AFRL-RY-WP-TR-2021-0167

SCENE MOTION DETECTION IN IMAGERY WITH ANISOPLANATIC OPTICAL TURBULENCE (Preprint)

Richard Lowell Van Hook
Electro-Optic Target Detection & Surveillance Branch
Multispectral Sensing & Detection Division

JULY 2021
Final Report

STINFO COPY

AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7320
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
July 2021	Dissertation	16 June 2021 –16 June 2021

4. TITLE AND SUBTITLE
SCENE MOTION DETECTION IN IMAGERY WITH ANISOPLANATIC OPTICAL TURBULENCE (Preprint)

5a. CONTRACT NUMBER
In-House

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
N/A

6. AUTHOR(S)
Richard Lowell Van Hook

5d. PROJECT NUMBER
N/A

5e. TASK NUMBER
N/A

5f. WORK UNIT NUMBER
N/A

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Sensors Directorate (AFRL/RYMT)
Wright-Patterson Air Force Base, OH 45433-7320
Air Force Materiel Command
United States Air Force

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Sensors Directorate
Wright-Patterson Air Force Base, OH 45433-7320
Air Force Materiel Command
United States Air Force

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
AFRL/RYMT

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RY-WP-TR-2021-0167

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
PAO case number AFRL-2021-1869, Clearance Date 16 June 2021. Submitted to the School of Engineering of the University of Dayton in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Electrical and Computer Engineering. This is a work of the U.S. Government and is not subject to copyright protection in the United States. Report contains color.

14. ABSTRACT
In long range imaging applications, anisoplanatic atmospheric optical turbulence imparts spatially-and temporally-varying blur and geometric distortions in acquired imagery. The ability to distinguish true scene motion from turbulence warping is important for many image processing and analysis tasks. I present two novel scene-motion detection algorithms specifically designed to operate in the presence of anisoplanatic optical turbulence.

15. SUBJECT TERMS
anisoplanatic imaging, image processing, motion detection, turbulence, turbulence mitigation, turbulence modeling

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
a. REPORT	b. ABSTRACT	c. THIS PAGE			Richard Lowell Van Hook
Unclassified	Unclassified	Unclassified	SAR	92	19b. TELEPHONE NUMBER *(Include Area Code)* N/A

SCENE MOTION DETECTION
IN IMAGERY WITH
ANISOPLANATIC OPTICAL TURBULENCE

Dissertation

Submitted to

The School of Engineering of the

UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for

The Degree of

Doctor of Philosophy in Electrical and Computer Engineering

By

Richard Lowell Van Hook

UNIVERSITY OF DAYTON

Dayton, Ohio

August, 2021

SCENE MOTION DETECTION IN IMAGERY WITH ANISOPLANATIC OPTICAL TURBULENCE

Name: Van Hook, Richard Lowell

APPROVED BY:

Russell C. Hardie, Ph.D. Advisor Committee Chairman Professor Electrical & Computer Engineering	Vijayan K. Asari, Ph.D. Committee Member Professor Electrical & Computer Engineering
John S. Loomis, Ph.D. Committee Member Professor Emeritus Electrical & Computer Engineering	Barry K. Karch, Ph.D. Committee Member Principal Electronics Research Engineer Air Force Research Laboratory

Thomas Wischgoll, Ph.D.
Committee Member
Professor
Wright State University

John G. Weber, Ph.D.
Associate Dean
School of Engineering

Eddy M. Rojas, Ph.D., M.A., P. E.
Dean, School of Engineering

ABSTRACT

SCENE MOTION DETECTION IN IMAGERY WITH ANISOPLANATIC OPTICAL TURBULENCE

Name: Van Hook, Richard Lowell
University of Dayton

Advisor: Dr. Russell C. Hardie

In long range imaging applications, anisoplanatic atmospheric optical turbulence imparts spatially- and temporally-varying blur and geometric distortions in acquired imagery. The ability to distinguish true scene motion from turbulence warping is important for many image processing and analysis tasks. I present two novel scene-motion detection algorithms specifically designed to operate in the presence of anisoplanatic optical turbulence.

The first method is based on modeling background intensity fluctuations with a Gaussian mixture model (GMM). The GMM parameters are formed using knowledge of the theoretical turbulence tilt variance statistics derived from the Fried parameter or refractive index structure function. Thus, this new method is referred to as the *Tilt Variance GMM* (TV-GMM) algorithm. While most prior intensity methods use empirical temporal data to estimate a background model, this approach is based on the theoretical atmospheric tilt variance statistics. This approach effectively avoids contamination in the background statistics when true scene motion is present. The approach also considers

the application of global image registration as a preprocessing step to improve performance by employing the recently developed residual tilt variance analysis that accounts for image registration as described in [1].

Rather than forming its statistcal model on the (optionally registered) input imagery directly, the second method uses a turbulence simulator. Multiple realizations of atmospheric turbulence are applied to a single prototype background image to create a non-contaminated image stack of the scene background. To incorporate the spatial relationship between neighboring pixels, each pixel's intensity is treated as an independent variable of a single Gaussian distribution. Although full anisoplanatic turbulence simulators are available, alternative approaches are sufficient provided that the anisoplanatic warping is accurate. In this work, a warping simulator based on the tilt field statistics described in [2] is used to provide a fast and reasonable approximation. Because this second method uses a warping simulator, it is referred to as the *Warping Simulator Gaussian Model* (WS-GM) algorithm.

Both quantitative and qualitative performance analyses are conducted, and the proposed methods are compared against several state-of-the art algorithms using both synthetic and real world data sets. The three synthetic image datasets are generated with an anisoplanatic numerical wave-propagation simulator ([3]) that enables per-pixel motion truth. Additionally, TV-GMM and WS-GM are extended to operate on imagery containing multiple color bands, referred to as *Multiband TV-GMM* (MB-TV-GMM) and *Multiband WS-GM* (MB-WS-GM), respectively. Both TV-GMM and WS-GM outperform the benchmark methods in the study on grayscale imagery, and MB-TV-GMM and MB-WS-GM outperform their grayscale counterparts.

I dedicate this dissertation to my parents Robert and Diane for instilling in me the value of education and to my wife Jennifer for all her support and patience.

ACKNOWLEDGMENTS

I would like to express my appreciation for the numerous people whom helped me along the way. Firstly, I thank Dr. Russell Hardie as my advisor for providing me initial concepts and guiding me throughout the effort. Additionally, I thank my other committee members, Dr. Vijay Asari, Dr. Barry Karch, Dr. John Loomis, and Dr. Thomas Wischgoll, for their continual support. Additionally, Mike Rucci provided suggestions that were both insightful and invaluable. Moreover, I would also like to thank everyone in AFRL/RYMT for allowing me the necessary time and resources to conduct this research.

Most importantlty, I thank my wife Jennifer for enduring my frequent absences, both physical and mental, and supporting me wholeheartedly throughout. This research would not have been possible without her encouragement and love.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

1.1 Problem Statement

In wide field-of-view video acquisition over long ranges, atmospheric optical turbulence tends to behave in an anisoplanatic manner. The anisoplanatic turbulence imparts quasi-periodic geometric distortions that are both spatially and temporally varying. There is also a spatially and temporally varying short-exposure atmospheric optical transfer function that causes blurring. The geometric distortions from turbulence cause stationary objects such as building edges to appear to move in video. When there is real in-scene motion, such as a moving car or pedestrian, it is often difficult to distinguish this from turbulence warping. The ability to accurately detect true scene motion and distinguish it from apparent motion due to turbulence is important for object detection and tracking, scene motion analysis, turbulence characterization, and robust turbulence mitigation (TM) processing.

1.2 Existing Work

A substantial amount of research has been focused on TM in the case of a static scene [4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. In many of the proposed TM methods, turbulence induced geometric warping is estimated in the short exposure frames of a video sequence. This is done so that some form

of dewarping may be applied to correct the geometry and allow for improved image restoration. However, many methods assume that all motion is turbulence motion. Thus, if left unchecked, real moving objects in the scene will tend to be heavily distorted or blurred beyond recognition using such methods. Therefore, a prerequisite step in restoring turbulent imagery with in-scene motion is determining which pixels represent moving objects and which pixels are part of a static background. By detecting and treating the moving object pixels differently, one can ensure the preservation of these important scene features during TM processing. As turbulence levels increase, this pixel classification problem becomes a highly non-trivial task.

Techniques for TM that are able to cope with dynamic scenes can largely be broken into two main categories: motion field and pixel intensity models. Motion field methods seek to discriminate between real motion and turbulent motion using estimated motion magnitude and/or direction with respect to a constructed reference frame. The method in [14] uses an adaptive control grid to estimate the turbulence motion vectors. A different approach in [15] creates a reference frame using rank filtering and then uses patch-based registration to estimate motion vectors. The method in [16] iterates between creating a reference with rank filtering and estimating residual motion using optical flow. Optical flow based on a reference frame is used in [17], with polar coordinates to determine motion direction. The methods in [18, 19] apply motion compensating averaging on images prior to establishing motion vectors from block-matching. Finally, the techniques in [20, 21] perform background subtraction based on a reference frame, but also include a motion tracker to build up a statistical history of all moving objects.

In contrast to the motion field methods, pixel intensity methods rely on fluctuations in brightness to distinguish between foreground motion from background variations. An excellent survey of intensity-based background subtraction methods is provided in [22]. The survey includes methods such as [23, 24, 25, 26] that employ various adaptive background modeling methods. However, none

of these method reported in the survey specifically consider turbulence. Conversely, [27, 28, 29] do explicitly consider geometric warping from turbulence when forming their background models. The technique in [30] creates a codebook based on each pixel's history, although no distinction is made between types of motion. Some techniques do not fall cleanly into either category. For instance, [31, 32, 33] all combine both motion field and pixel intensity methods. The use of multiple cameras and a non-conventional camera are explored in [34] and [35], respectively. The work in [36] proposes a new method entirely, attempting to use rank optimization to decompose a matrix of image columns into background, object motion, and turbulent motion.

1.3 Proposed Work

In this dissertation, I propose two new scene-motion detection algorithms specifically designed to operate in the presence of anisoplanatic optical turbulence. The first method is based on modeling background intensity fluctuations with a Gaussian mixture model (GMM). The GMM parameters are formed using knowledge of the theoretical turbulence tilt variance statistics derived from the Fried parameter or refractive index structure function. Thus, this new method is referred to as the *Tilt Variance GMM* (TV-GMM) algorithm. While most prior intensity methods use empirical temporal data to estimate a background model, this approach is based on the theoretical atmospheric tilt variance statistics. This approach effectively avoids contamination in the background statistics when true scene motion is present. The approach also considers the application of global image registration as a preprocessing step to improve performance by employing the recently developed residual tilt variance analysis that accounts for image registration as described in [1]. Rather than forming its statistcal model on the (optionally registered) input imagery directly, the second method uses a turbulence simulator. Multiple realizations of atmospheric turbulence are applied to a single prototype background image to create a non-contaminated image stack of the scene background. To incorporate the spatial relationship between neighboring pixels, each pixel's intensity is treated as

an independent variable of a single Gaussian distribution. Although full anisoplanatic turbulence simulators are available, alternative approaches are sufficient provided that the anisoplanatic warping is accurate. In this work, a warping simulator based on the tilt field statistics described in [2] is used to provide a fast and reasonable approximation. Because this second method uses a warping simulator, it is referred to as the *Warping Simulator Gaussian Model* (WS-GM) algorithm.

I provide a detailed quantitative performance analysis over a rage of turbulence levels and compare the proposed method to several state-of-the art foreground motion detection methods using both synthetic and real world data. The three synthetic datasets are generated with a recently developed anisoplanatic numerical wave-propagation simulator [3] that allows corresponding per-pixel truth information. To the best of our knowledge, other than the author's prior work in [28, 29] and provisional patent 63155404, this dissertation is the first reported research to employ this type of simulator to study scene motion detection in anisoplanatic turbulence. Additionally, TV-GMM and WS-GM are extended to operate on imagery containing multiple color bands, referred to as Multiband TVGMM (MB-TV-GMM) and *Multiband WS-GM* (MB-WS-GM), respectively. I demonstrate that TV-GMM and WS-GM both outperform the benchmark methods in this dissertation using grayscale imagery and that MB-TV-GMM and MB-WS-GM outperform their grayscale counterparts when color imagery is available.

The remainder of this dissertation is organized as follows. In Chapter II, I present the basic phenomenology of turbulence and the mathematics of tools used to model turbulence including the turbulence tilt statistics used in the classifier background models. The TV-GMM and WS-GM algorithms are formally presented in Chapters III and IV, respectively. The experimental setup is explained in Chapter V, and the experimental results are presented in Chapter VI. Finally, I offer my conclusions in Chapter VII.

CHAPTER II

TURBULENCE MODELING AND SIMULATION

2.1 Phenomenology

Atmospheric optical turbulence is often characterized by the refractive index structure parameter [37, 38]. This is denoted with the variable C_n^2 and has units of $\mathrm{m}^{-2/3}$. In some imaging scenarios, this parameter may vary along the optical path. In this case, it is represented as the function $C_n^2(z)$ where z is the distance from the source. Another key turbulence statistic is the atmospheric coherence diameter or Fried parameter [37, 38]. This parameter plays a major role in defining the level of blurring and warping due to atmospheric optical turbulence. The Fried parameter can be expressed in terms of $C_n^2(z)$ as

$$r_0 = \left[0.423\left(\frac{2\pi}{\lambda}\right)^2 \int_{z=0}^{z=L} C_n^2(z)\left(\frac{z}{L}\right)^{5/3} dz\right]^{-3/5}, \tag{2.1}$$

where λ is the wavelength and L is the optical path length.

What is critical to the proposed scene motion detection algorithms is the amount of warping motion caused by the turbulence. Using knowledge of the warping, I build a background model that captures temporal intensity variation due to turbulence. The turbulence warping motion can be characterized by the point source angular tilt variance statistic. For a given point source position, let the apparent tilt vector in units of radians be denoted $\tilde{\boldsymbol{\alpha}}_T = [\tilde{\alpha}_{T_x}, \tilde{\alpha}_{T_y}]^{\mathsf{T}}$. The one-axis Zernike

tilt (Z-tilt) variance [39] in units of radians squared is given by

$$\tilde{\sigma}_T^2 = \left\langle \tilde{\alpha}_{T_x}^2 \right\rangle = \left\langle \tilde{\alpha}_{T_y}^2 \right\rangle = 0.1820 \left(\frac{D}{r_0}\right)^{5/3} \left(\frac{\lambda}{D}\right)^2, \tag{2.2}$$

where D is the aperture diameter of the optics and $\langle \cdot \rangle$ is the ensemble mean operator. Note that the ratio of the aperture to r_0 plays a critical role in governing the level of tilt variance. As this ratio increases, so does the tilt variance. This increases the difficulty of performing scene motion detection as the turbulence motion leads to false positives.

Note also that using the small angle approximation, the tilt vector can be expressed in units of pixels by multiplying by the focal length and dividing by the pixel spacing, yielding $\boldsymbol{\alpha}_T = [\alpha_{T_x}, \alpha_{T_x}]^\mathsf{T} = (l/p)\tilde{\boldsymbol{\alpha}}_T$, where l is the focal length and p is the pixel detector pitch. Therefore, the tilt variance in units of pixel spacings squared is given by $\sigma_T^2 = (l/p)^2 \tilde{\sigma}_T^2$. Following the approach in [13, 1], I will model the tilt warping at each pixel as a 2D Gaussian random vector with variance given by the theoretical atmospheric tilt variance. Thus, the tilt vector is a zero-mean Normal random vector specified as $\boldsymbol{\alpha}_T \sim \mathcal{N}\left(\mathbf{0}, \sigma_T^2 \mathbf{I}\right)$.

If image registration is performed, turbulence warping is reduced, leaving some level of uncorrected residual tilt variance. The computation of the residual atmospheric tilt variance after image registration is addressed by Hardie *et al* in [1]. In the case of global registration, the global average image tilt is subtracted from the random tilt fields. The residual tilt variance then turns out to be spatially varying and different in the horizontal and vertical dimensions [1]. Let the horizontal and vertical residual tilt variances as a function of pixel location $\mathbf{n} = [n_1, n_2]^\mathsf{T}$ be denoted $\sigma_{R_{xx}}^2(\mathbf{n})$ and $\sigma_{R_{yy}}^2(\mathbf{n})$, respectively. There is also a relatively small correlation between the x and y residual tilts. The covariance will be denoted as $\sigma_{R_{xy}}^2(\mathbf{n})$, and the correlation coefficient is $\rho_{xy}(\mathbf{n})$. These statistics may be computed using the method introduced in [1] with knowledge of the basic optical parameters, the image size used for registration, and $C_n^2(z)$. Note that if no registration is performed, then the residual tilt variance is simply the input tilt variance, yielding $\sigma_{R_{xx}}^2(\mathbf{n}) = \sigma_{R_{yy}}^2(\mathbf{n}) = \sigma_T^2$

and $\sigma^2_{R_{xy}}(\mathbf{n}) = 0$. Let the residual random tilt vector be denoted $\boldsymbol{\alpha}_R(\mathbf{n}) = [\alpha_{R_x}(\mathbf{n}), \alpha_{R_y}(\mathbf{n})]^\mathsf{T}$. I treat this also as a zero-mean normal random vector such that $\boldsymbol{\alpha}_R(\mathbf{n}) \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_R(\mathbf{n}))$, where

$$\boldsymbol{\Sigma}_R(\mathbf{n}) = \begin{bmatrix} \sigma^2_{R_{xx}}(\mathbf{n}) & \sigma^2_{R_{xy}}(\mathbf{n}) \\ \sigma^2_{R_{xy}}(\mathbf{n}) & \sigma^2_{R_{yy}}(\mathbf{n}) \end{bmatrix}. \tag{2.3}$$

The optical parameters used for the simulation study presented in this dissertation are listed in Table 2.1. The simulated camera is operating in the green visible wavelength at a range of 7 km. The optical cut-off frequency ρ_c is one half that of the sampling frequency. This provides Nyquist rate sampling and no aliasing. The turbulence parameters used are given in Table 2.2. Note that here we use a constant C_n^2 profile and 5 levels of turbulence. Consider the case of level 4 turbulence as shown in Table 2.2 along with the optical parameters in Table 2.1. Here the RMS tilt is 4 pixels with no registration. For an image size of 1001×1001 pixels with ideal global registration, the residual RMS tilt is shown in Fig. 2.1. One can see that the RMS tilt varies across the image, with the highest residual in the corners and the lowest in the center. This is because the global image shifts are most representative of the bulk of the image in the center [1]. The average residual RMS tilt in Fig. 2.1 is 3.0268 pixels, compared with the original unregistered RMS tilt of 4. The reduction in RMS tilt can help improve scene motion detection by suppressing the global component of the turbulence motion in the imagery.

For the proposed motion detection algorithms, it is assumed that the optical parameters and $C_n^2(z)$ are known so that tilt statistics described in this section may be computed. In practice, $C_n^2(z)$ may be obtained using a scintillometer, differential image motion monitor, or scene-based method employing a standard camera [40, 41, 42, 43, 1].

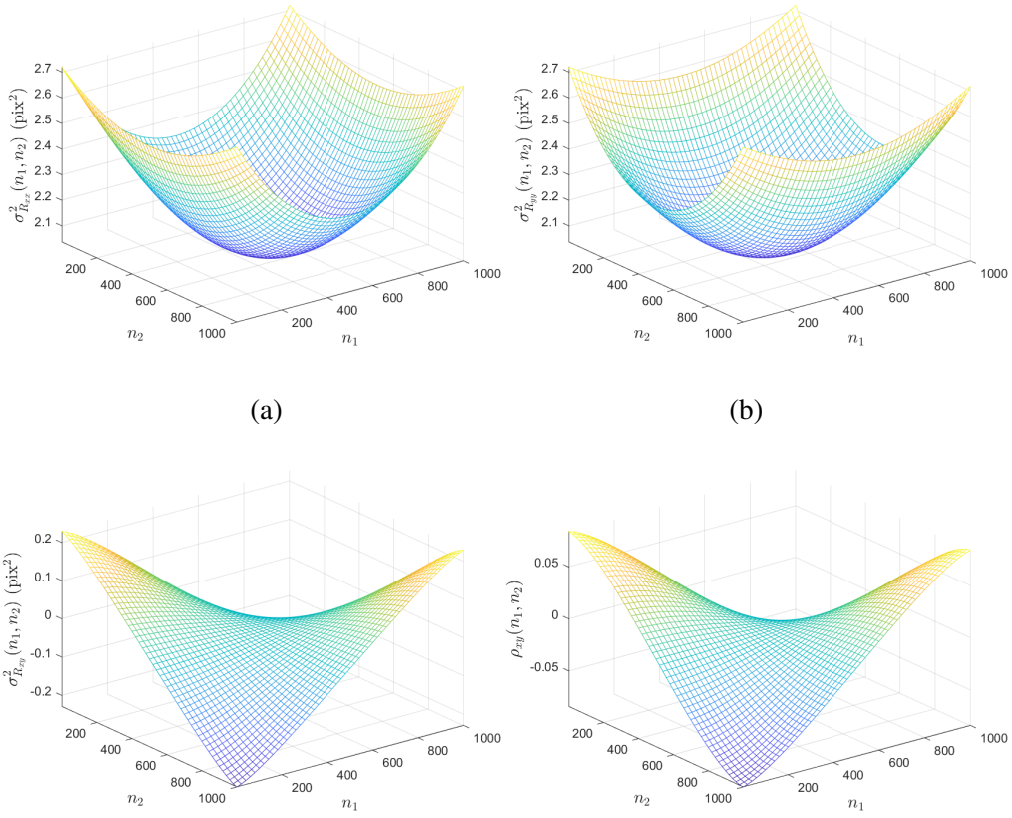


Figure 2.1: Residual tilt variance after global registration for an image of size of 1001×1001 for level 3 turbulence from Table 2.2 with a raw tilt variance of $\sigma_T^2 = 4.0000$ pixels^2 (RMS tilt of 2.0000 pixels). The horizontal or x residual tilt variance is shown in (a), and the vertical or y residual tilt variance is shown in (b). The covariance between x and y is shown in (c), and the correlation coefficient is shown in (d).

Table 2.1: Optical parameters used in simulation results.

Parameter	**Value**
Aperture	$D = 0.2034$ (m)
Focal length	l =1.2000 (m)
F-number	5.8997
Wavelength	λ =0.5250 (μm)
Object distance	L =7000 (m)
Pixel spacing (Nyquist)	p =1.5487 (μm)
Optical cut-off frequency	ρ_c =322.8571 (cycles/mm)

Table 2.2: Turbulence parameters used in the simulation results.

	Turbulence Degradation					
Parameter	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**	**Level 6**
Relative strength	**Light**		**Moderate**		**Heavy**	
$C_n^2 \times 10^{-15}$ ($\mathrm{m}^{-2/3}$)	0.0307	0.1227	0.4909	1.9638	4.4187	7.8554
Theoretical r_0 (m)	0.3863	0.1681	0.0732	0.0319	0.0196	0.0139
Theoretical D/r_0 (unitless)	0.5265	1.2097	2.7789	6.3843	10.3857	14.6331
Isoplanatic angle (pixels)	13.4454	5.8516	2.5472	1.1087	0.6816	0.4826
RMS Z-tilt σ_T (pixels)	0.5	1	2	4	6	8

2.2 Turbulence Simulation

With the phenomenology of atmospheric turbulence established, it is possible to add realistic anisoplantic turbuence effects to imagery. This may be useful when turbulent imagery is not available. An additional benefit of artifically adding turbulence is perfect knowledge of the turbulent effects that were added. The turbulence simulator used in this dissertation, described fully in [3], using split-step propagation to propagate a series of point spread functions (PSFs) from the object plane to the image through a series of phase screens. As shown in Figure 2.2, each point source in the object plane has a unique propagation path, which gives rise to spatially-correlated turbulence.

