2013

Signal Processing and Performance Evaluation Issues in Multi-Sensor Data Fusion

Chuanming Wei
Lehigh University

SIGNAL PROCESSING AND PERFORMANCE EVALUATION ISSUES IN MULTI-SENSOR DATA FUSION

by

Chuanming Wei

Presented to the Graduate and Research Committee

of Lehigh University

in Candidacy for the Degree of

Doctor of Philosophy

in

Electrical Engineering

Lehigh University

January 2013

Approved and recommended for acceptance as a dissertation in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Date

Prof. Rick S. Blum
(Dissertation Director)

Accepted Date

Committee Members:

Prof. Rick S. Blum
(Committee Chair)

Prof. Tiffany Jing Li

Prof. Shalinee Kishore

Dr. Lance M. Kaplan

Acknowledgements

First of all, I wish to express my sincere thanks to my advisor, the chair of my doctoral dissertation examining committee, Prof. Rick S. Blum. I consider myself very fortunate to have him as my advisor, who is always ready to offer the best guidance and help whenever I come to him with questions, regardless of whether these questions are about my research work or are about my life. Not only has Prof. Blum showed me his creativity and passion for research, he also encouraged my own interests and helped steer me towards successfully publishing our work. Because of Prof. Blum, I have been able to enjoy my stay at Lehigh University and the process of pursuing my Ph.D. degree. I believe in the future I will still benefit from the things I learnt from him.

I am also grateful to Prof. Tiffany Jing Li, Prof. Shalinee Kishore and Dr. Lance M. Kaplan, for serving on my doctoral dissertation examining committee and giving me many invaluable suggestions for the accomplishment of this dissertation. My special thanks go to Lance, who gave me the opportunity to collaborate with U.S Army Research Lab. Our work became a very important part of my dissertation.

I owe my gratitude to Prof. Xiaolei Huang from Computer Science and Engineering Department of Lehigh University, who helped me with the opportunity to work as an intern at Lister Hill National Center for Biomedical Communications, NIH. In addition, I am indebted to Dr. George Thoma, Dr. Sameer Antani, Rodney Long, May Cheh, Ed Kim, Beibei Cheng, Wen Cheng and many other great people during my internship at NIH.

I take this opportunity to extend my thanks to Dr. Ami Wiesel in the Hebrew University of Jerusalem, Israel. It is truly an honor to collaborate with Ami Wiesel. His insightful comments improved our work significantly.

My deepest thanks also go to my Master advisors Prof. Ling Qiu and Prof. Jinkang Zhu in the University of Science and Technology of China. I have benefited extremely from them during my Master study. Without their guidance, I might not have the chance to come to Lehigh University for my Ph.D. degree.

During my years in the Signal Processing and Communications Lab, many of my seniors and colleagues encouraged and supported me. I would like to specially mention the following: Jinzhong Yang, Yang Yang, Danny Safi, Zhemin Xu, Xun Chen, Qian He, Anand Srinivas Guruswamy, Liang Zhao, Xinda Ke, Jiangfan Zhang, Basel Alnajjab, and Raghusrinivas Vishnubhotla. I thank them all for their help.

This dissertation is dedicated to my dear dad and mom, my sister, Ranran, and my wife, Yuan, without whom none of this would have been possible. Thank you, my family.

Contents

List of Tables

List of Figures

List of Acronyms

CDWT color discrete wavelet transform

CLAV color averaging

CLMT color multiscale transform

CONA contrast pyramid A

CONB contrast pyramid B

CRB Cramer-Rao bound

CWD concealed weapon detection

DC direct current

DPMLR diffuse prior monotonic likelihood ratio

DPMLRT diffuse prior monotonic likelihood ratio test

DWT discrete wavelet transform

EIV errors-in-variables

FIM Fisher information matrix

FIQM fused image quality measure

G-GLRT global generalized likelihood ratio test

GGM Gaussian graphical model

GLR generalized likelihood ratio

GLRT generalized likelihood ratio test

HVS human visual system

II image intensified

IQ image quality

LC logistic correlation

L-GLRT local generalized likelihood ratio test

LRT likelihood ratio test

LWIR long-wave infrared

MC monotonic correlation

MI mutual information

ML maximum likelihood

MIMO multi-input and multi-output

MMW millimeter-wave

MOUT military operation in an urban terrain

MSE mean squared error

NIIRS national imagery interpretability ratings scale

P-GLRT pseudo generalized likelihood ratio test

PMU phasor measurement unit

RMSE root mean squared error

ROC receiver operating characteristic

SCNR signal-to-clutter-plus-noise ratio

TLS total least squares

TLS-GLRT total least squares based generalized likelihood ratio test

TML total maximum likelihood

TML-GLRT total maximum likelihood based generalized likelihood ratio test

UIQI universal image quality index

UMP uniformly most powerful

Abstract

Over the past few decades, multi-sensor data fusion has been applied to a broad range of problems in many different areas including object detection and recognition, target tracking, remote sensing, medical diagnosis, robotics, and autonomous vehicles. Researchers have recognized that the synergistic combination of data from multiple sensors can provide a more robust and complete view of the object of interest than can be achieved by a single-sensor system. Further advances require a better understanding of the science behind different multi-sensor data fusion systems. This dissertation presents our research on several selected issues concerning multi-sensor data fusion systems that have recently received significant attention. In particular, we focus on the novel signal processing design and performance evaluation techniques for three popular systems: the multi-sensor image fusion system, the multi-input and multi-output (MIMO) radar system and the distributed sensor network.

Quantitatively measuring the performance of a multi-sensor image fusion system is a complicated but important task. We focus on the theoretical analysis of three correlation-based fused image quality measures (FIQMs) when they are used to judge the performance of weighted averaging image fusion algorithms. Our analysis shows that when we change the power of the desired signal or the noise in the source images, these correlation-based FIQMs exhibit some undesired behaviors. In addition, we develop a novel statistic to score the effectiveness of FIQMs for the detection task in light of practical measurements from human perception experiments. The performance of the proposed monotonic test is demonstrated via Monte Carlo simulations. We also show the application of the proposed method to evaluate potential FIQMs in a specific target detection experiment.

The second part of this dissertation considers the joint location and velocity estimation

problem in a multi-target non-coherent MIMO radar system. The Cramer-Rao bound (CRB) is a useful tool for evaluating the performance of radar systems, as it provides the mean square error lower bound for any unbiased estimation. In this dissertation, we focus on a multi-target case, in which a non-coherent MIMO radar system is considered. This case has not yet been studied by others. We investigate the joint location and velocity estimation of multiple targets, and the Cramer-Rao bound for a two-target case is derived and evaluated. This bound gives us theoretically achievable joint estimation performance for a sufficient number of antennas.

The third part of the dissertation considers the design of change detection methods using observations from distributed sensor networks, where each node has access to local observations and is only allowed to communicate with its neighbors. Results apply to monitoring large systems like the electrical grid but they also apply generally to cases where sensors monitor changes in a random field. Using our algorithms, all the nodes will reach a consensus on the test in the end. First, we study the distributed change detection problem for distributions that can be represented as Gaussian graphical models. We propose two distributed tests. The first distributed test is a natural approach which simply applies the generalized likelihood ratio test (GLRT) to smaller size local clusters in the graph. The second method employs the pseudo-likelihood as a surrogate function for the global likelihood. Next, we consider the fault detection problem for measurements following the errors-in-variables (EIV) model. The standard approach to parameter estimation in such problems is known as total least squares (TLS). Recently, a competing approach known as total maximum likelihood (TML) was proposed and was shown to provide promising performance gains in various estimation problems. Following these works, we derive the TLS based GLRT and the TML based GLRT, which are specifically tailored for the smart grid structure.

Chapter 1

Introduction

1.1 Background

Over the past two decades, there has been growing interest in multi-sensor data fusion, an emerging technology that originated from the requirements of military applications, which was soon applied to a wide variety of nonmilitary uses. Multi-sensor data fusion, as defined in [1], is the process of combining sensor data from multiple sources, to provide a robust and complete description of an object of interest. There are two main reasons for the emerging interest in multi-sensor data fusion. First, the development of advanced computing and sensing techniques enables real-time processing of data from multiple sensors. Second and more importantly, researchers have recognized that the synergistic combination of data from multiple sensors can be used to overcome the limitations of a single-sensor system, i.e., it can provide more accurate or complete information than can be achieved by using one sensor alone.

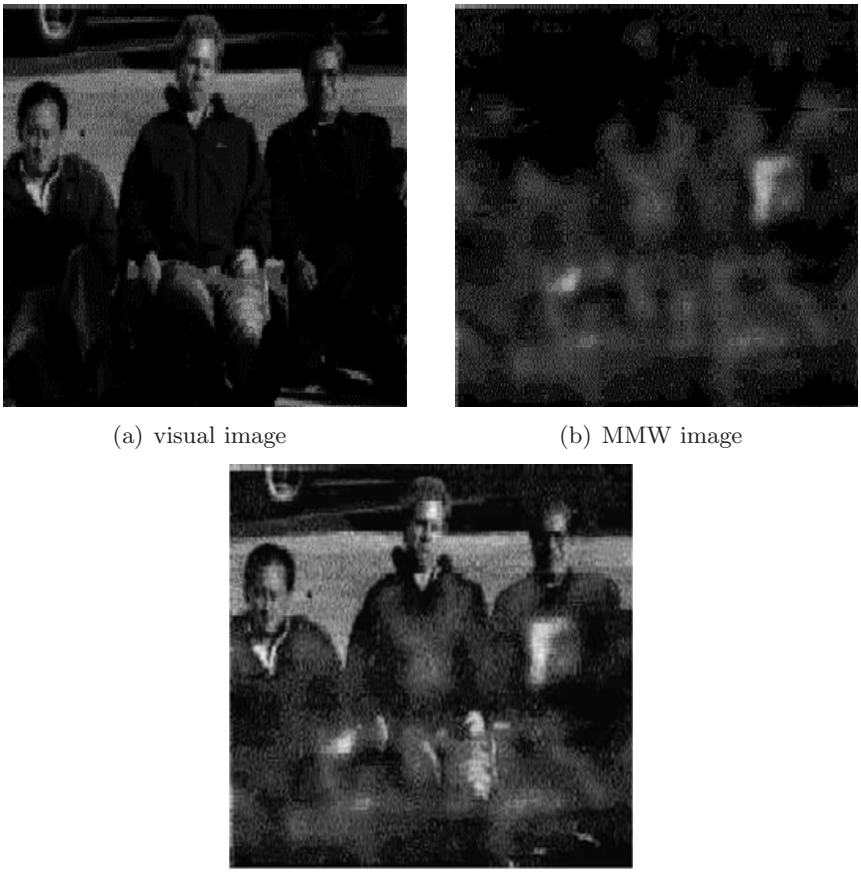
Strictly speaking, a multi-sensor data fusion system does not necessarily have to use

multiple sensor modalities. The observations that have to be combined may come from a group of sensors of the same or different types. They may also be produced by only one sensor but at different time instants. The main advantage of multi-sensor data fusion comes from the redundant information and the complementary information derived from multiple sources. The fusion of redundant information can serve to reduce the uncertainty of the sensor system and increase its reliability. The estimation error can be reduced if we have more independent observations and combine them in an optimal way. Complementary information may be obtained by various types of sensors or by the same type of sensors placed at different locations. Because of the ways these sensors work or their relative placement, each of such sensors could only provide a part of the information one wants to perceive (e.g., the image of some object). Therefore the integration of those data allows a more complete view of the object of interest.

Multi-sensor data fusion has been applied to a broad range of problems in many different areas including object detection and recognition, target tracking, remote sensing, medical diagnosis, robotics, and autonomous vehicles. In this dissertation, our research focuses on three multi-sensor data fusion systems that have recently received significant attention: the multi-sensor image fusion system, the MIMO radar system and the distributed sensor networks. The following subsections provide a brief introduction to these systems and a summary of our work.

1.2 Multi-Sensor Image Fusion

Multi-Sensor Image fusion refers to generating a combined image in which each pixel is determined from a set of pixels in each of the source images [2]. The fused image should



(a) visual image (b) MMW image

(c) fused image

Figure 1.1: Example of image fusion for a concealed weapon detection (CWD) application.

provide an easier view for a human to interpret the scene than any of the source images does, thus improving the performance of the human in accomplishing his/her task. Image fusion has been widely used in many applications such as concealed weapon detection (CWD) [3], remote sensing [4], intelligent robots [5], medical diagnosis [6], and surveillance [7], etc. The interested reader is referred to Chapter 1 of [2] for a survey of various image fusion algorithms developed in past years.

Fig.1.1 gives an example to illustrate the usage of image fusion in a CWD application

[8, 9]. Fig.1.1(a) and Fig.1.1(b) show the visual image and the 94 GHz millimeter-wave (MMW) image of the same scene respectively. The visual image provides the outline and the appearance of the people in the scene while the MMW image shows the existence of a concealed object (possibly a gun). The visual sensor can provide a good view of the people, but not of the concealed object. The MMW sensor can provide a good view of the object, but not of the people. Therefore, one would like to produce one single image combining all important information from the two source images, so that it will be easy to identify the people carrying weapons. Fig.1.1(c) provides an example of such a fused image, which presents considerable evidence to suspect that the person on the right has a concealed gun beneath his clothes.

Measuring the performance of image fusion algorithms is an extremely important task, which has received past study [10–25]. The performance of image fusion algorithms is primarily assessed by perceptual evaluation in the form of subjective human tests [16]. Typically in these tests, human observers are asked to view a series of fused images and rate them. Because images are fused for better human interpretation, it is important to judge fusion methods by how well humans are able to perform interpretation tasks. Examples of human interpretation studies for image fusion evaluations appear in [20, 25]. Regardless of the goal of the human perception test, these tests are inconvenient, expensive and time consuming.

It is clearly highly desirable to identify an objective performance measure that can accurately predict human perception by determining the quality of the fused image. The objective measure should be a feature that is obtained via an automatic computation employing the fused image and can serve as a surrogate for human perception results. We refer to such a feature as the *fused image quality measure* (FIQM). If a good FIQM can be devised, then one can compare image fusion algorithms without expensive perception experiments. Furthermore,

the measure can be used as a design criteria for an "optimal" image fusion algorithm.

In this dissertation, first we extend the study in [11] by performing theoretical analysis for three recently introduced FIQMs. These FIQMs are based on computing the correlation between the images input into the fusion algorithm and the image produced by the fusion algorithm, and therefore we call them *correlation-based* FIQMs. Since the FIQM is used to reveal how well a human can interpret a given fused image, we expect a good FIQM to be consistent with the human perception results. For example, if noise is added to the source images, we expect the FIQM for the fused image to decrease. We study the closed-form expressions for these FIQMs by employing the image formation model described in [26], and analyze quality changes with respect to changes in the noise power of the source images. Our interesting finding is that these correlation-based FIQMs sometimes behave opposite to what is expected. Sufficient conditions for the unexpected behavior along with an intuitive explanation are provided in the dissertation.

Next, we are also interested in which FIQM better describes the performance of the human interpreting the fused imagery. Our previous work indicates that we can judge a FIQM by how well it is consistent with actual human performance. We believe that the best FIQM should be task specific. Thus we focus on the case of a target detection application. In this case, performance is measured by the probability that a human observer can correctly detect certain objects in the fused image. The human perception experiments measure the number of observers that are able to correctly detect ground truthed targets as the human performance. This performance metric can be reasonably modeled by a binomial distribution. We develop a novel statistic to quantify how well a monotonic function explains the relation between the values of a FIQM and human performance over a series of fused imagery representing the same scene. The monotonic statistic is general and can be applied to other applications when

one may need to test for a monotonic relationship.

1.3 MIMO Radars

Multiple-input multiple-output (MIMO) radar systems have received a lot of study in recent years. A MIMO radar is broadly defined as a radar system that employs multiple transmitters/receivers and jointly processes the received signals [27]. Unlike the standard phased-array radar, which transmits the same waveforms with relatively different phases, a MIMO radar can choose the waveform for each of its transmit antennas more freely.

There are two types of MIMO radar systems: MIMO radar with co-located antennas and MIMO radar with widely separated antennas. For the former configuration, the target is modeled as a single point, i.e., the reflection coefficients of a target are the same for all transmit-receive paths. Because of the direct applicability of adaptive arrays techniques and the enhanced flexibility of transmit waveform design [28], there have been many waveform optimization methods proposed for the MIMO radar system with co-located antennas [28,29]. The approach enables this system to get superior capabilities over the standard phased-array radar system. The MIMO radar with widely separated antennas is able to view the target from several different directions and take advantage of the spatial properties of the target. Therefore, it has the ability of avoiding large attenuated target returns over all transmit-receiver paths, overcoming bandwidth limitations, and performing high resolution target detection and parameter estimation [27]. The MIMO radar system with widely separated antennas is considered in this dissertation.

Both coherent and non-coherent processing approaches can be used on the receiver side of MIMO radar systems. If the MIMO radar only uses the amplitude information of the received

signals while processing the data, the processing approach is non-coherent. Otherwise it is coherent. Coherent processing requires that the oscillators at each transmit and receive antenna be aligned in phase, which is usually not very easy to achieve.

The Cramer-Rao bound (CRB) is a useful tool for evaluating the performance of radar systems, as it provides the mean square error lower bound for any unbiased estimation. Recently, there have been various Cramer-Rao bound studies for the performance of MIMO radar systems in the literature [30–35]. In this dissertation, we focus on a multi-target case, in which a non-coherent MIMO radar system is considered. This case has not yet been studied by others. We investigate the joint location and velocity estimation of multiple targets, and the Cramer-Rao bound for a two-target case is derived and evaluated. This bound gives us theoretically achievable joint estimation performance. Numerical results show that the spatial advantage observed previously for single-target cases can also be observed in two-target cases.

1.4 Fault Detection Using Distributed Sensor Networks

We refer to an interconnected sensor network as a system consisting of multiple sensors that work together to perform a specific task. The deployment of large scale systems taking place today is accompanied with growing vulnerability and security concerns, which makes it desirable to recognize abrupt system failures and intrusions rapidly. We are interested in the design of methods to rapidly detect a change in the observation model of such systems. The change could be due to a fault event that occurs in the system [36–38], system parameter errors [39] or even malicious cyber-intrusions [40,41]. Change detection finds its application in a wide variety of areas such as remote monitoring, quality control, speech signal segmentation, system fault detection and seismic data processing. Many different algorithms have been

proposed to solve specific fault detection problems. Solutions often result to detecting a change in the parameters of a static or dynamic stochastic system model. A comprehensive survey on change detection techniques was provided in [42].

The traditional way to apply the change detection techniques using sensor networks is to utilize a fusion center that gathers all sensor observations and performs all computations. This problem has been well-studied and reported in the literature [36–38,43,44]. However, this centralized architecture is being replaced by the new distributed approaches, wherein each individual node has the ability to process data and cooperates with other sensors. There are several reasons why a distributed approach is preferred in many cases. Distributed processing does not need a central processing node, which makes the whole system more reliable to node failures. Distributed algorithms are usually more scalable than the centralized algorithms when the size of the sensor network increases rapidly. Furthermore, with proper design, each node in the distributed sensor network requires less communication and processing resources. This is very important, especially for the energy-constrained wireless sensor network where a large portion of the battery power is consumed by the communication and computing.

Our work focuses on failure detection in the large scale systems, with an emphasis on distributed and scalable methods. For this purpose, we have relied on probabilistic graphical models and error in variables models.

First, we studied the distributed change detection problem for distributions that can be represented as Gaussian graphical models (GGMs). A graphical model characterizes conditional dependencies within a multivariate distribution using undirected graphs [45, 46]. When the graph is sparse and the variables are jointly Gaussian, the graphical model imposes sparsity on the inverse covariance, also called the information, concentration or precision matrix. Graphical models are attractive in two complementary manners. First, they allow

the modeling of a large system using a smaller number of parameters. Second, inference methods can be performed as local decentralized computations with message passing between neighbors of the graph. For this purpose, it is common to assume that the topology of the statistical models matches the topology of the internode communications. Based on this framework, GGMs have been successfully applied to different statistical problems such as inference [47], parameter estimation [48, 49], and dimensionality reduction [50]. Applications range from wireless sensor networks [51,52], internet backbone networks [50] and recently the smart power grid [53].

The time-tested approach to change detection is via composite hypothesis testing, and in particular the generalized likelihood ratio test (GLRT) [54]. This method has been successfully applied to change detection in the Gaussian multivariate distribution, and is known as the Bartlett's test [55, 56]. Classical centralized results on testing in GGM are available in [45]. In contrast, our focus is on distributed solutions for change detection. We begin by deriving the global centralized GLRT, and then propose two distributed approximations based on aggregating multiple tests performed at each of the nodes in the graph. The aggregation can then be efficiently implemented using consensus methods [57]. The first distributed test is a natural approach which simply applies the Bartlett's test to smaller size local clusters in the graph. The second method employs the pseudo-likelihood [58] as a surrogate function for the global likelihood. The advantages of our proposed tests are demonstrated using numerical experiments, including one in the context of failure detection in smart power grids [53].

Next, we considered the change detection problem for measurements following the errors-in-variables (EIV) model. From a statistical perspective, the noisy measurements of voltages and currents in a smart grid system can be considered as a special case of linear EIV models [59]. The standard approach to parameter estimation in such problems is known as total

least squares (TLS), and generalizes classical least squares by allowing noise in both sides of the linear model. Hypothesis testing in EIV models has also been addressed. The GLRTs have been derived in the context of process monitoring [36] as well as array processing [44]. Following these works, we derive the TLS-GLRT which is specifically tailored for the smart grid structure and examine its detection performance.

Parameter estimation in EIV models is known to be difficult. TLS and its extensions are often unstable and may be improved. Recently, a competing approach known as total maximum likelihood (TML) was proposed in [60–62] and was shown to provide promising performance gains in various estimation problems. In particular, it was shown that TML can be interpreted as a regularized version of TLS with improved properties. As a continuation of the work of [60–62], we consider hypothesis testing within the TML framework and derive the TML-GLRT for detecting changes in such models. First, we express a smart grid fault detection problem as a hypothesis testing problem employing a linear EIV model. Then we derive the corresponding TLS-GLRT. Second, we propose a novel approach for testing in EIV models based on the TML methodology. Specifically, we consider this approach in the context of the smart grid, but the technique may be applied to other EIV models. Numerical results show the promising detection advantages TML-GLRT has with no increase in computational complexity.

1.5 Outline of the Dissertation

The rest of this dissertation is organized as follows.

Chapter 2 focuses on theoretical analysis of three correlation-based FIQMs when they are used to judge the performance of weighted averaging image fusion algorithms. Our analysis

shows that when we change the power of the desired signal or the noise in the source images, these correlation-based FIQMs exhibit some undesired behaviors. The sufficient conditions for when the undesired behaviors occur and the intuitive explanation for our observations are given in this chapter.

Chapter 3 develops a novel statistic to score the effectiveness of FIQMs for the detection task in light of practical measurements from human perception experiments. The performance of the proposed monotonic test is demonstrated via Monte Carlo simulations. We also show the application of the proposed method to evaluate potential FIQMs in a specific target detection experiment.

Chapter 4 considers the joint location and velocity estimation problem in a multi-target non-coherent MIMO radar system. The Cramer-Rao bound for a two-target case is derived and evaluated. This bound gives us theoretically achievable joint estimation performance for a sufficient number of antennas. Numerical results show that the spatial advantage observed previously for single-target cases can also be observed in two-target cases.

Chapter 5 studies the distributed change detection problem for distributions that can be represented as GGMs. We formulate the hypothesis testing problem and propose a global and centralized solution using the GLRT. We then provide two distributed approximations to this global test based on aggregation of multiple local or conditional tests. We compare the performance of these tests in the context of failure detection in smart grids.

Chapter 6 considers fault detection through apparent changes in the bus susceptance parameters of modern power grids. We formulate the problem using a linear EIV model and derive its corresponding GLRT based on the TLS methodology. Next, we propose a competing detection technique based on the recently proposed TML framework. We derive the so called TML-GLRT, and show that it can be interpreted as a regularized TLS-GLRT.

We also discussed the distributed implementation of TLS-GLRT.

The following notation is used in this dissertation. Boldface uppercase letters denote matrices. Boldface lowercase letters denote column vectors. Standard letters denote scalars. The superscript $(\cdot)^T$ denotes the transpose of a vector or a matrix, and $(\cdot)^{-1}$ denotes the matrix inverse. $|\mathbf{X}|$ is the determinant of matrix $\mathbf{X}$ and $\mathrm{Tr}\{\mathbf{X}\}$ represents its trace. $\mathbf{X} \succ \mathbf{0}$ means $\mathbf{X}$ is positive definite. The union of sets a and b is denoted by $a \cup b$, and $a \setminus b$ is the set of elements in a which are not in b. We use indices in the subscript $[\mathbf{x}]_a$ or $[\mathbf{X}]_{a,b}$ to denote sub-vectors or sub-matrices, respectively. Where possible, we omit the brackets and use $\mathbf{x}_a$ or $\mathbf{X}_{a,b}$ instead. $||\cdot||_F$ is the Frobenius norm of a matrix. $\mathbf{I}$ represents the identity matrix.

Chapter 2

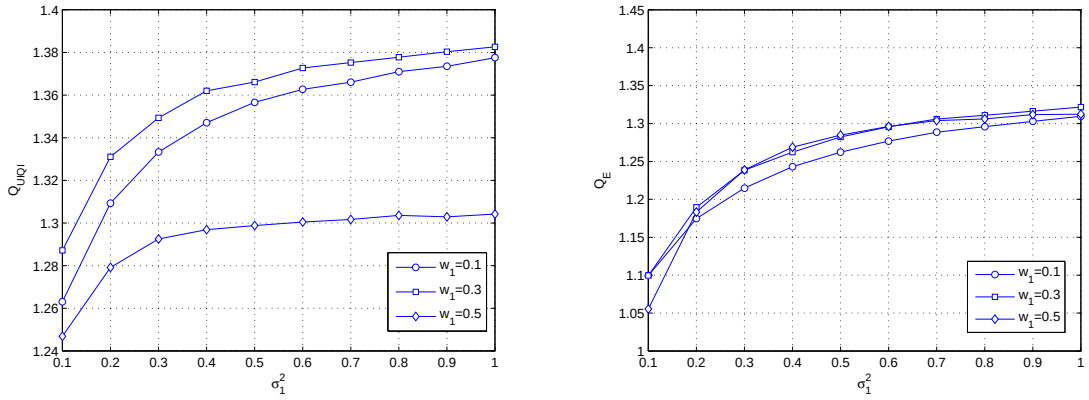
Theoretical Analysis of Correlation-Based FIQMs

2.1 Introduction

(a) image 1: IR image

(b) image 2: II image

Figure 2.1: Infrared and low light visual test images for a night vision application.



(a) $Q_{UIQI}, \sigma_2^2 = 0.07059$ (b) $Q_E, \sigma_2^2 = 0.08652$

Figure 2.2: Q_{UIQI} and Q_E with increasing σ_1^2 for night vision image fusion (Fig.2.1).

To make the purpose of this chapter very clear, we start with an example of the phenomenon we study. Consider the quality measures introduced in [12, 15] and the source images shown in Fig.2.1. Fig.2.2 shows these quality measures, calculated exactly as specified in [12, 15], when we add noise with power σ_1^2 to the first source image and noise with power σ_2^2 to the second source image. Note that for both quality measures[1], the quality measure increases, indicating higher quality, as we increase σ_1^2. Clearly, increasing the noise power in a source image will not really increase the quality of the fused image. Thus we are observing bad behavior in the quality measure. In particular we are able to recreate this behavior using a reasonable mathematical model which is not subject to any experimental inaccuracies or to any imperfections introduced through the capture of real images. Using the model we prove rigorously that these quality measures exhibit this bad behavior and we provide intuitive explanations for the behavior.

Measuring the performance of image fusion algorithms is an extremely important task

[1]The same things happens for the quality measure in [14]. See [11] for more detail.

which has received past study [12, 13, 15, 63–67]. The vast majority of investigations have been focused on developing approaches suitable for experimental evaluation of fused images. Thus, given a set of source images and a few fusion algorithms, the goal is to determine which fusion algorithm is more suitable for a given application. Such experimental procedures can be very useful and in fact we have done some research in this area [63, 68]. However, there are some inherent limitations to experimental tests. For example, experimental tests are only valid for the particular set of source images considered. Trying to argue that these results apply for a larger set of cases than the exact set considered (with the exact source images and any imperfections in them) is not possible in a rigorous way. Experimental studies can be considered only for a finite number of specific cases with a very specific set of conditions in each case. Thus trying to prove that a fusion approach is generally good is not possible using experimental tests. Further, whenever real data is taken there are always some unknown imperfections that enter into this data and these imperfections can influence the results. Of course one tries to minimize these imperfections, but typically they can not be completely eliminated. It is very easy to study a whole set of cases all at once using analysis and to vary any or all parameters in any way desired. Most importantly, using mathematical analysis one can *prove* that certain performance measures have some bad behavior, as we do in this chapter, without concern that imperfections in the experimental tests or the images used are causing this bad behavior. This can not be accomplished using the existing experimental approaches. Thus while experimental investigations can be useful, analytical investigations can be very helpful in filling in the gaps that such studies can never address. We are hoping this chapter will encourage further analytical image fusion research. We hope that our research will show that there is value in analytical image fusion studies. We acknowledge that we have only just started to study this topic and that many unanswered questions remain. We also acknowledge

that we consider a highly simplified model. Analysis with more comprehensive models is a great topic for future research.

Recently, several FIQMs have been introduced [12,14,15] which are all based on a similar philosophy. The philosophy involves considering correlations between the images input into the fusion algorithm and the image produced by the fusion algorithm, possibly after some preprocessing. The correlation is intended to provide a measure of the amount of information transferred from the source images to the fused image. Higher correlation implies more information transfer and thus is considered to result in better fused images. Here we shall refer to the methods from [12, 14, 15] as *correlation-based* FIQMs. The correlation-based FIQMs have received a lot of attention in the past few years because they don't require a ground-truth reference image which is usually difficult to obtain in most applications. Here we focus on three correlation metrics which are the mutual information [14] [11], the universal image quality index (UIQI) [12] and the edge-based correlation metric [13] [15].

In [11], the authors employed a statistical model [26] which approximates the observed sensor images as an additive sum of the scene of interest plus the random distortion part, which we call noise for simplicity. The authors of [11] studied the mutual information-based FIQM from [14]. In particular they studied how this FIQM reacted to increases in the noise power of the source images for weighted averaging image fusion algorithms. They made an interesting observation. In some cases the FIQM indicated higher quality for increased input noise power. In this chapter, we extend the study in [11] by performing theoretical analysis for three different correlation-based FIQMs applied to the weighted averaging fusion algorithm. We study the closed-form expressions for these FIQMs by employing the image formation model described in [26] and analyze quality changes with respect to changes in noise power. Our analysis shows that all of these three correlation-based FIQMs sometimes

behave opposite to what is expected. Sufficient conditions for the unexpected behavior along with an intuitive explanation are provided in this chapter.

The rest of this chapter is organized as follows. Section 2.2 presents the statistical model for the observed sensor images. Our main findings are stated in Section 2.3. In Section 2.4 we describe three correlation-based FIQMs and derive their closed-form expressions. An intuitive explanation for our findings is provided in Section 2.5.

2.2 Image Formation Model

To perform analytical studies, it is typically necessary to develop a model for the process under consideration. Here we employ a very simple model which was first proposed in [26]. We hope that other researchers will build on these ideas and propose more comprehensive models and perform more comprehensive studies based on these models.

Assume that we observe two registered sensor images. Let z_1 and z_2 represent two corresponding pixels from two registered sensor images to be fused. Assume these pixels can be modeled as

$$z_k = \beta_k s + n_k, k = 1, 2, \tag{2.1}$$

where s describes the underlying scene, β_k denotes the sensor selectivity factor of image k (its ability to "view" the underlying scene), and n_k represents all unwanted degradation which includes random distortion and noise. For brevity we call n_k noise in the sequel. This model assumes that certain objects may or may not be seen by certain sensors, as determined by the size of β_k which takes on real values.

The quantities s, n_1 and n_2 are modeled as independent Gaussian random variables such

that

$$s \sim N(\mu_s, \sigma_s^2), \tag{2.2}$$

$$n_k \sim N(0, \sigma_k^2), k = 1, 2 \tag{2.3}$$

where $N(\mu, \sigma^2)$ denotes a Gaussian random variable with mean μ and variance σ^2. As linear transforms of Gaussian random variables result in Gaussian random variables, this model also applies after any linear transforms, including wavelet and other multiscale transforms.

We focus on the weighted averaging fusion method because it is the most basic algorithm and is well-studied. The fused image using the weighted averaging fusion method is

$$f = w_1 z_1 + w_2 z_2 = (w_1\beta_1 + w_2\beta_2)s + w_1 n_1 + w_2 n_2, \tag{2.4}$$

where w_1 and w_2 are the weights for the two different source images. We note that equations (2.1)-(2.4) apply at each pixel and the values of z_k, s, n_k, β_k and w_k for $k = 1, 2$ can be different for each pixel.

Increasing the power of the noise added into either source image should lead to a loss of quality. Thus a good FIQM should satisfy the following desired behavior.

Desired Behavior 1. A FIQM for the fused image should decrease if σ_1^2 or σ_2^2, the power of the noise added to the source images, is increased.

Increasing the power of the signal of interest should lead to an increase in quality. Thus a good FIQM should satisfy the following desired behavior.

Desired Behavior 2. A FIQM for the fused image should increase if σ_s^2, the power of the signal of interest, is increased.

2.3 Correlation-Based FIQMs

Recently, several FIQMs have been introduced [12, 14, 15] which are all based on computing the correlation between the images input into the fusion algorithm and the image produced by the fusion algorithm. We call these *correlation-based* FIQMs. The general form of these FIQMs can be expressed as

$$Q = F(f, z_1) + F(f, z_2), \tag{2.5}$$

where $F(x, y)$ is the metric used to measure the correlation between the source images and the fused image.

These correlation-based FIQMs exhibit some bad behavior as described by the following Theorems.

Theorem 1 *Under the model in (2.1), the correlation-based FIQMs in [12, 14, 15] violate Desired Behavior 1 under some conditions.*

Theorem 2 *Under the model in (2.1), the correlation-based FIQMs in [12, 14, 15] violate Desired Behavior 2 under some conditions.*

The proofs of Theorem 1 and Theorem 2 are omitted due to space limits.

We can also describe conditions when other correlation-based FIQMs, that are described by (2.5), exhibit bad behavior. The following Theorem addresses this issue.

Theorem 3 *If $F(f, z_1)$ and $F(f, z_2)$ can be written as functions of σ_1^2, σ_2^2 and σ_s^2, then when $F(f, z_1)$ and $F(f, z_2)$ satisfy*

$$\frac{\partial F(f, z_1)}{\partial(\sigma_1^2)} + \frac{\partial F(f, z_2)}{\partial(\sigma_1^2)} \geqslant 0, \tag{2.6}$$

the correlation-based FIQM violates Desired Behavior 1. When $F(f, z_1)$ and $F(f, z_2)$ satisfy

$$\frac{\partial F(f, z_1)}{\partial(\sigma_s^2)} + \frac{\partial F(f, z_2)}{\partial(\sigma_s^2)} \leqslant 0, \tag{2.7}$$

the correlation-based FIQM violates Desired Behavior 2.

The proof of Theorem 3 is straightforward because equation (2.6) implies the correlation-based FIQM Q increases with increasing noise power and equation (2.7) means Q decreases with increasing signal power.

2.4 Description of Three Different Correlation-Based FIQMs

Here we provide more details on the three popular correlation-based FIQMs. Closed-form expressions for their quality estimates are also derived by employing the model described in (2.1).

2.4.1 Mutual information based FIQM

Mutual information is a quantity in information theory which is used to measure the amount of information that one random variable contains about another random variable [69]. The mutual information of two continuous random variables x and y is defined as:

$$F_{MI}(x, y) \quad = \quad \int_y \int_x p(x, y) \log_2 \frac{p(x, y)}{p_1(x) p_2(y)} dx dy, \tag{2.8}$$

where $p(x, y)$ is the joint probability density function of x and y, and $p_1(x)$ and $p_2(y)$ are the marginal probability density functions of x and y respectively. Assuming that x, y are

Gaussian random variables, $F_{MI}(x,y)$ can be further expressed as:

$$F_{MI}(x,y) = \frac{1}{2}\log_2\left(\frac{1}{1-\left(\frac{\sigma_{x,y}}{\sigma_x\sigma_y}\right)^2}\right), \tag{2.9}$$

where $\frac{\sigma_{x,y}}{\sigma_x\sigma_y}$ is the correlation coefficient between x and y.

Substituting $F_{MI}(x,y)$ for F in (2.5), $F_{MI}(x,y)$ was employed as the correlation metric in [14] to evaluate the amount of information transferred from the source images to the fused image. From equations (2.1) and (2.4) we can easily tell that f, z_1 and z_2 are Gaussian random variables. Using the model in (2.1) we can calculate (2.5) using results from [69] as shown in [11]. After this we substitute equation (2.8) into equation (2.5) and we can calculate the mutual information based FIQM as [11]

$$Q_{MI} = \frac{1}{2}\log_2\left(\frac{\left((w_1\beta_1+w_2\beta_2)^2\sigma_s^2+w_1^2\sigma_1^2+w_2^2\sigma_2^2\right)^2}{w_1^2w_2^2\left(\beta_1^2\sigma_s^2\sigma_2^2+\beta_2^2\sigma_s^2\sigma_1^2+\sigma_1^2\sigma_2^2\right)^2}\right) + \frac{1}{2}\log_2\left(\beta_1^2\sigma_s^2+\sigma_1^2\right) + \frac{1}{2}\log_2\left(\beta_2^2\sigma_s^2+\sigma_2^2\right). \tag{2.10}$$

2.4.2 UIQI based FIQM

The Universal Image Quality Index (UIQI) was proposed by Wang and Bovik in [70] to evaluate the similarity of two images. This metric outperforms the traditional mean squared error (MSE) based image quality metric because it considers the structural distortions of source images including loss of correlation, luminance distortion and contrast distortion [70]. To calculate the UIQI, we use

$$F_{UIQI}(x,y) = \frac{4\sigma_{x,y}\bar{x}\bar{y}}{(\sigma_x^2+\sigma_y^2)(\bar{x}^2+\bar{y}^2)} \tag{2.11}$$

to replace F in (2.5), where $\bar{x}, \sigma_x^2, \bar{y}, \sigma_y^2$ denote the mean values and variances of x and y respectively and $\sigma_{x,y}$ means the covariance between x and y.

In [12] and [15] the authors use the UIQI to assess the performance of image fusion algorithms. Using the model in (2.1) to calculate the quantities in (2.11) and then substituting equation (2.11) into equation (2.5), we can obtain the closed-form expression of UIQI based image fusion FIQM

$$Q_{UIQI} = \frac{\left(\left(w_1\beta_1 + w_2\beta_2\right)\beta_1\sigma_s^2 + w_1\sigma_1^2\right)\left(w_1\beta_1 + w_2\beta_2\right)\beta_1}{\left(\left(w_1\beta_1 + w_2\beta_2\right)^2\sigma_s^2 + w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + \beta_1^2\sigma_s^2 + \sigma_1^2\right)\left(\left(w_1\beta_1 + w_2\beta_2\right)^2 + \beta_1^2\right)} + \frac{\left(\left(w_1\beta_1 + w_2\beta_2\right)\beta_2\sigma_s^2 + w_2\sigma_2^2\right)\left(w_1\beta_1 + w_2\beta_2\right)\beta_2}{\left(\left(w_1\beta_1 + w_2\beta_2\right)^2\sigma_s^2 + w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + \beta_2^2\sigma_s^2 + \sigma_2^2\right)\left(\left(w_1\beta_1 + w_2\beta_2\right)^2 + \beta_2^2\right)}. \quad (2.12)$$

2.4.3 Edge based FIQM

Considering the fact that the human visual system (HVS) is quite sensitive to the edge information in an image, in [15] the authors suggested replacing F in (2.5) with

$$F_{UIQI}(x', y'), \quad (2.13)$$

where x' and y' denote the edge images corresponding to x and y, respectively. The "edge image" is defined as the Euclidean norm of the horizontal and vertical gradient edge information which can be achieved by applying a horizontal gradient convolution mask and a vertical gradient convolution mask to the original image respectively.

