

# Human skin detection in the visible and near infrared

MICHAEL J. MENDENHALL,\* ABEL S. NUNEZ, AND RICHARD K. MARTIN

Department of Electrical & Computer Engineering, Air Force Institute of Technology, WPAFB, Ohio 45433, USA

\*Corresponding author: michael.mendenhall.1@us.af.mil

Received 14 August 2015; revised 31 October 2015; accepted 3 November 2015; posted 3 November 2015 (Doc. ID 247950); published 10 December 2015

Skin detection is a well-studied area in color imagery and is useful in a number of scenarios to include search and rescue and computer vision. Most approaches focus on color imagery due to cost and availability. Many of the visible-based approaches do well at detecting skin (above 90%) but they tend to have relatively high false-alarm rates (8%–15%). This article presents a novel feature space for skin detection in visible and near infrared portions of the electromagnetic spectrum. The features are derived from known spectral absorption of skin constituents to include hemoglobin, melanin, and water as well as scattering properties of the dermis. Fitting a Gaussian mixture to skin and background distributions and using a likelihood ratio test detector, the features presented here show dominating performance when comparing receiver-operating characteristic curves (ROCs) and statistically significant improvement when comparing equal error rates and area under the ROC (AUC). A detection/false-alarm probability of 98.6%/1.1% is achieved for the averaged equal error rate (EER). EER values for the proposed feature space show a 5.6%–11.2% increase in detection probability with a 6.0%–11.6% decrease in false-alarm probability compared to well performing color-based features. The AUC shows a 0.034–0.173 increase in total area under the curve compared to well performing color-based features.

**OCIS codes:** (100.0100) Image processing; (100.5010) Pattern recognition; (280.0280) Remote sensing and sensors; (280.4991) Passive remote sensing; (330.0330) Vision, color, and visual optics; (330.6180) Spectral discrimination.

<http://dx.doi.org/10.1364/AO.54.010559>

## 1. INTRODUCTION

Hyperspectral sensors provide a great deal of spectral granularity and a potential means for improved detection and classification of select materials [1]. Exploitation of these images has proven useful in identifying materials of interest from airborne platforms over a large geographic area to include geologic and biologic surface cover as well as locate anomalous materials such as aircraft debris or in search and rescue (SAR) operations [2–4]. It is noted by experts that any hyperspectral system developed for use in SAR must be simple enough to operate for a non-hyperspectral-exploitation expert [5], the system must be able to discriminate small targets in a large scene [6], and real-time processing is essential [7]. Beyond search and rescue, a system that accurately detects skin is useful in providing the requisite spatial discriminant information for facial and hand-gesture recognition systems [8–11] and helping address the difficulties in their automation caused by varying illumination levels [12]. In reality, the acquisition and necessary post-processing of hyperspectral imagery are often not fast enough to support such real-time applications. A multispectral solution designed to capture the necessary spectra for skin detection and

false-alarm suppression could enable such real-time applications, if those spectra were known.

This article presents a novel two-dimensional (2D) feature space used for detecting skin in a suburban environment and suppressing common false-alarm sources. The feature definition is based on a thorough understanding of visible-to-near-infrared reflectance of human skin gained through analysis of, and experience with, human-tissue modeling. Tissue models include physical structure and the absorption and scattering effects of each constituent component in human skin [13] and provide the necessary insights to build a good skin-detection feature space. The features set shows excellent separability between skin and background pixels and performs the skin-detection task better than features derived from color-imagery alone. This is demonstrated through dominating receiver-operating characteristic (ROC) curves over color-based features. When comparing equal error rates (EER), a detection probability of 98.6% and a false-alarm probability of 1.1% is achieved for the averaged EER. A 5.6%–11.2% increase in detection probability and a 6%–11.6% decrease in false-alarm probability over color imagery-based features is achieved.

Furthermore, the area under the ROC curve shows a 0.034–0.173 increase in total area under.

The remainder of this article is arranged as follows. Section 2 provides a review of current skin-detection literature. Section 3 describes the reflectance of human skin that leads to the definition of the novel 2D feature space presented in this article. The detection methodology is described in Section 4. Section 5 presents the test data and describes required data preprocessing and Section 6 describes the color-image based features used for comparison purposes. Section 7 describes the experiments and presents results. Finally, concluding remarks and areas of future research are discussed in Section 8.

