SPECTRAL DETECTION OF HUMAN SKIN IN VIS-SWIR HYPERSPECTRAL IMAGERY WITHOUT RADIOMETRIC CALIBRATION

THESIS

Andrew P. Beisley, 2nd Lieutenant, USAF

AFIT/GE/ENG/12-03

AFIT/GE/ENG/12-03

SPECTRAL DETECTION OF HUMAN SKIN IN VIS-SWIR HYPERSPECTRAL

IMAGERY WITHOUT RADIOMETRIC CALIBRATION

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Electrical Engineering

Andrew P. Beisley, B.S.E.E.

2nd Lieutenant, USAF

March 2012

AFIT/GE/ENG/12-03

SPECTRAL DETECTION OF HUMAN SKIN IN VIS-SWIR HYPERSPECTRAL

IMAGERY WITHOUT RADIOMETRIC CALIBRATION

Andrew P. Beisley, B.S.E.E.
2nd Lieutenant, USAF

Approved:

<table>
<tr><td>//signed//</td><td>March 2012</td></tr>
<tr><td>―――――――――――――</td><td>―――――――――――――</td></tr>
<tr><td>Lt Col Jeffrey D. Clark, PhD
Chairman</td><td>Date</td></tr>
<tr><td>//signed//</td><td>March 2012</td></tr>
<tr><td>―――――――――――――</td><td>―――――――――――――</td></tr>
<tr><td>Dr. Michael A. Marciniak
Member</td><td>Date</td></tr>
<tr><td>//signed//</td><td>March 2012</td></tr>
<tr><td>―――――――――――――</td><td>―――――――――――――</td></tr>
<tr><td>Dr. Richard K. Martin
Member</td><td>Date</td></tr>
</table>

AFIT/GE/ENG/12-03

Abstract

Many spectral detection algorithms require precise ground truth measurements that are hand-selected in the image to apply radiometric calibration, converting image pixels into estimated reflectance vectors. That process is impractical for mobile, real-time hyperspectral target detection systems, which cannot empirically derive a pixel-to-reflectance relationship from objects in the image. Implementing automatic target recognition on high-speed snapshot hyperspectral cameras requires the ability to spectrally detect targets without performing radiometric calibration.

This thesis demonstrates human skin detection on hyperspectral data collected at a high frame rate without using calibration panels, even as the illumination in the scene changes. Compared to an established skin detection method that requires calibration panels, the illumination-invariant methods in this thesis achieve nearly as good detection performance in sunny scenes and superior detection performance in cloudy scenes.

The method in this thesis defines a pixel-to-reflectance relationship as an illumination transform acting on a reflectance vector. The illumination transform is estimated from a series of multi-environmental atmospheric radiative transfer function simulations. Applied to dismount detection, this transform produces an estimated skin signature that is used by three different hyperspectral detection algorithms. These algorithms consistently achieve false alarm rates below 1% while detecting 80% of skin pixels in a variety of illumination conditions. As the scene illumination changes from sunny to cloudy in a sequence of test images, these algorithms maintain a high average skin pixel detection rate with an AUC of over 95%.

iv

Acknowledgments

I would like to thank my advisor, Lt Col Jeffrey Clark, for the enormous amount of help he has given me in composing this thesis and submitting a conference paper. He never fails to impress me with his dedication to helping his students succeed, and his enthusiasm for our area of research is contagious.

I also want to thank my former advisor, Maj Michael Mendenhall, for setting me on a successful path. After introducing me to this area of research, he often let our meetings go longer than expected so I could discuss new concepts and develop my ideas.

Finally, I owe a lot of thanks to my wife, whose cheery attitude and incredible work ethic have given me invaluable encouragement. For the past year, she has been a continuing source of inspiration to me.

<div align="right">Andrew P. Beisley</div>

Table of Contents

List of Figures

List of Tables

List of Symbols

Symbol   Page

Symbol   Page

List of Abbreviations

Abbreviation		Page

SPECTRAL DETECTION OF HUMAN SKIN IN VIS-SWIR HYPERSPECTRAL

IMAGERY WITHOUT RADIOMETRIC CALIBRATION

1. Introduction

Automatic target recognition (ATR) is a key enabling concept in the defense community, where smart sensors are deployed to detect threats. Such threats could include vehicles, obstacles, or humans. Air crews and ground troops use infrared and night-vision cameras to detect threats in low-light conditions. More recently, computers are using these cameras to identify and track potential threats in imagery and video. Using computers for image-based threat detection requires an understanding of the characteristics that make these threats unique. For example, detecting dismounts can be aided by focusing on the unique characteristics of human skin.

Previous research has yielded a method of detecting human skin in hyperspectral imagery that requires only four spectral bands instead of hundreds [37]. That research led to the design of a prototype multispectral camera with optical filters tuned to view those four spectral bands [40]. This study advances hyperspectral dismount detection towards becoming a deployable capability by evaluating its performance under unknown outdoor illumination conditions.

1.1 Problem Statement

Snapshot hyperspectral image (HSI) cameras have recently become available on the market [55]. They capture visual information about a scene in two spatial dimensions, across multiple spectral channels at a fast frame rate. Such data can provide valuable information if it is processed effectively and efficiently. In particular, the

snapshot HSI camera has advantages as a new ATR sensor, as it captures high resolution information in spatial, spectral, and temporal dimensions. Threats can be identified not only by their size, shape, and spectral signature, but also by their motion across a scene.

Scanning HSI cameras take several seconds to capture each datacube, an image with spatial height and width and spectral depth. Material identification algorithms developed for these datacubes have the luxury of taking several seconds to process the data and identify objects by their spectral signature before the next datacube arrives from the camera. Snapshot HSI cameras produce datacubes at a faster rate, decreasing the available computation time to perform detection between frames. Snapshot HSI cameras require computationally fast detection algorithms to fully utilize their fast frame rates.

Dismount detection algorithms search a datacube for the spectral signature of human skin. The spectral reflectance of human skin has been thoroughly studied [37], and hyperspectral skin detection algorithms have been proposed [6, 38]. A pair of calibration panels is required in each scene to convert pixels into estimated reflectance vectors before comparing them to the known spectral reflectance of human skin. Calibration accounts for factors like atmospheric moisture and the angle of the sun in the sky.

This calibration step makes this skin detection algorithm inappropriate for use with snapshot HSI cameras for three reasons. First, the reference objects required for calibration may not appear in every scene, especially if the camera is moving. The second reason is that the pixels of the reference objects must be manually selected for each datacube, adding considerable processing time. The third reason is the processing time involved in calibrating each pixel, which contributes to the overall computational time allotment.

A real-time ATR sensor like the snapshot HSI camera needs a skin detection algorithm that does not rely on reference objects and minimizes the processing time between capturing a datacube and the detection of human skin. It should be invariant to different atmospheric and illumination conditions. This will save time by not requiring calibration of every pixel in the scene.

1.2 Justification

Using a specialized skin detection algorithm with a snapshot HSI camera offers several benefits. These algorithms are accurate, robust in different conditions, and usable in real-time.

Skin detection algorithms combined with HSI cameras are more accurate than the leading skin detection methods for color cameras [37]. Pixels in a color camera represent light in a red-green-blue (RGB) triad. Pinkish-brown objects are commonly misclassified as skin in RGB pixels. In contrast, HSI cameras sample the electromagnetic spectrum across many channels, picking up features that distinguish skin from other objects with greater accuracy than RGB cameras.

An effective skin detection method must be robust in different atmospheric and illumination conditions. It must also detect skin without including calibration panels in each scene. This skin detection method must produce accurate results in minimal processing time.

Real-time threat detection relies on fast, efficient, and accurate human skin detection. An ideal skin detection technique will minimize the number of computations between an observed image and a classified image. This necessitates the ability to detect skin in the raw image data instead of calibrated, preprocessed data.

1.3 Assumptions

Using a skin detection camera for threat detection applications assumes that human skin is an indicator of a human target. Conditions may arise in which a human target has no skin exposed. For example, in very cold weather conditions a human target may be covered in protective winter clothing. Since a skin detection camera cannot see though objects, it will only detect exposed skin.

Although threat detection could be applied to indoor settings, the skin detection camera is designed for natural illumination from the sun. Sunlight is far easier to model than artificial light sources. Numerous software models can generate sunlight profiles for different atmospheric conditions. Inconsistencies in artificial light sources, like street lamps or vehicle headlights, are too variable across all possible scenarios. The sun is a common light source that can be reliably modeled.

The need for sunlight illumination limits the skin detection camera to daytime use. While it may seem trivial to detect humans in daytime imagery with the naked eye, computers do not have as finely tuned target recognition as people; however, with advanced sensors, automated dismount detection is becoming a reality. Skin detection gives the computer a starting point from which to search for a human threat. Future applications may adapt the skin detection camera for use at night in order to enhance night vision devices.

1.4 Standards

Illumination-invariant skin detection methods will be evaluated on their accuracy across different scenarios.

A common metric for detection accuracy is the comparison of the correct classification rate to the incorrect classification rate. Correctly classifying a pixel as skin is called a detection, and the probability of detection is given by the number of correctly

identified skin pixels out of the total number of skin pixels in the image. When a background pixel, e.g. one that is not skin, is incorrectly classified as a skin pixel, it is called a false alarm. The probability of false alarm is the number of background pixels labeled as skin out of the total number of background pixels.

1.5 Approach

In order to understand the appearance of human skin in HSI cameras, the information represented in a hyperspectral pixel must be understood. First, a physical model for sensor-reaching radiance is needed. Such a model describes the light arriving at the camera from a scene. Next, the spectral characteristics of human skin are examined. Sensor-reaching radiance that originates from human skin has a unique spectral distribution. It is affected by the spectral reflectance of skin and the spectral distribution of the light source.

After characterizing the spectral reflectance of skin, the scene illumination must be understood. The scope of this study is limited to outdoor, natural, daytime illumination; therefore, the light being reflected off human skin originates from the sun. Sunlight under different conditions is simulated by a radiative transfer computer program. From these simulations, a common illumination estimate is derived that produces an estimated skin spectral radiance signature when combined with the known spectral reflectance of skin.

The estimated skin spectral radiance signature is then applied to three hyperspectral detection algorithms: the matched filter, the linear discriminant, and the adaptive coherence estimator. Data is collected with a HyperSpecTIR3 camera from a scene that includes a human subject with exposed skin under natural sunlight. The data includes a sunny scene and a cloudy scene. All three skin detection algorithms are applied to the images and the results are compared.

These algorithms maintain false alarm rates below 1% while detecting 80% of skin pixels across all test images. As the scene illumination changes from sunny to cloudy, these algorithms maintain a high average skin pixel detection rate with an AUC of over 95%.

1.6 Materials and Equipment

Collecting a human skin spectral reflectance measurement will require a spectrometer with a reflectance measurement contact probe. An Analytical Spectral Devices (ASD) FieldSpec® Pro spectrometer is used for this research. A radiative transfer code will be needed to generate a large and diverse set of atmospheric and illumination simulations. MODTRAN®5, a radiative transfer code often referenced in the literature, is used for this research. Data processing software is also necessary for this study. For this, MATLAB® is used for analyzing data simulations. Finally, the HyperSpecTIR®3 (HST3) camera is used for collecting data to evaluate the skin detection techniques.

2. Background

This chapter reviews works of literature that explain important concepts in detecting dismounts. A brief overview of dismount detection methods in Section 2.1 provides an understanding of spectral detection of human skin. A physical model is presented in Section 2.2 that describes the optical path of light entering a camera system to create an image. This is followed by general explanation of how cameras convert light into digital images in Section 2.3. General atmospheric correction techniques and their importance are detailed in Section 2.4.

A few important hyperspectral detection algorithms are introduced in Section 2.5, including the matched filter, the linear discriminant, the normalized difference index, and the adaptive coherence estimator. This chapter concludes with a discussion on human skin detection methods in RGB images and HSI datacubes in Section 2.6.

2.1 Dismount Detection

A dismount is a human who is on the ground and not in a vehicle [19]. Dismount detection and tracking has a variety of military and civilian applications. Military uses include crowd monitoring, weapon targeting, search and rescue, and collateral damage avoidance. Civilian uses include industrial safety, security systems, improved computer interfaces, and biometric identification. A variety of sensors are used in dismount detection: RGB cameras, HSI cameras, thermal infrared sensors, and synthetic aperture radar (SAR). Dismount detection may employ temporal, spatial, or spectral detection techniques [29].

Temporal techniques rely on the unique motion of human bodies. In SAR, moving dismounts are detected by their distinctive gait, breathing, and limb movement, each of which reflects a measurable Doppler signature [19, 57]. Video cameras can also

detect moving dismounts by the distinct periodicity of walking and running [36].

Spatial techniques for dismount detection often compare the observed scene to a template of the expected shape of the dismount. This can be accomplished with a histogram of oriented gradients (HOG), or searching for human-shaped geometric figures [13]. To reduce the search space and computational time, spatial dismount detectors are often cued to search in areas that consist of human-colored pixels, which is called spectral detection.

Spectral dismount detection techniques exploit the known spectral signature of dismounts. Camera based dismount detectors differentiate spectral signatures between pixels and mark areas of interest for spatial and temporal techniques. Depending on sensor-to-target range, the dismount may not occupy enough pixels to be detected spatially; however, the spectral detection technique is still a viable detection method.

2.2 Sensor-Reaching Radiance

In order to develop a spectral detector, it is important to understand the nature of imaging spectroscopy. The remote sensing community is concerned with images of the ground taken from an airborne or space-borne HSI sensor. In this context, light traveling from the Sun to the Earth passes through the Earth's atmosphere, where it is transmitted, scattered, and absorbed at different rates depending on the wavelength. Light that reaches through the atmosphere strikes the surface of an object, where it is again absorbed and reflected at different rates for different wavelengths depending on the material properties. Reflected light from the surface of the object travels back up through the atmosphere, encountering additional scattering and absorption before it is collected by the aperture of the sensor. This light energy at the sensor is called apparent radiance or sensor-reaching radiance (SRR).

SRR is composed of light rays reflected from the scene, as well as scattered from
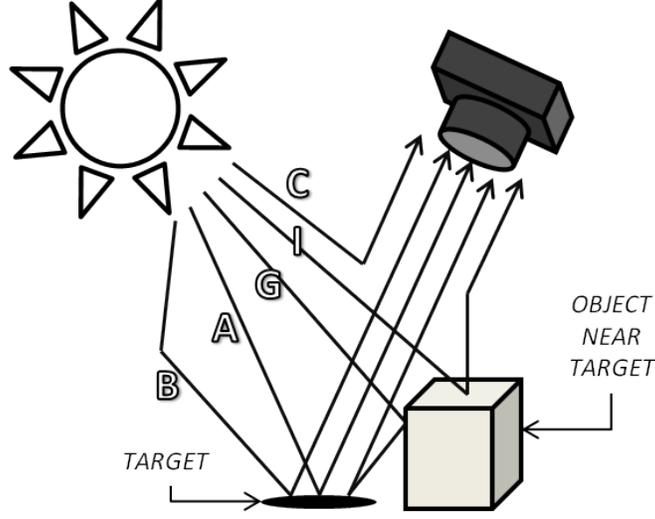
Figure 1. Sensor-reaching radiance in the visible to near-infrared spectrum is composed of reflected direct sunlight (L_A), scattered sunlight (L_B), multiple-bounce reflected light from the background (L_G), the adjacency effect (L_I), and path radiance (L_C).

the air molecules in front of the sensor. The latter is called path radiance, and is dependent on the atmosphere and illumination geometry. Path radiance and other factors that comprise SRR are illustrated in Fig. 1. The governing equation for SRR in the visible to infrared spectrum can be expressed as [17, 46]:

$$L = L_A + L_B + L_C + L_G + L_I, \tag{1}$$

where L represents the total SRR, L_A is the direct sunlight reflected off the target, L_B is sunlight scattered by the atmosphere and reflected off the target, L_C is the path radiance, L_G is a multiple-bounce of light from the background to the target to the sensor, and L_I is scattered light reflected off the adjacent background. Although each of these terms are a function of wavelength λ, the functional notation has been abbreviated in Eq. (1) for compactness.

The last term in Eq. (1), L_I, accounts for the adjacency effect, where light is reflected from an object near the target and is scattered into the sensor-target path by

the atmosphere, making it appear to come from the target. It is sometimes regarded as negligible, while some authors do include it in their models [1, 33, 34]. For the purpose of this study, the adjacency effect will be ignored.

Multiple-bounce reflected background light, L_G, is difficult to model [17]. Cases where the target is exposed to sunlight L_A or skylight L_B, the reflected background term L_G is typically dominated by L_A and L_B. Thus, for ease of calculation, the reflected background L_G will also be ignored.

Both L_A and L_B incorporate light that has been attenuated by the atmosphere and reflected off the target towards the sensor. For objects in direct light, L_A has a much greater contribution than L_B in Eq. (1). However, for an object in the shade where L_A is occluded, L_A goes to zero and L_B represents all of the reflected light from the object. Both of these terms will be described here.

2.2.1 Direct Reflected Sunlight

Direct reflected light (L_A) at the sensor, as a function of wavelength (λ), can be written as [46]:

$$L_A(\lambda) = E'_S(\lambda) \cos \sigma' \frac{r(\lambda)}{\pi} \tau_1(\lambda) \tau_2(\lambda), \tag{2}$$

where $E'_S(\lambda)$ is the irradiance distribution of the sunlight measured at the top of the Earth's atmosphere (illustrated in Fig. 2), σ' is the angle from the object to the sun relative to the surface normal of the object, $r(\lambda)$ is the spectral reflectance of the object, $\tau_1(\lambda)$ is the atmospheric transmittance along the sun-to-ground path, and $\tau_2(\lambda)$ is the atmospheric transmittance along the ground-to-sensor path. Since Eq. (2) applies to objects facing the sun, the angle σ' is limited to values between 0 and π such that $0 \leq \cos \sigma' \leq 1$.

The spectral reflectance ($r(\lambda)$) is the ratio of reflected to incident light from the surface of an object [46]. The spectral reflectance is a unitless measure ranging
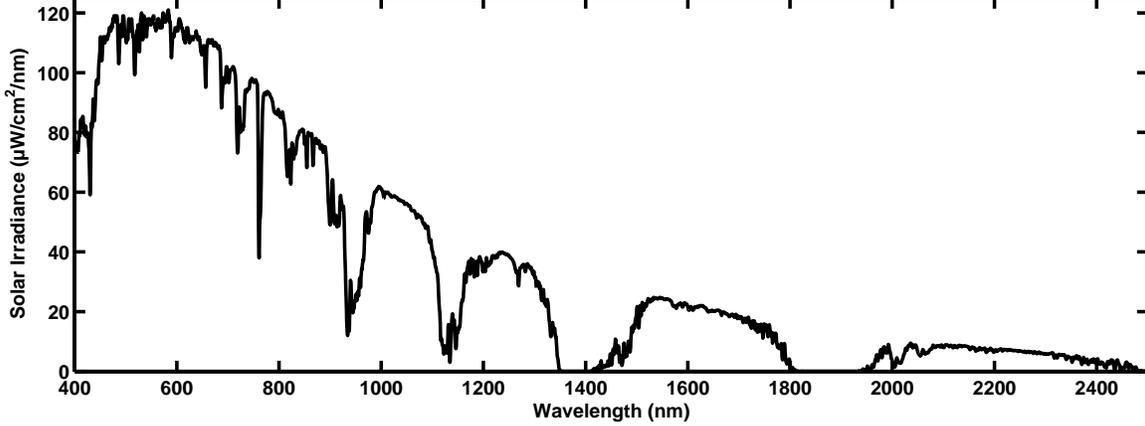
Figure 2. Solar irradiance distribution incident on the Earth's surface ($E'_S(\lambda)\tau_1(\lambda)$ in Eq. (2)) generated in MODTRAN®5, assuming a solar zenith angle of 10 degrees from nadir and a mid-latitude summer standard atmosphere.

from zero to one as a function of wavelength. In general, the radiance of reflected light changes with viewing angle according to the object's bidirectional reflectance distribution function (BRDF) [46]. For objects where viewing angle changes the spectral distribution of $r(\lambda)$ at different wavelengths, the BRDF is an important factor in SRR equations. However, accounting for variability in sun angle, sensor-to-target angle, and surface normal adds significant complexity. For simplicity, this study will assume that the BRDFs of all targets are uniform with respect to scene geometry. That is, light incident on a surface will be reflected diffusely, with equal magnitude in all directions. This is represented in Eq. (2) by the factor $\frac{1}{\pi}$, as reflected light is distributed equally over the entire reflectance hemisphere of π steradians.

2.2.2 Reflected Skylight

Reflected skylight (L_B) at the sensor, as a function of wavelength (λ), can be written as [46]:

$$L_B(\lambda) = FE_{ds}(\lambda)\frac{r(\lambda)}{\pi}\tau_2(\lambda), \tag{3}$$

where F is a form factor ranging from 0 to 1 representing the fraction of the total sky exposed to the object, E_{ds} is the average downwelled skylight in the hemisphere above the object, and $r(\lambda)$ and $\tau_2(\lambda)$ are the same as in Eq. (2). Combining Eq. (1), (2), and (3), and dropping the L_G and L_I terms, results in a simplified model for SRR [33]:

$$L(\lambda) = [E_S'(\lambda)\cos\sigma'\tau_1(\lambda) + FE_{ds}(\lambda)]\frac{1}{\pi}\tau_2(\lambda)r(\lambda) + L_C(\lambda), \tag{4}$$

which can be written compactly [31] as:

$$L(\lambda) = L_R(\lambda)r(\lambda) + L_C(\lambda). \tag{5}$$

The physical model summarized in Eq. (5) describes SRR as a linear combination of reflected sunlight ($L_R(\lambda)$ and path radiance ($L_C(\lambda)$). Information about the objects in the scene is carried by the signal $L_R(\lambda)r(\lambda)$, while the path radiance $L_C(\lambda)$ is independent of the objects in the scene.

2.2.3   Electromagnetic Spectrum

Light that humans can see with the naked eye comprises a segment of the electromagnetic spectrum known as the visible (VIS) spectrum. Light with wavelengths from about $\lambda = 0.39\mu$m to $\lambda = 0.75\mu$m fall within the VIS spectrum. Shorter wavelengths, including ultraviolet (UV) light and X-rays, are invisible to humans. Some cameras respond to UV light, but HSI cameras usually capture longer wavelengths.

Infrared light has longer wavelengths than visible light, and is separated into different sub-regions, shown in Fig. 3. The near infrared (NIR) and short-wave infrared (SWIR) regions are separated by an atmospheric absorption band at 1.4μm caused by water vapor. HSI cameras designed to capture reflected light typically respond to
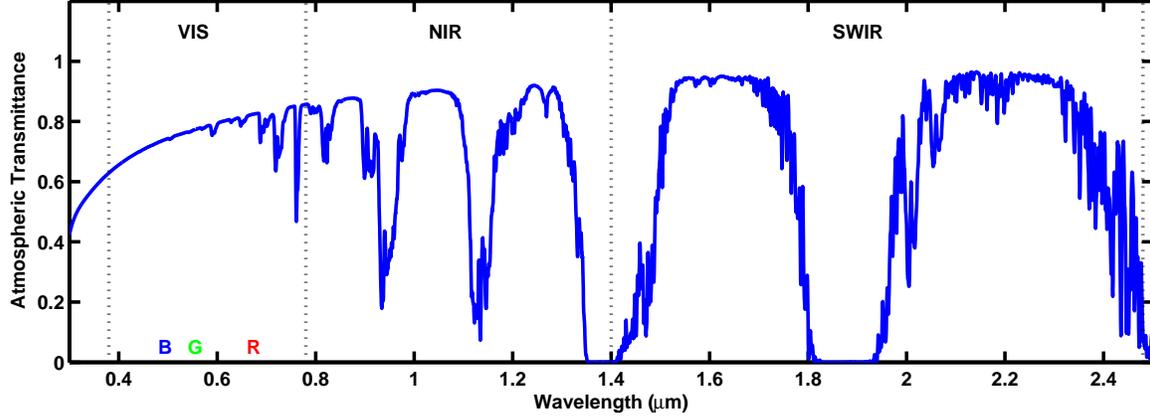
Figure 3. Atmospheric transmittance $(\tau_1(\lambda))$ of the electromagnetic spectrum from $\lambda = 0.3\mu\text{m}$ to $\lambda = 2.5\mu\text{m}$. Transitions between infrared sub-divisions are marked by absorption bands that are due to water vapor, ozone, carbon dioxide, and other atmospheric aerosols.

light in the VIS to SWIR spectrum. Thermal imagers capture infrared light at much longer wavelengths: 3-8μm for mid-wave infrared, and 8-15μm for long-wave infrared.