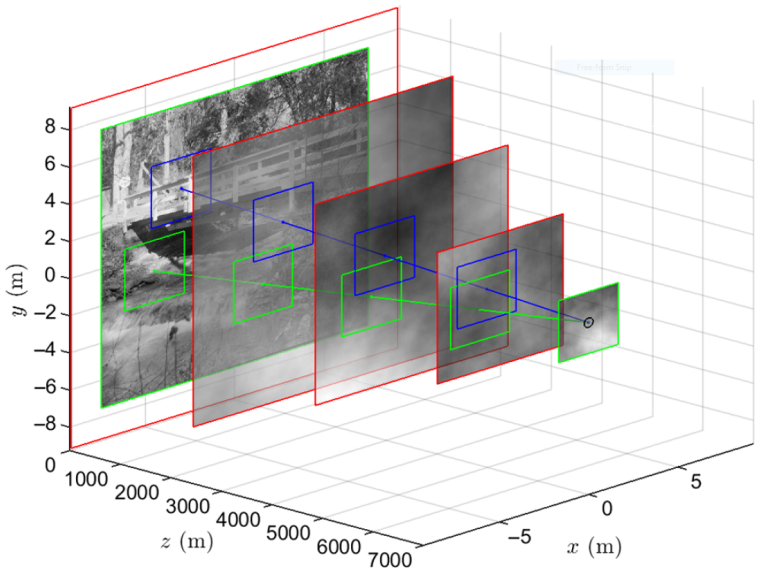


Figure 2.2: Phase screen geometry

The split-step process is shown in Figure 2.3. The point sources are modeled as identical 2-D Gaussian windowed sinc functions with quadratic phase and are propagated to each phase screen using Fresnel Angular Spectrum Propagation. A Gaussian window is used to absorb energy that reaches the edge of the simulation area. Phase screens are typically spaced equidistantly, and each phase screen is modeled as a modified Von Karmon power spectral density (PSD). Voelz's critical sampling criterion is used to set both phase screen and propagation sampling rates. Wavefronts are collimated at the pupil plane, and the PSFs are then used in a spatially-varying sum operation with an ideal image to produce a simulated image with realistic optical turbulence degradation. Figure 2.4 shows an example of a turbulent-free input image and a corresponding output image with artifical turbulence applied.

A drawback of the turbulence simulator described above is its computational complexity, driven primarily by numerical wave propagation. An alternative approach to modeling turbulence is to examine the theoretical tilt correlations. It is straightforward to filter white noise to generate x

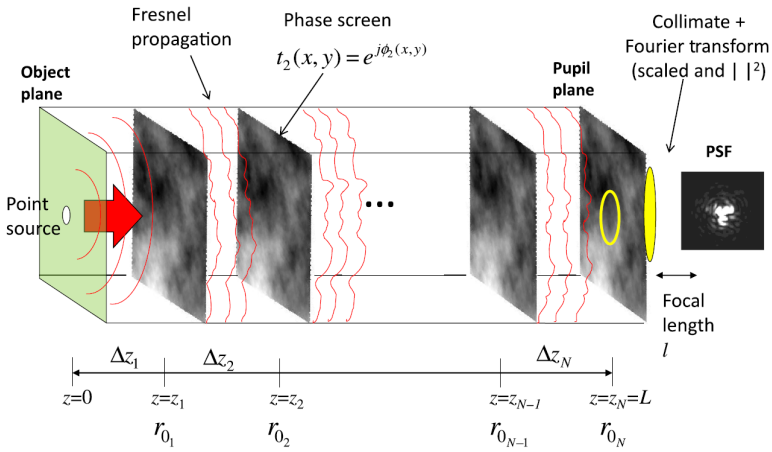
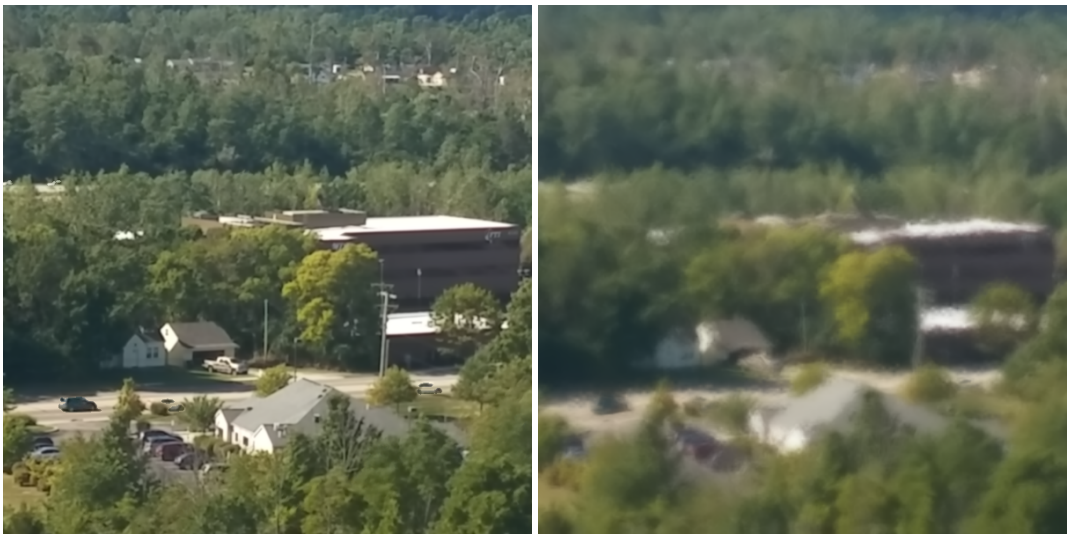


Figure 2.3: Split-step diagram



(a) (b)

Figure 2.4: An original image (a) and the resulting image with artifical turbulence (b).

and y motion fields with appropriate autocorrelation within each motion field. However, in reality, the x and y motion fields are correlated and so the warping must account for that correlation, the statistics of which is described in [3, 41]. To achieve blurring, the warped image is blurred with the average theoretical short exposure PSF. While the resulting simulation does not give the full anisoplanatic PSF blurring that is achieved with the numerical simulation in [3], it does give both accurate anisoplanatic warping and the correct average short and long exposure PSF blurring, and it does so orders of magnitude faster than [3]. Note that the proposed WS-GM motion classifier described in Chapter IV employs the warping portion of this turbulence simulation approach but not the PSF blurring aspect.

CHAPTER III

TILT VARIANCE-BASED GAUSSIAN MIXTURE MODEL

Let us now turn our attention to scene motion detection. Drawing upon the statistics from Chapter II, we propose an anomaly detection method that uses the turbulence tilt-variance to creat a Gaussian mixture model, hence the name *Tilt Variance-based GMM* (TV-GMM). The TV-GMM classifier is intended to model the intensity fluctuations that result exclusively from turbulence warping. This is done on a pixel-by-pixel basis. When an intensity value is not well described by that pixel's TV-GMM, the anomalous value is declared scene motion.

3.1 Baseline Algorithm

A block diagram of the proposed method is provided in Figure 3.1. The input to the system is a sequence of short-exposure images $f_k(\mathbf{n})$ for $k = 1, 2, ..., K$ as described in Section 2.2. The next step in Figure 3.1 is an optional global image registration. This may be done to compensate for camera platform motion or simply to reduce the turbulence warping to remove the global component to reduce the tilt variance as described in Section II. In our implementation, we use a normalized cross-correlation operation to estimate the shift between a reference frame and all of the other frames. The frames are then aligned using an integer shift. We limit the cross-correlation search window to twice the RMS pixel Z-tilt to mitigate maximums whose displacement is unlikely.

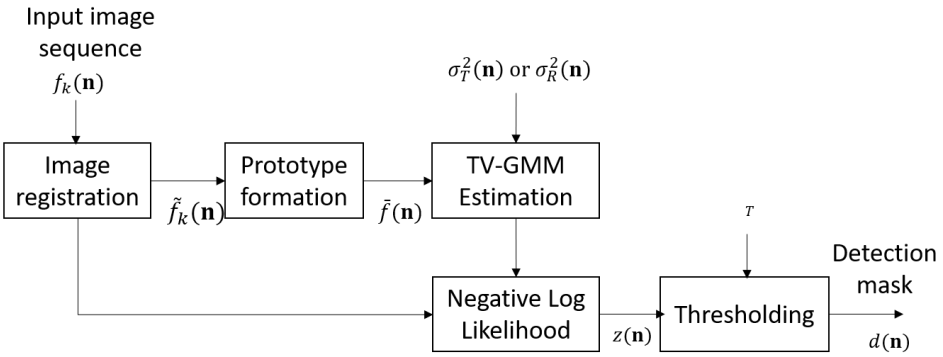


Figure 3.1: Block diagram of the proposed tilt variance based Gaussian mixture model scene motion detection algorithm.

Figure 3.2 illustrates the actual horizontal and vertical displacements of a single frame for a turbulence level of 6, before and after global registration is applied. One can see that the global frame registration shifts the tilt distribution towards zero and lowers the mean tilt for the frame. This type of global registration reduces false positive scene-motion detections for most algorithms at a small increase in computational burden.

After applying global image registration, we shall represent these frames as $\tilde{f}_k(\mathbf{n})$. Next, we form a prototype image $\bar{f}(\mathbf{n})$. This image is intended to be an estimate of the warp-free background image with no moving objects present. We have found that a temporal median is highly effective in providing an image with the approximately correct background geometry, and it is highly robust in eliminating foreground scene motion. This is represented as

$$\bar{f}(\mathbf{n}) = \text{median}\left\{\tilde{f}_1(\mathbf{n}), \tilde{f}_2(\mathbf{n}), ..., \tilde{f}_K(\mathbf{n})\right\}. \tag{3.1}$$

Given the temporal median prototype image and the tilt variance model from Section II, we form the background TV-GMM intensity fluctuation model. The idea is that this model predicts intensity fluctuations due exclusively to turbulence warping of the background. Anomalies relative to this model would be classified as scene motion.

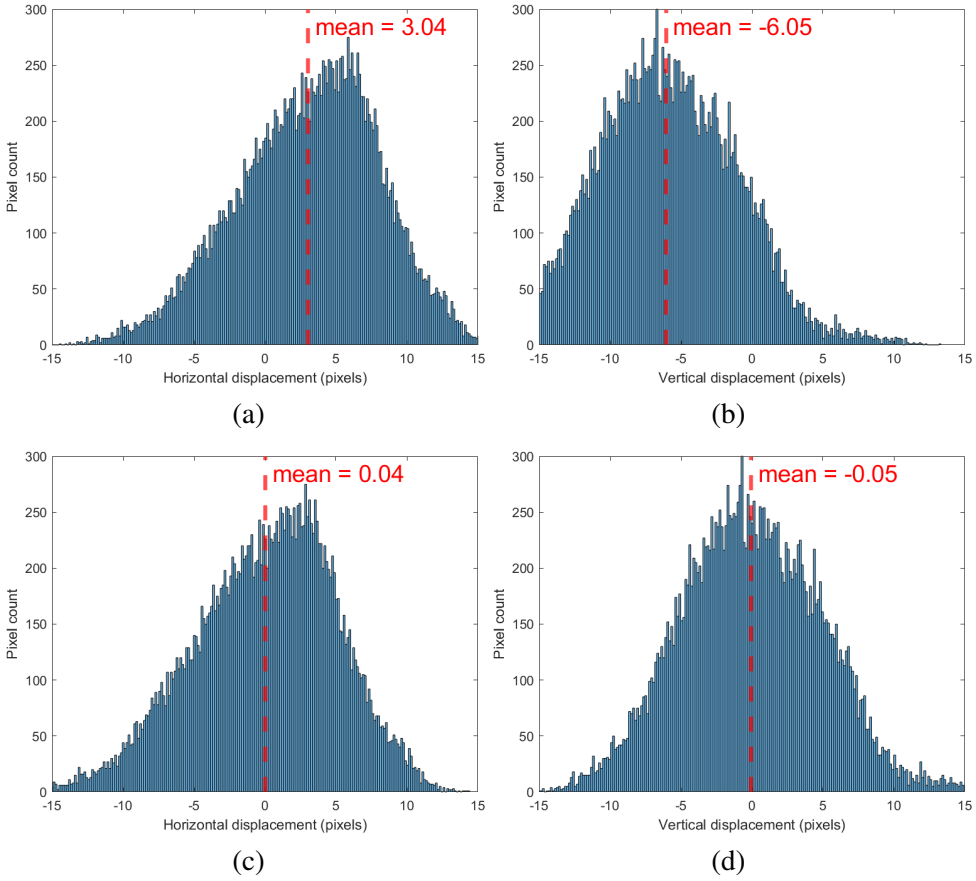


Figure 3.2: Histograms of turbulence driven pixel displacements before and after global registration that performs an integer pixel shift of $x = -3$ and $y = 6$. (a) Horizontal displacements before registration, (b) vertical displacements before registration, (c) horizontal displacements after registration, and (d) vertical displacements after registration.

To define the TV-GMM model, consider an $M \times M$ window about pixel $\mathbf{n}$. This window size is selected to be a multiple of the RMS tilt. In our implementation, we use $M = 2\lceil 3\mathsf{max}(\sigma_{R_x}(\mathbf{n}), \sigma_{R_y}(\mathbf{n}), 4)\rceil + 1$ with registration or $M = 2\lceil 3\mathsf{max}(\sigma_T, 4)\rceil + 1$ without. Using lexicographical notation, we define the samples in this window as

$$\bar{\mathbf{f}}(\mathbf{n}) = \left[\bar{f}_1(\mathbf{n}), \bar{f}_2(\mathbf{n}), ..., \bar{f}_{M^2}(\mathbf{n})\right]^{\mathsf{T}}. \tag{3.2}$$

The TV-GMM probability density function (PDF) is formed using one mode for every sample in the local window defined in Equation (3.2). The mean of mode m is the sample $\bar{f}_m(\mathbf{n})$, for $m = 1, 2, ..., M^2$. The weight of each mode is based on the sample's distance to the window center at location $\mathbf{n}$ scaled by the tilt variance. The weight function itself is a Gaussian to reflect the assumed Gaussian turbulence tilt distribution. The variance for all of the Gaussian modes in the window centered at pixel $\mathbf{n}$, denoted $\sigma(\mathbf{n})^2$, is set to be the sample variance of the samples in $\bar{\mathbf{f}}(\mathbf{n})$. While a constant value can be effective here also, we have observed improved performance using the local window variance.

Using the approach described above, the resulting TV-GMM PDF may be expressed as

$$p_{\mathbf{n}}(x) = \frac{1}{W(\mathbf{n})} \sum_{m=1}^{M^2} w_m(\mathbf{n}) \frac{1}{\sqrt{2\pi\sigma(\mathbf{n})^2}} \exp\left\{ -\frac{\left(x - \bar{f}_m(\mathbf{n})\right)^2}{2\sigma(\mathbf{n})^2} \right\}. \tag{3.3}$$

In the case of global registration, the mode weights have a Gaussian form using the residual tilt covariance matrix

$$w_m(\mathbf{n}) = \frac{1}{2\pi|\mathbf{\Sigma}_R(\mathbf{n})|^{1/2}} \exp\left\{ -\frac{1}{2}[x(m), y(m)]\mathbf{\Sigma}_R^{-1}(\mathbf{n})[x(m), y(m)]^T \right\}. \tag{3.4}$$

In the case of no registration, the mode weight function simplifies to an isotropic 2D Gaussian that does not depend on $\mathbf{n}$. The variance of the Gaussian is the turbulence tilt variance, yielding

$$w_m(\mathbf{n}) = \frac{1}{2\pi\sigma_T^2} \exp\left\{ -\frac{x(m)^2 + y(m)^2}{2\sigma_T^2} \right\}. \tag{3.5}$$

In either case, the weights are then normalized using the sum term

$$W(\mathbf{n}) = \sum_{m=1}^{M^2} w_m(\mathbf{n}). \tag{3.6}$$

The variables $x(m)$ and $y(m)$ in Eqs. (3.4) are the x and y displacements of sample $\bar{f}_m(\mathbf{n})$ relative to the the window center at location $\mathbf{n}$. The TV-GMM captures neighboring pixel values in proportion to the likelihood that the turbulence warping will bring each if the neighboring pixels to the window center. The impact of the tilt variance on the mode weights in Equation (3.5) and TV-GMM

PDF in Equation (3.3) is illustrated in Figure 3.3 for two different windows and two different tilt variances. Note that the tilt variance determines the spatial area that the GMM modes cover and the likelihood that the neighboring pixels at various distances will be displaced by warping to the window center. Thus, the PDF for a small tilt variance tends to be more compact and dominated by samples very close to $\mathbf{n}$. As the tilt variance increases, the local window and mode weight function expands and may end up spanning an edge or other structure. This can result in the addition of modes with very different means. This can be seen with the red PDFs in Figure 3.3. Thus, one can see that the tilt variance plays a critical role in forming the GMM PDF.

We use the GMM PDF in Equation (3.3) as the background intensity fluctuation model due to turbulence for pixel $\mathbf{n}$ in all frames of a given sequence. The likelihood value for pixel $\mathbf{n}$ and frame k is obtained by evaluating Equation (3.3) at $x = \tilde{f}_k(\mathbf{n})$ yielding,

$$p_{\mathbf{n}}\left(\tilde{f}_k(\mathbf{n})\right) = \frac{1}{W(\mathbf{n})}\sum_{m=1}^{M^2} w_m(\mathbf{n})\frac{1}{\sqrt{2\pi\sigma(\mathbf{n})^2}}\exp\left\{-\frac{\left(\tilde{f}_k(\mathbf{n}) - \bar{f}_m(\mathbf{n})\right)^2}{2\sigma(\mathbf{n})^2}\right\}. \tag{3.7}$$

Note that the negative log likelihood (NLL) value is a more convenient detection statistic. Using the NLL, a large value corresponds to an anomalous intensity value relative to the background model. This would indicate a likely moving object or other in-scene motion. The NLL is given by

$$z_k(\mathbf{n}) = -\ln\left(p_{\mathbf{n}}\left(\tilde{f}_k(\mathbf{n})\right)\right). \tag{3.8}$$

Finally, a threshold is applied to produce the final binary detection value for each pixel location in every frame

$$d_k(\mathbf{n}) = \begin{cases} 1 & z_k(\mathbf{n}) > T \\ 0 & \text{otherwise} \end{cases} \tag{3.9}$$

Figure 3.4 illustrates the major steps of the proposed TV-GMM scene motion detection process using an example frame from the 300 frame sequence Intersection05. Figure 3.4(a) shows one short exposure frame with a moving car on the road. The frame includes simulated level 1 turbulence as described in Table 2.2. The temporal median image described in Equation (3.1) is shown in Figure

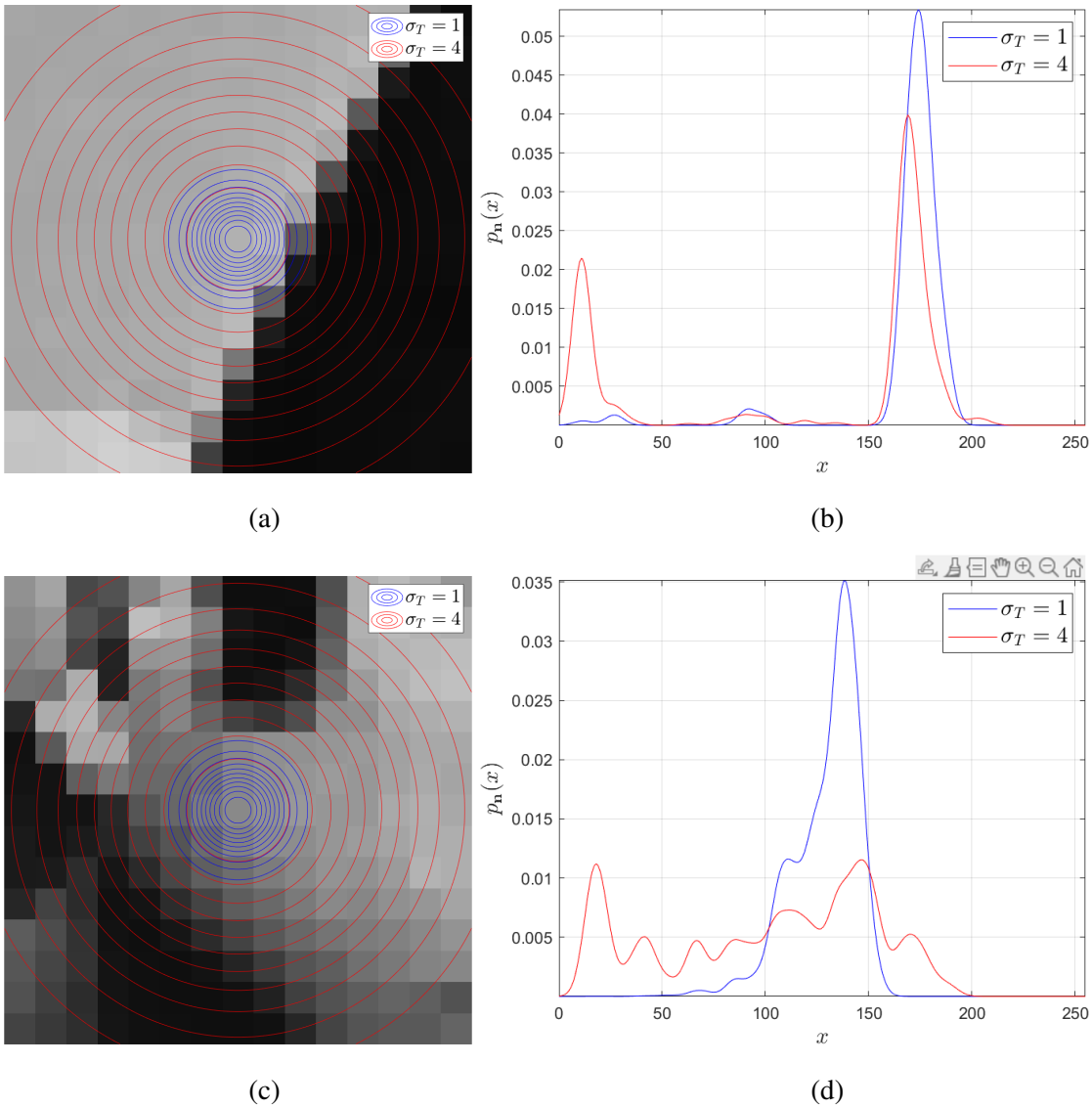


Figure 3.3: TV-GMM PDF formation from two local prototype image windows. The local windows with isocountours for Gaussian mode weighting with $\sigma_T = 1$ pixel (blue) and $\sigma_T = 4$ pixels (red) are shown in (a) and (c). The corresponding TV-GMM PDFs are shown in (b) and (d), respectively.