## 2. RELATED WORK

Detection of human skin in color imagery is challenging as many materials have a color similar to one of the many shades of skin. Skin-detection methods in color imagery vary based on manipulating color-space channels [14] to more sophisticated statistical modeling and machine learning methods [14–16]. Regardless of the methodology used, the end result is often a probability of detection ( $P_D$ ) above 90% and a probability of false alarm ( $P_{FA}$ ) on the order of 8%–15% [14,17,18] with the work in [19] as low as 2%.

Color-space channel methods typically use two channels for detection. For example, the full range of skin colors has red-to-green, red-to-blue, and green-to-blue ratios greater than one [18]. These features result in a significant number of false alarms, and some have attempted to reduce them using rule sets that combine ratios, color-space channel thresholds, and differences [16,20]. These methods essentially define a volume of the 3D color-space that encompasses all possible skin colors. Other methods reduce false alarms by examining how skin pixels cluster spatially and then attempt to determine if the spatial clustering resembles a body part [21,22].

Other skin-detection methods use three-channel color-space examples as a training set to train a binary classification system. Some approaches project pixels onto a plane within the color-space that provides the furthest separation between skin and non-skin pixels [18]. This technique in particular provides a slight improvement over the ratio-based methods.

The red-green-blue (RGB) color space is most common in the literature but may not be ideal for the skin-detection task where the primary disadvantage is the lack of separation of luminance from chrominance and the strong correlation between the channels [16]. Color spaces that separate luminance and chrominance often result in a better clustering of skin-colored pixels [23] while [24] (and references therein) shows best results using cylindrical color spaces (hue, saturation, value; hue, saturation, intensity). Although the selection of a color-space may allow for a simpler or more intuitive algorithm, others have shown that an optimum three-channel color-space for skin detection does not exist if the optimal skin detector for that color-space is used [25].

Some statistical-based approaches analyze training images for the probability of skin occurring given a pixel in the image, the probability of a skin color occurring given a pixel is skin, and the probability of skin color occurring in an image overall.

From these quantities the probability a pixel is skin given a specific skin color is calculated from the training images [18].

The same optical parameters that affect the color of human skin in the visible affect its appearance in the near infrared. A useful observation in skin reflectance is that it is high between 800–1100 nm and low beyond 1400 nm, which has been noted and exploited by others. Skin detection using two near-infrared channels is used for the purpose of counting occupants in a vehicle [26] and for face detection [27]. Both works use bands that are several hundred nanometers wide in the near-infrared.

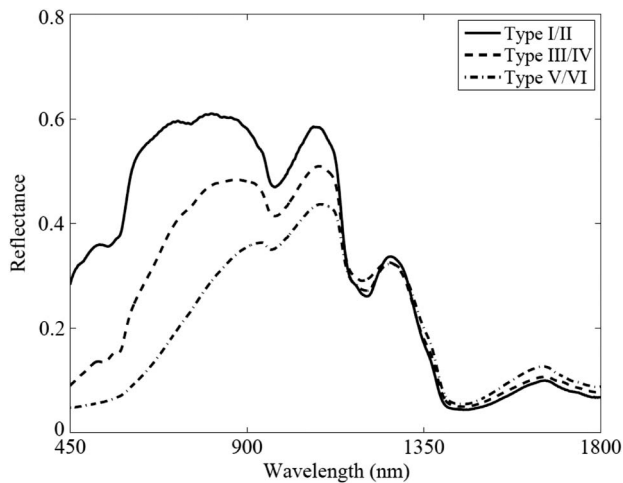
The patent described in [28] exploits the absorptive and reflective properties of skin operating in the range of 800–1400 nm for the lower wavelengths and 1400–2500 nm for the upper wavelengths. Although not explicitly shown in that patent, the authors describe a scaled distance between the upper and lower wavelengths and threshold that scaled distance to declare the presence or absence of skin. This article defines a 2D feature space that is conceptually similar. It is based on the normalized difference skin index (NDSI) defined in [13] and a normalized difference green-red index (NDGRI) used to suppress common false-alarm sources in the natural and urban environments. These two indices are described in detail in Sections 3.B.1 and 3.B.2, respectively.