2.3   Camera Response

When the SRR enters the aperture of the camera, it propagates through a set of optics and filters before striking a focal plane to form an image [46]. Each pixel of an image, $\mathbf{x} = [x_1, ..., x_N]^T$, is comprised of N channels. The value of each channel is proportional to the SRR as processed by a different spectral filter. The equation that describes $\mathbf{x}$ in terms of SRR from Eq. (5) is [46]:

$$x_i = \int_0^{t_{exp}} \int_\lambda L(\lambda)C_i(\lambda)d\lambda dt, \tag{6}$$

where x_i is the digital number value of the ith channel of pixel $\mathbf{x}$, t_{exp} is the exposure time, $L(\lambda)$ is the SRR from Eq. (5), $C_i(\lambda)$ is the spectral response of the ith channel filter, and λ is wavelength. Equation (6) ignores camera noise that may affect the pixel values, such as shot noise or thermal noise.

Table 1. Channel responses of common imaging sensors

Camera	Type	Spectral Range	Channels	Channel FWHM
Common RGB camera	VIS	0.39-0.75 μm	3	$\sim$0.1μm
CAP's ARCHER	VIS-NIR	0.5-1.1 μm	52	12nm
NRL's HYDICE	VIS-SWIR	0.4-2.5 μm	210	10nm
NASA's AVIRIS	VIS-SWIR	0.4-2.5 μm	224	9.5nm
SpecTIR's HST3	VIS-SWIR	0.45-2.45 μm	227	8.2nm - 12nm

Usually the exposure time (t_{exp}) is too short to capture temporal changes in SRR. Since many cameras are able to automatically set the exposure time to prevent saturation, it is assumed that t_{exp} is set such that the sensor operates in its linear region. Thus, t_{exp} can be factored out of the integral and treated as a constant:

$$x_i = t_{exp} \int_{\lambda} L(\lambda)C_i(\lambda)d\lambda. \tag{7}$$

Channel spectral response $C_i(\lambda)$ incorporates the spectral response of the optical filter, the throughput efficiency of the camera optics, and the sensitivity of the film or charge-coupled device (CCD). Each of the N channels has a different spectral response $C_i(\lambda)$; therefore, the pixel **x** is a sampling of different parts of the spectrum of $L(\lambda)$. Hyperspectral cameras often have channel responses that are Gaussian shaped. The spectral bandwidth of a Gaussian filter is called its Full Width Half Maximum (FWHM). For channels with a narrow FWHM, $C_i(\lambda)$ approximates a weighted delta function $\delta(\lambda_i)$, in which case Eq. (7) is reduced to [31]:

$$x_i \approx c_i L(\lambda_i), \tag{8}$$

with the constant c_i replacing the constants for exposure time and filter width, and where λ_i is the center wavelength of channel filter $C_i(\lambda)$. Omitted from Eq. (8) is a quantization noise term that is a consequence of storing the result x_i as a hardware-

defined data type. For cameras with narrow FWHM filters, Eq. (8) provides a convenient linear relationship between SRR ($L(\lambda)$) and pixel values ($\mathbf{x}$). Table 1 lists the channel responses for different types of cameras.

2.4 Atmospheric Correction

The spectral reflectance ($r(\lambda)$) component of SRR carries information about objects in the scene, as described in Section 2.2; therefore, it is the term used to perform spectral identification. The illumination components of SRR, $L_R(\lambda)$ and $L_C(\lambda)$ in Eq. (5), must be factored out to estimate the object's spectral reflectance.

2.4.1 Radiative Transfer Codes

Radiative transfer (RT) codes such as MODTRAN [4], FLAASH [34], HATCH [41], and ATREM [15], take in a set of atmospheric parameters and generate the terms in Eq. (4). These are then used to factor out atmospheric effects in an image and estimate the spectral reflectances of objects.

The input parameters to RT codes are often based on atmospheric measurements at the time the image was taken. Measurements can be taken with meteorological sensors or with precision laser instruments [11]. When direct measurements are not available, the input parameters can be estimates of the past measurements of the atmosphere. Mis-estimating the atmospheric parameters in RT codes can cause diminished performance in spectral detection algorithms [56].

2.4.2 Empirical Line Method

$L_R(\lambda)$ and $L_C(\lambda)$ can be deduced by observing the appearance of objects in the scene that have known spectral reflectances. Instead of requiring a complete atmospheric measurement at the time the image is taken, this method requires a "ground-

truth" spectral reflectance measurement of objects in the scene.

A well-known implementation of the process mentioned above is the Empirical Line Method (ELM) [25, 48]. Adopting the narrowband assumption from Section 2.3, ELM is a combination of Eq. (5) and Eq. (8) that provides a linear mapping from pixel values (x_i) to the estimated spectral reflectance (r'_i).

Applying the SRR model in Eq. (5) to the pixel simplification in Eq. (8) gives,

$$x_i = c_i L_R(\lambda_i) r'_i + c_i L_C(\lambda_i), \tag{9}$$

which is solved for r'_i,

$$r'_i = \frac{1}{c_i L_R(\lambda_i)} x_i + \frac{-L_C(\lambda_i)}{L_R(\lambda_i)}. \tag{10}$$

Equation 10 in linear form is $r'_i = m_i x_i + b_i$, where the slope m_i and offset b_i are functions of reflected and upwelled radiance through the ith channel. By selecting two pixels $(\mathbf{x}_w, \mathbf{x}_g)$ that correspond to white and gray calibration panels with known spectral reflectances $(r_w(\lambda), r_g(\lambda))$, and assuming that the same illumination conditions $L_R(\lambda)$ and $L_C(\lambda)$ apply to all pixels in the scene, the values of m and b are determined for the ith channel by solving the linear equation:

$$\begin{bmatrix} r_w(\lambda_i) \\ r_g(\lambda_i) \end{bmatrix} = \begin{bmatrix} x_{w,i} & 1 \\ x_{g,i} & 1 \end{bmatrix} \begin{bmatrix} m_i \\ b_i \end{bmatrix}. \tag{11}$$

Once m and b are found for all N channels in Eq. (11), each pixel $\mathbf{x}$ in the image can be converted to estimated spectral reflectance by applying the ELM transform:

$$\mathbf{r}' = \mathbf{M}\mathbf{x} + \mathbf{b}, \tag{12}$$

where $\mathbf{r}' = [r'(\lambda_1), \ldots, r'(\lambda_N)]^T$, $\mathbf{x} = [x_1, \ldots, x_N]^T$, $\mathbf{b} = [b_1, \ldots, b_N]^T$, and $\mathbf{M}$ is a

diagonal matrix with entries equal to each slope m_i:

$$\mathbf{M} = \begin{bmatrix} m_1 & 0 & \cdots & 0 \\ 0 & m_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & m_N \end{bmatrix}.$$ (13)

The result of Eq. (12) is only an estimate of the spectral reflectance because of the inaccuracies introduced by the filter bandwidth $C_i(\lambda)$ and imaging constants c_i. Also, the unknown orientation of objects in the scene makes a full BRDF estimate very difficult, so ELM must assume that all objects are diffuse.

2.5 Spectral Detection

The steps involved in spectral detection can be described as an algorithm chain as shown in Fig. 4 [16]. Raw image data, a collection of pixels ($\mathbf{x}$) of a scene, undergoes a series of preprocessing stages prior to the application of the target detection algorithm. These steps may include atmospheric compensation, dimensionality reduction, noise whitening, background characterization, and adaptive filter generation. Each step is intended to improve the accuracy and robustness of the detection algorithm to be applied, often at the cost of added complexity and computation time. In real-time target detection, the algorithm chain must maximize both computational speed and detection accuracy.

Hyperspectral detection algorithm chains may be divided into two groups: those that employ radiance-based detection, and those that employ reflectance-based detection. Reflectance-based detection includes atmospheric compensation as one of its preprocessing steps, producing an estimated reflectance vector ($\mathbf{r}'$) out of each pixel ($\mathbf{x}$) as described in Section 2.4. A target detection algorithm searches these estimated
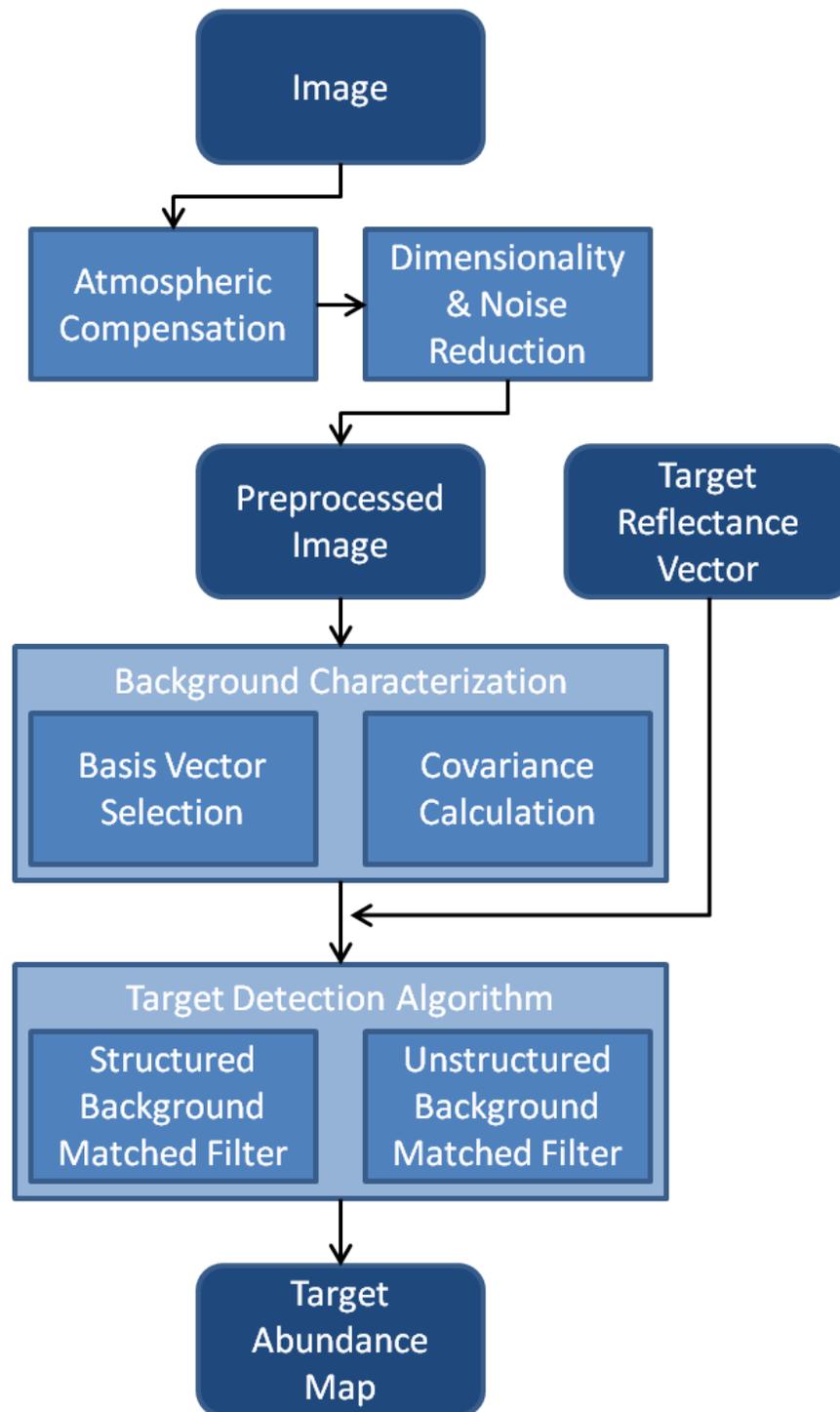
Figure 4. Common steps in the algorithm chain for spectral detection as presented in [16].

reflectance vectors for the known spectral reflectance of the target.

Radiance-based detection skips the atmospheric compensation step and performs other preprocessing steps, such as noise whitening and normalization, prior to applying the detection algorithm. Ientilucci *et al.* have claimed that reflectance-based and radiance-based methods perform comparably well, as long as atmospheric effects are considered at some point in the algorithm chain [21].

Many different detection algorithms exist for hyperspectral data [31]. In general, they compare an observation, $\mathbf{x}$ or $\mathbf{r}'$, to a target spectral signature, expected distribution of signatures, or a set of detection rules, in order to classify the observation. Some detection algorithms do not require prior knowledge of any spectral signatures in the scene. These anomaly detectors highlight regions of interest in the image where the spectrum is statistically distinct from the rest of the background [50]. Other detectors use *a priori* knowledge of specific signatures either taken from a spectral library [8] or measured directly [53].

For target recognition purposes, hyperspectral detection algorithms of interest are those that search for a specific target signature. These detection methods compare each pixel in an image to a known target signature, and label the pixels with a score based on their similarity to the target signature. Comparing the score to a threshold (γ) results in pixel classification. The performance of a binary classifier, which classifies each pixel as either target or non-target, is evaluated by the probability of detection (P_D) and the probability of false alarm (P_{FA}). Setting different thresholds (γ) on the score gives different (P_D, P_{FA}) pairs, and plotting those pairs makes a receiver operating characteristics (ROC) curve.

The following sections describe basic detection methods that will be used in this study.

2.5.1 Linear Discriminant

Fisher's linear discriminant is a classification method for multidimensional data [12]. It is an optimal classifier for a two-class problem set for which the first and second order statistical moments are known [10]. The linear discriminant analysis (LDA) method projects multidimensional data down to a single dimension, along which the two classes are most separated.

The LDA equation for a pixel vector ($\mathbf{x}$) is:

$$\gamma \underset{H_0}{\overset{H_1}{\gtrless}} \mathbf{w}^T \mathbf{x}, \tag{14}$$

which classifies a pixel as either target (positive hypothesis H_1) or background (null hypothesis H_0) by applying a threshold γ to the result of a dot product $\mathbf{w}^T \mathbf{x}$. The vector $\mathbf{w}$ is normal to the hyperplane separating the target and background pixel distributions in multidimensional space. The optimal $\mathbf{w}$ is given by [3]:

$$\mathbf{w} = (\boldsymbol{\Sigma}_b + \boldsymbol{\Sigma}_t)^{-1}(\mu_t - \mu_b), \tag{15}$$

where $\boldsymbol{\Sigma}_b$ is the background covariance matrix, $\boldsymbol{\Sigma}_t$ is the target covariance matrix, μ_b is the mean background pixel, and μ_t is the mean target pixel.

Target statistics $\boldsymbol{\Sigma}_t$ and μ_t can be found by hand-selecting target pixels from a database of images. Background statistics $\boldsymbol{\Sigma}_b$ and μ_b are harder to define unless specific scenes and conditions are chosen. With empirical data, the background statistics can be computed from all of the remaining pixels after target pixels have been removed.

2.5.2 Matched Filter

The matched filter (MF) algorithm is a well-known hyperspectral detection method. Because of its simplicity, it is often used as a baseline for comparison when authors propose more complex detection methods. The MF equation is:

$$d_{MF}(\mathbf{x}) = \frac{\mathbf{s}^T \mathbf{x}}{\|\mathbf{s}\|\|\mathbf{x}\|}, \tag{16}$$

where $d_{MF}(\mathbf{x})$ is the score given to a pixel $(\mathbf{x})$, $\mathbf{s}$ is the target signature, and the $\|\cdot\|$ operation is the l_2 norm of a vector.

The score is maximum at $d_{MF}(\mathbf{x}) = 1$ for a perfect match, and minimum at $d_{MF}(\mathbf{x}) = 0$ for orthogonal vectors. The MF method is invariant to changes in magnitude because it normalizes each pixel, by dividing by the l_2 norm, before taking the dot product with the normalized target signature.

A common shorthand notation for a normalized vector is $\hat{\mathbf{x}} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$. Using this notation, the MF formula is:

$$d_{MF}(\mathbf{x}) = \hat{\mathbf{s}}^T \hat{\mathbf{x}}, \tag{17}$$

where $\hat{\mathbf{x}}$ is a normalized pixel and $\hat{\mathbf{s}}$ is the normalized target signature.

The MF method can be used for both radiance-based and reflectance-based detection. In reflectance-based detection, the target signature is a reflectance vector and the pixels are the estimated reflectance vectors produced by atmospheric correction. In radiance-based detection, the target signature is the expected value of a target pixel taken from the unprocessed HSI datacube.

2.5.3 Adaptive Coherence Estimator

The Adaptive Coherence Estimator (ACE) detector relies on a geometric interpretation of hyperspectral data. It finds the squared cosine of the angle between a

Table 2. Common Normalized Difference Indices and their associated wavelengths

Material	Abbreviation	λ_1 (μm)	λ_2 (μm)
Vegetation	NDVI	0.6-0.7	0.75-1.35
Vegetative Water	NDWI	0.86	1.24
Snow	NDSI	0.555	1.64
Human Skin	NDSI	1.08	1.58

pixel and a target signature [45]. Like the LDA detector, the ACE detector adapts to each image by accounting for the covariance of the image data. The ACE equation is:

$$d_{ACE}(\mathbf{x}) = \frac{|\hat{\mathbf{s}}^T \mathbf{\Sigma}_b^{-1} \hat{\mathbf{x}}|^2}{(\hat{\mathbf{s}}^T \mathbf{\Sigma}_b^{-1} \hat{\mathbf{s}})(\hat{\mathbf{x}}^T \mathbf{\Sigma}_b^{-1} \hat{\mathbf{x}})}, \tag{18}$$

where $d_{ACE}(\mathbf{x})$ is the score given to a pixel $\mathbf{x}$, $\hat{\mathbf{s}}$ is the normalized target signature, $\hat{\mathbf{x}}$ is a normalized pixel, and $\mathbf{\Sigma}_b$ is the background covariance matrix, which can estimated by taking the covariance of all normalized pixels in an image about their mean [28].

2.5.4 Normalized Difference Index

Another well known detection approach is the use of the Normalized Difference Index (NDI). A reflectance-based detection method, the NDI identifies pixels of a hyperspectral image that have specific spectral properties associated with a known material. It has been utilized in remote sensing applications to identify vegetation [43], vegetative water [14], urban build-up [58], snow [44], vegetative moisture [24], and barrenness [59]. NDI applications to human skin detection have been proposed [37, 38], and will be discussed in Section 2.6.

The NDI compares the values of an estimated reflectance vector at two wavelengths. The general form of the NDI equation is:

$$d_{NDI}(\mathbf{r}') = \frac{r'(\lambda_1) - r'(\lambda_2)}{r'(\lambda_1) + r'(\lambda_2)}, \tag{19}$$

where $d_{NDI}(\mathbf{r}')$ is the score between -1 and 1 given to an estimated reflectance vector $\mathbf{r}'$, and the two wavelengths λ_1 and λ_2 reveal features that distinguish a certain material from the rest of the image. Several NDI methods and their wavelengths are listed in Table 2.

2.6 Detecting Human Skin

Detecting the spectral signature of human skin is necessary for ATR systems designed to track dismounts. Skin detection precedes spatial feature detectors that search for the shape of a human [9]. Further analysis can then indicate other relevant features of the dismount including pose, skin type, or clothing type. Ultimately these features provide information to determine if a dismount should be cued for further investigation. The development of some early skin detection methods with RGB cameras are explained in this section, followed by some recent efforts to detect skin with multispectral and hyperspectral cameras.

2.6.1 RGB Skin Detection

Human skin color detection in RGB images has been studied extensively [54]. It has applications that include facial recognition [20], pose recognition [22], content filtering [13], and security systems [60].

RGB pixels have three channels, described in Table 1. The three channels capture red, green, and blue light, similar to the spectral response of the human eye. Skin detection methods rely on the clustering of skin pixels apart from background pixels in the 3-dimensional RGB vector space, shown in Fig. 5. One basic way to detect skin is to observe that skin pixels are more red than green [5], and score pixels based on their red to green ratio. However, this method has a high false-alarm rate because many other objects that are not skin produce a high red to green ratio.
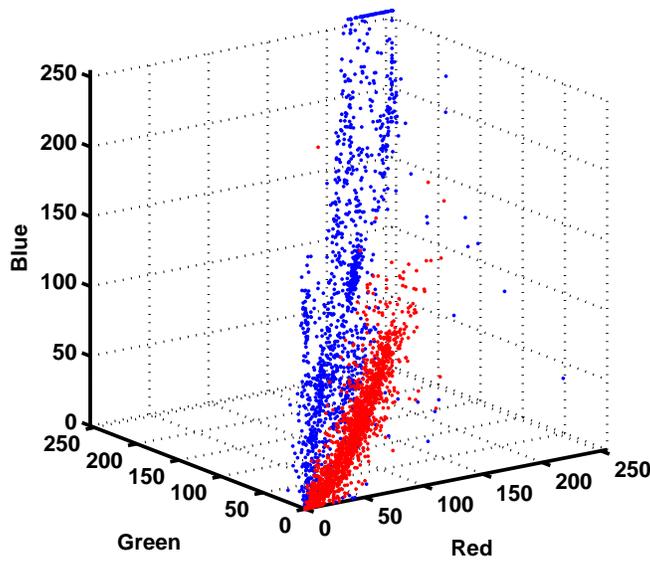
Figure 5. Three dimensional scatter plot of skin (red) and background (blue) pixels in the RGB colorspace. The skin pixels have a higher red component than the background pixels.

A more accurate way to detect skin in RGB images is to estimate the probability distribution functions (pdf) of both skin and background pixels and perform the likelihood ratio test (LRT) of the form,

$$\frac{p(\mathbf{x}|\text{skin})}{p(\mathbf{x}|\text{background})} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma, \tag{20}$$

where $p(\mathbf{x}|\Theta)$ is the conditional pdf of a pixel vector $\mathbf{x}$ given the condition Θ that the pixel is either skin or background. The pdfs can be calculated using a large training set, which leads to an optimal LRT, or estimated as multidimensional Gaussian distributions [20].

Other skin detection methods transform the RGB pixels into colorspaces that have less correlation between components or are less varied by illumination. Some of these colorspaces are normalized RGB [49], Hue Saturation Value (HSV) [47], and the luma-chroma colorspace YC_rC_b [20]. These colorspaces allow for detection rules in

fewer dimensions. However, these colorspaces are rearrangements of the same RGB data, and do not distinguish skin pixels from background pixels any better than the RGB colorspace. In [2] it is shown that the optimal LRT detector performs as well or better in the original RGB colorspace as compared to any other derived colorspace.

The LRT yields better skin detection performance than the red to green ratio, but it still has a high false-alarm rate. Some background objects, like cardboard, flesh-colored mannequins, sand, tan shirts, and khaki pants have a similar color to skin in the RGB colorspace. These objects are often misclassified as skin even with the optimal LRT detector. This suggests that there is inherently an upper limit to the performance of RGB skin detection methods. The limit comes from the fact that RGB cameras collect light in only three channels, and only within the VIS range of the spectrum.

2.6.2   Hyperspectral Skin Detection

To overcome the limits of skin detection in RGB cameras, recent research has investigated the appearance of human skin and background objects in cameras that collect electromagnetic energy outside the range of human vision [27, 37]. While some flesh-colored background objects have the same red, green, and blue components as human skin, they reflect light much differently than skin in the NIR and SWIR ranges of the spectrum. Multispectral and hyperspectral cameras observe electromagnetic energy across many spectral channels, creating multidimensional clusters of skin and background pixels that are more distinct than in RGB imagery.