However, the Euclidean norm operation makes it difficult to analytically calculate the statistical characteristics (mean values, variances and covariance) of x' and y'. To derive the closed-form expression of Q_E, we focus on a simplified case in which only the horizontal gradient edge information is used to calculate the edge image. This approximation is reasonable if we consider the situation where the edge in the horizontal direction is much more obvious

than that in the vertical direction. In this case Q_E can be expressed as

$$Q_E = \frac{\left(\left(w_1\beta_1 + w_2\beta_2\right)\beta_1\sigma_s'^2 + w_1\sigma_1'^2\right)\left(w_1\beta_1 + w_2\beta_2\right)\beta_1}{\left(\left(w_1\beta_1 + w_2\beta_2\right)^2\sigma_s'^2 + w_1^2\sigma_1'^2 + w_2^2\sigma_2'^2 + \beta_1^2\sigma_s'^2 + \sigma_1'^2\right)\left(\left(w_1\beta_1 + w_2\beta_2\right)^2 + \beta_1^2\right)} + \frac{\left(\left(w_1\beta_1 + w_2\beta_2\right)\beta_2\sigma_s'^2 + w_2\sigma_2'^2\right)\left(w_1\beta_1 + w_2\beta_2\right)\beta_2}{\left(\left(w_1\beta_1 + w_2\beta_2\right)^2\sigma_s'^2 + w_1^2\sigma_1'^2 + w_2^2\sigma_2'^2 + \beta_2^2\sigma_s'^2 + \sigma_2'^2\right)\left(\left(w_1\beta_1 + w_2\beta_2\right)^2 + \beta_2^2\right)} \quad (2.14)$$

by taking the steps explained to calculate Q_{UIQI}.

Notice that the expression for Q_E is almost the same as that in equation (2.12), except that σ_s^2, σ_1^2 and σ_2^2 are replaced by $\sigma_s'^2$, $\sigma_1'^2$ and $\sigma_2'^2$ which denotes the variances of the signal and the noise components of the pixels in the edge images. Detailed definitions of $\sigma_s'^2$, $\sigma_1'^2$ and $\sigma_2'^2$ and the derivation of equation (2.14) are omitted due to space limits.

Here we should note that all these three FIQMs are direct functions of the correlation coefficient. According to equations (2.9) and (2.11), Q_{MI} and Q_{UIQI} can be expressed as functions of the correlation coefficient. After we obtain the edge information, we calculate Q_E using exactly the same approach used to calculate Q_{UIQI}. Thus Q_E is also correlation-based. It just simply operates on images preprocessed by an edge filter. This is why we call these three the correlation-based FIQMs in this chapter.

2.5 Intuitive Explanations

In this section we explain intuitively the poor behavior of the aforementioned three correlation-based image fusion quality measures. Let us start with the mutual information based quality measure Q_{MI}.

According to equation (2.9), $F_{MI}(f, z_k)$ $(k = 1, 2)$ can be calculated as:

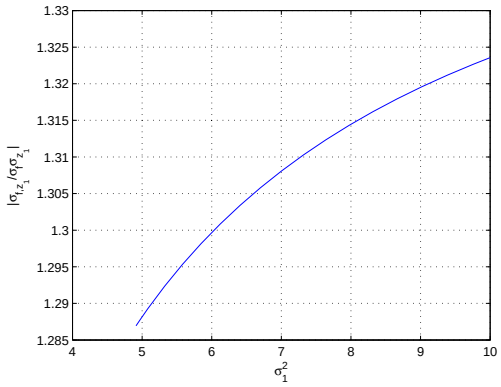
$$F_{MI}(f, z_k) = \frac{1}{2}\log_2\left(\frac{1}{1-\left(\frac{\sigma_{f,z_k}}{\sigma_f\sigma_{z_k}}\right)^2}\right), \tag{2.15}$$

where $\frac{\sigma_{f,z_k}}{\sigma_f\sigma_{z_k}}$ is the correlation coefficient between f and z_k. Apparently $\left|\frac{\sigma_{f,z_k}}{\sigma_f\sigma_{z_k}}\right| \in (0, 1)$, so we have that $F_{MI}(f, z_k)$ is monotonically increasing with $\left|\frac{\sigma_{f,z_k}}{\sigma_f\sigma_{z_k}}\right|$ which can be written as:

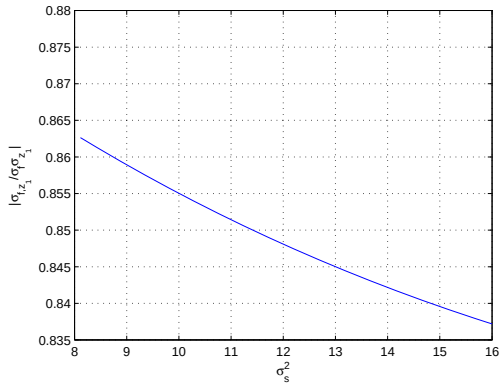
$$\left|\frac{\sigma_{f,z_k}}{\sigma_f\sigma_{z_k}}\right| = \frac{|\beta_k(w_1\beta_1 + w_2\beta_2)\sigma_s^2 + w_k\sigma_k^2|}{\sqrt{\beta_k^2\sigma_s^2 + \sigma_k^2}\sqrt{(w_1\beta_1 + w_2\beta_2)^2\sigma_s^2 + w_1^2\sigma_1^2 + w_2^2\sigma_2^2}}, k = 1, 2. \tag{2.16}$$

Since $F_{MI}(f, z_1)$, the first term of Q_{MI} in (2.5), depends on $\left|\frac{\sigma_{f,z_1}}{\sigma_f\sigma_{z_1}}\right|$ and $F_{MI}(f, z_2)$, the second term of Q_{MI} in (2.5), is determined by $\left|\frac{\sigma_{f,z_2}}{\sigma_f\sigma_{z_2}}\right|$, we can discuss how Q_{MI} changes with σ_1^2 and σ_s^2 by just analyzing $\left|\frac{\sigma_{f,z_1}}{\sigma_f\sigma_{z_1}}\right|$ and $\left|\frac{\sigma_{f,z_2}}{\sigma_f\sigma_{z_2}}\right|$.

2.5.1 The change in Q_{MI} with increasing σ_1^2



(a) $\left|\frac{\sigma_{f,z_1}}{\sigma_f\sigma_{z_1}}\right|$ increases with increasing σ_1^2, for $w_1 = 0.9501$; $w_2 = 0.0499$; $\beta_1 = 0.2311$; $\beta_2 = 0.7689$; $\sigma_s^2 = 6.0684$; $\sigma_2^2 = 8.913$.



(b) $\left|\frac{\sigma_{f,z_1}}{\sigma_f\sigma_{z_1}}\right|$ decreases with increasing σ_s^2, for $w_1 = 0.6412$; $w_2 = 0.3588$; $\beta_1 = 0.0162$; $\beta_2 = 0.8369$; $\sigma_1^2 = 6.9778$; $\sigma_2^2 = 4.6189$.

Figure 2.3: $\left|\frac{\sigma_{f,z_1}}{\sigma_f\sigma_{z_1}}\right|$ increases with increasing σ_1^2 and decreases with increasing σ_s^2.

When σ_1^2 increases, $\left|\frac{\sigma_{f,z_2}}{\sigma_f\sigma_{z_2}}\right|$ always decreases because its numerator doesn't change and

its denominator keeps increasing. On the other hand, from equation (2.16) we can see that the numerator and denominator of $\left|\frac{\sigma_{f,z_1}}{\sigma_f \sigma_{z_1}}\right|$ both increase when σ_1^2 increases. If the numerator part dominates, then $\left|\frac{\sigma_{f,z_1}}{\sigma_f \sigma_{z_1}}\right|$ increases with increasing σ_1^2 which is shown in Fig.2.3(a). Now when the first term in (2.5), $F_{MI}(f, z_1)$, increases faster than the second term, $F_{MI}(f, z_2)$, decreases, then we can conclude that Q_{MI} increases with increasing noise power σ_1^2. Some sufficient conditions are omitted due to space limits. Intuitively, if a larger amount of noise is added to a source image, then this same noise ends up in the fused image when weighted averaging image fusion is employed (see (2.4)). This can increase the correlation between the source and fused image, one of the terms in (2.5), as we have just explained, which can thus increase (2.5).

2.5.2 The change in Q_{MI} with increasing σ_s^2

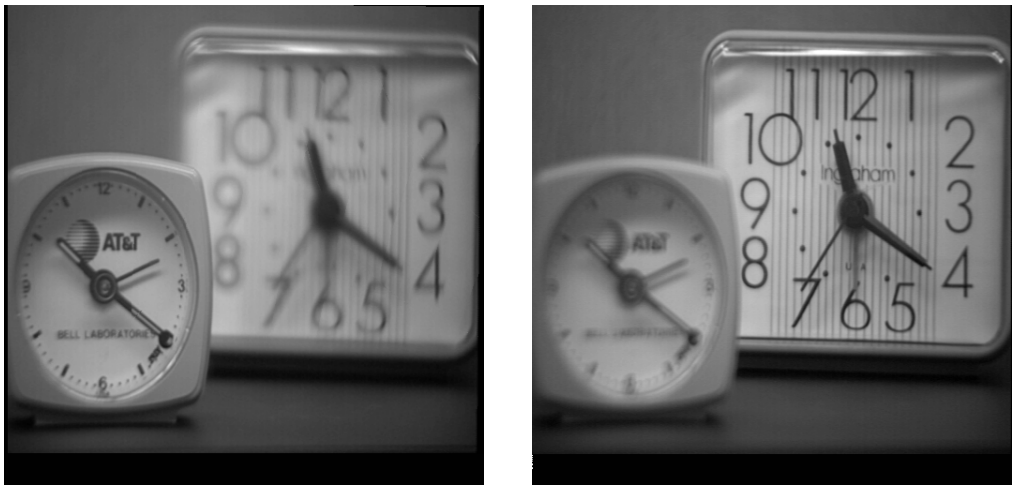
When σ_s^2 increases, the numerator and denominator of $\left|\frac{\sigma_{f,z_k}}{\sigma_f \sigma_{z_k}}\right|$ both increase. If the denominator part increases faster, then $\left|\frac{\sigma_{f,z_k}}{\sigma_f \sigma_{z_k}}\right|$ decreases as shown in Fig.2.3(b). When a decreasing term in (2.5) dominates, Q_{MI} decreases with increasing σ_s^2. Some sufficient conditions are omitted due to space limits.

For the UIQI based FIQM, equation (2.11) can be expanded as:

$$F_{UIQI}(f, z_k) = \frac{4\sigma_{f,z_k}}{\sigma_f^2 + \sigma_{z_k}^2} \frac{(w_1\beta_1 + w_2\beta_2)\beta_k}{(w_1\beta_1 + w_2\beta_2)^2 + \beta_k^2}, k = 1, 2. \tag{2.17}$$

Notice that $F_{UIQI}(f, z_k)$ is monotonically increasing with a similar quantity $\frac{\beta_k \sigma_{f,z_k}}{\sigma_f^2 + \sigma_{z_k}^2}$ which increases with the noise power and decreases with the power of the signal of interest in some cases. Therefore the poor behavior of the UIQI based FIQM can be explained similarly. The same explanation also applies to the edge based FIQM because Q_E and Q_{UIQI} have almost the same closed-form expression.

2.6 Experimental Results Using Real Images



(a) image 1: focus on the left (b) image 2: focus on the right

Figure 2.4: Test clock images for a multifocus image fusion application.

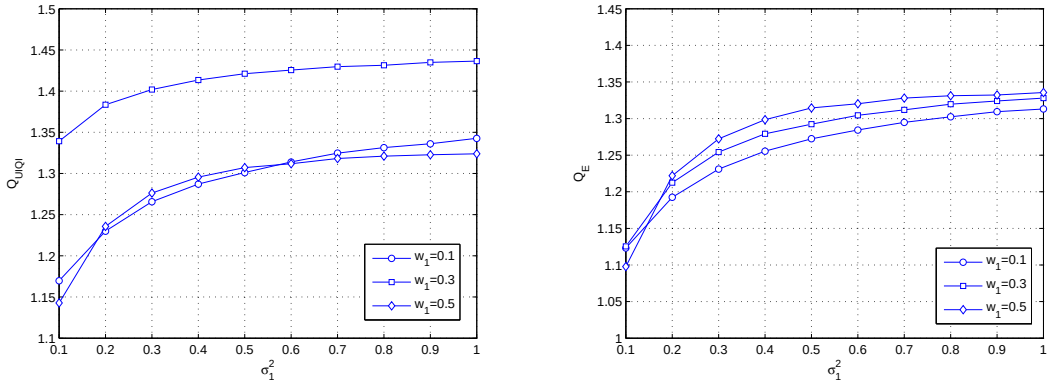
Finally, we demonstrate the utility of our theoretical analysis by observing the behavior of the correlation-based quality measures for real images with respect to **Desired Behavior 1** and **Desired Behavior 2**. The proofs of Theorem 1 and Theorem 2 in the Appendix predict that if the noise power in one input image is large enough, then the correlation-based quality measure increases when more noise is added to that input image. Here we show that such undesired behavior can be observed in real tests by adding Gaussian noise to each input image and evaluating the correlation-based quality measure.

We consider three different image fusion applications, which include CWD [8, 9], night vision [71] and multi-focus image fusion [2]. The three sets of images shown in Fig.1.1, Fig.2.1 and Fig.2.4 are used as the input images. For each input image set, we add Gaussian noise with variance σ_1^2 and σ_2^2 to image 1 and image 2 respectively. The fused image is obtained by using the weighted averaging approach in (2.4). For simplicity, we apply weight w_1 on every

pixel of image 1 and weight $w_2 = 1 - w_1$ on every pixel of image 2.

Since the effects of noise on the mutual information based quality measure Q_{MI} is already well studied in [11], here we only provide experimental results for the UIQI based quality measure Q_{UIQI} and the edge based quality measure Q_E. We calculate the quality measures using the exact procedures recommended by their inventors. Q_{UIQI} is evaluated using the procedure proposed in [15]. A sliding window is run over the whole image pixel by pixel. For each area within the sliding window, the mean values of the subimage and the other statistical parameters in equation (2.11) are estimated to calculate a local quality measure. Then the overall quality measure Q_{UIQI} is computed by averaging all local quality measures. The window size is 10 by 10 in our examples. For the edge based quality measure, we first use a Sobel edge detector to obtain the edge images for the input and fused images. Then the same procedure used for Q_{UIQI} is applied to the edge images.

Fig.2.5 shows the undesired behavior of Q_{UIQI} and Q_E for the CWD application when more noise is added to the first input image (Fig.1.1 (a)). As shown in Fig.2.5 (a), the noise power of the second input image σ_2^2 is fixed at 0.09227, and the noise power of the first input image σ_1^2 varies from 0.1 to 1. We choose different weights $w_1 = 0.1, 0.3, 0.5$, but Q_{UIQI} always indicates better image fusion quality when σ_1^2 is larger which corresponds to higher noise power. We can see from Fig.2.5 (b) that Q_E also increases if σ_1^2 increases from 0.1 to 1 when σ_2^2 is fixed at 0.03322. Fig.2.6 shows that for the multifocus application, Q_{UIQI} and Q_E increase with σ_1^2 when σ_2^2 is fixed at 0.05341 and 0.02972 respectively. Fig.2.2 shows the same trend of Q_{UIQI} and Q_E for the night vision application when we set σ_2^2 to be 0.07059 and 0.08652. These experimental results match our theoretical analysis very well.

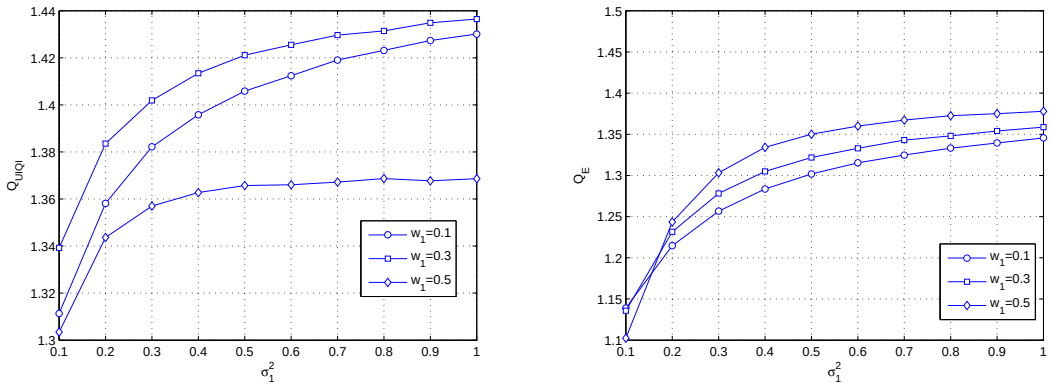


(a) $Q_{UIQI}, \sigma_2^2 = 0.09227$ (b) $Q_E, \sigma_2^2 = 0.03322$

Figure 2.5: Q_{UIQI} and Q_E with increasing σ_1^2 for concealed weapon detection (Fig.1.1).

2.7 Summary

This chapter generalizes the work in [11] which shows bad behavior for an information-based quality measure when the input images are more distorted. The generalization holds for a class of correlation-based quality measures. We focus on the properties of three popular correlation-based quality measures when used to judge the quality of weighted averaging image fusion algorithms. By employing a model for the input images, closed-form expressions for these quality measure are derived. Further, these quality measures are shown to indicate higher quality in some cases where lower quality is evident. Sufficient conditions for these cases are provided, along with intuitive explanations on why this bad behavior occurs. Quality calculations with real images, using the procedures recommended by their inventors, also document the predicted bad behavior of the correlation-based image fusion quality measures, which demonstrates the utility of our theoretical analysis.



(a) $Q_{UIQI}, \sigma_2^2 = 0.05341$ (b) $Q_E, \sigma_2^2 = 0.02972$

Figure 2.6: Q_{UIQI} and Q_E with increasing σ_1^2 for multifocus image fusion (Fig.2.4).

2.8 Appendix

Proof of Theorem 1 and Theorem 2 *Proof:* To prove that the three correlation-based quality measures violate the desire behavior, we just need to show that they increase with increasing noise power σ_1^2 and decrease with increasing σ_s^2 in some cases.

2.8.1 Analytical results for Q_{MI}

In [11], the authors have already proven that Q_{MI} increases with increasing σ_1^2 in some cases. As for the second desired behavior, taking the partial derivative of Q_{MI} with respect to σ_s^2, we obtain

$$\frac{\partial Q_{MI}}{\partial(\sigma_s^2)} = \frac{A\sigma_s^6 + B\sigma_s^4 + C\sigma_s^2 + D}{2\ln(2)\,E}, \tag{2.18}$$

where

$$\begin{aligned}
A &= 2\,\beta_2^2\beta_1^2\left(w_1\beta_1+w_2\beta_2\right)^2\left(\sigma_1^2\beta_2^2+\beta_1^2\sigma_2^2\right)\\
B &= \left(6\,\beta_1^2\sigma_2^2\sigma_1^2\beta_2^2+\sigma_1^4\beta_2^4+\beta_1^4\sigma_2^4\right)\left(w_1\beta_1+w_2\beta_2\right)^2\\
C &= w_1\beta_1^3w_2\beta_2\sigma_1^2\sigma_2^4+6w_1\beta_1w_2\beta_2^3\sigma_1^4\sigma_2^2+2\beta_1^4\sigma_2^4w_1^2\sigma_1^2+3w_1^2\sigma_1^4\beta_2^2\beta_1^2\sigma_2^2\\
&\quad +2\beta_2^4\sigma_1^4w_2^2\sigma_2^2+3w_2^2\sigma_2^4\beta_1^2\sigma_1^2\beta_2^2-w_2^2\sigma_2^6\beta_1^4\\
D &= \sigma_1^2\sigma_2^2\left(\left(\sigma_2w_2\sigma_1\beta_2+\sigma_2w_1\sigma_1\beta_1\right)^2-\left(\sigma_2^2w_2\beta_1-w_1\sigma_1^2\beta_2\right)^2\right)\\
E &= \left(\beta_1^2\sigma_s^2\sigma_2^2+\beta_2^2\sigma_s^2\sigma_1^2+\sigma_1^2\sigma_2^2\right)\left(\left(w_1\beta_1\sigma_s+w_2\beta_2\sigma_s\right)^2\right.\\
&\quad \left.+w_1^2\sigma_1^2+w_2^2\sigma_2^2\right)\left(\beta_1^2\sigma_s^2+\sigma_1^2\right)\left(\beta_2^2\sigma_s^2+\sigma_2^2\right).
\end{aligned}$$

It can be easily seen that $E>0$. So the sign of $\frac{\partial Q_{MI}}{\partial(\sigma_s^2)}$ depends on the sign of the numerator part of equation (2.18). As $A>0$ and $\sigma_s^2>0$, $A\sigma_s^6+B\sigma_s^4+C\sigma_s^2+D$ must be positive if σ_s^2 is large enough.

However, according to equation (2.19), when

$$\left(\sigma_2w_2\sigma_1\beta_2+\sigma_2w_1\sigma_1\beta_1\right)^2-\left(\sigma_2^2w_2\beta_1-w_1\sigma_1^2\beta_2\right)^2<0 \tag{2.19}$$

we have $D<0$. Thus, there must exist very small positive σ_s^2 which makes $A\sigma_s^6+B\sigma_s^4+C\sigma_s^2$ approach zero. Thus $\frac{\partial Q_{MI}}{\partial(\sigma_s^2)}$ is negative and Q_{MI} decreases with increasing σ_s^2 in such cases.

2.8.2 Analytical results for Q_{UIQI}

Equation (2.12) can be rewritten as:

$$Q_{UIQI}=\frac{A_1\sigma_s^2+B_1}{C_1\sigma_s^2+D_1}+\frac{A_2\sigma_s^2+B_2}{C_2\sigma_s^2+D_2}, \tag{2.20}$$

where

$$\begin{aligned}
A_1 &= (w_1\beta_1 + w_2\beta_2)^2 \beta_1^2 \\
A_2 &= (w_1\beta_1 + w_2\beta_2)^2 \beta_2^2 \\
B_1 &= w_1\sigma_1^2\beta_1 (w_1\beta_1 + w_2\beta_2) \\
B_2 &= w_2\sigma_2^2\beta_2 (w_1\beta_1 + w_2\beta_2) \\
C_1 &= \left((w_1\beta_1 + w_2\beta_2)^2 + \beta_1^2\right)^2 \\
C_2 &= \left((w_1\beta_1 + w_2\beta_2)^2 + \beta_2^2\right)^2 \\
D_1 &= \left(w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + \sigma_1^2\right)\left((w_1\beta_1 + w_2\beta_2)^2 + \beta_1^2\right) \\
D_2 &= \left(w_1^2\sigma_1^2 + w_2^2\sigma_2^2 + \sigma_2^2\right)\left((w_1\beta_1 + w_2\beta_2)^2 + \beta_2^2\right).
\end{aligned}$$

Take the partial derivative of Q_{UIQI} with respect to σ_s^2 to obtain

$$\frac{\partial Q_{UIQI}}{\partial(\sigma_s^2)} = \frac{E\sigma_s^4 + F\sigma_s^2 + G}{(C_1\sigma_s^2 + D_1)^2 (C_2\sigma_s^2 + D_2)^2}, \tag{2.21}$$

where

$$\begin{aligned}
E &= A_1D_1C_2^2 - C_2B_2C_1^2 + A_2C_1^2D_2 - C_1B_1C_2^2 \\
F &= 2\,A_2C_1D_1D_2 - 2\,C_1B_1C_2D_2 + 2\,A_1D_1C_2D_2 - 2\,C_2B_2C_1D_1 \\
G &= A_2D_1^2D_2 - C_1B_1D_2^2 + A_1D_1D_2^2 - C_2B_2D_1^2.
\end{aligned}$$

Obviously, the denominator of equation (2.21) is positive. Thus the sign of $\frac{\partial Q_{UIQI}}{\partial(\sigma_s^2)}$ depends on the numerator part of equation (2.21). When $E < 0$, $\frac{\partial Q_{UIQI}}{\partial(\sigma_s^2)}$ is negative for sufficiently large σ_s^2.

Equation (2.12) can be rewritten as

$$Q_{UIQI} = \frac{A_1'\sigma_1^2 + B_1'}{C_1'\sigma_1^2 + D_1'} + \frac{B_2'}{C_2'\sigma_1^2 + D_2'}, \tag{2.22}$$

where

$$\begin{aligned}
A_1' &= (w_1\beta_1 + w_2\beta_2)w_1\beta_1 \\
B_1' &= (w_1\beta_1 + w_2\beta_2)^2 \beta_1^2\sigma_s^2 \\
B_2' &= (w_1\beta_1 + w_2\beta_2)^2 \beta_2^2\sigma_s^2 + (w_1\beta_1 + w_2\beta_2)\, w_2\beta_2\sigma_2^2 \\
C_1' &= \left((w_1\beta_1 + w_2\beta_2)^2 + \beta_1^2\right)\left(w_1^2 + 1\right) \\
C_2' &= \left((w_1\beta_1 + w_2\beta_2)^2 + \beta_2^2\right) w_1^2 \\
D_1' &= \left((w_1\beta_1 + w_2\beta_2)^2 + \beta_1^2\right)^2 \sigma_s^2 + \left((w_1\beta_1 + w_2\beta_2)^2 + \beta_1^2\right) w_2^2\sigma_2^2 \\
D_2' &= \left((w_1\beta_1 + w_2\beta_2)^2 + \beta_2^2\right)^2 \sigma_s^2 + \left((w_1\beta_1 + w_2\beta_2)^2 + \beta_2^2\right)\left(w_2^2 + 1\right)\sigma_2^2.
\end{aligned}$$

Take the partial derivative of Q_{UIQI} with respect to σ_1^2 to obtain

$$\frac{\partial Q_{UIQI}}{\partial(\sigma_1^2)} = \frac{E'\sigma_1^4 + F'\sigma_1^2 + G'}{\left(C_1'\sigma_1^2 + D_1'\right)^2\left(C_2'\sigma_1^2 + D_2'\right)^2}, \tag{2.23}$$

where

$$\begin{aligned}
E' &= A_1'D_1'C_2'^2 - B_2'C_2'C_1'^2 - C_1'B_1'C_2'^2 \\
F' &= 2\,A_1'D_1'C_2'D_2' - 2\,C_1'B_1'C_2'D_2' - 2\,B_2'C_2'C_1'D_1' \\
G' &= C_1'B_1'D_2'^2 + A_1'D_1'D_2'^2 - C_2'B_2'D_1'^2.
\end{aligned}$$

Similarly, the sign of $\frac{\partial Q_{UIQI}}{\partial(\sigma_1^2)}$ depends on the numerator part of equation (2.23). When $E' < 0$, $\frac{\partial Q_{UIQI}}{\partial(\sigma_1^2)}$ is negative for sufficiently large σ_1^2. When $E' > 0$, $\frac{\partial Q_{UIQI}}{\partial(\sigma_1^2)}$ is positive for sufficiently large σ_1^2.

2.8.3 Analytical results for Q_E

The "edge image" is defined as the Euclidean norm of the horizontal and vertical gradient edge information. The edge information can be obtained by applying a horizontal gradient 2-D convolution mask and a vertical gradient 2-D convolution mask to the original image respectively. The resulting edge image is described by

$$i'(m,n) = \sqrt{\left(\sum_{p=-L}^{L}\sum_{q=-L}^{L} h_x(p,q)i(m-p,n-q)\right)^2 + \left(\sum_{p=-L}^{L}\sum_{q=-L}^{L} h_y(p,q)i(m-p,n-q)\right)^2} \quad (2.24)$$

where $i'(m,n)$ denotes the pixel at location (m,n) of the edge image, $i(m,n)$ denotes the corresponding quantity for the original image, and the $(2L+1)\times(2L+1)$ matrices $h_x(p,q)$ and $h_y(p,q)$ implement the horizontal and vertical gradient operations respectively.

Substituting $i(m,n)$ in equation (2.24) with either an observed source image from equation (2.1) or the fused image from (2.4), we can write the edge information of the sensor images and the fused image as

$$\begin{aligned} z_k'(m,n)^2 &= \left(\sum_{p=-L}^{L}\sum_{q=-L}^{L} h_x(p,q)\big(\beta_k s(m-p,n-q) + n_k(m-p,n-q)\big)\right)^2 \\ &\quad + \left(\sum_{p=-L}^{L}\sum_{q=-L}^{L} h_y(p,q)\big(\beta_k s(m-p,n-q) + n_k(m-p,n-q)\big)\right)^2, k = 1,2 \end{aligned} \quad (2.25)$$

$$\begin{aligned} f'(m,n)^2 &= \Big(\sum_{p=-L}^{L}\sum_{q=-L}^{L} h_x(p,q)\big((w_1\beta_1 + w_2\beta_2)s(m-p,n-q) + w_1 n_1(m-p,n-q) \\ &\quad + w_2 n_2(m-p,n-q)\big)\Big)^2 + \Big(\sum_{p=-L}^{L}\sum_{q=-L}^{L} h_y(p,q)\big((w_1\beta_1 + w_2\beta_2)s(m-p,n-q) \\ &\quad + w_1 n_1(m-p,n-q) + w_2 n_2(m-p,n-q)\big)\Big)^2 \end{aligned} \quad (2.26)$$

Here we assume β_k, w_k, μ_s, σ_s^2 and σ_k^2 $(k = 1,2)$ are constant within the small region

surrounding location (m, n).

To obtain $F_{UIQI}(z'_k, f')$ $(k = 1, 2)$, we need to calculate some statistical characteristics of z'_k and f', such as their mean values, variances and covariance. We first simplify equations (2.25) and (2.26) by using only the horizontal gradient edge value to replace the magnitude of the horizontal and vertical gradient edge information. This approximation is reasonable if we consider the case in which the edge along the horizonal direction is much more obvious than the edge in the vertical direction. We will derive and analyze the closed-form expression of Q_E for this simplified case. Numerical results which show the undesired behavior of the general Q_E will also be provided next.

Under the stated assumption, z'_k and f' can be written as linear combinations of Gaussian random variables as

$$z'_k(m,n) = \beta_k s'(m,n) + n'_k(m,n), \tag{2.27}$$

$$f'(m,n) = (w_1\beta_1 + w_2\beta_2)s'(m,n) + w_1 n'_1(m,n) + w_2 n'_2(m,n), \tag{2.28}$$

where

$$\begin{aligned} s'(m,n) &= \sum_{p=-L}^{L}\sum_{q=-L}^{L} h_x(p,q)s(m-p,n-q), \\ n'_k(m,n) &= \sum_{p=-L}^{L}\sum_{q=-L}^{L} h_x(p,q)n_k(m-p,n-q), k = 1,2. \end{aligned}$$

According to the distributions of $s(m-p, n-q)$ and $n_k(m-p, n-q)$ $(-L \leqslant p \leqslant L, -L \leqslant q \leqslant L)$, we have

$$s'(m,n) \sim N(\mu'_s, \sigma'^2_s), \tag{2.29}$$

$$n'_k(m,n) \sim N(0, \sigma'^2_k), k = 1,2. \tag{2.30}$$

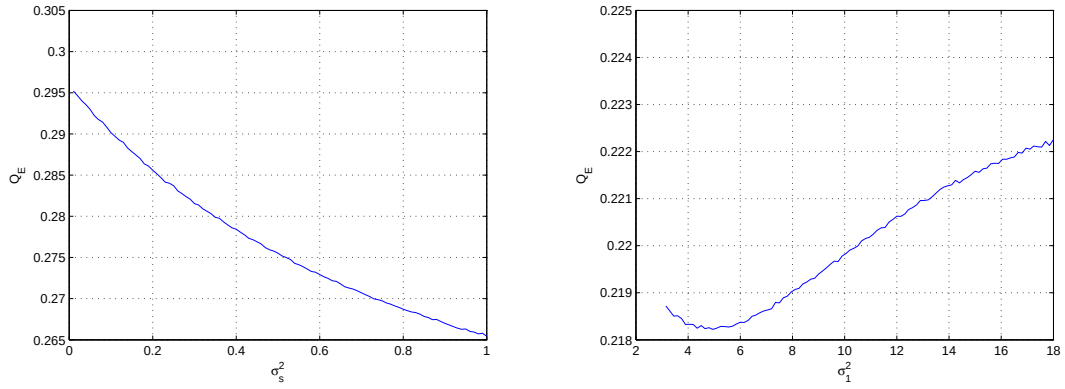
where

$$\mu'_s = \mu_s \sum_{p=-L}^{L} \sum_{q=-L}^{L} h_x(p,q),$$

$$\sigma'^2_s = \sigma_s^2 \sum_{p=-L}^{L} \sum_{q=-L}^{L} h_x(p,q)^2, \tag{2.31}$$

$$\sigma'^2_k = \sigma_k^2 \sum_{p=-L}^{L} \sum_{q=-L}^{L} h_x(p,q)^2, k = 1, 2. \tag{2.32}$$

Equations (2.27), (2.28), (2.29) and (2.30) have the same forms as equations (2.1)-(2.4). Thus Q_E has the same closed-form expression as Q_{UIQI} (equation (2.12)) except that σ_s, σ_1 and σ_2 in equation (2.12) are replaced by σ'_s, σ'_1 and σ'_2.



(a) Q_E decreases with increasing σ_s^2, for $w_1 = 0.9498; w_2 = 0.0502; \beta_1 = 0.8518; \beta_2 = 0.7595; \sigma_1^2 = 0.7117; \sigma_2^2 = 29.8089$.

(b) Q_E increases with increasing σ_1^2, for $w_1 = 0.4387; w_2 = 0.5613; \beta_1 = 0.9901; \beta_2 = 0.7889; \sigma_s^2 = 0.0498; \sigma_2^2 = 32.1746$.

Figure 2.7: Numerical results for the undesired behavior of the general Q_E.

We can see that although the images are filtered by an edge detection operator, the noise still occurs in the edge information of the input and fused images inevitably. According to equations (2.31) and (2.32), increasing σ_s^2 or σ_k^2 will directly increase σ'^2_s or σ'^2_k. We can conclude that the edge based quality measure Q_E have the same problem as Q_{UIQI} when the power of the desired signal or noise varies.

For the general Q_E which uses the Euclidean norm of the horizontal and vertical gradient edge information, although the closed-form expression is hard to get, here we can provide some numerical results to show that it also violates the desired behavior 1 and 2.

Consider a gradient edge detector with horizontal gradient operator $h_x = [1, -1]$ and vertical gradient operator $h_y = [1, -1]^T$. Given realizations of Gaussian random variables s, n_1 and n_2, f', z_1' and z_2' in equations (2.25) and (2.26) can be calculated and their statistics can be estimated using the Monte Carlo method. We calculate Q_E for different parameters and obtain the two curves plotted in Fig.2.7. In Fig.2.7(a) Q_E decreases with increasing σ_s^2. In Fig.2.7(b) Q_E increases with increasing σ_1^2 when σ_1^2 is larger than 5. The values of other parameters are shown in the caption part of these two figures. ■

Chapter 3

A New Method for Evaluation of FIQMs

3.1 Introduction

Measuring the performance of image fusion algorithms is an extremely important task, which has received past study [10–25]. The performance of image fusion algorithms is primarily assessed by perceptual evaluation in the form of subjective human tests [16]. Typically in these tests, human observers are asked to view a series of fused images and rate them. Because images are fused for better human interpretation, it is more important to judge fusion methods by how well humans are able to perform interpretation tasks. Examples of human interpretation studies for image fusion evaluations appear in [20, 25]. No mater the goal of the human perception test, these tests are inconvenient, expensive and time consuming.

It is clearly highly desirable to identify an objective performance measure that can accurately predict human perception by determining the quality of the fused image. The objective

measure should be a feature that is obtained via an automatic computation employing the fused image and can serve as a surrogate for human perception results. We refer to such a feature as the *fused image quality measure* (FIQM). If a good FIQM can be devised, then one can compare image fusion algorithms without expensive perception experiments. Furthermore, the measure can be used as a design criteria for an "optimal" image fusion algorithm.

In the literature, three broad classes of FIQMs have been proposed. The first class represents full-reference measures. They require a reference fused image (or the ground truth image) that represents the "ideal" image of the scene. Once the ground truth image is given, one can use existing quality metrics such as the mean square error, the peak signal to noise ratio, or more sophisticated measures such as structure similarity [72] to compare the fused images with the reference. In the image compression application, the uncompressed image represents the ideal, and it has been demonstrated that the structure similarity is a meaningful full-reference measure [72]. For the image fusion application, it is only possible to generate a reference image for some special cases (for instance, the multi-focus image fusion [11]). In most cases, one has to resort to other classes of FIQMs that do not require a reference image. The second class of FIQMs represents source comparative measures that utilize partial information about the scene, e.g., the source images that were collected and utilized as input to the image fusion process. This class of FIQMs has recently received a great deal of attention [12–15]. These measures quantify the amount of information transferred from the source images to the fused image by considering the sum of correlations between each source image and the fused image. An analysis of this class of FIQMs is provided in [17]. The third class of FIQMs represents no-source comparative measures. These measures attempt to extract the salient features, such as the structure, texture, contrast and edge information, directly from the fused image without regard to the source images [20–24].

Quantitatively evaluating the image fusion performance is a complicated issue because of the lack of a complete understanding of the human visual system (HVS), and because of the variety of image fusion applications [18]. We expect that the FIQM should be task specific, and the best measure changes from task to task. Given an image fusion application and many kinds of proposed FIQMs, we are interested in which quality measure better describes the performance of the human interpreting the fused imagery.

Ideally, the FIQM for a given image would reveal how well a human can interpret the image for a given task, i.e., it can predict human performance. One can achieve this aim by inventing a measure that linearly fits the human perception performance. In [73], the authors have shown an evidence of the approximately linear fitness between image quality (IQ) measures and the subjective rating of image distortions. However, an image is a projection of a particular scene, and the context in the scene, i.e., the relationship of the objects in the scene, can affect the ability of human to perform a *particular task* (target detection for example). Since the linearity is a stricter requirement than monotonicity for a FIQM and is harder to achieve under various context, we believe that it will be more difficult to guarantee linearity when the IQ is used to predict the ability of a human to interpret the image for a given task. Thus, we focus on the monotonicity criterion in this chapter.

By monotonicity we mean that a realistic FIQM can determine the relative ranking of human performance over a series of fused images derived from the same exact source images, which we now refer to as a scene. For a given scene, as FIQM increases over a series of fused images, human performance over these images should also increase. If the human performance is consistently decreasing, the measure is still good as it can be trivially transformed into a proper FIQM via a reciprocal operation. Thus, a potential FIQM should be judged by how well a monotonic function (ascending or descending) explains the relation between the

FIQM and human performance over a variety of fused imagery representing the same scene. In addition, the nature of the monotonic relationship (ascending or descending) should be consistent from scene to scene. Overall, a statistic that quantifies how well different FIQMs are consistent with actual human performance is necessary.

This chapter focuses on scoring FIQMs for the case of the detection task. Performance is measured by the probability that a human observer can correctly detect certain objects in the fused image. The human perception experiments measure the number of observers that are able to correctly detect ground truthed targets as the human performance. This performance metric can be reasonably modeled by a binomial distribution. This chapter introduces a new monotonic statistic for the object detection task that is applicable when the underlying perception results are derived from a small number of human observers. To handle a small number of observers, this statistic does not make Gaussian assumptions about the performance measurements.

Previous work does exist to objectively score the effectiveness of FIQMs. In [19], Pearson (or linear) correlation and root mean squared error (RMSE) are used to score potential FIQMs. The Pearson correlation quantifies how well a straight line fits the mapping between the input and output sequences. Unfortunately, when the relationship between the quality measure and the human performance is nonlinear, the value of Pearson correlation can be small despite the fact that the sequences are still monotonically related. In essence, a proper statistic needs to determine if the ordering of a quality measure preserves the ordering of the corresponding human performance measures.