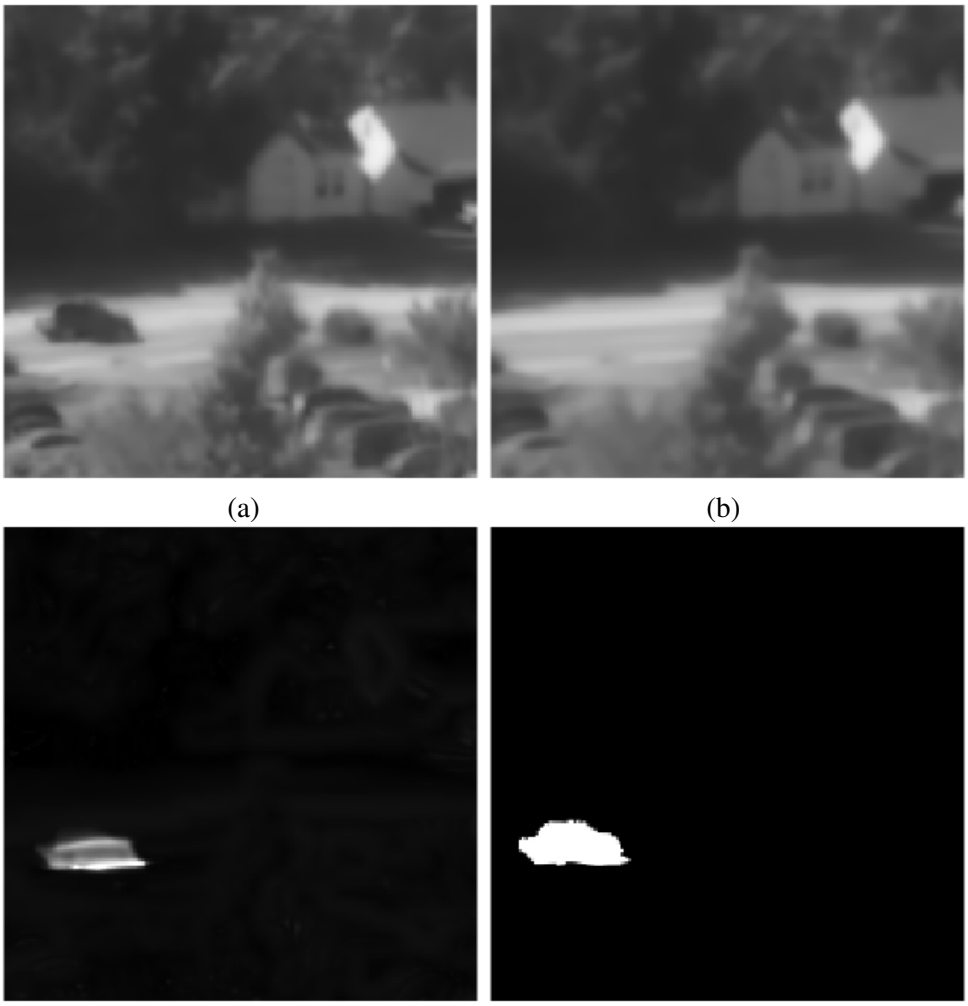
3.4(b). Note that the car is absent in the temporal median image. This is because the car is only present in this region for a few frames out of the 300 frame sequence. Thus, the temporal median takes on the more representative road pixel values. The NLL image is shown in Figure 3.4(c). Note that the car stands out strongly here, but there are also some non-zero values near the strong background edges in the scene from turbulence. The final detection mask is shown in Figure 3.4(d) after thresholding the NLL image.

In cases of heavy turbulence, the local window defined in Equation 3.2 becomes larger and more likely to contain a number of pixels of very similar or identical intensities. This gives rise to multiple GMM modes with similar or identical means, but likely different weights. To reduce computational requirements and reduce redundancy, one may use Lloyd's quantization [44] to combine modes with similar means to produce a smaller total number of modes. The resulting mode means are the representation values from the quantizer. The resulting mode weights are the sums of the original constituent mode weights within each quantization bin. Alternatively, we have found that combining modes with duplicate means (using an 8-bit image) results in a substantially more compact statistical model with negligible impact to motion detection accuracy.

3.2 Multiband TV-GMM

The baseline TV-GMM classifer described above assumes that $f_k(\mathbf{n})$ is a grayscale image with a single intensity value for each pixel location $\mathbf{n}$. However, it can be extended to work on any arbitrary number of bands B by treating each color band as a separate variable, referred to as *Multiband TV-GMM* (MB-TV-GMM). TV-GMM uses a univariate Gaussian Mixture Model whose PDF from Equation 3.3 is repeated here for convenience.

$$p_{\mathbf{n}}(x) = \frac{1}{W(\mathbf{n})} \sum_{m=1}^{M^2} w_m(\mathbf{n}) \frac{1}{\sqrt{2\pi\sigma(\mathbf{n})^2}} \exp\left\{ -\frac{\left(x - \bar{f}_m(\mathbf{n})\right)^2}{2\sigma(\mathbf{n})^2} \right\}.$$



(a) (b)

(c) (d)

Figure 3.4: Region of interest of size 200×200 pixels in the 300 frame sequence Intersection05 showing a moving car in level 3 turbulence to illustrate the TV-GMM detection algorithm processing chain. (a) Short exposure frame, (b) temporal median prototype, (c) negative log likelihood image, (d) final detection mask.

The weighting term $w_m(\mathbf{n})$ and the weighting normalization term $W(\mathbf{n})$ do not change. I define $\bar{\mathbf{f}}_m(\mathbf{n})$ to be a B-element vector containing the pixel intensities for all B bands at pixel location $\mathbf{n}$. Because MB-TV-GMM is multivariate, the grayscale variance σ must be extended to be the covariance of each color band

$$\boldsymbol{\Sigma}(\mathbf{n}) = \hat{\boldsymbol{\Sigma}}(\mathbf{n}) + \sigma_\eta^2 \mathbf{I}. \tag{3.10}$$

$\hat{\boldsymbol{\Sigma}}(\mathbf{n})$ is the $B \times B$ covariance of $\bar{\mathbf{f}}$ which is now of size $M^2 \times B$. σ_η^2 is an estimate of the noise that was present in f_j but suppressed with registration and the temporal median operation. In practice, the noise would be different for each mode (i.e., at each pixel location) and for each color band. However, to keep the number of parameter estimates manageable, I assume that σ_η^2 is identical both across all color channels and for all pixels in the image stack. Although this assumption is simplistic in nature, it works very well across all datasets tested. For the synthetic data set, good results were obtained by setting σ_η^2 to be equal to the Gaussian noise injected during dataset creation.

The MB-TV-GMM PDF can therefore be expressed as

$$p_{\mathbf{n}}(\mathbf{x}) = \frac{1}{W(\mathbf{n})} \sum_{m=1}^{M^2} w_m(\mathbf{n}) \frac{\exp\left\{-\frac{1}{2}\left(\mathbf{x} - \bar{\mathbf{f}}_m(\mathbf{n})\right)^{\mathsf{T}} \boldsymbol{\Sigma}(\mathbf{n})^{-1} \left(\mathbf{x} - \bar{\mathbf{f}}_m(\mathbf{n})\right)\right\}}{\sqrt{(2\pi)^B \left|\boldsymbol{\Sigma}(\mathbf{n})\right|}}. \tag{3.11}$$

Because this is now a multivariate PDF, the likelihood value for pixel $\mathbf{n}$ and frame k is obtained by evaluating Equation 3.11 at $\mathbf{x} = \tilde{\mathbf{f}}_k(\mathbf{n})$, resulting in

$$p_{\mathbf{n}}(\tilde{\mathbf{f}}_k(\mathbf{n})) = \frac{1}{W(\mathbf{n})} \sum_{m=1}^{M^2} w_m(\mathbf{n}) \frac{\exp\left\{-\frac{1}{2}\left(\tilde{\mathbf{f}}_k(\mathbf{n}) - \bar{\mathbf{f}}_m(\mathbf{n})\right)^{\mathsf{T}} \boldsymbol{\Sigma}(\mathbf{n})^{-1} \left(\tilde{\mathbf{f}}_k(\mathbf{n}) - \bar{\mathbf{f}}_m(\mathbf{n})\right)\right\}}{\sqrt{(2\pi)^B \left|\boldsymbol{\Sigma}(\mathbf{n})\right|}}. \tag{3.12}$$

The negative log likelihood for MB-TV-GMM therefore becomes

$$z_k(\mathbf{n}) = -\ln\left(p_{\mathbf{n}}\left(\tilde{\mathbf{f}}_k(\mathbf{n})\right)\right) \tag{3.13}$$

and the final detection mask is again obtained by

$$d_k(\mathbf{n}) = \begin{cases} 1 & z_k(\mathbf{n}) > T \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

CHAPTER IV

WARPING SIMULATOR GAUSSIAN MODEL

As described in the previous chapter, TV-GMM creates a Gaussian Mixture Model based on a single estimate of the contamination-free background. Rather than just a single image of the background, the 2nd motion classifier proposed under this research estimates a complete contamination-free image sequence. It does this by employing the warping simulator described in Section **??** on the prototype image $\bar{f}(\mathbf{n})$ to create a non-contaminated image stack representative of a static background subjected to different realizations of atmospheric turbulence. Because the classifier makes use of the warping simulator, it is referred to as the *Warping Simulator Gaussian Model* (WS-GM).

4.1 Baseline Algorithm

A block diagram of the proposed WS-GM method is provided in Fig. 4.1. As with TV-GMM, we assume an input image sequence $f_k(\mathbf{n})$ is globally registered as described in Chapter III, producing $\tilde{f}(\mathbf{n})$. Median filtering is again used to create the single non-contaminated prototype image $\bar{f}(\mathbf{n})$. The fast warping simulator described in Section 2.2 is used to create J independent realizations of $\bar{f}$ which are geometrically warped but not blurred. Each realization is a non-contaminated output frame $g_j(\mathbf{n})$ of size $N_1 \times N_2$ where $j = 1, 2, ..., J$. $g_j(\mathbf{n})$ is then globally registered to produce $\tilde{g}_j(\mathbf{n})$ in a manner similar to the formation of $\tilde{f}_k(\mathbf{n})$ from $f_k(\mathbf{n})$.

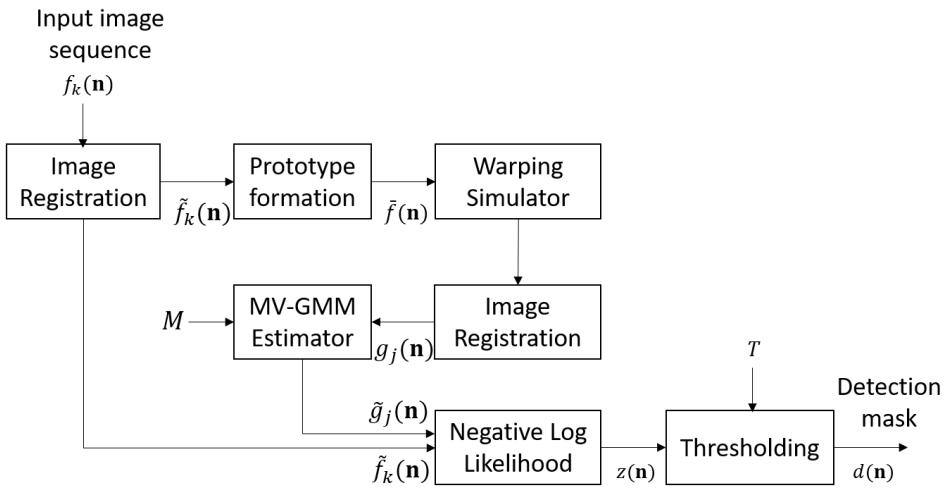


Figure 4.1: Block diagram

I define $\mathbf{g}_j(\mathbf{n})$ to be an $M \times M$ superpixel of $\tilde{g}_j$ centered on pixel location n and $\bar{g}(\mathbf{n})$ to be the mean of $\mathbf{g}_j(\mathbf{n})$ for all j. Note that although this mean is taken on a pixel by pixel basis across all J frames, this is *not* a "temporal" mean because the frames of g_j (and thus $\tilde{g}_j$) are independent realizations of the background image. Using lexographical notation, we define the samples in this window as

$$\bar{\mathbf{g}}(\mathbf{n}) = [\bar{g}_1, \bar{g}_2, ..., \bar{g}_{M^2}]^{\mathsf{T}}. \tag{4.1}$$

Each of the M^2 pixel locations within the superpixel is taken as an independent variable in a single Gaussian distribution, and $\bar{\mathbf{g}}(\mathbf{n})$ is taken to be that mode's multivariate mean. The covariance of the $M \times M$ superpixel is

$$\mathbf{\Sigma}(\mathbf{n}) = \hat{\mathbf{\Sigma}}(\mathbf{n}) + \sigma_\eta^2 \mathbf{I} \tag{4.2}$$

where $\hat{\mathbf{\Sigma}}(\mathbf{n})$ is the sample covariance estimate from all J frames of $\mathbf{g}_j(\mathbf{n})$. σ_η^2 is an estimate of the total noise and error that is present $f_k(\mathbf{n})$ but which has been suppressed or attenuated during the

(a) (b) (c)

Figure 4.2: The image $\bar{f}(\mathbf{n})$ (a) is input to the warping simulator. A sample from the resulting output image stack $g_j(\mathbf{n})$ is shown in (b), and a quiver plot of the difference to the input image is shown in (c).

production of $\tilde{g}_j(\mathbf{n})$. For the synethetic data sets, σ_η^2 is taken as three times the noise variance that was added during dataset generation. $\mathbf{I}$ is the identity matrix.

Using the approach described above, the resulting WS-GM PDF may be expressed as

$$p_{\mathbf{n}}(\mathbf{x}) = \frac{\exp\Big(-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{g}}(\mathbf{n}))^{\mathsf{T}}(\mathbf{\Sigma}(\mathbf{n}))^{-1}(\mathbf{x} - \bar{\mathbf{g}}(\mathbf{n}))\Big)}{\sqrt{(2\pi)^{M^2}|\mathbf{\Sigma}(\mathbf{n})|}}. \tag{4.3}$$

We use the PDF in Equation 4.3 as the background intensity fluctuation model due to turbulence for pixel $\mathbf{n}$ in all frames of a given sequence. Let us define $\tilde{\mathbf{f}}_k(\mathbf{n})$ to be the M^2 observation vector from the $M \times M$ superpixel in $\tilde{f}_k(\mathbf{n})$ for each pixel location $\mathbf{n}$. The likelihood value for pixel $\mathbf{n}$ and frame k is obtained by evaluating Equation 4.3 at $x = \tilde{\mathbf{f}}_k(\mathbf{n})$

$$p_{\mathbf{n}}(\tilde{\mathbf{f}}_k(\mathbf{n})) = \frac{\exp\Big(-\frac{1}{2}(\tilde{\mathbf{f}}_k(\mathbf{n}) - \bar{\mathbf{g}}(\mathbf{n}))^{\mathsf{T}}(\mathbf{\Sigma}(\mathbf{n}))^{-1}(\tilde{\mathbf{f}}_k(\mathbf{n}) - \bar{\mathbf{g}}(\mathbf{n}))\Big)}{\sqrt{(2\pi)^{M^2}|\mathbf{\Sigma}(\mathbf{n})|}} \tag{4.4}$$

where $|\mathbf{\Sigma}(\mathbf{n})|$ is the determinant of $\mathbf{\Sigma}(\mathbf{n})$.

As with TV-GMM, the negative log likelihood (NLL) is used as the detection statistic. Using the NLL, a large value corresponds to an anomalous intensity value relative to the background model.

This would indicate a likely moving object or other in-scene motion. The NLL for WS-GM is given by

$$z_k(\mathbf{n}) = -\ln\left(p_{\mathbf{n}}\left(\tilde{\mathbf{f}}_k(\mathbf{n})\right)\right). \tag{4.5}$$

Finally, a threshold is again applied to produce the final binary detection value for each pixel location in every frame

$$d_k(\mathbf{n}) = \begin{cases} 1 & z_k(\mathbf{n}) > T \\ 0 & \text{otherwise} \end{cases} \tag{4.6}$$

Although larger values of M provide marginally better results, the computation time grows on the order of $O(M^2)$. As such, even slightly larger values of M result in exponentially-increased computation times. I have found that using a value of $M = 3$ provides a good balance between performance and computational requirements. Note that for all cases, M is constrained to be an odd integer to ensure there is a pixel centered in the $M \times M$ superpixel.

Additionally, I have found that $J = 500$ to be a sufficient number of realizations for $g_j(\mathbf{n})$. Fewer than 500 realizations results in substantially-degraded performance. Using more than 500 realizations results in diminishing returns and increases the computational complexity linearly.

4.2 Multiband WS-GM

Next, let us consider how WS-GM can be extended to any number of color bands, which will be referred to as the Multiband WS-GM (MB-WS-GM). The baseline WS-GM PDF from Equation 4.3 is repeated here for convenience.

$$p_{\mathbf{n}}(\mathbf{x}) = \frac{\exp\Big(-\frac{1}{2}(\mathbf{x}-\bar{\mathbf{g}}(\mathbf{n}))^{\mathsf{T}}(\boldsymbol{\Sigma}(\mathbf{n}))^{-1}(\mathbf{x}-\bar{\mathbf{g}}(\mathbf{n}))\Big)}{\sqrt{(2\pi)^{M^2}|\boldsymbol{\Sigma}(\mathbf{n})|}}.$$

Recall that each pixel within the superpixel is already treated as an independent variable. Extension to multiple color bands additionally requires that each band within each pixel is also treated

as an independent variable. Therefore, $\bar{\mathbf{g}}(\mathbf{n})$ is redefined as

$$\bar{\mathbf{g}}(\mathbf{n}) = [\bar{g}_1, \bar{g}_2, ..., \bar{g}_{BM^2}] \tag{4.7}$$

and the updated WS-GM PDF can be expressed as

$$p_{\mathbf{n}}(\mathbf{x}) = \frac{\exp\Big(-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{g}}(\mathbf{n}))^{\mathsf{T}} (\mathbf{\Sigma}(\mathbf{n}))^{-1} (\mathbf{x} - \bar{\mathbf{g}}(\mathbf{n}))\Big)}{\sqrt{(2\pi)^{BM^2}|\mathbf{\Sigma}(\mathbf{n})|}}. \tag{4.8}$$

The matrix $\mathbf{\Sigma}(\mathbf{n})$ remains defined as

$$\mathbf{\Sigma}(\mathbf{n}) = \hat{\mathbf{\Sigma}}(\mathbf{n}) + \sigma_\eta^2 \mathbf{I} \tag{4.9}$$

where $\hat{\mathbf{\Sigma}}(\mathbf{n})$ is the covariance of $\bar{\mathbf{g}}(\mathbf{n})$ and σ_η^2 is an estimate of the noise present in $f_k(\mathbf{n})$ that is not present $\tilde{g}_j$. However, since $\bar{\mathbf{g}}(\mathbf{n})$ is now of size $J \times BM^2$, both $\mathbf{\Sigma}(\mathbf{n})$ and $\hat{\mathbf{\Sigma}}(\mathbf{n})$ are therefore now of size $BM^2 \times BM^2$. Pixel scoring is still done via the negative log likelihood

$$z_k(\mathbf{n}) = -\ln\left(p_{\mathbf{n}}\left(\tilde{\mathbf{f}}_k(\mathbf{n})\right)\right), \tag{4.10}$$

and thresholding is still applied to produce the final binary detection mask for each pixel location $\mathbf{n}$ in every frame

$$d_k(\mathbf{n}) = \begin{cases} 1 & z_k(\mathbf{n}) > T \\ 0 & \text{otherwise} \end{cases} \tag{4.11}$$

CHAPTER V

TESTING METHODOLOGY

To evaluate the effectiveness of the proposed algorithms, both the baseline TV-GMM and WS-GM methods and their extensions were implemented in engineering-level code for both quantitative and qualitative analysis. The methods were tested and compared against other turbulence motion classification techniques developed for similar applications. The testing methodology discussed in this section includes descriptions of the datasets used for evaluation, the performance metrics used for comparative assessment, and an overview of the algorithms' implementations.

5.1 Image Data

While it is desirable to test on real-world imagery, there is often a lack of per-pixel truth, hindering the development of quantitative results. Conversely, although purely synthetic imagery has corresponding truth, the degree to which that synthetic data resembles read-world data creates uncertainty in algorithm performance under real-world conditions.

In order to obtain realistic data sets with corresponding per-pixel truth, we create a pseudo-synthetic data set as follows. Our approach begins with a turbulence-free input image sequence. We compensate for global camera motion by first computing SURF features [45] from each input frame. The features for each frame are matched with those from the first frame of the image sequence. This enables the formation of a similarity transformation to spatially register each frame to the first. This

intermediate product is a stabilized image sequence that is free of platform motion but which still contains in-scene motion. Next we use a temporal median filter to create a single static background image from the sequence. We then use a combination of frame-differencing and morphological operators to segment movers within each frame. Pixels corresponding to local motion are then overlaid onto the static background image, creating a new image sequence nearly identical to the original sequence, but with a well defined per-pixel truth map for each frame. Finally, we apply the anisoplanatic optical turbulence simulator described in [3] to degrade the image sequence in a realistic manner. The simulator also outputs a turbulence motion field that is used to warp the truth maps so as to maintain a proper spatial correspondence between the movers in the degraded imagery and the truth maps. In the turbulence simulation, we use the optical parameters listed in Table 2.1 and the turbulence parameters provided in Table 2.2. The detailed simulation tuning parameters match those originally used in [3].

The key steps in the pseudo-synthetic data generation process are illustrated in Fig. 5.1. An original frame is shown in Fig. 5.1(a). The static background image with the segmented movers inserted is shown in Fig. 5.1(b). Finally, the image after the anisoplanatic turbulence simulation degradation for level 3 turbulence is shown in Fig. 5.1(c). Using the method described above, three 300-frame data sets are generated and used to generate the experimental results in this dissertation. The first, denoted *Intersection05*, is a foliage-heavy scene with sparse motion. The sequence *Thailand03* is a daytime scene of a city street with lots of moving objects. Finally, the sequence *Thailand06* exhibits motion similar to Thailand03, but with additional challenges due to glare from recent rainfall and nighttime ambient lighting. A representative original, undegraded frame from each of the sequences is shown in Fig. 5.2.

The pseudo-synthetic datasets are complemented by a real-world image sequenced denoted *Football*. It depicts two people passing an American football back and forth in light turbulence

(a) (b) (c)

Figure 5.1: Key steps in the pseudo-synthetic data generation process. (a) original turbulence-free real-world image, (b) static background image with the segmented movers inserted, (c) image with movers degraded by the anisoplanatic turbulence simulator with level 3 turbulence from Table 2.2.

conditions (Fig. 5.3). The optical parameters are provided in Table 5.1. During data collection, a scintillometer operating at 880 nm provided a path- and time-average refractive index structure of approximately $C_n^2 = 8.99 \times 10^{-15}$ m$^{-2/3}$, corresponding to an RMS Zernike tilt variance of $\sigma_T = 2.1536$ pixels. A sample frame of the Football image sequence is provided in Fig. 5.3.

Table 5.1: Optical parameters for Football dataset.