## 3. 2D FEATURE SPACE FOR SKIN DETECTION AND FALSE-ALARM SUPPRESSION

The following sections describe a 2D feature space used to separate skin pixels from background pixels. One coordinate is used for skin detection while a second coordinate is used for false-alarm suppression. The feature definitions take advantage of *Meade's Maxim*: “You are absolutely unique, just like everyone else” [29]. (This quote is often attributed to Anthropologist Margaret Mead. However, there is no evidence that supports she said it. *Meade's Maxim* appears in a 1979 humor book by Peers [29].) Although the previous quote is intended to be humorous, it applies to the problem at hand. That is to say, although all human-skin tissue is uniquely different at a micro scale, it is essentially the same at a macro scale. As such, skin detection and a false-alarm suppression feature are defined through a fundamental understanding of reflected spectra based on tissue structure and the absorption and scattering of its constituent components. This understanding is used to describe specific absorption features in human skin as a function of its primary chromophores, leading to the skin-detection feature definition. Skin-reflectance measurements are then compared to known false-alarm sources for color-image based skin detection as cited in the literature to demonstrate resiliency to these common pitfalls. False-alarm sources for the skin-detection feature are discussed and the false-alarm suppression feature is described. Finally, separability of skin pixels from background pixels in this new 2D feature space is demonstrated.

### A. Observations from Skin-Reflectance Measurements

A comparison of measured diffuse reflectance spectra of different skin types is shown in Fig. 1. They are labeled based on Fitzpatrick's skin characterization shown in Table 1 [30].

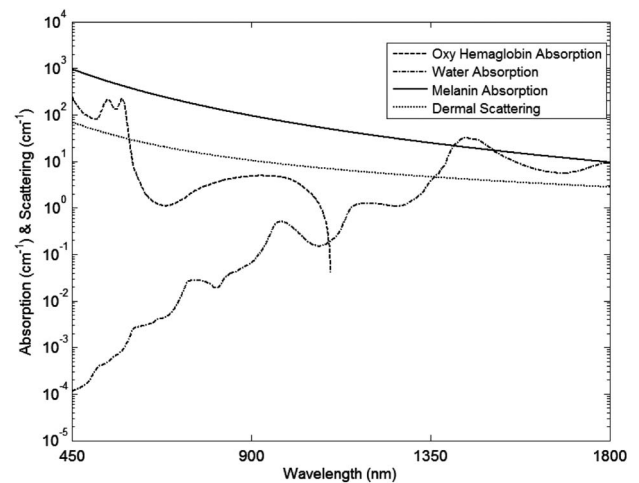


**Fig. 1.** Spectra of Type I/II, Type III/IV, and Type V/VI skin. Spectral measurements were acquired by the authors using analytical spectral devices (ASD) field spectroradiometer (FieldSpec3).

The Fitzpatrick scale declares six numerical skin types where each skin type is characterized by its likelihood of burning. (Note that the likelihood of burning is directly related to the skin's melanin content—more melanin, the darker the skin, the less likely it will burn.)

The skin spectra in Fig. 1 get their spectral shape as a result of a layered structure where each layer has a specified thickness and a baseline scattering and an absorption profile based on collagen and water. Additional absorption is caused by *varying amounts* of the following chromophores: blood, oxygenated and deoxygenated hemoglobin, bilirubin, beta-carotene, and melanin [13] (and references therein). Melanin, oxygenated hemoglobin, and water absorption are the dominate absorbers in the spectra while scattering due to collagen fibers in the epidermis and dermis is the primary scatterer in skin. The absorption and scattering coefficients are shown in Fig. 2 to facilitate discussion on the skin spectra shown in Fig. 1. (Note that approximately 75% of the hemoglobin in blood is oxygenated at all times [31] and therefore deoxygenated hemoglobin is not considered in this analysis.)

Skin is a highly forward scattering material where collagen fibers and water in the epidermis and dermis are primary contributors to the scattering of light. At smaller wavelengths, light is impacted by Mie scattering and as the wavelength increases Rayleigh scattering dominates. This mostly accounts for the increased reflectance of skin as the wavelength increases in



**Fig. 2.** (Left y-axis) Absorption coefficients of melanin [32], oxygenated hemoglobin [33], and water [34,35]. (Right y-axis) Scattering coefficient of skin from [36,37].

the visible. The scattering coefficient is what helps define the maximum possible reflectance of skin at different wavelengths while the absorption of the chromophores is what gives the reflectance spectra of skin its characteristic features.