Figure 6 illustrates the spectral reflectance of human skin with different melanin levels. One paper proposes a Normalized Difference Skin Index (NDSI), a variant of the NDI described in Section 2.5.4, to detect skin in hyperspectral images [37]. The NDSI compares estimated reflectance vectors at wavelengths $\lambda_1 = 1.08\mu$m and
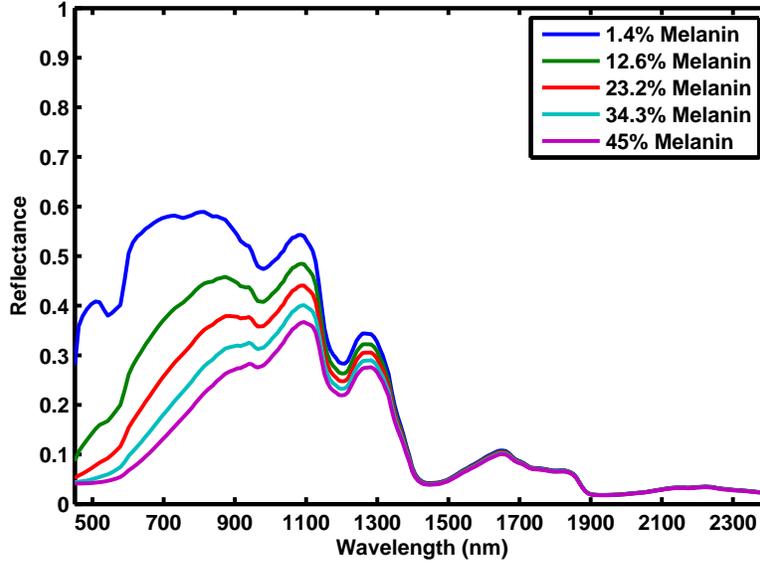
Figure 6. Spectral reflectance of human skin at VIS-SWIR wavelengths. Skin with less melanin appears brighter because it has higher reflectance.

$\lambda_2 = 1.58\mu$m. Pixels containing human skin have high NDSI values. However, the leaves of many plants and trees also have high NDSI values, and can cause false alarms. To mitigate this, the paper suggests false-alarm suppression using the Normalized Difference Green Red Index (NDGRI) with $\lambda_1 = 0.54\mu$m and $\lambda_2 = 0.66\mu$m. Green vegetation scores high on the NDGRI, while skin does not because skin is more red than green as discussed in Section 2.6.1. Plotting the NDSI against the NDGRI, as shown in Fig. 7, reveals the separation between skin and background clusters. Using these metrics for skin detection yields a much lower false alarm rate than the optimal LRT detection method in RGB images.

Hyperspectral dismount detectors have been designed that employ the NDSI with NDGRI false alarm suppression [6, 9]. They search for a human shape by starting at pixels marked as human skin, and are accelerated due to the false-alarm suppression discussed previously. In another paper, the spectral reflectance of human skin in the VIS to SWIR spectrum has been modeled as a function of several biometrics such as melanin level and blood oxygen level [37]. The results have been used to estimate the
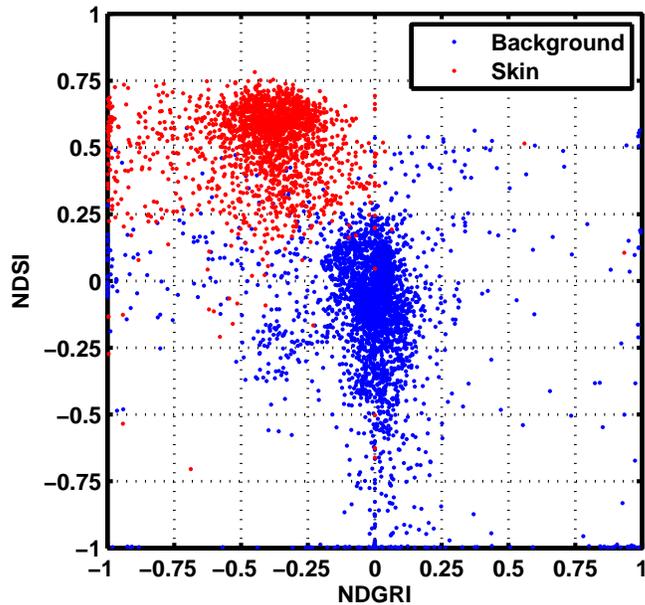
Figure 7. Skin and background pixels are separated in NDSI vs NDGRI space. Skin pixels, shown in red, receive a high NDSI score and a low NDGRI score.

melanin level and skin type of detected dismounts in hyperspectral images [39].

2.7   Summary

Designing a dismount detection camera that identifies the spectral signature of human skin requires an understanding of the underlying physical model. SRR contains information about objects in the scene, but that information must be extracted from the atmospheric and illumination components. HSI cameras convert SRR into pixel vectors ($\mathbf{x}$), which can be converted into estimated reflectance vectors ($\mathbf{r}'$) by applying atmospheric correction. Spectral detection algorithms may include atmospheric correction as one of the preprocessing steps in an algorithm chain. Shortening the algorithm chain leads to a faster detection algorithm, but may degrade performance. The Matched Filter, Linear Discriminant, and Adaptive Coherence Estimator are three detection algorithms that have short algorithm chains.

Human skin detection in RGB images is difficult because many objects have the same color as human skin in the VIS region of the electromagnetic spectrum. Hyperspectral skin detection performs better than RGB because it incorporates spectral features in the NIR and SWIR bands. However, recent hyperspectral skin detection methods can be improved by removing time-consuming calibration steps from the algorithm chain.

3. Methodology

The sensor-reaching radiance (SRR) from an object can vary depending on the type of illumination, the composition of the atmosphere, and the angle at which it is observed. Accounting for these factors can dramatically shorten computation time when performing human skin detection on hyperspectral image (HSI) datacubes. This thesis suggests a process to derive an estimated target signature that provides accurate target detection across multiple illumination scenarios. The process will be applied primarily to skin detection in this thesis.

First, differences between dismount detection and remote sensing are discussed in Section 3.1. Next, the physical model for SRR is adapted from the general remote sensing case to the specific dismount detection case in Section 3.2. Then in Section 3.3, the Matched Filter (MF), Linear Discriminant Analysis (LDA), and Adaptive Coherence Estimation (ACE) detection methods that were introduced in Section 2.5 are each combined with the SRR model to create illumination-invariant dismount detectors. These require an estimate of the scene illumination, which is simulated in MODTRAN. The MODTRAN simulation setup is described in Section 3.4, and the procedure for processing the results is detailed in Section 3.5.

3.1 Hyperspectral Dismount Detection

Hyperspectral dismount detection is a specific application of the more general hyperspectral material identification problem that has been thoroughly studied by the remote sensing community. Many fundamental assumptions that form the basis of hyperspectral detection algorithms in remote sensing scenarios remain pertinent to dismount detection. These assumptions help to shape a model framework that can be extended to the dismount detection problem.

The physics model underlying remote sensing algorithms is introduced in Section 2.2. Each HSI pixel represents SRR that is convolved through a channel filter bank, scaled by a gain factor, and quantized to form a vector of non-negative digital numbers. The SRR itself is a function of the solar irradiance distribution, atmospheric transmission, scene geometry, object spectral reflectance, and scene-to-sensor path radiance. The channel filter bank collects the SRR over hundreds of narrow adjacent spectral bands to provide high-dimensional HSI data, which often has a much lower inherent dimensionality due to correlation between channels. Remote sensing algorithms search for HSI pixels that most likely represent materials with known spectral reflectances under the given illumination conditions. All of these remote sensing concepts can be extended to hyperspectral dismount detection.

Some aspects of dismount detection are different from the general remote sensing model. One difference is that dismount detection is limited by a maximum effective sensor-to-target range. The HSI camera must be close enough for a dismount to occupy multiple pixels in order to apply any sort of spatial recognition method. Imposing the more restrictive condition that at least some dismount skin pixels must be spectrally pure, that an area of exposed skin must take up at least a full pixel, limits the range even further. An estimate of a 10cm^2 pixel covering the exposed face of a dismount puts the maximum effective range for full-pixel skin detection with the HST3 camera at 100 meters, given its instantaneous field of view (IFOV) of 1 millirad [23]. This is a much closer range than the high-altitude scenarios for which the remote sensing algorithms were developed. At this close range, there is relatively little atmosphere between the sensor and the target, so the path radiance (L_C) becomes insignificant. The governing equation for SRR in Eq. (4) reduces to:

$$L(\lambda) = [E'_S(\lambda) \cos \sigma' \tau_1(\lambda) + FE_{ds}(\lambda)] \frac{1}{\pi} r(\lambda), \tag{21}$$

where $E'_S(\lambda)$ is the irradiance distribution from the sun measured at the top of the Earth's atmosphere, σ' is the angle from the object to the sun relative to the surface normal of the object, $\tau_1(\lambda)$ is the atmospheric transmittance along the sun-to-ground path, F is the fraction of the total sky exposed to the object, $E_{ds}(\lambda)$ is the average downwelled skylight in the hemisphere above the object, and $r(\lambda)$ is the spectral reflectance of the object. Equation (21) reduces to $L(\lambda) = L_R(\lambda)r(\lambda)$, the product of the scene illumination and the spectral reflectance of an object. This simplification assumes Lambertian surfaces.

Another distinction between remote sensing and dismount detection is the viewing angle of the sensor. Remote sensing algorithms assume a downward-looking sensor observing the surface of the Earth from a high-altitude platform. A dismount detecting sensor, in contrast, will likely be mounted on a ground vehicle, a structure, or a low-altitude aircraft. The maximum effective altitude will be limited both by the range to target and by the fact that the camera must maintain a low viewing angle. Too high of a viewing angle, i.e. looking straight down, will not only make it hard to discern the humanoid shape of standing dismounts, but will shrink the apparent surface area of the dismount and potentially block the visibility of any exposed skin. Restricting the sensor to low altitudes removes any atmospheric boundary layers from the sensor-target path, like clouds, which can cause large variations in SRR.

One consequence of a low-altitude sensor with a low viewing angle is a large variety of illumination angles. In a remote sensing scenario, the sensor looks down perpendicular to the surface of the Earth and the sunlight strikes from a known angle. Variation in exposure angle (σ') in Eq. (21) is due to surface texture and shadows that occur near tall obstructions like trees or buildings [1]. For a dismount detection sensor at a low viewing angle, there are more possibilities for scene geometry. In one scenario, the sun is behind the sensor, directly illuminating the front of a standing

dismount. In another, the sun illuminates a dismount from either side. The sun could also be behind the dismount, and the only surface visible to the sensor is shadowed and illuminated by skylight. In each of these cases, the sun angle from zenith could range from very low at nadir to very high at early morning or late evening. All cases must be assumed to be equally probable, along with every other possible combination of sun azimuth and zenith angles. Since the sun could illuminate the scene from any position in the hemisphere overhead, there is a wide range of possible illumination angles on a potential dismount target.

Even if the position of the sun relative to the sensor and dismount are known, there are still a wide range of possible exposure angles for all visible surfaces of human skin in the scene. This is because dismounts have surface normals pointing in many different directions. For a direct illumination source like the sun, few surfaces of a dismount are directly illuminated, that is, having surface normals pointing parallel to the illumination path. Many surfaces are angled away from the illumination source, which increases the exposure angle σ'. Accounting for all of these possible dismount surface normals and illumination positions requires a 3D model of dismounts in different positions. This type of data has recently been studied but is outside the scope of this research [30]. Overcast skies may offer better scenarios for dismount detection, as cloudy illumination is more evenly distributed from all directions. However, clouds attenuate much of the SWIR spectrum.

To understand the impact of exposure angle on radiance from an object, consider the example in Fig. 8. In Fig. 8, there are two objects in a scene, a flat rectangle and a sphere, that are directly illuminated by sunlight. Both objects have the same Lambertian spectral reflectance across all wavelengths, and differ only in shape. The flat rectangle will appear to have a constant brightness across its entire surface, while the sphere will be brighter in some areas and dimmer in others. Although they both

3-4
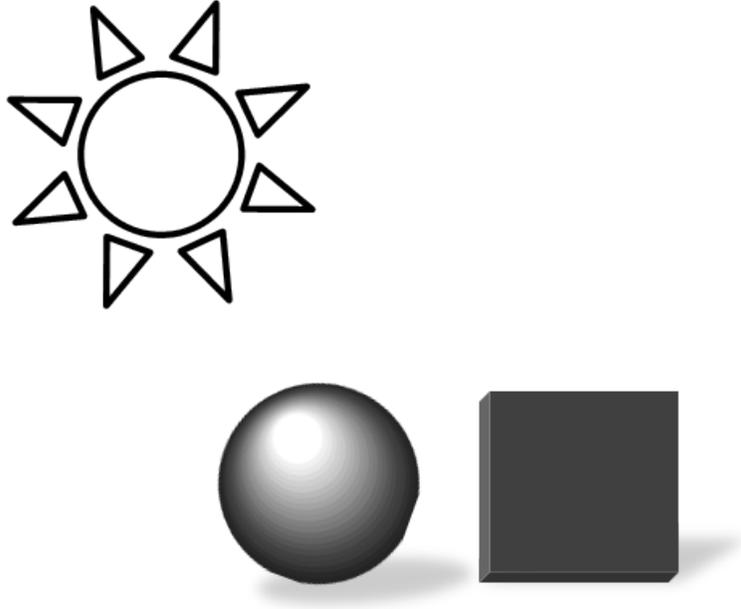
Figure 8. Example illustrating uneven illumination across a scene caused by different surface orientations.

have Lambertian surfaces, they will reflect light at different intensities. This is not because of specularity, but exposure angle σ'. Along the outside edges of the sphere, the exposure angle is close to 90 degrees, bringing the $\cos \sigma'$ term in Eq. (21) close to zero and making those parts of the sphere appear dark. Towards the center of the sphere, the exposure angle is close to zero, like the rectangle, so the two objects have the same brightness.

The preceding example demonstrates another assumption of remote sensing algorithms that cannot be applied to dismount detection. Many remote sensing algorithms search for a known reflectance vector by backing out the atmosphere and illumination from the image pixels. As discussed in Section 2.4, a common method of converting pixels to estimated reflectance vectors is to employ ELM correction. However, ELM assumes a linear relationship between the SRR from an object and that object's inherent spectral reflectance. While that assumption may hold true for remote sensing of the Earth's flat surface, the above example illustrates that it does not apply to

short-range, low-angle dismount detection. A dismount may be illuminated unevenly and have different brightness levels at different points, but that does not mean each point has a different inherent spectral reflectance, as an ELM-corrected HSI datacube would suggest. Thus, a dismount detection method should not search for a known reflectance vector like many remote sensing methods. It should instead search for an estimated target signature that is invariant to the unknown exposure angle.

Uneven illumination also justifies the requirement of a full-pixel on skin. A pixel mixture model used in remote sensing will not work for dismount detection because it could be looking at the side of a face that is blocked from the sun and catch some of the background behind it that is fully exposed to the sun. A mixture model assumes that the spectral reflectances of both objects within the pixel are illuminated equally, which may not be the case.

Other desirable features in dismount detection are to decrease computational complexity and to increase robustness in different environments. Computational complexity is rarely a concern with remote sensing algorithms because they are often applied to individual images during post-flight analysis using plentiful computing resources with no time constraints. In contrast, a dismount detection algorithm should operate quickly on real-time snapshot hyperspectral images using embedded computing resources. The usefulness of a dismount detection sensor for target recognition is directly tied to its speed and agility. It also depends on the variety of environments in which it can be used. A dismount detection sensor should remain functional in situations where there are no ground truth objects to measure illumination.

3.2 Radiance-Based Illumination-Invariant Detection

Understanding the aspects that make dismount detection unique from remote sensing helps derive relevant requirements for a detection method. The new dismount

detection method must be radiance-based and illumination-invariant. The first requirement, radiance-based, is an alternative to reflectance-based detection which is popular in remote sensing. It bypasses much of the algorithm chain in Fig. 4 by applying the target detection algorithm directly to the unprocessed image. The difference is how the target detection rules are determined. In reflectance-based target detection, ELM conversion produces estimated reflectance vectors that are searched to find the known spectral reflectance of the target. The accuracy of the ELM conversion depends on the availability of ground truth objects in the scene that must be manually selected. Conversely, in radiance-based target detection, unprocessed pixels are searched to locate the target signature as it would appear under the given illumination conditions. The detection rule adjusts to the data instead of the data adjusting to the detection rule.

The second requirement, illumination-invariance, means that the dismount detection method should be robust in different environments. Different atmospheres (clear, cloudy, dry, humid, urban or rural) combined with various sun angles (high in the sky, low on the horizon, behind the camera, in front of the camera, or to the side) are some of the various environments that are encountered in dismount detection. Illumination-invariance means that the dismount detector should function accurately with no prior knowledge of these conditions. Instead of adjusting the detection rule to fit the environment, unprocessed pixels are subjected to the same detection rule with minimal impact on detection accuracy. An illumination-invariant, radiance-based detection method is developed in this thesis. This method will enable real-time hyperspectral dismount detection.

3.3 Dismount Detection Model

The pixel model in Section 2.2 can be tailored to the dismount detection problem by incorporating the assumptions detailed previously. Extending the pixel formula from Eq. (8) to a vector form yields:

$$\mathbf{x} = a\mathbf{Cl_{SRR}},\tag{22}$$

where the vector $\mathbf{l_{SRR}} = [L(\lambda_1), \ldots, L(\lambda_N)]^T$ is the SRR sampled at the center wavelengths of the channel filters, the entries of the diagonal transform matrix $\mathbf{C}$ compensate for the width of each channel filter specific to the HSI camera, and the constant term a accounts for the exposure time and gain factor set by the camera. The gain factor (a) is set to minimize the overall signal-to-quantization-noise (SNR_q) power while preventing pixel saturation.

The product $\mathbf{Cl_{SRR}}$ is an approximation of the integral in Eq. (7) and is appropriate for cameras with narrow channel bandwidths. A spectral filter with a narrow bandwidth can be approximated by a weighted delta function, which reduces an integral to a product by the sifting property.

Writing the SRR vector $\mathbf{l_{SRR}}$ in terms of the dismount detection scenario from Eq. (21), with no path radiance term, gives:

$$\mathbf{l_{SRR}} = \mathbf{L_R r},\tag{23}$$

which has a form similar to Eq. (22) in that the reflectance vector $\mathbf{r}$ is the object's spectral reflectance, and the entries of the diagonal transform matrix $\mathbf{L_R}$ are the SRR, both sampled at the center wavelengths of the channel filters. The reflectance vector $\mathbf{r}$ in Eq. (23) is different than the estimated reflectance vector $\mathbf{r}'$ in Eq. (12) because it represents the sampled inherent spectral reflectance of the object and not

the estimated reflectance vector derived from ELM correction. The reflected radiance term $\mathbf{L_R}$ is detailed in Eq. (21). In matrix form, it is a matrix transformation operator that transforms a reflectance vector into a SRR vector.

Equation (23) can be substituted for the SRR vector $\mathbf{l_{SRR}}$ in Eq. (22) to produce

$$\mathbf{x} = a\mathbf{CL_R}\mathbf{r}, \tag{24}$$

which is simplified to

$$\mathbf{x} = \mathbf{Tr}. \tag{25}$$

In this case, $\mathbf{T} = a\mathbf{CL_R}$ can be described as a transformation matrix that transforms the N-dimensional reflectance vector $\mathbf{r}$ into a pixel $\mathbf{x}$. It is a diagonal matrix, and its diagonal entries are the point-wise products of SRR $\mathbf{L_R}$, channel response $\mathbf{C}$, and gain factor a. The entries of the diagonal matrix $\mathbf{T}$ are:

$$T_{ii} = aC_{ii}L_{R,ii}, \tag{26}$$

where T_{ii} is the ith diagonal entry of $\mathbf{T}$, a is the camera gain factor, C_{ii} is the channel width factor of the ith channel, and $L_{R,ii}$ is the SRR intensity at the center wavelength of the ith channel.

Another representation of the diagonal entries of $\mathbf{T}$ is the SRR entering each channel from an object with a constant, unity spectral reflectance at all wavelengths. From Eq. (7), this can be written as:

$$T_{ii} = a \int_{\lambda} L(\lambda)C_i(\lambda)d\lambda \tag{27}$$

where where T_{ii} is the ith diagonal entry of the illumination transform $\mathbf{T}$, a is the camera gain factor (related to exposure time t_{exp} in Eq. (7)), $L(\lambda)$ is the SRR as a

function of wavelength λ, and $C_i(\lambda)$ is the spectral response of the ith channel as a function of wavelength λ.

Expressing the relationship between spectral reflectance and pixel vectors as a matrix transform further justifies the preference for radiance-based detection over reflectance-based detection. To estimate the spectral reflectance of each pixel using ELM correction, the inverse of $\mathbf{T}$ must be applied to each pixel:

$$\mathbf{r}' = \mathbf{T}^{-1}\mathbf{x}, \tag{28}$$

where $\mathbf{r}'$ is the estimated reflectance vector of the object at pixel $\mathbf{x}$. However, the inverse of $\mathbf{T}$ may not always exist. HSI cameras with channels that collect light at wavelengths where the atmospheric transmission $\tau(\lambda)$ is low will have zero entries along the diagonal of $\mathbf{T}$. These zero entries cause the determinant of $\mathbf{T}$ to be zero, making it singular and non-invertible. There will be errors when comparing the estimated reflectance vectors to known target spectral reflectances because of the lack of illumination at some wavelengths. Instead, detection can be performed on the unprocessed pixels $\mathbf{x}$ if the target spectral reflectance undergoes the same illumination transform $\mathbf{T}$ as the pixels in the image to create a target signature $\mathbf{s}$.

3.3.1 Target Signature Generation

The challenge in radiance-based detection is choosing an appropriate target signature for the unknown illumination conditions. For a target with a known spectral reflectance, an estimated target signature can be generated by applying Eq. (25) with an estimate of the illumination transform $\hat{\mathbf{T}}$:

$$\hat{\mathbf{s}} = \frac{\hat{\mathbf{T}}\mathbf{r}}{\|\hat{\mathbf{T}}\mathbf{r}\|}, \tag{29}$$

where $\hat{\mathbf{s}}$ is the normalized target signature, $\hat{\mathbf{T}}$ is the normalized illumination transform, and $\mathbf{r}$ is the target reflectance vector. A target reflectance vector is derived by sampling the measured target spectral reflectance at the center wavelengths of the channels of the hyperspectral sensor. The normalized illumination transform $\hat{\mathbf{T}}$ is a diagonal matrix that is normalized such that its diagonal entries sum to one:

$$\sum_{i=1}^{N} \hat{T}_{ii} = 1. \tag{30}$$

3.3.2   Linear Discriminant Detection

The LDA detector in Eq. (14) can be adapted for unknown illumination. Assuming a normalized target signature ($\hat{\mathbf{s}}$) from Eq. (29), the normalized LDA detector is:

$$d_{LDA}(\mathbf{x}) = \mathbf{w}^T \hat{\mathbf{x}}, \tag{31}$$

where $d_{LDA}(\mathbf{x})$ is the score given to a pixel ($\mathbf{x}$) by the LDA detector, $\mathbf{w}$ is the linear discriminant vector, and $\hat{\mathbf{x}}$ is a normalized pixel. The linear discriminant vector $\mathbf{w}$ can be computed from:

$$\mathbf{w} = \mathbf{\Sigma}_b^{-1}(\hat{\mathbf{s}} - \hat{\boldsymbol{\mu}}_b) \tag{32}$$

where $\mathbf{\Sigma}_b$ is the covariance and $\hat{\boldsymbol{\mu}}_b$ is the mean of the set of all normalized pixel vectors in an image, and $\hat{\mathbf{s}}$ is the normalized target signature. Equation (32) is an adaptation of Eq. 15 using the target signature ($\hat{\mathbf{s}}$) instead of the target class mean ($\hat{\boldsymbol{\mu}}_t$), and ignoring the target class covariance.

Because it incorporates the statistics of the image, this detector adapts to each scene. The background statistics are derived from every pixel in the image under the assumption that a target occupies only a small fraction of the pixels and does not significantly bias the statistics. A similar assumption underlies the anomaly detector

algorithm used by the CAP ARCHER sensor [42, 51].

As discussed in [51], the background covariance ($\boldsymbol{\Sigma}_b$) may not be invertible when using full-spectrum pixels, because hyperspectral data has a low inherent dimensionality. Hyperspectral pixels consist of hundreds of distinct spectral bands, but the responses from many of the bands are highly correlated. In order to have an invertible covariance matrix, the data must be projected onto a lower dimensional subspace using a kernel (e.g. principle component analysis (PCA)).

PCA derives orthogonal dimensions for the normalized hyperspectral pixels by performing eigen decomposition on the covariance matrix $\boldsymbol{\Sigma}_b$ [35]. Principal components are the eigenvectors ranked in descending order by their eigenvalues. If the columns of the $N \times K$ projection matrix $\mathbf{P}$ are the first K principal components, with $N > K$, then the matrix operation,

$$\mathbf{y} = \mathbf{P}^T \hat{\mathbf{x}} \tag{33}$$

represents a $N \times 1$ normalized pixel ($\hat{\mathbf{x}}$) as a $K \times 1$ vector ($\mathbf{y}$) in an orthogonal subspace. When the number of principle components retained (K) is less than the rank of the covariance matrix ($\boldsymbol{\Sigma}_b$), the covariance of the normalized pixels projected into this subspace will be invertible. This new covariance ($\boldsymbol{\Sigma}_\mathbf{y}$) is a diagonal matrix with non-zero entries on its main diagonal.