Parameter	**Value**
Aperture	$D = 0.0762$ (m)
Focal length	l =0.925 (m)
F-number	12.1391
Wavelength	λ =0.785 (μm)
Object distance	L =5,090 (m)
Pixel spacing	p =6.5. (μm)
Optical cut-off frequency	ρ_c =104.94 (cycles/mm)

(a) (b) (c)

Figure 5.2: Representative undegraded frame from each of the three sequences generated and used in this study: (a) Intersection05, (b) Thailand03, and (c) Thailand06.

Figure 5.3: Representative frame from the Football image sequence.

5.2 Other motion classifiers

We compare our proposed TV-GMM and WS-GM anomaly detection algorithm against three other algorithms. The first benchmark method is the adaptive mixture learning method described by Lee *et al* in [26]. This method adaptively estimates a GMM to model the intensity fluctuation for each pixel based solely on empirical pixel measurements. The model accounts for both the background and the foreground by employing multiple modes. Our implementation uses 3 modes

for the GMM. In contrast with the TV-GMM approach, Lee's method does not exploit any *a priori* knowledge of the turbulence statistics. It should be noted that Lee's method requires a certain number of input frames during its preliminary learning phase to allow the adaptive model to initially converge and begin to perform its best. Thus, the the first 40 output frames are excluded from the performance metrics for all of the methods studied.

The second technique is the SS-4 technique described in [46]. It is similar to Lee's algorithm with two significant exceptions. The first is that SS-4 uses all of the frames at one time to form the GMM parameter estimation. The second difference is that the number of GMM modes is adjusted dynamically based on the Akaike Information Criterion (AIC) [47].

The final technique that we use for comparison is Oreifej *et al*'s Three Way Decomposition (3WD) [36]. This method seeks to use low-rank matrix decomposition to distinguish between background, turbulence motion, and in-scene motion. The tuning parameters for 3WD were optimized separately for each turbulence level and for each of the 3 synthetic datasets used in this work. The resulting 18 sets of tuning parameters ensure the best possible performance for the algorithm.

All algorithms were implemented in MATLAB® R2021a within a Windows 10 environment. The computer contains an Intel® Core(TM) i9-9980XE processor, 128 GB of RAM, and a pair of NVIDIA Titan RTX GPUs connected with an NVLink bridge.

CHAPTER VI

EXPERIMENTAL RESULTS

In this section, I present experimental results of the proposed TV-GMM and WS-GM algorithms presented in Chapters III and IV, respectively. I first present results on the synthetic data in Section 6.1. In Section 6.2 analyze results on real data. Finally, in Section 6.3, I examine MB-TV-GMM's and MB-WS-GM's performance on multiband imagery.

6.1 Synthetic Data

The NLL from Equation 3.8 is used as the decision statistic for per-pixel labeling of motion detection. Receiver Operating Characteristic (ROC) curves [48] generated from the NLL images and truth maps are used to study classification performance at all possible operating points for both of the proposed algorithms (TV-GMM and WS-GM) and the baseline algorithms (Lee, SS-4, and 3WD). I then compute the area under the (ROC) curves (AUCs) to quantitatively compare the performance of the studied methods. A higher AUC represents better classification performance, with an area of one being ideal.

Figure 6.1 shows the area under the ROC curves for all turbulence levels listed in Table 2.2. Intersection05, Thailand03, and Thailand06 are shown in (a), (b), and (c), respectively. As expected, we see in Figure 6.1 that the performance of all the algorithms declines as the turbulence strength increases. This is due to smaller C_n^2 values corresponding to higher turbulence Z-tilt. This leads to

increased false positives from turbulence-driven apparent motion, particularly near edges and high contrast regions of the background.

Although the detailed analyses are deferred until later in this section, there are some other trends worth mentioning. Firstly, the proposed TV-GMM and WS-GM motion classifiers are the top two performers in nearly every case. In the few instances where they are not the top two performers, the performance of the lower method is effectively equal to a third classifier's performance. Furthermore, although SS-4 and Lee are similar, SS-4 outperforms Lee in all but hte lightest turbulence where they are effectively equal. This is likely due to Lee's online approach where no frames are known *a priori*, which places it at a disadvantage at classifying motion in early frames. Note that the early frames that Lee uses for training are discarded for all techniques in order to maintain a fair comparison. Lastly, compared to all other algorithms, the performance of 3WD both is the worst of all classifiers and degrades the fastest as turbulence levels increase. 3WD has the most tuning parameters of all the algorithms in this dissertation, and so extensive effort was spent on 3WD's parameter tuning. Three tuning approaches were employed for each of the 3 synthetic datasets and each of the 6 turbulence levels: 1) manual tuning, 2) an exhaustive search of the parameter space, and 3) Matlab's `fminsearch()`. All three tuning methods were in close agreement with each other for all 18 parameter sets.

While the AUC allows for a generalized performance evaluation across all operating conditions, practical applications typically operate in a narrow regime, such as ensuring that probability of true detections P_D is above some threshold or that the probability of false detections/alarms P_{FA} is below some threshold. For all three synthetic datasets analyzed in this section, I set $P_D = 0.95$. Let us focus on each synthetic data set individually for a more extensive performance analysis.

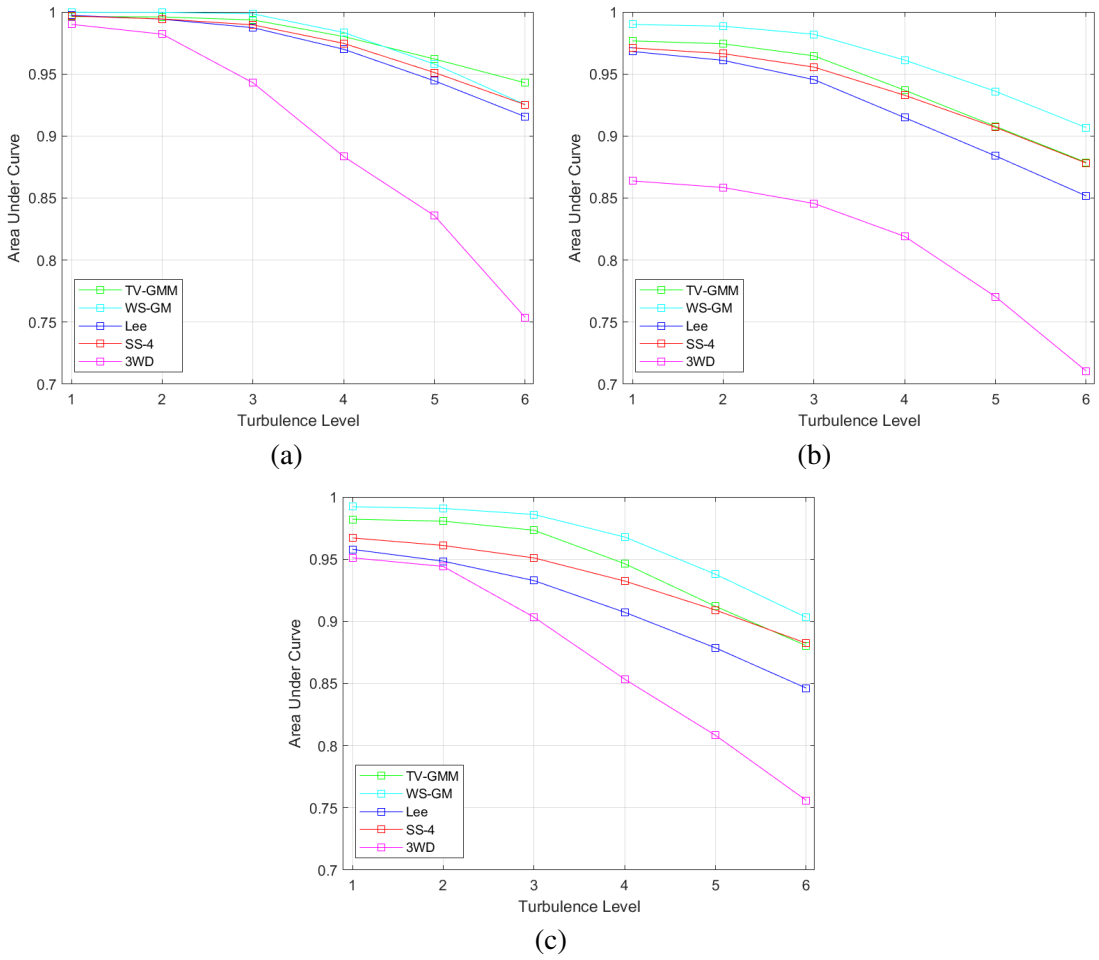


Figure 6.1: Area under ROC curves for TV-GMM, WS-GM, and the benchmark methods using imagery from (a) Intersection05, (b) Thailand03, and (c) Thailand06.

6.1.1 Intersection05

Figure 6.2 shows the individual ROC curves for all six turbulence levels. While all algorithms perform well in light turbulence, WS-GM performs almost perfectly and outperforms all other algorithms at all operating points through turbulence level 4. However, in heavy turbulence, WS-GM has the lowest P_D at lower P_{FA}s. This is likely due to the high turbulence levels greatly displacing small movers. WS-GM's spatial nature makes it more conservative in claiming detections and so in sparse data sets may require a higher false alarm rate to make all the correct positive detections.

In lighter and moderate turbulence, TV-GMM is consistently the 2nd best classifier behind WS-GM. With heavy turbulence, TV-GMM consistently maintains the highest P_D, suggesting that it is robust to turbulence-driven apparent motion. Both SS-4 and Lee perform similarly, typically falling below both TV-GMM and WS-GM, but generally meeting or outperforming 3WD. Although SS-4 and Lee have knees in the their respective curve at roughly the same false alarm rate, SS-4 exhibits a higher P_D than Lee at all false alarm rates. This is particularly noticeable as turbulence strength increases. 3WD has a very competitive P_D at very low false alarm rates, but exhibits a sharp decline much sooner than other algorithms, particularly at higher turbulence levels. This suggests that 3WD struggles to discriminate between true in-scene motion and turbulence-driven apparent motion.

Visual inspection of detection masks at various operating points is a useful analysis. Figure 6.3 shows detection masks for a sample frame of Intersection05 with a turbulence level of 1. There are three vehicles in plain view corresponding to the three large white blobs and a fourth vehicle on the right which is mostly obscured by folliage. In all detection masks shown in this dissertation, black pixels indicate true background, white pixels indicate true motion, and the colored contours show the detections for each algorithm. Because the turbulence is very light, it comes as no surprise that contours around moving objects are generally tight and conforming. Rather, it is the false

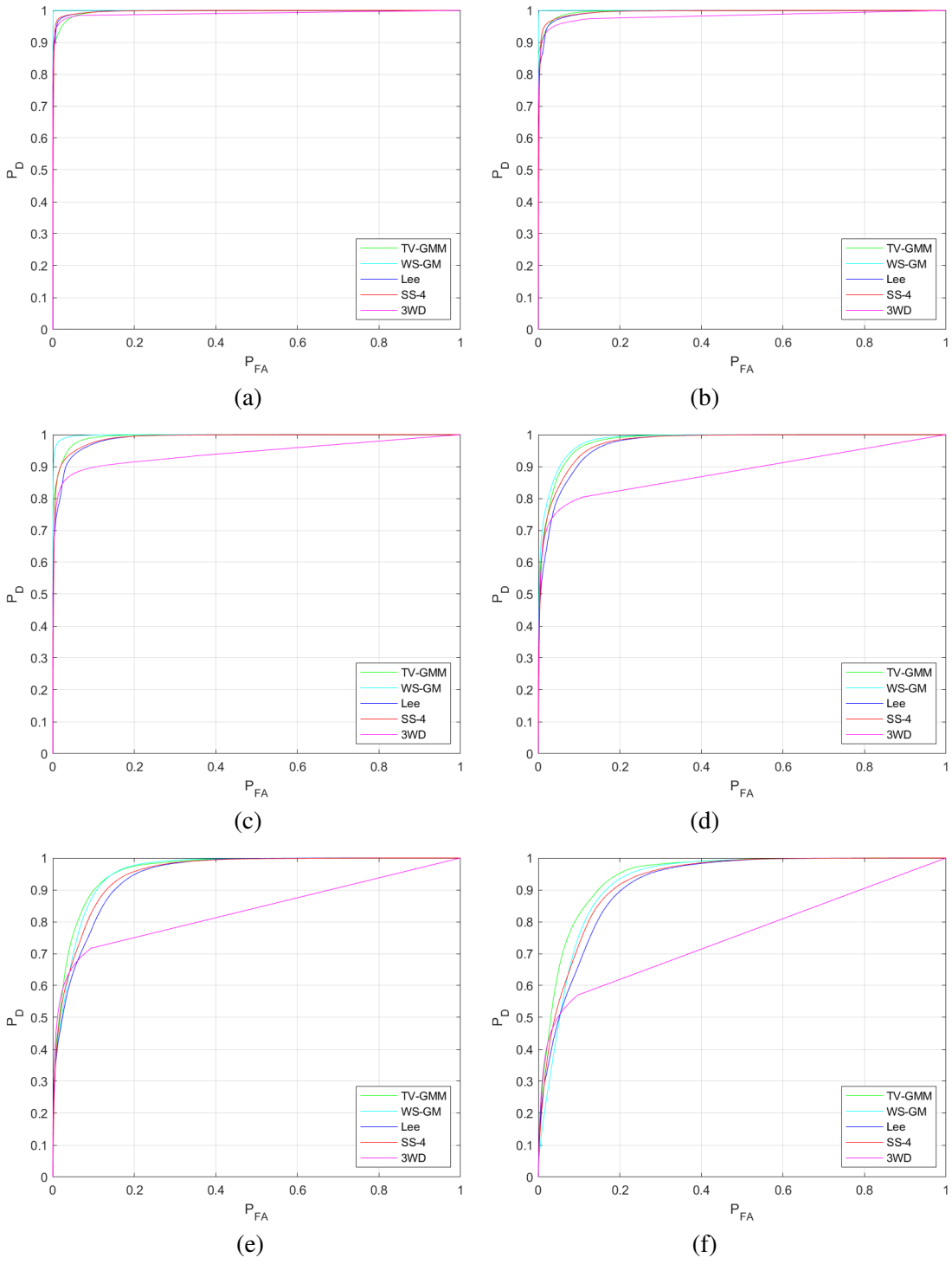


Figure 6.2: ROC curves for Intersection05 for (a)-(f) corresponding to turbulence levels 1-6, respectively.

detections for each algorithm that are revealing. WS-GM performs nearly perfectly. Both WS-GM and 3WD aren't far behind, but both have picked up a number of false detections spanning near-field, mid-field, and far-field areas of the scene. These detections generally correspond to the sharp edges between bright building surfaces and dark surfaces/folliage. Even in very light turbulence conditions, high-contrast boundaries are difficult for single-pixel detectors to correctly classify. However, WS-GM's detection mask indicates that appropriate spatial considerations may be effective at addressing such conditions.

Both Lee and SS-4 have very similar detections masks which is not surprising as they are similar techniques. While they both are able to detect all four moving vehicles in the sample frame (with tight contours), they both pick up a substantial amount of false detections grouped in two clusters. The first cluster is in the center of the image in the scene's mid-field, containing primarily folliage and a verticle pole. The second cluster is in the upper-right quarantant of the sample image and represents primarily motion in the distance and consists again primarily of folliage. The folliage is not obscuring anything in the sample image that is visible in other frames of the image sequence, and there was no apparent wind in the image sequence. Therefore, the folliage is taken as the the true static background. This indicates that Lee and SS-4 may be sensitive to any type of motion, whether true moving objects or turbulence-driven apparent motion.

Figure 6.3 is under very light turbulent conditions (Z-tilt of half a pixel) and thus shows relatively clean detection masks. To understand each classifier's effectiveness at higher turbulence levels, we examine the same sample frame under a turbulence level of 4 (corresponding to a Z-tilt of 4 pixels), shown in Figure 6.4. Although all four moving objects are generally detected by each algorithm, the contours no longer conform to the vehicles' outlines. This is due to the higher turbulence causing high-contrast edges to displace more throughout the vehicle sequence. TV-GMM explicitly addresses the probability of such displacements by use of the weighting term $w_m(\mathbf{n})$ for

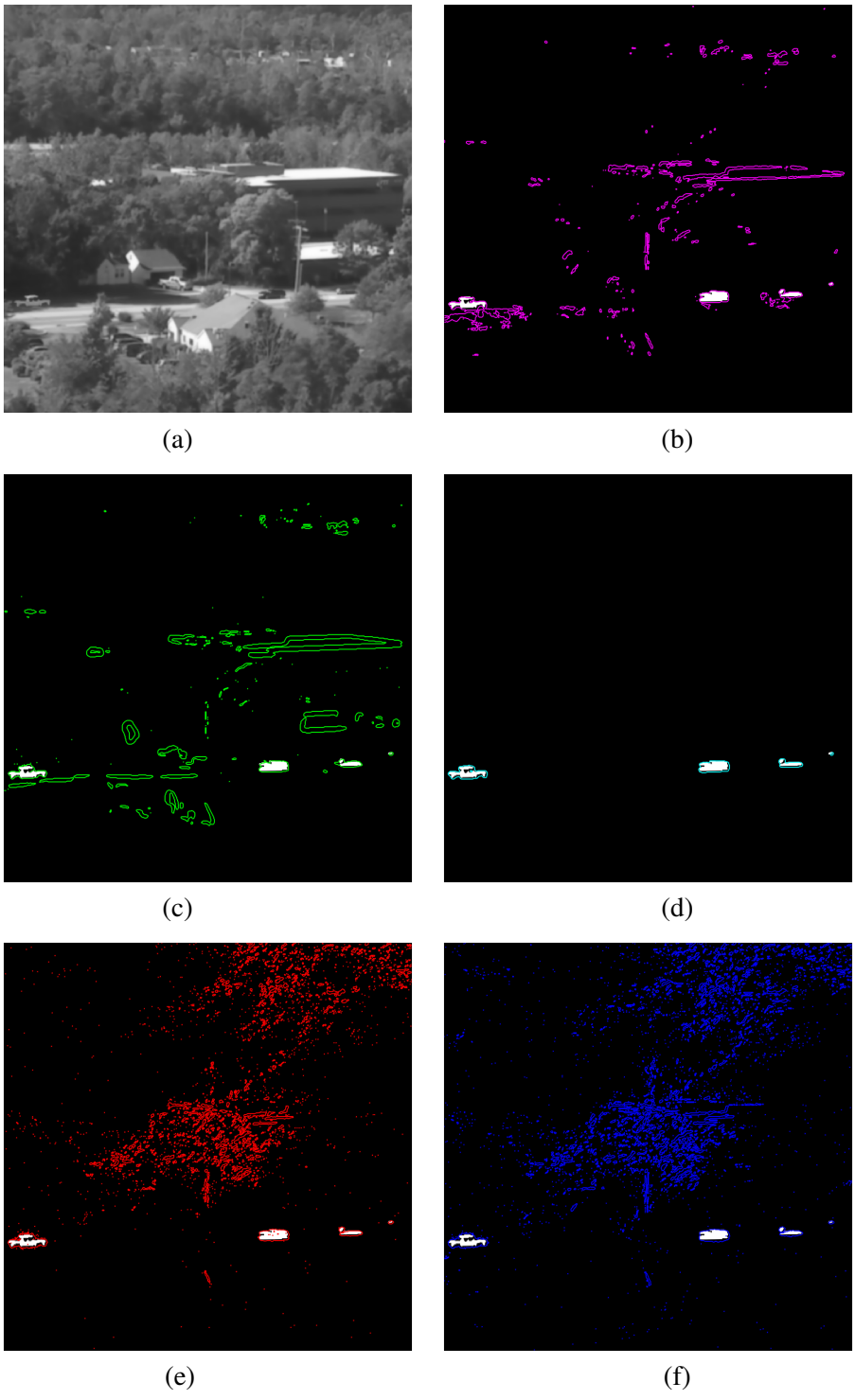


Figure 6.3: Sample turbulence level 1 frame of the Intersection05 dataset (a) with corresponding detection masks by 3WD (b), TV-GMM (c), WS-GM (d), SS-4 (e), and Lee (f) using a $P_D = 0.95$.

each mode m in the $M \times M$ window. Although there are still false detections around the vehicles and on the high-contrast edges of buildings, TV-GMM is reasonably robust around lower-contrast edges, correctly identifying the majority of the folliage as background. WS-GM has very similar performance to TV-GMM, which is expected given their ROC curves for turbulence level 2 are nearly identical. MV-GMM's patch-based approach does allow some contouring around moving objects, but like TV-GMM, it does have substantial false detections near the periphery of moving objects. Similarly, high-contrast edges like those on and near buildings are incorrectly detected as movers.

Lee, SS-4, and 3WD all detect the four moving objects with little trouble. However, the number of false detections required to make those positive detections is very high. Like TV-GMM and WS-GM, there are false detections around both the periphery of the movers and the high-contrast edges on buildings. Unlike TV-GMM and WS-GM, these three classifiers have substantial false detections on folliage in the scene. Overall, this suggests that in higher turbulence levels, Lee, SS-4, and 3WD all struggle to differentiate real-motion from turbulence-driven apparent motion.

Note that for the results shown in Figure 6.3, the P_D is maintained at 0.95. Lowering the P_D does lower the false detections, but it also removes some of the correctly-detected pixels with movers. This behavior is expected because all ROC curves in Figure 6.2(d) except 3WD have a slope of approximately 1 at $P_D = 0.95$. Therefore, a reduction of false detections comes with a corresponding reduction in true positives. The exception is 3WD whose ROC curve exhibits a noticeable bend at $P_{FA} \approx 0.1$. Its false detections can be substantially lowered with only a minor loss in detected positives only because of the abnormally high number of false detections to begin with. However, reducing the false detection rate to the bend in 3WD's ROC curve ($P_D \approx 0.75$) still results in substantial loss of detected movers.