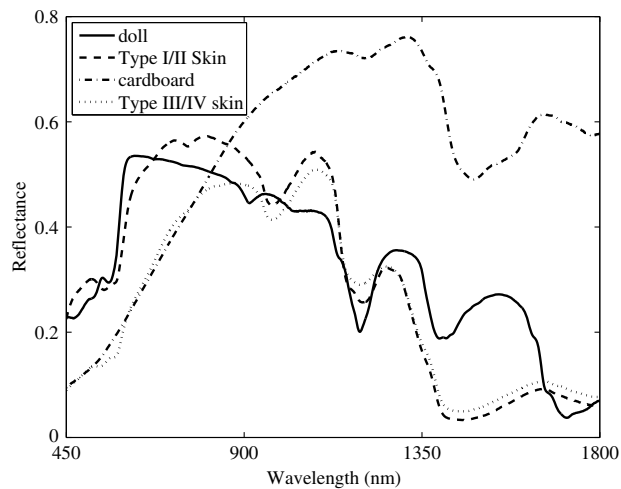
Considering Fig. 2, as melanin content increases (more melanin is present in darker skin), the reflectance of skin in Fig. 1 decreases over the visible and near-infrared portions of the spectrum. As the wavelength increases, the difference between the reflectance of skin with different amounts of melanin decreases due to melanin absorption decreasing as the wavelength increases [32]. Beyond 1300 nm melanin absorption is not significant and skin reflectance is approximately the same [38]. In the visible region (oxygenated) hemoglobin significantly affects the spectrum [39], accounting for the *m*-shaped absorption feature seen in Fig. 2 around 570 nm and the decreased reflectance in the visible up to 600 nm. This *m*-shaped absorption feature results in a *w*-shaped reflectance feature often seen in fair skin (note that it is masked in Type I/II skin in Fig. 1, likely due to the presence of melanin and possibly due to the resolution of the spectro-radiometer). This *w*-shaped reflectance feature becomes insignificant as melanin increases. Water absorption becomes significant in the near infrared accounting for the reduced reflectance beyond 1150 nm, the local maxima at 1080 and 1250 nm, and the local minima at 1200 and 1400 nm [38] as can be seen in Fig. 1. Water absorption becomes so significant at the longer wavelengths that it dominates the scattering affect of skin, dramatically reducing the reflectance of the skin.

Several materials have colors similar to one of the wide varieties of skin tones. Some of this color similarity is by design, such as with mannequins and dolls [40], and others by coincidence, such as brown cardboard, wood, leather, and some metals [41–43]. In some cases the natural environment is rich in colors similar to skin such as a desert with various shades of brown, red, and yellow [44].

A comparison of the diffuse reflectance of Type I/II skin with a plastic flesh-colored doll is shown in Fig. 3. Like skin,

**Table 1. Fitzpatrick Six-Level Characterization of Human Skin Based on Its Likelihood of Burning [30]**

Skin Type	Characteristic
Type I	<i>Always</i> burns
Type II	<i>Usually</i> burns
Type III	<i>Sometimes</i> burns
Type IV	<i>Rarely</i> burns
Type V	<i>Very rarely</i> burns
Type VI	<i>Never</i> burns



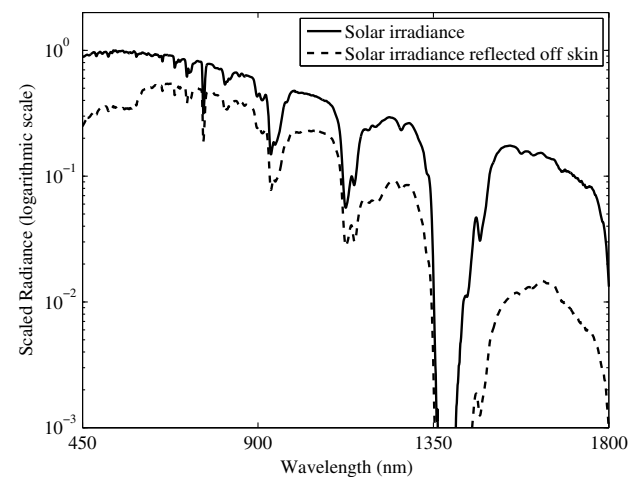
**Fig. 3.** Spectra of Type I/II skin, Type III/IV skin, plastic doll, and brown cardboard. Spectral measurements were acquired by the authors using an ASD FieldSpec3.

the reflectance of the flesh-colored doll rapidly increases as the wavelength increases in the visible portion of the spectrum. Beyond 1200 nm, the reflectance of the flesh-colored doll is significantly greater than skin. The difference between these two measurements is largely due to the water in skin that is not present in the doll. A comparison of the diffuse reflectance of Type III/IV skin with that of brown cardboard is also shown in Fig. 3. Cardboard and Type III/IV skin exhibit an increase in reflectance as the wavelength increases in the visible portion of the spectrum. The reflectance of cardboard remains relatively high while the reflectance of skin is much lower due to water absorption.