Rewriting the LDA equation in the PCA subspace gives:

$$d_{LDA}(\mathbf{x}) = (\hat{\mathbf{s}} - \hat{\mu}_b)^T \mathbf{P} \boldsymbol{\Sigma}_\mathbf{y}^{-1} \mathbf{P}^T \hat{\mathbf{x}}, \tag{34}$$

where $d_{LDA}(\mathbf{x})$ is the score given to the pixel $\mathbf{x}$, $\hat{\mathbf{s}}$ is the normalized target signature, $\hat{\mu}_b$ is the background mean, $\mathbf{P}$ is the PCA projection operator from N dimensions to K orthogonal dimensions, $\boldsymbol{\Sigma}_\mathbf{y}$ is the background covariance matrix in the PCA

subspace, and $\hat{\mathbf{x}}$ is the normalized pixel.

3.3.3 Adaptive Coherence Estimator Detection

Like the LDA detector in Eq. (34), the ACE detector from Section 2.5.3 also requires an invertible covariance matrix. PCA will be used to project the normalized pixel data into an orthogonal subspace with an invertible covariance matrix for ACE detection. In the PCA subspace, the ACE detector from Eq. (18) is:

$$d_{ACE}(\mathbf{x}) = \frac{|\hat{\mathbf{s}}^T \mathbf{P} \mathbf{\Sigma_y}^{-1} \mathbf{P}^T \hat{\mathbf{x}}|^2}{(\hat{\mathbf{s}}^T \mathbf{P} \mathbf{\Sigma_y}^{-1} \mathbf{P}^T \hat{\mathbf{s}})(\hat{\mathbf{x}}^T \mathbf{P} \mathbf{\Sigma_y}^{-1} \mathbf{P}^T \hat{\mathbf{x}})}, \tag{35}$$

where $d_{ACE}(\mathbf{x})$ is the score given to the pixel $\mathbf{x}$, $\hat{\mathbf{s}}$ is the normalized target signature computed by Eq. (29), the matrix $\mathbf{P}$ is the PCA projection of the top K principal components, $\mathbf{\Sigma_y}$ is the background covariance matrix in the PCA subspace, and $\hat{\mathbf{x}}$ is the normalized pixel.

Section 3.2 calls for a radiance-based illumination-invariant dismount detection method. Using any one of the previously discussed detectors, MF, LDA, or ACE, radiance-based detection can be performed by searching a raw image for an estimated target signature. To achieve illumination-invariance, an estimate of the normalized illumination transform ($\hat{\mathbf{T}}$) must be found that allows these detectors to perform well under a wide range of illumination conditions. This will enable dismount detection with no knowledge of the illumination conditions.

3.4 Simulation Setup

The normalized illumination transform ($\hat{\mathbf{T}}$) used in Eq. (29) is a function of the sensor channel responses ($\mathbf{C}$) and the SRR ($\mathbf{l_{SRR}}$) reflected from a theoretical object with a unity reflectance vector ($\mathbf{r} = \mathbf{1}$). Specific transforms can be generated for a given HSI camera in a particular scenario using a radiative transfer (RT) code

to simulate the SRR in Eq. (21). The SRR from the theoretical "white" object is integrated through the channel filters of the sensor as in Eq. (27) and normalized to produce a specific illumination transform ($\hat{\mathbf{T}}$). The illumination transforms for many scenarios can be collected by varying the radiative transfer code over a range of different input conditions. From this collection, an estimate of the illumination transform can be derived to use in Eq. (29).

3.4.1 MODTRAN®5

The RT code MODTRAN® is used to simulate the SRR in Eq. (21). MODTRAN® is a collaborative development between Spectral Sciences, Inc., and the Air Force Research Lab Space Vehicles Directorate. It is an atmospheric simulation tool used in the remote sensing community to account for the effects of the Earth's atmosphere on imaging spectroscopy measurements. Many researchers use MODTRAN® to demonstrate their spectral detection algorithms on simulated radiance data [18].

MODTRAN®5, the most recent release, is used for these simulations. The wide range of input parameters for MODTRAN®5 gives the user precise control over many different aspects of the simulation, such as the sensor response, target properties, scene geometry, and atmospheric composition.

3.4.2 HST3 Sensor Response

The sensor to be simulated in MODTRAN®5 is the HST3 hyperspectral camera from HyperSpecTIR [23]. The HST3 pixels capture 227 spectral channels between 0.45μm and 2.45μm, with narrow channel bandwidths of 12nm at VIS wavelengths and 8.2nm at SWIR wavelengths. Narrow channel bandwidths are necessary to assume the illumination transform model discussed in Section 3.3.

3.4.3   Scene Geometry

Each MODTRAN®5 run is simulated using a theoretical "white" panel as the target of interest. The target has a unity spectral reflectance across VIS-SWIR wavelengths. The MODTRAN®5 runs are performed over various scene geometries and atmospheric conditions. In each simulation, both the sensor and the target are at sea level altitude. The sensor is geographically located at 45°N latitude, aimed towards the target 100 meters away to the north. The zenith angle of the sun from nadir varies between 5, 15, 30, 45, 60, and 75 degrees to simulate different times of day. The angle between the sun and the surface normal of the target, σ' in Eq. (4), is varied between 0, 45, and 90 degrees, such that $\cos \sigma'$ becomes 1, $\frac{1}{2}$, and 0, respectively. These different angles simulate the target being exposed to direct sunlight, indirect sunlight, and shaded from sunlight. The form factor (F) in Eq. (4) is held constant at 1, under assumption that the area of visible sky above the target remains constant.

3.4.4   Atmospheric Parameters

Although MODTRAN®5 enables fine-tuning of many atmospheric parameters, it also includes several preset atmospheres that replicate common conditions. Some of the preset atmospheres are used in these simulations. The MODEL parameter, which creates various preset geographical and seasonal model atmospheres, will take on values for a tropical atmosphere at 15°N latitude, mid-latitude summer at 45°N latitude, mid-latitude winter at 45°N latitude, and 1976 US standard atmosphere. Also, the IHAZE parameter, determining the type of aerosol extinction model, will be changed between rural, maritime, urban, and desert. The VIS parameter, which scales the aerosol extinction content to create a given meteorological visibility range in kilometers, will be set to 4km, 8km, 12km, 16km, 20km, and 24km.

The different MODTRAN®5 parameters and their ranges of values for these sim-

Table 3. Names and ranges of MODTRAN®5 parameters varied in these simulations.

Parameter	Values
Sensor	HST3 camera
Target	unity reflectance
Sensor altitude	sea level
Target altitude	sea level
Range	100 meters
Sun zenith angle	5°, 15°, 30°, 45°, 60°, 75°
Exposure angle σ'	0°, 45°, 90°
Atmosphere model	tropical, mid-lat summer, mid-lat winter, 1976 US std.
Aerosol extinction	rural, maritime, urban, desert
Visibility	4km, 8km, 12km, 16km, 20km, 24km

ulations are shown in Table 3.

3.5   Estimated Illumination Transform

For each combination of parameters in Table 3, the MODTRAN®5 simulation produces a normalized illumination transform ($\hat{\mathbf{T}}$) that describes the relationship between a reflectance vector ($\mathbf{r}$) and a pixel vector ($\mathbf{x}$) for the HST3 camera in a specific scenario. The mean normalized transform ($\hat{\mathbf{T}}_\mu$) is the element-wise average of the collection of M simulated normalized transforms $\hat{\mathbf{T}}_i, i \in \{1, M\}$:

$$\hat{\mathbf{T}}_\mu = \frac{1}{M} \sum_{i=1}^{M} \hat{\mathbf{T}}_i. \tag{36}$$

This normalized transform is used in Eq. (29) to create an estimated target radiance signature for the MF, LDA, and ACE detectors.

Choosing a single normalized transform to use in all scenarios has two benefits. First, a scene does not need to contain a set of ground truth objects to directly measure the illumination. Eliminating this requirement increases the usability of the dismount detection camera. Second, the target signature remains the same in all cases, not needing to be updated to the illumination conditions in each scene. This

improvement would reduce computing time and increase detection speed.

3.6   Summary

Radiance-based spectral detection can perform as well as reflectance-based detection if the illumination conditions are taken into account. Rather than converting HSI pixels into estimated reflectance vectors, which lengthens the algorithm chain, a radiance-based detector can operate directly on the unprocessed pixels data with an estimate of the illumination transform.

Dismount detection allows certain assumptions to be made that simplify the SRR model from remote sensing. A pixel vector is assumed to be the product of scene illumination and an object's spectral reflectance. Normalizing the pixel vectors in an image reduces the variability caused by different surface orientations of objects in the scene. A normalized target signature for human skin is generated from an illumination transform and a measurement of skin's spectral reflectance. Using that signature, the MF, LDA, and ACE detectors can search an image for skin pixels.

Instead of measuring the illumination transform with calibration panels, which lengthens the skin detection algorithm chain, it is estimated from a series of MOD-TRAN simulations. The mean of a large dataset of MODTRAN simulations for the illumination transform in various scenes provides a common illumination estimate. In the next chapter, the MF, LDA, and ACE detectors will be applied to hyperspectral images using a skin signature derived from the estimated illumination transform. The results will demonstrate the ability to perform dismount detection without measuring scene illumination.

4. Results and Analysis

This chapter presents the results of the illumination-invariant dismount detect strategy outlined in Chapter 3. First, the MODTRAN simulation results for the normalized illumination transform are presented in Section 4.1. Then the hyperspectral data collection process is described in Section 4.2, including the scene setup and the measured illumination conditions. A dismount is shown in a series of hyperspectral images, and the spectral reflectance of the exposed skin of the dismount is measured. In Section 4.3, the estimated illumination transform is compared to the measured illumination transform in each scene as measured from a white Spectralon calibration panel. The procedure for generating a normalized target signature for human skin is illustrated, and this skin signature is compared against the actual observed skin pixels from each image in Section 4.4.

The reflectance-based skin detection method from [37] is used as a baseline for comparing the performance of the three illumination-invariant methods from Section 3.3. Section 4.5 describes the pixel scoring procedure for each method. Finally, the detection results for all four methods are discussed in Section 4.6. Skin detection images are qualitatively compared based on how well they display the outline of the dismount while suppressing false-alarms. Then the detection methods are quantitatively compared with a series of ROC curves.

4.1 MODTRAN Simulations

The MODTRAN simulations in Section 3.4 produce 1728 normalized illumination transforms for the range of different scenarios. Each transform represents the response from each spectral channel of the HST3 camera to an object with a unity, Lambertian spectral reflectance being illuminated by natural daylight under specific atmospheric
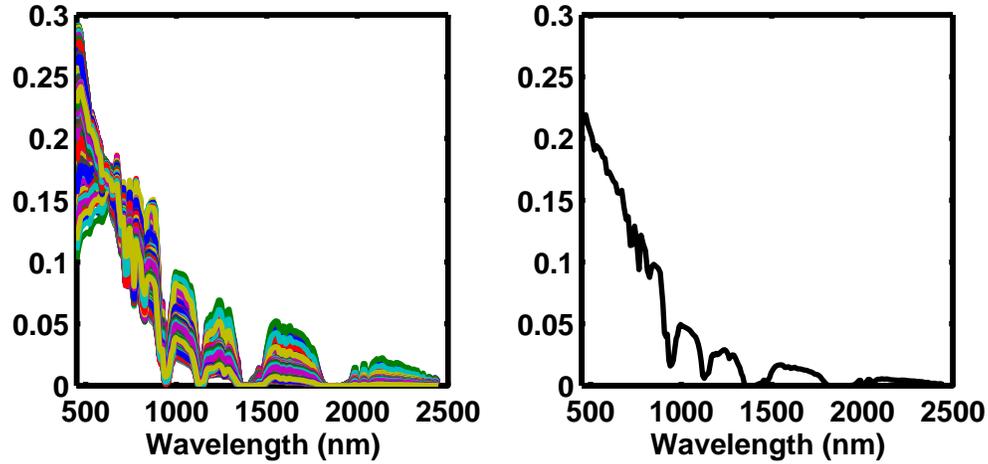
Figure 9. Left: 1728 different normalized illumination transforms from the simulations described in Section 3.4. Right: Mean normalized illumination transform derived using Eq. (36)**.**

conditions. Section 3.3 explains how a normalized illumination transform is multiplied by a reflectance vector to predict the normalized signature of a target with a known spectral reflectance. To create a general-purpose illumination transform when the illumination conditions are unknown, the 1728 simulation results are applied to Eq. (36) in Section 3.5, creating an estimate of the normalized illumination transform $\hat{\mathbf{T}}$.

All 1728 individual simulation results are plotted in Fig. 9. Although the conditions were unique for each simulation, they all produced similar illumination transforms. The illumination transforms have the highest magnitude in the VIS portion of the spectrum, from 450nm to 750nm, and taper off in the NIR and SWIR portions of the spectrum, from 750nm to 2500nm, similar to the solar irradiance distribution plot

in Fig. 2. HST3 channels in the VIS portion of the spectrum have a wider FWHM (12nm) than those in the rest of the spectrum (8.2nm), allowing more light energy to pass through those channels. That is why the VIS channels are higher than the NIR-SWIR channels in Fig. 9 compared to the solar irradiance curve in Fig. 2.

The widths of the HST3 channels also lend to the consistent shape of the illumination transforms throughout different atmospheric conditions. Slight changes in the levels of certain atmospheric aerosols can significantly effect atmospheric transmission ($\tau_1(\lambda)$) at the sub-nanometer scale. MODTRAN can simulate these effects at high spectral resolution. Hyperspectral cameras, however, are limited in spectral resolution by the FWHM of their channel filters, which are on the order of 10nm wide. HST3 channels are too wide to capture the fine spectral resolution where slight changes in atmospheric aerosols appear. As the atmospheric content changes, the illumination shape has little variance about its mean, making it acceptable to estimate the illumination transform by taking the mean of all simulation results.

The right side of Fig. 9 shows diagonal entries of the mean normalized illumination transform calculated from Eq. (36). Atmospheric attenuation bands, mentioned in Section 2.2, are noticable around 1350nm-1430nm and 1800nm-1950nm, just as in Fig. 3. A relatively low amount of electromagnetic energy reaches through the atmosphere to the surface of the Earth at these wavelengths. This creates a problem when converting pixels into reflectance vectors using ELM as described in Section 2.4.2. Since the pixel values of the white and gray calibration panels are nearly zero at the channels of the atmospheric attenuation bands, the solution to Eq. (11), finding the ELM conversion constants for those channels, has a large error. Estimated reflectance vectors using those conversion constants would be significantly inaccurate in the atmospheric attenuation bands. To empirically estimate an object's spectral reflectance at a given band, there must be sufficient illumination energy at that band. Natural

illumination, due to the atmospheric attenuation bands, is not ideal for full-spectrum reflectance estimation. This justifies the need for radiance-based dismount detection when relying on natural illumination.

4.2 Data Collection

A dataset of hyperspectral images is collected with the HST3 camera in Dayton, OH, on 6 March 2009. The intent of the data collect is to capture the same dismount detection scene multiple times under different illumination conditions. This is accomplished by taking images at different times of the day as the sun is partially blocked by passing clouds.

The hyperspectral dataset consists of four sequences, with 12 images in each sequence, shown in Fig. 10. The first sequence is taken at 0945, the second at 1000, the third at 1100, and the fourth at 1430. Each sequence captured the scene illuminated by the sun at a different zenith angle. The weather is partly cloudy on the date of the data collect, with mostly clear skies in the morning and overcast skies in the afternoon. During the first two sequences at 0945 and 1000, the sun is partially blocked by passing clouds, producing a different illumination for each image in each sequence. In the third sequence, at 1100, the sky is mostly clear and the sun is high in the sky, so the scene illumination is composed of direct sunlight. By the time the fourth sequence is collected at 1430, the sky is overcast. Illumination in the fourth sequence is dimmer and more indirect than in the previous sequences.

The scene configuration is similar to the conditions in the MODTRAN simulations described in Section 3.4. The HST3 camera is set up at ground level, with the aperture facing northeast. The setting is a suburban area that includes houses, trees, and a road. Two Spectralon calibration panels are placed on the ground eight meters in front of the HST3 camera. The panels stand upright, with the surface normal of each
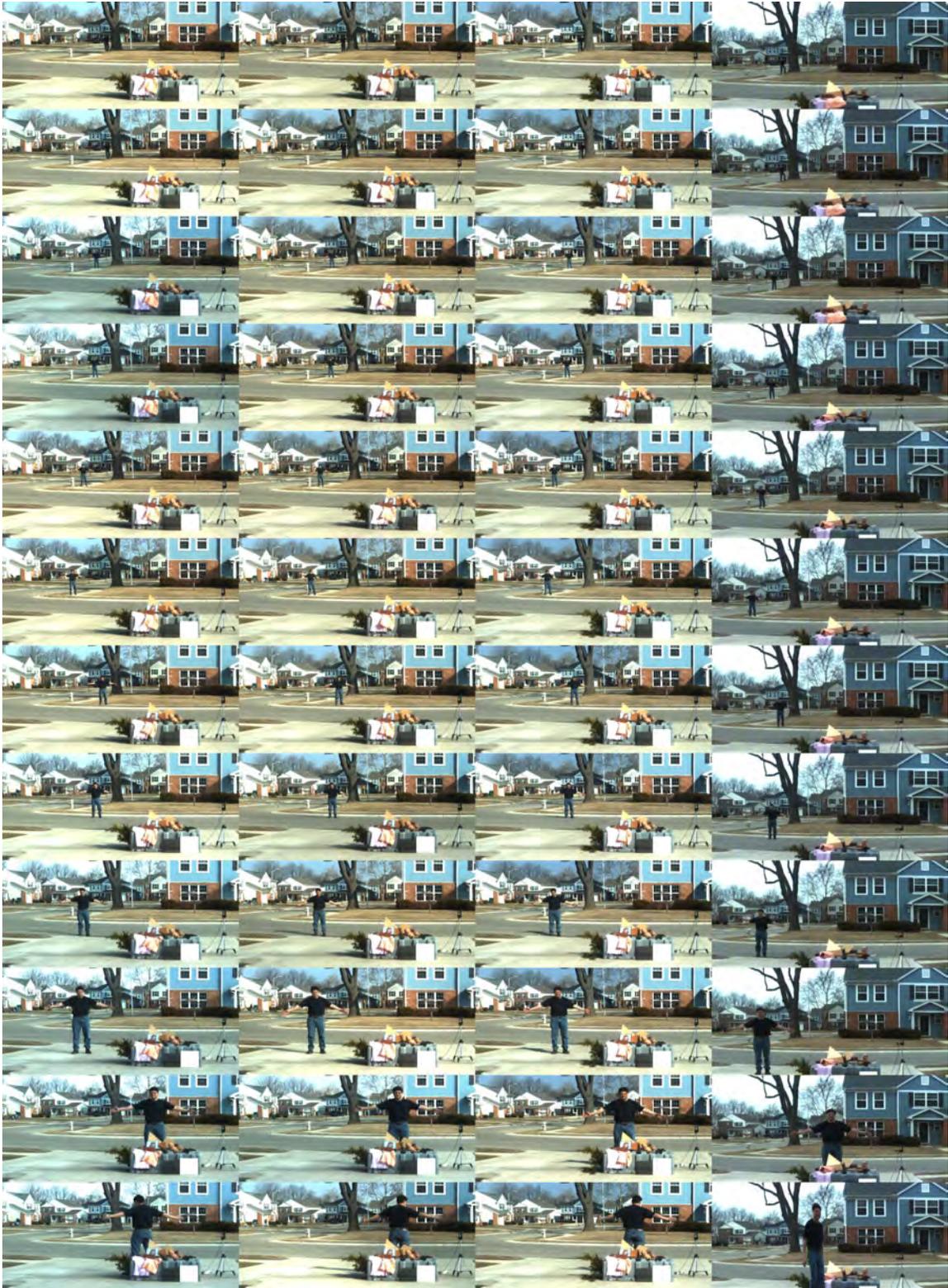
Figure 10. Hyperspectral data collected for this study, converted to RGB for display. The first column is the sequence collected at 0945, the second at 1000, the third at 1100, and the fourth at 1430.

panel pointing towards the camera to expose as much surface area as possible. These panels are included in the scene to provide a spectral reflectance reference as required in Section 2.4.2.

The images from each sequence contain a dismount with arms extended out to the side. The dismount, Caucasian male with dark brown hair, is wearing a black, cotton, short-sleeved t-shirt and blue, denim jeans. Skin is exposed on the face and arms. Starting from a point 50 meters away from the camera in the first image, the dismount walks towards the camera. In each image, the dismount progresses closer to the camera, finally stopping eight meters in front of the camera in the twelfth image.

After the images are collected, skin pixels are selected by hand to create truth masks. Also, the pixels of both of the Spectralon calibration panels are selected for their own truth masks. The spectral reflectance of the exposed skin on the forearms and face of the dismount is measured 100 times with an ASD contact probe. These measurements are sampled at the center wavelengths of the HST3 channel filters as described in Section 3.3, and averaged to create a skin reflectance vector. The skin reflectance measurements and the resulting mean skin reflectance vector are shown in Fig. 11.