The Spearman and Kendall correlations [74, 75] are common statistics to quantify how well the output sequence is ordered. In fact, the Spearman correlation has been used to evaluate the quality measures for video streams [76]. Both quantities are invariant to monotonic

transformations of both the input and output sequences because monotonic transformations preserve the rank order of the sequences. For evaluation of the utility of FIQMs, a miss-ordering of human performance values that are nearly identical should not lower the correlation value too much. Because only ranks and not actual values are considered, the reduction in correlation score due to these miss-orderings can be identical or even greater than that of miss-orderings of widely varying human performance values. Clearly, measurement noise can greatly impact the correlation scores. Therefore, these rank-order correlations are not appropriate for seeking out good FIQMs.

In [72, 76], a nonlinear regression fit to a logistic function followed by linear correlation is used to compare various FIQMs in order to accommodate the nonlinear, but monotonic, relationships. Recently, the monotonic correlation (MC), which uses isotonic regression followed by linear correlation has been proposed in [20]. As demonstrated in [20], the MC is more flexible than linear correlation or the logistic analysis in [72,76]. Like linear and logistic correlation, the MC assumes that the perception error is Gaussian, which is inappropriate for the detection task when the number of observers is small.

To our knowledge, this chapter represents the first attempt to score the effectiveness of FIQMs for the detection task in light of practical measurements from human perception experiments. To this end, the chapter develops a novel statistic to test whether or not a monotonic relationship exists between the proposed FIQM and the human performance. The monotonic statistic is general and can be applied to other applications when one may need to test for a monotonic relationship.

The chapter is organized as follows. Section 3.2 presents the perception model and introduces the new monotonic statistic. Section 3.3 demonstrates the effectiveness of the new statistic via Monte Carlo simulations. The statistic is used to score potential FIQMs against

actual perception results for fused images in Section 3.4. Finally, Section 4.5 provides some concluding remarks.

3.2 Statistical Monotonic Analysis

The chapter focuses on the detection task and measures the performance of image fusion algorithms by the probability that a human observer can correctly detect certain objects in the fused image. This section develops the test statistic that compares the hypothesis that the relationship between human detection performance and FIQM values are monotonic to the hypothesis that the relationship is random. The statistic is based on the model that each image exhibits a ground truth quality score, which is the probability that any human can detect the object in it. Section 3.2.1 derives the likelihoods for each hypothesis conditioned on these ground truth quality scores. Then, Section 3.2.2 uses an uninformative prior for the ground truth quality scores to define the likelihood ratio so that it is computationally feasible as demonstrated in Section 3.2.3. Finally, Section 3.2.4 presents properties of the test statistic.

3.2.1 Data Models

A scene is a realization of F source images, and we consider N different fusion algorithms. The existence (or lack) of a monotonic relationship between measured human performance and computed FIQMs can be inferred over S scenes. To this end, this subsection provides the data models that enable this inference.

For a given scene, let the $N \times 1$ vector $\tilde{\mathbf{p}} = (\tilde{p}_1, \tilde{p}_2, \ldots, \tilde{p}_N)^T$ denote the actual performance for all fusion methods, where $\tilde{p}_i$ is the object detection probability, i.e., the ground

truth quality score, associated with the image obtained from the i-th fusion algorithm. Let a given FIQM evaluated over N fusion algorithms be denoted as a $N \times 1$ vector $\tilde{\mathbf{x}} = (\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_N)^T$. The computed value $\tilde{x}_i$ is a deterministic function of the image obtained from the i-th fusion algorithm and the F source images. The proposed monotonic hypothesis test evaluates how well a FIQM monotonically relates to human object detection performance. Under the monotonic hypothesis, there is a monotonic function that maps the measure value $\tilde{x}_i$ associated with the i-th fusion algorithm to the detection probability $\tilde{p}_i$, i.e.,

$$\tilde{p}_i = g(\tilde{x}_i), \tag{3.1}$$

where $g(x)$ is a monotonic increasing or decreasing function of x. Let $\mathbf{p}$ and $\mathbf{x}$ denote a reordering of $\tilde{\mathbf{p}}$ and $\tilde{\mathbf{x}}$ such that the measure values are in ascending order. i.e. $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_N$. Note that $\mathbf{p} = P_k\tilde{\mathbf{p}}$ and $\mathbf{x} = P_k\tilde{\mathbf{x}}$ where P_k is one of a possible $N!$ permutation matrices. This chapter uses the convention that P_1 is the identity matrix and $P_{N!}$ reverses the original ordering, i.e., the anti-diagonal matrix of ones. Now, we consider two alternative H_1 hypotheses: $H_\uparrow$ for ascending p_i's and $H_\downarrow$ for descending p_i's. On the other hand, the null hypothesis is that over the ensemble of possible fused imagery, the $\tilde{x}_i$'s are *i.i.d.* samples. Thus, the p_i's are in random order where the probability of any permutation of the order is equal. In other words, P_k is the permutation matrix that orders the p_i's for the H_1 hypotheses, and P_k is randomly chosen via a uniform distribution over the $N!$ possible permutation matrices under the null (H_0) hypothesis. Namely, the conditional probability mass functions (pmfs) of the permutations conditioned on $\tilde{\mathbf{p}}$ and the hypotheses for $k = 1, \ldots, N!$ are

$$f_\pi(P_k|\tilde{\mathbf{p}}, H_\uparrow) = \begin{cases} 1 & \text{if } P_k\tilde{\mathbf{p}} \in \mathcal{P}_\uparrow, \\ 0 & \text{otherwise,} \end{cases} \qquad f_\pi(P_k|\tilde{\mathbf{p}}, H_\downarrow) = \begin{cases} 1 & \text{if } P_k\tilde{\mathbf{p}} \in \mathcal{P}_\downarrow, \\ 0 & \text{otherwise,} \end{cases} \tag{3.2}$$

$$f_\pi(P_k|\tilde{\mathbf{p}}, H_0) = \tfrac{1}{N!},$$

where

$$\mathcal{P}_{\uparrow} = \{\mathbf{p} : 0 \leqslant p_1 \leqslant \ldots \leqslant p_N \leqslant 1\}, \quad \mathcal{P}_{\downarrow} = \{\mathbf{p} : 1 \geqslant p_1 \geqslant \ldots \geqslant p_N \geqslant 0\}. \tag{3.3}$$

For this discusion, it is also convenient to define $\mathcal{P}_0$ as the set of all possibe $\mathbf{p}$'s, i.e.,

$$\mathcal{P}_0 = \{\mathbf{p} : 0 \leqslant p_1, \ldots, p_N \leqslant 1\}. \tag{3.4}$$

If $\mathbf{p} = P_k\tilde{\mathbf{p}}$ is observed, then the likelihoods of the hypotheses, i.e., $l(H_i|\mathbf{p}) = f(P_k|\tilde{\mathbf{p}}, H_i)$ for $i \in \{\uparrow, \downarrow, 0\}$ demonstrate that if $\mathbf{p}$ is not in ascending (or descending) order, then the ascending (or descending) likelihood (and likelihood ratio) is zero, and the $H_{\uparrow}$ (or $H_{\downarrow}$) hypothesis must be incorrect. Also, if $\mathbf{p}$ happens to be in ascending (or descending) order, then either the $H_{\uparrow}$ (or $H_{\downarrow}$) hypothesis is true or the ordering of $\mathbf{p}$ is due to random luck under the null hypothesis, which occurs with a probability of $1/N!$. Thus, for $\mathbf{p} \in \mathcal{P}_{\uparrow}$ (or $\mathbf{p} \in \mathcal{P}_{\downarrow}$), the likelihood ratio is not infinite, i.e., a sure monotonic relationship. Rather, it is $N!$ due to the fact that the random $\mathbf{x}$ can order $\mathbf{p}$ by chance.

Unfortunately, the value of $\tilde{\mathbf{p}}$ (or $\mathbf{p}$) is unobservable. It can only be inferred via perception experiments that measure $\mathbf{y} = (y_1, y_2, \ldots, y_N)^T$ where y_i is the number of observers that correctly detect the targets in the image obtained from the i-th fusion algorithm.[1] We use o_i to represent the number of observers that participate in the detection experiment for the image formed by the i-th fusion image. Under the assumption that all human are equally capable, it is reasonable to model $\mathbf{y}$ as a random vector whose elements are statistically independent where y_i is drawn from a binomial distribution with parameters o_i and p_i so that the pmf of $\mathbf{y}$ conditioned on $\mathbf{o}$ and $\mathbf{p}$ is

$$\mathbf{y} \sim f_y(\mathbf{y}|\mathbf{o}, \mathbf{p}) = \prod_{i=1}^{N} \binom{o_i}{y_i} p_i^{y_i}(1-p_i)^{o_i - y_i}. \tag{3.5}$$

[1]For variables that do not use the tilde, the indices for the images is such that x_i's are in ascending order.

Here we represent the o_i's in an $N \times 1$ vector $\mathbf{o}$ for notational convenience. Since $\mathbf{p} = P_k\tilde{\mathbf{p}}$, one can define $f_y(\mathbf{y}|\mathbf{o}, P_k, \tilde{\mathbf{p}}) = f_y(\mathbf{y}|\mathbf{o}, \mathbf{p})$.

The joint pmf of the observations $\mathbf{y}$ and the permutations P_k can be written as

$$f_{y\pi}(\mathbf{y}, P_k|\mathbf{o}, \tilde{\mathbf{p}}, H_i) = f(\mathbf{y}|P_k, \mathbf{o}, \tilde{\mathbf{p}}, H_i) f(P_k|\mathbf{o}, \tilde{\mathbf{p}}, H_i). \tag{3.6}$$

Because $\mathbf{y}$ conditioned on $\mathbf{o}$ and $\tilde{\mathbf{p}}$ is independent of H_i, $f(\mathbf{y}|P_k, \mathbf{o}, \tilde{\mathbf{p}}, H_i) = f_y(\mathbf{y}|P_k, \mathbf{o}, \tilde{\mathbf{p}})$ for all H_i's. Furthermore, $f(P_k|\mathbf{o}, \tilde{\mathbf{p}}, H_i) = f_\pi(P_k|\tilde{\mathbf{p}}, H_i)$ because P_k does not depend on $\mathbf{o}$. Thus, $f_{y\pi}(\mathbf{y}, P_k|\mathbf{o}, \tilde{\mathbf{p}}, H_i)$ is obtained by the multiplication of (3.2) and (3.5) so that

$$\begin{aligned}
f_{y\pi}(\mathbf{y}, P_k|\mathbf{o}, \tilde{\mathbf{p}}, H_\uparrow) &= \begin{cases} f_y(\mathbf{y}|\mathbf{o}, P_k, \tilde{\mathbf{p}}) & \text{if } P_k\tilde{\mathbf{p}} \in \mathcal{P}_\uparrow, \\ 0 & \text{otherwise,} \end{cases} \\
f_{y\pi}(\mathbf{y}, P_k|\mathbf{o}, \tilde{\mathbf{p}}, H_\downarrow) &= \begin{cases} f_y(\mathbf{y}|\mathbf{o}, P_k, \tilde{\mathbf{p}}) & \text{if } P_k\tilde{\mathbf{p}} \in \mathcal{P}_\downarrow, \\ 0 & \text{otherwise,} \end{cases} \\
f_{y\pi}(\mathbf{y}, P_k|\mathbf{o}, \tilde{\mathbf{p}}, H_0) &= \tfrac{1}{N!} f_y(\mathbf{y}|\mathbf{o}, P_k, \tilde{\mathbf{p}}).
\end{aligned} \tag{3.7}$$

Then, a hypothesis test to distinguish $H_\uparrow$ or $H_\downarrow$ from H_0 using the observed values can be derived from the likelihoods $l(H_i|\mathbf{y}, \mathbf{o}, P_k, \tilde{\mathbf{p}}) = f_{y\pi}(\mathbf{y}, P_k|\mathbf{o}, \tilde{\mathbf{p}}, H_i)$. Because $\mathbf{p} = P_k\tilde{\mathbf{p}}$ is not observed, the hypothesis test is a composite test. It is unclear whether a uniformly most powerful (UMP) test exists. A common test to apply is the generalized likelihood ratio test (GLRT). This requires one to compute the maximum likelihood (ML) estimates $\hat{\mathbf{p}}_\uparrow$, $\hat{\mathbf{p}}_\downarrow$, $\hat{\mathbf{p}}_0$ for the $H_\uparrow$, $H_\downarrow$, and H_0 hypotheses, respectively. For the two H_1 hypotheses, the ML estimates can be obtained by the $O(N)$ pool adjacent violators algorithm [20, 77, 78]. For the null hypothesis, $\hat{p}_i = \frac{y_i}{o_i}$. The GLRT has the property that for any ascending (or descending) $\mathbf{y}$, the ascending (or descending) generalized likelihood ratio (GLR) is $N!$. However, if the y_i's are close in values, the ordering is more likely to be due to luck than when the y_i's are

well spread. However, the GLRT is unable to make this distinction between different ordered $\mathbf{y}$'s. A different approach that accounts for the relative spread of the observations values is needed.

3.2.2 Diffuse Prior Monotonic Likelihood Ratio Test

A given scene is a realization from the ensemble of possible source images. Therefore, it is reasonable to model the detection probabilities as being drawn from a random distribution, i.e., $\tilde{\mathbf{p}} \sim f_{\tilde{p}}(\tilde{\mathbf{p}})$. The *diffuse prior monotonic likelihood ratio test* (DPMLRT) assumes that for a given scene, $\tilde{\mathbf{p}}$ is a realization of an uninformative (or diffuse) prior distribution, i.e., the elements $\tilde{p}_i$ are *i.i.d.* uniform $[0, 1)$ so that $f_{\tilde{p}}(\tilde{\mathbf{p}}) = 1$. The uniform distribution models the fact that the imagery are collected in various conditions where the ability to detect the objects can be easy, hard, or somewhere in between. The independence between fusion methods is a simplifying assumption that leads to a computationally feasible test. Because the prior on $\tilde{\mathbf{p}}$ is independent of the hypothesis H_i and $\mathbf{o}$, we have $f(\tilde{\mathbf{p}}|\mathbf{o}, H_i) = f_{\tilde{p}}(\tilde{\mathbf{p}}) = 1$. Then, $\tilde{\mathbf{p}}$ is marginalized so that the expected likelihood for the i-th hypothesis is

$$\tilde{l}(H_i|\mathbf{y}, \mathbf{o}, P_k) = \int_{\mathcal{P}_0} f_{y\pi}(\mathbf{y}, P_k|\mathbf{o}, \tilde{\mathbf{p}}, H_i) f(\tilde{\mathbf{p}}|\mathbf{o}, H_i) d\tilde{\mathbf{p}} = \int_{\mathcal{P}_0} f_{y\pi}(\mathbf{y}, P_k|\mathbf{o}, \tilde{\mathbf{p}}, H_i) d\tilde{\mathbf{p}}. \tag{3.8}$$

Now the expected likelihoods do not depend on any unobservable parameters. The integral in (3.8) can be simplified by noting that the integrand is given by (3.7) and using the change of variable $\tilde{\mathbf{p}} \mapsto P_k^{-1}\mathbf{p}$. Then, it is easy to see that

$$\tilde{l}(H_\uparrow|\mathbf{y}, \mathbf{o}, P_k) = \textstyle\int_{\mathcal{P}_\uparrow} f_y(\mathbf{y}|\mathbf{o}, \mathbf{p}) d\mathbf{p}, \qquad \tilde{l}(H_\downarrow|\mathbf{y}, \mathbf{o}, P_k) = \textstyle\int_{\mathcal{P}_\downarrow} f_y(\mathbf{y}|\mathbf{o}, \mathbf{p}) d\mathbf{p}, \tag{3.9}$$

$$\tilde{l}(H_0|\mathbf{y}, \mathbf{o}, P_k) = \tfrac{1}{N!} \textstyle\int_{\mathcal{P}_0} f_y(\mathbf{y}|\mathbf{o}, \mathbf{p}) d\mathbf{p}.$$

Now, the tests to distinguish the H_1 hypotheses from the null hypothesis are simple

hypothesis tests, and the likelihood ratio test (LRT) is the most powerful test. Namely, given that for each scene $\tilde{\mathbf{p}}$ is drawn from the uninformative prior, then the following LRTs are optimal in the Neyman-Pearson sense [79] for distinguishing the monotonically ascending or descending hypothesis from the null hypothesis,[2]

$$\lambda_N^{\uparrow}(\mathbf{y},\mathbf{o}) = \frac{N!\int_{\mathcal{P}_{\uparrow}} f(\mathbf{y}|\mathbf{o},\mathbf{p})d\mathbf{p}}{\int_{\mathcal{P}_0} f(\mathbf{y}|\mathbf{o},\mathbf{p})d\mathbf{p}}, \quad \lambda_N^{\downarrow}(\mathbf{y},\mathbf{o}) = \frac{N!\int_{\mathcal{P}_{\downarrow}} f(\mathbf{y}|\mathbf{o},\mathbf{p})d\mathbf{p}}{\int_{\mathcal{P}_0} f(\mathbf{y}|\mathbf{o},\mathbf{p})d\mathbf{p}}. \tag{3.10}$$

We refer to $\lambda_N^{\uparrow}$ and $\lambda_N^{\downarrow}$ as the ascending and descending diffuse prior monotonic likelihood ratio (DPMLR), respectively.

For multiple scenes, the nature of the monotonicity (ascending or descending) should be consistent from scene to scene. Therefore, one must consider the cumulative likelihoods for the ascending, descending, and null hypotheses. Since we assume that the $\mathbf{y}$'s and $\mathbf{p}$'s are statistically independent from scene to scene, the likelihoods for each hypothesis accumulate via the product operation. The cumulative likelihood ratios are then proportional to the geometric mean of the likelihood ratios for each scene. The geometric mean provides a convenient way to normalize the score against the number of scenes. The overall likelihood ratio for the monotonic relationship over S scenes is formally defined as

$$\Lambda_N = \left(\max \left\{ \prod_{s=1}^{S} \lambda_N^{\uparrow}(\mathbf{y}_s,\mathbf{o}_s), \prod_{s=1}^{S} \lambda_N^{\downarrow}(\mathbf{y}_s,\mathbf{o}_s) \right\} \right)^{1/S}, \tag{3.11}$$

where $\mathbf{y}_s$ and $\mathbf{o}_s$ are the number of correct detections and observers for the s-th scene, respectively. Note that Λ_N is agnostic to the nature of the monotonicity. Unless it is required, the scene index is implicit for the sake of notational brevity. We refer to Λ_N as the composite DPMLR. When $\Lambda_N > 1$ the evidence in support of the monotonic hypothesis is greater than that of the null hypothesis where the FIQM behaves as noise with respect to human

[2]For notational convenience, the dependency of λ to the ordering P_k is left implicit since λ is actually invariant to P_k except in how it orders $\mathbf{y}$.

performance. As Λ_N increases, so does the evidence that the FIQM under test is actually a good measure. The DPMLRT is simply accepting the monotonic hypothesis if the DPMLR exceeds a given threshold value. Usually, the threshold is greater than one.

3.2.3 Recursive Computation

To our knowledge, a closed form expression for (3.10) does not exist, and numerical integration quickly becomes infeasible as N increases. Fortunately, it is possible to calculate the diffuse likelihood ratios numerically. However, due to the multi-variable integration involved in the expression, the calculation requires large computational cost, especially when N and the o_i's are large. This subsection provides a recursion to calculate these diffuse likelihood ratios.

The diffuse likelihood for H_0 can be simply expressed as:

$$\tilde{l}(H_0|\mathbf{y},\mathbf{o}) = \prod_{i=1}^{N} \binom{o_i}{y_i} \beta(y_i+1, o_i-y_i+1) \tag{3.12}$$

where

$$\beta(a,b) \quad = \quad \int_0^1 z^{a-1}(1-z)^{b-1}\,dz \tag{3.13}$$

is the Beta function.

Substituting equations (3.5), (3.8) and (3.12) into (3.10), the ascending diffuse likelihood ratio can be expressed as:

$$\lambda_N^{\uparrow}(\mathbf{y},\mathbf{o}) = \frac{N!\int_{\mathcal{P}_\uparrow} h(p_N;y_N,o_N)\ldots h(p_1;y_1,o_1)d\mathbf{p}}{\prod_{i=1}^{N}\beta(y_i+1,o_i-y_i+1)}, \tag{3.14}$$

where

$$h(p;y,o) = p^y(1-p)^{o-y}. \tag{3.15}$$

By considering the power series expansion of the regularized incomplete Beta function, the calculation of $\lambda_N^{\uparrow}(\mathbf{y}, \mathbf{o})$ can be simplified in a recursive way. Specifically, the regularized incomplete Beta function is defined as

$$I(y; a, b) = \frac{\int_0^y z^{a-1}(1-z)^{b-1}\, dz}{\beta(a, b)}, \tag{3.16}$$

and the power series expansion for $I(y; a, b)$ is

$$I(y; a, b) = \frac{1}{a+b} \sum_{j=a}^{a+b-1} \frac{1}{\beta(j+1, a+b-j)} y^j (1-y)^{a+b-1-j}. \tag{3.17}$$

Then, (3.14) can be written as:

$$\begin{aligned} & \frac{N! \int_0^1 \ldots \int_0^{p_3} h(p_N; y_N, o_N) \ldots h(p_2; y_2, o_2) \left(\int_0^{p_2} h(p_1; y_1, o_1) dp_1 \right) dp_2 \ldots dp_N}{\beta(y_1+1, o_1-y_1+1) \prod_{i=2}^N \beta(y_i+1, o_i-y_i+1)} \\ = & \frac{N! \int_0^1 \ldots \int_0^{p_3} h(p_N; y_N, o_N) \ldots h(p_2; y_2, o_2) I(p_2; y_1+1, o_1-y_1+1) dp_2 \ldots dp_N}{\prod_{i=2}^N \beta(y_i+1, o_i-y_i+1)} \end{aligned} \tag{3.18}$$

Now substituting (3.17) into (3.18), we obtain:

$$\begin{aligned} \lambda_N^{\uparrow}(\mathbf{y}, \mathbf{o}) &= \frac{N!}{o_1+2} \sum_{j=y_1+1}^{o_1+1} \frac{\beta(j+y_2+1, o_1+o_2+2-y_2-j)}{\beta(j+1, o_1+2-j)\beta(y_2+1, o_2-y_2+1)} \cdot \\ & \quad \frac{\int_0^1 \ldots \int_0^{p_3} h(p_N; y_N, o_N) \ldots h(p_2; j+y_2, o_1+o_2+1)\, dp_2 \ldots dp_N}{\prod_{i=3}^N \beta(y_i+1, o_i-y_i+1)\beta(j+y_2+1, o_1+o_2+2-y_2-j)} \\ &= \frac{N!}{o_1+2} \sum_{j=y_1+1}^{o_1+1} \frac{\beta(j+y_2+1, o_1+o_2+2-y_2-j)}{\beta(j+1, o_1+2-j)\beta(y_2+1, o_2-y_2+1)} \cdot \\ & \quad \lambda_{N-1}^{\uparrow}\left([j+y_2, y_3, \ldots, y_N]', [o_1+o_2+1, o_3, \ldots, o_N]' \right). \end{aligned} \tag{3.19}$$

Also from (3.3), one can see that $\mathcal{P}_{\uparrow}$ and $\mathcal{P}_0$ are the same when $N = 1$. Therefore by definition, we have

$$\lambda_1^{\uparrow}(y_1, o_1) = 1, \tag{3.20}$$

and the ascending diffuse likelihood ratio can be computed numerically via the recursion

defined in (3.19) and (3.20). A similar recursion can compute the descending diffuse likelihood ratio. Alternatively, one can use the symmetry property (see Property 2 in the next subsection) to derive $\lambda_N^{\downarrow}$ from the computation of $\lambda_N^{\uparrow}$.

3.2.4 Properties

The diffuse likelihood ratios demonstrate a number of interesting properties than can easily be proven. Some of these properties are for the general case where the number of observers can vary over the different fused images. Other properties are for the case that the number of observers is constant, i.e., $o_i = o$. This more specific case that $\mathbf{o} = o\mathbf{1}$ is common for perception experiments where one would expect the evaluation of the fused imagery over the same number of observers. In addition to these provable properties, we have discovered other interesting attributes for the DPMLR by exhaustively computing the DPMLRs for all $(o+1)^N$ values of $\mathbf{y}$ for manageable, i.e., small, values of o and N. These attributes make sense based upon the intuition of how the DPMLRT should behave; we speculate that these attributes are preserved for larger values of o and N; and we are willing to go out on a limb by disseminating them as *conjectures* in this subsection. We hope that proofs will be discovered in the future so that the conjectures can become properties.

This section first presents the properties that are valid for general values of $\mathbf{o}$.

Property 1 $\lambda_N^{\uparrow}(\mathbf{y}, \mathbf{o}), \lambda_N^{\downarrow}(\mathbf{y}, \mathbf{o}), \Lambda_N \in (0, N!)$.

The proof of this property can be found in Appendix 3.6.1. The property bounds the possible values of the diffuse likelihood ratios. As the number of objects N to consider increases, the upper bound for the likelihood ratios grows fast. For a given value of N and o, the bounds of zero and $N!$ are loose since the set of all possible values of $\mathbf{y}$ is finite. However,

as demonstrated later in this subsection, as the number of observers increases, one can find a $\mathbf{y}$ that corresponds to a likelihood ratio value that is arbitrarily close to either bound. In other words, as the number of observers increases and the y_i's have sufficient spread, the likelihood ratio becomes as if $\mathbf{p}$ is observable (see Section 3.2.1).

Property 2 $\lambda_N^{\downarrow}(\mathbf{y}, \mathbf{o}) = \lambda_N^{\uparrow}(P_{N!}\mathbf{y}, P_{N!}\mathbf{o}) = \lambda_N^{\uparrow}(\mathbf{o} - \mathbf{y}, \mathbf{o})$.

Proof: The first equality is the result of a simple change of variables $\mathbf{p} \mapsto P_{N!}\mathbf{p}$ in (3.14). Likewise, the second equality is the result of the change of variables $p_i \mapsto 1 - p_i$ for $i = 1, \ldots, N$ in (3.14) followed by a reversal of the order of integration. ■

This property demonstrates a symmetry between $\lambda_N^{\uparrow}$ and $\lambda_N^{\downarrow}$. The symmetry provides a convenient way to derive the descending likelihood ratio via the computation of the ascending likelihood ratio and vice versa.

The first two properties are valid for a variable amount of observers per a fused image. The final set of properties are specific for the case that a constant number of observers o are utilized for the N fused images, i.e., $\mathbf{o} = o\mathbf{1}$.

Property 3 *If* $y_1 = y_2 = \ldots = y_N$, *then* $\lambda_N^{\downarrow}(\mathbf{y}, o\mathbf{1}) = \lambda_N^{\uparrow}(\mathbf{y}, o\mathbf{1}) = 1$.

The proof of this property is given in Appendix 3.6.2. The property states that when all observations are equal, one cannot distinguish between the ascending, descending, and null hypotheses because all orderings of the observations are indistinguishable. Clearly, when all observations are the same, it is an ill-posed problem to determine whether or not the FIQMs are ordering the fused imagery in any special manner.

Property 4 *If the* y_i*'s are in ascending order and they are not constant then* $\lambda_N^{\uparrow}(P_{N!}\mathbf{y}, P_{N!}\mathbf{o}) < \lambda_N^{\uparrow}(P_k\mathbf{y}, P_k\mathbf{o}) < \lambda_N^{\uparrow}(\mathbf{y}, \mathbf{o})$ *for* $1 < k < N!$. *Likewise, if the* y_i*'s are in descending order and*

they are not constant then $\lambda_N^{\downarrow}(P_{N!}\mathbf{y}, P_{N!}\mathbf{o}) < \lambda_N^{\downarrow}(P_k\mathbf{y}, P_k\mathbf{o}) < \lambda_N^{\downarrow}(\mathbf{y}, \mathbf{o})$ *for* $1 < k < N!$.

Property 5 *If the* y_i*'s are in ascending order and they are not constant, then* $\lambda_N^{\uparrow}(\mathbf{y}, \mathbf{o}) > 1$ *and* $\lambda_N^{\downarrow}(\mathbf{y}, \mathbf{o}) < 1$. *Likewise, if the* y_i*'s are in descending order and they are not constant, then* $\lambda_N^{\downarrow}(\mathbf{y}, \mathbf{o}) > 1$ *and* $\lambda_N^{\uparrow}(\mathbf{y}, \mathbf{o}) < 1$.

The proof of these two properties is provided in Appendix 3.6.3. Property 4 states that if the observations demonstrate a perfect monotonic ascending relationship with the FIQM, then the ascending likelihood ratio is larger than that for any other ordering of the observations. Furthermore, the descending order of observations demonstrates the lowest ascending likelihood ratio of all possible orderings. Since it can be shown that the average likelihood ratio over all possible orderings of the observations is one, Property 5 is a corollary of Property 4. The property states that as long as the human performance $\mathbf{y}$ is increasing in concert with $\mathbf{x}$, the diffuse likelihood ratio will always favor the ascending $H_{\uparrow}$ and disfavor the descending $H_{\downarrow}$ hypotheses over the null hypothesis H_0. Similarly, as long as the human performance $\mathbf{y}$ is decreasing in concert with $\mathbf{x}$, the diffuse likelihood ratio will always favor the descending $H_{\downarrow}$ and disfavor the ascending $H_{\uparrow}$ hypotheses over the null hypothesis H_0. Clearly, these two properties are both intuitively appealing.

Conjecture 1 *The product* $\lambda_N^{\uparrow}(\mathbf{y}, o\mathbf{1}) \cdot \lambda_N^{\downarrow}(\mathbf{y}, o\mathbf{1}) \leqslant 1$ *where equality occurs if and only if* $\lambda_N^{\uparrow}(\mathbf{y}, o\mathbf{1}) = \lambda_N^{\downarrow}(\mathbf{y}, o\mathbf{1}) = 1$.

As stated earlier, this conjecture is the result of searching through an exhaustive list of $(o+1)^N$ monotonic likelihood ratio values for manageable values of o and N. This conjecture states that the ascending and descending hypotheses can never both be favored over the null hypothesis. In other words, $\lambda_N^{\uparrow} > 1$ implies $\lambda_N^{\downarrow} < 1$, and $\lambda_N^{\downarrow} > 1$ implies $\lambda_N^{\uparrow} < 1$. However,

the converse is not true. It is possible that for a given $\mathbf{y}$ both $\lambda_N^{\uparrow}$ and $\lambda_N^{\downarrow}$ can be less than one. As a simple example, consider $\mathbf{y} = [\ 0 \quad 2 \quad 0\ ]$ for $o = 2$. Because of the symmetry property, $\lambda_N^{\uparrow} = \lambda_N^{\downarrow}$. At best, a symmetric $\mathbf{y}$ can have a monotonic likelihood ratio of one when all the y_i's are constant. Otherwise, the symmetric $\mathbf{y}$ is neither ascending or descending and should not provide evidence to support $H_{\uparrow}$ or $H_{\downarrow}$ over H_0. For this case, the ascending, descending, and composite DPMLRs are all 0.2286.

Conjecture 2 $\lambda_N^{\uparrow}(\mathbf{y}, o\mathbf{1}) = 1$ *(or* $\lambda_N^{\downarrow}(\mathbf{y}, o\mathbf{1}) = 1$*) if and only if the* y_i*'s are constant.*

This conjecture states that the only way for the ascending (or descending) hypothesis to be indistinguishable from the null hypothesis is for all the observations y_i to be the same. Furthermore, if the ascending hypothesis cannot be distinguished from the null hypothesis then the same is true for the descending hypothesis.

Conjecture 3 *For a given* N*, the bounds in Property 1 are tight in the sense that one can identify a value of* o *and corresponding* $\mathbf{y}$ *whose monotonic likelihood ratio is arbitrarily close to either the lower bound of zero or the upper bound of* $N!$*.*

Inspection of the exhaustive list of monotonic likelihood ratios of possible $\mathbf{y}$'s for small values of N and o has revealed that

$$\overline{y}_i = \left\lfloor \frac{i-1}{N-1} o \right\rfloor \quad \text{and} \quad \underline{y}_i = \begin{cases} o & i < N/2 \\ 0 & i \geqslant N/2 \end{cases} \tag{3.21}$$

achieve close to the maximum and minimum values of $\lambda_N^{\uparrow}$, respectively, for a given value of N and o. A different rounding function in (3.21) may lead to a higher $\lambda_N^{\uparrow}$. Intuitively, as the values of the y_i's spread apart, the discriminability between the hypotheses improves. If the observations use the entire dynamic range of o and they increase linearly with respect

	o				
N	5	10	20	40	80
$\lambda_N^{\uparrow}(\overline{\mathbf{y}}, o\mathbf{1})$					
3	5.27	5.93	6.00	6.00	6.00
5	33.88	71.59	104.38	117.30	119.87
$\lambda_N^{\uparrow}(\underline{\mathbf{y}}, o\mathbf{1})$					
3	1.62e-004	1.55e-008	1.09e-016	3.93e-033	3.71e-066
5	1.16e-007	7.69e-015	2.58e-029	2.12e-058	1.04e-116
$\Lambda_N(\breve{\mathbf{y}}, o\mathbf{1})$					
3	6.17e-003	8.47e-006	1.11e-011	1.41e-023	1.64e-047
5	4.39e-005	9.14e-011	1.71e-022	2.91e-046	4.08e-094

Table 3.1: $\lambda_N^{\uparrow}$ and Λ_N can approach their bounds of zero and $N!$ as the number of observers o increases.

to the rank order, then it makes sense that $\lambda_N^{\uparrow}$ is as large as possible. Since maximizing $\lambda_N^{\uparrow}$ also maximizes Λ_N due to (3.11) and the symmetry property, $\overline{\mathbf{y}}$ also achieves close to the maximum of Λ_N. For a small $\lambda_N^{\uparrow}$, the y_i's should be decreasing and $\underline{\mathbf{y}}$ has the maximum drop possible. While $\underline{\mathbf{y}}$ leads to a small $\lambda_N^{\uparrow}$, its corresponding Λ_N value is greater than one because it is monotonically descending (see (3.11)). The observation sequence

$$\breve{y}_i = (1 - (-1)^i)o/2 \tag{3.22}$$

achieves close to the minimum value of Λ_N for a given value of N and o. It is neither increasing nor decreasing and utilizes the dynamic range of o. Table 3.1 demonstrates how these sequence are converging to the lower and upper bounds for $\lambda_N^{\uparrow}$ and Λ_N as o increases for a given N. The symmetry properties can be used to show similar results for $\lambda_N^{\downarrow}$.

In summary, the evidence to accept the H_1 hypothesis (DPMLR > 1) or the null hypothesis (DPMLR < 1) increases as the number of observers increases because the spread of possible DPMLRs increases. Furthermore, if $\mathbf{y}$ happens to exhibit a perfect monotonic ordering, then the evidence to support H_1 also increases as the spread of the y_i's increases.

In other words, the chances of measurement errors leading to errors in inferring the wrong hypothesis decreases as the number of observers increases. The performance of the DPMLRT in terms of hypothesis errors is evaluated by Monte Carlo simulations in the next section.

3.3 DPMLRT Performance Analysis

In this section, we justify the performance of the proposed DPMLRT. To this end, we generate Monte Carlo realizations of $\mathbf{y}$, $\tilde{\mathbf{x}}$, and $\tilde{\mathbf{p}}$. Specifically, the $\tilde{p}_i$'s are generated uniformly over $[0, 1)$. For the monotonic hypothesis, $\tilde{x}_i = (\tilde{p}_i)^\alpha$. For the null hypothesis, the $\tilde{x}_i$'s are *i.i.d.* from a uniform distribution. For either hypothesis, the y_i's are random realizations of the binomial distribution (see (3.5)). For a given hypothesis and values of $o\mathbf{1}$, N, and α, we generated 10^6 realizations of $\mathbf{y}$, $\tilde{\mathbf{x}}$, and $\tilde{\mathbf{p}}$, and we computed the associated DPMLR given one scene, i.e. $S = 1$. Then, we use the histograms of the DPMLR to generate receiver operating characteristic (ROC) curves by varying the acceptance threshold and tabulating the number of acceptances under the monotonic hypothesis, i.e. probability of detection (P_d), and under the null hypothesis, i.e. probability of false alarms (P_f). As a means of comparison, we also compute ROC curves associated with some other correlation tests in a similar fashion over the same simulations.

Fig. 3.1 includes ROC curves of the DPMLR, the monotonic correlation [20], the Pearson correlation and the logistic correlation [72, 76] tests for various values of α when $N = 10$ and $o = 5$. Interested readers are referred to [20] for a detailed description of the monotonic and logistic correlations. In Fig. 3.1(a) where $\alpha = 1$, the Pearson correlation performs better than the others. This is explained by the fact that the relationship between $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{p}}$ is actually linear, and Pearson correlation exploits the actual values of $\tilde{\mathbf{x}}$ and not just the ordering. In

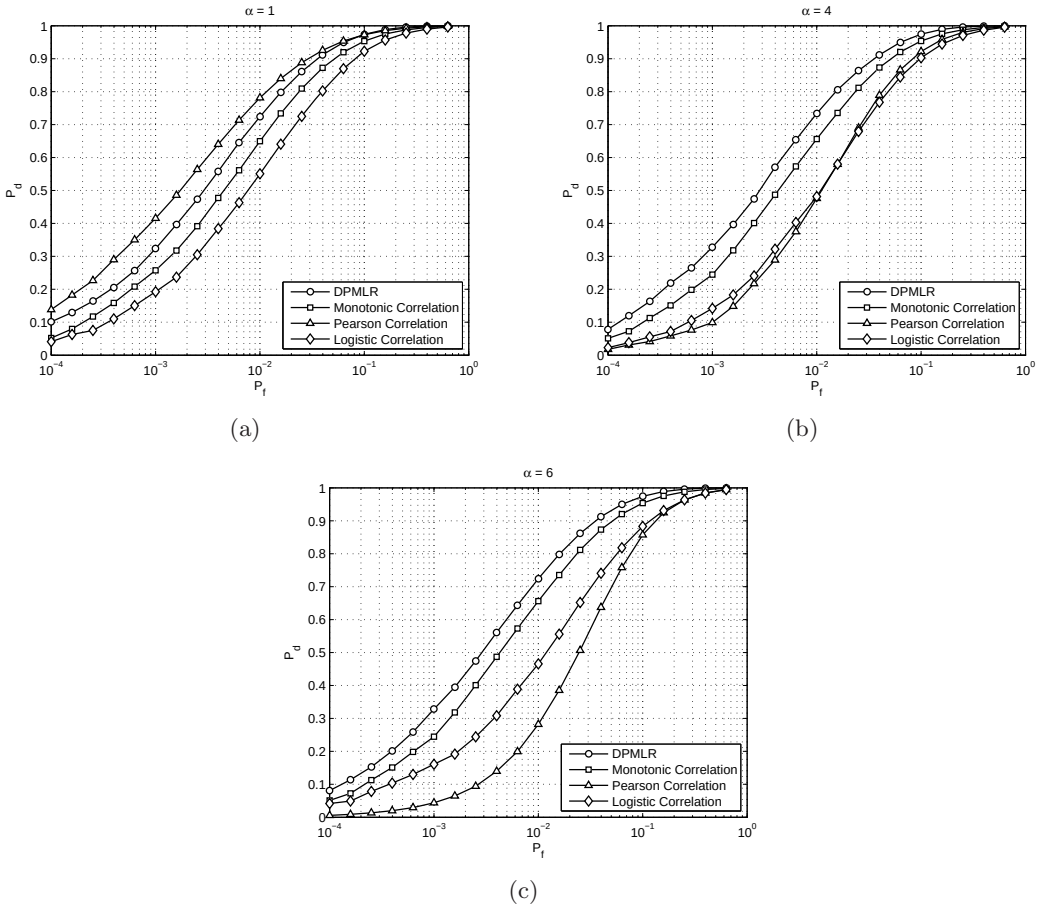
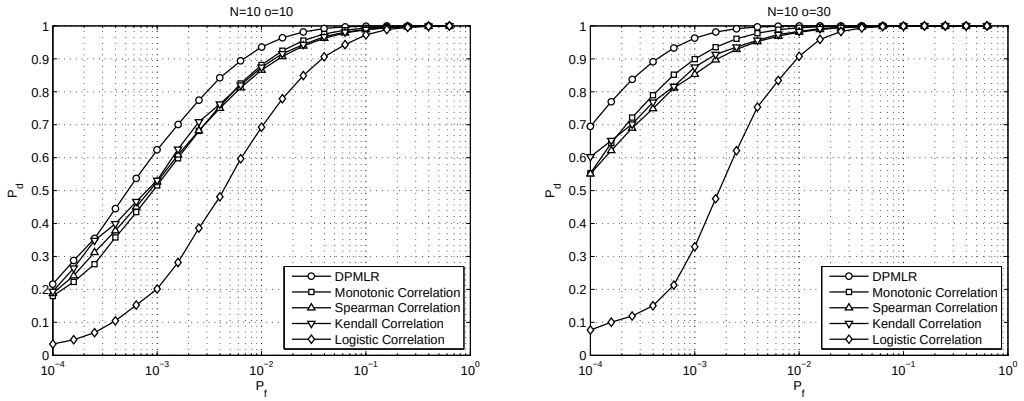


Figure 3.1: ROC curves for DPMLR, monotonic correlation, Pearson correlation and logistic correlation tests: (a) $\alpha = 1$, (b) $\alpha = 4$, and (c) $\alpha = 6$.

essence, the test for linearity is better in this case than the more general test of monotonicity because it exploits more information. As the $g(x)$ function becomes more nonlinear (i.e., α increases), the performance of the Pearson correlation degrades significantly. Clearly, the logistic correlation is more robust to the nonlinearity than the Pearson correlation, but since not all monotonic relations follow a logistic function, the logistic correlation performs worse than the monotonic correlation. Note that α does not change the ordering of $\tilde{x}_i$'s. Therefore, the performance of the DPMLRT and the monotonic correlation is invariant to the nonlinearity. The DPMLRT always outperforms the monotonic correlation because for



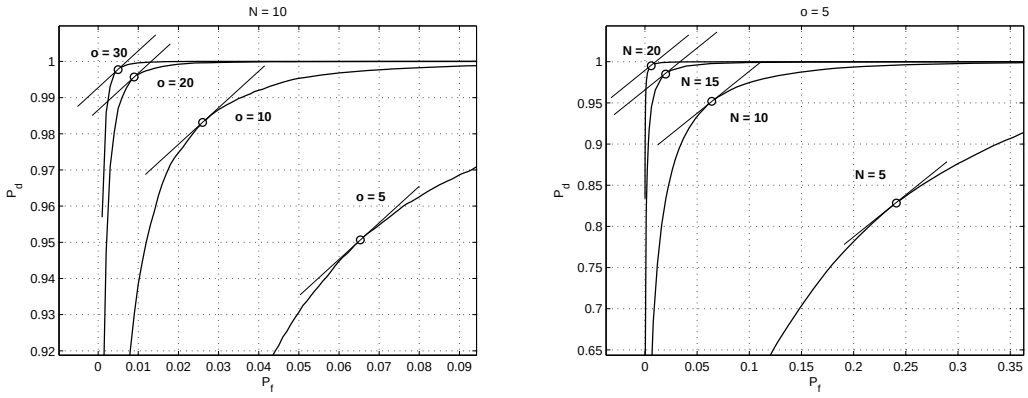
(a) (b)

Figure 3.2: ROC curves for DPMLR, monotonic correlation, Spearman correlation, Kendall correlation and logistic correlation tests: (a) $o = 10$ and (b) $o = 30$.

the case of the uniform prior on $\tilde{\mathbf{p}}$, the DPMLRT is the most powerful test of monotonicity.