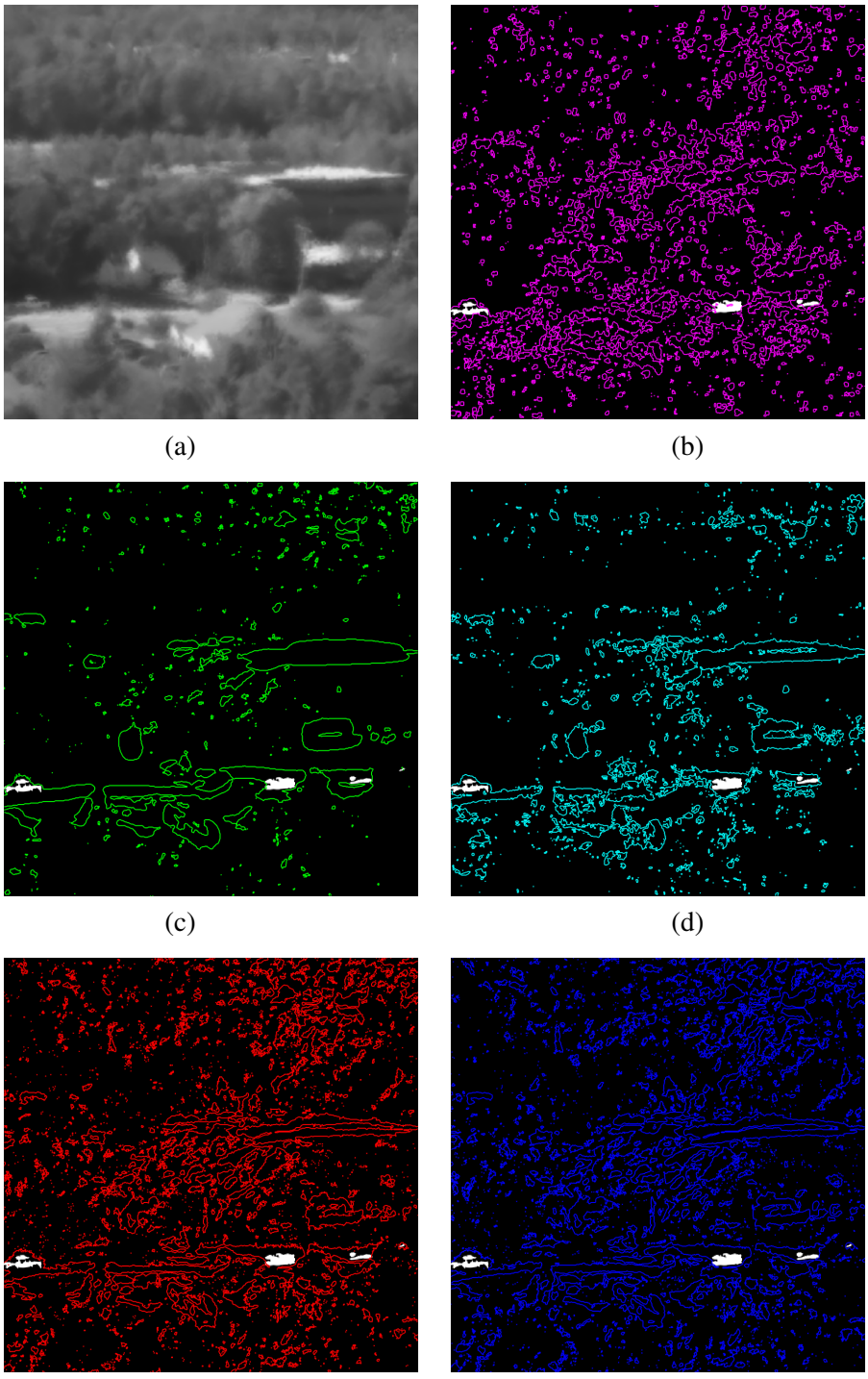


Figure 6.4: Sample turbulence level 4 frame of the Intersection05 dataset (a) with corresponding detection masks by 3WD (b), TV-GMM (c), WS-GM (d), SS-4 (e), and Lee (f) using a $P_D = 0.95$

6.1.2 Thailand03

Intersection05 was a folliage-heavy scene with relatively few movers in the mid-field. I now perform a similar analysis on an image sequence with very different content. Thailand03 is an city scene with dense vehicle traffic and pedestrians walking along the sidewalk. Because the camera is overhead looking along the street, the vehicles vary in size as they transit between near field of view and far field of view. As the vehicles move into the far field of view, the farthest vehicles becomes partially obscured by slightly-closer vehicles.

Figure 6.5 shows the ROC curves for all six turbulence levels. WS-GM either meets or outperforms all other techniques at nearly all operating points for all turbulence levels. At extremely low false alarm rates ($P_{FA} \approx 0.005$), 3WD has the highest P_D, but the corresponding probability of detection is still relatively low at $P_D < 0.5$ in light turbulence and particularly low at $P_D < 0.3$ in heavy turbulence. Although it's preferable to operate at such low false alarm rates, the corresponding detection probabilities are too low to be actionable for most applications. Furthermore, the knee in 3WD's ROC curve occurs at only slightly higher false detection rates, meaning that a useful positive detection rate cannot be obtained without a substantial increase in false positives.

TV-GMM has has similar performannce to WS-GM, but hits its knee in the curve slightly sooner. It outperforms both Lee and SS-4 at lower false alarm rates, but falls behind first SS-4 and and then Lee as the false alarm rate is increased. In lower turbulence levels, the crossover point between TV-GMM and SS-4 occurs at $P_{FA} \approx 0.075$ and gradually increases to $P_{FA} \approx 0.25$ at turbulence level 6. SS-4 is once again slightly superior to Lee at all operating points, likely as a result of its offline-style processing.

Figure 6.6 shows detection masks for a sample frame of turbulence level 1. Although the turbulence is light, it is a very busy scene with lots of moving vehicles and pedestrians on the sidewalk. WS-GM has the cleanest detection mask, showing generally-tight contours around movers, although

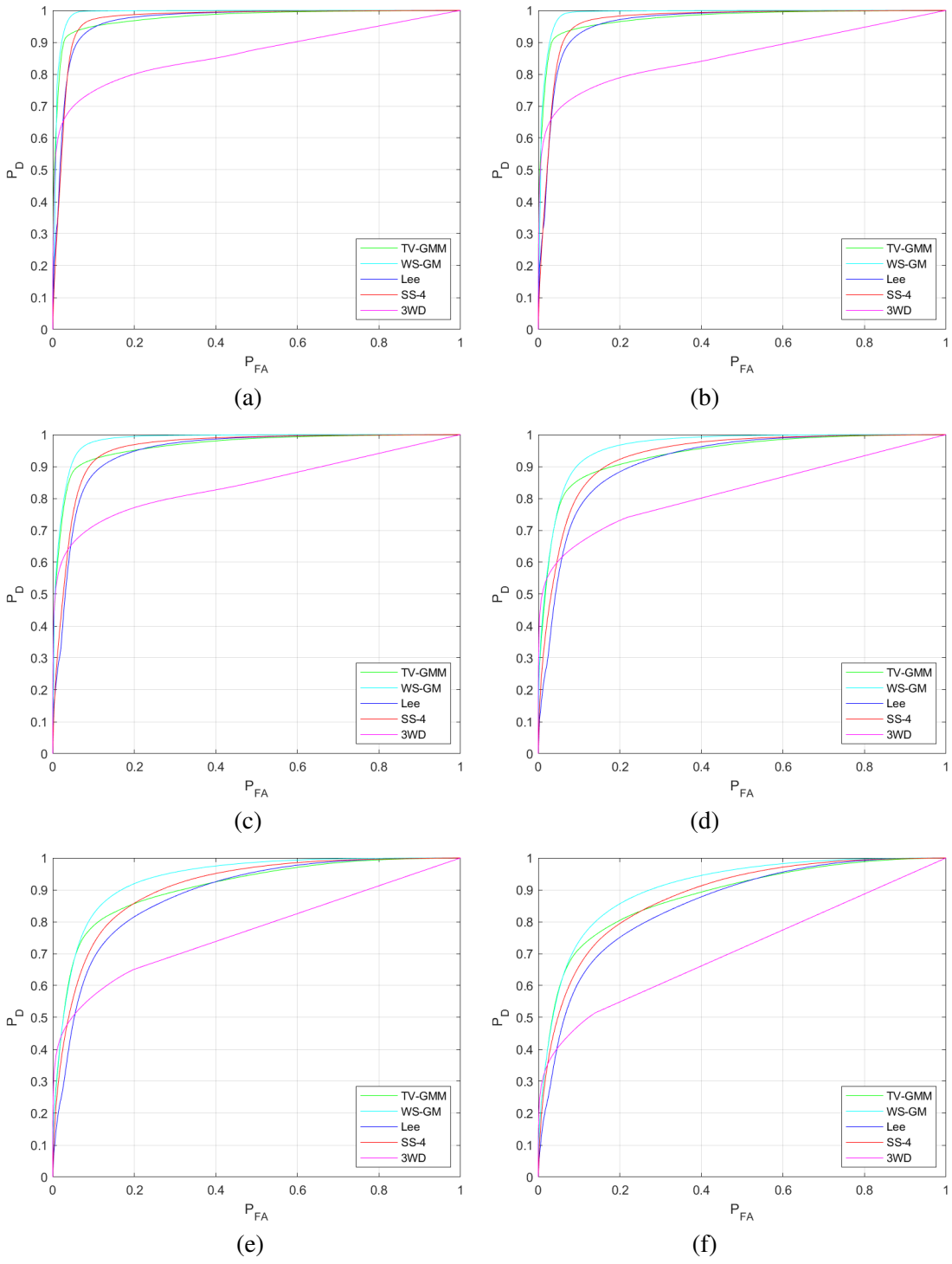


Figure 6.5: ROC curves for Thailand03 for (a)-(f) corresponding to turbulence levels 1-6, respectively.

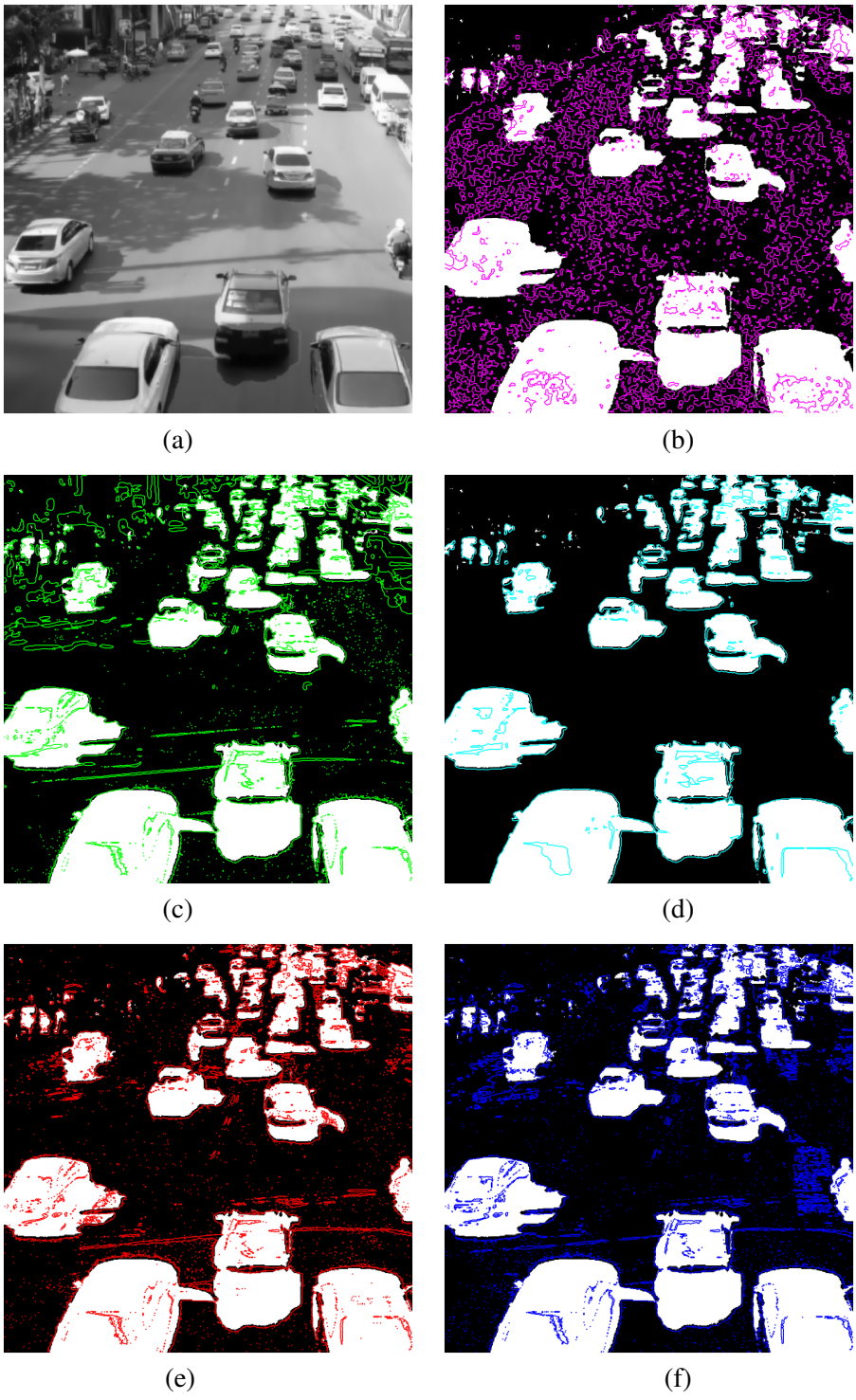


Figure 6.6: Sample turbulence level 1 frame of the Thailand03 dataset (a) with corresponding detection masks by 3WD (b), TV-GMM (c), WS-GM (d), SS-4 (e), and Lee (f) using a $P_D = 0.95$.

it does not consistently detect the interiors of moving objects. WS-GM does fail to detect some of the smallest movers that consist of only several pixels, but that's not surprising given that it examines $M \times M$ patches, and this detection mask corresponds to $M = 3$. Therefore, movers that are smaller than a 3×3 block of pixels are unlikely to be detected.

TV-GMM, SS-4, and Lee all have similar detection masks. For instance, they all detect the shadow that a vertical post casts on the road as the shadow creates high-contrast edges with the road surface. Similarly, they all detect the many of the white painted dashed lines delineating individual lanes on the road. Where the detection masks differ is in the type of false positives. TV-GMM's incorrect detections tend to have straight sides which correspond to structures in the scene such as street signs, building/window edges, stationary vehicles, etc. Conversely, SS-4's and Lee's false positives tend to be concentrated in complex shadows on the roadways. 3WD has marginal positive detections on vehicles but has a very high number of false positives. This is commiserate with 3WD's high false detection rate in the ROC curve at the given $P_D = 0.95$.

Figure 6.7 shows the same sample from for a turbulence level of 4. The same trends still hold for each of the classifiers at this higher level of turbulence. All vehicles are detected reasonably well. Additionally, at this higher turbulence level, detections generally over-extend moving objects slightly. A turbulence level of 4 corresponds to a Z-tilt of 4 pixels. The high-constrast edges of vehicles become somewhat blurred and displaced within a small area. This in turn causes the background models to be less rigid.

WS-GM once again has the fewest number of false positives. TV-GMM's false positives still tend to focus on sharp structures. In particular, at this higher turbulence level, TV-GMM is (incorrectly) picking up detections on most of the edges of shadows on the road, even if they're not straight. Conversely, SS-4 and Lee's false shadow detections are generally limited to straight, relatively-sharp edges. 3WD's false positives occur throughout the scene and have no apparent

correlation with any structure. The only location in the image where 3WD has correctly avoiding detecting motion is in the interior of the central shadow on the road.

6.1.3 Thailand06

The Thailand06 dataset is also in the middle of a city, looking down a busy street. However, it was collected at night shortly after rain shower. This results in multiple specular surfaces that may reflect moving objects in the scene but which are not actually moving themselves. Additionally, there is a large digital advertising board in the upper-left portion of the image.

Similar to before, Figure 6.8 shows the ROC curves for all motion classifiers, this time for the Thailand06 dataset. As before, WS-GM is the best performer across nearly all operating conditions. In this more complex dataset, TV-GMM has separated from SS-4 and Lee, distinguishing itself as the second best performer for $P_{FA} < 0.25$. SS-4 once again outperforms Lee across all operating conditions. Interestingly enough, 3WD maintains relatively comparable performance at lighter turbulence levels, although the familiar degradation occurs at moderate and severe turbulence levels.

The detection masks for the lowest turbulence level are shown in Figure 6.9. Both WS-GM and TV-GMM readily handle the specular reflections in the scene while still maintaining tight contours around moving objects. WS-GM has very few false detections while TV-GMM again incorrectly detects motion around high-constrast structures such as lighted advertisement signs in the upper right and the elevated walkway across the road. SS-4 and Lee both have tight contours around moving objects. Like TV-GMM, they detection a lot of the structure in the scene. Unlike TV-GMM, though, they also incorrectly detect numerous, small detections on the roadway, likely as a result of the specular reflections. 3WD struggles with these specular reflections with an extremely high number of false detections on the roadway.

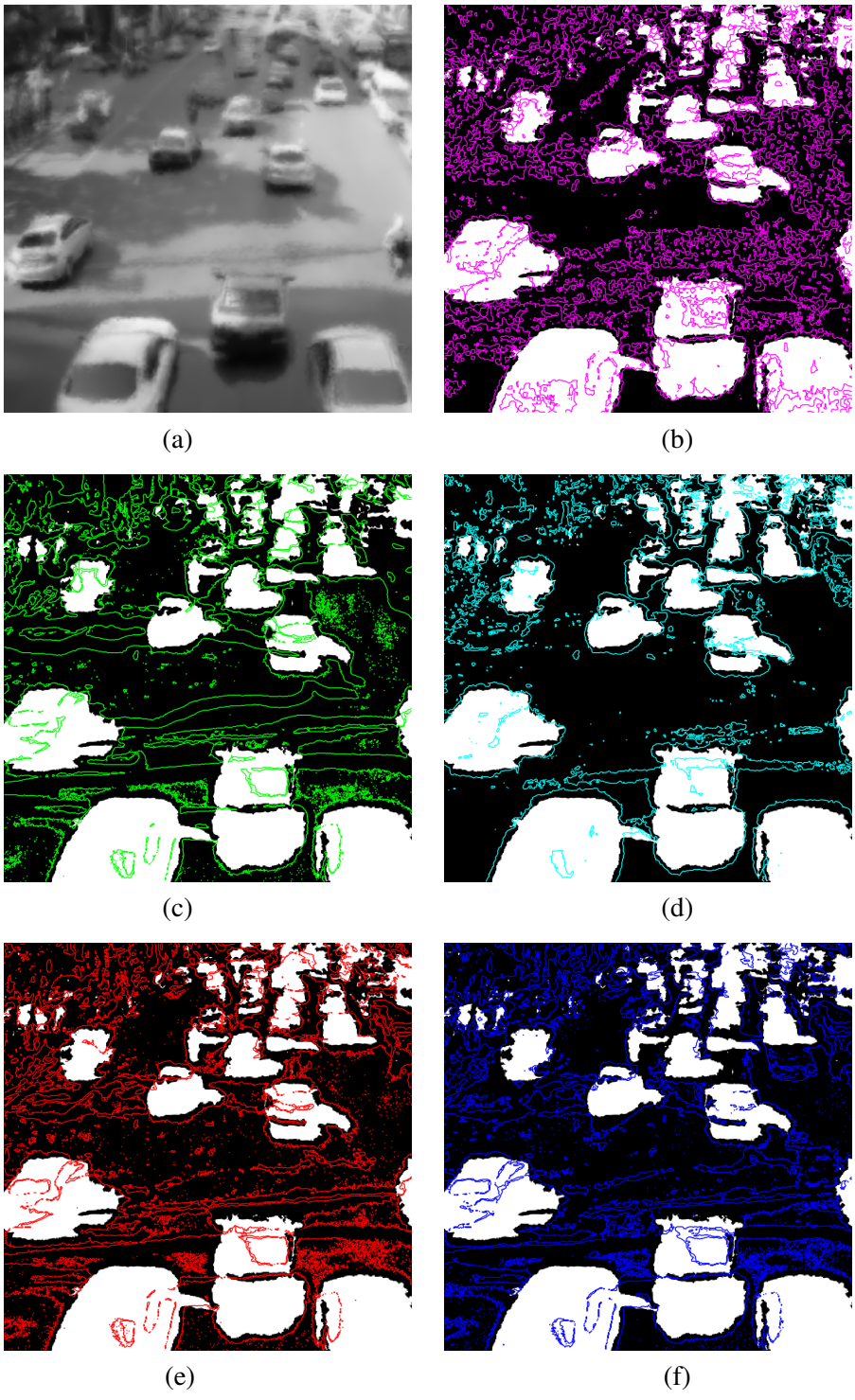


Figure 6.7: Sample turbulence level 4 frame of the Thailand03 dataset (a) with corresponding detection masks by 3WD (b), TV-GMM (c), WS-GM (d), SS-4 (e), and Lee (f) using a $P_D = 0.95$.

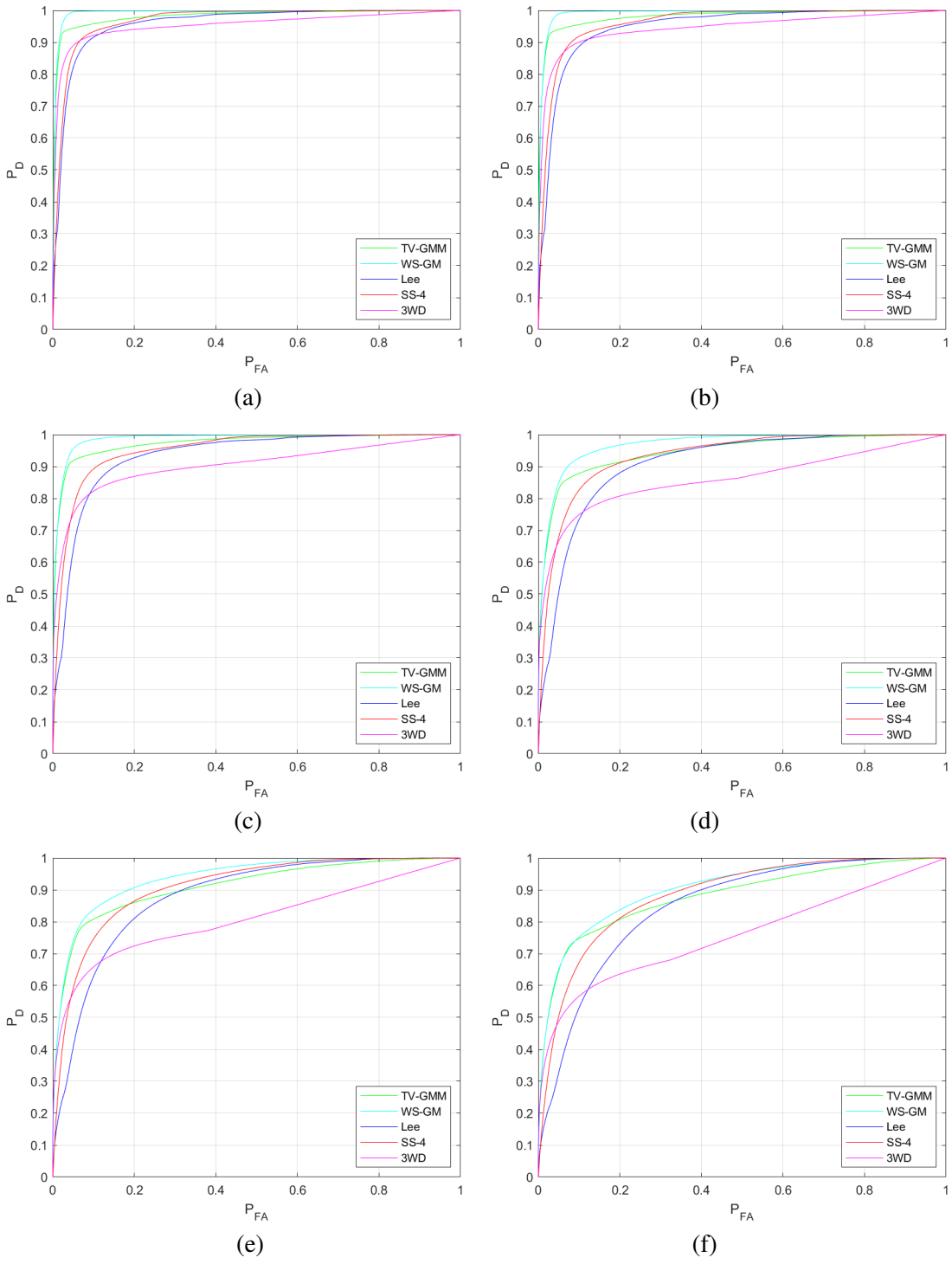


Figure 6.8: ROC curves for Thailand06 for (a)-(f) corresponding to turbulence levels 1-6, respectively.