4.3   Illumination Comparison

The MODTRAN simulation results in Section 4.1 display the response of the HST3 camera to the radiance from an object with a constant, unity spectral reflectance under simulated conditions. A similar process is used to measure the normalized illumination transforms from each image in Fig. 10. The white Spectralon calibration panel has a spectral reflectance that is nearly one at wavelengths between 400nm and 2500nm. Its spectral reflectance, as collected by the ASD field spectrometer, is shown in Fig. 12. After dividing by its known spectral reflectance, pixels of the white
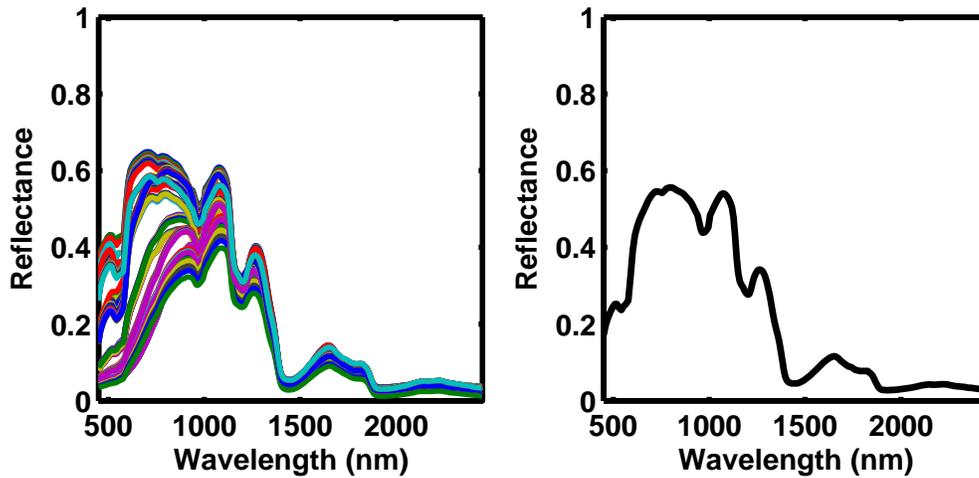
Figure 11. Left: 100 skin reflectance measurements taken with an ASD field spectrometer of the exposed skin on the forearms and face of the dismount shown in the HST3 images in Fig. 10. Right: Mean skin reflectance vector used to generate a target signature.
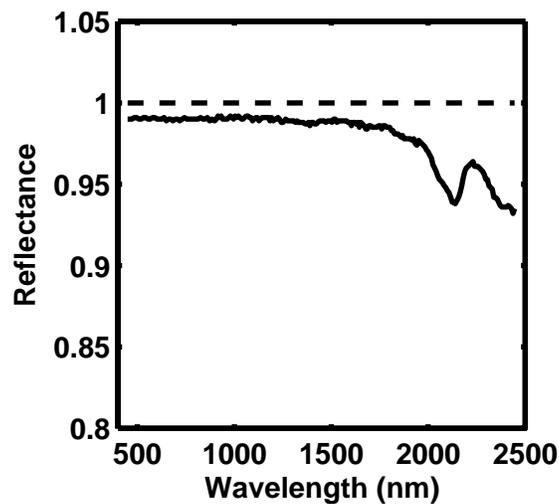
Figure 12. Spectral reflectance of the white Spectralon calibration panel used to measure scene illumination in each HST3 image. Since its spectral reflectance is very close to one for wavelengths between 450nm and 2500nm, the white Spectralon panel makes a good reference object for measuring scene illumination.
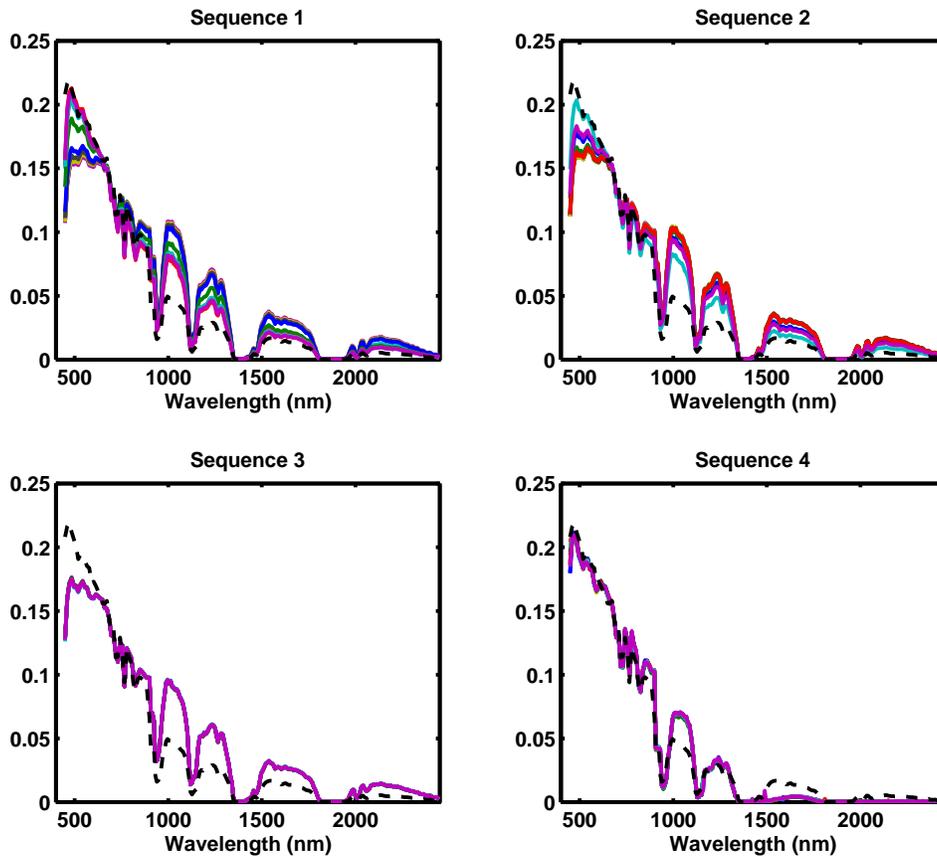
Figure 13. Normalized illumination transforms measured from the white Spectralon calibration panel in each sequence. All twelve measurements for each sequence are plotted as the colored lines. The dashed line is the estimated illumination transform in Section 4.1 obtained from MODTRAN simulations.

panel estimate the scene illumination transform in the same way as the MODTRAN simulations.

The normalized illumination transform for each image is measured by averaging all of the pixels from the white Spectralon panel, dividing by the white panel's spectral reflectance, and normalizing the result. These measurements are shown in Fig. 13. During sequences 1 and 2, the sun was partially blocked by passing clouds, causing the measured illumination to change between images. This appears in Fig. 13 as the distribution of colored curves. Sequences 3 and 4 had more consistent illumi-
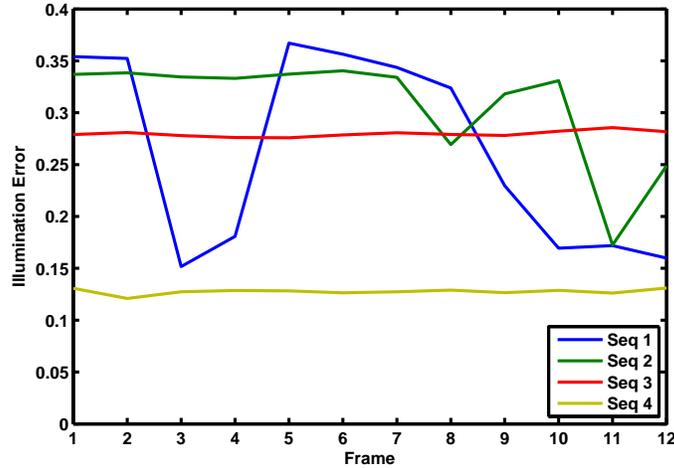
Figure 14. Illumination error given by the Euclidean distance between the estimated illumination and measured illumination in each frame. Illumination stayed consistently sunny in sequence 3 and cloudy in sequence 4. Passing clouds blocked the sunlight in frames 3,4,9-12 of sequence 1, and frames 9-12 of sequence 2.

nation between images. For comparison, the estimated illumination transform from Section 4.1 is also plotted in Fig. 13 as a black dashed line.

The error between the estimated illumination transform used to generate the target signature in Eq. (29) and the measured illumination in each image is plotted in Fig. 14. The error is computed as the Euclidean distance between the diagonal entries of both transforms:

$$\text{Illumination Error} = \sqrt{\sum_{i=1}^{N}(\hat{T}_{A,ii} - \hat{T}_{B,ii})^2}, \tag{37}$$

where $\hat{T}_{A,ii}$ is the ith diagonal entry of the estimated illumination transform and $\hat{T}_{B,ii}$ is the ith diagonal entry of the measured illumination transform.

The estimated illumination transform most closely matches the measured illumination transform in sequence 4, when the sky is overcast. In sequences 1 and 2, the illumination error drops as the sunlight is partially blocked by passing clouds between
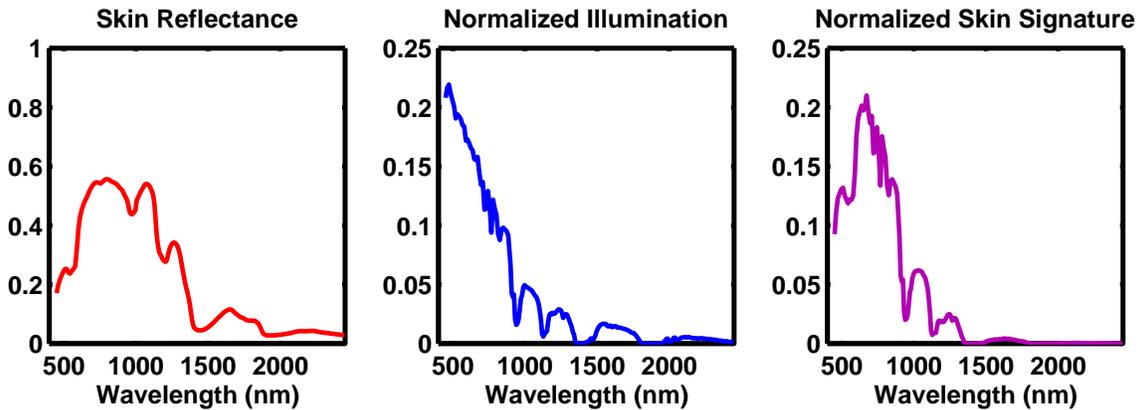
Figure 15. Generating a normalized skin signature using Eq. (29). The skin reflectance vector (r, left) of the dismount is transformed by the estimated illumination transform ($\hat{\mathbf{T}}$, middle) into an estimated skin signature ($\hat{\mathbf{s}}$, right).

frames.

4.4 Skin Signature

In Section 2.5, the Matched Filter (MF), Linear Discriminant Analysis (LDA), and Adaptive Coherence Estimator (ACE) spectral detection algorithms were introduced. These algorithms search hyperspectral pixels for a normalized target signature ($\hat{\mathbf{s}}$), assigning scores to pixels based on how similar they are to the target signature. For a target with a known reflectance vector ($\mathbf{r}$), its normalized signature can be generated using the estimated illumination transform in Eq. (29). This section compares the empirical appearance of human skin pixels in the HST3 images shown in Fig. 10 to the estimated skin signature generated by Eq. (29).

The diagram in Fig. 15 illustrates the concept of generating a skin signature with the estimated illumination transform. On the left is a plot of the average spectral reflectance of skin shown in Fig. 11. In the middle is a plot of the estimated illumination transform shown in Fig. 9. Both of these are known *a priori*, before any of the
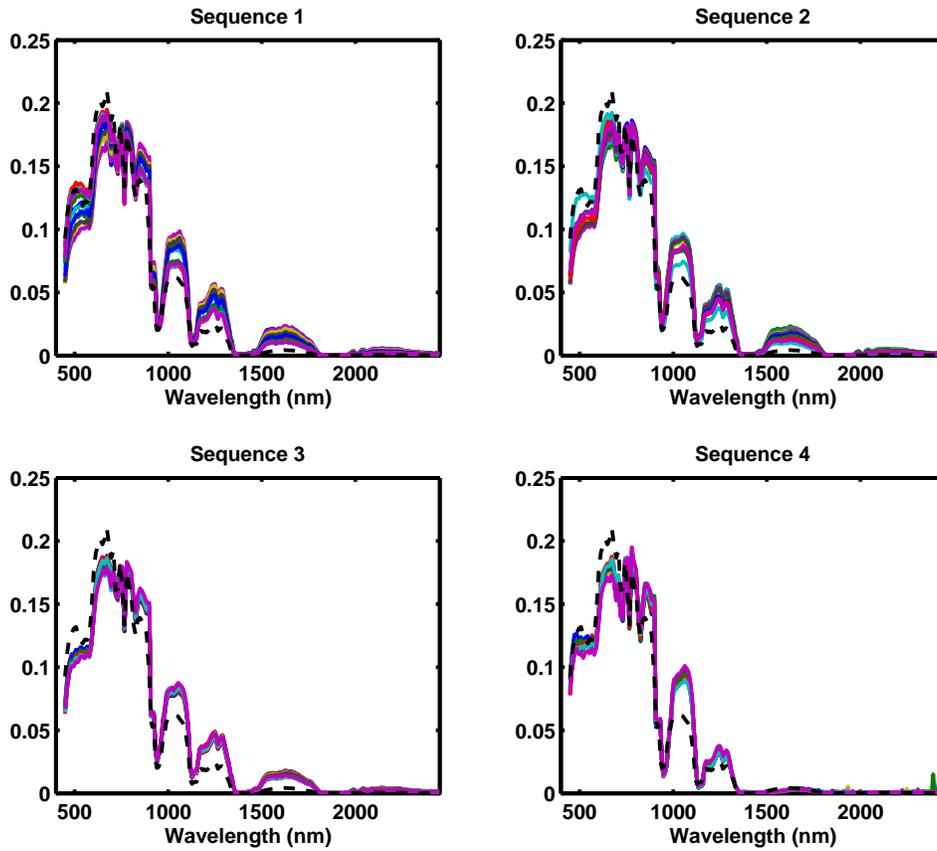
Figure 16. Normalized skin pixels hand-selected from each sequence of images. All twelve mean normalized skin pixels are plotted for each sequence as the colored lines. The dashed line is the estimated skin signature ($\hat{\mathbf{s}}$) generated from Eq. (29) using the estimated illumination transform.

images in Fig. 10 are taken. When these are combined in Eq. (29), the result is an estimated skin signature ($\hat{\mathbf{s}}$) shown on the right of Fig. 15. This skin signature requires no knowledge of the atmospheric or illumination conditions in the scene because it is generated from the estimated illumination transform in Fig. 9.

To verify that $\hat{\mathbf{s}}$ is an acceptable estimate of the actual normalized skin signature, skin pixels collected by the HST3 imager are compared to this estimate in Fig. 16. Each colored line represents the average of all normalized skin pixels from one image. Twelve colored lines, one for each image, are plotted for each sequence. The dashed

black line is the estimated skin signature ($\hat{\mathbf{s}}$), shown for reference. In all four sequences, the estimated and actual normalized skin pixels are very similar, especially at VIS-NIR wavelengths from 400nm to 900nm. Between 900nm and 1500nm, $\hat{\mathbf{s}}$ has a smaller magnitude than do the actual skin pixels, because the estimated illumination transform at those wavelengths is lower than what is measured. At SWIR wavelengths from 1500nm to 2500nm, both the estimated and actual skin pixels are nearly zero. Natural illumination is weak at these wavelengths compared to VIS wavelengths, and so is the spectral reflectance of human skin.

4.5   Pixel Scoring

4.5.1   Reflectance-Based Method

The NDSI skin detection method described in Section 2.6.2 requires ELM correction to convert pixels into estimated reflectance vectors. ELM correction uses a linear transform that is derived by observing the radiance from two Spectralon calibration panels with known spectral reflectances. For each HST3 image, the pixels of the white and gray Spectralon panels are selected by hand and averaged. These average pixels are applied to Eq. (11) to generate the ELM transform matrices $\mathbf{M}$ and $\mathbf{b}$ in Eq. (12). The ELM transform is applied to the pixels to create estimated reflectance vectors.

Once the pixels of an image are converted into estimated reflectance vectors, they are scored by the NDSI and NDGRI equations, written here:

$$\text{NSDI} = \frac{r'(1080\text{nm}) - r'(1580\text{nm})}{r'(1080\text{nm}) + r'(1580\text{nm})}, \tag{38}$$

$$\text{NDGRI} = \frac{r'(660\text{nm}) - r'(540\text{nm})}{r'(660\text{nm}) + r'(540\text{nm})}, \tag{39}$$

where $r'(\lambda)$ is the value of the estimated reflectance vector ($\mathbf{r}'$) at the spectral channel

closest to λ.

When using the NDSI method of skin detection, independent thresholds must be applied to both the NDSI and NDGRI scores, creating two degrees of freedom for the classifier. This makes it difficult to generate a ROC curve, which is a function of a single threshold. To compare the performance of the NDSI skin detection method with the other algorithms by their ROC curves, the NDSI and NDGRI values for each pixel must be combined into a single score.

In [37], the NDSI was proposed as a score for how likely a pixel is skin. It was noted that green vegetation created false alarms because it also scored highly on the NDSI. The NDGRI was introduced for false-alarm suppression because vegetation is more green than red, while skin is more red than green. A skin reflectance vector has a high NDSI and a negative NDGRI. From this context, a single score is used in this thesis that incorporates both the NDSI and NDGRI values. The score, denoted d_{NDSI}, is equal to the NDSI if the NDGRI is negative, and is equal the NDSI minus one if the NDGRI is positive:

$$d_{NDSI} = \begin{cases} \text{NDSI}, & \text{if NDGRI} < 0 \\ \text{NDSI} - 1, & \text{if NDGRI} > 0 \end{cases}, \qquad (40)$$

where NDSI and NDGRI are calculated from an estimated reflectance vector using Eq. (38) and Eq. (39).

The output of Eq. (40) ranges between -2 and 1. Equation (40) is a fair representation of the NDSI skin detection method in one dimension instead of two. Skin pixels are given high scores because they have a high NDSI, while the NDGRI carries out its false alarm suppression purpose by reducing the score of green vegetation.

4.5.2   Illumination-Invariant Methods

The three illumination-invariant detection algorithms applied to the HST3 images are the Matched Filter detector in Eq. (17), the Linear Discriminant Analysis detector in Eq. (34), and the Adaptive Coherence Estimation detector in Eq. (35). The MF detector assigns a score (d_{MF}) to each pixel according to its correlation with the estimated skin signature $(\hat{\mathbf{s}})$. The value of d_{MF} ranges from 0 to 1.

Unlike the MF detector, the LDA and ACE detectors require an invertible background covariance matrix $\boldsymbol{\Sigma}_b$. PCA is used to project the hyperspectral data onto its K principle components, making the background covariance matrix invertible. For this thesis, $K = 12$ is chosen as the number of principle components to retain. This value is chosen after trial and error, where it is found that the performance of $K = 12$ is better than $K = 4$ and $K = 8$, but similar to $K = 16$.

4.6   Detection Results

4.6.1   False Alarm Comparison

Skin detection results from four of the hyperspectral images are compared in Figs. 17, 18, 19, and 20. The four images are the tenth frames from each of the four sequences. These four images are compared because the dismount is in a similar position close to the camera, and because the illumination is different between them. The normalized illumination transforms from each image, measured from the white Spectralon panel, are shown together in Fig. 21. The sky is partly cloudy in the first image, mostly clear in the second image, cloudless in the third image, and overcast in the fourth image. Color versions of the images are in the top left, and truth masks of hand-selected skin pixels are in the top right of each figure. White pixels are those that the detection method classifies as skin.

Sequence 1, Frame 10

Truth
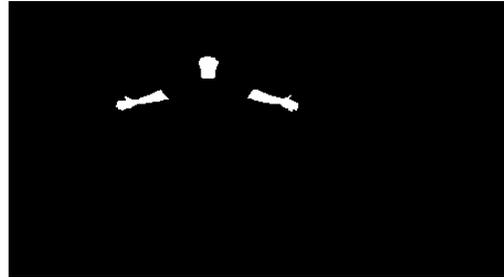
NDSI: P$_{FA}$ = 0.0009
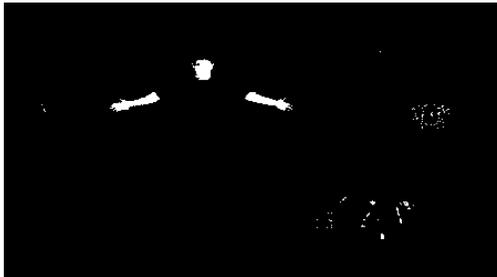
MF: P$_{FA}$ = 0.0007

LDA: P$_{FA}$ = 0.0034

ACE: P$_{FA}$ = 0.0011

Figure 17. Frame 10 from the first sequence of images. The sky is partly cloudy. Detection results are thresholded such that $P_D = 0.8$. Illumination-invariant methods (middle right, bottom left, bottom right) have a P_{FA} as low as the reflectance-converted NDSI method (middle left).

Sequence 2, Frame 10

Truth
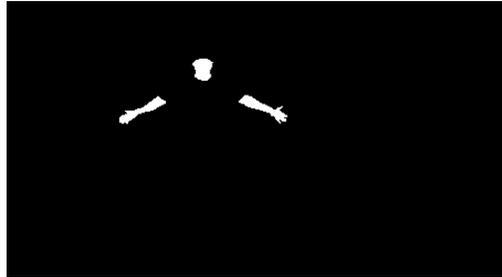
NDSI: P_{FA} = 0.0006
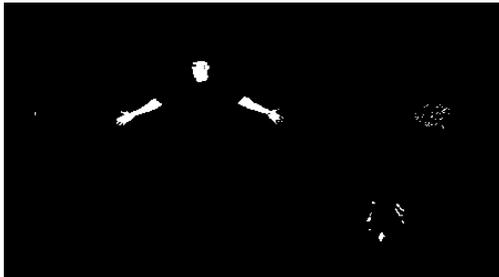
MF: P_{FA} = 0.0134

LDA: P_{FA} = 0.0057

ACE: P_{FA} = 0.0059
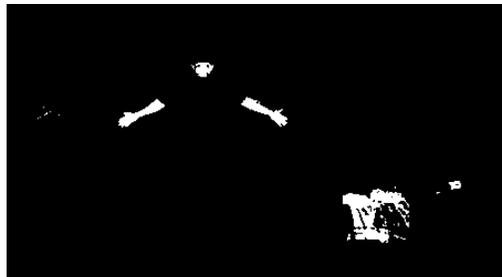
Figure 18. Frame 10 from the second sequence of images. The sky is mostly clear. Detection results have been thresholded such that $P_D = 0.8$. Illumination-invariant methods (middle right, bottom left, bottom right) have a noticeably higher P_{FA} than the reflectance-converted NDSI method (middle left).

Sequence 3, Frame 10

Truth
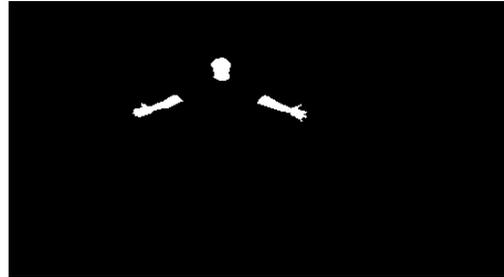
NDSI: P$_{FA}$ = 0.0005
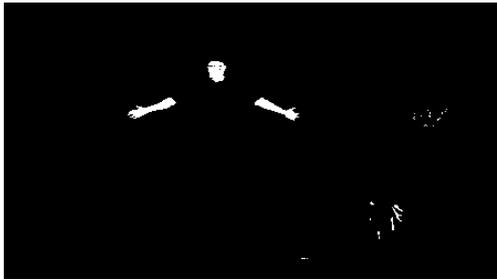
MF: P$_{FA}$ = 0.0094

LDA: P$_{FA}$ = 0.0054
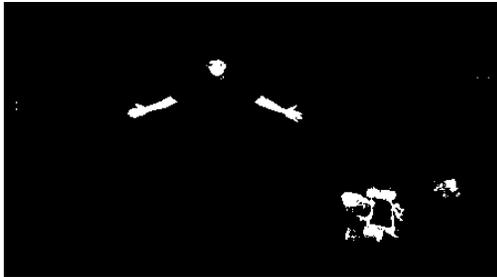
ACE: P$_{FA}$ = 0.0022
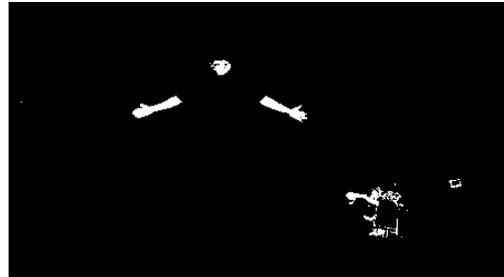
Figure 19. Frame 10 from the third sequence of images. The sky is cloudless. Detection results have been thresholded such that $P_D = 0.8$. Illumination-invariant methods (middle right, bottom left, bottom right) have a noticeably higher P_{FA} than the reflectance-converted NDSI method (middle left).

Sequence 4, Frame 10

Truth
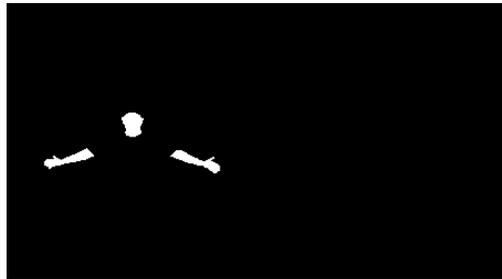
NDSI: P$_{FA}$ = 0.052

MF: P$_{FA}$ = 0.005

LDA: P$_{FA}$ = 0.0035

ACE: P$_{FA}$ = 0.0006

Figure 20. Frame 10 from the fourth sequence of images. The sky is overcast. Detection results have been thresholded such that $P_D = 0.8$. Illumination-invariant methods (middle right, bottom left, bottom right) have a much lower P_{FA} than the reflectance-converted NDSI method (middle left).
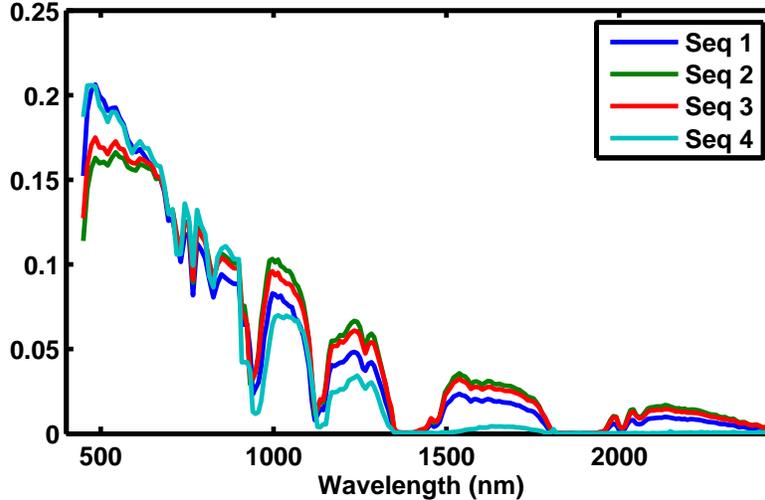
Figure 21. Normalized illumination measured from the white Spectralon panel for frame 10 in all four sequences. The illumination conditions in these four images were unique, as indicated by the non-overlapping illumination vectors above.