We also consider some common rank-order correlations — the Spearman correlation and the Kendall correlation when $N = 10$ and $\alpha = 6$. The ROC curves for different tests are shown in Fig. 3.2. Fig. 3.2(a) corresponds to a case where $o = 30$ and Fig. 3.2(b) corresponds to a case where $o = 10$. The DPMLRT always outperforms the others as expected given that the $\tilde{\mathbf{p}}$'s are generated by the assumed prior distribution. As the number of observers increases, the gap between the ROC curves of the DPMLRT and the rank-order correlation tests becomes larger. When $o = 10$, the performance of the monotonic correlation is a little poorer than that of the rank-order correlations. For larger o ($o = 30$), the monotonic correlation outperforms the rank-order correlations because it takes advantage of the values of y_i's while the rank-order correlations only use their rank information. The logistic correlation exhibits the worst performance because of the limitation of the logistic regression fitting.

Fig. 3.3 provides the ROC curves of the DPMLRT for different o's and N's. The circle on each curve denotes the operating point when the threshold is set to one. As shown in [79],



(a) (b)

Figure 3.3: ROC curves for DPMLRT for various values of N and o: (a) $N = 10$, and $o = 5, 10, 20$ or 30 and (b) $o = 5$ and $N = 5, 10, 15$ or 20.

the slope of the ROC curve for a LRT is equal to the corresponding threshold value. Thus, when the threshold is one, the slope is one corresponding to the "knee" of the ROC curve as demonstrated in Fig. 3.3, which uses a linear scale for the P_f-axis. As one increases the number of observers, the knee of the ROC curve shifts to the top left corner, which means higher P_d and lower P_f can be achieved for a threshold of one. As expected, as the number of fused images N or the number of observers o increases, the efficacy of the DMPLR improves.

The next set of simulations consider how the DPMLR performs when the model assumptions do not match the data. For these simulations, $N = 10$, $o = 10$, and $\alpha = 6$. The first case considers uniform random variables $\tilde{p}_i$'s with a prespecified correlation matrix Σ, whose (m, n)-th element denotes the correlation coefficient of $\tilde{p}_m$ and $\tilde{p}_n$ $(1 \leqslant m, n \leqslant N)$. The method for generating such $\tilde{p}_i$'s is from [80]. In this case we denote the nondiagonal elements of Σ by ρ (the diagonal elements equal 1). The $\tilde{p}_i$'s are completely correlated or independent for $\rho = 1$ or $\rho = 0$, respectively. Fig. 3.4 compares the ROC curves of DPMLR with the other correlations for different ρ's. Fig. 3.4(a), (b) and (c) correspond to $\rho = 0.1, 0.5$ and 0.9,

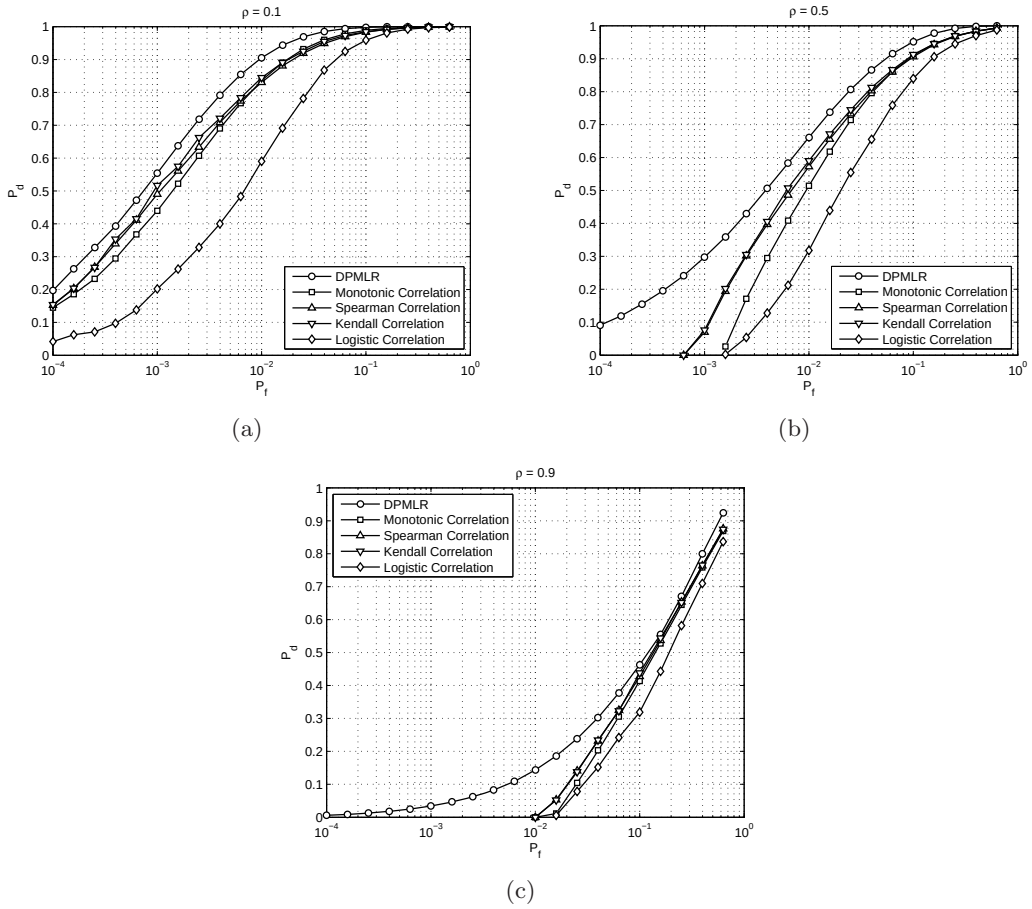
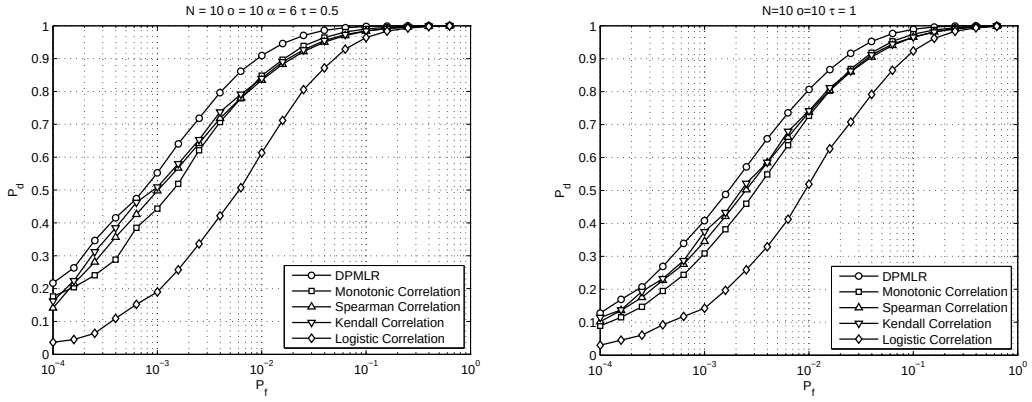


Figure 3.4: ROC curves for correlated $\tilde{p}_i$'s: (a) $\rho = 0.1$, (b) $\rho = 0.5$ and (c) $\rho = 0.9$.

respectively. By comparing these ROC curves to Fig. 3.2, we can see that the gap between the DPMLRT and the others decreases as ρ increases. But clearly the DPMLRT exhibits the best performance among these correlations. In the limit, as ρ goes to 1, the monotonic evaluation is moot as all values of the $\tilde{p}_i$'s are equal.

The next case considers the effect when the model of human performance does not match the binomial distribution. We consider the generalized binomial distribution [81] to incorporate diversity in the capabilities of humans. Specifically, the nominal human performance $\tilde{p}_i$ and associated FIQM $\tilde{x}_i = (\tilde{p}_i)^\alpha$ are generated as usual. Then, the realized mean performance



(a) (b)

Figure 3.5: ROC curves under generalized binomial distribution: (a) $\tau = 0.5$, and (b) $\tau = 1$.

for the observers $\hat{p}_i$ is drawn from the uniform distribution over $[\tilde{p}_i - \tau\tilde{p}_i(1-\tilde{p}_i), \tilde{p}_i + \tau\tilde{p}_i(1-\tilde{p}_i)]$ and y_i is drawn from a binomial distribution with parameters o and $\hat{p}_i$.[3] Here $\tau \in [0, 1]$ is referred to as the spread parameter, which denotes the deviation of y_i's distribution from the binomial distribution. Note that for $\tau = 0$, y_i still follows the binomial distribution with parameters o and $\tilde{p}_i$. Fig. 3.5 shows the ROC curves of the DPMLRT, the monotonic, the rank-order and the logistic correlation tests for different spread parameters τ. This figure demonstrates that the DPMLRT is robust to τ and still outperforms the others even when τ is as large as one.

The final case demonstrates that the DPMLR is not the UMP for any arbitrary prior distribution. Consider a pathological case in which for odd i, $\tilde{p}_i = \tilde{p}_{i+1}$ and $\tilde{p}_i$ is drawn from a uniform distribution over $[0, 1)$. In practice, this case is unlikely because it means that two different fusion methods provide images with equivalent performance over multiple scenes. Nevertheless, Fig. 3.6 compares the DPMLRT with the other correlations. The figure shows that the DPMLRT outperforms the others when the P_d is high. But the other correlations

[3] As discussed in [81], any pmf of y_i over $[0, o]$ can be generated by choosing a specific pdf to generate $\hat{p}_i$.

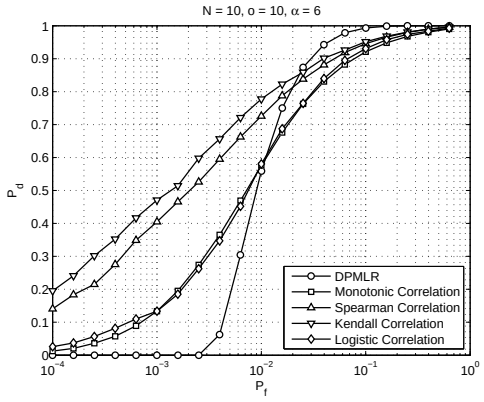


Figure 3.6: An example of the case in which $\tilde{p}_i$'s are at the edge of $\mathcal{P}_{\uparrow}$ under H_1.

all achieve higher detection probability than the DPMLRT when P_f is less than 0.01.

3.4 FIQM Evaluation via the DPMLRT

This section demonstrates the application of the DPMLRT to score potential FIQMs. The DPMLRT and some other correlations are used to evaluate the monotonic relationships between 17 different FIQMs proposed in the literature and the human detection results in a specific target detection experiment. Details of the experiment and the discussion about the evaluation results are provided in the following subsections.

3.4.1 Experimental Setup

Long-wave infrared (LWIR) and image intensified (II) imagery were collected in a simulated military operation in an urban terrain (MOUT) environment. The imagery includes six interior and exterior locations, where four scenarios were collected for each location. The four scenarios represent cases where zero, one, two, and three people are within the field of regard of the camera. Individuals who were in the field of regard were typically obscured

by objects in the scene, such as doorways, windows, furniture, and tables. For each of the scenarios, a horizontal pan of 150 images was then used to create a larger mosaic of imagery in both the LWIR and II bands.

The perceptual goal for the human observers is to detect the target in the scenes by interrogating the fused imagery. To generate the imagery, the LWIR and II images were registered, bore-sighted and fused via six different algorithms: 1) Contrast Pyramid A (CONA), 2) Contrast Pyramid B (CONB) [82], 3) Discrete Wavelet Transform (DWTT) [3, 83, 84] 4) Color Discrete Wavelet Transform (CDWT), 5) Color Averaging (CLAV) and 6) Color Multiscale Transform (CLMT) [85]. The first three algorithms generate grayscale fused images, and the final three methods generate color fused images. It is worth mentioning that the distinction between CONA and CONB is which image (LWIR or II) populates the coarsest coefficients in the pyramid. Also, the color methods generate a grayscale fusion method for the luminance component, map the differences in the image coefficients in the saturation component, and encode the source of the largest coefficient (LWIR vs II) in the hue component. The CDWT uses this coloring scheme for the DWT coefficients, the CLAV uses simple averaging for the luminance and the raw pixels for the color components, and the CLMT uses the coloring scheme for the multiscale fusion method defined in [85]. Finally, it is instructive to compare the fused imagery against the source imagery. Therefore, we consider eight fused image displays: 1) II, 2) LWIR, 3) CONA, 4) CONB, 5) DWTT, 6) CDWT, 7) CLAV and 8) CLMT.

Fig. 3.7 shows an example of the resulting eight fused image displays for a typical scenario in our experiment. In this scenario, there are two target persons which are highlighted by the boxes in each image. As seen in Fig. 3.7(b), the human targets stand out in the LWIR imagery because they are usually hotter than the background. For the most part, detection

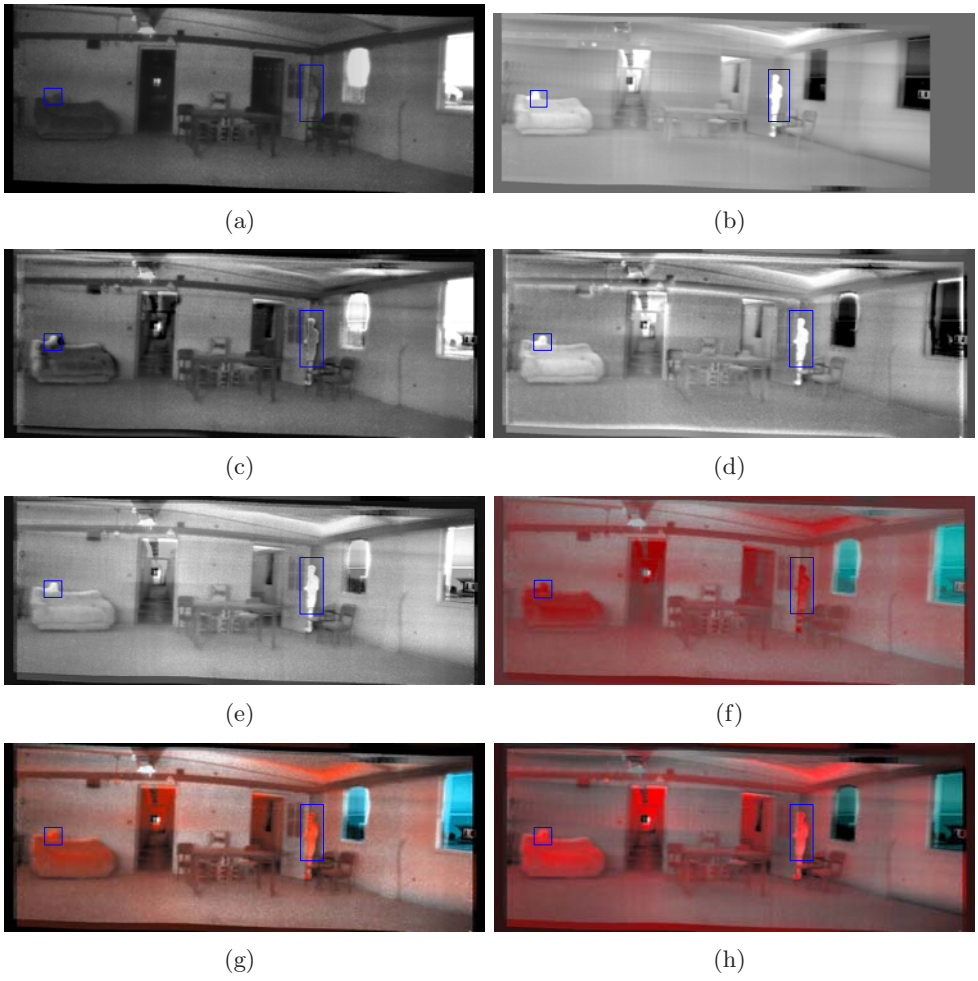
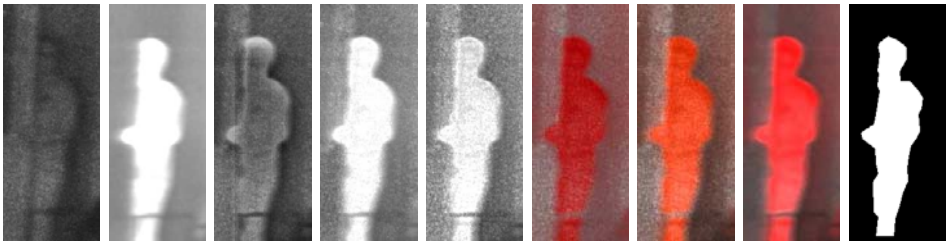


Figure 3.7: Eight fused image displays for one of the 18 scenarios: (a) II, (b) LWIR, (c) CONA, (d) CONB, (e) DWTT, (f) CDWT, (g) CLAV, and (h) CLMT.

performance is best on the LWIR only band because the search task can often be reduced to simply finding the white hot object on a grey background. However, the II band has the potential to add context to the LWIR band as the objects like tables and chairs are easier to distinguish in the II band (see Figs. 3.7(a) and (b)). Therefore, there can be value in fusing the two bands.

A perception test was set up whereby observers were asked to try to find the human



(a) (b) (c) (d) (e) (f) (g) (h) (i)

Figure 3.8: An example of the eight fused image displays and corresponding silhouette for a single scene, i.e., a target instance, in Fig. 3.7: (a) II, (b) LWIR, (c) CONA, (d) CONB, (e) DWTT, (f) CDWT, (g) CLAV, (h) CLMT, and (i) silhouette.

targets in a "field of regard" search. An observer's display was calibrated to look as though it were seeing a single field of regard of a given scene, and the observer had to navigate across the scene and detect human targets. Observers could mark as many as three places on the display as detections for human targets (as they were told that the images could contain between zero and three humans hiding in the scene). At any point an observer could push a button to indicate that they either did not detect any targets in the scene or that there were no other targets in the scene. In the end, the detection performance of the humans was recorded over the eight image displays.

Overall, $o = 8$ observers evaluated 18 scenarios that contained 35 human targets. We treat each target and its surrounding area as a scene for every scenario. For example, the inside of each box in Fig. 3.7 represents a scene, as shown in Fig. 3.8(a)–(h). Then, $\mathbf{y}_s$ is the number of observers that correctly detected the target located in the s-th scene for $s = 1, \ldots, 35$.

3.4.2 The Evaluated FIQMs

We test 17 potential FIQMs over each scene. These FIQMs are listed in Table 3.2 with corresponding citations. Most measures listed in Table 3.2 were also evaluated in [20] for a recognition task. All the measures except the first are computed automatically. The first ten measures are simply complexity features that do not consider the source images (the no-source comparative class according to the classification in Section 6.1). They represent the structure, texture, contrast and/or edge intensities in the image in order to characterize the complexity of the image. Such measures have already been used to evaluate the quality of image fusion algorithms [20,21,86]. Most of these measures have been inspired by work to develop clutter complexity measures [22,87]. These works search for features that characterize the degree to which the background appears target-like [87]. Ideally, the clutter complexity determines how hard it is to detect or classify a target in the scene due to the complexity of the background. The last seven measures compare how well the salient features in the two source imagery are transferred into the fused image (the source comparative class). For the most part, the distinction between these comparative measures is in the definition of saliency.

Ideally, the FIQM should be computed automatically from the fused and source images. The contrast measure is considered because it is one of the measures that is averaged in an objective National Imagery Interpretability Ratings Scale (NIIRS) rating [89]. Furthermore, it is intuitive that the contrast between the target and the background facilitates ease of detection. The contrast is computed by manually segmenting human silhouettes for each scene. Fig. 3.8(i) shows an example of the silhouette that separates the target from the background. The white part in the silhouette denotes target pixels, and the black part denotes background pixels. The measure is equivalent to the percent contrast used in [90].

Category	Index	Measure Description
Contrast	1	Difference of intensity or color between the target and the background
Saturation [20]	2	Normalized histogram peak
STD	3	Standard deviation
Schmieder Weathersby [22]	4	Block average local standard deviation
fBm [23]	5	Hurst parameter for fBm model
TIR [24]	6	Block average target interference ratio (contrast)
Energy [24]	7	Block average energy of histogram
Entropy [24]	8	Block average entropy of histogram
Homogeneity [24]	9	Block average pixel variation
Block Outlier [24]	10	Block average number of outliers
Universal Quality Index [72]	11	Average Structure SIMilarity (SSIM) index between fused and reference images
Information Measures [14]	12	Average mutual information between fused and reference images (bin size = 16)
Objective Measure [13]	13	Average objective edge information between fused and reference images
Salient Quality Index [15]	14	Weighted average salient quality index of edge intensities between fused and reference images
	15	Weighted average salient quality index between fused and reference images
	16	Average salient quality index between fused and reference images
Harris Response based quality metric [88]	17	Difference of Harris response between fused and reference images

Table 3.2: List of the evaluated FIQMs.

For grayscale imagery, it is defined as

$$\text{contrast} = \frac{|I_t - I_b|}{d}, \tag{3.23}$$

where I_t and I_b are the mean target and background intensities, respectively, and d denotes the dynamic range, i.e. the intensity difference between the brightest and darkest pixels in a scene. For color imagery, the RGB coordinates are converted to the CIE $L^*a^*b^*$ color space [91] and the single band contrast is calculated independently over the L^*, a^*, and b^* bands via (3.23). Then the root sum square of the three single band contrasts is reported as the overall contrast. Since the information about the color is given in the a^* and b^* bands, these bands exhibit zero contrast for grayscale imagery, and the color version of contrast is a consistent generalization of the grayscale definition, i.e., it provides the same answer if the RGB image contains no

color. Intuitively, the color version of contrast integrates the contrast that exists in all ways the eye can distinguish the foreground from the background, i.e., lightness and color. It might be possible to generate an automated contrast measure by incorporating automated image segmentation techniques. This is a matter of future investigation.

While the generalization of contrast for color imagery is straightforward, it is not clear how to best extend the definition of the other automatic FIQMs to accommodate color imagery. To this end, we follow the convention in [87] where for the color images, one generates four color measures for a given grayscale measure. Namely, the grayscale measure is computed over each RGB band and summarized by the 1) maximum, 2) minimum and 3) median values over all bands. The fourth measure is computed by converting the RGB image into a grayscale image before calculating the measure.

3.4.3 Evaluation Results and Discussion

First, we evaluated the consistency of the FIQMs with human detection performance over the five grayscale fused images displays: 1) II, 2) LWIR, 3) CONA, 4) CONB, and 5) DWTT. Then, we considered scoring the FIQMs generalized for color using all eight fused image displays.

Table 3.3 provides the composite DPMLR score over the five grayscale displays of the 35 scenes for each of the 17 grayscale measures as well as the corresponding p-values. Note that for each FIQM, the p-value is evaluated by calculating the probability of obtaining a result with the DPMLR larger than the composite DPMLR score listed in Table 3.3 when the H_0 hypotheses is true. Furthermore, the table also includes the average values and average absolute values of the monotonic, logistic, Spearman and Kendall correlations.

The second column of Table 3.3 shows that the composite DPMLR scores for all but

Index	Λ_N	p-value	Mean MC	Mean |MC|	Mean LC	Mean |LC|
1	1.3291	0.1221	0.5835	0.8276	0.5628	0.8069
2	0.0204	0.4284	0.0153	0.6962	0.0105	0.6382
3	0.0301	0.3982	0.1291	0.7958	0.1374	0.7714
4	0.0340	0.3884	0.0192	0.7965	0.0180	0.7911
5	0.5925	0.1741	0.5322	0.8564	0.5330	0.8398
6	0.3637	0.2073	0.1380	0.5493	0.1381	0.5490
7	0.0376	0.3803	0.0546	0.7670	0.0405	0.7248
8	0.0392	0.3768	0.0921	0.7636	0.0883	0.7386
9	0.0382	0.3789	-0.0106	0.8068	-0.0012	0.7851
10	0.0422	0.3709	0.1453	0.7723	0.1385	0.7555
11	0.0316	0.3943	-0.0292	0.7543	-0.0268	0.7339
12	0.0362	0.3833	0.2030	0.7262	0.1667	0.6771
13	0.0479	0.3607	0.1018	0.7863	0.1063	0.7622
14	0.0252	0.4120	0.1454	0.7668	0.1426	0.7581
15	0.0242	0.4150	0.1541	0.7756	0.1402	0.7558
16	0.0387	0.3778	0.3124	0.7694	0.2969	0.7391
17	0.3407	0.2122	0.1973	0.5980	0.1931	0.5910

Index	Λ_N	p-value	Mean SC	Mean |SC|	Mean KC	Mean |KC|
1	1.3291	0.1221	0.5210	0.6168	0.4700	0.5602
2	0.0204	0.4284	0.0053	0.4158	0.0190	0.3425
3	0.0301	0.3982	0.0868	0.4981	0.0891	0.4285
4	0.0340	0.3884	0.0720	0.4498	0.0711	0.3807
5	0.5925	0.1741	0.3608	0.5730	0.3094	0.4933
6	0.3637	0.2073	0.1381	0.5490	0.0649	0.4158
7	0.0376	0.3803	0.0748	0.4433	0.0729	0.3780
8	0.0392	0.3768	0.0896	0.4434	0.0745	0.3730
9	0.0382	0.3789	-0.0360	0.5160	-0.0292	0.4384
10	0.0422	0.3709	0.0927	0.5002	0.0807	0.4197
11	0.0316	0.3943	0.0130	0.4066	0.0345	0.3494
12	0.0362	0.3833	0.1363	0.3950	0.1315	0.3457
13	0.0479	0.3607	0.0647	0.4677	0.0651	0.4094
14	0.0252	0.4120	0.0792	0.4340	0.0870	0.3743
15	0.0242	0.4150	0.0873	0.4421	0.0934	0.3807
16	0.0387	0.3778	0.1534	0.4396	0.1430	0.3898
17	0.3407	0.2122	0.1424	0.3792	0.1427	0.3527

Table 3.3: List of DPMLR scores, associated p-values, and average values of some other correlations for 17 grayscale FIQMs tested over five image displays.

the grayscale contrast measure are significantly less than one. This means that the evidence points to the fact that these potential FIQMs are viewed as noise with respect to ordering the detection probabilities of the imagery. The poor performance of the source comparative measures may be explained by structure in the fused and source images that leads to good inter-image correlation but that has no (or even negative) effect on human performance. Examples of the pitfalls of source comparative measures when the ideal image is unknown are provided in [17].

For the grayscale contrast measure, the composite DPMLR score is still modest at 1.3291 and the p-value is not very low. In fact, the perfect FIQM that consistently ordered the number of detections $\mathbf{y}$ over all 35 scenes would provide a composite DPMLR of 9.632. This means that while there is evidence to reject the null hypothesis, the evidence to support

the monotonic hypothesis is not compelling. However, the composite DPMLR score for the grayscale contrast measure is much greater than the scores for the others. Thus, the contrast measure may be a key aspect to a proper FIQM.

From Table 3.3, one can see that the orderings of the FIQMs via the DPMLR and the other correlations differ. Also note that for each FIQM, the differences between the average correlations and the average absolute correlations indicate a consistency issue for the nature of monotonicity over the 35 scenes. The contrast measure exhibits by far the largest DPMLR. However, its average absolute values of the monotonic correlation and the logistic correlation (Mean |MC| and Mean |LC| in Table 3.3) are less than those of the fBm, respectively. Furthermore, the other average correlations of the contrast measure are only slightly larger than those of the fBm. To better compare these two measures, and to show how differently the DPMLR and the other four correlations evaluate a FIQM based upon the human perception results, we present the human detection results and the scores of the DPMLRT, monotonic, logistic, Spearman, and Kendall correlation tests for each scene for the contrast and the fBm measures.

Fig. 3.9 graphically depicts the relationship between the aforementioned two measures and the human performance over all 35 scenes. The lines marked by the asterisk correspond to the contrast measure and the lines marked by the circle correspond to the fBm. Since only the five gray fused image displays are considered here, for each scene and each FIQM, we have five detection numbers $y_i \in [0, 8]$ and five FIQM values $\tilde{x}_i$ $(1 \leqslant i \leqslant 5)$. In each plot of Fig. 3.9, the vertical axis denotes the number of humans that detected the target, while the horizontal axis stands for the rank of the $\tilde{x}_i$'s sorted in ascending order. The shade of the background of each plot indicates the significance of the monotonic ordering for each scene. The significance value is obtained by calculating the DPMLR of the given y_i's for an

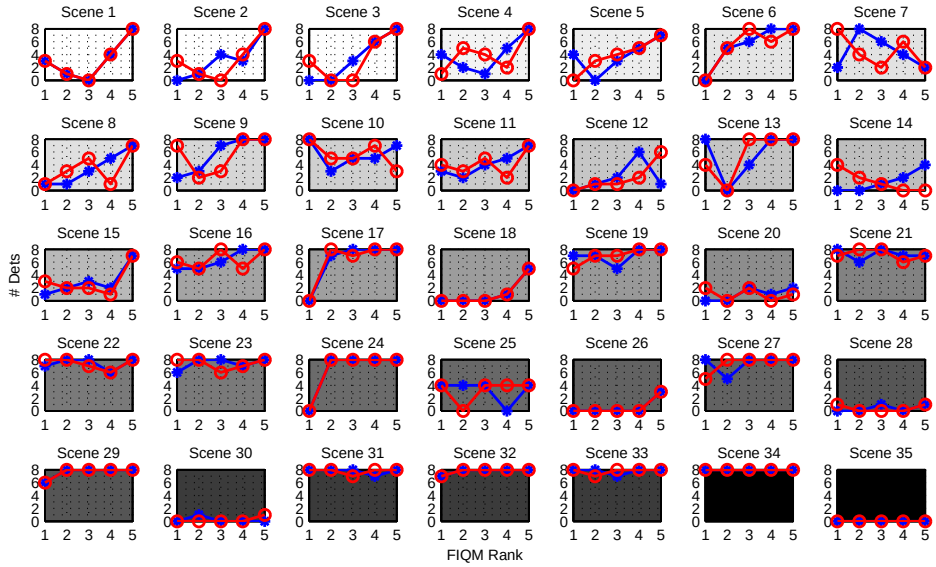


Figure 3.9: Scatter plots of the number of detections versus the quality rank order over 35 scenes ('$*$' represents the contrast measure, and '$\circ$' represents the fBm measure).

imaginary FIQM whose values perfectly match the y_i's in the monotonically increasing order.

Tables 3.4 and 3.5 provide the ascending and descending DPMLRs as well as the other four correlations (monotonic, logistic, Spearman, and Kendall) over each scene for the contrast measure and the fBm measure, respectively. Note that in Scenes 34 and 35, the same number of detections are obtained for five different displays. Because of the fact that the target is so obvious in Scene 34, all eight observers detected it successfully. Similarly, no one detected the target in Scene 35 because it is so unclear. Both cases are naturally ignored as they don't provide any information on the monotonicity.

One very important property of the DPMLR is that it can capture the significance of a scene based upon the human detection results, and accordingly adjust its score to provide a more precise evaluation. The significance, as defined, is determined by the number of unique

human detection values and the spread of these values over the dynamic range from 0 to 8 detections. Essentially, the significance describes how easy (or difficult) it is for random noise to affect the order of the human detection results. The more unique values that the human detection results take in a scene, the less likely that random noise will order the human detection results. The monotonic and logistic correlations give a value of one whenever a scene's scatter plot is perfectly monotonic, as observed from Scenes $3, 6, 8, 9, 14, 16, 17, 18, 24, 26, 29,$ and 32 in Fig. 3.9 and the corresponding statistics in Table 3.4. The Spearman and the Kendall correlations give a value of one whenever a scene's scatter plot is strictly monotonic, as observed from the scatter plot of Scene 5 and the corresponding statistics in Table 3.5. However, the DPMLR gives a score much greater than one in the significant scenes. Specifically, in Scenes $3, 6, 8, 9$ and 14, the likelihoods for noise to order the data are much slimmer than those in Scenes $24, 26, 29$ and 32. As a result, the DPMLR provides significantly higher scores in the former than in the later as seen in Table 3.4.

One can also observe that the miss-ordering for the more significant scenes causes lower DPMLR scores than those of the less significant scenes. For instance, we compare the scatter plots of Scenes 25 and 31 in Fig. 3.9 as well as the corresponding descending DPMLR values in Table 3.4. The descending DPMLR for Scene 25 is much smaller than that of Scene 31 because the DPMLR treats the miss-ordering in Scene 31 as due to the measurement noise. On the other hand, the other four correlations are equally unforgiving of the miss-ordering regardless of the significance of the scene. This is because that the correlations are invariant to linear scaling of the human detection results, whereas the DPMLR uses the binomial measurement model to determine whether or not the scale of the miss-ordering is significant.

Once we realize the DPMLR's ability to incorporate the significance of each scene into the statistical test, it is easy to see why the DPMLR provides the significantly high score

for the contrast measure. Comparing the scatter plots of these two measures in the first 20 scenes, we can see that seven of them, i.e., Scenes $3, 6, 8, 9, 14, 16$ and 17, exhibit the perfect monotonicity for the contrast measure and some miss-orderings for the fBm measure. In fact, the nature of the monotonic relationship for the fBm feature flips for Scene 14, i.e., it is perfectly decreasing. Both of these factors lead to the significantly higher DPMLR for the contrast measure. Because the contrast measure is still not nearly monotonically related to the perception results of many of the significant scenes, the composite DPMLR score is only slightly greater than one.

Next, we used all $N = 8$ fused image displays and ran another DPMLRT for color-based FIQMs when the human detection results were collected over $o = 8$ observers. The composite DPMLR scores of the 64 color-based FIQMs derived from the 16 automated grayscale FIQMs are low and not included here for the sake of brevity. On the other hand, the color-based contrast measure achieved a composite DPMLR score of 1.4000, which is slightly greater than that of the contrast computed only over the five grayscale fused image displays. Because the number of fused images N has increased, the significance of this "greater than one" score increases and the p-value is 0.041972, which gives stronger support for the monotonic hypothesis. Certainly, the color-based contrast is able to incorporate the contrast from both the luminance and color components in an RGB image and serves as a potential FIQM that is able to explain some of the human performance. Again, the perfect FIQM would provide a composite DPMLR of 54.5150, and contrast is only one aspect of a good FIQM, which has yet to be identified.

3.5 Summary

This chapter proposes the composite DPMLR to quantify how consistent the values of a FIQM are with measured human performance represented by the probability of detection. Specifically, the DPMLR can be used to test whether or not a monotonic relationship exists between the FIQM and the underlying human detection performance that is measured via a perception experiment. The resulting test is designed to be applicable even when the number of observers is small so that the measurement errors from the perceptual experiment are not necessarily Gaussian. The chapter discusses some interesting properties of the DPMLR, and simulation results demonstrate the advantages of the DPMLR over other monotonic statistics. Unlike the monotonic correlation in [20], the DPMLR seamlessly accounts for the spread of the human observations and the number of fused images. It indicates to what degree the ordering of the human observations by the FIQM is not by random chance. The DPMLRT is a general test of monotonicity that can be used to evaluate monotonic relationships beyond the image fusion application. Finally, the DPMLR was used to score a number of potential FIQMs using real image data with a corresponding perception study.

The DPMLR scores reveal that a proper FIQM for the detection task is not yet available. The comparative measures may have scored poorly because the salient features exploited by these measures may not have captured the context in II imagery that humans exploit for detection. Of note, the contrast measure does demonstrate some utility based on its DPMLR score, and is clearly one aspect that drives human detection performance. Future work is needed to identify a more meaningful FIQM. Such a measure may incorporate aspects of the contrast as well as other quality features of both the luminance and color components of the image. However, we expect that a measure needs to understand what context is available in

the image, which makes the search for a good FIQM very challenging.

The chapter revealed many interesting properties of the DPMLR and conjectured many more properties. Future work is necessary to prove (or disprove) these conjectured properties. Furthermore, one can further study over what values of $\tilde{\mathbf{p}}$ the DPMLRT is the most powerful test.

The DPMLRT does incorporate some simplifying assumptions that could be relaxed for a more robust test. For instance, not all human observers are created equal and the binomial distribution may not be the best model for the perception results. Furthermore, the values of $\tilde{p}_i$ are not independent since all fusion algorithms attempt to provide a good image for human perception. The chapter does demonstrate that the DPMLRT is robust as these model assumptions are relaxed. In addition, the DPMLRT assumes that the observers' probability of false alarms are calibrated, and it ignores the impact of contextual information, which may be known *a priori* or obtained in the image, on human detection performance. Future research can also focus on statistical scoring mechanisms that account for increasingly realistic data models.