One notable complexity in this dataset is the large display sign in the upper left. Although the sign is stationary itself, the content it displays changes throughout the image sequence. As such, pixel values within that display area may stay relatively constant only for short durations (e.g., over several seconds) before drastically different content is displayed. The practical implication is that all pixels contained within the display area are effectively the result of a pseudo-random signal generator. As such, it is an incredibly challenging aspect of the Thailand06 dataset, and it is hard to fault any of the five algorithms in this dissertation for incorrect classification of these pixels.

For the higher turbulence level shown in Figure 6.10, the specular reflection is still a major issue for 3WD, but it also picks up numerous false detections on surfaces throughout that image that aren't reflective. SS-4 and Lee still have some specular reflection detections on the roadway, but the major source of false detections for them (as well as WS-GM and TV-GMM) is high-contrast edges in the scene. TV-GMM, SS-4, and Lee all detect the white dashed line delinating adjacent lanes, but WS-GM correctly rejects those pixels as movers. Overall, WS-GM provides the cleanest detection mask. Note that once again the detected contours around the moving vehicles are looser than before; this is a direct result of the increased turbulence level and the relatively high P_D.

6.2 Real Data

The truth data that inherently accompanies the creation of synthetic data allows a thorough analysis of the classifiers' performance. However, synthetic data often does not have the complexities and unusual behaviors that can be present in real world data. Therefore, we analyze the performance of all five classifiers against the Football dataset. Figure 6.11(a) shows frame #502, and Figure 6.12 shows frame #631. The detection masks in the left and right columns of those figures are based upon two different thresholds. The left detection masks are based on detecting (correctly or incorrectly) 1% of the pixels in the image. It is important to note that the *exact* number of detected pixels may

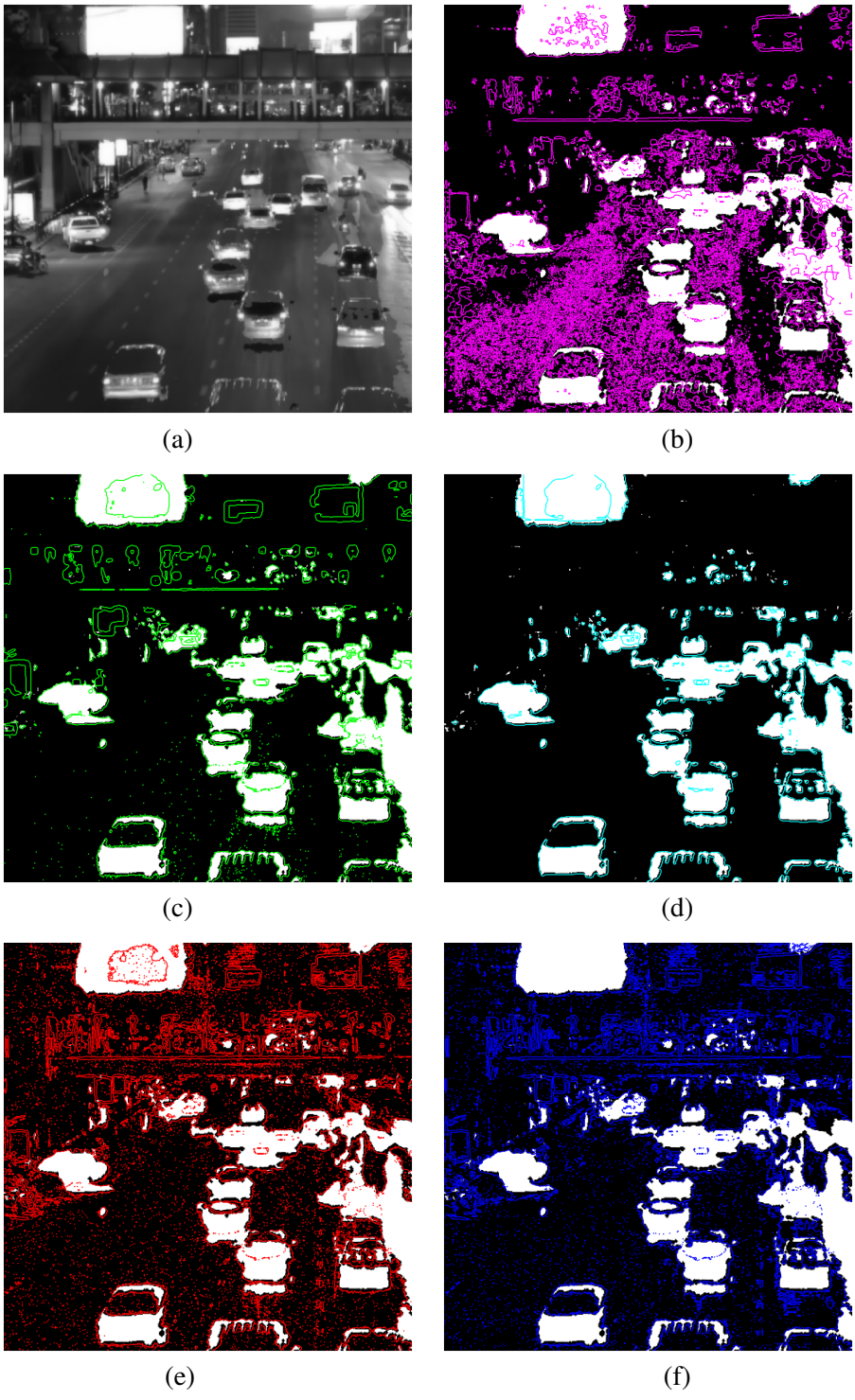


Figure 6.9: Sample turbulence level 1 frame of the Thailand06 dataset (a) with corresponding detection masks by 3WD (b), TV-GMM (c), WS-GM (d), SS-4 (e), and Lee (f) using a $P_D = 0.95$.

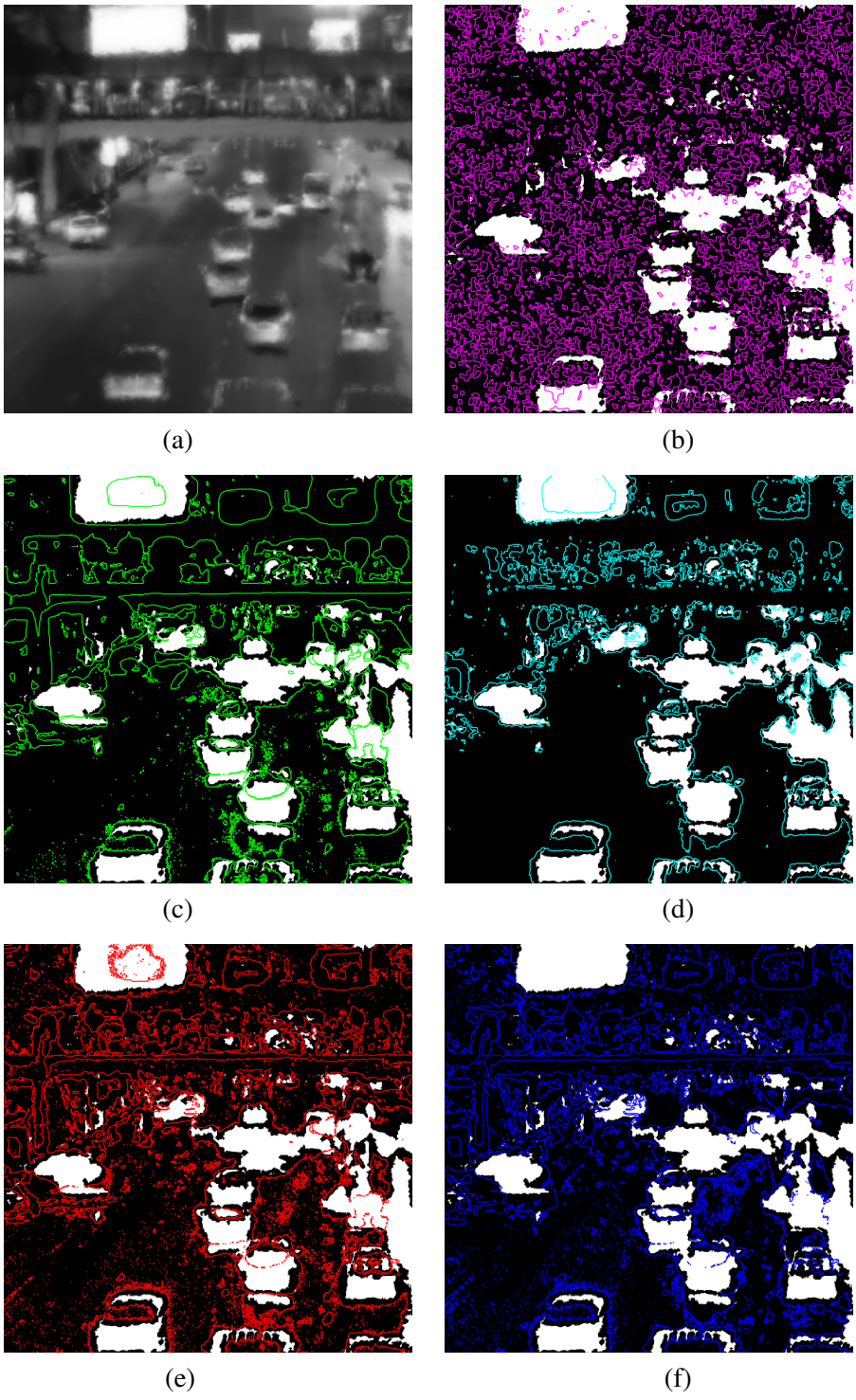


Figure 6.10: Sample turbulence level 4 frame of the Thailand06 dataset (a) with corresponding detection masks by 3WD (b), TV-GMM (c), WS-GM (d), SS-4 (e), and Lee (f) using a $P_D = 0.95$.

differ between classifiers due to that classifiers distribution of scores. For instance, Lee's detection mask in Figure 6.11(f) has many more detected pixels than the other classifiers because the highest 1% of Lee's scores are all the same value, so the threshold T was lowered to just below that value.

The detection masks on the right side of the two figures are based upon the same P_D from ROC curves. Because no per-pixel truth is available for this real world dataset, the image sequence was visually analyzed, and a hand truth map was created as shown in Figures 6.11(g) and 6.12(g). As before, white pixels represent known movement, and black pixels represent known background. Known movement was outlined conservatively to avoid incorrect labeling of pixels. Additionally, "don't care" boundary boxes (shown in gray) were created around the known truth to prevent any missed truth detections from negatively affecting any classifier's performance.

I refer to detections masks based on the 1% thresholding strategy as *Similar Pixel* (SP) detection masks and detection masks based on the hand-truth as *ROC* detection masks. Both detection masks will show the pixels that each classifier deems to be most likely to be movers. However, the SP detection mask is intended to show a comparable number of detected pixels. Although the ROC mask will show the same P_D across all classifiers, each classifier's ROC detection mask may contain more or fewer pixels than the corresponding SP mask and more of fewer pixels than other classifier's ROC detection mask.

Figure 6.11(a) shows a single person on the right side of the image, and no other moving objects are present in this frame. Regardless of whether SP and ROC detection masks are used, 3WD does not detect the person, instead detecting 2 circular patches on the ground and several locations on the box truck, including the passenger door, the fender, the hubcap, and the edge of the door. The four detections on the box truck are all locations where specular reflection (i.e., glare) is likely to be high, particularly the door as it is likely to be reflecting light from the passenger mirror. Although

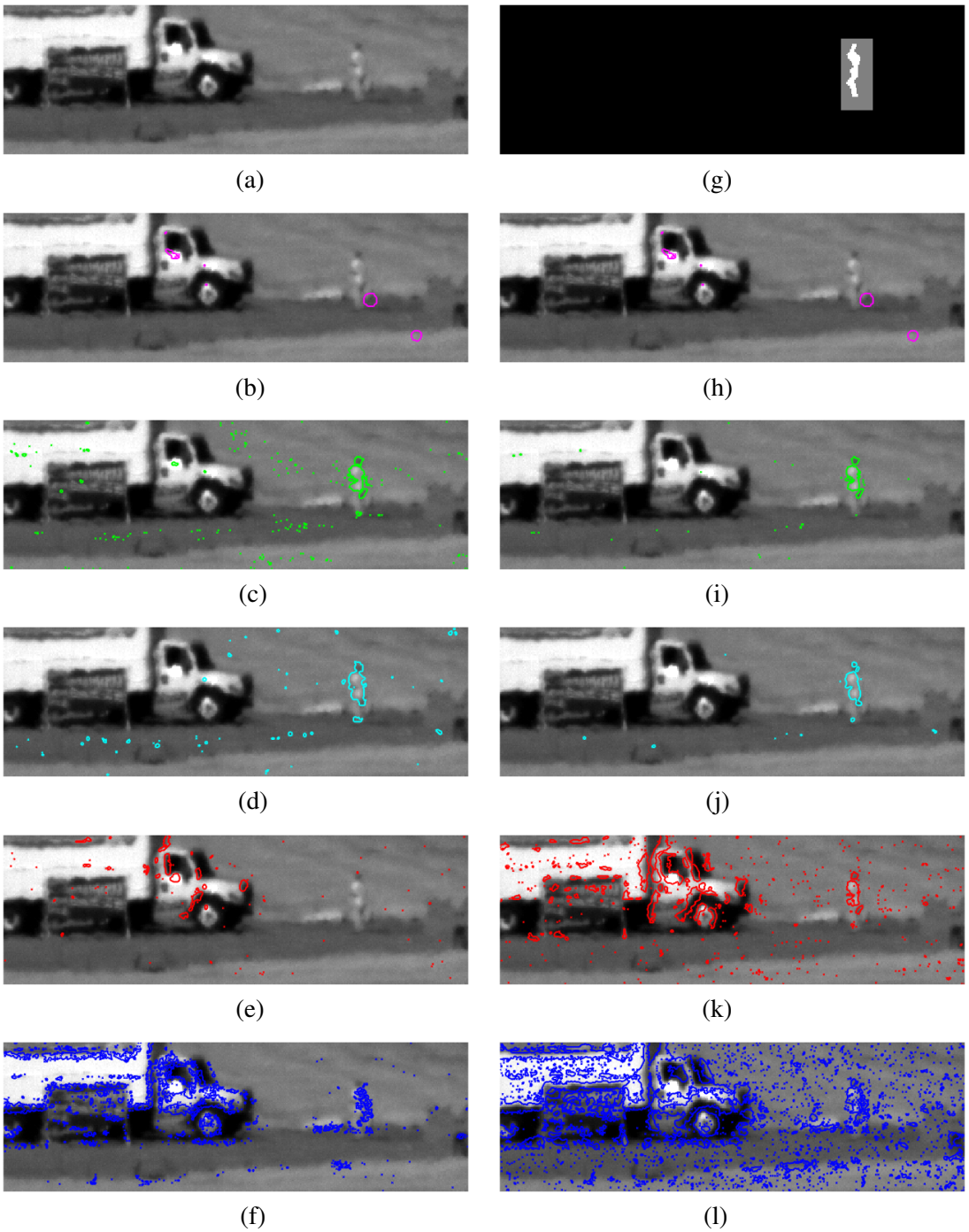


Figure 6.11: Frame 502 of the Football dataset (a) with corresponding detection masks using similar number of detected pixels by 3WD (b), TV-GMM (c), WS-GM (d), SS-4 (e), and Lee (f). A hand-truthed mask (g) with the same P_D for 3WD (h), TV-GMM (i), WS-GM (j), SS-4 (k), and Lee (l)

both masks agree that these pixels are the most likely to be movers, the true moving object is not detected.

SS-4 barely detects the person in the SP detection mask. Rather, the detections are predominantely on high-contrast edges on the box truck. The ROC detection mask has an increased number of pixels, some of which do correspond to the person. This does, however, come with a substantial increase in the number of false positives, indicating that the algorithm is having a hard time distinguishing between high-contrast edges on the truck and the true moving object.

Although Lee's SP detection mask does detect the person, it only does so with an extroidinarily high number of false alarms. As previously mentioned, the top scores for Lee all had an identical value indicating that Lee believes all the piexels shown are equally likely to be movers. The ROC detection mask allows an even higher number of pixels, but comes with a corresponding increase in the number of false alarms.

Conversely, both TV-GMM and WS-GM are able to detect the person with the fewest false alarms in their SP detection masks. The ROC detection masks show fewer pixels without any substantial degradation to the positive detections, indicating that those detected pixels were correctly determined to be among the most likely to be movers. It is possible that the false few remaining false alarms would be eliminated if the P_D were decreased slightly.

Figure 6.12 shows a second sample frame from the Football dataset. In this, the same person is seen throwing a football to a second individual who is outside the field of view. Each of the five classifiers has similar detection performance for the person in this frame as they did in the frame #501 (Figure 6.11). What is new in this sample frame is the football that is in mid-air. Every classifier except 3WD detects the football in both SP and ROC detection masks despite the football's small size.

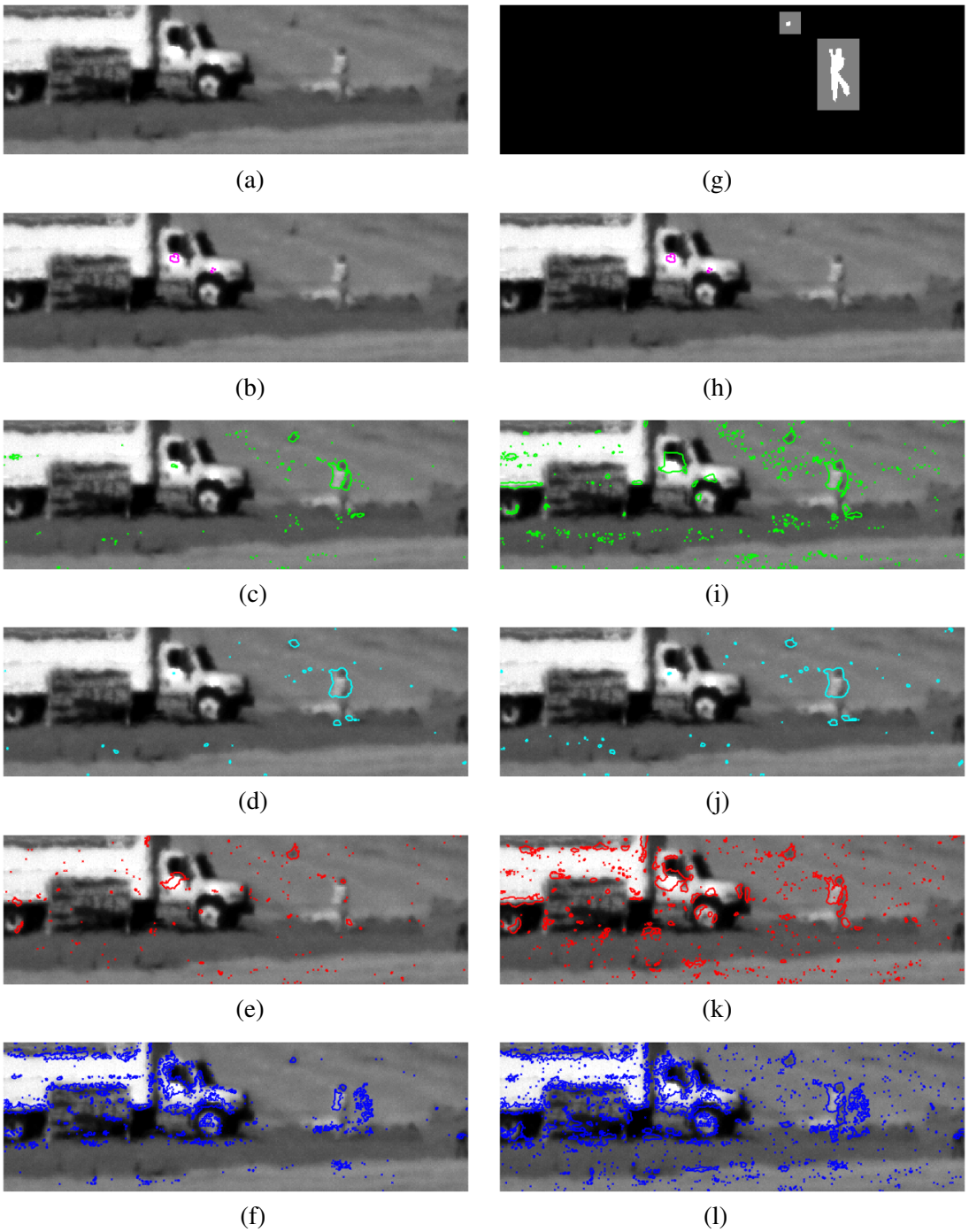


Figure 6.12: Frame 631 of the Football dataset (a) with corresponding detection masks using similar number of detected pixels by 3WD (b), TV-GMM (c), WS-GM (d), SS-4 (e), and Lee (f). A hand-truthed mask (g) with the same P_D for 3WD (h), TV-GMM (i), WS-GM (j), SS-4 (k), and Lee (l)

6.3 Multiband study

TV-GMM and WS-GM operate on monochromatic imagery, and the MB-TV-GMM and MB-WS-GM classifiers extend their respective single-band relatives to an arbitrary number of color bands. It is assumed that additional color bands better describe each pixel than a single band, thus helping to distinguish movers and background. In this study, I examine the performance diffences between TV-GMM and MB-TV-GMM as well as between WS-GM and MB-WS-GM using the synthetic datasets. Further, I use the same sample frames, turbulence levels, and P_D threshold that was used in Section 6.1 to ensure a fair comparison.

Let us first examine the ROC curves for the four classifiers using the Thailand03 dataset as shown in Figure 6.13. As before, TV-GMM-type classifiers are in green,and WS-GM-type classifiers are in cyan. The single-band versions are represented by dashed lines, and the multiband versions are represented by solid lines. There are several generic trends worth mentioning. Firstly and as expected, performance degrades as turbulence strength increases. More interestingly, like their single-band versions, MB-TV-GMM and MB-WS-GM performance is nearly identical at lower P_{FA} levels. Examining the MB-TV-GMM ROC curves, there is a performance gain compared to TV-GMM across all operating conditions for all turbulence levels. This indicates that the additional color bands are enabling a more complex GMM that is able to better identify true movers.

MB-WS-GM does not exhibit any advantage over WS-GM in light turbulence, meaning the additional color bands are not providing additional information. In moderate and heavy turbulence, however, there is a noticeable performance difference. This difference become more substantial as turbulence increases. This is natural because higher turbulence levels correspond to both higher magnitude displacements as well as more several blurring. This leads to more contamination from neighboring pixels, and the color bands enable better distinction between the moving targets and the background pixels.