These measurements are indicative of skin versus other common false-alarm sources found in visible light-based skin-detection methods (e.g., based on RGB image data). Exploiting the water absorption and melanin absorption features in the near infrared should provide a feature that yields improved skin detection. Due to other materials having water content (such as vegetation), false-alarm mitigation is likely required. Combining both should enable more accurate skin detection while producing fewer false alarms than those approaches that only exploit color-image data.

## B. Feature Definitions

The efficacy of any detection algorithm is based on the quality of the received signal. Since the need is to detect human skin under solar illumination, consideration must be given to the irradiance of the sun through Earth's atmosphere. The irradiance on a sunny day in Dayton, Ohio (scaled so its maximum value is one) is shown in Fig. 4. The water vapor absorption bands, nominally at 1400 (shown) and 1900 nm (not shown), need to be avoided since little to no solar energy reaches the surface of the earth. The object of interest further imposes constraints on the received spectra and hence impacts the wavelengths used in defining the features. As an example, Fig. 4 shows the reflected radiance of Type I/II skin under solar illumination scaled by the same factor as the solar irradiance.



**Fig. 4.** Solar irradiance scaled by the maximum irradiance and the radiance spectra of Type I/II skin (from Fig. 1) illuminated by sunlight scaled by the same maximum irradiance. Data were collected by the authors on a sunny day in Dayton, OH (USA) using an ASD FieldSpec3 and a dome style-reflective cosine receptor for measuring full-sky irradiance.

The local minimums of the skin-reflectance measurement corresponds to water absorption at approximately 950, 1150, and 1400 nm. The spectra beyond 1600 nm is dominated by water absorption. As such, the location of the local minimums and maxima in the near-infrared portion of skin's reflected radiance corresponds to the locations of the local minimums and maxima of skin reflectance as seen in Fig. 3.

Since there is a need for illumination invariance in the detection of skin, the concept of the normalized difference index is used. Normalized difference indices further provide a mechanism to exploit known absorption characteristics in defining the feature. As an example the normalized difference vegetation index (NDVI) [45] is a popular normalized index used in a remote sensing community for the detection and characterization of vegetation. The NDVI exploits the chlorophyll absorption feature and the intense reflection of the so-called vegetation red-edge.

### 1. NDSI for Skin Detection

The NDSI is a function of reflectance at 1080 and 1580 nm. Skin's local maxima at 1080 nm is where melanin absorption dominates. Beyond 1080 nm, water absorption becomes more significant until the local minimum at approximately 1400 nm (known atmospheric water-vapor absorption region). A stable yet low-valued reflectance feature in skin spectra is noted at 1580 nm (beyond the atmospheric water-vapor band). The diffuse skin-reflectance measurements in Fig. 1 show that the difference in reflectance for dark- to light-skin types is fairly large at 1080 versus 1580 nm, which is consistent with reflectance measurements in the literature. Furthermore, according to the measured (and known theoretical) solar-irradiance curves in Fig. 4, a significant amount of solar irradiance reaches the surface of the earth ensuring a strong signal-to-noise ratio (SNR) at that longer wavelength. This ensures that the derivative is large between a melanin-dominated and water-dominated portion of



the spectra. Such conditions make normalized indices good feature choices for the detection problem. The (NDSI) is defined as [13]

$$\gamma^i = \frac{\hat{\rho}_{\lambda_1=1080 \text{ nm}}^i - \hat{\rho}_{\lambda_2=1580 \text{ nm}}^i}{\hat{\rho}_{\lambda_1=1080 \text{ nm}}^i + \hat{\rho}_{\lambda_2=1580 \text{ nm}}^i}, \quad (1)$$

where  $\gamma^i$  is the NDSI value for the  $i$ th pixel, and  $\hat{\rho}$  is the estimated reflectance at wavelength  $\lambda$ .