3.6 Appendix

3.6.1 Proof of Property 1

Proof: $\lambda_N^{\uparrow}(\mathbf{y}, \mathbf{o})$ and $\lambda_N^{\downarrow}(\mathbf{y}, \mathbf{o})$ can be expressed as:

$$\lambda_N^{\uparrow}(\mathbf{y}, \mathbf{o}) = \frac{N! \int_{\mathcal{P}_{\uparrow}} \prod_{i=1}^{N} p_i^{y_i} (1-p_i)^{o-y_i} d\mathbf{p}}{\int_{\mathcal{P}_0} \prod_{i=1}^{N} p_i^{y_i} (1-p_i)^{o-y_i} d\mathbf{p}}, \text{ and } \lambda_N^{\downarrow}(\mathbf{y}, \mathbf{o}) = \frac{N! \int_{\mathcal{P}_{\downarrow}} \prod_{i=1}^{N} p_i^{y_i} (1-p_i)^{o-y_i} d\mathbf{p}}{\int_{\mathcal{P}_0} \prod_{i=1}^{N} p_i^{y_i} (1-p_i)^{o-y_i} d\mathbf{p}}. \tag{3.24}$$

Note that the integrands in the numerator and denominator are the same. This integrand is strictly positive for all $\mathbf{p}$ except for a finite set of points of measure zero, namely $\{\mathbf{p} : p_i \in \{0, 1\}\}$. Any integral of the integrand over $\mathcal{P}_0$, $\mathcal{P}_\uparrow$, $\mathcal{P}_\downarrow$, $\mathcal{P}_0 \backslash \mathcal{P}_\uparrow$, and $\mathcal{P}_0 \backslash \mathcal{P}_\downarrow$ must be strictly positive. Thus, the integrals in the numerator of (3.24) are strictly less than the integrals in the denominator. Furthermore, all the integrals are strictly positive. Thus,

$$0 < \frac{\int_{\mathcal{P}_\uparrow} \prod_{i=1}^{N} p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p}}{\int_{\mathcal{P}_0} \prod_{i=1}^{N} p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p}} < 1, \text{ and } 0 < \frac{\int_{\mathcal{P}_\downarrow} \prod_{i=1}^{N} p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p}}{\int_{\mathcal{P}_0} \prod_{i=1}^{N} p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p}} < 1. \tag{3.25}$$

Multiplication by $N!$ leads to $0 < \lambda_N^{\uparrow}(\mathbf{y}, \mathbf{o}) < N!$ and $0 < \lambda_N^{\downarrow}(\mathbf{y}, \mathbf{o}) < N!$. Because the ascending and descending DPMLRs are bounded by zero and $N!$ for each scene, it is clear by (3.11) that the composite DPMLR is also bounded by 0 and $N!$. ■

3.6.2 Proof of Property 3

Proof: Let $\pi_k : \{1, 2, \ldots, N\} \longmapsto \{1, 2, \ldots, N\}$ be a permutation mapping such that $\pi_k(i) \neq \pi_k(j)$ when $i \neq j$. There are $N!$ such mappings, and let each mapping be identified with a unique index k where $k = 1, 2, \ldots, N!$. As a matter of convention, $k = 1$ is the identity mapping, i.e., $\pi_1(i) = i$, and $k = N!$ is the reverse sort, i.e., $\pi_{N!}(i) = N + 1 - i$. Each permutation function allows one to define an ordering of the coordinates, i.e.,

$$\mathcal{R}_k = \{\mathbf{p} : 0 \leqslant p_{\pi_k(1)} \leqslant p_{\pi_k(2)} \leqslant \ldots \leqslant p_{\pi_k(N)} \leqslant 1\} \tag{3.26}$$

such that the collection of all $N!$ orderings defines any possible sequence of coordinate values, i.e.,

$$\mathcal{P}_0 = \bigcup_{k=1}^{N!} \mathcal{R}_k. \tag{3.27}$$

Furthermore, $\mathcal{P}_\uparrow = \mathcal{R}_1$ and $\mathcal{P}_\downarrow = \mathcal{R}_{N!}$. As a result,

$$\int_{\mathcal{P}_0} \prod_{i=1}^{N} p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p} = \sum_{k=1}^{N!} \int_{\mathcal{R}_k} \prod_{i=1}^{N} p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p}. \tag{3.28}$$

Using the change of variable $p_{\pi_k(i)} \mapsto p_i$, the right hand side of (3.28) can be rewritten as

$$\int_{\mathcal{P}_0} \prod_{i=1}^{N} p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p} = \sum_{k=1}^{N!} \int_{\mathcal{P}_\uparrow} \prod_{i=1}^{N} p_i^{y_{\pi_k(i)}}(1-p_i)^{o-y_{\pi_k(i)}} d\mathbf{p}. \tag{3.29}$$

If $y_1 = y_2 = \ldots = y_N$, then we have $y_{\pi_1(i)} = y_{\pi_2(i)} = \ldots = y_{\pi_{N!}(i)} = y_i$ for $i = 1, 2, \ldots, N$. It follows that

$$\int_{\mathcal{P}_0} \prod_{i=1}^{N} p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p} = N! \int_{\mathcal{P}_\uparrow} \prod_{i=1}^{N} p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p}. \tag{3.30}$$

According to (3.24), we have $\lambda_N^\downarrow(\mathbf{y}, o\mathbf{1}) = 1 = \lambda_N^\uparrow(\mathbf{y}, o\mathbf{1})$ ■

3.6.3 Proof of Properties 4 and 5

Proof: First we want to show that if the y_i's are in ascending order and not constant, then

$$\prod_{i=1}^{N} p_i^{y_i}(1-p_i)^{o-y_i} > \prod_{i=1}^{N} p_i^{y_{\pi_k(i)}}(1-p_i)^{o-y_{\pi_k(i)}}, \tag{3.31}$$

when $k \neq 0$.

To this end, we transform the permutation π_k back to the identity. Let's define the permutation function $g_0(i) = \pi_k(i)$. For the first step, the value of $g_0(1)$ is switched with the value $g_0(j)$ where $g_0(j) = 1$ to form g_1. The process repeats itself for $N-1$ steps such that for the n-th step, the value of $g_{n-1}(n)$ is switched with $g_{n-1}(j)$ where $g_{n-1}(j) = n$ to form g_n. Formally, at the n-th step we have $g_n(n) = n, g_n(j) = g_{n-1}(n)$, and $g_n(i) = g_{n-1}(i)$, where $j = g_{n-1}^{-1}(n)$, $i > n$ and $i \neq j$.

Note that $j \geqslant n$ and $g_{n-1}(n) \geqslant n$ because $g_{N-1}(i) = i$ for $i = 1, 2, \ldots, n-1$. After $N-1$ steps, $\pi_1 = g_{n-1}$. After the n-th step, the ratio of the likelihoods associated with permutations g_n and g_{n-1}, i.e.,

$$\frac{\prod_{i=1}^N p_i^{y_{g_n(i)}}(1-p_i)^{o-y_{g_n(i)}}}{\prod_{i=1}^N p_i^{y_{g_{n-1}(i)}}(1-p_i)^{o-y_{g_{n-1}(i)}}} = \left(\frac{p_n(1-p_j)}{p_j(1-p_n)}\right)^{y_n - y_{g_{n-1}}(n)} \tag{3.32}$$

is greater than or equal to unity because $y_n \leqslant y_{g_{n-1}(n)}$ and $p_n \leqslant p_j$ over $\mathcal{P}_\uparrow$. By taking the product of (3.32) for $n = 1, 2, \ldots, N-1$, we have

$$\frac{\prod_{i=1}^N p_i^{y_{g_{N-1}(i)}}(1-p_i)^{o-y_{g_{N-1}(i)}}}{\prod_{i=1}^N p_i^{y_{g_0(i)}}(1-p_i)^{o-y_{g_0(i)}}} \geqslant 1. \tag{3.33}$$

The equality occurs only if $y_n = y_{g_{n-1}(n)}$ for $n = 1, \ldots, N$, which means y_i's are equal. Because $g_0 = \pi_k$ and g_{N-1} is the identity map and the y_i's are not constant, (3.31) is proven.

Now, integrating both sides of (3.31) over $\mathcal{P}_\uparrow$ leads to

$$\int_{\mathcal{P}_\uparrow} \prod_{i=1}^N p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p} > \int_{\mathcal{P}_\uparrow} \prod_{i=1}^N p_i^{y_{\pi_k(i)}}(1-p_i)^{o-y_{\pi_k(i)}} d\mathbf{p} \tag{3.34}$$

when $k \neq 1$. Similarly, one can show that

$$\int_{\mathcal{P}_\uparrow} \prod_{i=1}^N p_i^{y_{\pi_{N!}(i)}}(1-p_i)^{o-y_{\pi_{N!}(i)}} d\mathbf{p} < \int_{\mathcal{P}_\uparrow} \prod_{i=1}^N p_i^{y_{\pi_k(i)}}(1-p_i)^{o-y_{\pi_k(i)}} d\mathbf{p} \tag{3.35}$$

when $k \neq N!$. The division of (3.34) and (3.35) by $\prod_{i=1}^N \beta(y_i+1, o-y_i+1)$ leads to the first statement in Property 4. Similar arguments prove the second statement in Property 4.

Summing (3.34) for $k = 1, \ldots, N!$ leads to

$$\int_{\mathcal{P}_0} \prod_{i=1}^N p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p} < N! \int_{\mathcal{P}_\uparrow} \prod_{i=1}^N p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p}. \tag{3.36}$$

Then $\lambda_N^\uparrow(\mathbf{y}, \mathbf{o}) > 1$. Similarly, (3.35) can be reexpressed as

$$\int_{\mathcal{P}_\downarrow} \prod_{i=1}^N p_i^{y_i}(1-p_i)^{o-y_i} d\mathbf{p} < \int_{\mathcal{P}_\downarrow} \prod_{i=1}^N p_i^{y_{\pi_k(i)}}(1-p_i)^{o-y_{\pi_k(i)}} d\mathbf{p}, \tag{3.37}$$

so that $\lambda_N^{\downarrow}(\mathbf{y}, \mathbf{o}) < 1$. This completes the proof of the first statement in Property 5. The proof of the second statement can be proven by similar arguments. ■

Scene 1	Scene 2	Scene 3	Scene 4	Scene 5
$\lambda_5^{\uparrow} = 0.7043$	$\lambda_5^{\uparrow} = 23.6503$	$\lambda_5^{\uparrow} = 45.2298$	$\lambda_5^{\uparrow} = 1.1634$	$\lambda_5^{\uparrow} = 0.6730$
$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0001$
MC = 0.9379	MC = 0.9935	MC = 1.0000	MC = 0.9189	MC = 0.8376
LC = 0.9379	LC = 0.9699	LC = 0.9990	LC = 0.9189	LC = 0.8376
SC = 0.6000	SC = 0.9000	SC = 0.9747	SC = 0.6000	SC = 0.7000
KC = 0.4000	KC = 0.8000	KC = 0.9487	KC = 0.4000	KC = 0.6000
Scene 6	Scene 7	Scene 8	Scene 9	Scene 10
$\lambda_5^{\uparrow} = 31.4151$	$\lambda_5^{\uparrow} = 0.0010$	$\lambda_5^{\uparrow} = 29.1496$	$\lambda_5^{\uparrow} = 24.7705$	$\lambda_5^{\uparrow} = 0.0274$
$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0137$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0746$
MC = 1.0000	MC = -0.5601	MC = 1.0000	MC = 1.0000	MC = -0.6882
LC = 1.0000	LC = -0.5601	LC = 0.9826	LC = 0.8778	LC = -0.6882
SC = 0.9747	SC = -0.2052	SC = 0.9747	SC = 0.9747	SC = -0.0513
KC = 0.9487	KC = -0.3162	KC = 0.9487	KC = 0.9487	KC = 0.1054
Scene 11	Scene 12	Scene 13	Scene 14	Scene 15
$\lambda_5^{\uparrow} = 11.1297$	$\lambda_5^{\uparrow} = 0.2919$	$\lambda_5^{\uparrow} = 0.0005$	$\lambda_5^{\uparrow} = 17.3610$	$\lambda_5^{\uparrow} = 9.4196$
$\lambda_5^{\downarrow} = 0.0014$	$\lambda_5^{\downarrow} = 0.0006$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0009$	$\lambda_5^{\downarrow} = 0.0002$
MC = 0.9830	MC = 0.6571	MC = 0.6124	MC = 1.0000	MC = 0.9886
LC = 0.9558	LC = 0.5026	LC = 0.6124	LC = 1.0000	LC = 0.9708
SC = 0.9000	SC = 0.5643	SC = 0.3354	SC = 0.9747	SC = 0.8208
KC = 0.8000	KC = 0.5270	KC = 0.3586	KC = 0.9487	KC = 0.7379
Scene 16	Scene 17	Scene 18	Scene 19	Scene 20
$\lambda_5^{\uparrow} = 11.8403$	$\lambda_5^{\uparrow} = 10.8893$	$\lambda_5^{\uparrow} = 10.4186$	$\lambda_5^{\uparrow} = 1.0475$	$\lambda_5^{\uparrow} = 4.7336$
$\lambda_5^{\downarrow} = 0.0022$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0001$	$\lambda_5^{\downarrow} = 0.0376$	$\lambda_5^{\downarrow} = 0.0250$
MC = 1.0000	MC = 1.0000	MC = 1.0000	MC = 0.7454	MC = 0.9354
LC = 1.0000	LC = 1.0000	LC = 1.0000	LC = 0.7454	LC = 0.8815
SC = 0.9487	SC = 0.8944	SC = 0.8944	SC = 0.6325	SC = 0.7906
KC = 0.8944	KC = 0.8367	KC = 0.8367	KC = 0.4472	KC = 0.6708
Scene 21	Scene 22	Scene 23	Scene 24	Scene 25
$\lambda_5^{\uparrow} = 0.1784$	$\lambda_5^{\uparrow} = 0.3496$	$\lambda_5^{\uparrow} = 1.7957$	$\lambda_5^{\uparrow} = 4.9996$	$\lambda_5^{\uparrow} = 0.0108$
$\lambda_5^{\downarrow} = 0.7175$	$\lambda_5^{\downarrow} = 0.2204$	$\lambda_5^{\downarrow} = 0.0501$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.1740$
MC = -0.5345	MC = -0.4082	MC = 0.8898	MC = 1.0000	MC = -0.6124
LC = -0.5345	LC = -0.4082	LC = 0.8458	LC = 1.0000	LC = -0.6124
SC = -0.2635	SC = 0.1118	SC = 0.4472	SC = 0.7071	SC = -0.3536
KC = -0.2236	KC = 0.1195	KC = 0.3586	KC = 0.6325	KC = -0.3162
Scene 26	Scene 27	Scene 28	Scene 29	Scene 30
$\lambda_5^{\uparrow} = 4.3533$	$\lambda_5^{\uparrow} = 0.5042$	$\lambda_5^{\uparrow} = 2.0380$	$\lambda_5^{\uparrow} = 3.5970$	$\lambda_5^{\uparrow} = 0.4156$
$\lambda_5^{\downarrow} = 0.0042$	$\lambda_5^{\downarrow} = 0.0259$	$\lambda_5^{\downarrow} = 0.1714$	$\lambda_5^{\downarrow} = 0.0296$	$\lambda_5^{\downarrow} = 1.2533$
MC = 1.0000	MC = 0.6124	MC = 0.7638	MC = 1.0000	MC = -0.6124
LC = 1.0000	LC = 0.6124	LC = 0.5417	LC = 1.0000	LC = -0.6124
SC = 0.7071	SC = 0.3536	SC = 0.5774	SC = 0.7071	SC = -0.3536
KC = 0.6325	KC = 0.3162	KC = 0.5164	KC = 0.6325	KC = -0.3162
Scene 31	Scene 32	Scene 33	Scene 34	Scene 35
$\lambda_5^{\uparrow} = 0.4156$	$\lambda_5^{\uparrow} = 2.4100$	$\lambda_5^{\uparrow} = 0.7393$	$\lambda_5^{\uparrow} = 1.0000$	$\lambda_5^{\uparrow} = 1.0000$
$\lambda_5^{\downarrow} = 1.2533$	$\lambda_5^{\downarrow} = 0.1818$	$\lambda_5^{\downarrow} = 0.7393$	$\lambda_5^{\downarrow} = 1.0000$	$\lambda_5^{\downarrow} = 1.0000$
MC = -0.6124	MC = 1.0000	MC = 0.4082	MC = NaN	MC = NaN
LC = -0.6124	LC = 1.0000	LC = 0.4082	LC = NaN	LC = NaN
SC = -0.3536	SC = 0.7071	SC = 0.0000	SC = NaN	SC = NaN
KC = -0.3162	KC = 0.6325	KC = 0.0000	KC = NaN	KC = NaN

Table 3.4: Statistics for the contrast measure where MC, LC, SC and KC are the monotonic, logistic, Spearman, and Kendall correlations, respectively.

Scene 1	Scene 2	Scene 3	Scene 4	Scene 5
$\lambda_5^{\uparrow} = 0.7043$	$\lambda_5^{\uparrow} = 0.7043$	$\lambda_5^{\uparrow} = 0.3819$	$\lambda_5^{\uparrow} = 1.5343$	$\lambda_5^{\uparrow} = 35.3484$
$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0000$
MC = 0.9379	MC = 0.9379	MC = 0.9396	MC = 0.9189	MC = 1.0000
LC = 0.9379	LC = 0.9379	LC = 0.9396	LC = 0.8385	LC = 0.9766
SC = 0.6000	SC = 0.6000	SC = 0.6669	SC = 0.6000	SC = 1.0000
KC = 0.4000	KC = 0.4000	KC = 0.5270	KC = 0.4000	KC = 1.0000
Scene 6	Scene 7	Scene 8	Scene 9	Scene 10
$\lambda_5^{\uparrow} = 6.8845$	$\lambda_5^{\uparrow} = 0.0001$	$\lambda_5^{\uparrow} = 0.8749$	$\lambda_5^{\uparrow} = 0.0708$	$\lambda_5^{\uparrow} = 0.0011$
$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 1.0015$	$\lambda_5^{\downarrow} = 0.0001$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 2.5692$
MC = 0.9766	MC = -0.8402	MC = 0.8402	MC = 0.7605	MC = -0.9081
LC = 0.9686	LC = -0.7796	LC = 0.7956	LC = 0.7605	LC = -0.8147
SC = 0.8208	SC = -0.6156	SC = 0.5643	SC = 0.6669	SC = -0.6669
KC = 0.7379	KC = -0.5270	KC = 0.5270	KC = 0.5270	KC = -0.5270
Scene 11	Scene 12	Scene 13	Scene 14	Scene 15
$\lambda_5^{\uparrow} = 0.6460$	$\lambda_5^{\uparrow} = 19.4999$	$\lambda_5^{\uparrow} = 0.2934$	$\lambda_5^{\uparrow} = 0.0009$	$\lambda_5^{\uparrow} = 0.7544$
$\lambda_5^{\downarrow} = 0.0064$	$\lambda_5^{\downarrow} = 0.0001$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 17.3610$	$\lambda_5^{\downarrow} = 0.0007$
MC = 0.8137	MC = 1.0000	MC = 0.9186	MC = -1.0000	MC = 0.9535
LC = 0.8137	LC = 0.9939	LC = 0.9186	LC = -0.9801	LC = 0.9535
SC = 0.3000	SC = 0.9747	SC = 0.7826	SC = -0.9747	SC = 0.0513
KC = 0.2000	KC = 0.9487	KC = 0.5976	KC = -0.9487	KC = -0.1054
Scene 16	Scene 17	Scene 18	Scene 19	Scene 20
$\lambda_5^{\uparrow} = 0.5725$	$\lambda_5^{\uparrow} = 5.2378$	$\lambda_5^{\uparrow} = 10.4186$	$\lambda_5^{\uparrow} = 9.9553$	$\lambda_5^{\uparrow} = 0.0779$
$\lambda_5^{\downarrow} = 0.0099$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0001$	$\lambda_5^{\downarrow} = 0.0071$	$\lambda_5^{\downarrow} = 0.5866$
MC = 0.6757	MC = 0.9949	MC = 1.0000	MC = 1.0000	MC = -0.6124
LC = 0.6233	LC = 0.9897	LC = 1.0000	LC = 0.9576	LC = -0.5001
SC = 0.3162	SC = 0.6708	SC = 0.8944	SC = 0.9487	SC = -0.3162
KC = 0.2236	KC = 0.5976	KC = 0.8367	KC = 0.8944	KC = -0.2236
Scene 21	Scene 22	Scene 23	Scene 24	Scene 25
$\lambda_5^{\uparrow} = 0.1318$	$\lambda_5^{\uparrow} = 0.1434$	$\lambda_5^{\uparrow} = 0.2252$	$\lambda_5^{\uparrow} = 4.9996$	$\lambda_5^{\uparrow} = 0.1740$
$\lambda_5^{\downarrow} = 0.8082$	$\lambda_5^{\downarrow} = 1.0109$	$\lambda_5^{\downarrow} = 0.6329$	$\lambda_5^{\downarrow} = 0.0000$	$\lambda_5^{\downarrow} = 0.0108$
MC = -0.7638	MC = -0.6124	MC = -0.6124	MC = 1.0000	MC = 0.6124
LC = -0.7638	LC = -0.6124	LC = -0.6124	LC = 1.0000	LC = 0.6124
SC = -0.3689	SC = -0.3354	SC = -0.2236	SC = 0.7071	SC = 0.3536
KC = -0.2236	KC = -0.3586	KC = -0.1195	KC = 0.6325	KC = 0.3162
Scene 26	Scene 27	Scene 28	Scene 29	Scene 30
$\lambda_5^{\uparrow} = 4.3533$	$\lambda_5^{\uparrow} = 4.3533$	$\lambda_5^{\uparrow} = 0.4822$	$\lambda_5^{\uparrow} = 3.5970$	$\lambda_5^{\uparrow} = 2.4100$
$\lambda_5^{\downarrow} = 0.0042$	$\lambda_5^{\downarrow} = 0.0042$	$\lambda_5^{\downarrow} = 0.4822$	$\lambda_5^{\downarrow} = 0.0296$	$\lambda_5^{\downarrow} = 0.1818$
MC = 1.0000	MC = 1.0000	MC = 0.6124	MC = 1.0000	MC = 1.0000
LC = 1.0000	LC = 1.0000	LC = 0.6124	LC = 1.0000	LC = 1.0000
SC = 0.7071	SC = 0.7071	SC = 0.0000	SC = 0.7071	SC = 0.7071
KC = 0.6325	KC = 0.6325	KC = 0.0000	KC = 0.6325	KC = 0.6325
Scene 31	Scene 32	Scene 33	Scene 34	Scene 35
$\lambda_5^{\uparrow} = 0.7393$	$\lambda_5^{\uparrow} = 2.4100$	$\lambda_5^{\uparrow} = 1.2533$	$\lambda_5^{\uparrow} = 1.0000$	$\lambda_5^{\uparrow} = 1.0000$
$\lambda_5^{\downarrow} = 0.7393$	$\lambda_5^{\downarrow} = 0.1818$	$\lambda_5^{\downarrow} = 0.4156$	$\lambda_5^{\downarrow} = 1.0000$	$\lambda_5^{\downarrow} = 1.0000$
MC = 0.4082	MC = 1.0000	MC = 0.6124	MC = NaN	MC = NaN
LC = 0.4082	LC = 1.0000	LC = 0.6124	LC = NaN	LC = NaN
SC = 0.0000	SC = 0.7071	SC = 0.3536	SC = NaN	SC = NaN
KC = 0.0000	KC = 0.6325	KC = 0.3162	KC = NaN	KC = NaN

Table 3.5: Statistics for the fBm where MC, LC, SC and KC are the monotonic, logistic, Spearman, and Kendall correlations, respectively.

Chapter 4

Cramer-Rao Bound for Multi-Target Non-Coherent MIMO Radars

4.1 Introduction

The Cramer-Rao bound (CRB) is a useful tool for evaluating the performance of radar systems, as it provides the mean square error lower bound for any unbiased estimation. Recently, there have been various Cramer-Rao bound studies for the performance of MIMO radar systems in the literature [30–35]. In [35], the joint CRB for the single target location and velocity estimation in a non-coherent MIMO radar system is calculated and analyzed. The CRB for the multi-target localization in a coherent MIMO radar system is investigated in [34].

We consider a multi-target case in a non-coherent MIMO radar system. In this chapter,

we investigate the joint location and velocity estimation of multiple targets, and the Cramer-Rao bound for a two-target case is derived and evaluated. This bound gives us theoretically achievable joint estimation performance. Numerical results show that the spatial advantage observed previously for single-target cases can also be observed in two-target cases.

4.2 System Model

We consider a MIMO radar system with M transmitters, N receivers and Q targets, which are placed in a two-dimensional plane. We denote by (x_k^t, y_k^t) and (x_l^r, y_l^r) the locations of the k-th ($1 \leqslant k \leqslant M$) transmitter and the l-th ($1 \leqslant l \leqslant N$) receiver, respectively. The location and velocity of the i-th ($1 \leqslant i \leqslant Q$) target are (x^i, y^i) and (v_x^i, v_y^i), which are unknown deterministic parameters. The lowpass equivalent waveform transmitted from the k-th transmitter is denoted by $\sqrt{\frac{E}{M}}s_k(t)$, where E is the total transmitted energy from all M transmitters, and the waveform is normalized as $\int_T |s_k(t)|^2 dt = 1$, where T is the signal duration time. The reflection coefficient corresponding to the (l, i, k)-th path is modeled as a zero-mean complex Gaussian random variable $\zeta_{lk}^i \sim CN(0, \sigma^2)$, which is constant over the observation interval. For different l, k or i, the coefficients ζ_{lk}^i ($1 \leqslant i \leqslant Q, 1 \leqslant l \leqslant N$ and $1 \leqslant k \leqslant M$) are independent of each other. Assume that the waveforms transmitted from the M transmitters are approximately orthogonal and maintain approximate orthogonality for time delays and Doppler shifts of interest, i.e.,

$$\int_T s_k(t - \tau_m) s_{k'}^*(t - \tau_{m'}) e^{j2\pi(f_m - f_{m'})t} dt = 0, \text{if } k \neq k'. \tag{4.1}$$

Thus the signals transmitted from different transmitters can be separated at each receiver. The clutter-plus-noise received at the l-th receiver and corresponding to the k-th

transmitter is a temporally white, zero-mean complex Gaussian random process, i.e.,

$$E\left\{w_{lk}(t)w^*_{l'k'}(u)\right\} = \begin{cases} \sigma_w^2\delta(t-u), & \text{if } l=l' \text{ and } k=k', \\ 0, & \text{otherwise,} \end{cases} \tag{4.2}$$

where σ_w^2 is the variance of the clutter-plus-noise and $\delta(t)$ is an unit impulse function. The signal-to-clutter-plus-noise ratio (SCNR) is defined as the average signal power received at each receiver divided by the power of the clutter-plus-noise. Without loss of generality, we normalize the noise power and the variance of the reflection coefficient by setting $\sigma_w^2 = 1$ and $\sigma^2 = 1$, so that the SCNR is written as $10\log(\frac{E}{M})$ in dB.

Under the previously stated assumptions, the signal received at receiver l due to the waveform transmitted from transmitter k and the reflection of the Q targets is written as

$$r_{lk}(t) = \sqrt{\frac{E}{M}}\sum_{i=1}^{Q}\zeta_{lk}^i s_k(t-\tau_{lk}^i)e^{j2\pi f_{lk}^i t} + w_{lk}(t), \tag{4.3}$$

where τ_{lk}^i and f_{lk}^i denote the time delay and the Doppler shift corresponding to the (l,i,k)-th path, respectively. The time delays and Doppler shifts are functions of the unknown parameters (x^i, y^i) and (v_x^i, v_y^i), i.e.,

$$\tau_{lk}^i = \frac{\sqrt{(x_k^t - x^i)^2 + (y_k^t - y^i)^2} + \sqrt{(x_l^r - x^i)^2 + (y_l^r - y^i)^2}}{c}, \tag{4.4}$$

and

$$f_{lk}^i = \frac{v_x^i(x_k^t - x^i) + v_y^i(y_k^t - y^i)}{\lambda\sqrt{(x_k^t - x^i)^2 + (y_k^t - y^i)^2}} + \frac{v_x^i(x_l^r - x^i) + v_y^i(y_l^r - y^i)}{\lambda\sqrt{(x_l^r - x^i)^2 + (y_l^r - y^i)^2}}, \tag{4.5}$$

where c is the speed of light and λ is the wavelength of the carrier.

4.3 Cramer-Rao Bound

We are interested in the accuracy of joint location and velocity estimation for multiple targets in this MIMO radar system. To simplify the notation, we define the following vectors

$$\mathbf{r} = [r_{11}(t), r_{12}(t), \ldots, r_{NM}(t)]^T, \tag{4.6}$$

$$\boldsymbol{\theta} = [x^1, y^1, v_x^1, v_y^1, \ldots, x^Q, y^Q, v_x^Q, v_y^Q]^T, \tag{4.7}$$

and

$$\boldsymbol{\phi} = [\tau_{11}^1, \tau_{12}^1, \ldots, \tau_{NM}^1, f_{11}^1, \ldots, f_{NM}^1, \ldots, \tau_{11}^Q, \ldots, f_{NM}^Q]^T. \tag{4.8}$$

The MN-dimensional vector $\mathbf{r}$ contains the observed signals at the receivers. The $4Q$-dimensional vector $\boldsymbol{\theta}$ contains the unknown location and velocity parameters. The $2MNQ$-dimensional vector $\boldsymbol{\phi}$ contains the unknown time delays and Doppler shifts. As we know, the CRB provides a bound on the smallest mean-square error which can be achieved by any unbiased estimator. The $4Q \times 4Q$ Cramer-Rao bound matrix $\boldsymbol{C}_{CRB}$ is expressed as:

$$\boldsymbol{C}_{CRB} = \begin{bmatrix} c_x^1 & C_{1,2} & C_{1,3} & C_{1,4} & \cdots & \cdots & \cdots & \cdots & C_{1,4Q} \\ C_{2,1} & c_y^1 & C_{2,3} & C_{2,4} & \cdots & \cdots & \cdots & \cdots & C_{2,4Q} \\ C_{3,1} & C_{3,2} & c_{v_x}^1 & C_{3,4} & \cdots & \cdots & \cdots & \cdots & C_{3,4Q} \\ C_{4,1} & C_{4,2} & C_{4,3} & c_{v_y}^1 & \cdots & \cdots & \cdots & \cdots & C_{4,4Q} \\ C_{5,1} & C_{5,2} & C_{5,3} & C_{5,4} & c_x^2 & \cdots & \cdots & \cdots & C_{5,4Q} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ C_{8,1} & C_{8,2} & C_{8,3} & C_{8,4} & \cdots & c_{v_y}^2 & \cdots & \cdots & C_{8,4Q} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ C_{4Q,1} & C_{4Q,2} & C_{4Q,3} & C_{4Q,4} & \cdots & \cdots & \cdots & \cdots & c_{v_y}^Q \end{bmatrix}, \tag{4.9}$$

where $c_x^q, c_y^q, c_{v_x}^q$ and $c_{v_y}^q$ $(1 \leqslant q \leqslant Q)$, the diagonal elements of $\boldsymbol{C}_{CRB}$, give bounds of the variances for the estimates of parameters x^q, y^q, v_x^q and v_y^q, respectively. To obtain $\boldsymbol{C}_{CRB}$, we need the Fisher Information Matrix (FIM) defined by:

$$\mathbf{J}(\boldsymbol{\theta}) = E_{\mathbf{r}} \left\{ \frac{\partial \log p(\mathbf{r}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \left(\frac{\partial \log p(\mathbf{r}|\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^H \right\}, \tag{4.10}$$

where $p(\mathbf{r}|\boldsymbol{\theta})$ is the joint pdf of vector $\mathbf{r}$ conditioned on $\boldsymbol{\theta}$. We have:

$$\boldsymbol{C}_{CRB} = \mathbf{J}(\boldsymbol{\theta})^{-1}. \tag{4.11}$$

According to equation (4.3), $\mathbf{r}$ can be written as a function of $\boldsymbol{\phi}$. Therefore, considering $\boldsymbol{\phi}$ as an intermediate variable and applying the chain rule, $\mathbf{J}(\boldsymbol{\theta})$ can be expressed as:

$$\mathbf{J}(\boldsymbol{\theta}) = \frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\theta}} \mathbf{J}(\boldsymbol{\phi}) \left(\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\theta}} \right)^T, \tag{4.12}$$

where $\mathbf{J}(\boldsymbol{\phi})$ is the FIM for the unknown vector $\boldsymbol{\phi}$. To calculate $\mathbf{J}(\boldsymbol{\phi})$, we need to know $p(\mathbf{r}|\boldsymbol{\phi})$, the joint pdf of $\mathbf{r}$ conditioned on the unknown parameter $\boldsymbol{\phi}$. From equation (4.3), it is easy to see that for different l's or k's, the received signals $r_{lk}(t)$'s are independent of each other, because ζ_{lk}^i's are independent of each other, and so are the $w_{lk}(t)$'s. Therefore, we have

$$\log p(\mathbf{r}|\boldsymbol{\phi}) = \sum_{l=1}^{N} \sum_{k=1}^{M} \log p(r_{lk}(t)|\boldsymbol{\phi}). \tag{4.13}$$

For given ζ_{lk}^i's, we have

$$p(r_{lk}(t)|\boldsymbol{\phi}, \boldsymbol{\zeta}_{lk}) = C_1 \exp\Big(- \int_T \big| r_{lk}(t) - \sqrt{\frac{E}{M}} \sum_{i=1}^{Q} \zeta_{lk}^i s_k(t - \tau_{lk}^i) e^{j2\pi f_{lk}^i t} \big|^2 dt \Big), \tag{4.14}$$

where $\boldsymbol{\zeta}_{lk} = [\zeta_{lk}^1, \zeta_{lk}^2, \ldots, \zeta_{lk}^Q]^T$ and C_1 is independent of $\boldsymbol{\theta}$ and $\boldsymbol{\zeta}_{lk}$. Since $\boldsymbol{\zeta}_{lk}$ is a complex Gaussian random vector, we can obtain $p(\mathbf{r}|\boldsymbol{\phi})$ by taking the expectation of equation (4.14) with respect to the distribution of the vector $\boldsymbol{\zeta}_{lk}$. Note that similar derivations can be found

in [92]. It turns out equation (4.13) can be expressed as:

$$\begin{aligned} \log p(\mathbf{r}|\boldsymbol{\phi}) &= C_2 + \sum_{l=1}^{N}\sum_{k=1}^{M}\Big(\log|\Psi_{lk}| + \sum_{i=1}^{Q}\sum_{h=1}^{Q}\Psi_{lk}(i,h)\frac{E}{M} \\ &\int_T r_{lk}^*(t)s_k(t-\tau_{lk}^i)e^{j2\pi f_{lk}^i t}dt \int_T r_{lk}(t)s_k^*(t-\tau_{lk}^h)e^{-j2\pi f_{lk}^h t}dt\Big), \end{aligned} \tag{4.15}$$

where C_2 is independent of $\boldsymbol{\theta}$, $\Psi_{lk}(i,h)$ is the (i,h)-th element of a $Q \times Q$ matrix Ψ_{lk}, $|\Psi_{lk}|$ represents the determinate of Ψ_{lk} and Ψ_{lk} is defined as

$$\Psi_{lk} = \Omega_{lk}^{-1}, \tag{4.16}$$

where Ω_{lk} is also a $Q \times Q$ matrix and its (i,h)-th element is defined as

$$\Omega_{lk}(i,h) = \begin{cases} \frac{E}{M}\int_T s_k^*(t-\tau_{lk}^i)s_k(t-\tau_{lk}^h) \\ e^{j2\pi t(f_{lk}^h - f_{lk}^i)}dt, \text{ if } i \neq h, \\ 1+\frac{E}{M}, \text{ otherwise.} \end{cases} \tag{4.17}$$

As shown in equation (4.15), $\log p(\mathbf{r}|\boldsymbol{\phi})$ is a function of $\boldsymbol{\phi}$. Therefore, we can obtain $\mathbf{J}(\boldsymbol{\phi})$ by first taking second order partial derivatives of $\log p(\mathbf{r}|\boldsymbol{\phi})$ with respect to $\boldsymbol{\phi}$, then calculating the expectation of its Hessian matrix with respect to the distribution of $\mathbf{r}$:

$$\mathbf{J}(\boldsymbol{\phi}) = E_{\mathbf{r}}\left\{\frac{\partial \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}\left(\frac{\partial \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}\right)^H\right\}, \tag{4.18}$$

which can be further expressed as:

$$\mathbf{J}(\boldsymbol{\phi}) = \begin{bmatrix} B_{11} & B_{12} & \cdots & B_{1Q} \\ B_{21} & B_{22} & \cdots & B_{2Q} \\ \cdots & \cdots & \cdots & \cdots \\ B_{Q1} & B_{Q2} & \cdots & B_{QQ} \end{bmatrix}, \tag{4.19}$$

where for $1 \leqslant i \leqslant Q$ and $1 \leqslant h \leqslant Q$,

$$B_{ih} = \begin{bmatrix} C_{ih} & D_{ih} \\ D_{ih} & E_{ih} \end{bmatrix}, \tag{4.20}$$

$$C_{ih} = \mathrm{diag}\left(E_{\mathbf{r}} \left\{ - \left[\frac{\partial^2 \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial \tau_{11}^i \partial \tau_{11}^h}, \frac{\partial^2 \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial \tau_{12}^i \partial \tau_{12}^h}, \ldots, \frac{\partial^2 \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial \tau_{NM}^i \partial \tau_{NM}^h} \right] \right\} \right), \tag{4.21}$$

$$D_{ih} = \mathrm{diag}\left(E_{\mathbf{r}} \left\{ - \left[\frac{\partial^2 \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial \tau_{11}^i \partial f_{11}^h}, \frac{\partial^2 \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial \tau_{12}^i \partial f_{12}^h}, \ldots, \frac{\partial^2 \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial \tau_{NM}^i \partial f_{NM}^h} \right] \right\} \right), \tag{4.22}$$

$$E_{ih} = \mathrm{diag}\left(E_{\mathbf{r}} \left\{ - \left[\frac{\partial^2 \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial f_{11}^i \partial f_{11}^h}, \frac{\partial^2 \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial f_{12}^i \partial f_{12}^h}, \ldots, \frac{\partial^2 \log p(\mathbf{r}|\boldsymbol{\phi})}{\partial f_{NM}^i \partial f_{NM}^h} \right] \right\} \right). \tag{4.23}$$

Note that the detailed expressions of equations (4.21)–(4.23) are given in the appendix of this chapter. $\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\theta}}$ from equation (4.12) is given by

$$\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\theta}} = \mathrm{diag}\left[A_1, A_2, \ldots, A_Q \right], \tag{4.24}$$

where A_i is given as follows:

$$A_i = \begin{bmatrix} \frac{\partial \tau_{11}^i}{\partial x^i} & \frac{\partial \tau_{12}^i}{\partial x^i} & \cdots & \frac{\partial \tau_{NM}^i}{\partial x^i} & \frac{\partial f_{11}^i}{\partial x^i} & \frac{\partial f_{12}^i}{\partial x^i} & \cdots & \frac{\partial f_{NM}^i}{\partial x^i} \\ \frac{\partial \tau_{11}^i}{\partial y^i} & \frac{\partial \tau_{12}^i}{\partial y^i} & \cdots & \frac{\partial \tau_{NM}^i}{\partial y^i} & \frac{\partial f_{11}^i}{\partial y^i} & \frac{\partial f_{12}^i}{\partial y^i} & \cdots & \frac{\partial f_{NM}^i}{\partial y^i} \\ 0 & 0 & \cdots & 0 & \frac{\partial f_{11}^i}{\partial v_x^i} & \frac{\partial f_{12}^i}{\partial v_x^i} & \cdots & \frac{\partial f_{NM}^i}{\partial v_x^i} \\ 0 & 0 & \cdots & 0 & \frac{\partial f_{11}^i}{\partial v_y^i} & \frac{\partial f_{12}^i}{\partial v_y^i} & \cdots & \frac{\partial f_{NM}^i}{\partial v_y^i} \end{bmatrix}, \tag{4.25}$$

where for $1 \leqslant i \leqslant Q$, $1 \leqslant l \leqslant N$ and $1 \leqslant k \leqslant M$,

$$\frac{\partial \tau_{lk}^i}{\partial x^i} = \frac{1}{c} \left(\frac{x^i - x_k^t}{\sqrt{(x_k^t - x^i)^2 + (y_k^t - y^i)^2}} + \frac{x^i - x_l^r}{\sqrt{(x_l^r - x^i)^2 + (y_l^r - y^i)^2}} \right), \tag{4.26}$$

$$\frac{\partial \tau_{lk}^i}{\partial y^i} = \frac{1}{c} \left(\frac{y^i - y_k^t}{\sqrt{(x_k^t - x^i)^2 + (y_k^t - y^i)^2}} + \frac{y^i - y_l^r}{\sqrt{(x_l^r - x^i)^2 + (y_l^r - y^i)^2}} \right), \tag{4.27}$$

$$\frac{\partial f_{lk}^i}{\partial x^i} = (x_k^t - x^i)\frac{v_x^i(x_k^t - x^i) + v_y^i(y_k^t - y^i)}{\lambda\left((x_k^t - x^i)^2 + (y_k^t - y^i)^2\right)^{\frac{3}{2}}} - \frac{v_x^i}{\lambda\sqrt{(x_k^t - x^i)^2 + (y_k^t - y^i)^2}}$$
$$+(x_l^r - x^i)\frac{v_x^i(x_l^r - x^i) + v_y^i(y_l^r - y^i)}{\lambda\left((x_l^r - x^i)^2 + (y_l^r - y^i)^2\right)^{\frac{3}{2}}} - \frac{v_x^i}{\lambda\sqrt{(x_l^r - x^i)^2 + (y_l^r - y^i)^2}}, \quad (4.28)$$

$$\frac{\partial f_{lk}^i}{\partial y^i} = (y_k^t - y^i)\frac{v_x^i(x_k^t - x^i) + v_y^i(y_k^t - y^i)}{\lambda\left((x_k^t - x^i)^2 + (y_k^t - y^i)^2\right)^{\frac{3}{2}}} - \frac{v_y^i}{\lambda\sqrt{(x_k^t - x^i)^2 + (y_k^t - y^i)^2}}$$
$$+(y_l^r - y^i)\frac{v_x^i(x_l^r - x^i) + v_y^i(y_l^r - y^i)}{\lambda\left((x_l^r - x^i)^2 + (y_l^r - y^i)^2\right)^{\frac{3}{2}}} - \frac{v_y^i}{\lambda\sqrt{(x_l^r - x^i)^2 + (y_l^r - y^i)^2}}, \quad (4.29)$$

$$\frac{\partial f_{lk}^i}{\partial v_x^i} = \frac{x_k^t - x^i}{\lambda\sqrt{(x_k^t - x^i)^2 + (y_k^t - y^i)^2}} + \frac{x_l^r - x^i}{\lambda\sqrt{(x_l^r - x^i)^2 + (y_l^r - y^i)^2}}, \quad (4.30)$$

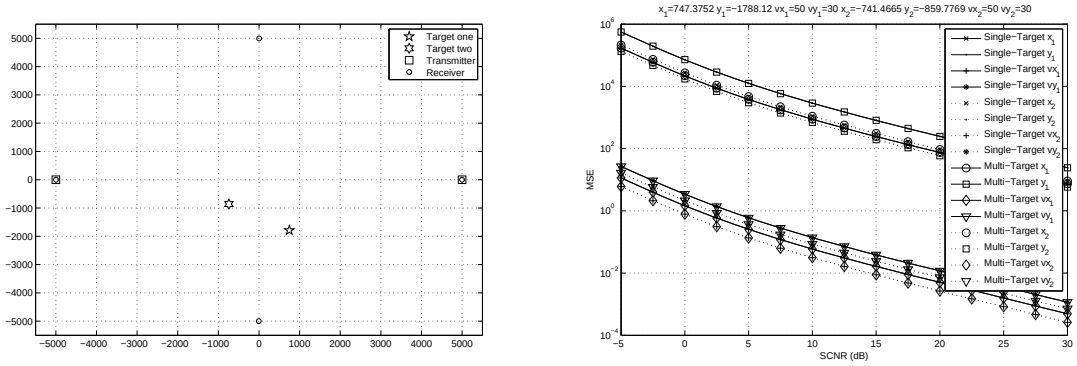
$$\frac{\partial f_{lk}^i}{\partial v_y^i} = \frac{y_k^t - y^i}{\lambda\sqrt{(x_k^t - x^i)^2 + (y_k^t - y^i)^2}} + \frac{y_l^r - y^i}{\lambda\sqrt{(x_l^r - x^i)^2 + (y_l^r - y^i)^2}}. \quad (4.31)$$

Substituting equations (4.19)–(4.31) into equation (4.10), we obtain the closed-form expression of the Cramer-Rao bound for joint location and velocity estimation in a multi-target MIMO Radar system. In the rest of this chapter, we only consider the two-target case, i.e., $Q = 2$.