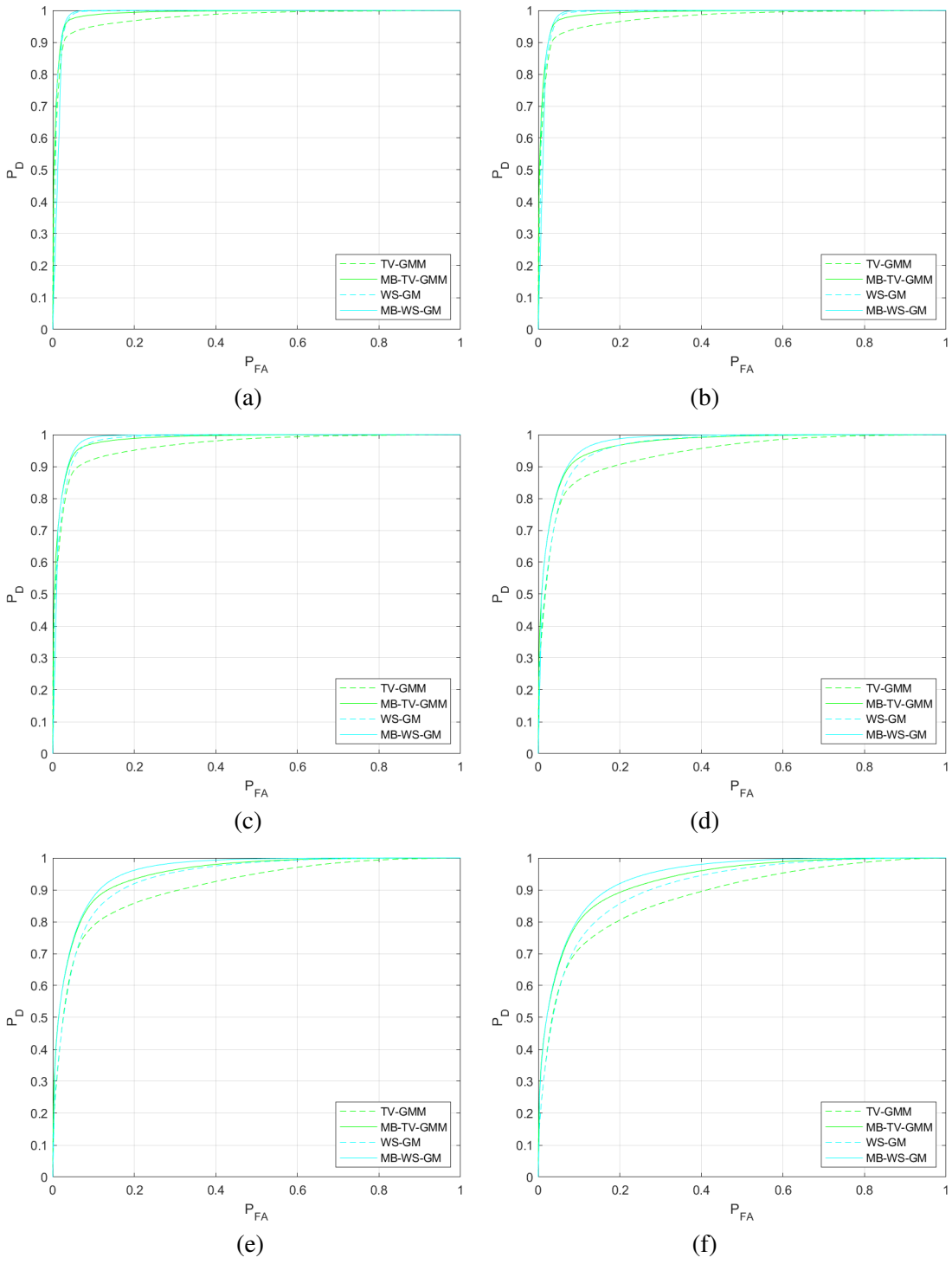


Figure 6.13: Multiband ROC curves for Thailand03 for (a)-(f) corresponding to turbulence levels 1-6, respectively.

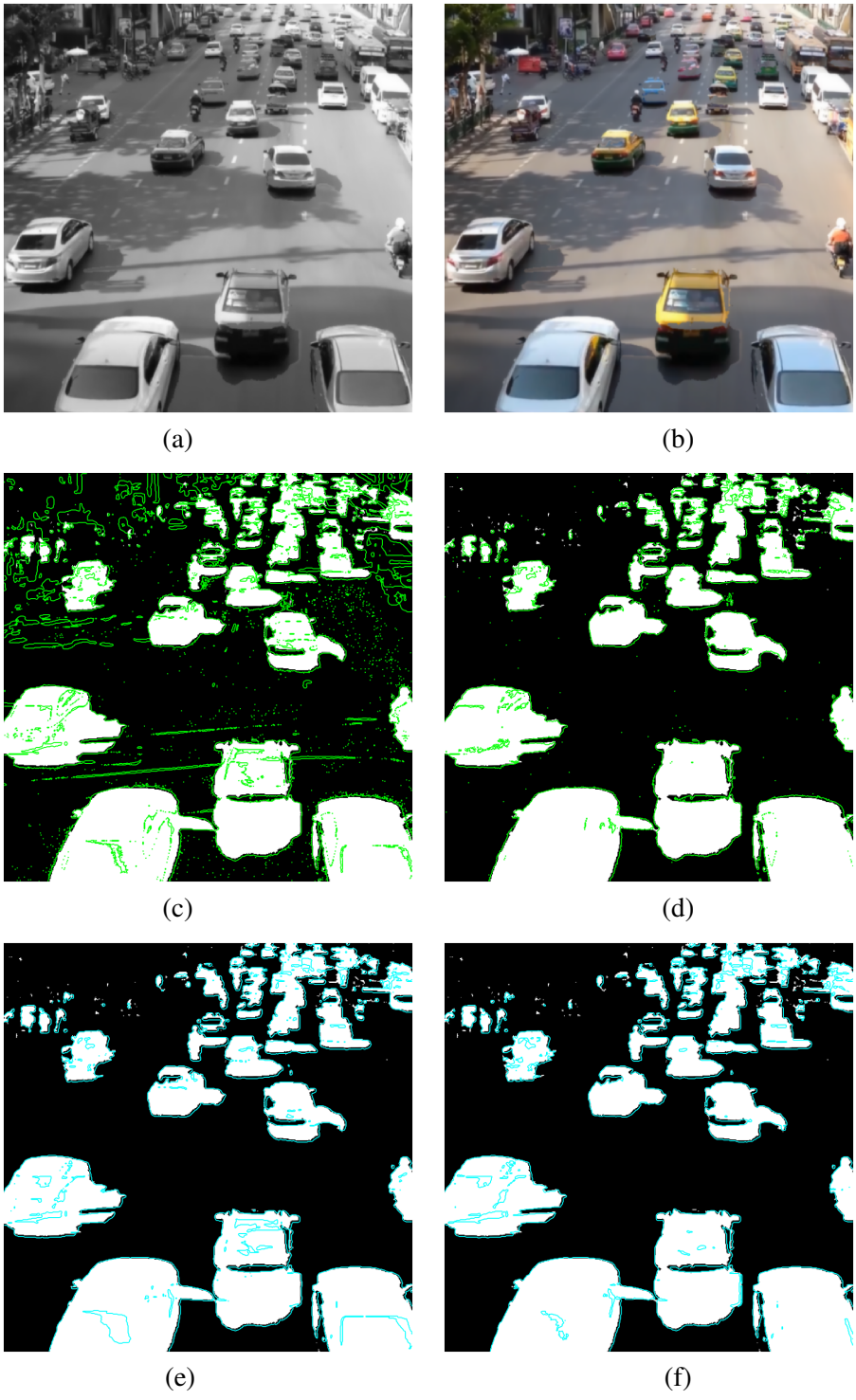


Figure 6.14: Sample grayscale (a) and color (b) images of turbulence level 1 frame from the Thailand03 datasetfor the TV-GMM (c), MB-TV-GMM (d), WS-GM (e), and MB-WS-GM (f) motion classifers.

Let us now examine the individual detection masks of Figure 6.14. The left and right columns represent the input image and corresponding output detection masks for grayscale and color, respectively. The top row represents the input images, the middle row is detection masks from TV-GMM-type classifiers, and the bottom row contains detection masks from WS-GM-type classifiers. Note that all grayscale results presented in this section are a repeat from Section 6.1 and are repeated here only for side-by-side comparison with the new multiband results.

We saw from Figure 6.13(a) that MB-WS-GM and WS-GM were effectively identical, and we see that same behavior in Figures 6.14(e) and 6.14(f). MB-WS-GM still has tight contours around its detections. Some of the false negatives that are interior to vehicles in the bottom-right of the image have been attenuated, but is offset by some additional false negatives on vehicles in the bottom-left. Overall, WS-GM and MB-WS-GM generate very similar detections masks at this low turbulence level.

Conversely, MB-TV-GMM shows a substantial improvement over TV-GMM. Contours around moving objects is still form-fitting, and nearly all the false detections have been suppressed. Additionally, false negatives interior to vehicles have been drastically reduced. The MB-TV-GMM detection mask is comparable to both WS-GM and MB-WS-GM detection masks. Based on visual inspection of the number and size of false negative patches, MB-TV-GMM is arguably the best performer at this turbulence level.

Figure 6.15 shows the detection masks for turbulence level 4. As expected, multiband imagery does improve detection masks, most notably in the reduction of false alarms. MB-WS-GM is now able to correctly reject many of the shadows on the road. Although MB-TV-GMM is not as able to reject all the shadows, it does break large patches of false detections into many smaller groupings, providing an overall reduction in false alarms. MB-WS-GM successfully rejects many of those same false alarms at edges of shadows completely as a result of its patch-based approach.

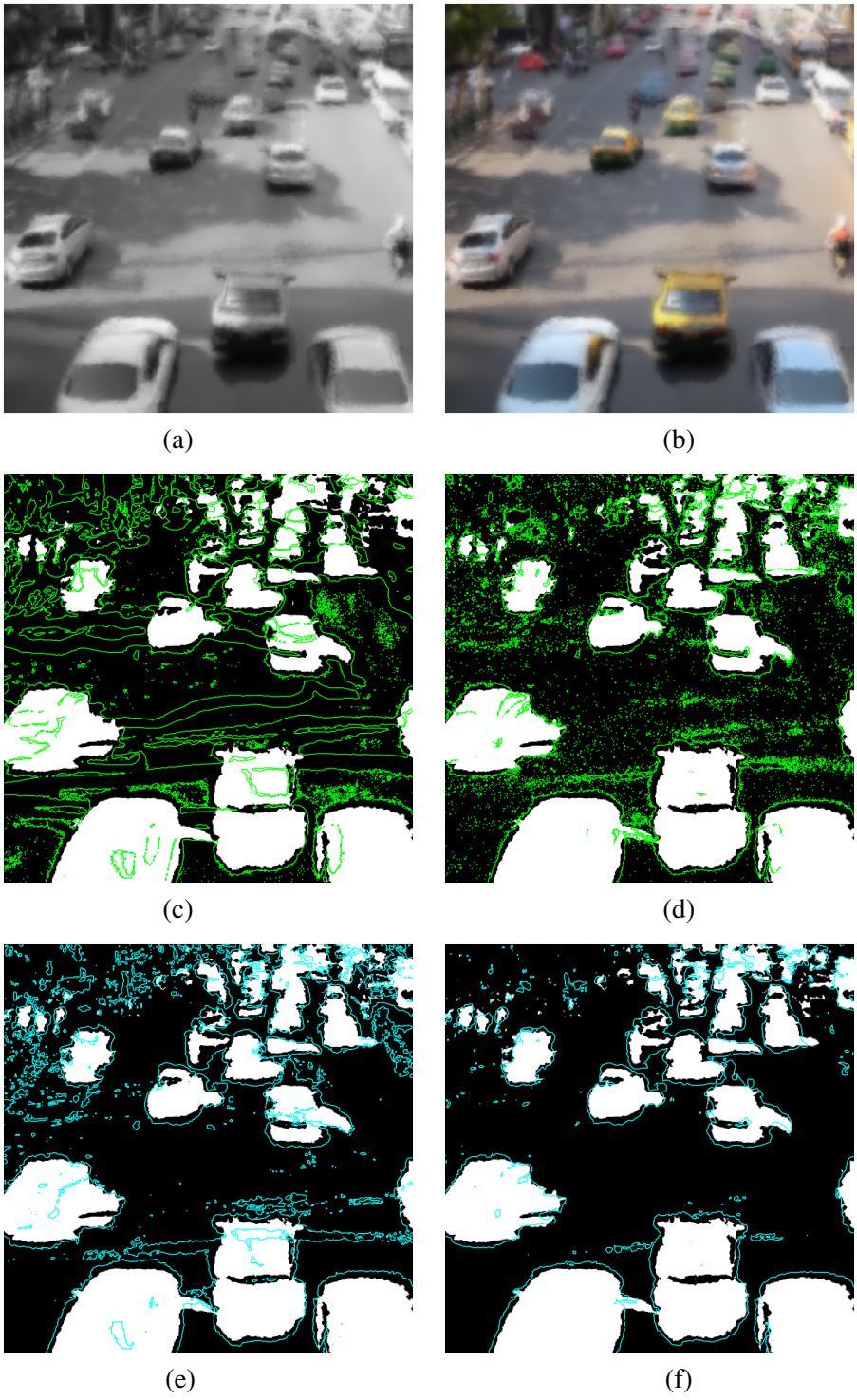


Figure 6.15: Sample grayscale (a) and color (b) images of turbulence level 4 frame from the Thailand03 datasetfor the TV-GMM (c), MB-TV-GMM (d), WS-GM (e), and MB-WS-GM (f) motion classifers.

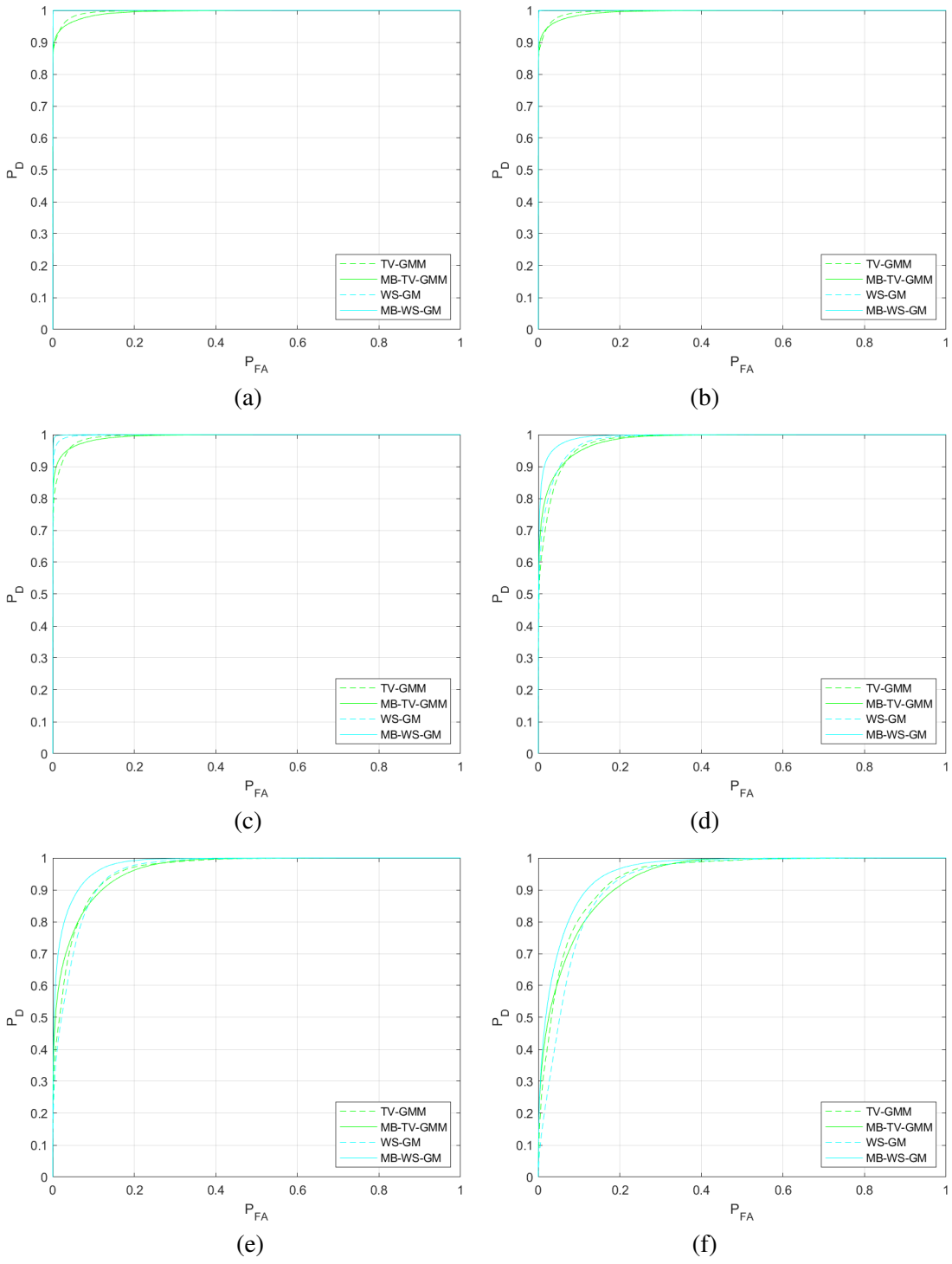


Figure 6.16: Multiband ROC curves for Intersection05 for (a)-(f) corresponding to turbulence levels 1-6, respectively.

As previously mentioned, it is assumed that adding additional color bands will result in superior classifier performance, and the results of Thailand03 color imagery support that statement. An example where that assumption is violated is the Intersection05 dataset whose ROC curves are shown in Figure 6.16. As with Thailand03, MB-WS-GM does not distinguish itself from WS-GM in lighter turbulence levels, and then has a growing separation as turbulence strength increases. Conversely, although MB-TV-GMM is slightly better at lower P_{FA} levels, TV-GMM becomes superior at all other P_{FA} levels, including for the $P_D = 0.95$ used in this dissertation.

Figure 6.17 shows detection masks for a Z-tilt of 4 pixels. While MB-TV-GMM does correctly identify the four movers in the scene, it has a substantial increase in the number of false alarms compared to TV-GMM. This is explained by looking at the differences between the grayscale and color input images. There are two main types of scene content. The first is the folliage which is predominantly various tones of green. The other main type of scene content is manmade objects such as vehicles, roads/sidewalks, and vehicles; these objects are visually very similar between the grayscale and color images, meaning that they are primarily various tones of gray in the real world. MB-TV-GMM has 3 times the number of variables to solve, but the grayscale and color input imagery provide similar information. Therefore, MB-TV-GMM's background model is, in a sense, the result of an undersolved system of equations. This illustrates the importance of band selection with respect to anticipated scene content (both background and movers). Note that while this dissertation shows the potential benefits of exploiting multiband imagery, the selection and analysis of various band combination is not considered here.

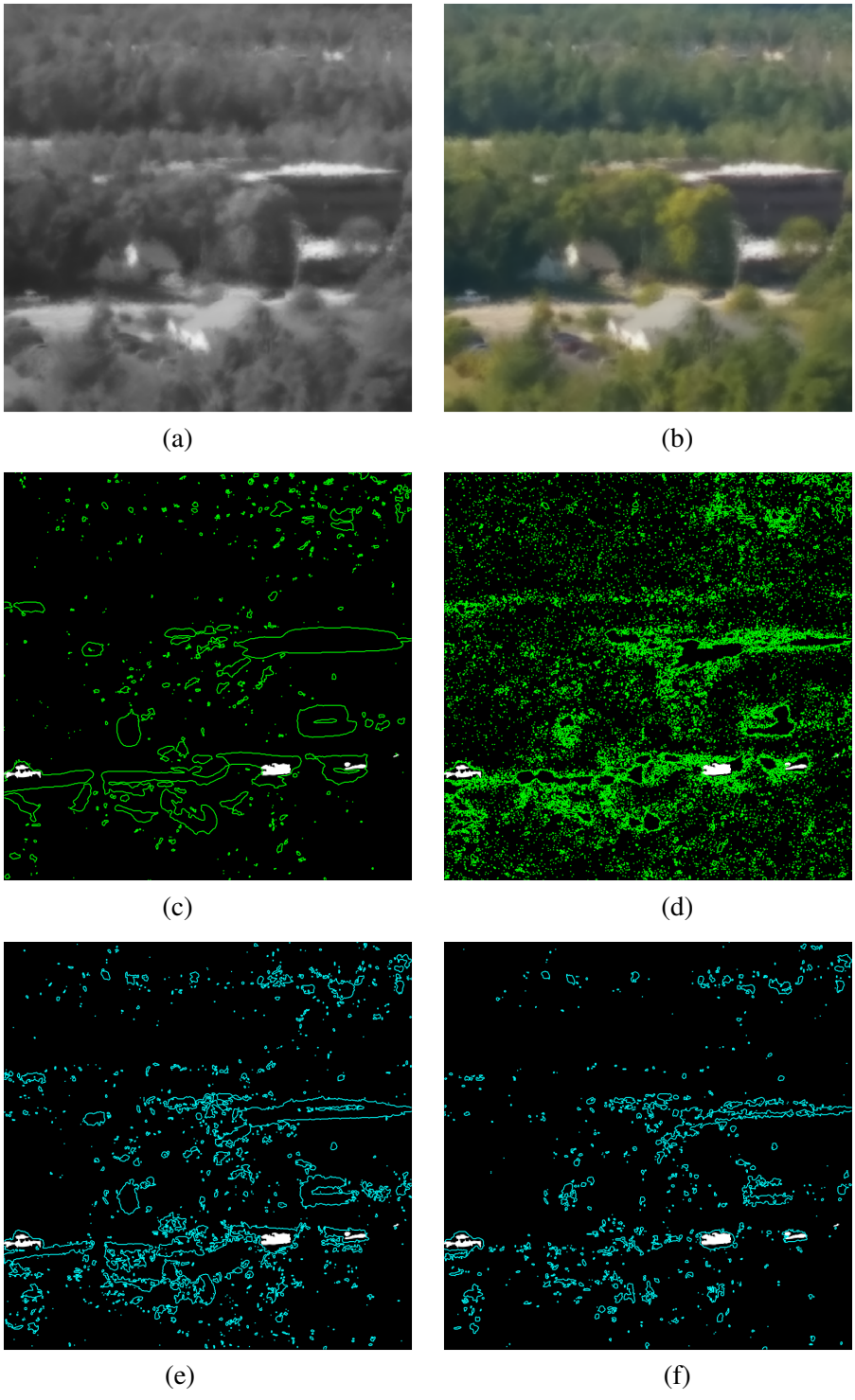


Figure 6.17: Sample grayscale (a) and color (b) images of turbulence level 4 frame from the Intersection05 datasetfor the TV-GMM (c), MB-TV-GMM (d), WS-GM (e), and MB-WS-GM (f) motion classifers.

CHAPTER VII

CONCLUSIONS AND FUTURE WORK

For long range imaging applications, acquired images may be degraded by anisoplanatic atmospheric optical turbulence, including spatially- and temporally-varying blurring as well as geometric distortions. Such effects make many image processing tasks much more difficult. An example of this is sharpening an image via multi-frame superresolution techniques where movers that exist in an image location for only few frames are suppressed in the processed output frame. Such restoration algorithms can be modified to operate only on pixels that are background, but atmospheric turbulence makes differentiating between true in-scene motion and turbulence-driven apparent motion difficult.

There have been a number of techniques developed to make such distinctions by using motion fields, pixel intensity fluctuations, low-rank matrix decomposition. Regardless of approach, however, these techniques are all based on using the input imagery which contain the background and all moving objects, both of which are affected by turbulence. In this dissertation, I proposed two techniques that model the scene background that model pixel intensity fluctuations while avoiding any issues of contamination from moving objects.