Skin detection images for the NDSI, MF, LDA, and ACE algorithms are created by applying a threshold to scored pixels such that 80% of the hand-selected skin pixels are correctly classified. At a constant detection rate of $P_D = 0.8$, the corresponding false-alarm rate indicates how well each algorithm suppresses false alarms while detecting skin.

In the partially cloudy scene (Fig. 17), the modified NDSI method from Eq. (40) produces a false-alarm rate of $P_{FA} = 0.0009$. The illumination-invariant algorithms have similar false-alarm rates on the same image: MF has $P_{FA} = 0.0007$, LDA has $P_{FA} = 0.0034$, and ACE has $P_{FA} = 0.0011$. The outline of the dismount is discernible in all four detection images.

Results on the mostly-clear scene (Fig. 18) are similar to those in the partially cloudy scene (Fig. 17), except that the illumination-invariant algorithms produce more false alarms. The modified NDSI method has $P_{FA} = 0.0006$, while MF has $P_{FA} = 0.0134$, LDA has $P_{FA} = 0.0057$, and ACE has $P_{FA} = 0.0059$. The incorrectly detected pixels obscure the outline of the dismount in the MF results. A pink t-shirt

and a red brick are incorrectly classified as skin in the LDA and ACE results.

The modified NDSI method continues to have a lower false-alarm rate in the cloudless scene (Fig. 19), with a false-alarm rate of $P_{FA} = 0.0005$, compared to MF with $P_{FA} = 0.0094$, LDA with $P_{FA} = 0.0054$, and ACE with $P_{FA} = 0.0022$. However, all four algorithms have a lower P_{FA} than in the mostly-clear scene (Fig. 18). This could be due to an artifact from the hand-selected skin truth images possibly containing mislabeled pixels.

Out of the four frames displayed, the results of the overcast scene (Fig. 20) are the most dramatic. In this scene, the modified NDSI method performs worse, with $P_{FA} = 0.052$, than the illumination-invariant methods: MF has $P_{FA} = 0.005$, LDA has $P_{FA} = 0.0035$, and ACE has $P_{FA} = 0.0006$. The false-alarms in the NDSI image obscure the shape of the dismount and appear spread out across the image. There is not one particular object that caused its high P_{FA}, like the t-shirt and brick did for the LDA and ACE methods in the other scenes.

ROC curves displayed in Fig. 22 compare the performance of all four detection methods on the images in Figs. 17, 18, 19, and 20. A horizontal line drawn through the ROC curves at $P_D = 0.8$ represents their operating points. Although a detection rate of only 80% is somewhat low for a target detection algorithm, it is used in this thesis because it accomplishes two objectives for a dismount detection system. First, even with 80% of the skin pixels detected the shape of the dismount is still noticeable in the thresholded images. Second, limiting the detection rate to $P_D = 0.8$ keeps the false alarm rate very low. In a skin-cued dismount detection system like the one in [6], a low false alarm rate is more important than a high detection rate, because false alarms increase the search time.

Out of the three illumination-invariant methods, the MF performs the worst. Its detection rate is lower than the LDA and ACE methods at false-alarm rates above
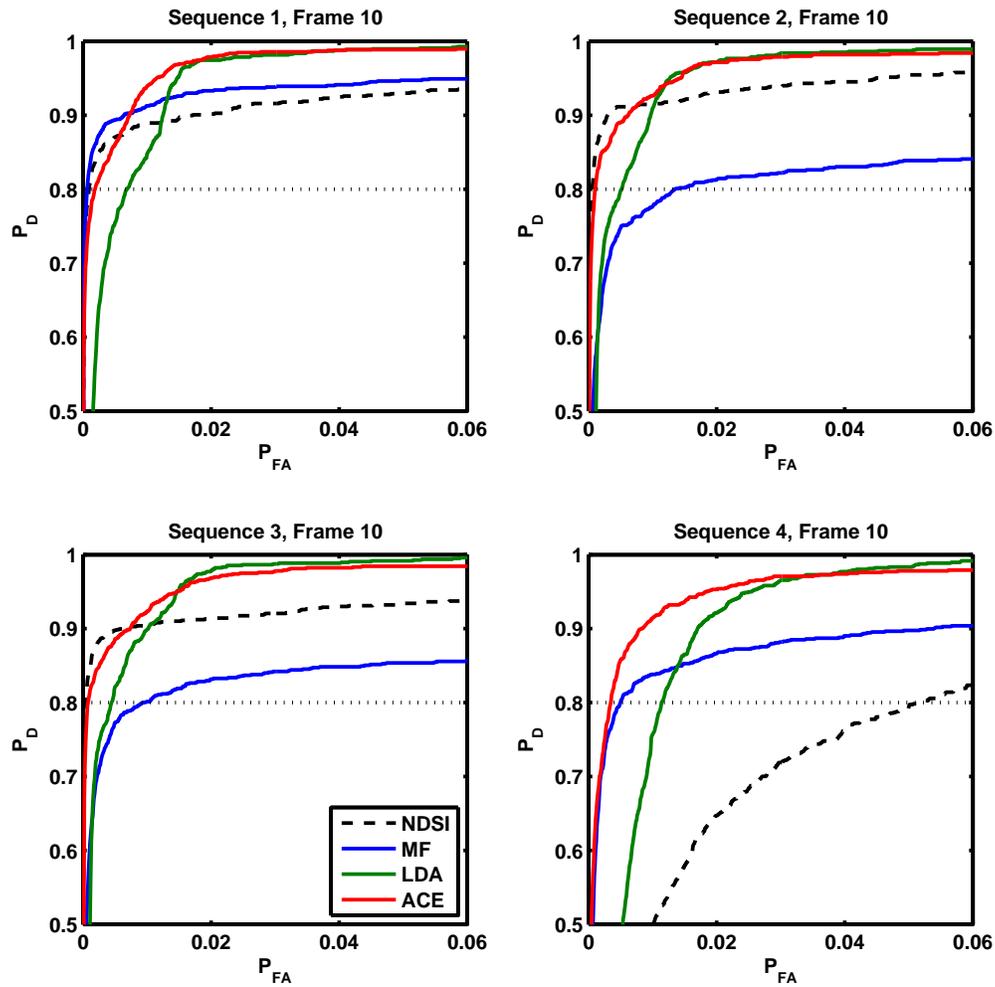
Figure 22. ROC curves from the four skin detection methods compared in Fig. 17, 18, 19, and 20. The dotted line at $P_D = 0.8$ represents the operating point of the four methods, and where it crosses the ROC curves is the corresponding P_{FA} for that method.

$P_{FA} = 0.02$, and it only outperforms the modified NDSI method on two of the four images. The LDA and ACE methods have the highest detection rate at false-alarm rates above $P_{FA} = 0.02$. Below $P_{FA} = 0.02$, the ACE method outperforms the LDA method, but the modified NDSI method performs better than both. The exception is in the overcast scene (Fig. 20), where the modified NDSI method detects fewer skin pixels than the illumination-invariant methods at all false-alarm rates.

What is notable about these results is that the NDSI method and the illumination-invariant methods produce visually similar detection results. The NDSI method requires the white and gray Spectralon calibration panels in order to perform ELM correction on the pixels, while the illumination-invariant algorithms do not.

4.6.2   Effect of Illumination Error

The four images discussed in the previous section are chosen because each are subject to different illumination conditions. The illumination-invariant skin detection methods have lower false alarm rates in the cloudy scene (Fig. 20) than in the other scenes. It is also shown in Fig. 14 that the error between the estimated and measured illumination transform is the lowest in the cloudy scene. This section investigates the effect of illumination error on the performance of the detection methods.

The four graphs in Fig. 23 plot false-alarm rate against illumination error for all 48 images. The false-alarm rate is drawn from the ROC curve at the $P_D = 0.8$ operating point. The illumination error is calculated from Eq. (37).

Neither the NDSI method nor the MF method appear to show a correlation between illumination error and false-alarm rate. This is not surprising for the NDSI method, which does not use the estimated illumination transform. However, it suggests that the MF method performs consistently, although poorly, even when the scene illumination varies. The LDA and ACE methods, on the other hand, show a general
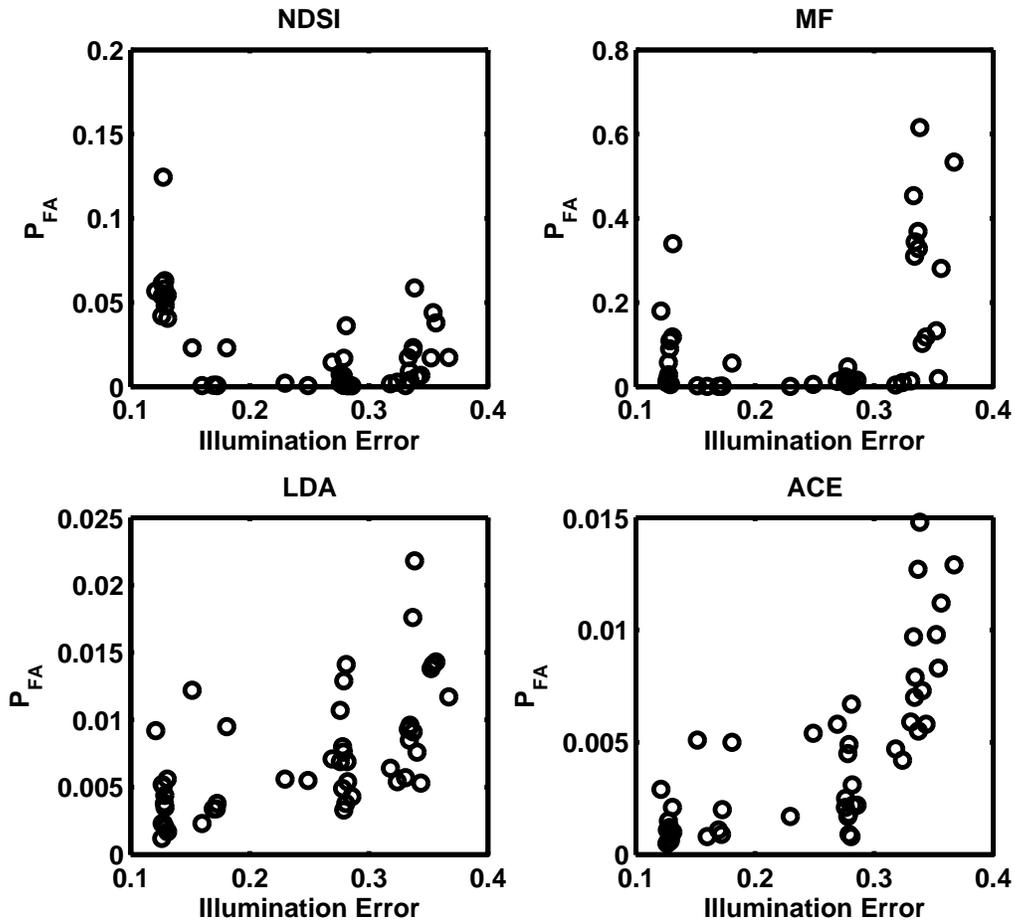
Figure 23. False-alarm rate at $P_D = 0.8$ plotted against illumination error for each method. The two factors appear independent from one another in the NDSI and MF methods, while they are somewhat correlated in the LDA and ACE methods.

trend of an increasing false alarm rate as the illumination error grows. Also, their highest false-alarm rates are lower than the highest of the NDSI and MF methods. Still, the LDA and ACE methods both detect human skin without using calibration panels at performances comparable to the panel-dependent NDSI method.

4.6.3   Constant Threshold Comparison

In Figs. 24, 25, 26, and 27, the skin detection methods are applied to each sequence of images to illustrate what they would display in a real-time dismount detection scenario. Unlike the previous sections where the thresholds were set to achieve a desired detection rate, the thresholds in this section are held constant throughout all four sequences. This is more realistic for a real-time dismount detection scenario. The modified NDSI scores are thresholded at 0.25, the MF scores at 0.975, the LDA scores at 37, and the ACE scores at 0.983. Pixels that score above these thresholds are classified as skin. Detection results for the modified NDSI method are included for reference, even though that method cannot operate in real-time because it relies on the hand-selected calibration panels.

The LDA method mistakes the pink t-shirt and the red brick in the center of the scene as a dismount in the first three sequences, but not in the fourth sequence. It is capable of a lower false-alarm rate in the sunny scenes, demonstrated in Fig. 17, but that requires adjusting the detection threshold. Ideally, a real-time dismount detection algorithm should maintain consistent performance between different illumination conditions without changing the detection threshold.

The pink t-shirt also causes false alarms for the ACE method in sunny scenes but not as many as the LDA method. While the false alarm rate of the ACE method decreases in cloudy skies, the opposite is true for the MF method. When the sky is cloudy in the fourth sequence (Fig. 20), the MF method produces false alarms on

Sequence 1

Frame: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

NDSI · MF · LDA · ACE
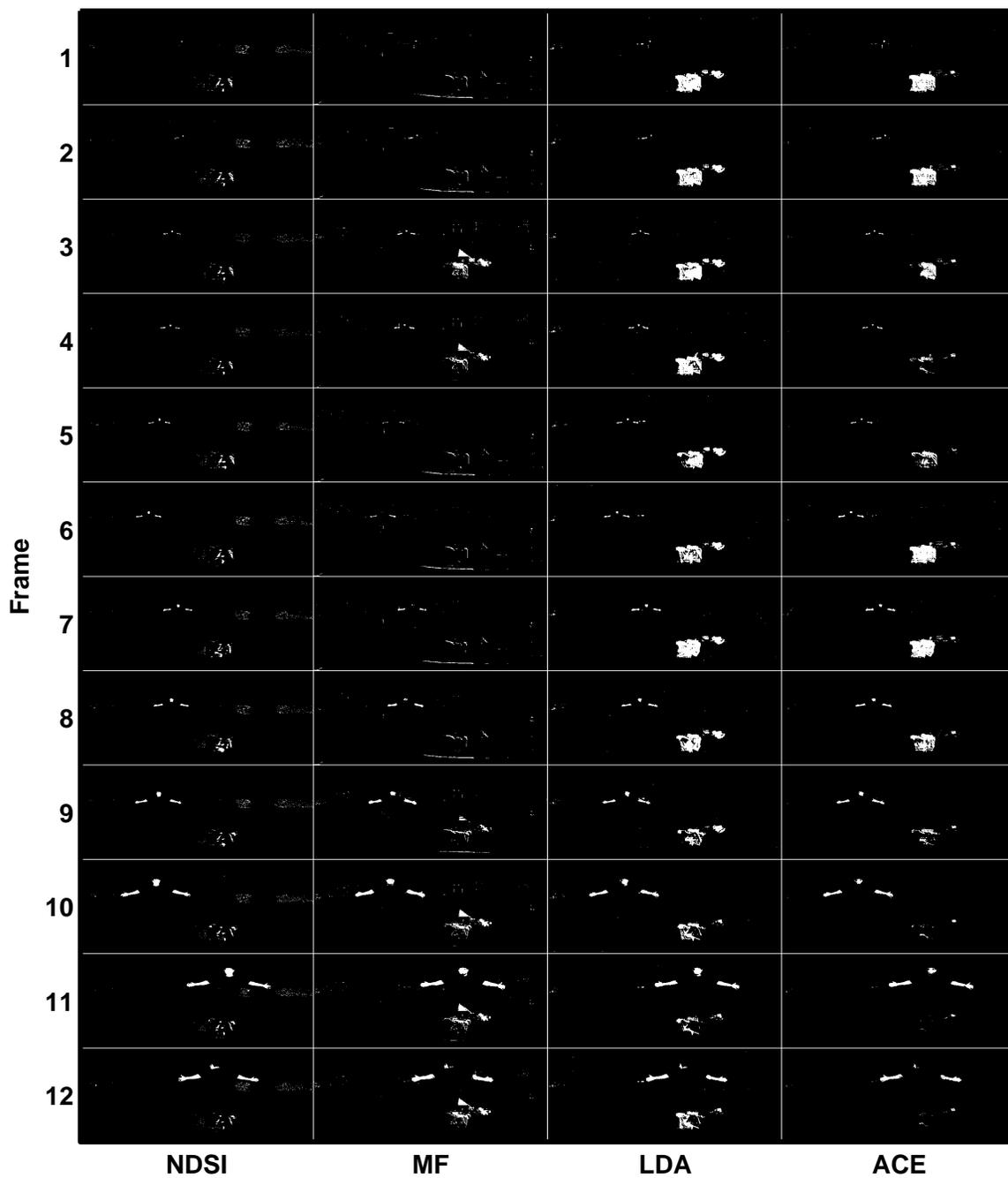
Figure 24. Detection results for the first sequence with constant detection thresholds of 0.25 for NDSI, 0.975 for MF, 37 for LDA, and 0.983 for ACE. As a passing cloud blocks the sunlight in frames 3 and 4, the number of false alarms increases in the MF method and decreases in the ACE method.

Sequence 2
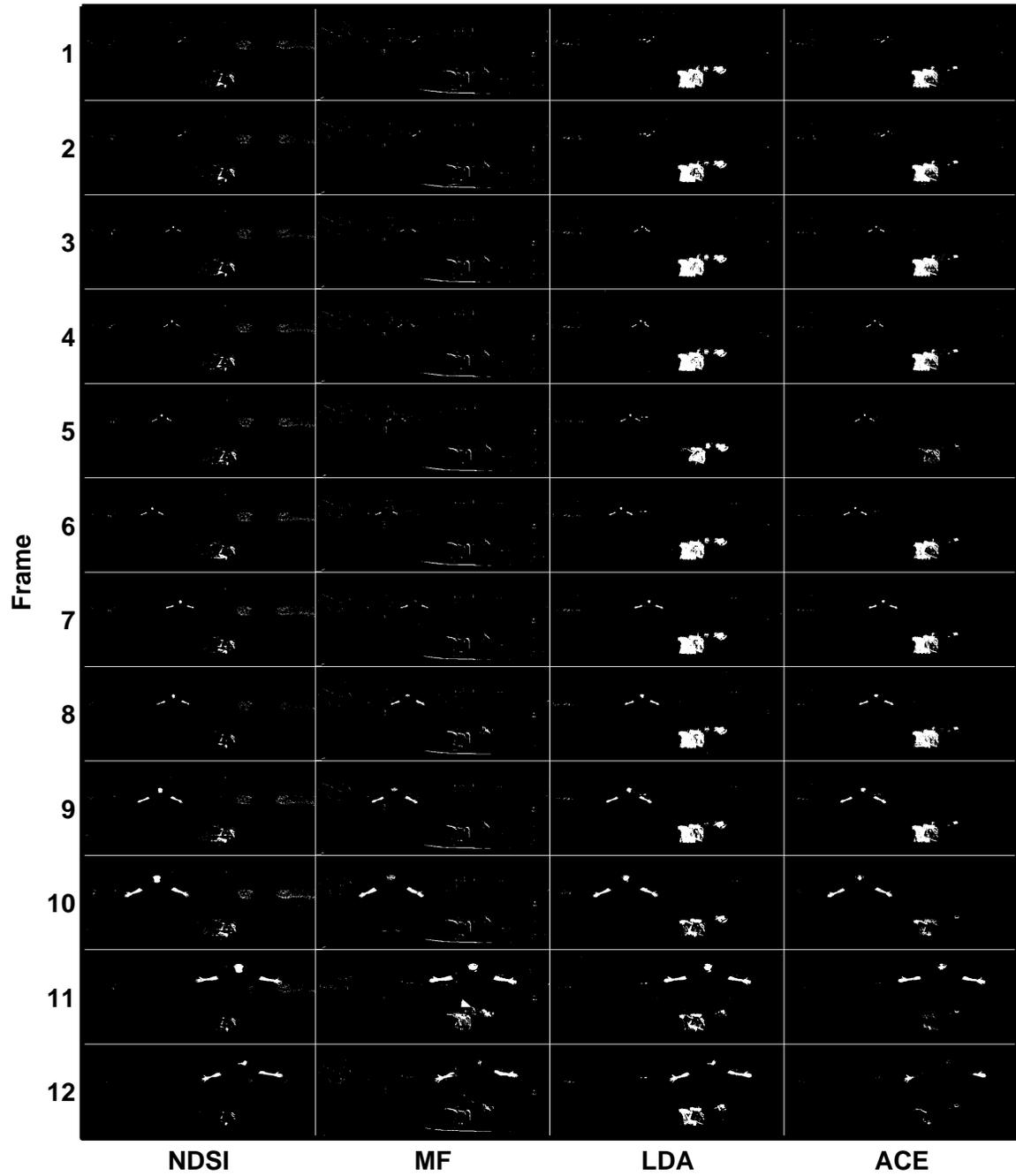
Figure 25. Detection results for the second sequence with constant detection thresholds of 0.25 for NDSI, 0.975 for MF, 37 for LDA, and 0.983 for ACE. A pink t-shirt and a red brick cause false alarms in the LDA and ACE methods.

Sequence 3
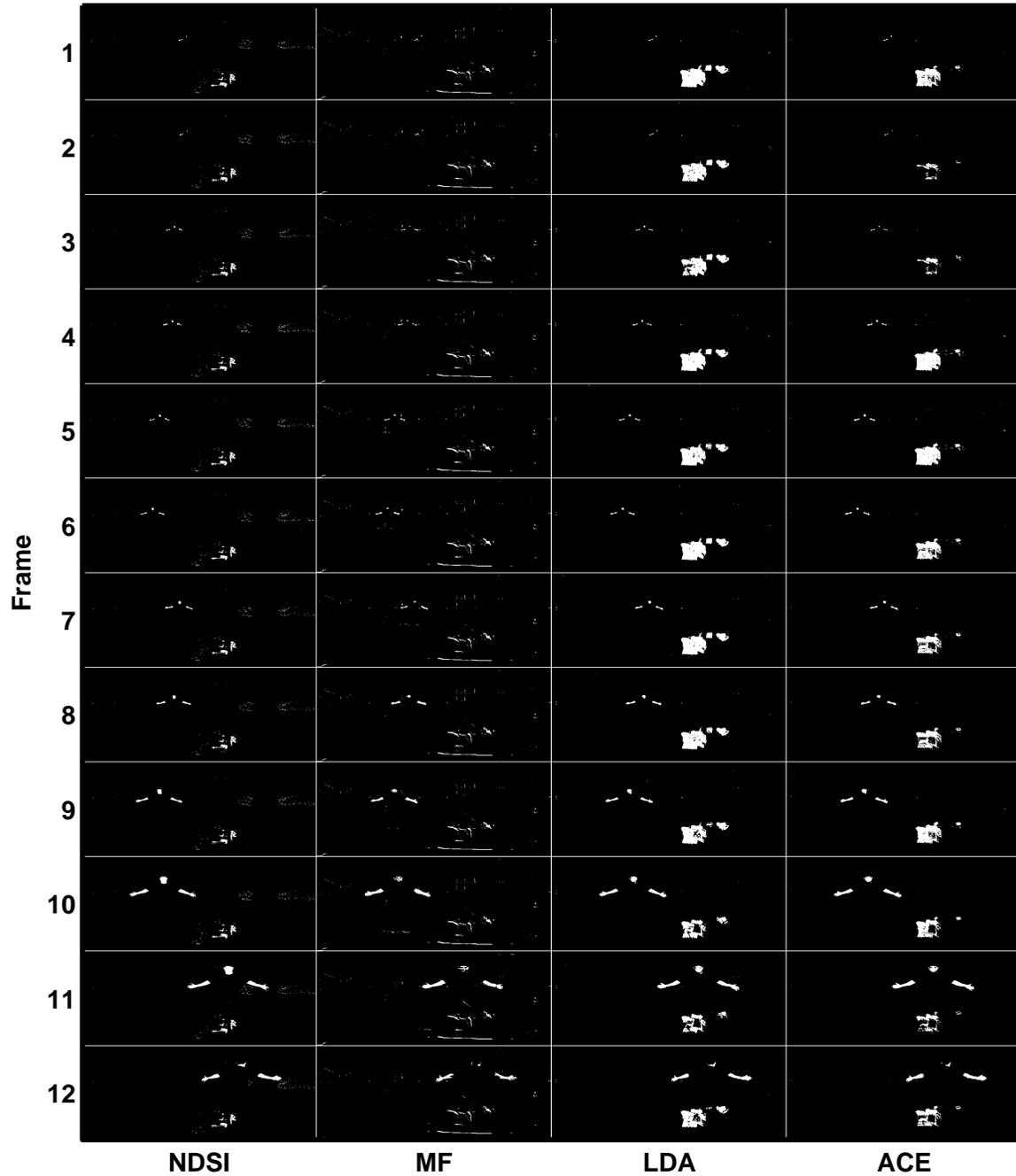
Figure 26. Detection results for the third sequence with constant detection thresholds of 0.25 for NDSI, 0.975 for MF, 37 for LDA, and 0.983 for ACE. The MF method does not require in-scene calibration panels, yet its skin detection results look similar to the NDSI method, which does require the panels.