4.4 Numerical Analysis

In this section, we calculate the Cramer-Rao bounds derived in this chapter and use them to study the behavior of a non-coherent MIMO radar with two targets. The following system parameters are used in the whole section. The carrier frequency is $1GHz$. The transmitted lowpass equivalent waveform is $s_k(t) = (\frac{2}{T^2})^{\frac{1}{4}} \exp\left(\frac{-\pi t^2}{T^2} j2\pi k\Delta f t\right)$, where T is proportional to the pulse width and $\Delta f = f_{k+1} - f_k$ is the frequency increment between $s_{k+1}(t)$ and $s_k(t)$.

Theorem 4 *Consider a MIMO radar system with M transmitters, N receivers and Q tar-*

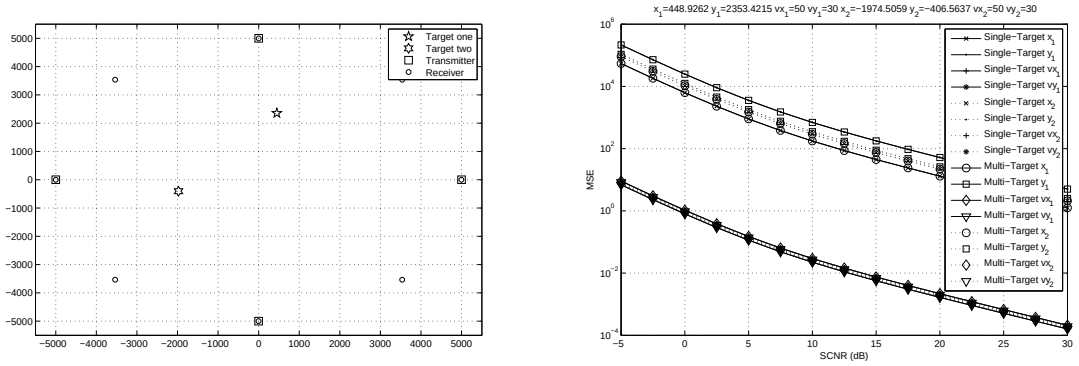


(a) (b)

Figure 4.1: Comparison between multi-target CRBs and single-target CRBs for a 2×4 MIMO radar system: (a) system layout, (b) CRBs.

gets. $c_x^q, c_y^q, c_{v_x}^q$ *and* $c_{v_y}^q$ $(1 \leqslant q \leqslant Q)$ *represent the Cramer-Rao bounds for the estimates of parameters* x^q, y^q, v_x^q *and* v_y^q*, respectively. They are called the multi-target Cramer-Rao bounds and can be calculated by using equation (4.11). Then we keep only one target in the system and remove the others, respectively.* $c_x^{q\prime}, c_y^{q\prime}, {c_{v_x}^q}'$ *and* ${c_{v_y}^q}'$ $(1 \leqslant q \leqslant Q)$ *indicate the Cramer-Rao bounds for target q when only this target is left in the system. They are called the single-target Cramer-Rao bounds and can be calculated by using results from [35]. If in the multi-target case, all targets are far away from each other so that the time delay differences between any two targets are large enough, i.e., for* $1 \leqslant i, j \leqslant Q$ *and* $(i \neq j)$*,* $|\tau_{lk}^i - \tau_{lk}^j| \gg 0$ *for any* $1 \leqslant l \leqslant N$ *and* $1 \leqslant k \leqslant M$*, then we have:* $c_x^q \approx c_x^{q\prime}$*,* $c_y^q \approx c_y^{q\prime}$*,* $c_{v_x}^q \approx {c_{v_x}^q}'$ *and* $c_{v_y}^q \approx {c_{v_y}^q}'$ *(*$1 \leqslant q \leqslant Q$*).*

Theorem 4 says, if the distances between the targets are large enough, then the interactions between the multiple targets can be ignored. Thus each target can be treated separately as it is treated in a single-target joint estimation problem. The proof of Theorem 4 is in the appendix of this chapter. We use this property to validate, at least in this one case, the two



(a) (b)

Figure 4.2: Comparison between multi-target CRBs and single-target CRBs for a 4×8 MIMO radar system: (a) system layout, (b) CRBs.

target joint location and velocity estimation Cramer-Rao bound derived in this chapter.

Consider a 2×4 MIMO radar system with two targets moving at the velocity of $(50, 30)(m/s)$. The transmitters and receivers are uniformly spaced along a circle of radius $5000(m)$. The locations of each antenna and target are shown in Fig.4.1 (a). Note that the distance between the two targets are chosen specifically to make the interactions between them small, which are determined by $\int_T s_k^*(t-\tau_{lk}^i)s_k(t-\tau_{lk}^j)e^{j2\pi t(f_{lk}^j - f_{lk}^i)}dt$ from equation (4.17). Here a special parameter ρ is defined as:

$$\rho = \max_{l,k,i\neq j} \int_T s_k^*(t-\tau_{lk}^i)s_k(t-\tau_{lk}^j)e^{j2\pi t(f_{lk}^j - f_{lk}^i)}dt. \tag{4.32}$$

The smaller the value of ρ is, the little influence each target has on the estimation performance of the other. In this case, $\rho = 6.2407 \times 10^{-5}$.

Fig.4.1 (b) provides a set of Cramer-Rao bound versus SCNR curves. The curves using circles and polygons as their markers represent the multi-target Cramer-Rao bounds. Those curves with other markers denote the single-target Cramer-Rao bounds with one target still at its location and the other one being removed, which are calculated using results from [35].

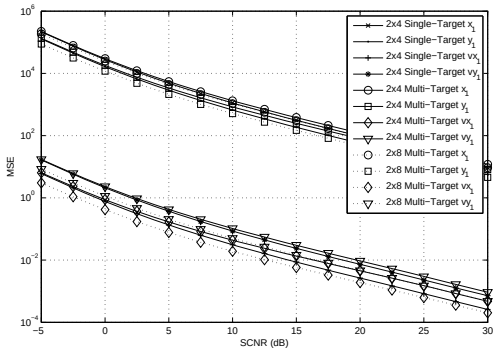


Figure 4.3: CRB versus SCNR curves for close targets.

The solid lines denotes the bounds for the first target and the dotted lines the second. Note that there are four bounds for each target because we are estimating the location and velocity in both x and y coordinates. As seen in Fig.4.1 (b), the multi-target bounds and the corresponding single-target bounds for each target are so close that they seem to overlap each other, which is exactly what Theorem 4 predicts.

We also perform the comparison for a 4×8 system, whose layout is shown in Fig.4.2 (a). Velocities of the targets are still $(50, 30)(m/s)$. The multi-target and corresponding single-target Cramer-Rao bounds are drawn in Fig.4.2 (b). In this case, $\rho = 2.3831 \times 10^{-14}$. Without any surprise, the multi-target bounds still match the single-target bounds very well.

When the targets are close to each other, often the estimation performance for each target is worse than that in the single-target case. We analyze such a situation and present the results in Fig.4.3. First, a 2×4 MIMO radar system is considered. The transmitter and receiver settings are the same as in Fig.4.1 (a), but the locations of the targets are changed. In this case, the first target is at $(800, 800)(m)$ and the second target is at $(1000, 1000)(m)$. Their velocities are $(50, 30)(m/s)$. $\rho = 0.7636$ in this case. In Fig.4.3, the solid curves with circles and polygons represent the multi-target Cramer-Rao bounds for the first target (Here

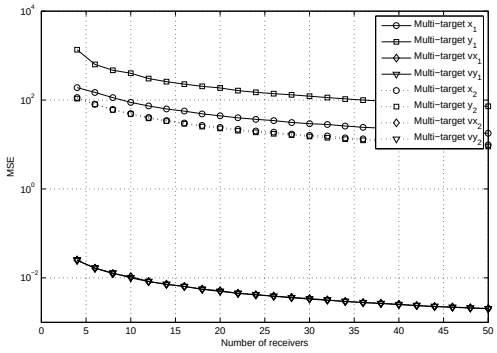


Figure 4.4: CRB versus the number of receivers for a MIMO radar system with four transmitters.

we only analyze the first target because results for the second target are very similar). The solid curves with other markers provide the single-target bounds for the first target. Clearly, the single-target bounds are a lot better than the multi-target ones. Then we add four more receivers to the system (the total eight receivers are uniformly distributed in the original circle) and study how the estimation performance will change in the multi-target case. The dotted curves in Fig.4.3 represent the multi-target Cramer-Rao bounds for the first target in this 2×8 MIMO radar system. We can see that they are below the solid curves, which means in this case, the performance degradation introduced by the multiple targets is compensated by using more antennas at the receiver.

Now let us take a deeper look into the how the number of receivers improves the estimation accuracy. Consider a MIMO radar with transmitters and receivers still uniformly distributed in a circle of radius $5000(m)$. The four transmitter are located as shown in Fig.4.2 (a). Two targets are located at $(150, 150)(m)$ and $(-800, -100)(m)$, respectively. Their velocities are $(50, 50)(m/s)$ and $(70, -70)(m/s)$, respectively. We fix the number of transmitters at four and the SCNR at $30(dB)$, and observe how the performance changes as

we increase the number of receivers. The performance of MIMO radar systems with different numbers of receivers are compared and the Cramer-Rao bounds for each target are provided in Fig.4.4. In Fig.4.4, the solid curves are for the the first target and the dotted curves are for the second. As we see, increasing the number of receivers gives rise to lower Cramer-Rao bounds. This indicates that the spatial advantage of MIMO radars can also be observed in this two-target case.

4.5 Summary

In this chapter, we investigate the joint location and velocity estimation of multiple targets in a non-coherent MIMO radar system. The Cramer-Rao bound for a two-target case is derived and evaluated. This bound gives us theoretically achievable joint estimation performance of a multi-target non-coherent MIMO radar. Numerical simulations show that the spatial advantage proposed by previous research can also be observed in a two-target example.

4.6 Appendix

4.6.1 Detailed expressions of equations (4.21)–(4.23)

Let us define

$$L = \log p(\mathbf{r}|\boldsymbol{\phi}). \tag{4.33}$$

$$-E\left\{\frac{\partial^2 L}{\partial {\tau_{lk}^i}^2}\right\} = -\frac{\partial^2 \ln|\Psi_{lk}|}{\partial {\tau_{lk}^i}^2} - \frac{E}{M}\sum_{m=1}^{Q}\sum_{h=1}^{Q}\frac{\partial^2 \Psi_{lk}(m,h)}{\partial {\tau_{lk}^i}^2} E\left\{a_{lk}(m,h)\right\} \tag{4.34}$$
$$-\frac{2E}{M}\sum_{m=1}^{Q}\sum_{h=1}^{Q}\frac{\partial \Psi_{lk}(m,h)}{\partial \tau_{lk}^i} E\left\{\frac{\partial a_{lk}(m,h)}{\partial \tau_{lk}^i}\right\}$$
$$-\frac{E}{M}\sum_{m=1}^{Q}\sum_{h=1}^{Q}\Psi_{lk}(m,h) E\left\{\frac{\partial^2 a_{lk}(m,h)}{\partial {\tau_{lk}^i}^2}\right\}$$

$$a_{lk}(m,h) = \int_T r_{lk}^*(t) s_k(t-\tau_{lk}^m) e^{j2\pi f_{lk}^m} dt \int_T r_{lk}(t) s_k^*(t-\tau_{lk}^h) e^{-j2\pi f_{lk}^h} dt \tag{4.35}$$

$$\frac{\partial^2 \ln|\Psi_{lk}|}{\partial {\tau_{lk}^i}^2} = \frac{1}{|\Psi_{lk}^{-1}|^2}\left|\frac{\partial|\Psi_{lk}^{-1}|}{\partial \tau_{lk}^i}\right|^2 - \frac{1}{|\Psi_{lk}^{-1}|}\frac{\partial^2|\Psi_{lk}^{-1}|}{\partial {\tau_{lk}^i}^2} \tag{4.36}$$

$$|\Psi_{lk}^{-1}| = (1+\frac{E}{M})^2 - \tag{4.37}$$
$$\frac{E^2}{M^2}\int_T s_k^*(t-\tau_{lk}^1) s_k(t-\tau_{lk}^2) e^{j2\pi t(f_{lk}^2 - f_{lk}^1)} dt \int_T s_k^*(t-\tau_{lk}^2) s_k(t-\tau_{lk}^1) e^{j2\pi t(f_{lk}^1 - f_{lk}^2)} dt$$

$$\Psi_{lk}^{-1} = \frac{1}{|\Psi_{lk}^{-1}|} \tag{4.38}$$
$$\begin{bmatrix} 1+\frac{E}{M} & -\frac{E}{M}\int_T s_k^*(t-\tau_{lk}^1) s_k(t-\tau_{lk}^2) e^{j2\pi t(f_{lk}^2 - f_{lk}^1)} dt \\ -\frac{E}{M}\int_T s_k^*(t-\tau_{lk}^2) s_k(t-\tau_{lk}^1) e^{j2\pi t(f_{lk}^1 - f_{lk}^2)} dt & 1+\frac{E}{M} \end{bmatrix}$$

$$\frac{\partial|\Psi_{lk}^{-1}|}{\partial \tau_{lk}^1} = \frac{2E^2}{M^2} \tag{4.39}$$
$$\Re\left\{\int_T s_k^{*\prime}(t-\tau_{lk}^1) s_k(t-\tau_{lk}^2) e^{j2\pi t(f_{lk}^2 - f_{lk}^1)} dt \int_T s_k^*(t-\tau_{lk}^2) s_k(t-\tau_{lk}^1) e^{j2\pi t(f_{lk}^1 - f_{lk}^2)} dt\right\}$$

$$\frac{\partial|\Psi_{lk}^{-1}|}{\partial \tau_{lk}^2} = \frac{2E^2}{M^2} \tag{4.40}$$
$$\Re\left\{\int_T s_k^*(t-\tau_{lk}^1) s_k'(t-\tau_{lk}^2) e^{j2\pi t(f_{lk}^2 - f_{lk}^1)} dt \int_T s_k^*(t-\tau_{lk}^2) s_k(t-\tau_{lk}^1) e^{j2\pi t(f_{lk}^1 - f_{lk}^2)} dt\right\}$$

$$\frac{\partial^2|\Psi_{lk}^{-1}|}{\partial {\tau_{lk}^1}^2} = -\frac{2E^2}{M^2} \tag{4.41}$$

$$\Re\Big\{ \int_T s_k^{*\prime\prime}(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-flk^1)}dt \int_T s_k^*(t-\tau_{lk}^2)s_k(t-\tau_{lk}^1)e^{j2\pi t(f_{lk}^1-flk^2)}dt$$
$$+ \int_T s_k^{*\prime}(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-flk^1)}dt \int_T s_k^*(t-\tau_{lk}^2)s_k'(t-\tau_{lk}^1)e^{j2\pi t(f_{lk}^1-flk^2)}dt\Big\}$$

$$\frac{\partial^2|\Psi_{lk}^{-1}|}{\partial {\tau_{lk}^2}^2} = -\frac{2E^2}{M^2} \tag{4.42}$$

$$\Re\Big\{ \int_T s_k^*(t-\tau_{lk}^1)s_k^{\prime\prime}(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-flk^1)}dt \int_T s_k^*(t-\tau_{lk}^2)s_k(t-\tau_{lk}^1)e^{j2\pi t(f_{lk}^1-flk^2)}dt$$
$$+ \int_T s_k^{*\prime}(t-\tau_{lk}^1)s_k'(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-flk^1)}dt \int_T s_k^{*\prime}(t-\tau_{lk}^2)s_k(t-\tau_{lk}^1)e^{j2\pi t(f_{lk}^1-flk^2)}dt\Big\}$$

$$\frac{\partial\Psi_{lk}(1,1)}{\partial\tau_{lk}^i} = \frac{\partial\Psi_{lk}(2,2)}{\partial\tau_{lk}^i} = -(1+\frac{E}{M})\frac{1}{|\Psi_{lk}^{-1}|^2}\frac{\partial|\Psi_{lk}^{-1}|}{\partial\tau_{lk}^i} \tag{4.43}$$

$$\frac{\partial\Psi_{lk}(1,2)}{\partial\tau_{lk}^1} = \frac{1}{|\Psi_{lk}^{-1}|^2}\frac{\partial|\Psi_{lk}^{-1}|}{\partial\tau_{lk}^1}\frac{E}{M}\int_T s_k^*(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt + \frac{1}{|\Psi_{lk}^{-1}|}\frac{E}{M}\int_T s_k^{*\prime}(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt \tag{4.44}$$

$$\frac{\partial\Psi_{lk}(1,2)}{\partial\tau_{lk}^2} = \frac{1}{|\Psi_{lk}^{-1}|^2}\frac{\partial|\Psi_{lk}^{-1}|}{\partial\tau_{lk}^2}\frac{E}{M}\int_T s_k^*(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt + \frac{1}{|\Psi_{lk}^{-1}|}\frac{E}{M}\int_T s_k^*(t-\tau_{lk}^1)s_k'(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt \tag{4.45}$$

$$\frac{\partial\Psi_{lk}(2,1)}{\partial\tau_{lk}^i} = \left(\frac{\partial\Psi_{lk}(1,2)}{\partial\tau_{lk}^i}\right)^* \tag{4.46}$$

$$\frac{\partial^2\Psi_{lk}(1,1)}{\partial {\tau_{lk}^i}^2} = \frac{\partial^2\Psi_{lk}(2,2)}{\partial {\tau_{lk}^i}^2} = (1+\frac{E}{M})\left(\frac{2}{|\Psi_{lk}^{-1}|^3\left(\frac{\partial|\Psi_{lk}^{-1}|}{\partial\tau_{lk}^i}\right)^2} - \frac{1}{|\Psi_{lk}^{-1}|^2}\frac{\partial^2|\Psi_{lk}^{-1}|}{\partial {\tau_{lk}^i}^2}\right) \tag{4.47}$$

$$\frac{\partial^2 \Psi_{lk}(1,2)}{\partial {\tau_{lk}^1}^2} = -\frac{2}{|\Psi_{lk}^{-1}|^3}\left(\frac{\partial|\Psi_{lk}^{-1}|}{\partial \tau_{lk}^1}\right)^2 \frac{E}{M}\int_T s_k^*(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt \tag{4.48}$$
$$+\frac{1}{|\Psi_{lk}^{-1}|^2}\frac{\partial^2|\Psi_{lk}^{-1}|}{\partial {\tau_{lk}^1}^2}\frac{E}{M}\int_T s_k^*(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt$$
$$-\frac{2}{|\Psi_{lk}^{-1}|^2}\frac{\partial|\Psi_{lk}^{-1}|}{\partial \tau_{lk}^1}\frac{E}{M}\int_T s_k^{*\prime}(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt$$
$$-\frac{1}{|\Psi_{lk}^{-1}|}\frac{E}{M}\int_T s_k^{*\prime\prime}(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt$$

$$\frac{\partial^2 \Psi_{lk}(1,2)}{\partial {\tau_{lk}^2}^2} = -\frac{2}{|\Psi_{lk}^{-1}|^3}\left(\frac{\partial|\Psi_{lk}^{-1}|}{\partial \tau_{lk}^2}\right)^2 \frac{E}{M}\int_T s_k^*(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt \tag{4.49}$$
$$+\frac{1}{|\Psi_{lk}^{-1}|^2}\frac{\partial^2|\Psi_{lk}^{-1}|}{\partial {\tau_{lk}^2}^2}\frac{E}{M}\int_T s_k^*(t-\tau_{lk}^1)s_k(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt$$
$$-\frac{2}{|\Psi_{lk}^{-1}|^2}\frac{\partial|\Psi_{lk}^{-1}|}{\partial \tau_{lk}^2}\frac{E}{M}\int_T s_k^*(t-\tau_{lk}^1)s_k'(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt$$
$$-\frac{1}{|\Psi_{lk}^{-1}|}\frac{E}{M}\int_T s_k^*(t-\tau_{lk}^1)s_k''(t-\tau_{lk}^2)e^{j2\pi t(f_{lk}^2-f_{lk}^1)}dt$$

$$\frac{\partial^2 \Psi_{lk}(2,1)}{\partial {\tau_{lk}^i}^2} = \left(\frac{\partial^2 \Psi_{lk}(1,2)}{\partial {\tau_{lk}^i}^2}\right)^* \tag{4.50}$$

$$E\left\{a_{lk}(m,h)\right\} = \frac{E}{M}\sum_{i=1}^{Q}\int_T s_k^*(t-\tau_{lk}^i)s_k(t-\tau_{lk}^m)e^{j2\pi t(f_{lk}^m-f_{lk}^i)}dt \tag{4.51}$$
$$\int_T s_k(t-\tau_{lk}^i)s_k^*(t-\tau_{lk}^h)e^{j2\pi t(f_{lk}^i-f_{lk}^h)}dt + \int_T s_k(t-\tau_{lk}^m)s_k^*(t-\tau_{lk}^h)e^{j2\pi t(f_{lk}^m-f_{lk}^h)}dt$$

If $m = i$ and $h \neq i$,

$$E\left\{\frac{\partial a_{lk}(m,h)}{\partial \tau_{lk}^i}\right\} = -\frac{E}{M}\sum_{n=1}^{Q}\int_T s_k^*(t-\tau_{lk}^n)s_k'(t-\tau_{lk}^m)e^{j2\pi t(f_{lk}^m-f_{lk}^n)}dt \tag{4.52}$$
$$\int_T s_k(t-\tau_{lk}^n)s_k^*(t-\tau_{lk}^h)e^{j2\pi t(f_{lk}^n-f_{lk}^h)}dt - \int_T s_k'(t-\tau_{lk}^m)s_k^*(t-\tau_{lk}^h)e^{j2\pi t(f_{lk}^m-f_{lk}^h)}dt$$

If $m \neq i$ and $h = i$,

$$E\left\{\frac{\partial a_{lk}(m,h)}{\partial \tau_{lk}^{i}}\right\} = -\frac{E}{M}\sum_{n=1}^{Q}\int_{T} s_k^*(t-\tau_{lk}^{n}) s_k(t-\tau_{lk}^{m}) e^{j2\pi t(f_{lk}^{m}-f_{lk}^{n})}dt \tag{4.53}$$
$$\int_{T} s_k(t-\tau_{lk}^{n}) s_k^{*\prime}(t-\tau_{lk}^{h}) e^{j2\pi t(f_{lk}^{n}-f_{lk}^{h})}dt - \int_{T} s_k(t-\tau_{lk}^{m}) s_k^{*\prime}(t-\tau_{lk}^{h}) e^{j2\pi t(f_{lk}^{m}-f_{lk}^{h})}dt$$

If $m = i$ and $h = i$,

$$E\left\{\frac{\partial a_{lk}(m,h)}{\partial \tau_{lk}^{i}}\right\} = -2\Re\Big\{\frac{E}{M}\sum_{n=1}^{Q}\int_{T} s_k^*(t-\tau_{lk}^{n}) s_k^{\prime}(t-\tau_{lk}^{m}) e^{j2\pi t(f_{lk}^{m}-f_{lk}^{n})}dt \tag{4.54}$$
$$\int_{T} s_k(t-\tau_{lk}^{n}) s_k^{*}(t-\tau_{lk}^{h}) e^{j2\pi t(f_{lk}^{n}-f_{lk}^{h})}dt + \int_{T} s_k^{\prime}(t-\tau_{lk}^{m}) s_k^{*}(t-\tau_{lk}^{h}) e^{j2\pi t(f_{lk}^{m}-f_{lk}^{h})}dt\Big\}$$

If $m = i$ and $h \neq i$,

$$E\left\{\frac{\partial^2 a_{lk}(m,h)}{\partial {\tau_{lk}^{i}}^2}\right\} = \frac{E}{M}\sum_{n=1}^{Q}\int_{T} s_k^*(t-\tau_{lk}^{n}) s_k^{\prime\prime}(t-\tau_{lk}^{m}) e^{j2\pi t(f_{lk}^{m}-f_{lk}^{n})}dt \tag{4.55}$$
$$\int_{T} s_k(t-\tau_{lk}^{n}) s_k^{*}(t-\tau_{lk}^{h}) e^{j2\pi t(f_{lk}^{n}-f_{lk}^{h})}dt + \int_{T} s_k^{\prime\prime}(t-\tau_{lk}^{m}) s_k^{*}(t-\tau_{lk}^{h}) e^{j2\pi t(f_{lk}^{m}-f_{lk}^{h})}dt$$

If $m \neq i$ and $h = i$,

$$E\left\{\frac{\partial^2 a_{lk}(m,h)}{\partial {\tau_{lk}^{i}}^2}\right\} = \frac{E}{M}\sum_{n=1}^{Q}\int_{T} s_k^*(t-\tau_{lk}^{n}) s_k(t-\tau_{lk}^{m}) e^{j2\pi t(f_{lk}^{m}-f_{lk}^{n})}dt \tag{4.56}$$
$$\int_{T} s_k(t-\tau_{lk}^{n}) s_k^{*\prime\prime}(t-\tau_{lk}^{h}) e^{j2\pi t(f_{lk}^{n}-f_{lk}^{h})}dt + \int_{T} s_k(t-\tau_{lk}^{m}) s_k^{*\prime\prime}(t-\tau_{lk}^{h}) e^{j2\pi t(f_{lk}^{m}-f_{lk}^{h})}dt$$

If $m = i$ and $h = i$,

$$E\left\{\frac{\partial^2 a_{lk}(m,h)}{\partial {\tau_{lk}^i}^2}\right\} = -2\Re\Big\{\frac{E}{M}\sum_{n=1}^{Q}\int_T s_k^*(t-\tau_{lk}^n)s_k''(t-\tau_{lk}^m)e^{j2\pi t(f_{lk}^m - f_{lk}^n)}dt \tag{4.57}$$
$$\int_T s_k(t-\tau_{lk}^n)s_k^*(t-\tau_{lk}^h)e^{j2\pi t(f_{lk}^n - f_{lk}^h)}dt + \frac{E}{M}\sum_{n=1}^{Q}\int_T s_k^*(t-\tau_{lk}^n)s_k'(t-\tau_{lk}^m)e^{j2\pi t(f_{lk}^m - f_{lk}^n)}dt$$
$$\int_T s_k(t-\tau_{lk}^n)s_k^{*\prime}(t-\tau_{lk}^h)e^{j2\pi t(f_{lk}^n - f_{lk}^h)}dt + \int_T s_k''(t-\tau_{lk}^m)s_k^*(t-\tau_{lk}^h)dt$$
$$+\int_T s_k'(t-\tau_{lk}^m)s_k^{*\prime}(t-\tau_{lk}^h)dt\Big\}$$

4.6.2 Proof of Theorem 4

Proof: For any $1 \leqslant l \leqslant N, 1 \leqslant k \leqslant M$ and $1 \leqslant i, j \leqslant Q$ $(i \neq j)$, if $|\tau_{lk}^i - \tau_{lk}^j|$ is large enough, we have $\int_T s_k^*(t-\tau_{lk}^i)s_k(t-\tau_{lk}^j)e^{j2\pi t(f_{lk}^j - f_{lk}^i)}dt \approx 0$. According to equation (4.17), $\Psi_{lk} = \Omega_{lk}^{-1} = (1+\frac{E}{M})I_{Q\times Q}$, where $I_{Q\times Q}$ is the Q dimensional identity matrix. Then equation (4.15) can be written as:

$$\log p(\mathbf{r}|\boldsymbol{\phi}) \quad = \quad C_3 + \sum_{q=1}^{Q} L_q, \tag{4.58}$$

where $C_3 = C_2 - MNQ\log\left(1+\frac{E}{M}\right)$ is another constant and

$$L_q = \frac{E}{E+M}\sum_{l=1}^{N}\sum_{k=1}^{M}\Big|\int_T r_{lk}^*(t)s_k(t-\tau_{lk}^q)e^{j2\pi f_{lk}^q t}dt\Big|^2. \tag{4.59}$$

According to [35], L_q is the likelihood of the received signal in a non-coherent MIMO radar system with only one target q.

$\boldsymbol{\phi}$ can be rewritten as;

$$\boldsymbol{\phi} = [\boldsymbol{\phi}_1^T, \boldsymbol{\phi}_2^T, \ldots, \boldsymbol{\phi}_Q^T]^T, \tag{4.60}$$

where $\boldsymbol{\phi}_q = [\tau_{11}^q, \tau_{12}^q, \ldots, \tau_{NM}^q, f_{11}^q, \ldots, f_{NM}^q]^T, (1 \leqslant q \leqslant Q)$. Because $\frac{\partial \boldsymbol{\phi}}{\partial \boldsymbol{\theta}}$ is a block diagonal matrix (equation (4.24)), according to equations (4.12) and (4.19), only the diagonal blocks

B_{qq}'s in equation (4.19) contribute to the FIM matrix $\mathbf{J}(\boldsymbol{\theta})$. We need the partial derivatives of $\log p(\mathbf{r}|\boldsymbol{\phi})$ with respect to $\boldsymbol{\phi}_q$ to calculate B_{qq} (equations (4.20)–(4.23)). Since $\log p(\mathbf{r}|\boldsymbol{\phi})$ is written as the sum of L_q's ($1 \leqslant q \leqslant Q$), and L_q is a function of $\boldsymbol{\phi}_q$, we have

$$B_{qq} = E_{\mathbf{r}} \left\{ \frac{\partial L_q}{\partial \boldsymbol{\phi}_q} \left(\frac{\partial L_q}{\partial \boldsymbol{\phi}_q} \right)^H \right\}, \tag{4.61}$$

and

$$\mathbf{J}(\boldsymbol{\theta}) = \text{diag}\left[A_1 B_{11} A_1^T, A_2 B_{22} A_2^T, \ldots, A_Q B_{QQ} A_Q^T \right], \tag{4.62}$$

where the q-th diagonal block of $\mathbf{J}(\boldsymbol{\theta})$ is the single-target FIM matrix for target q [35].

The inverse of $\mathbf{J}(\boldsymbol{\theta})$ is the multi-target Cramer-Rao bound matrix C_{CRB}, each of whose diagonal blocks is the single-target Cramer-Rao bound matrix for the corresponding target. Therefore, to obtain C_{CRB}, we can treat the targets independently, as if every one of them is the only target in a non-coherent MIMO radar system, and calculate each target's Cramer-Rao bound matrix using results in [35]. ■

Chapter 5

Distributed Change Detection in Gaussian Graphical Models

5.1 Introduction

Change detection is a classical and fundamental problem in statistical signal processing. Detection is usually performed by testing the null hypothesis that the observed samples originate from a known nominal distribution. In the Gaussian case this reduces to testing whether the sample mean or the sample covariance deviates from some given nominal parameters. Such tests are well known and can easily detect a change when the number of samples n is sufficiently large in comparison to the dimension p, and there are no practical complexity or communication constraints. Unfortunately, these assumptions are impractical when monitoring modern large scale networks, which involve high dimensional parametric models (large p) and dynamic behavior in short stationary periods (small n). Centralized processing is difficult and sensitive to attacks in large networks. This leads to an ongoing

search for efficient distributed change detection techniques. We propose to address both of these problems through the use of Gaussian graphical models (GGM) which allow accurate and robust statistical analysis in the "large p small n" regime via message passing algorithms.

A graphical model characterizes conditional dependencies within a multivariate distribution using undirected graphs [45,46]. When the graph is sparse and the variables are jointly Gaussian, the graphical model imposes sparsity on the inverse covariance, variously called the information, concentration or precision matrix. Graphical models are attractive in two complementary manners. First, they allow the modeling of a large scale distribution using a smaller number of parameters which may be easier to handle. Second, inference methods can be performed as local decentralized computations with message passing between neighbors of the graph. For this purpose, it is common to assume that the topology of the statistical models matches the topology of the internode communications. Based on this framework, GGMs have been successfully applied to different statistical problems such as inference [47], parameter estimation [48, 49], and dimensionality reduction [50]. Applications range from wireless sensor networks [51, 52], internet backbone networks [50] and recently the smart power grid [53].

The time-tested approach to change detection is via composite hypothesis testing, and in particular the generalized likelihood ratio test (GLRT) [54]. This method has been successfully applied to change detection in the Gaussian multivariate distribution, and is known as the Bartlett's test [55, 56]. Classical centralized results on testing in GGM are available in [45]. In contrast, our focus is on distributed solutions for change detection. We begin by deriving the global centralized GLRT, and then propose two distributed approximations based on aggregating multiple tests performed at each of the nodes in the graph. The aggregation can then be efficiently implemented using modern consensus methods [57]. The

first distributed test is a natural approach which simply applies the Bartlett's test to smaller size local clusters in the graph. The second method employs the pseudo-likelihood [58] as a surrogate function for the global likelihood. The advantages of our proposed tests are demonstrated using numerical experiments, including one in the context of failure detection in smart power grids [53].

For completeness, two comments are in order. First, composite hypothesis testing via GLRT relies on a first phase of parameter estimation, usually based on ML techniques. This chapter only focuses on the detection phase. Distributed implementations of ML estimation in GGM and their approximations are provided in [49]. Second, we emphasize that we consider the problem of change detection with a known graphical model, i.e., the conditional independence graph topology is known a priori and does not change even when a change occurs.

5.2 Problem Formulation

In this section, we provide a brief introduction to GGM [45], and then formulate the distributed change detection problem in GGM.

Graphical models are intuitive characterizations of conditional independence structures exhibited by variables with a joint probability distribution function (pdf) $f(\mathbf{x}_1, \cdots, \mathbf{x}_p)$. Specifically, define an undirected graph $\mathcal{G} = (V, E)$ with a set of nodes $V = \{1, \cdots, p\}$ connected by undirected edges $E = \{(i_1, j_1), \cdots (i_{|E|}, j_{|E|})\}$, where we use the convention that each node is connected to itself, i.e., $(i, i) \in E$ for all $i \in V$. We define N_i as the set of neighbors of the i'th node, $i = 1, \cdots, p$, i.e., $N_i = \{j | (i, j) \in E, j \neq i\}$.

Let $\mathbf{x}$ be a length p random vector, called the vector of node states, whose elements are

indexed by the nodes in V. The vector $\mathbf{x}$ satisfies the Markov property with respect to $\mathcal{G}$, if for any pair of non-adjacent nodes the corresponding pair of elements in $\mathbf{x}$ are conditionally independent given the remaining elements:

$$\begin{aligned} & f\left(\mathbf{x}_i, \mathbf{x}_j | \mathbf{x}_{V\setminus i,j}\right) \\ =& f\left(\mathbf{x}_i | \mathbf{x}_{V\setminus i,j}\right) f\left(\mathbf{x}_j | \mathbf{x}_{V\setminus i,j}\right) \forall \{i,j\} \notin E. \end{aligned} \tag{5.1}$$

The class of GGMs are graphical models over the multivariate Gaussian distribution. This distribution is appealing due to the fact that it is completely characterized by the mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma} \succ \mathbf{0}$ parameters:

$$\frac{1}{(2\pi)^{\frac{p}{2}} |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \mathbf{e}^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})}, \tag{5.2}$$

In the sequel, it will be more natural to use the canonical parameters, $\mathbf{h} = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ and $\mathbf{J} = \boldsymbol{\Sigma}^{-1} \succ \mathbf{0}$. These lead to the following representation

$$\mathcal{N}(\mathbf{x}; \mathbf{h}, \mathbf{J}) = \frac{\mathbf{e}^{-\frac{1}{2}\mathbf{h}^T \mathbf{J}^{-1} \mathbf{h}}}{(2\pi)^{\frac{p}{2}} |\mathbf{J}|^{-\frac{1}{2}}} \mathbf{e}^{-\frac{1}{2}\mathbf{x}^T \mathbf{J} \mathbf{x} + \mathbf{x}^T \mathbf{h}}, \tag{5.3}$$

in which the exponent exhibits a linear dependency on both $\mathbf{h}$ and $\mathbf{J}$. This formulation is attractive since the graph $\mathcal{G}$ is directly related to the sparsity of $\mathbf{J}$. Indeed, simple algebraic manipulations reveal that specializing the factorization property (5.1) to the jointly Gaussian case yields

$$[\mathbf{J}]_{i,j} = 0 \qquad \forall \{i,j\} \notin E. \tag{5.4}$$

The change detection problem in GGM is formulated as follows. Let $\mathbf{x}$ be a Gaussian random vector with parameters $\mathbf{h} = \mathbf{0}$ and $\mathbf{J} \succ \mathbf{0}$. Given n independent and identically distributed (i.i.d.) realizations of $\mathbf{x}$, denoted by $\{\mathbf{x}[l]\}_{l=1}^n$, and knowledge of the conditional

independence structure through the topology of $\mathcal{G} = (V, E)$, the goal is to test whether or not the inverse covariance is different from some nominal value, denoted by $\mathbf{J}_0$. Mathematically, we would like to test which of the following hypotheses holds

$$\begin{cases} H_0 : \mathbf{J} = \mathbf{J}_0 \\ H_1 : \mathbf{J} \neq \mathbf{J}_0 \end{cases} . \tag{5.5}$$

Specifically, we address this change detection problem in a practical and distributed setting. We assume that the topology of communication links between these nodes matches the topology of $\mathcal{G}$, which means each node can only communicate with its neighbors, i.e., node i can only access $\{\mathbf{x}_{[i\ N_i]}[l]\}_{l=1}^n$ and exchange messages with nodes in N_i. Under these restrictions, we would like all the nodes in the system to reach a consensus on test (5.5).