The first technique *Tilt Variance Gaussian Mixture Model* (TV-GMM) takes the median of the input images to create a single prototype frame of the scene background. Each pixel within an $M \times M$ patch centered on a center pixel is treated as the mean of one of M^2 modes in a univariate

Gaussian mixture model. The weight of each mode is determined by a Gaussian weighting term which reflects the zero-mean displacement of the geometric warping from turbulence. Modes with identical means are optionally combined to reduce computational burden. The variance of the GMM is the variance of the pixel intensity within the patch plus the estimate of the noise variance.

The second technique *Warping Simulator Gaussian Mode* (WS-GM) forms the same prototype as in TV-GMM and then uses it as an input frame to a warping simulator which applies geometric distortions following the statistics of [3, 41]. The output of the warping simulator is a non-contaminated image stack whose J frames contain only the background scene but still exhibit the warping of the original images. Each pixel location within an $M \times M$ patch of the non-contaminated image stack is treated as a separate variable in a multivariate Gaussian Model. The gaussian's covariance is the covariance of the $M \times M$ patch for all J frames with added variance on the diagonal to account for image noise and other error.

The TV-GMM and WS-GM motion classifiers were compared against three other motion classifers: *Lee*'s online, per-pixel GMM; the offline, per-pixel adaptive GMM *SS-4* classifier, and the low-rank matrix decomposition *3WD*. Three synthetic datasets were generated, each with the same 6 levels of turbulence. The per-pixel truth enabled quantitative analysis, including ROC curves across all turbulence levels. Additionally, one real-world dataset with hand truth was used to evaluate classifier performance. A qualitative analysis was performed based on detection masks generated from each motion classifier.

With grayscale imagery, TV-GMM generally had better detection masks than Lee, SS-4, and 3WD. TV-GMM also had better ROC curves at low false alarm rates, but fell behind first SS-4 and Lee at higher false alarm rates. Given that most applications necessitate low false alarm rates, it is unlikely they would operate in the regimes where SS-4 and Lee were superior. WS-GM was superior to all other techniques discussed in this dissertation. In low turbulence levels, it provides

high detection rates with very few false alarms and is more robust to turbulence affects than the other classifiers. This advantage is likely due to WS-GM's multivariate approach where the spatial orientation of neighboring pixels is maintained.

Both TV-GMM and WS-GM were extended to work on imagery with multiple color bands, resulting in the *Multiband TV-GMM* (MB-TV-GMM) and *Multiband WS-GM* (MB-WS-GM) motion classifiers. Additional color bands typically provide additional information about each pixel, thus allowing MB-TV-GMM and MB-WS-GM to better distinguish the background from moving objects. The end result is a reduction in false alarms for a given probability of detection P_D. As expected, performance gains due to color imagery are not realized if there is substantial redundancy in the color bands such as with the Intersection05 dataset. Although MB-WS-GM had similar performance to the grayscale WS-GM, MB-TV-GMM was substantially worse than TV-GMM, which is attributed to the background model being undersolved.

The analysis in this dissertation indicate several likely areas of future investigation. To begin, let's review the WS-GM motion classifier's covariance. It was defined in Equation 4.2 as

$$\mathbf{\Sigma}(\mathbf{n}) = \hat{\mathbf{\Sigma}}(\mathbf{n}) + \sigma_\eta^2 \mathbf{I}.$$

Recall that $\hat{\mathbf{\Sigma}}(\mathbf{n})$ was the covariance of the $M \times M$ patch $\bar{\mathbf{g}}(\mathbf{n})$ and that σ_η^2 is an estimate of the noise present in $\tilde{f}_k(\mathbf{n})$ that is not present $\tilde{g}_j(\mathbf{n})$. When the noise is not known, σ_η^2 can be treated as a tuning parameter. For the synthetic datasets used in this dissertation, σ_η^2 was taken as three times the noise injected into the input imagery (1 digital unit), and the results were favorable. However, a different covariance model may be more appropriate. Recall that $\tilde{g}_j(\mathbf{n})$ is the theoretical background of $\tilde{f}_k(\mathbf{n})$ with all true moving objects removed. Therefore, it is expected that the covariance of $\tilde{g}_j(\mathbf{n})$ should be approximately equal to the covariance of $\tilde{f}_k(\mathbf{n})$. The use of image registration prior to the WS-GM estimator suppresses some of the covariance, which is then partially recovered with σ_η^2. It

is suspected that the appropriate covariance can be recovered via a model such as

$$\mathbf{\Sigma}(\mathbf{n}) = \beta\hat{\mathbf{\Sigma}}(\mathbf{n}) + \sigma_\eta^2\mathbf{I} \tag{7.1}$$

where β is the ratio of variance in $\tilde{f}_k(\mathbf{n})$ to the variance in $\tilde{g}_j(\mathbf{n})$. σ_η^2 would likely still need to account for other sources of error such as registration.

Another area of potential research is multiscale motion classifiers. The WS-GM motion classifier outperforms the other four classifiers examined in this dissertation, and it is the only classifier that maintains a spatial orientation of neighboring pixel. To better determine if this is the primary factor in WS-GM's performance, it would be beneficial to compare against another classifier that also considered the spatial orientation of neighboring ppixels. Although the baseline TV-GMM does have a spatial element to it, it is a per-pixel detector and can be susceptible to single-pixel outlier intensity values caused by noise. To suppress false detections from single-pixel outliers, the baseline TV-GMM algorithm could be extended to support patches of various sizes, creating the *Multiscale TV-GMM* (MS-TV-GMM). To begin, I examine how the TV-GMM likelihood $z_k(\mathbf{n})$ is calculated. The TV-GMM NLL of Equation 3.8 is repeated here for convenience.

$$z_k(\mathbf{n}) = -\ln\left(p_{\mathbf{n}}\left(\tilde{f}_k(\mathbf{n})\right)\right)$$

Rather than using only a single pixel's likelihood, let us consider a superpixel of size $S \times S$ centered on pixel $\mathbf{n}$ from frame k. I define in lexicographical notation samples in this window as

$$\tilde{\mathbf{f}}_k(\mathbf{n}) = \left[\tilde{f}_{k,1}(\mathbf{n}), \tilde{f}_{k,2}(\mathbf{n}), ..., \tilde{f}_{k,S^2}(\mathbf{n})\right]^\mathsf{T}. \tag{7.2}$$

To determine the score of pixel $\mathbf{n}$ in frame k, we take the product of the likelihoods of all samples within $\tilde{\mathbf{f}}_k(\mathbf{n})$

$$z_k(\mathbf{n}) = -\ln\left(\prod_{s=1}^{S^2} p_{\mathbf{n}}(\tilde{f}_{k,s}(\mathbf{n}))\right). \tag{7.3}$$

Each individual pixel's probability is in the range of zero (i.e., 0% likelihood of being part of the GMM) and one (i.e., a perfect fit with the GMM). Therefore, the product of these likelihoods is also

in the range of [0,1]. Given an $S \times S$ superpixel, as S increases, the impact of any individual pixel is attenuated, providing some robustness against single-pixel outliers. For both WS-GM and MS-TV-GMM, detecting movers smaller than the scale is unlikely. As such, the scale of the classifier would be adjusted based on the expected size of movers. Additionally, pixels along the edge of moving objects are also unlikely to be detected without additional consideration.

Lastly, the classification algorithms presented in this research were developed in order to support turbulence mitigation. It was demonstrated in [49] that turbulence mitigation techniques can be tuned on a per-pixel basis where image restoration can be more aggressive on background pixels and less agressive on movers. The proposed TV-GMM and WS-GM (and their multiband versions) provide just such a metric. As such, future work should validate the ability of these motion classifers to enhance image restoration techniques.

APPENDIX A

Acronyms

Table A.1: List of acronyms

Acronym	Stands for
AIC	Akaike Information Criterion
AUC	Area Under the Curve
GMM	Gaussian mixture model
MB-WS-GM	Multiband WS-GM
MB-TV-GMM	Multiband TV-GMM
MS-WS-GM	Multiscale WS-GM
MS-TV-GMM	Multiscale TV-GMM
NLL	Negative Log Likelihood
PDF	Probability Density Function
PSF	Point Spread Function
ROC curve	Receiver Operating Characteristic curve
RMS	Root Mean Square
SURF	Speeded-Up Robust Features
TM	Turbulence Mitigation
TV-GMM	Tilt Variance-based Gaussian Mixture Model
WS-GM	Warping Simulator Gaussian Model
Z-tilt	Zernike tilt
2D	2-dimensional
3WD	Three Way Decomposition

APPENDIX B

Symbols

Table B.1: List of symbols

Symbol	Meaning
B	Number of bands for each pixel $\mathbf{n}$
C_n^2	Refractive index structure parameter ($\mathrm{m}^{-2/3}$)
d_k	Detection mask for frame k
D	Aperture size (m)
$f/\#$	F-number
f_k	Input image k
$\tilde{f}_k$	Registered image k
$\tilde{\mathbf{f}}_k$	$M^2 \times 1$ observation vector containing samples from $\tilde{f}_k$
$\bar{f}$	Prototype image
g_j	Output frame j from warping simulator
$\tilde{g}_j$	$M \times M$ superpixel of g_j centered on pixel location $\mathbf{n}$
$\bar{g}$	The mean of $\tilde{g}_j$ for all J frames
$\bar{\mathbf{g}}$	Samples in $\bar{g}$
$h_{k,\mathbf{n}}[\cdot]$	Spatially and temporally varying blur for frame k, pixel $\mathbf{n}$
$\mathbf{I}$	Identity matrix
j and J	Warping simulator output frame numbers for $j = 1, 2, 3, ..., J$

Continued on next page

Table B.1 – *Continued from previous page*

Symbol	Meaning
k and K	Input frame $k = 1, 2, 3, ..., K$
k_0	Wavenumber $k_0 = \frac{2\pi}{\lambda}$ (1/m)
l	Focal length (m)
L	Total distance between camera and object (m)
M	Superpixel/window size (pixels)
$\mathbf{n}$	Pixel index $\mathbf{n} = [n_1, n_2]^\mathsf{T}$
$\mathcal{N}(\cdot)$	Normal random vector function
n_1 and n_2	Pixel location $n_1 = 1, 2, 3, ..., N_1$; $n_2 = 1, 2, 3, ..., N_2$
N_1 and N_2	Image dimensions (pixels)
$O(\cdot)$	"Big O" notation for computational runtime analysis
p	Pixel spacing (m) - may or may not be Nyquist
q_k	Ideal image k
R	Correlation coefficients of $\tilde{q}$
r_0	Fried parameter (m)
$s_{k,\mathbf{n}}[\cdot]$	Spatially and temporally varying warping for frame k, pixel $\mathbf{n}$
S	In MS-TV-GMM and MS-MV-GM, the size of the $S \times S$ superpixel
T	Threshold for NLL scoring
w_m	TV-GMM weighting for mode m
W	TV-GMM weight normalization term
z	Incremental distance between camera and object (m)
z_k	Negative log likelihood score for frame k
$\tilde{\boldsymbol{\alpha}}_T$	$\tilde{\boldsymbol{\alpha}}_T = [\tilde{\alpha}_{T_x}, \tilde{\alpha}_{T_y}]^\mathsf{T}$
$\tilde{\alpha}_{T_x}$ and $\tilde{\alpha}_{T_y}$	Apparent tilt vector (radians)
δ	Nyquist pixel spacing (m)
η_k	Additive noise term for frame k
λ	Wavelength (m)
ρ_c	Cut-off frequency (cycles/mm)
$\tilde{\sigma}_T^2$	Z-tilt variance (rad^2)

Continued on next page

Table B.1 – *Continued from previous page*

Symbol	Meaning
$\sigma^2_{R_x}$	Residual tilt variance in horizontal dimension (pixels^2)
$\sigma^2_{R_y}$	Residual tilt variance in vertical dimension (pixels^2)
σ^2_η	Estimated noise
$\mathbf{\Sigma}(\mathbf{n})$	Covariance of local patch with added noise σ^2_η
$\hat{\mathbf{\Sigma}}(\mathbf{n})$	Covariance of image patch
Σ_R	Covariance of the residual tilt variance

BIBLIOGRAPHY

[1] R. C. Hardie, M. A. Rucci, S. Bose-Pillae, and R. L. Van Hook, "Application of tilt correlation statistics to anisoplanatic optical turbulence modeling and mitigation," *submitted to OSA Applied Optics*, December 2021.

[2] A. Schwartzman, M. Alterman, R. Zamir, and Y. Y. Schechner, "Turbulence-induced 2d correlated image distortion," in *2017 IEEE International Conference on Computational Photography (ICCP)*, 2017, pp. 1–13.

[3] R. C. Hardie, J. D. Power, D. A. LeMaster, D. R. Droege, S. Gladysz, and S. Bose-Pillai, "Simulation of anisoplanatic imaging through optical turbulence using numerical wave propagation with new validation analysis," *Optical Engineering*, vol. 56, no. 7, p. 071502, 2017.

[4] M. Tahtali, D. Fraser, and A. Lambert, "Restoration of non-uniformly warped images using a typical frame as prototype," in *TENCON 2005-2005 IEEE Region 10 Conference*. IEEE, 2005, pp. 1–6.

[5] M. Tahtali, A. Lambert, and D. Fraser, "Progressive restoration of non-uniformly warped images by shiftmap prediction using kalman filter," in *Signal Recovery and Synthesis*. Optical Society of America, 2007, p. SMC5.

[6] D. Li, R. M. Mersereau, and S. Simske, "Atmospheric turbulence-degraded image restoration using principal components analysis," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 3, pp. 340–344, 2007.

[7] M. Tahtali, A. Lambert, and D. Fraser, "Self-tuning kalman filter estimation of atmospheric warp," in *Image Reconstruction from Incomplete Data V*, vol. 7076. International Society for Optics and Photonics, 2008, p. 70760F.

[8] C. S. Huebner and M. Greco, "Blind deconvolution algorithms for the restoration of atmospherically degraded imagery: a comparative analysis," in *Optics in Atmospheric Propagation and Adaptive Systems XI*, vol. 7108. International Society for Optics and Photonics, 2008, p. 71080M.

[9] C. S. Huebner, "Compensating image degradation due to atmospheric turbulence in anisoplanatic conditions," in *Mobile Multimedia/Image Processing, Security, and Applications 2009*, vol. 7351. International Society for Optics and Photonics, 2009, p. 735106.

[10] X. Zhu and P. Milanfar, "Image reconstruction from videos distorted by atmospheric turbulence," in *Visual Information Processing and Communication*, vol. 7543. International Society for Optics and Photonics, 2010, p. 75430S.

[11] W. W. Arrasmith and S. F. Sullivan, "Novel wavelength diversity technique for high-speed atmospheric turbulence compensation," in *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IX*, vol. 7666. International Society for Optics and Photonics, 2010, p. 76661R.

[12] Y. Mao and J. Gilles, "Non rigid geometric distortions correction-application to atmospheric turbulence stabilization," *Inverse Problems & Imaging*, vol. 6, no. 3, p. 531, 2012.

[13] R. C. Hardie, M. A. Rucci, A. J. Dapore, and B. K. Karch, "Block matching and Wiener filtering approach to optical turbulence mitigation and its application to simulated and real imagery with quantitative error analysis," *Optical Engineering*, vol. 56, no. 7, pp. 1 – 16, 2017. [Online]. Available: https://doi.org/10.1117/1.OE.56.7.071503

[14] D. H. Frakes, J. W. Monaco, and M. J. Smith, "Suppression of atmospheric turbulence in video using an adaptive control grid interpolation approach," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No. 01CH37221)*, vol. 3. IEEE, 2001, pp. 1881–1884.

[15] S. Gepshtein, A. Shtainman, B. Fishbain, and L. P. Yaroslavsky, "Restoration of atmospheric turbulent video containing real motion using rank filtering and elastic image registration," in *2004 12th European Signal Processing Conference*. IEEE, 2004, pp. 477–480.

[16] H. He and L. P. Kondi, "A super-resolution technique with motion estimation considering atmospheric turbulence," in *Passive Millimeter-Wave Imaging Technology VIII*, vol. 5789. International Society for Optics and Photonics, 2005, pp. 135–144.

[17] B. Fishbain, L. P. Yaroslavsky, and I. A. Ideses, "Real-time stabilization of long range observation system turbulent video," *Journal of Real-Time Image Processing*, vol. 2, no. 1, pp. 11–22, 2007.

[18] C. S. Huebner, "Software-based turbulence mitigation of short exposure image data with motion detection and background segmentation," in *Optics in Atmospheric Propagation and Adaptive Systems XIV*, vol. 8178. International Society for Optics and Photonics, 2011, p. 81780K.

[19] ——, "Turbulence mitigation of short exposure image data using motion detection and background segmentation," in *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXIII*, vol. 8355. International Society for Optics and Photonics, 2012, p. 83550I.

[20] E. Chen, O. Haik, and Y. Yitzhaky, "Classification of moving objects in atmospherically degraded video," *Optical Engineering*, vol. 51, no. 10, p. 101710, 2012.

[21] ——, "Detecting and tracking moving objects in long-distance imaging through turbulent medium," *Applied optics*, vol. 53, no. 6, pp. 1181–1190, 2014.

[22] M. Piccardi, "Background subtraction techniques: a review," in *2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583)*, vol. 4. IEEE, 2004, pp. 3099–3104.

[23] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, vol. 2, 1999, pp. 246–252 Vol. 2.

[24] P. KaewTraKulPong and R. Bowden, "An improved adaptive background mixture model for real-time tracking with shadow detection," in *Video-based surveillance systems*. Springer, 2002, pp. 135–144.

[25] P. W. Power and J. A. Schoonees, "Understanding background mixture models for foreground segmentation," in *Proceedings image and vision computing New Zealand*, vol. 2002, 2002, pp. 10–11.

[26] D.-S. Lee, "Effective gaussian mixture learning for video background subtraction," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, no. 5, pp. 827–832, 2005.

[27] A. S. Deshmukh, S. S. Medasani, and G. Reddy, "Moving object detection from images distorted by atmospheric turbulence," in *2013 International Conference on Intelligent Systems and Signal Processing (ISSP)*. IEEE, 2013, pp. 122–127.

[28] R. Van Hook and R. Hardie, "Scene motion detection in the presence of atmospheric optical turbulence with performance anslysis using numerical wave simulation data and ground truth," in *9th International Symposium on Optronics in Defence and Security*. Association Aéronautique et Astronautique de France, 2020.

[29] R. L. Van Hook and R. C. Hardie, "Patch-based gaussian mixture model for scene motion detection in the presence of atmospheric optical turbulence," in *Automatic Target Recognition XXX*, vol. 11394. International Society for Optics and Photonics, 2020, p. 1139414.

[30] L. Fei, H. Wen-lu, and Z. Ze, "Underwater moving object detection based on codebook model," in *2013 International Conference on Computational and Information Sciences*. IEEE, 2013, pp. 1319–1322.

[31] K. K. Halder, M. Tahtali, and S. G. Anavatti, "Turbulence mitigation and moving object detection for underwater imaging," in *2015 International Conference on Optical Instruments and Technology: Optoelectronic Imaging and Processing Technology*, vol. 9622. International Society for Optics and Photonics, 2015, p. 96220C.

[32] P. E. Robinson and A. L. Nel, "Foreground segmentation in atmospheric turbulence degraded video sequences to aid in background stabilization," *Journal of Electronic Imaging*, vol. 25, no. 6, p. 063010, 2016.

[33] M. T. E. Elahi and K. K. Halder, "Detecting moving objects from long-range atmospheric turbulence degraded videos," in *2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT)*. IEEE, 2018, pp. 260–263.

[34] T. A. Underwood, J. Stufflebeam, D. Soules, M. Kircher, M. Roberts, J. Lohn-Jaramillo, G. Knox, and J. Shankle, "Exploring mitigation of image blur due to atmospheric turbulence by utilizing multiple sensors to achieve optical path diversity," in *Long-Range Imaging III*, vol. 10650. International Society for Optics and Photonics, 2018, p. 1065005.

[35] J. P. Bos, "Imaging theory and mitigation in extreme anisoplanatism," in *Technologies for Optical Countermeasures XV*, vol. 10797. International Society for Optics and Photonics, 2018, p. 107970B.

[36] O. Oreifej, X. Li, and M. Shah, "Simultaneous video stabilization and moving object detection in turbulence," *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 2, pp. 450–462, 2012.

[37] J. D. Schmidt, *Numerical Simulation of Optical Wave Propagation with Examples in MATLAB*, ser. SPIE Press Monograph. SPIE Press, August 2010.

[38] M. C. Roggemann and B. M. Welsh, *Imaging through Turbulence*, ser. Laser and Optical Science and Technology. CRC Press, January 1996.

[39] R. K. . Tyson, *Adaptive Optics Engineering Handbook*. CRC Press, August 1999.

[40] F. Molina-Martel, R. Baena-Gallé, and S. Gladysz, "Fast PSF estimation under anisoplanatic conditions," in *Optics in Atmospheric Propagation and Adaptive Systems XVIII*, K. U. Stein and J. D. Gonglewski, Eds., vol. 9641, International Society for Optics and Photonics. SPIE, 2015, pp. 141 – 151. [Online]. Available: https://doi.org/10.1117/12.2194570

[41] S. Basu, J. E. McCrae, and S. T. Fiorino, "Estimation of the path-averaged atmospheric refractive index structure constant from time-lapse imagery," *Proc. SPIE*, vol. 9465, pp. 94 650T–94 650T–9, 2015. [Online]. Available: http://dx.doi.org/10.1117/12.2177330

[42] J. E. McCrae, S. Basu, and S. T. Fiorino, "Estimation of atmospheric parameters from time-lapse imagery," *Proc. SPIE*, vol. 9833, pp. 983 303–983 303–8, 2016. [Online]. Available: http://dx.doi.org/10.1117/12.2223986

[43] S. R. Bose-Pillai, J. E. McCrae, C. A. Rice, R. A. Wood, C. E. Murphy, and S. T. Fiorino, "Estimation of atmospheric turbulence using differential motion of extended features in time-lapse imagery," *Optical Engineering*, vol. 57, no. 10, pp. 1 – 12, 2018. [Online]. Available: https://doi.org/10.1117/1.OE.57.10.104108

[44] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.

[45] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European conference on computer vision*. Springer, 2006, pp. 404–417.

[46] R. L. Van Hook and R. C. Hardie, "Patch-based gaussian mixture model for scene motion detection in the presence of atmospheric optical turbulence," in *Automatic Target Recognition XXX*, vol. 11394. International Society for Optics and Photonics, 2020, p. 1139414.

[47] H. Akaike, "A new look at the statistical model identification," *IEEE transactions on automatic control*, vol. 19, no. 6, pp. 716–723, 1974.

[48] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861 – 874, 2006, rOC Analysis in Pattern Recognition. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S016786550500303X

[49] D. R. Droege, R. C. Hardie, B. S. Allen, A. J. Dapore, and J. C. Blevins, "A real-time atmospheric turbulence mitigation and super-resolution solution for infrared imaging systems," in *Infrared Imaging Systems: Design, Analysis, Modeling, and Testing XXIII*, G. C. Holst and K. A. Krapels, Eds., vol. 8355, International Society for Optics and Photonics. SPIE, 2012, pp. 234 – 250. [Online]. Available: https://doi.org/10.1117/12.920323