Sequence 4
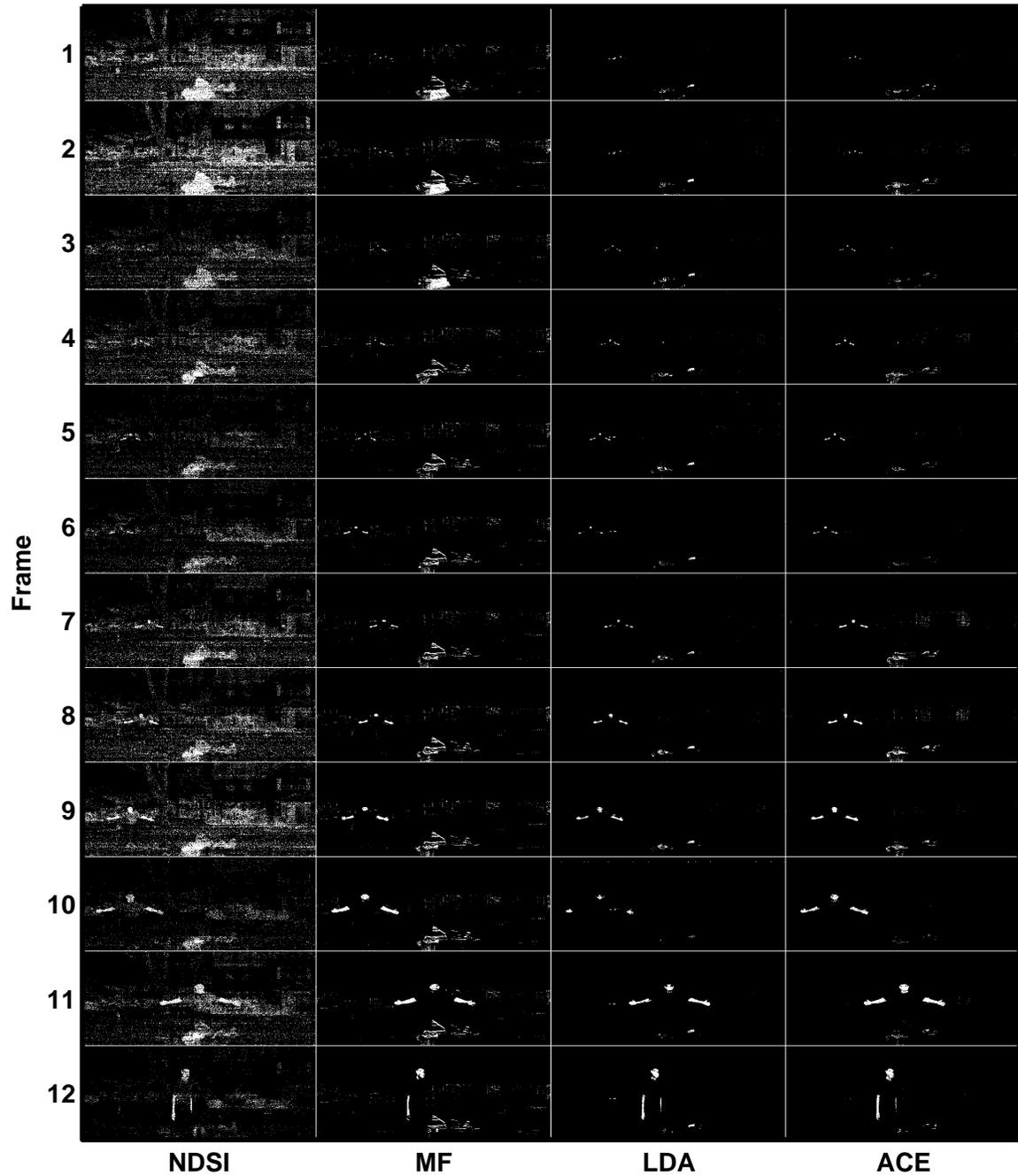
Figure 27. Detection results for the fourth sequence with constant detection thresholds of 0.25 for NDSI, 0.975 for MF, 37 for LDA, and 0.983 for ACE. In these cloudy scenes, the illumination-invariant methods outperform the reflectance-based NDSI method.

the pink t-shirt and on a piece of cardboard above the t-shirt. This also occurs as a cloud momentarily blocks the sun in the first sequence (Fig. 17). A passing cloud changes the scene illumination in frame 3 and 4, causing false alarms in the MF method. Conversely, the false alarm rate decreases for the ACE method in the same two frames.

The performance of the four skin detection methods are quantitatively compared in ROC curves in Figs. 28, 29, 30, and 31. These ROC curves reveal a few trends. One is that the false-alarm rate for the NDSI method increases in the cloudy fourth sequence. Also, the LDA and ACE methods have a higher detection rate than the NDSI method at false-alarm rates above $P_{FA} = 0.01$, but the MF method does not. The MF method only occasionally outperforms the NDSI method.

The Area Under Curve (AUC) for the ROC curves are plotted in Fig. 32. As the name suggests, the AUC is the area under the ROC curve. It represents the average P_D across all P_{FA} between zero and one. A high AUC is a quality of a good detector. Across all four sequences, the LDA and ACE methods have a high AUC. It remains high when the illumination changes. The NDSI method also has a high AUC, but not in the cloudy fourth sequence. The MF method consistently has the lowest AUC.

Another trend in the ROC curves is that the LDA method has a lower detection rate than the ACE method at very low false-alarm rates. The ACE method is better than the LDA method at false-alarm suppression. Furthermore, in sunny scenes when the NDSI method performs well, it has a higher detection rate when $P_{FA} < 0.01$ than any of the illumination-invariance methods. The NDSI method was designed to suppress false alarms, and it does this better then the other methods when the scene is brightly illuminated. The illumination-invariant methods were designed to detect skin pixels under a wide variety of illumination conditions. The ACE and LDA methods accomplish this by having a high detection rate even as the illumination changes.
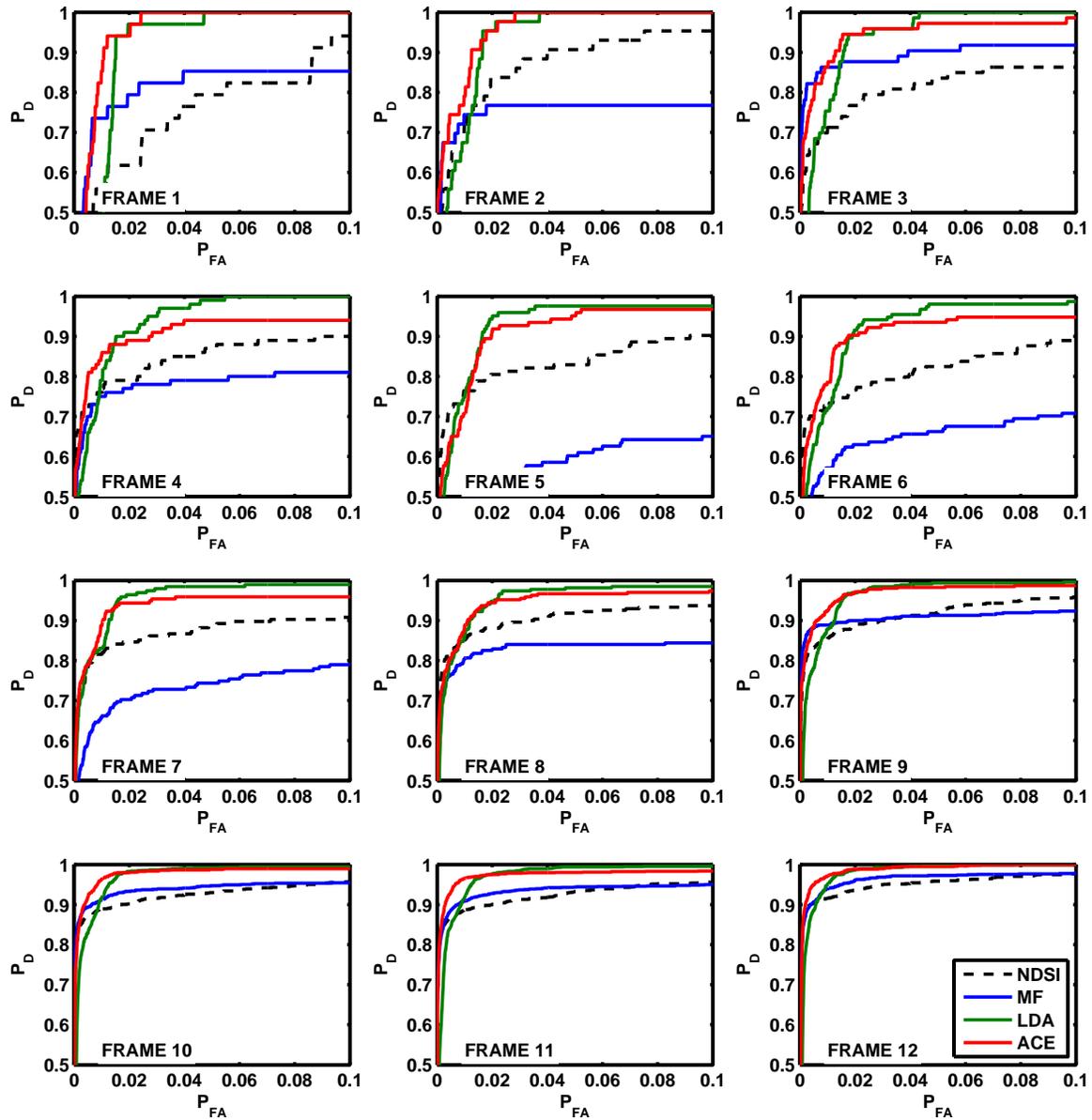
Figure 28. ROC curves comparing the NDSI, MF, LDA, and ACE detection methods on the images in sequence 1.
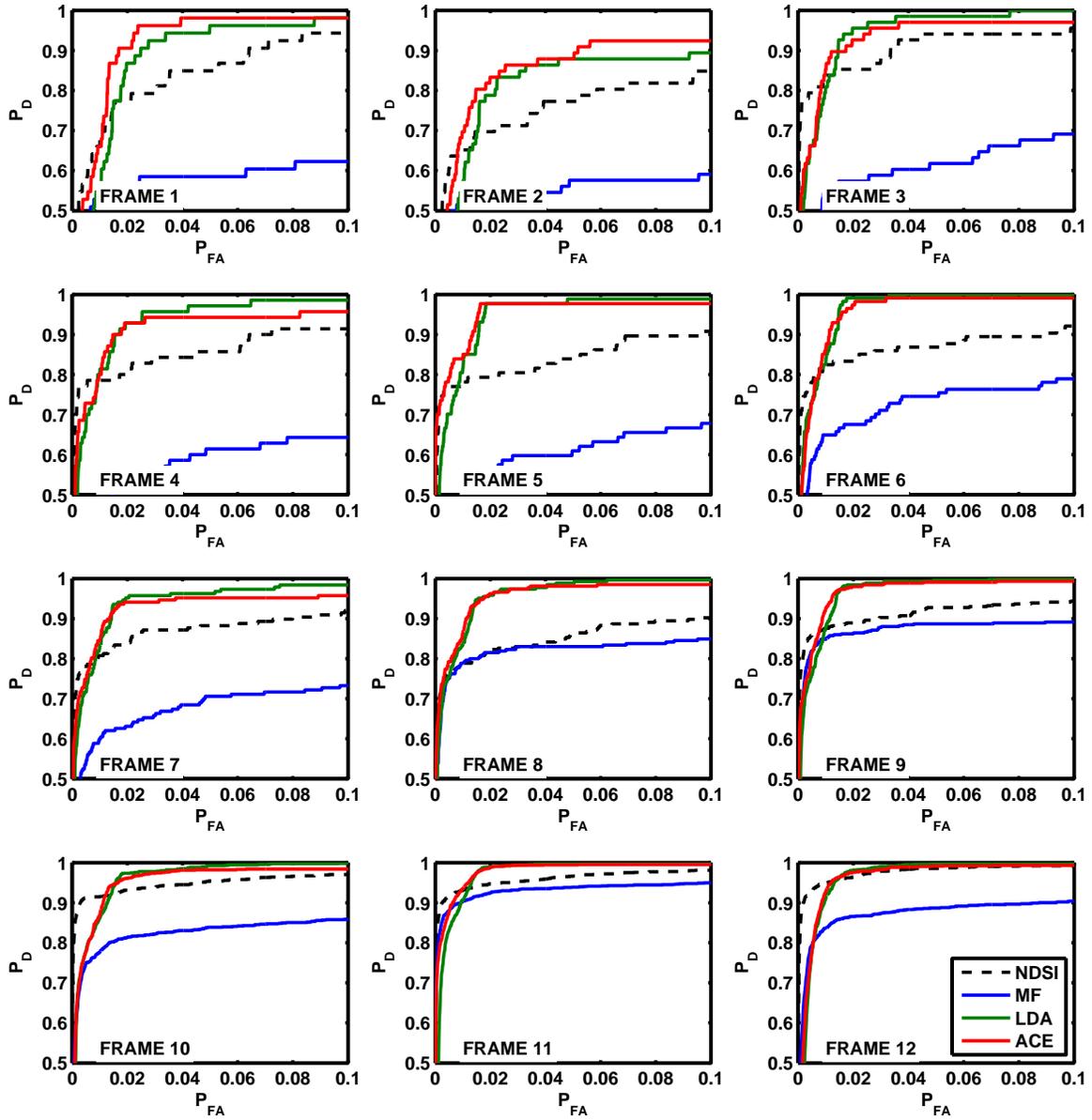
Figure 29. ROC curves comparing the NDSI, MF, LDA, and ACE detection methods on the images in sequence 2.
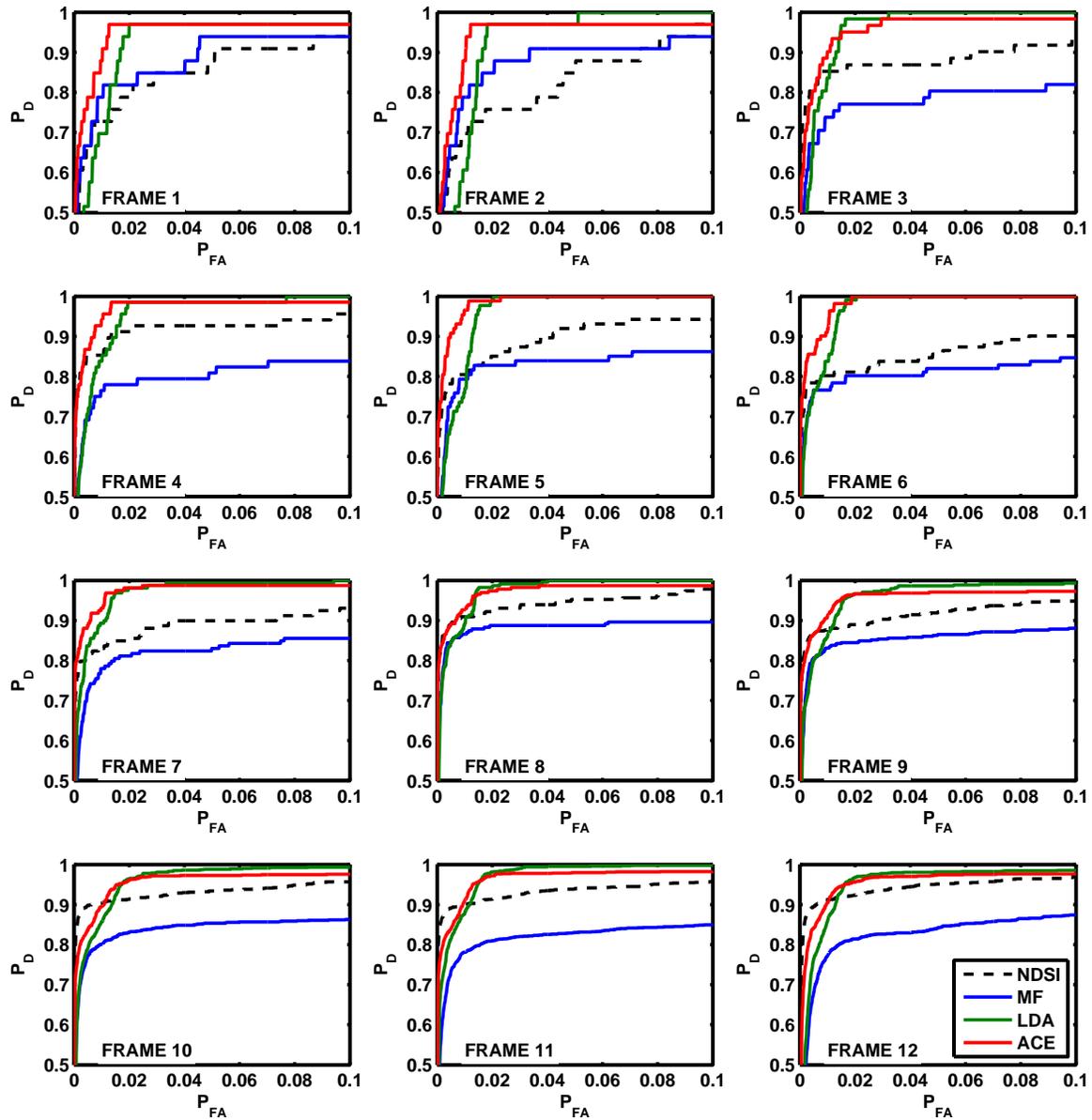
Figure 30. ROC curves comparing the NDSI, MF, LDA, and ACE detection methods on the images in sequence 3.
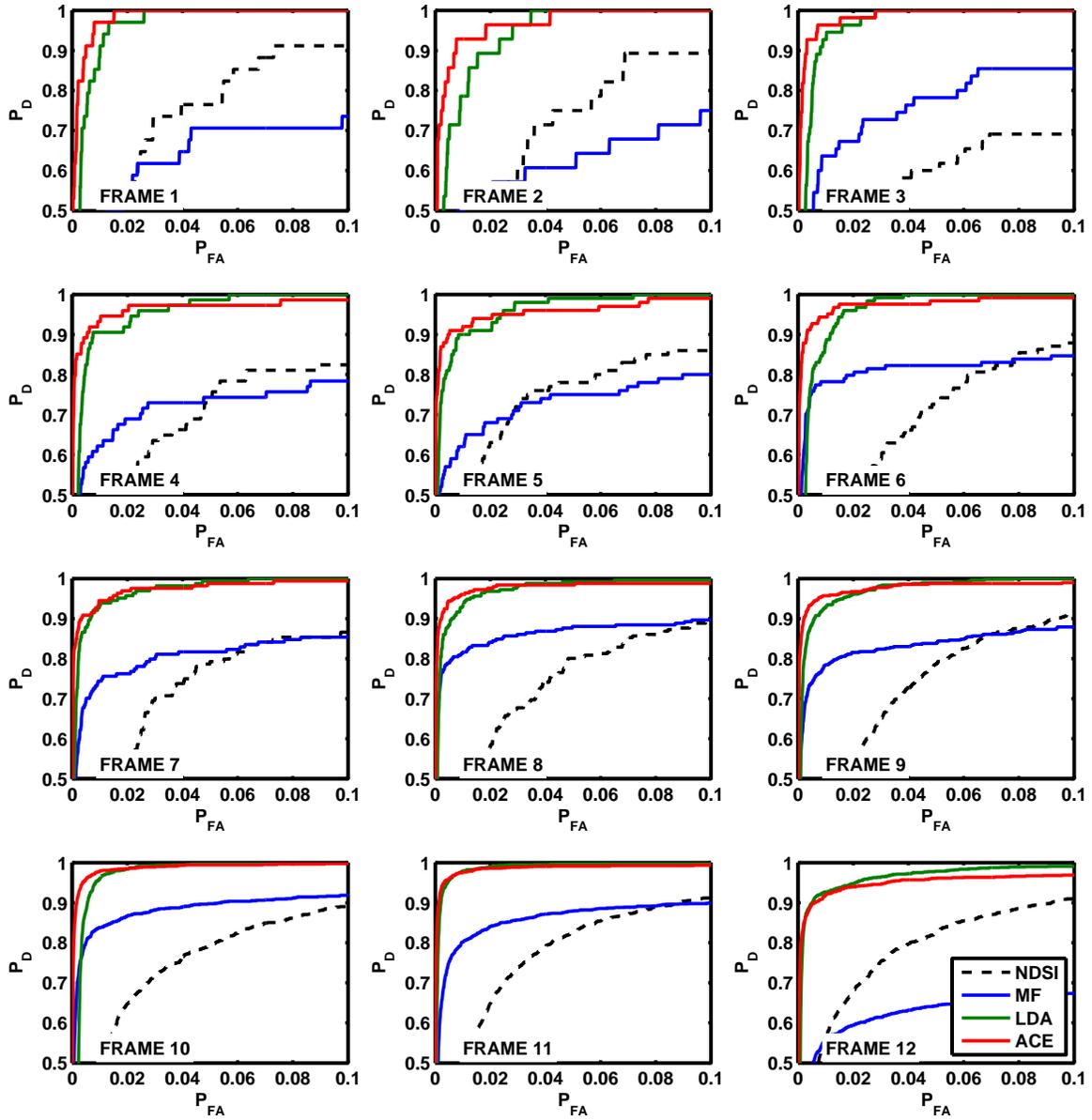
Figure 31. ROC curves comparing the NDSI, MF, LDA, and ACE detection methods on the images in sequence 4.
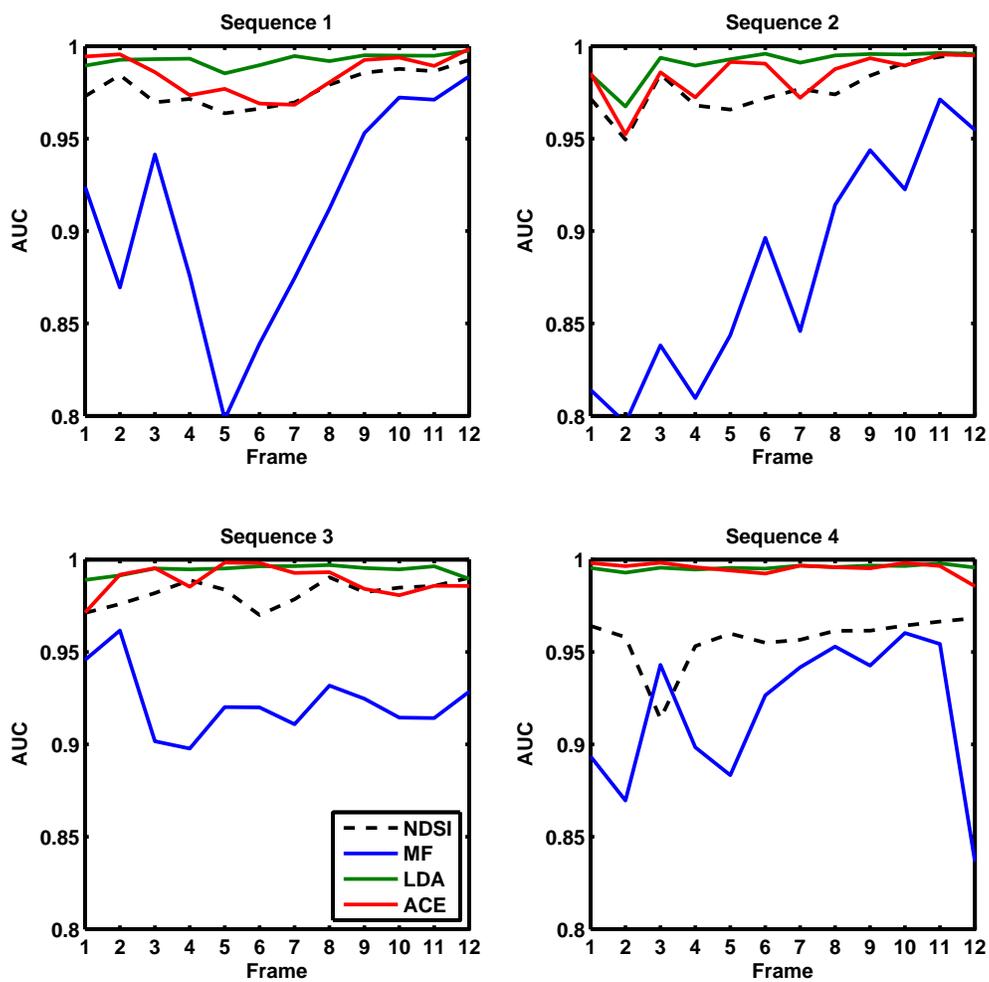
Figure 32. AUC comparison of the ROC curves in Fig. 28, 29, 30, and 31. The LDA and ACE methods have consistently high AUCs across different illumination conditions, while that of the NDSI method drops in the cloudy fourth sequence.