For later use, we define the Kullback-Leibler (KL) divergence between two k-dimensional multivariate normal distributions parameterized by $\{\mathbf{h}_1, \mathbf{J}_1\}$ and $\{\mathbf{h}_2, \mathbf{J}_2\}$ as

$$\begin{aligned} D_{\mathrm{KL}}\left(\mathbf{h}_1, \mathbf{J}_1; \mathbf{h}_2, \mathbf{J}_2\right) = \frac{1}{2}\big(&\mathrm{Tr}\left\{\mathbf{J}_2\mathbf{J}_1^{-1}\right\} - \log|\mathbf{J}_2\mathbf{J}_1^{-1}| \\ &+(\mathbf{J}_2^{-1}\mathbf{h}_2 - \mathbf{J}_1^{-1}\mathbf{h}_1)^T\mathbf{J}_2(\mathbf{J}_2^{-1}\mathbf{h}_2 - \mathbf{J}_1^{-1}\mathbf{h}_1) - k\big) . \end{aligned} \tag{5.6}$$

5.3 Change Detection Methods

5.3.1 Global GLRT

The classical centralized solution to composite hypothesis testing is the GLRT methodology [54]. Specifically, the GLRT for (5.5), which we call the global GLRT (G-GLRT), is defined as

$$t_G = \log\frac{\max_{\mathbf{J}\in\mathcal{J}}\prod_{l=1}^n \mathcal{N}\left(\mathbf{x}[l]; \mathbf{0}, \mathbf{J}\right)}{\prod_{l=1}^n \mathcal{N}\left(\mathbf{x}[l]; \mathbf{0}, \mathbf{J}_0\right)} \tag{5.7}$$

where

$$\mathcal{J} = \left\{ \mathbf{J} \neq \mathbf{J}_0, \mathbf{J} \succ 0 \text{ and } [\mathbf{J}]_{i,j} = 0 \; \forall \{i,j\} \notin E \right\}. \tag{5.8}$$

Note that the optimal solution to the maximization in the numerator of (5.7) is exactly the global ML estimator of $\mathbf{J}$, which we denote by $\hat{\mathbf{J}}$. This estimate satisfies [45]

$$\mathrm{Tr}\left\{\mathbf{S}\mathbf{J}_0\right\} = \mathrm{Tr}\left\{\hat{\mathbf{J}}^{-1}\mathbf{J}_0\right\} \text{ and } \mathrm{Tr}\left\{\mathbf{S}\hat{\mathbf{J}}\right\} = p, \tag{5.9}$$

where $\mathbf{S}$ is the sample covariance defined as

$$\mathbf{S} \;\; = \;\; \frac{1}{n}\sum_{l=1}^{n} \mathbf{x}[l]\mathbf{x}[l]^T. \tag{5.10}$$

Using (5.9), the G-GLRT can be simplified to

$$t_{\mathrm{G}} = \frac{n}{2}\left(\mathrm{Tr}\left\{\hat{\mathbf{J}}^{-1}\mathbf{J}_0\right\} - \log|\hat{\mathbf{J}}^{-1}\mathbf{J}_0| - p\right) \tag{5.11}$$

$$= nD_{\mathrm{KL}}\left(\mathbf{0}, \hat{\mathbf{J}}; \mathbf{0}, \mathbf{J}_0\right). \tag{5.12}$$

Intuitively, the decision is made based on the KL distance between the estimated $\hat{\mathbf{J}}$ and the nominal $\mathbf{J}_0$. As a classical GLRT, this test enjoys all of its well known attractive properties [54]. Its main drawback is that it requires centralized processing and is difficult to implement in a distributed setting.

5.3.2 Local GLRT

Our first distributed test is based on aggregating multiple local GLRTs in each of the nodes. Specifically, the i'th local network consists of nodes given by $\{i, N_i\}$. The GLRT for testing a change in this small network is defined as

$$t_{\mathrm{L},i} = \log \frac{\max_{\mathbf{J}^i \succ \mathbf{0}} \prod_{l=1}^{n} \mathcal{N}\left(\mathbf{x}_{[i \;\; N_i]}[l]; \mathbf{0}, \mathbf{J}^i\right)}{\prod_{l=1}^{n} \mathcal{N}\left(\mathbf{x}_{[i \;\; N_i]}[l]; \mathbf{0}, \mathbf{J}_0^i\right)}, \tag{5.13}$$

where

$$\mathbf{J}_0^i = \left([\mathbf{\Sigma}_0]_{[i\ N_i],[i\ N_i]}\right)^{-1}. \tag{5.14}$$

The solution to the numerator of (5.13) is the unconstrained ML estimate

$$\hat{\mathbf{J}}^i = \left(\mathbf{S}_{[i\ N_i],[i\ N_i]}\right)^{-1}. \tag{5.15}$$

Plugging it into (5.13) leads to

$$\begin{aligned} t_{\mathrm{L},i} &= \frac{n}{2}\left(\log|\hat{\mathbf{J}}^i| - \log\left|\mathbf{J}_0^i\right| + \mathrm{Tr}\left\{\mathbf{S}_{[i\ N_i],[i\ N_i]}\left(\mathbf{J}_0^i - \hat{\mathbf{J}}^i\right)\right\}\right) \\ &= \frac{n}{2}\left(\mathrm{Tr}\left\{(\hat{\mathbf{J}}^i)^{-1}\mathbf{J}_0^i\right\} - \log\left|(\hat{\mathbf{J}}^i)^{-1}\mathbf{J}_0^i\right| - 1 - |N_i|\right) \\ &= nD_{\mathrm{KL}}\left(\mathbf{0}, \hat{\mathbf{J}}^i; \mathbf{0}, \mathbf{J}_0^i\right), \end{aligned} \tag{5.16}$$

which is independently performed at the i'th node of the network. The overall decision is then obtained by aggregating these results via standard consensus methods [57]

$$t_{\mathrm{L}} = \sum_{i=1}^{p} t_{\mathrm{L},i} = n\sum_{i=1}^{p} D_{\mathrm{KL}}\left(\mathbf{0}, \hat{\mathbf{J}}^i; \mathbf{0}, \mathbf{J}_0^i\right). \tag{5.17}$$

5.3.3 Pseudo GLRT

Recently, it was shown that the local aggregation approach can be formulated using a similar approach known as pseudo-likelihood [49]. In brief, the idea is to approximate the global likelihood function with a product of conditional likelihoods [58, 93]

$$\prod_{l=1}^{n}\mathcal{N}\left(\mathbf{x}[l]; \mathbf{0}, \mathbf{J}\right) \approx \prod_{l=1}^{n}\prod_{i=1}^{p}\mathcal{N}\left(\mathbf{x}_i[l]; h^i[l], \mathbf{J}_{i,i}\right) \tag{5.18}$$

where

$$h^i[l] = -\mathbf{J}_{i,N_i}\mathbf{x}_{N_i}[l] \tag{5.19}$$

and $\mathbf{J}_{i,i}$ are the canonical parameters of the conditional pdf of $\mathbf{x}_i[l]$ given $\mathbf{x}_{V\setminus i}[l]$. Note that the $h^i[l]$ parameter depends on the values of $\mathbf{J}_{i,N_i}$. Maximizing this pseudo-likelihood instead of the likelihood has shown promising complexity, communication and accuracy tradeoffs in the context of estimation. Thus, a natural candidate for distributed hypothesis testing is to replace the local tests with tests using conditional likelihoods. Specifically, the i'th node computes

$$\begin{aligned} t_{\mathrm{P},i} &= \log \frac{\max_{\mathbf{J}_{i,[i\ N_i]}} \prod_{l=1}^{n} \mathcal{N}\left(\mathbf{x}_{[i\ N_i]}[l]; h^i[l], \mathbf{J}_{i,i}\right)}{\prod_{l=1}^{n} \mathcal{N}\left(\mathbf{x}_{[i\ N_i]}[l]; h_0^i[l], [\mathbf{J}_0]_{i,i}\right)} \\ &= \sum_{l=1}^{n} D_{\mathrm{KL}}\left(\hat{h}^i[l], \hat{\mathbf{J}}_{i,i}; h_0^i[l], [\mathbf{J}_0]_{i,i}\right), \end{aligned} \tag{5.20}$$

where

$$\hat{h}^i[l] = -\hat{\mathbf{J}}_{i,N_i}\mathbf{x}_{N_i}[l], \tag{5.21}$$

$$h_0^i[l] = -[\mathbf{J}_0]_{i,N_i}\mathbf{x}_{N_i}[l], \tag{5.22}$$

and $\hat{\mathbf{J}}_{i,[i\ N_i]}$ is the first row of $\hat{\mathbf{J}}^i$ (5.15). As before, the overall P-GLRT is obtained by aggregating these local decisions via consensus methods [57]

$$t_{\mathrm{P}} = \sum_{i=1}^{p} t_{\mathrm{P},i}. \tag{5.23}$$

5.3.4 Choice of estimators

The above tests have been derived as appropriate GLRTs. Each of them uses different optimizations in its numerator, and has different definitions for $\hat{\mathbf{J}}$. That is, G-GLRT assumes that $\hat{\mathbf{J}}$ is the global constrained ML estimate in the GGM. L-GLRT and P-GLRT assume that $\hat{\mathbf{J}}^i$ $(1 \leqslant i \leqslant p)$ are simply the inverses of the local sample covariances. However, the resulting tests are sums of KL distances between $\hat{\mathbf{J}}$ and $\mathbf{J}_0$. In practice, these tests can all be used,

with other estimates, by plugging in any estimator $\hat{\mathbf{J}}$ as long as these can be implemented in a distributed manner. A few promising choices are detailed in [49].

5.4 Numerical Results

In this section, simulation results of two experiments are provided to demonstrate the performance of the proposed tests.

As a promising application, we first consider a practical fault detection problem in smart power grids. We follow the formulation in [53], which shows that the bus susceptance matrix of an electrical power system is directly related to the covariance matrix of a GGM. Therefore, the proposed tests in this chapter can be used to detect any susceptance change that affects the covariance matrix of the network. The graph and parameters used in this simulation are based on the IEEE 30-bus test system [94], whose diagram is given by Fig. 5.1. We consider a case where the transmission line between bus 2 and bus 5 breaks so that the susceptance between these two buses changes from 0.1983 siemens to 0 siemens. The sample number is 20 ($n = 20$), which is smaller than the dimension ($p = 30$). The results are summarized in Fig. 5.2. As these ROC curves show, G-GLRT, P-GLRT and L-GLRT all successfully detect this fault.

Next, we create a GGM by building a network with 100 nodes ($p = 100$), whose locations are uniformly distributed within the unit square. Then every node in this network is connected to the nodes within a certain distance to it. Fig.5.3 illustrates the topology of this GGM. In this case on average every node has 3 neighbors and the maximal degree $k = 5$. Under H_0 the nominal inverse covariance matrix $\mathbf{J}_0$ is obtained using the Matlab function "sprandsym", which randomly generates a positive definite matrix whose sparsity matches

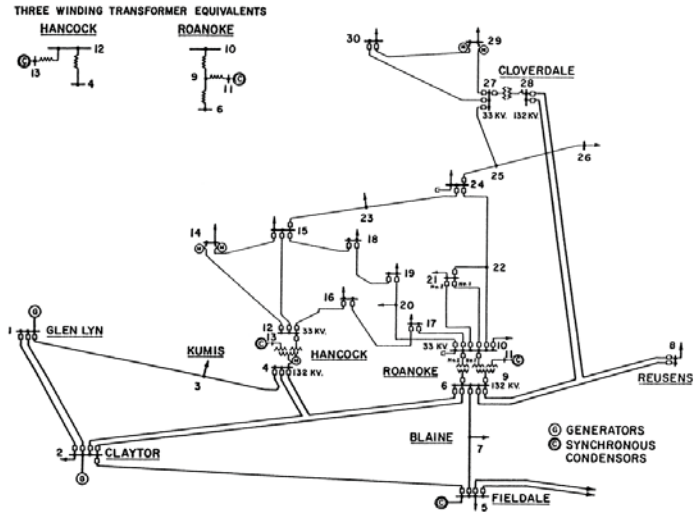


Figure 5.1: IEEE 30-bus test system.

the topology of Fig.5.3. Under H_1 every non-diagonal element (the non-zero ones) of $\mathbf{J}_0$ is increased by 0.1.

We choose different sample numbers ($n = 8, 15$ and 25) and the corresponding ROC curves are provided in Fig. 6.3. Notice that all of the sample numbers we used here are larger than k but much smaller than p. As we can see, the ROC curves of G-GLRT, P-GLRT and L-GLRT in Fig. 6.3 are very close. In this case, the performance of G-GLRT is always better than L-GLRT. The performance of P-GLRT is close to L-GLRT and is slightly worse than G-GLRT when $n = 8$. But as n increases, P-GLRT becomes better (when $n = 15$) and even outperforms G-GLRT (when $n = 25$).

5.5 Summary

This chapter studies the distributed change detection problem in GGMs. Statistical analysis in GGM leads to several advantages, including a smaller number of parameters to model a large scale distribution, less samples required for the detection, faster detection and

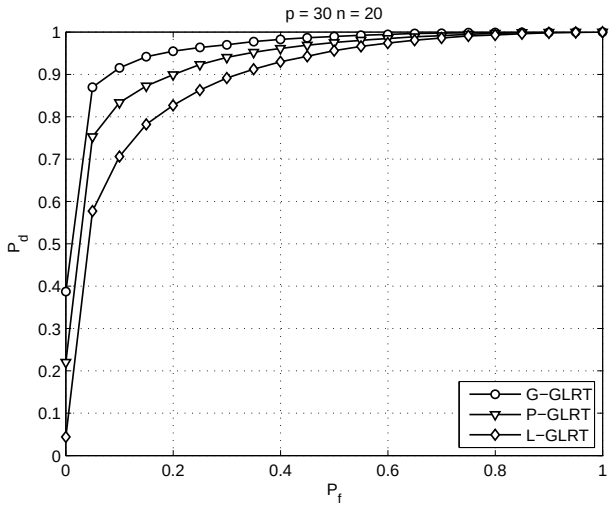


Figure 5.2: Comparison of tests for the Smart Grids example.

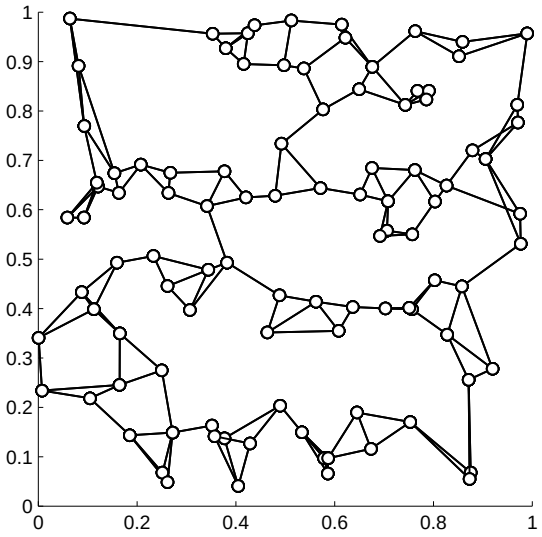


Figure 5.3: Topology of the 100-node GGM.

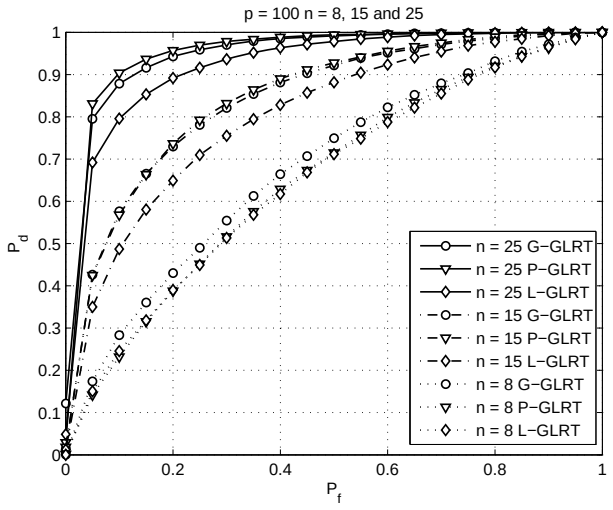


Figure 5.4: Comparison of tests for the 100-node GGM example.

less communication costs. We formulate the hypothesis testing problem for change detection in GGMs and propose a global and centralized solution using the GLRT. We then provide two distributed approximations to this global test based on aggregation of multiple local or conditional tests. We compare the performance of these tests in the context of failure detection in smart grids.

Chapter 6

Change Detection in Smart Grids Using Errors In Variables Models

6.1 Introduction

The deployment of large scale smart grid systems taking place today is accompanied with growing vulnerability and security concerns, which makes it desirable to recognize abrupt system failures and intrusions rapidly. Various work on anomaly detection in smart grids has been proposed recently. Fault detection and localization in power lines under steady state models has been studied in [37, 95, 96]. Detection techniques have been derived based on voltage and current measurements from one bus [95,97], both buses of the faulted transmission line [38], and even measurements from other buses [96]. Recently, identification of the outaged lines using the LASSO method has been considered in [98]. In a different line of work, the problem is expressed as change detection in Gaussian graphical models [53, 99]. These work did not cope with the noise characteristics of practical smart grid systems. Thus,

in this chapter, we will consider change detection in the bus susceptance based on noisy measurements of both the voltages and the power flows.

This chapter focuses on detecting a change in the susceptance parameters of the smart grid system. The direct current (DC) power flow model assumes a simple linear relation between the active power flow on a transmission line and the difference of the voltage angles on the two corresponding buses [100]

$$\text{power flow} \;=\; \text{susceptance} \;\times\; \text{voltage angle difference}. \tag{6.1}$$

In practical systems, noisy measurements of the flows and voltages are obtained via Phasor Measurement Units (PMUs). Given these observations, we seek to detect changes in the susceptance parameters. Specifically, we need to decide whether the susceptances are equal to some nominal known values, or have changed.

From a statistical perspective, the noisy smart grid system can be considered as a special case of linear errors-in-variables (EIV) models [59]. The standard approach to parameter estimation in such problems is known as total least squares (TLS), and generalizes classical least squares by allowing noise in both sides of the linear model. Hypothesis testing in EIV models has also been addressed. Generalized likelihood ratio tests (GLRT) have been derived in the context of process monitoring [36] as well as array processing [44]. Following these works, we derive the TLS-GLRT which is specifically tailored for the smart grid structure and examine its detection performance.

Parameter estimation in EIV models is known to be difficult. TLS and its extensions are often unstable and may be improved. Recently, a competing approach known as total maximum likelihood (TML) was proposed in [60–62] and was shown to provide promising performance gains in various estimation problems. In particular, it was shown that TML can

be interpreted as a regularized version of TLS and is therefore more stable. As a continuation of the work of [60–62], we consider hypothesis testing within the TML framework and derive the TML-GLRT for detecting changes in such models.

The main contributions of this chapter are twofold. First, we express the smart grids fault detection as a hypothesis testing problem employing a linear EIV model. Then we derive the corresponding TLS-GLRT. Second, we propose a novel approach for testing in EIV models based on the TML methodology. Specifically, we consider this approach in the context of the smart grid, but the technique may be applied to other EIV models. Initial numerical results show the promising detection advantages the TML-GLRT has in some cases with no increase in computational complexity. We also provide a distributed version of the TLS-GLRT.

The remainder of this chapter is organized as follows. Section 6.2 describes the rigorous formulation of the hypothesis testing problem. In Section 6.3 we derive the TLS-GLRT and the TML-GLRT. The distributed implementation of TLS-GLRT is provided in Section 6.4.

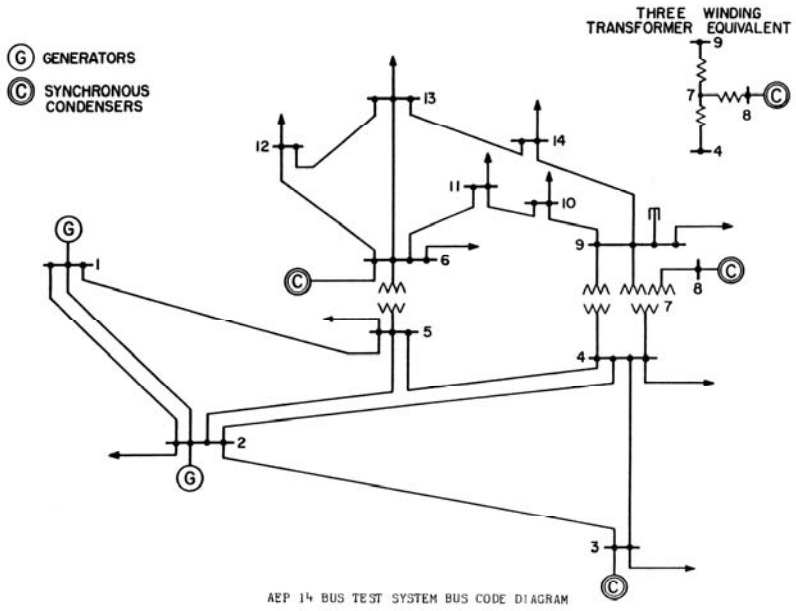


Figure 6.1: IEEE 14-bus test system.

6.2 Problem Formulation

We model a power system with N buses and M transmission lines using an undirected graph $\mathcal{G} = (V, E)$ with the node set $V = \{1, \cdots, N\}$ representing the buses, and the edge set $E = \{(i_1, j_1), \cdots, (i_M, j_M)\}$ representing the transmission lines. We assume that no node is connected to itself. An example consisting of 14 buses and 21 transmission lines is shown in Fig. 6.1. The length N vector $\mathbf{x}[t]$ $(t = 1, \cdots, T)$ models the voltage angles at time slot t. The $N \times N$ skew-symmetric matrix $\mathbf{Y}[t]$ $(t = 1, \cdots, T)$ models the active power flow between the buses at time slot t. The $N \times N$ matrix $\mathbf{B}$ is the susceptance matrix on the transmission lines, i.e., $\mathbf{B}_{i,j} = \mathbf{B}_{j,i}$ is the susceptance between buses i and j if $(i, j) \in E$ and $\mathbf{B}_{i,j}$ is zero otherwise. The DC power flow model states

$$\mathbf{Y}_{i,j}[t] = \mathbf{B}_{i,j}(\mathbf{x}_i[t] - \mathbf{x}_j[t]), t = 1, \cdots, T. \tag{6.2}$$

To simplify the notation, we define $\mathbf{b}$ as a length M vector with the elements $\mathbf{B}_{i,j}$ for $(i,j) \in E$ and $i > j$. Similarly, a length M vector $\mathbf{z}[t]$ collects $\mathbf{Y}_{i,j}[t]$ for $(i,j) \in E$ and $i > j$. Thus, (6.2) can be expressed as

$$\mathbf{z}[t] = \text{diag}\{\mathbf{b}\}\,\mathbf{D}\mathbf{x}[t] \tag{6.3}$$

where the $M \times N$ mapping matrix $\mathbf{D}$ is defined as follows. For the k'th line $(i_k, j_k) \in E$ and $i_k > j_k$, we have $\mathbf{D}_{k,i_k} = 1$ and $\mathbf{D}_{k,j_k} = -1$. The other elements in the k'th row of $\mathbf{D}$ are zero.

Future power grids will allow real time monitoring of the power via noisy measurements of the power flows

$$\tilde{\mathbf{z}}[t] = \mathbf{z}[t] + \mathbf{w}_z[t] \tag{6.4}$$

and noisy measurements of the voltage angles

$$\tilde{\mathbf{x}}[t] = \mathbf{x}[t] + \mathbf{w}_x[t] \tag{6.5}$$

where $\{\mathbf{w}_z[t]\}_{t=1}^T$ and $\{\mathbf{w}_x[t]\}_{t=1}^T$ are independent Gaussian noise processes:

$$\mathbf{w}_z[t] \sim \mathcal{N}(\mathbf{0}, \sigma_z^2\mathbf{I}) \tag{6.6}$$

$$\mathbf{w}_x[t] \sim \mathcal{N}(\mathbf{0}, \sigma_x^2\mathbf{I}). \tag{6.7}$$

Using matrix notations, we construct $M \times T$ matrices $\mathbf{Z}$ and $\mathbf{W}_z$, and $N \times T$ matrices $\mathbf{X}$ and $\mathbf{W}_x$ by stacking the T samples of $\mathbf{z}[t]$, $\mathbf{w}_z[t]$, $\mathbf{x}[t]$ and $\mathbf{w}_x[t]$ $(1 \leqslant t \leqslant T)$ together,

respectively. The matrix power flow EIV model can then be expressed as

$$\begin{cases} \mathbf{Z} = \text{diag}\{\mathbf{b}\}\,\mathbf{D}\mathbf{X} \\ \tilde{\mathbf{Z}} = \mathbf{Z} + \mathbf{W}_z \\ \tilde{\mathbf{X}} = \mathbf{X} + \mathbf{W}_x \end{cases} . \tag{6.8}$$

In this chapter, we consider the problem of detecting changes in the susceptance vector $\mathbf{b}$ based on the noisy observations $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Z}}$. Specifically, we assume the knowledge of a vector $\mathbf{b}_0$ corresponding to the nominal behavior of the grid and test the following hypotheses

$$\begin{cases} H_0: & \mathbf{b} = \mathbf{b}_0 \\ H_1: & \mathbf{b} \neq \mathbf{b}_0 \end{cases} . \tag{6.9}$$

We emphasize that under both hypotheses $\mathbf{Z}$ and $\mathbf{X}$ are unknown and have to be estimated as well.

6.3 Centralized Solutions

In this section, we propose two possible solutions to the hypothesis test in (6.9). Both solutions are based on the GLRT methodology but differ in their statistical model for the unknown parameters $\mathbf{Z}$ and $\mathbf{X}$.

6.3.1 Total least squares

In the classical statistical literature, the EIV model (6.8) assumes $\mathbf{X}$, and therefore also $\mathbf{Z}$, are deterministic unknown vectors. ML estimation of $\mathbf{b}$, $\mathbf{X}$ and $\mathbf{Z}$ in this model is known as TLS [59]. Following this terminology, we now derive the TLS-GLRT. The joint distribution

of the observations is

$$p(\tilde{\mathbf{z}}[t], \tilde{\mathbf{x}}[t]; \mathbf{b}, \mathbf{x}[t]) = p(\tilde{\mathbf{z}}[t]; \mathbf{b}, \mathbf{x}[t]) p(\tilde{\mathbf{x}}[t]; \mathbf{x}[t]). \tag{6.10}$$

where

$$\begin{aligned} p(\tilde{\mathbf{z}}[t]; \mathbf{b}, \mathbf{x}[t]) &= \mathcal{N}(\text{diag}\{\mathbf{b}\}\mathbf{D}\mathbf{x}[t], \sigma_z^2\mathbf{I}) \quad (6.11) \\ p(\tilde{\mathbf{x}}[t]; \mathbf{x}[t]) &= \mathcal{N}(\mathbf{x}[t], \sigma_x^2\mathbf{I}). \quad (6.12) \end{aligned}$$

Thus, the TLS-GLRT is

$$t_{\text{TLS}} = \log \frac{\max_{\mathbf{b},\mathbf{X}} \prod_{t=1}^{T} p(\tilde{\mathbf{z}}[t], \tilde{\mathbf{x}}[t]; \mathbf{b}, \mathbf{x}[t])}{\max_{\mathbf{X}} \prod_{t=1}^{T} p(\tilde{\mathbf{z}}[t], \tilde{\mathbf{x}}[t]; \mathbf{b}_0, \mathbf{x}[t])} \underset{H_1}{\overset{H_0}{\lessgtr}} \rho, \tag{6.13}$$

where ρ is a fixed threshold.

We now simplify the expression of (6.13). The denominator of (6.13) is a simple weighted LS problem

$$\max_{\mathbf{X}} \sum_{t=1}^{T} \log p(\tilde{\mathbf{z}}[t], \tilde{\mathbf{x}}[t]; \mathbf{b}_0, \mathbf{x}[t]) = C - \min_{\mathbf{X}} \frac{1}{2\sigma_z^2} ||\tilde{\mathbf{Z}} - \text{diag}\{\mathbf{b}_0\}\mathbf{D}\mathbf{X}||_F^2 + \frac{1}{2\sigma_x^2} ||\tilde{\mathbf{X}} - \mathbf{X}||_F^2, \tag{6.14}$$

where

$$C = \frac{(M+N)T}{2} \log(2\pi) - MT \log \sigma_z - NT \log \sigma_x. \tag{6.15}$$

The closed-form solution to (6.14) is

$$\hat{\mathbf{X}} = \left(\sigma_z^2\mathbf{I} + \sigma_x^2\mathbf{D}^T \text{diag}\{\mathbf{b}_0\} \text{diag}\{\mathbf{b}_0\} \mathbf{D}\right)^{-1} \left(\sigma_x^2\mathbf{D}^T \text{diag}\{\mathbf{b}_0\} \tilde{\mathbf{Z}} + \sigma_z^2\tilde{\mathbf{X}}\right). \tag{6.16}$$

We define

$$\mathbf{H}(\mathbf{b}) = \sigma_z^2\mathbf{I} + \sigma_x^2 \text{diag}\left\{\mathbf{b}\right\}\mathbf{D}\mathbf{D}^T\text{diag}\left\{\mathbf{b}\right\} \tag{6.17}$$

and

$$\mathbf{A}\left(\mathbf{b}\right) = \tilde{\mathbf{Z}} - \text{diag}\left\{\mathbf{b}\right\}\mathbf{D}\tilde{\mathbf{X}}. \tag{6.18}$$

Plugging (6.16) into (6.14) and applying the following matrix inverse lemma

$$\left(\sigma_z^2\mathbf{I} + \sigma_x^2\mathbf{D}^T\text{diag}\left\{\mathbf{b}_0\right\}\text{diag}\left\{\mathbf{b}_0\right\}\mathbf{D}\right)^{-1} = \frac{1}{\sigma_z^2}\mathbf{I} - \frac{\sigma_x^2}{\sigma_z^2}\mathbf{D}^T\text{diag}\left\{\mathbf{b}_0\right\}\mathbf{H}^{-1}(\mathbf{b}_0)\text{diag}\left\{\mathbf{b}_0\right\}\mathbf{D}, \tag{6.19}$$

we reduce the denominator of (6.13) to

$$C - \frac{1}{2}\text{Tr}\left\{\mathbf{A}^T(\mathbf{b}_0)\mathbf{H}^{-1}(\mathbf{b}_0)\mathbf{A}(\mathbf{b}_0)\right\}. \tag{6.20}$$

The numerator of (6.13) is a structured TLS problem. It can be simplified by expressing it as a double optimization problem and solving the inner problem in closed form. Following the same steps as we did in the denominator, we obtain

$$\begin{aligned}
&\max_{\mathbf{b},\mathbf{X}} \sum_{t=1}^{T} \log p(\tilde{\mathbf{z}}[t], \tilde{\mathbf{x}}[t]; \mathbf{b}, \mathbf{x}[t]) \\
=&C - \min_{\mathbf{b}} \min_{\mathbf{X}} \frac{1}{2\sigma_z^2}||\tilde{\mathbf{Z}} - \text{diag}\left\{\mathbf{b}\right\}\mathbf{D}\mathbf{X}||_F^2 + \frac{1}{2\sigma_x^2}||\tilde{\mathbf{X}} - \mathbf{X}||_F^2 \\
=&C - \frac{1}{2}\min_{\mathbf{b}} \text{Tr}\left\{\mathbf{A}^T(\mathbf{b})\mathbf{H}^{-1}(\mathbf{b})\mathbf{A}(\mathbf{b})\right\}.
\end{aligned} \tag{6.21}$$

Altogether, the TLS-GLRT is defined as

$$t_{\mathrm{TLS}} = \frac{1}{2}\mathrm{Tr}\left\{\mathbf{A}^T(\mathbf{b}_0)\mathbf{H}^{-1}(\mathbf{b}_0)\mathbf{A}(\mathbf{b}_0)\right\} - \frac{1}{2}\min_{\mathbf{b}}\mathrm{Tr}\left\{\mathbf{A}^T(\mathbf{b})\mathbf{H}^{-1}(\mathbf{b})\mathbf{A}(\mathbf{b})\right\} \underset{H_1}{\overset{H_0}{\lessgtr}} \rho. \tag{6.22}$$

6.3.2 Total maximum likelihood

Recently, an alternative statistical model for (6.8) was proposed in [60–62] by assuming that $\mathbf{x}[t]$ are random variables centered around $\tilde{\mathbf{x}}[t]$. ML estimation in this formulation was denoted by TML and was shown to outperform TLS under various settings. In particular, it was shown that TML can be interpreted as a regularized version of TLS which is more stable. Following these works, we now consider TML in the context of detection and derive the TML-GLRT.

In the TML model, the observed $\tilde{\mathbf{x}}[t]$ are deterministic parameters whereas $\mathbf{x}[t]$ are random:

$$p(\mathbf{x}[t];\tilde{\mathbf{x}}[t]) = \mathcal{N}(\tilde{\mathbf{x}}[t],\sigma_x^2\mathbf{I}). \tag{6.23}$$

Conditioned on $\mathbf{x}[t]$, we have

$$p(\tilde{\mathbf{z}}[t]|\mathbf{x}[t];\mathbf{b}) = \mathcal{N}(\mathrm{diag}\left\{\mathbf{b}\right\}\mathbf{D}\mathbf{x}[t],\sigma_z^2\mathbf{I}). \tag{6.24}$$

The marginal distribution of $\tilde{\mathbf{z}}[t]$ is therefore

$$p(\tilde{\mathbf{z}}[t];\mathbf{b},\tilde{\mathbf{x}}[t]) = \mathcal{N}(\mathrm{diag}\left\{\mathbf{b}\right\}\mathbf{D}\tilde{\mathbf{x}}[t],\mathbf{H}\left(\mathbf{b}\right)). \tag{6.25}$$

where $\mathbf{H}(\mathbf{b})$ is defined in (6.17) above. Thus, TML-GLRT is defined as

$$t_{\mathrm{TML}} = \log \frac{\max_{\mathbf{b}} \prod_{t=1}^{T} p(\tilde{\mathbf{z}}[t]; \mathbf{b}, \tilde{\mathbf{x}}[t])}{\prod_{t=1}^{T} p(\tilde{\mathbf{z}}[t]; \mathbf{b}_0, \tilde{\mathbf{x}}[t])} \begin{array}{c} H_0 \\ \lessgtr \\ H_1 \end{array} \rho. \tag{6.26}$$

Plugging in the distributions yields

$$\begin{aligned} t_{\mathrm{TML}} &= \max_{\mathbf{b}} \sum_{t=1}^{T} \log p(\tilde{\mathbf{z}}[t]; \mathbf{b}, \tilde{\mathbf{x}}[t]) - \sum_{t=1}^{T} \log p(\tilde{\mathbf{z}}[t]; \mathbf{b}_0, \tilde{\mathbf{x}}[t]) \\ &= \frac{1}{2}\mathrm{Tr}\left\{\mathbf{A}^T(\mathbf{b}_0)\mathbf{H}^{-1}(\mathbf{b}_0)\mathbf{A}(\mathbf{b}_0)\right\} + \frac{T}{2}\log|\mathbf{H}(\mathbf{b}_0)| - \\ &\min_{\mathbf{b}} \left\{\frac{1}{2}\mathrm{Tr}\left\{\mathbf{A}^T(\mathbf{b})\mathbf{H}^{-1}(\mathbf{b})\mathbf{A}(\mathbf{b})\right\} + \frac{T}{2}\log|\mathbf{H}(\mathbf{b})|\right\} \begin{array}{c} H_0 \\ \lessgtr \\ H_1 \end{array} \rho. \end{aligned} \tag{6.27}$$

Interestingly, test (6.27) can be interpreted as a regularized version of test (6.22) with additional logarithmic penalties. Previous work in the context of estimation suggested that this structure is more stable than the classical TLS approach [60–62]. Simulation results in the end of this chapter will show similar results in the context of detection.

6.3.3 Threshold setting

The proposed tests (6.22) and (6.27) compare their statistics to a threshold ρ. We choose H_0 if the statistic is smaller than ρ, and H_1 otherwise. Given enough samples, the asymptotic performance of both TLS-GLRT and TML-GLRT under H_0 is given by [101]

$$2t_{\mathrm{TLS/TML}} \sim \mathcal{X}^2_M, \tag{6.28}$$

where $\mathcal{X}^2_M$ is a Chi Squared random variable with M degrees of freedom, and M is the dimension of $\mathbf{b}$. Note that this result is independent of the specific value of the unknown $\mathbf{X}$. Thus, a reasonable approach to choosing the threshold for a given false alarm rate α is

$$\rho_\alpha = \frac{1}{2} F^{-1}_{\mathcal{X}^2_M}(\alpha). \tag{6.29}$$

where $F^{-1}_{\mathcal{X}^2_M}(\cdot)$ is the inverse cumulative distribution function of the Chi Squared distribution with M degrees of freedom.

6.3.4 Simulations

In this section, simulation results in a typical power system are provided to demonstrate the performance of the proposed tests. The susceptance matrix used in the simulation is defined by the IEEE 14-bus test system, whose diagram is shown in Fig. 6.1. The other parameters are $N = 14$, $M = 21$, $T = 20$, $\sigma_x^2 = 0.01$ and $\sigma_z^2 = 0.01$. Numerical solutions[1] to the minimizations in (6.22) and (6.27) are obtained using the Maltab function 'fminunc'. The following results are based on 10^4 Monte Carlo simulations. At each realization, the elements of $\mathbf{X}$ are generated independently as standard normal random variables.

In the first experiment, we verify the accuracy of the threshold setting. Under H_0, $\mathbf{b}_0$ uses the susceptance values provided in the IEEE 14-bus test system profile. We choose 50 different false alarm rates equally spaced within $[0, 0.2]$, which we call the theoretical false alarm rates. For each false alarm rate, the corresponding threshold is calculated using (6.29). Then we apply these thresholds to the proposed tests and count the simulated false alarm rates. The solid line in Fig. 6.2 represents the theoretical false alarm rates. The circles and the stars denote the simulated false alarm rates for TLS-GLRT and TML-GLRT, respectively.

[1]Note that the optimizations are non-convex and there is no guarantee that these solutions will be globally optimal.

