support is either 1 (when the decision rendered by the k-th biometric matcher is the i-th class) or 0. During training, a decision template (DT^i) is constructed for each enrolled individual, as the average of the decision profiles from the training examples belonging to that individual [100].

During testing, after computing the decision profile (DP) for the test signal, a measure of similarity between the decision profile and the stored decision templates of the enrolled individuals is computed as [100]

$$S_{k,i} = \frac{\left(1 + \left(||DP_k - DT_k^i||\right)^2\right)^{-1}}{\sum_{j=1}^{I}\left(1 + \left(\left|\left|DP_k - DT_k^j\right|\right|\right)^2\right)^{-1}}, \tag{B.16}$$

where DP_k and DT_k^i are the k-th row of DP and DT^i, respectively, and $||.||$ represents the l^2-norm.

Based on the similarity measure, the degree of belief $b_{k,i}$ of the i-th class using the k-th classifier can be computed by [100]

$$b_{k,i} = \frac{S_{k,i}\left[\prod_{j=1,j\neq i}^{I}(1-S_{k,j})\right]}{1-S_{k,i}\left[\prod_{j=1,j\neq i}^{I}(1-S_{k,j})\right]}. \tag{B.17}$$

Finally, the degrees of belief of all the biometric matchers can be consolidated into a single degree of belief for each class, by multiplying them together, and the final decision of the multibiometric system is given as the class with the highest value of the combined degree of belief.

Although majority voting is not directly applicable to fusing two modalities, we consider majority voting with test sample rejection and majority voting with random tie breaks. When the two modalities differ in their decisions, we can randomly select one of the decisions (random tie breaks) or completely reject (no decision) the test sample. When rejecting the discordant test samples, the identification performance is based only on the test samples that are not rejected.

Appendix C

Notes

C.1 Fan Beam Source, Flat-panel Detector Array

After profiling the reconstruction algorithm in MATLAB, we found that the backward operator computations were about 4 times slower than those of the forward operator. The culprit for this mismatch in computational time was the process of finding the detector location for each scatter angle as shown in Pseudocode 7.8. To make the speed of the forward and backward operators comparable, we precomputed the detector location for each predefined scatter angle.

Vita

Ikenna C. Odinaka

Degrees

Ph.D., Electrical Engineering, Washington University in St. Louis, May 2014

M.Sc., Electrical Engineering, Washington University in St. Louis, Saint Louis, MO, October 2013

B.Sc., Summa Cum Laude, Physics and Mathematics, Illinois Wesleyan University, Bloomington, IL, May 2008

Honors & Awards

Donald Wann Fellowship — Awarded to the most outstanding first-year graduate student in the Department of Electrical and Systems Engineering, Washington University in St. Louis, 2009

Annabelle Scrogin Anderson award — Awarded to one student for distinguished achievement in multiple fields of study, Illinois Wesleyan University, 2008

Order of Titans — Awarded to a few students on the basis of academic accomplishments, athletic contributions, and general citizenship, Illinois Wesleyan University, 2008

Professional Societies

Phi Kappa Phi
Phi Beta Kappa
Institute of Electrical and Electronics Engineers (IEEE)

Certifications

Health Insurance Portability and Accountability Act (HIPAA)

Publications

I. Odinaka, J. A. O'Sullivan, D.G. Politte, K. MacCabe, K. Krishnamurthy, M. Lakshmanan, A. Kapadia, D.J. Brady, "Computational Imaging Approach for Fanbeam X-ray Coherent Scatter Imaging," Pending Internal Revision.

I. Odinaka, J. A. O'Sullivan, E.J. Sirevaag, J.W. Rohrbaugh, "Cardiovascular Biometrics: Combining Mechanical and Electrical Signals," Submitted to *IEEE Transactions on Information Forensics*

and Security, 2014.

I. Odinaka, P.-H. Lai, A. D. Kaplan, J. A. O'Sullivan, E.J. Sirevaag, J.W. Rohrbaugh, "ECG Biometric Recognition: A Comparative Analysis," *IEEE Transactions on Information Forensics and Security* 2012.

I. Odinaka, P.-H. Lai, A. D. Kaplan, J. A. O'Sullivan, E.J. Sirevaag, S.D. Kristjansson, A.K. Sheffield, J.W. Rohrbaugh, ECG biometrics: A robust short-time frequency analysis, *IEEE International Workshop on Information Forensics and Security (WIFS)* 2010.

I. Odinaka, P.-H. Lai, A. D. Kaplan, J. A. O'Sullivan, On estimating biometric capacity: An example based on LDV biometrics, *Allerton Conference on Communication, Control, and Computing*, 2010.

Service Activities

Volunteer Math Tutor, Urban Future, Saint Louis, MO, Fall 2011

Extracurricular Activities

Men's Club Soccer, Washington University in St. Louis, 2010 - 2014
Final four — NIRSA national soccer championship & offensive player of the year, 2012

Men's Varsity Soccer, Illinois Wesleyan University, 2005 - 2008
First Team Academic All-District, 2007
Second Team Academic All-American, 2007
Second Team All-Conference in the College Conference of Illinois and Wisconsin, 2007
Record for scoring two goals in 17 seconds against Wheaton College, 2005

May 2014