4.7 Summary

This chapter demonstrates the results of the using the target signature generation process outlined in Chapter 3. The MODTRAN simulations of normalized illumination transforms follow a general pattern of high magnitude in VIS and low magnitude in SWIR with atmospheric attenuation bands characteristic of natural illumination. The estimated illumination transform is compared to the measured illumination transform from four sequences of hyperspectral images, and it has the lowest error in the cloudy scenes. This estimated illumination transform is used to generate a target signature from the measured spectral reflectance of the dismount's exposed skin. The target signature has low magnitude at SWIR wavelengths where both the illumination transform and the spectral reflectance of human skin are low.

The reflectance-based skin detection method from [37] is modified to produce a scalar score for each pixel, so that it can be thresholded to form a ROC curve. ROC curves are used to evaluate it against the illumination-invariant methods.

The detection results of all four detection methods are qualitatively evaluated by comparing their ability to suppress false alarms at a constant detection rate. The modified NDSI method produces low false-alarm rates ($P_{FA} < 0.001$) in sunny scenes, but has a higher false-alarm rate than the others ($P_{FA} > 0.05$) in the cloudy scene. The illumination-invariant methods did not need the calibration panels in the scene like the NDSI method, and yet they achieved comparable skin detection. The ACE and LDA methods show a pattern of increasing the false-alarm rates as the illumination error increases, but their average false-alarm rates are still lower than the MF method.

The four methods are quantitatively compared by observing their detection performance across each frame while the detection thresholds are held constant. This demonstrates what would be seen in a real-time dismount detection scenario. The

LDA and ACE methods are able to maintain consistent average detection rates, as characterized by their AUC value, even as the illumination conditions of the scene change from sunny to cloudy.

5. Conclusions and Future Work

This chapter concludes the thesis by summarizing the important points and proposing ideas for further research on this topic. The process of characterizing a hyperspectral pixel and predicting a skin pixel from a simulated illumination transform is described in Section 5.1. Then a few possible ideas for future research in real-time hyperspectral dismount detection are discussed in Section 5.2. Finally, the main contributions of this thesis to the ATR community are explained in Section 5.3.

5.1   Summary of Methods and Conclusions

The goal of this thesis is to shorten the dismount detection algorithm chain by identifying a method that does not require in-scene calibration panels for radiometric calibration. The new method should be robust in different illumination conditions, and achieve similar detection performance to the reflectance-based method that requires the calibration panels.

This goal is met by understanding the appearance of skin in hyperspectral imagery and predicting its appearance across a wide range of different illumination conditions. A hyperspectral image pixel is expressed as a function of an object's spectral reflectance, the scene illumination, and the spectral response of the hyperspectral camera. In vector notation, this expression reduces to a transform between a reflectance vector, which is an object's spectral reflectance sampled at the center wavelengths of the sensor's channel filters, and a hyperspectral pixel. The transform is referred to as an illumination transform because it accounts for the effects of the illumination conditions on the appearance of an object in the scene. It is simulated in MODTRAN over 1728 different scenarios by independently varying a set of input parameters. The resulting set of illumination transforms reveals that the spectral

distribution of natural daylight follows a predictable pattern and can be adequately estimated by taking the mean of that set.

The mean illumination transform is used to convert a measured skin reflectance vector into an estimated skin signature. This skin pixel estimate is used by three different detection methods on a set of test images. They are considered illumination-invariant detection methods because they use a skin pixel estimate that does not change between scenes. They are evaluated against the NDSI skin detection method, which is a reflectance-based method that requires a pair of calibration panels in each scene in order to perform ELM correction.

The Matched Filter achieves good false-alarm suppression in some scenes, with a P_{FA} as low as 0.0007 in Fig. 17, but it has a lower average skin pixel detection rate (AUC < 0.95) than the other methods. In sunny scenes, the Adaptive Coherence Estimator and the method based on Linear Discriminant Analysis have a higher skin pixel detection rate ($P_D > 0.95$) than the NDSI method when $P_{FA} > 0.02$, but have a much lower detection rate than the NDSI method when $P_{FA} < 0.01$. In cloudy scenes, however, the illumination-invariant skin detection methods have noticeably lower false-alarm rates ($P_{FA} < 0.01$) than the NDSI method when the detection rate is set to $P_D = 0.8$.

5.2   Future Work

This section proposes further areas of research that are pertinent to hyperspectral dismount detection.

5.2.1   Combine Skin and Illumination Covariance

A hyperspectral target detection method called the Kelly detector accounts for variations in the target's spectral reflectance [32]. In the Kelly detector, the spectral

reflectance of the target is expressed as a random vector with a known mean and covariance. After the image is converted into estimated reflectance vectors, the mean target vector is subtracted from all pixels, which are then divided by the target covariance. Target pixel vectors are detected by selecting the pixels with low absolute magnitudes at the end of this process.

This detector can be applied to dismount detection by using human skin as the target signature. Different skin types have different spectral reflectances, as pointed out in [37]. If one were to express the covariance of human skin's spectral reflectance, then a skin detection algorithm could be created that is invariant to skin type. It would detect fair skin and dark skin with the same likelihood, making it useful for scenarios when a potential dismount may be of unknown ethnicity.

Furthermore, the covariance of the normalized illumination transform used to generate a target radiance signature could also be derived and included in the detector. The MODTRAN-simulated illumination transforms in Fig. 9 could serve as a starting point for deriving the covariance of a natural, normalized illumination transform. The detector could account for variability in the target's spectral reflectance and the scene illumination to determine the expected variability of the target signature. This could result in a dismount detector that is more robust in different illumination conditions and more diverse in the types of skin that it detects.

5.2.2 Common False-Alarm Sources

Whether employing the NDSI method or the illumination-invariant methods from Chapter 4 for dismount detection, there are some materials that are more likely than others to induce false alarms. Certain types of vegetation, for example, cause high false-alarm rates with the NDSI detector. The pink t-shirt and the red brick in Fig. 18 cause false alarms for the LDA and ACE detectors. These materials are misclassified

because they have a similar spectral signature to human skin.

Future work could identify common sources of false alarms in dismount detection algorithms by looking for spectrally similar materials. It is important to know if some detection algorithms are more sensitive than others to certain false alarm sources. This would help predict which algorithms would be most suitable for different settings. A dismount detector that is sensitive to bricks and other synthetic materials should not be used in an urban environment. A dismount detector that commonly mistakes tree bark for human skin should not be used in a forest. Spectral databases like the USGS spectral library in [8] could offer a starting point to find common false-alarm sources for dismount detectors.

5.2.3   Band Selection for Skin Detection

The NDSI skin detection method with NDGRI false-alarm suppression only requires four spectral bands [37]. Implementing the method on a full hyperspectral camera with hundreds of spectral bands is a waste of sensor capability. A specialized multispectral camera responsive to only the required spectral bands can perform dismount detection more efficiently [40]. In order to justify special multispectral cameras for dismount detection, those four bands should be tested against other possible sets of spectral bands for optimality.

Band selection methods search for the optimal set of spectral bands for a specific target signature and detection algorithm [7, 26, 52]. They incorporate information about the target-to-background contrast and the availability of natural illumination at each spectral band. It is possible that there exists a three-band combination that could provide dismount detection results comparable to the NDSI method. There could also be a five-band combination that detects dismounts so well that it is worth the cost of the additional spectral channel.

The number of spectral bands is one factor that drives the cost of multispectral cameras. A band selection study could find the optimal dismount detection performance for a given number of spectral bands.

5.2.4   Band Width vs Feature Retention

Similar to the band selection study suggested in Section 5.2.3, the appropriate band widths should also be considered for optimal target detection. When designing a multispectral detection system, the width of each band should be considerate of two opposing factors. On one hand, the bands should be narrow enough to retain the spectral features that distinguish the target from the background. On the other hand, the bands should be wide enough to allow plenty of light energy to pass through and increase the SNR. There may be valuable features at wavelengths that receive very weak natural illumination. It would be acceptable to widen the bands at such wavelengths.

The four bands used in the NDSI skin detection method may yield better performance if they were wider. A trade off study could find the optimal band widths at these wavelengths for maximizing the average skin detection performance.

5.2.5   Incorporate Image Processing Techniques

The spectral detection methods discussed in this study evaluate each image pixel independently, without considering their relation to other pixels in the image. Image processing techniques like morphological filters and median filters can correct many of the false alarms generated by spectral detection methods. A single stray pixel that is marked as skin can be suppressed if its neighboring pixels are not skin. This might limit the effective range of the dismount detector, but it would ensure a lower false alarm rate.

Additionally, a real-time dismount detection camera can compare the skin pixels between consecutive frames to filter out false alarms caused by noise. A temporal low-pass filter can suppress stray skin pixels that appear and disappear between frames. Combining spectral, spatial, and temporal techniques could yield better dismount detection performance than each of them individually.

5.3 Contributions

This thesis has shown that, with only a slight decrease in performance, dismount detection can be applied in real-time to hyperspectral data without using calibration panels, even as the illumination in the scene changes. By incorporating some assumptions about the expected operating conditions of a real-time dismount detection system, the effects of changing scene illumination can be mitigated. This development advances hyperspectral dismount detection from only being applicable in stationary scenes with calibration panels, to being implemented on a snapshot HSI camera as it pans across a scene in unknown illumination conditions. Detecting human skin without performing radiometric calibration moves real-time hyperspectral dismount detection towards becoming a deployable capability.

Bibliography

[1] Adler-Golden, S.M., R.Y. Levine, M.W. Matthew, S.C. Richtsmeier, L.S. Bernstein, J. Gruninger, G. Felde, M. Hoke, G.P. Anderson, and A. Ratkowski. "Shadow-insensitive material detection/classification with atmospherically corrected hyperspectral imagery". *Proc. SPIE*, volume 4381, 460–469. 2001.

[2] Albiol, A., L. Torres, and E.J. Delp. "Optimum color spaces for skin detection". *Proc. International Conference on Image Processing*, volume 1, 122–124. Oct 2001.

[3] Anderson, T.W. and R.R. Bahadur. "Classification into two multivariate normal distributions with different covariance matrices". *The Annals of Mathematical Statistics*, 420–431, 1962.

[4] Berk, A., G.P. Anderson, P.K. Acharya, L.S. Bernstein, L. Muratov, J. Lee, M.J. Fox, S.M. Adler-Golden, J.H. Chetwynd, M.L. Hoke, et al. "MODTRAN 5: a reformulated atmospheric band model with auxiliary species and practical multiple scattering options". *Proc. SPIE*, volume 5655, 88–95. 2004.

[5] Brand, J. and J.S. Mason. "A comparative assessment of three approaches to pixel-level human skin-detection". *Proc. 15th International Conference on Pattern Recognition*, volume 1, 1056–1059. Sep 2000.

[6] Brooks, Adam L. *Improved Multispectral Skin Detection and Its Application to Search Space Reduction For Dismount Detection Based on Histograms of Oriented Gradients*. Master's thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 2010.

[7] Chen, Y., N.M. Nasrabadi, and T.D. Tran. "Sparse representation for target detection in hyperspectral imagery". *IEEE Journal of Selected Topics in Signal Processing*, 5(3):629–640, 2011.

[8] Clark, R.N., G.A. Swayze, R.A. Wise, K.E. Livo, T.M. Hoefen, R.F. Kokaly, S.J. Sutley, RN Clark, GA Swayze, R. Wise, et al. *USGS Digital Spectral Library*. US Geological Survey, 2000.

[9] Climer, J. *Overcoming Pose Limitations of a Skin-Cued Histograms of Oriented Gradients Dismount Detector Through Contextual Use of Skin Islands and Multiple Support Vector Machines*. Master's thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 2011.

[10] Cooke, T. and M. Peake. "The optimal classification using a linear discriminant for two point classes having known mean and covariance". *Journal of multivariate analysis*, 82(2):379–394, 2002.

[11] Fiorino, S.T., R.J. Bartell, M.J. Krizo, G.L. Caylor, K.P. Moore, T.R. Harris, and S.J. Cusumano. "A first principles atmospheric propagation & characterization tool: the laser environmental effects definition and reference (LEEDR)". *Proc. of SPIE*, volume 6878, 68780B. 2008.

[12] Fisher, R.A. "The use of multiple measurements in taxonomic problems". *Annals of Human Genetics*, 7(2):179–188, 1936.

[13] Fleck, M.M., D.A. Forsyth, and C. Bregler. "Finding Naked People". *Proc. 4th European Conference on Computer Vision*, volume 2, 593–602. Springer-Verlag, 1996.

[14] Gao, B.C. "NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space". *Remote sensing of environment*, 58(3):257–266, 1996.

[15] Gao, B.C., K.B. Heidebrecht, and A.F.H. Goetz. "Derivation of scaled surface reflectances from AVIRIS data". *Remote Sensing of Environment*, 44(2-3):165–178, 1993.

[16] Grimm, D.C., D.W. Messinger, J.P. Kerekes, and J.R. Schott. "Hybridization of hyperspectral imaging target detection algorithm chains". *Proc. SPIE*, volume 5806, 753. 2005.

[17] Haavardsholm, T.V., T. Skauli, and I. Kasen. "A physics-based statistical signature model for hyperspectral target detection". *IGARSS: International Geoscience and Remote Sensing Symposium*, 3198–3201. Jul 2007.

[18] Healey, G. and D. Slater. "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions". *IEEE Transactions on Geoscience and Remote Sensing*, 37(6):2706–2717, 1999.

[19] Hersey, R.K., W.L. Melvin, and E. Culpepper. "Dismount modeling and detection from small aperture moving radar platforms". *IEEE Radar Conference*, 1–6. May 2008.

[20] Hsu, R.L., M. Abdel-Mottaleb, and A.K. Jain. "Face detection in color images". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):696–706, 2002.

[21] Ientiluccia, E.J., S. Matteolib, and J.P. Kerekesa. "Tracking of vehicles across multiple radiance and reflectance hyperspectral datasets". *Proc. SPIE*, volume 7334, 73340A–1.

[22] Ioffe, S. and D.A. Forsyth. "Probabilistic methods for finding people". *International Journal of Computer Vision*, 43(1):45–68, 2001.

[23] Jengo, C.M. and J. LaVeigne. "Sensor performance comparison of HyperSpecTIR instruments 1 and 2". *Proc. IEEE Aerospace Conference*, volume 3. Mar 2004.

[24] Jin, S. and S.A. Sader. "Comparison of time series tasseled cap wetness and the normalized difference moisture index in detecting forest disturbances". *Remote Sensing of Environment*, 94(3):364–372, 2005.

[25] Karpouzli, E. and T. Malthus. "The empirical line method for the atmospheric correction of IKONOS imagery". *International Journal of Remote Sensing*, 24(5):1143–1150, 2003.

[26] Keshava, N. "Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries". *IEEE Transactions on Geoscience and Remote Sensing*, 42(7):1552–1565, 2004.

[27] Koch, B. *A multispectral bidirectional reflectance distribution function study of human skin for improved dismount detection.* Master's thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 2011.

[28] Kraut, S., L.L. Scharf, and R.W. Butler. "The adaptive coherence estimator: A uniformly most-powerful-invariant adaptive detection statistic". *IEEE Transactions on Signal Processing*, 53(2):427–438, 2005.

[29] Licata, W.H. "Automatic Target Recognition (ATR) Beyond the Year 2000". *Technologies for Future Precision Strike Missile Systems*, volume 1, 2001.

[30] Maier, M.J. *Estimating Anthropometric Marker Locations from 3-D LADAR Point Clouds.* Master's thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 2011.

[31] Manolakis, D., R. Lockwood, T. Cooley, and J. Jacobson. "Is there a best hyperspectral detection algorithm?" *Proc. SPIE*, volume 7334, 733402. 2009.

[32] Manolakis, D., D. Marden, and G.A. Shaw. "Hyperspectral image processing for automatic target detection applications". *Lincoln Laboratory Journal*, 14(1):79–116, 2003.

[33] Matteoli, S., E.J. Ientilucci, and J.P. Kerekes. "Comparison of radiative transfer in physics-based models for an improved understanding of empirical hyperspectral data". *WHISPERS: First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 1–4. IEEE, 2009.

[34] Matthew, M.W., S.M. Adler-Golden, A. Berk, G.W. Felde, G.P. Anderson, D. Gorodetzky, S.E. Paswaters, and M. Shippert. "Atmospheric correction of spectral imagery- Evaluation of the FLAASH algorithm with AVIRIS data". *Proc. SPIE*, volume 5093, 474–482. 2003.

[35] Moore, B. "Principal component analysis in linear systems: Controllability, observability, and model reduction". *IEEE Transactions on Automatic Control*, 26(1):17–32, 1981.

[36] Niyogi, S.A. and E.H. Adelson. "Analyzing and recognizing walking figures in XYT". *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 469–474. Jun 1994.

[37] Nunez, Abel S. *A Physical Model of Human Skin and Its Application for Search and Rescue.* Ph.D. thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 2009.

[38] Nunez, A.S. and M.J. Mendenhall. "Detection of human skin in near infrared hyperspectral imagery". *IGARSS: International Geoscience and Remote Sensing Symposium*, volume 2, 621. Jul 2008.

[39] Nunez, A.S., M.J. Mendenhall, and K. Gross. "Melanosome level estimation in human skin from hyperspectral imagery". *WHISPERS: First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 1–4. Aug 2009.

[40] Peskosky, Kieth R. *Design of a Monocular Multi-Spectral Akin Detection, Melanin Estimation, and False Alarm Supression System.* Master's thesis, Graduate School of Engineering, Air Force Institute of Technology (AETC), Wright-Patterson AFB OH, 2010.

[41] Qu, Z., B.C. Kindel, and A.F.H. Goetz. "The high accuracy atmospheric correction for hyperspectral data (HATCH) model". *IEEE Transactions on Geoscience and Remote Sensing*, 41(6):1223–1231, 2003.

[42] Reed, I.S. and X. Yu. "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution". *IEEE Transactions on Acoustics, Speech and Signal Processing*, 38(10):1760–1770, 1990.

[43] Rouse, J.W., R.H. Haas, J.A. Schell, and D.W. Deering. "Monitoring vegetation systems in the great plains with ERTS". *Earth Resources Technology Satellite-1 Symposium*, volume 3, 309–317. 1973.

[44] Salomonson, V.V. and I. Appel. "Estimating fractional snow cover from MODIS using the normalized difference snow index". *Remote Sensing of Environment*, 89(3):351–360, 2004.

[45] Scharf, L.L. and L.T. McWhorter. "Adaptive matched subspace detectors and adaptive coherence estimators". *Conference Record of the 30th Asilomar Conference on Signals, Systems and Computers*, 1114–1117. Nov 1996.

[46] Schott, John R. *Remote Sensing The Image Chain Approach.* Oxford University Press, New York, New York, 2007.

[47] Sigal, L., S. Sclaroff, and V. Athitsos. "Estimation and prediction of evolving color distributions for skin segmentation under varying illumination". *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, 152–159. IEEE, 2000.

[48] Smith, G.M. and E.J. Milton. "The use of the empirical line method to calibrate remotely sensed data to reflectance". *International Journal of Remote Sensing*, 20(13):2653–2662, 1999.

[49] Soriano, M., B. Martinkauppi, S. Huovinen, and M. Laaksonen. "Skin detection in video under changing illumination conditions". *Proc. 15th International Conference on Pattern Recognition*, volume 1, 839–842. Sep 2000.

[50] Stein, D.W.J., S.G. Beaven, L.E. Hoff, E.M. Winter, A.P. Schaum, and A.D. Stocker. "Anomaly detection from hyperspectral imagery". *Signal Processing Magazine*, 19(1):58–69, 2002.

[51] Stevenson, B., R. O'Connor, W. Kendall, A. Stocker, W. Schaff, D. Alexa, J. Salvador, M. Eismann, K. Barnard, and J. Kershenstein. "Design and performance of the Civil Air Patrol Archer hyperspectral processing system". *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI*, 5806:731–742, June 2005.

[52] Theiler, J. and K. Glocer. "Sparse linear filters for detection and classification in hyperspectral imagery". *Proc. SPIE*, volume 6233.

[53] Thiemann, S. and H. Kaufmann. "Determination of chlorophyll content and trophic state of lakes using field spectrometer and IRS-1C satellite data in the Mecklenburg Lake District, Germany". *Remote Sensing of Environment*, 73(2):227–235, 2000.

[54] Vezhnevets, V., V. Sazonov, and A. Andreeva. "A survey on pixel-based skin color detection techniques". *Proc. Graphicon*, volume 3. 2003.

[55] Wong, G. "Snapshot hyperspectral imaging and practical applications". *Journal of Physics: Conference Series*, volume 178, 012048. Issue 1, 2009.

[56] Yarbrough, A.W., M.J. Mendenhall, and R.K. Martin. "The effects of atmospheric mis-estimation on hyperspectral-based adaptive matched filter target detection as measured by the Bhattacharyya coefficient". *WHISPERS: Second Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 1–4. Jun 2010.

[57] Yarovoy, A.G., L.P. Ligthart, J. Matuzas, and B. Levitas. "UWB radar for human being detection". *Aerospace and Electronic Systems Magazine*, 21(3):10–14, 2006.

[58] Zha, Y., J. Gao, and S. Ni. "Use of normalized difference built-up index in automatically mapping urban areas from TM imagery". *International Journal of Remote Sensing*, 24(3):583–594, 2003.

[59] Zhao, H. and X. Chen. "Use of normalized difference bareness index in quickly mapping bare areas from TM/ETM+". *IGARSS: International Geoscience and Remote Sensing Symposium*, volume 3, 1666–1668. Jul 2005.

[60] Zhou, J. and J. Hoang. "Real time robust human detection and tracking system". *IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, 149. Jun 2005.

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
22–03–2012	Master's Thesis	Aug 2010 — Mar 2012

4. TITLE AND SUBTITLE

Spectral Detection of Human Skin in VIS-SWIR Hyperspectral Imagery Without Radiometric Calibration

5a. CONTRACT NUMBER

-

5b. GRANT NUMBER

-

5c. PROGRAM ELEMENT NUMBER

-

6. AUTHOR(S)

Beisley, Andrew P., 2d Lt USAF

5d. PROJECT NUMBER

-

5e. TASK NUMBER

-

5f. WORK UNIT NUMBER

-

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GE/ENG/12-03

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFRL Human Signatures Branch
David R. Bowden, Electrical Engineer
2800 Q Street, Bldg 824
WPAFB, USA 45433
(937)255-8515
david.bowden@wpafb.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

711 HPW/RHXBA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approval for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

"This material is declared a work of the U.S. Government and is not subject to copyright protection in the United States."

14. ABSTRACT

Many spectral detection algorithms require precise ground truth measurements that are hand-selected in the image to apply radiometric calibration, converting image pixels into estimated reflectance vectors. That process is impractical for mobile, real-time hyperspectral target detection systems, which cannot empirically derive a pixel-to-reflectance relationship from objects in the image. Implementing automatic target recognition on high-speed snapshot hyperspectral cameras requires the ability to spectrally detect targets without performing radiometric calibration. This thesis demonstrates human skin detection on hyperspectral data collected at a high frame rate without using calibration panels, even as the illumination in the scene changes. Compared to an established skin detection method that requires calibration panels, the illumination-invariant methods in this thesis achieve nearly as good detection performance in sunny scenes and superior detection performance in cloudy scenes.

15. SUBJECT TERMS

Hyperspectral, Skin Detection, Automatic Target Recognition

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Jeffrey D. Clark, Lt Col, USAF, AFIT/ENG
U	U	U	UU	113	19b. TELEPHONE NUMBER *(include area code)* (937)255-3636x4614 Jeffrey.Clark@afit.edu