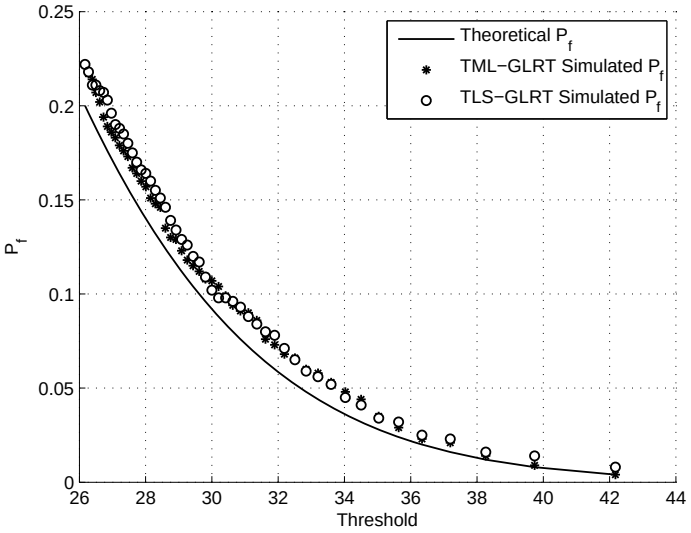


Figure 6.2: Theoretical vs. simulated false alarm rates.

As Fig. 6.2 indicates, the approximation in (6.29) is quite accurate.

In the second experiment, we consider the ROC curves of the proposed tests. Under H_0, we use the susceptance parameters given by the test profile. Under H_1, we apply a change of -2% to every element of $\mathbf{b}$. The other parameters are the same as in the previous simulation. The ROC curves as shown in Fig. 6.3 are generated for both tests by varying the acceptance threshold and tabulating the detection rate under H_1 and the false alarm rate under H_0. We can see that both tests succeed in detecting the change, and that TML-GLRT is significantly better than TLS-GLRT for all false alarm rates. Note that these results do not imply that TML-GLRT is always better than TLS-GLRT. In some other cases, the results show that TLS-GLRT performs better. In future work, we will try to provide a more thorough analysis and comparison of these tests.

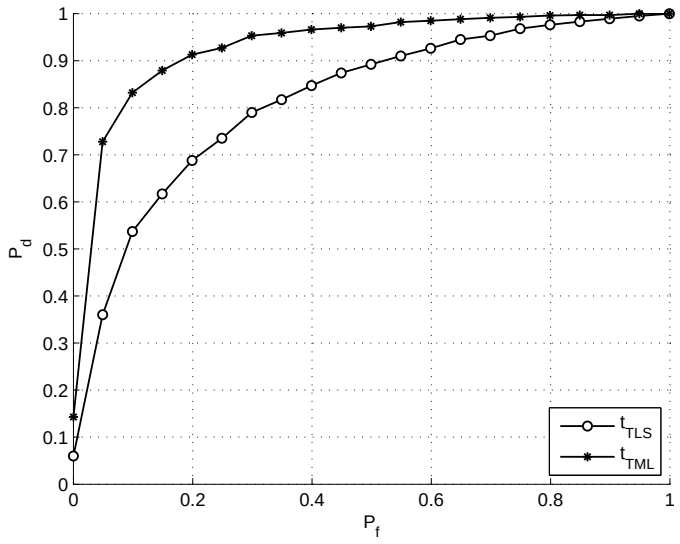


Figure 6.3: The ROC curves of the proposed tests.

6.4 Distributed TLS-GLRT

In this section we provide a distributed implementation of the centralized TLS-GLRT proposed in Section 6.3. The distributed setting is similar to that in Chapter 5. Specifically, we assume there is a communication link between any two nodes that are connected by a power transmission line, i.e., the topology of communication links between these nodes matches the topology of $\mathcal{G}$. If we look at a real power system, these connected nodes will also be in close physical proximity since a physical line or device must be used to connect these two nodes. In this case, it would be easy to run a communication line between these nodes using either power line communications or some other wireless or wired technology. As before, the objective is to compute (6.22) using only local communications and let every node in the system reach the consensus.

To make the distributed computation easier, first we use a diagonal matrix as an ap-

proximation to $\mathbf{D}\mathbf{D}^T$. According to the definition of the mapping matrix $\mathbf{D}$, we can tell that $\mathbf{D}\mathbf{D}^T$ is a sparse matrix. Fig. 6.4 shows the values of the elements of $\mathbf{D}\mathbf{D}^T$ for the IEEE 14-bus, IEEE 30-bus and IEEE 57-bus test systems, respectively. In each subfigure, the x and y axes denote the column and row indices of $\mathbf{D}\mathbf{D}^T$. The colors represent different values of each element of the matrix. Brown represents value 2. Light blue represents value 0. Dark blue and yellow denote values 1 and -1, respectively. Notice that not only are the non-zero off diagonal elements of $\mathbf{D}\mathbf{D}^T$ smaller than the diagonal elements, the matrix $\mathbf{D}\mathbf{D}^T$ also becomes sparser as we consider larger sized networks.

Replacing $\mathbf{D}\mathbf{D}^T$ in (6.17) with a matrix consisting of only its diagonal elements, i.e., $2\mathbf{I}$, we obtain a diagonal approximation to $\mathbf{H}(\mathbf{b})$. Plugging this approximation into (6.22), we have

$$t_{TLS} \approx \sum_{k=1}^{M}\sum_{t=1}^{T} \frac{(\tilde{\mathbf{z}}_k(t) - [\mathbf{b}_0]_k(\tilde{\mathbf{x}}_{j_k}(t) - \tilde{\mathbf{x}}_{i_k}(t)))^2}{\sigma_z^2 + 2[\mathbf{b}_0]_k^2\sigma_x^2} + \sum_{k=1}^{M}\min_{\mathbf{b}_k}\sum_{t=1}^{T} \frac{(\tilde{\mathbf{z}}_k(t) - \mathbf{b}_k(\tilde{\mathbf{x}}_{j_k}(t) - \tilde{\mathbf{x}}_{i_k}(t)))^2}{\sigma_z^2 + 2\mathbf{b}_k^2\sigma_x^2}, \tag{6.30}$$

which can be written as the summation of M components

$$t_{TLS} \approx \sum_{k=1}^{M} t_{TLS,k}, \tag{6.31}$$

where

$$t_{TLS,k} = \sum_{t=1}^{T} \frac{(\tilde{\mathbf{z}}_k(t) - [\mathbf{b}_0]_k(\tilde{\mathbf{x}}_{j_k}(t) - \tilde{\mathbf{x}}_{i_k}(t)))^2}{\sigma_z^2 + 2[\mathbf{b}_0]_k^2\sigma_x^2} + \min_{\mathbf{b}_k}\sum_{t=1}^{T} \frac{(\tilde{\mathbf{z}}_k(t) - \mathbf{b}_k(\tilde{\mathbf{x}}_{j_k}(t) - \tilde{\mathbf{x}}_{i_k}(t)))^2}{\sigma_z^2 + 2\mathbf{b}_k^2\sigma_x^2}. \tag{6.32}$$

We can obtain the second term of $t_{TLS,k}$ by solving a quadratic equation in $\mathbf{b}_k$, which has a closed form solution. It is easy to see that the computation of $t_{TLS,k}$ only involves observations from nodes i_k and j_k, which are connected by transmission line k, thus it can be performed at

either node i_k or node j_k using local communications. In the end, the test (6.30) is obtained by aggregating the local components via a consensus algorithm.

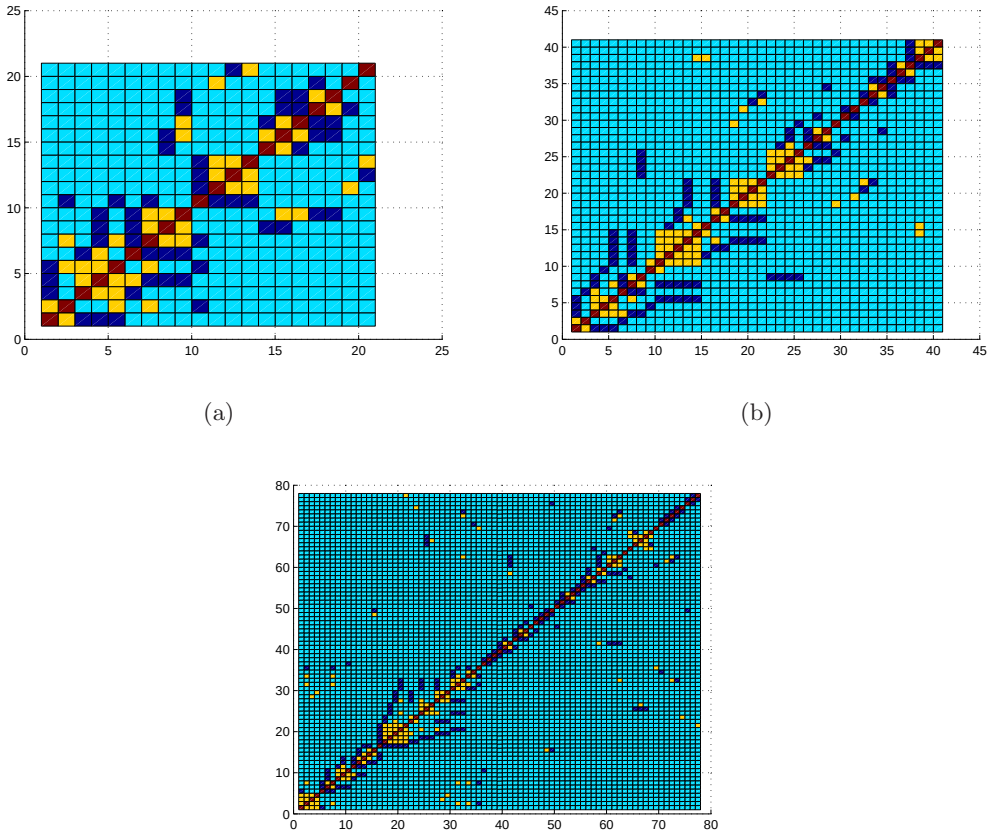


Figure 6.4: $\mathbf{D}\mathbf{D}^T$ for three test systems: (a) IEEE 14-bus (b) IEEE 30-bus and (c) IEEE 57-bus.

Using the relatively small IEEE 14-bus model, the diagonal approximation provided performance indistinguishable from that of the centralized GLR test to the accuracies of our calculations in a number of particular cases we studied. We show some of these results in Fig. 6.5. Under H_0, we use the susceptance parameters given by the test profile. Under

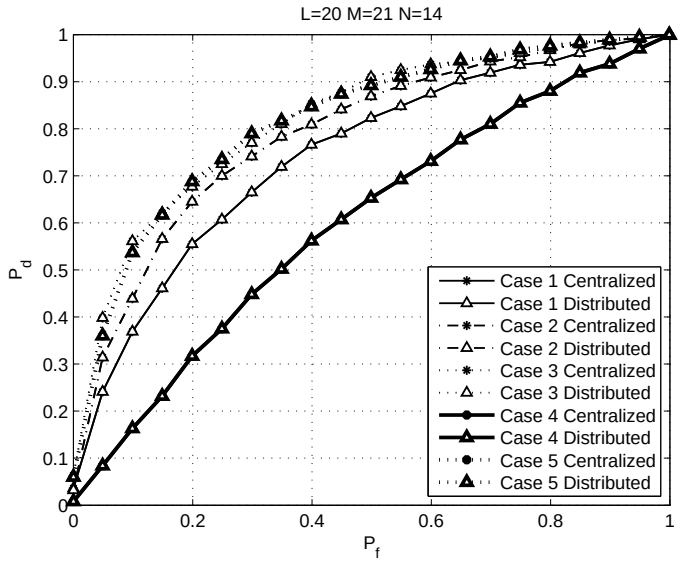


Figure 6.5: Comparison of distributed TLS-GLRT and centralized TLS-GLRT.

H_1, we apply various changes to $\mathbf{b}$. The other parameters are the same as in Section 6.3.4 unless otherwise noted. In case 1, $\sigma_x^2 = \sigma_z^2 = 0.01$; $\mathbf{b}_5$ is decreased by 1%. In case 2, $\sigma_x^2 = 0.0001, \sigma_z^2 = 0.01$; $\mathbf{b}_5$ is decreased by 1%. In case 3, $\sigma_x^2 = 0.00001, \sigma_z^2 = 0.01$; $\mathbf{b}_5$ is decreased by 1%. In case 4, $\sigma_x^2 = \sigma_z^2 = 0.01$; every element of $\mathbf{b}$ is decreased by 1%. In case 5, $\sigma_x^2 = \sigma_z^2 = 0.01$; every element of $\mathbf{b}$ is decreased by 2%. In all cases, the estimates for both the distributed and centralized test are very close to the true values.

6.5 Summary

We investigated the problem of detecting an unknown change in the susceptance parameters of the smart grid power system. The change detection problem was formulated as a GLRT defined in the structured linear EIV model. Inspired by the standard TLS formulation and the recently proposed TML approach, we derived two competing tests, the TLS-GLRT

and the TML-GLRT. Numerical results illustrated the significantly better performance of the TML-GLRT in some cases. We also discussed the distributed implementation of TLS-GLRT.

Chapter 7

Conclusions

This dissertation presents our research on several selected issues concerning multi-sensor data fusion systems. In particular, we focus on the novel signal processing design and performance evaluation techniques for the following three multi-sensor data fusion systems: the multi-sensor image fusion system, the MIMO radar system and the distributed sensor network.

Quantitatively measuring the performance of a multi-sensor image fusion system is a complicated but important task. Chapter 2 generalizes the work in [11] which shows bad behavior for an information-based quality measure when the input images are more distorted. The generalization holds for a class of correlation-based quality measures. We focus on the properties of three popular correlation-based quality measures when used to judge the quality of weighted averaging image fusion algorithms. By employing a model for the input images, closed-form expressions for these quality measure are derived. Further, these quality measures are shown to indicate higher quality in some cases where lower quality is evident. Sufficient conditions for these cases are provided, along with intuitive explanations on why this bad

behavior occurs. Quality calculations with real images, using the procedures recommended by their inventors, also document the predicted bad behavior of the correlation-based image fusion quality measures, which demonstrates the utility of our theoretical analysis.

In Chapter 3, we develop a novel statistic to score the effectiveness of FIQMs for the detection task in light of practical measurements from human perception experiments. The performance of the proposed monotonic test is demonstrated via Monte Carlo simulations. We also show the application of the proposed method to evaluate potential FIQMs in a specific target detection experiment. This chapter proposes the composite DPMLR to quantify how consistent the values of a FIQM are with measured human performance represented by the probability of detection. Specifically, the DPMLR can be used to test whether or not a monotonic relationship exists between the FIQM and the underlying human detection performance that is measured via a perception experiment. The resulting test is designed to be applicable even when the number of observers is small so that the measurement errors from the perceptual experiment are not necessarily Gaussian. The chapter discusses some interesting properties of the DPMLR, and simulation results demonstrate the advantages of the DPMLR over other monotonic statistics. Unlike the monotonic correlation in [20], the DPMLR seamlessly accounts for the spread of the human observations and the number of fused images. It indicates to what degree the ordering of the human observations by the FIQM is not by random chance. The DPMLRT is a general test of monotonicity that can be used to evaluate monotonic relationships beyond the image fusion application. Finally, the DPMLR was used to score a number of potential FIQMs using real image data with a corresponding perception study.

In Chapter 4, we investigate the joint location and velocity estimation of multiple targets in a non-coherent MIMO radar system. The Cramer-Rao bound for a two-target case

is derived and evaluated. This bound gives us theoretically achievable joint estimation performance of a multi-target non-coherent MIMO radar with a sufficient number of antennas. Numerical simulations show that the spatial advantage proposed by previous research can also be observed in a two-target example.

The rest of this dissertation considers the design of change detection methods using distributed sensor networks, where each node is only allowed to communicate with its neighbors. Using our algorithms, all the nodes in the system will ultimately reach a consensus on the test. For this purpose, we have relied on probabilistic graphical models and EIV models. In Chapter 5, we study the distributed change detection problem for distributions that can be represented as GGMs. The time-tested approach to change detection is via composite hypothesis testing, and in particular the GLRT [54]. This method has been successfully applied to change detection in the Gaussian multivariate distribution, and is known as the Bartlett's test [55, 56]. Classical centralized results on testing in GGM are available in [45]. In contrast, our focus is on distributed solutions for change detection. We begin by deriving the global centralized GLRT, and then propose two distributed approximations based on aggregating multiple tests performed at each of the nodes in the graph. The aggregation can then be efficiently implemented using consensus methods [57]. The first distributed test is a natural approach which simply applies the Bartlett's test to smaller size local clusters in the graph. The second method employs the pseudo-likelihood [58] as a surrogate function for the global likelihood. The advantages of our proposed tests are demonstrated using numerical experiments, including one in the context of failure detection in smart power grids [53].

Chapter 6 considers the fault detection problem for measurements following the EIV model. From a statistical perspective, the noisy measurements of voltages and currents in a smart grid system follow a special case of linear EIV models [59]. The standard approach

to parameter estimation in such problems is known as TLS, and generalizes classical least squares by allowing noise in both sides of the linear model. Hypothesis testing in EIV models has also been addressed. The GLRTs have been derived in the context of process monitoring [36] as well as array processing [44]. Following these works, we derive the TLS-GLRT which is specifically tailored for the smart grid structure and examine its detection performance. Parameter estimation in EIV models is known to be difficult. TLS and its extensions are often unstable and may be improved. Recently, a competing approach known as TML was proposed in [60–62] and was shown to provide promising performance gains in various estimation problems. In particular, it was shown that TML can be interpreted as a regularized version of TLS with improved properties. As a continuation of the work of [60–62], we consider hypothesis testing within the TML framework and derive the TML-GLRT for detecting changes in such models. Numerical results show the promising detection advantages the TML-GLRT has in some cases with no increase in computational complexity. We also discussed the distributed implementation of TLS-GLRT.

Bibliography

[1] D. L. Hall and J. Llinas, *Handbook of multisensor data fusion.* CRC Press, 2001.

[2] R. S. Blum and Z. Liu, *Multi-sensor Image Fusion and Its Applications.* CRC Press, 2006.

[3] Z. Zhang and R. S. Blum, "A region-based image fusion scheme for concealed weapon detection," in *In Proc. of the 31st Annual Conf. on Inf. Sci. and Syst.*, 1997, pp. 168–173.

[4] G. Simone, A. Farina, F. C. Morabito, S. B. Serpico, and L. Bruzzone, "Image fusion techniques for remote sensing applications." *Inf. Fusion*, vol. 3, pp. 3–15, 2002.

[5] J. A. Castellanos, J. Neira, and J. D. Tardos, "Multisensor fusion for simultaneous localization and map building," *IEEE Trans. Robot. Autom.*, vol. 17, no. 6, pp. 908–914, 2001.

[6] S. P. Constantinos, M. S. Pattichis, and E. Mitheli-Tzanakou, "Medical imaging fusion applications: An overview," in *Proc. of the 35th Asilomar Conf.on Signals, Syst. and Computers*, vol. 2, 2001, pp. 1263–1267.

[7] R. R. Murphy, "Sensor and information fusion improved vision-based vehicleguidance," *IEEE Intell. Syst.*, vol. 13, no. 6, pp. 49–56, 1998.

[8] J. Yang and R. S. Blum, "A statistical signal processing approach to image fusion for concealed weapon detection," in *IEEE International Conference on Image Processing*, Rochester, NY, 2002, pp. 513–516.

[9] Z. Xue and R. S. Blum, "Concealed weapon detection using color image fusion," in *The Sixth International Conference on Image Fusion*, 2003.

[10] R. S. Blum, "On multisensor image fusion performance limits from an estimation theory perspective," *Inf. Fusion*, vol. 7, no. 3, pp. 250–263, 2006.

[11] Y. Chen, Z. Xue, and R. S. Blum, "Theoretical analysis of an information-based quality measure for image fusion," *Inf. Fusion*, vol. 9, no. 2, pp. 161–175, 2008.

[12] N. Cvejic, A. Loza, D. Bull, and N. Canagarajah, "A similarity metric for assessment of image fusion algorithms," *Int. Journal of Signal Process.*, vol. 2, no. 3, pp. 178–182, 2005.

[13] V. Petrovic and C. Xydeas, "Objective image fusion performance measure," *Electronics Lett.*, vol. 36, no. 4, pp. 308–309, 2000.

[14] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electronics Lett.*, vol. 38, no. 7, pp. 313–315, Mar. 2002.

[15] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. Int. Conf. on Image Process.*, 2003, pp. III–173–176.

[16] A. Toet, N. Schoumans, and J. K. Uspeert, "Perceptual evaluation of different nighttime imaging modalities," in *Proc. of the 3rd Int. Conf. on Inf. Fusion*, vol. 1, 2000, pp. TUD3/17–TUD3/23.

[17] C. Wei and R. S. Blum, "Theoretical analysis of correlation-based quality measures for weighted averaging image fusion," *Inf. Fusion*, to be published.

[18] J. Puzicha, J. M. Buhmann, Y. Rubner, and C. Tomasi, "Empirical evaluation of dissimilarity measures for color and texture," in *Proc. of the 7th IEEE Intl. Conf. on Computer Vision*, vol. 2, 1999, pp. 1165–1173.

[19] Y. Chen and R. S. Blum, "A new automated quality assessment algorithm for image fusion," *Image and vision computing*, vol. 27, no. 10, pp. 1421–1432, Sep. 2009.

[20] L. M. Kaplan, R. S. Blum, and S. D. Burks, "Analysis of image quality for image fusion via monotonic correlation," *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 2, pp. 222–235, Apr. 2009.

[21] F. Sadjadi, "Comparative image fusion analysis," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 3, San Diego, CA, Jun. 2005.

[22] D. Schmieder and M. Weathersby, "Detection performance in clutter with variable resolution," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-19, no. 4, pp. 622–630, 1983.

[23] L. M. Kaplan, "Extended fractal analysis for texture classification and segmentation," *IEEE Trans. Image Process.*, vol. 8, no. 11, pp. 1572–1585, 1999.

[24] M. J. T. Smith and A. Docef, *A study guide for Digital Image Processing.* GA Scientific Publishers, 1997.

[25] C. Howell, R. Moore, S. Burks, and C. Halford, "An evaluation of fusion algorithms using image fusion metrics and human identification performance," in *Infrared Imaging Syst.: Design, Analysis, Modeling, and Testing XVIII. Edited by Holst, Gerald C. Proc. of the SPIE*, vol. 6543, 2007, p. 65430V.

[26] R. K. Sharma, T. K. Leen, and M. Pavel, "Probabilistic image sensor fusion," in *Proc. of the 1998 conf. on Advances in neural inf. Process. syst. II.* Cambridge, MA, USA: MIT Press, 1999, pp. 824–830.

[27] A. M. Haimovich, R. S. Blum, and L. Cimini, "Mimo radar with widely separated antennas," *IEEE Signal Processing Magazine*, 2008.

[28] J. Li and P. Stoica, "Mimo radar with colocated antennas," *IEEE Signal Proc. Magazine*, pp. 106–114, Sep 2007.

[29] D. R. Fuhrmann and G. S. Antonio, "Transmit beamforming for mimo radar systmes using partial signal correlations," in *Proc. of 38th Asilomar Conf. on Signals, Systems and Computers*, Nov 2004, pp. 295–299.

[30] N. Lehman, A. M. Haimovich, R. S. Blum, and L. Cimini, "High resolution capabilities of mimo radar," in *Proc. 40th Asilomar Conf. Singal, System and Computers*, Nov 2006, pp. 25–30.

[31] H. Godrich, A. M. Haimovich, and R. S. Blum, *Concepts and applications of a MIMO radar sysem with widely separated antennas.* Book Chapter, in preparation, 2007.

[32] E. Fishler, A. M. Haimovich, R. S. Blum, D. Chizhik, L. Cimini, and R. Valenzuela, "Mimo radar: an idea whose time has come," in *Proc. IEEE Radar Conf.*, Apr 2004, pp. 71–78.

[33] N. Lehman, E. Fishler, A. M. Haimovich, R. S. Blum, L. Cimini, and R. Valenzuela, "Evaluation of transmit diversity in mimo-radar direction finding," *IEEE Trans. on Signal Processing*, vol. 55, no. 5, pp. 2215–2225, May 2007.

[34] H. Godrich, A. Haimovich, and R. Blum, "Target localization accuracy and multiple target localization: Tradeoff in mimo radars," in *Proc. Asilomar Conf. Signals, Systems, and Computers*, Oct 2008.

[35] Q. He, R. S. Blum, and A. M. Haimovich, "Non-coherent mimo radar for target estimation: More antennas means better performance," *IEEE Trans. on Signal Processing*, 2010.

[36] B. Huang, "Detection of abrupt changes of total least squares models and application in fault detection," *IEEE Transactions on Control Systems Technology*, vol. 9, no. 2, pp. 357 – 367, 2001.

[37] S. Lotfifard, M. Kezunovic, and M. J. Mousavi, "Voltage sag data utilization for distribution fault location," *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 1239 –1246, 2011.

[38] J.-A. Jiang, J.-Z. Yang, Y.-H. Lin, C.-W. Liu, and J.-C. Ma, "An adaptive pmu based fault detection/location technique for transmission lines. i. theory and algorithms," *IEEE Transactions on Power Delivery*, vol. 15, no. 2, pp. 486–493, 2000.

[39] J. Zhu and A. Abur, "Identification of network parameter errors," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 586 – 592, 2006.

[40] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "On malicious data attacks on power system state estimation," in *Universities Power Engineering Conference (UPEC), 2010 45th International*, 2010, pp. 1–6.

[41] ——, "Limiting false data attacks on power system state estimation," in *Information Sciences and Systems (CISS), 2010 44th Annual Conference on*, 2010, pp. 1–6.

[42] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes: Theory and Application*. previously published by Prentice-Hall, Inc., 1993.

[43] J. J. Fuchs, "Matched detector and estimator with signature uncertainty," in *the Forty-First Asilomar Conference on Signals, Systems and Computers*, 2007, pp. 2177–2181.

[44] ——, "A robust matched detector," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5133–5142, 2007.

[45] S. L. Lauritzen, *Graphical models*. New York: Oxford Statistical Science Series, 1996, vol. 17.

[46] A. P. Dempster, "Covariance selection," *Biometrics*, vol. 28, pp. 157–175, 1972.

[47] Y. Weiss and W. Freeman, "Correctness of belief propagation in Gaussian graphical models of arbitrary topology," *Neural Computation*, vol. 13, no. 10, pp. 2173–2200, 2001.

[48] A. Wiesel, Y. C. Eldar, and A. O. Hero, "Covariance estimation in decomposable Gaussian graphical models," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1482 –1492, march 2010.

[49] A. Wiesel and A. O. Hero, "Distributed covariance estimation in gaussian graphical models," *Signal Processing, IEEE Transactions on*, vol. 60, no. 1, pp. 211 –220, jan. 2012.

[50] ——, "Decomposable principal component analysis," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4369 –4377, nov. 2009.

[51] M. Cetin, L. Chen, J. W. Fisher, A. T. Ihler, R. L. Moses, M. J. Wainwright, and A. S. Willsky, "Distributed fusion in sensor networks: A graphical models perspective," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 42– 55, July 2006.

[52] C. Guestrin, P. Bodik, R. Thibaux, M. Paskin, and S. Madden, "Distributed regression: an efficient framework for modeling sensor network data," in *Proceedings of the 3rd international symposium on Information processing in sensor networks*. ACM, 2004, pp. 1–10.

[53] M. He and J. Zhang, "A dependency graph approach for fault detection and localization towards secure smart grid," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 342–351, 2010.

[54] P. J. Bickel and K. A. Doksum, *Mathematical Statistics: Basic Ideas and Selected Topics*. Holden-Day, CA, 1977.

[55] M. S. Bartlett, "Properties of sufficiency and statistical tests," in *Proceedings of the Royal Society of London*, vol. 160, no. 901, May 1937, pp. 268–282.

[56] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis 3rd Edition*. Wiley-Interscience, 2003.

[57] L. Xiao, S. Boyd, and S.-J. Kim, "Distributed average consensus with least-mean-square deviation," *Journal of Parallel and Distributed Computing*, vol. 67, no. 1, pp. 33–46, 2007.

[58] J. Besag, "Statistical analysis of non-lattice data," *Journal of the Royal Statistical Society. Series D (The Statistician)*, vol. 24, no. 3, pp. 179–195, 1975.

[59] S. V. Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*. Frontier in Applied Mathematics, SIAM, 1991.

[60] A. Wiesel, Y. C. Eldar, and A. Beck, "Maximum likelihood estimation in linear models with a gaussian model matrix," *IEEE Signal Processing Letters*, vol. 13, no. 5, pp. 292–295, 2006.

[61] A. Wiesel, Y. C. Eldar, and A. Yeredor, "Linear regression with gaussian model uncertainty: Algorithms and bounds," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2194–2205, June 2008.

[62] A. Beck and Y. C. Eldar, "Structured total maximum likelihood: An alternative to structured total least squares," *SIAM J. on Matrix Analysis and Applications*, vol. 31, no. 5, pp. 2623–2649, 2010.

[63] Z. Xue, R. S. Blum, and Y. Li, "Fusion of visual and ir images for concealed weapon detection," in *The fifth International Conference on Image Fusion*, Annapolis, Maryland, 2002, pp. 1198–1205.

[64] V. Petrovic, "Multilevel image fusion," *B.V. Dasarathy (Ed.), Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications, SPIE-5099, The International Society for Optical Engineering*, pp. 87–96, 2003.

[65] V. Petrovic and C. S. Xydeas, "Optimizing multiresolution pixel-level image fusion," *B.V. Dasarathy (Ed.), Sensor Fusion: Architectures, Algorithms, and Applications, SPIE-4385, The International Society for Optical Engineering*, pp. 96–107, 2001.

[66] ——, "Sensor noise effects on signal-level image fusion performance," *Information Fusion*, vol. 4, no. 3, pp. 167–183, 2003.

[67] A. Toet and M. A. Hogervorst, "Performance comparison of different gray-level image fusion schemes through a universal image quality index," *Ivan Kadar (Ed.), Signal Processing, Sensor Fusion, and Target Recognition XII, SPIE-5096, The International Society for Optical Engineering*, pp. 552–561, 2003.

[68] L. M. Kaplan, R. S. Blum, and S. D. Burks, "Analysis of image quality for image fusion via monotonic correlation," *accepted by IEEE Journal of Selected Topics in Signal Processing special issue on Visual Media Quality Assessment (April 2009) and 26th Army Science Conference*, 2008.

[69] T. M. Cover and J. A. Thomas, *Elements of information theory.* New York, NY, USA: Wiley-Interscience, 1991.

[70] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Process. Lett.*, vol. 9, no. 3, pp. 81–84, 2002.

[71] Y. Chen and R. S. Blum, "Experimental tests of image fusion for night vision," in *The Eighth International Conference on Information Fusion*, 2005.

[72] Z. Wang, A. C. Bovik, H. R. Sheikh, S. Member, and E. P. Simoncelli, "Image quality assessment: from error measurement to structural similarity," *IEEE Trans. Image Process.*, vol. 13, pp. 600–612, 2004.

[73] V. Laparra, J. Munoz-Mari, and J. Malo, "Divisive normalization image quality metric revisited," *Journal of the Optical Society of America A*, vol. 27, no. 4, pp. 852–864, 2010.

[74] W. Mendenhall, R. L. Scheaffer, and D. D. Wackerly, *Mathematical Statistics with Applications*, 3rd ed. Boston: Duxbury Press, 1986.

[75] M. G. Kendall and A. Stuart, *The Advanced Theory of Statistics*. New York: Halner Press, 1961.

[76] VQEG, "Final report from the video quality experts group on validation of objective models of video quality assessment," in *[online] Available: http://www.vqeg.org/*, Mar. 2000.

[77] R. Barlow, D. Barholomew, J. Bremner, and H. Brunk, *Statistical Infernece under Order Restrictions*. New York: Wiley, 1972.

[78] M. Best and N. Chakravarti, "Active set algorithms for isotonic regression: A unifying approach," *Mathematical Programming*, vol. 47, pp. 425–439, May 1990.

[79] T. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*. Upper Saddle River, NJ: Prentice Hall, Inc., 2000.

[80] E. Schumann, "Generating correlated uniform variates," in *http://comisef.wikidot.com/tutorial:correlateduniformvariates*, 2009.

[81] K. L. Kowalskia, "Generalized binomial distributions," *Journal of Mathematical Physics*, vol. 41, no. 4, pp. 2375–2382, 2000.

[82] A. Toet, "Image fusion by a ratio of low-pass pyramid," *Pattern Recognition Lett.*, vol. 9, no. 4, pp. 245–253, 1989.

[83] T. Huntsberger and B. Jawerth, "Wavelet based sensor fusion," in *Proc. SPIE*, vol. 2059, 1993, pp. 488–498.

[84] C. Lejeune, "Wavelet transform for infrared application," in *Proc. SPIE*, vol. 2552, 1995, pp. 313–324.

[85] L. Jiang, F. Tian, L. E. Shen, S. Wu, S. Yao, Z. Lu, and L. Xu, "Perceptual-based fusion of ir and visual images for human detection," in *Proc. of Int. Symposium on Intelligent Multimedia, Video and Speech Process.*, 2004.

[86] H. Chen and P. K. Varshney, "A human perception inspired quality metric for image fusion based on regional information," *Image Fusion*, vol. 8, no. 2, pp. 193–207, Apr. 2007.

[87] O. O. Fadiran, P. Molnar, and L. M. Kaplan, "A statistical approach to quantifying clutter in hyperspectral infrared images," in *Proc. of the IEEE Aerospace Conf.*, Big Sky, MT, Mar. 2006.

[88] D.-O. Kim and R.-H. Park, "New image quality metric using the Harris response," *IEEE Signal Process. Lett.*, vol. 16, no. 7, pp. 616–619, Jul. 2009.

[89] J. C. Leachtenauer, W. Malila, J. Irvine, L. Colburn, and N. Salvaggio, "General image-quality equation: GIQE," *Applied Optics*, vol. 36, no. 32, pp. 8322–8328, 1997.

[90] J. P. Estrera, "Localized signal-to-noise ratio of man and vehicle size targets," in *Infrared Technol. and Appl. XXXV, Proc. of the SPIE*, vol. 7298, Orlando, FL, Apr. 2009, pp. 72 983M–72 983M–14.

[91] J. Schanda, Ed., *Colorimetry: Understanding the CIE System.* Hoboken, NJ: John Wiley & Sons, Inc., 2007.

[92] E. Fishler, A. M. Haimovich, R. S. Blum, L. Cimini, D. Chizhik, and R. Valenzuela, "Spatial diversity in radars - models and detection performance," *IEEE Trans. on Signal Processing*, vol. 54, no. 3, pp. 823–838, 2006.

[93] P. Liang and M. I. Jordan, "An asymptotic analysis of generative, discriminative, and pseudolikelihood estimators," in *Proceedings of the 25th international conference on Machine learning.* ACM, 2008, pp. 584–591.

[94] U. of Washington, "Power systems test case archive," at webpage http://www.ee.washington.edu/research/pstca/.

[95] N. C. Woolley, M. Avenda-Mora, J. V. Milanovic, and A. P. Woolley, "Probabilistic fault location using erroneous measurement devices," in *IEEE International Conference on Smart Measurements for Future Grids*, 2011, pp. 101 – 106.

[96] Y. Liao, "Fault location for single-circuit line based on bus-impedance matrix utilizing voltage measurements," *IEEE Transactions on Power Delivery*, vol. 23, no. 2, pp. 609–617, 2008.

[97] M. Djuric, Z. Radojevic, and V. Terzija, "Distance protection and fault location utilizing only phase current phasors," *IEEE Transactions on Power Delivery*, vol. 13, no. 4, pp. 1020 – 1026, 1998.

[98] H. Zhu and G. B. Giannakis, "Lassoing line outages in the smart power grid," in *IEEE International Conference on Smart Grid Communications*, 2011, pp. 570–575.

[99] C. Wei, A. Wiesel, and R. Blum, "Distributed change detection in gaussian graphical models," in *46th Annual Conference on Information Sciences and Systems*, 2012.

[100] B. W. R. Christie and I. Wangensteen, "Transmission management in the deregulated environment," in *Proceedings of the IEEE*, vol. 88, Feb. 2000, pp. 170–195.

[101] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory.* Prentice Hall, 1998.

Vita

Chuanming Wei was born in 1984 in Shandong Province, China. He received the B.Eng. and M.Eng. degrees in Electrical Engineering from University of Science and Technology of China (USTC), Hefei, China, in 2003 and 2006, respectively. Since 2007, he has been working towards the Ph.D. degree in Electrical Engineering at Lehigh University, Bethlehem, PA. From 2003 to 2007, he worked as an intern engineer at UT Starcom, Inc. During summer 2010, he was an intern at Lister Hill National Center for Biomedical Communications, NIH, Bethesda MD. His research interests include signal and image processing, multisensor data fusion, pattern recognition, radar and sensor networking.

List of Publications

Journal Papers

1. Chuanming Wei, Lance M. Kaplan, Stephen D. Burks and Rick S. Blum, "Diffuse Prior Monotonic Likelihood Ratio Test for Evaluation of Fused Image Quality Measures," IEEE Transactions on Image Processing, vol. 20, no. 2, pp. 327 C 344, Feb. 2011.

2. Chuanming Wei and Rick S. Blum, "Theoretical analysis of correlation-based quality measures for weighted averaging image fusion," Information Fusion, vol. 11, no. 4, pp. 301 C 310, Oct. 2010.

3. Jinzhong Yang, Chuanming Wei, Lifei Zhang, Yongbin Zhang, Rick S. Blum and Lei Dong, "A statistical modeling approach for evaluating auto-segmentation methods for image-guided radiotherapy," Comp. Med. Imag. and Graph. 36(6): 492-500, 2012.

4. Chuanming Wei, Ling Qiu, and Jinkang Zhu, "Margin Adaptive Optimization in Multi-User MISO-OFDM Systems under Rate Constraint," Journal of Communications and Networks (JCN), vol. 9, no. 2, pp. 112 C 117, Jun. 2007.

5. Chuanming Wei, Ling Qiu, and Jinkang Zhu, "A Power Control Method Based on Post-Detection SINR Balancing in Cellular V-BLAST Systems," Journal of Applied Sciences, vol. 25, no. 1, 2007.

Conference Papers

1. Chuanming Wei, Ami Wiesel and Rick S. Blum, "Distributed Change Detection in Gaussian Graphical Models," the Conference on Information Sciences and Systems (CISS), Princeton, NJ, March 2012.

2. Chuanming Wei, Ami Wiesel and Rick S. Blum, "Change Detection in Smart Grids Using Errors In Variables Models," SAM 2012 Conference.

3. Qian He, Chuanming Wei and Rick S. Blum, "Effects of Unsuccessful Transmissions of Measurements on Uncertainty in Power Flow Calculations," SAM 2012 Conference.

4. Jiangfan Zhang, Chuanming Wei and Rick S. Blum, "Ordering for Shift-in-Mean of Gaussian Markov Random Fields with Dependent Observations," SAM 2012 Conference.

5. Chuanming Wei, Qian He and Rick S. Blum, "Cramer-Rao bound for joint location and velocity estimation in multi-target non-coherent MIMO radars," in proceeding of the Conference on Information Sciences and Systems (CISS), Princeton, NJ, March 2010.

6. Chuanming Wei and Rick S. Blum, "Theoretical analysis of correlation-based quality measures for weighted averaging image fusion," in proceeding of the Conference on Information Sciences and Systems (CISS), Baltimore, MD, March 2009.

7. Chuanming Wei, Lance M. Kaplan, Stephen D. Burks and Rick S. Blum, "Diffuse prior monotonic likelihood ratio test for evaluation of fused image quality metrics," in proceeding of the 12th International Conference on Information Fusion, Seattle, WA, Jul. 6-9, 2009, pp. 1076-1083.

8. Chuanming Wei, Ling Qiu, and Jinkang Zhu, "User Selection and Resource Allocation for Multi-User MIMO-OFDM Systems with Downlink Beamforming," Proceeding of the First International Conference on Communications and Networking in China, ChinaCom 2006, Beijing, China, October 2006.

9. Chuanming Wei, Ling Qiu, and Jinkang Zhu, "A Power Control Method Based on Post-Detection SINR Balancing in Cellular V-BLAST Systems," Proceeding of IEEE 9th International Symposium on Spread Spectrum Techniques and Applications, ISSSTA 2006, Manaus, Brazil, August 2006.