## 2. NDGRI for False-Alarm Suppression

Since the primary detection feature relies on the water-absorption characteristics of skin, there must be concern regarding in-scene objects that also contain a significant amount of water (e.g., vegetation and water itself). In early experiments (using image data as well as hyperspectral measurements from the United States Geological Survey spectral library [46]), the most notable false-alarm sources were heavy water-bearing objects such as vegetation (specifically conifers) and highly forward scattering water masses (e.g., snow and murky water). In order to reduce the impact of common false-alarm sources, the fact that hemoglobin absorption causes skin to be more red than green (see Fig. 1) is exploited. The red-green relationship is inverted with the NDGRI. The specific bands are chosen to coincide with specific absorption and reflectance features associated with skin and green vegetation. More specifically, chlorophyll in vegetation causes green light to reflect with local maxima at 550 nm and the absorption of visible red light at 660 nm (based on the observation of green-vegetation spectra in [46]). Skin has the  $w$ -shaped reflection ( $m$ -shaped absorption) with local minimums at 540 nm (more readily seen in fair skin where dark skin already strongly absorbs at 540 nm) and a strong reflection at 660 nm regardless of skin tone. Since the minimums at 540 nm due to hemoglobin in skin is near the maxima in vegetation due to chlorophyll, it is chosen for the green portion of the NDGRI. Since vegetation has a characteristic absorption at 660 nm and skin has a strong reflection at that wavelength, it is chosen for the red portion of the NDGRI. The NDGRI is defined as

$$\beta^i = \frac{\hat{\rho}_{\lambda_1=540 \text{ nm}}^i - \hat{\rho}_{\lambda_2=660 \text{ nm}}^i}{\hat{\rho}_{\lambda_1=540 \text{ nm}}^i + \hat{\rho}_{\lambda_2=660 \text{ nm}}^i}, \quad (2)$$

where  $\beta^i$  is the NDGRI value for the  $i$ th pixel and  $\hat{\rho}$  is the estimated reflectance at wavelength  $\lambda$ .

## 4. SKIN DETECTION USING THE LIKELIHOOD RATIO TEST

Although there are several detection schemes that may be employed for a 2D feature space, the likelihood ratio test (LRT) is used in this work. Given two probability distribution functions, one for the object of interest and one for the background, the LRT minimizes the Bayes Risk and provides flexibility for choosing a detection threshold to meet specific criterion (e.g., equal error rate or fixed probability of detection or false alarm). The LRT is defined as

$$\Lambda_{\Theta}(\theta) = \frac{\hat{f}_1(\theta)^{H_1}}{\hat{f}_0(\theta)^{H_0}} \lesseqgtr \eta, \quad (3)$$

where  $H_0$  is the hypothesis that the sample is not skin,  $H_1$  is the hypothesis that the sample is skin, and

$$\begin{aligned} \hat{f}_0(\theta) &= P[\Theta = \theta | \text{not skin}], \\ \hat{f}_1(\theta) &= P[\Theta = \theta | \text{skin}], \\ \Theta &= \{B, \Gamma\}, \\ \theta &= \{\beta, \gamma\}, \end{aligned}$$

are sets of parameters based on the (NDGRI, NDSI)-based detector,  $\hat{f}_1(\theta)$  is the estimated probability density function (pdf) of human skin, and  $\hat{f}_0(\theta)$  is the estimated pdf of the background.

The functional forms of  $\hat{f}_1(\theta)$  and  $\hat{f}_0(\theta)$  are estimated by Gaussian mixture models parameterized using expectation maximization [47] such that

$$\hat{f}_j(\theta) = \sum_{k=1}^{K_j} \pi_{j,k} N(\mu_{j,k}, \Sigma_{j,k}), \quad j \in \{0, 1\}, \quad (4)$$

where  $K_j$  is the number of Gaussian distributions utilized to estimate  $\hat{f}_j(\theta)$ ,  $\pi_{j,k}$  is the weighted value of each Gaussian such that  $\pi_{j,k} \in [0, 1]$  and  $\sum_{k=1}^{K_j} \pi_{j,k} = 1$ . The parameters of each Gaussian are represented by mean vector  $\mu_{j,k}$  and covariance matrix  $\Sigma_{j,k}$ . The likelihood ratio represents a 2D decision surface for the feature spaces evaluated in this work.

The imaged skin-reflectance values are used to compute  $\hat{f}_1(\theta)$ . Similarly,  $\hat{f}_0(\theta)$  is generated using the background pixels in the image. The number of Gaussian distributions ( $K_j$ ) in the mixture model is dependent upon the features used [e.g., the (NDGRI, NDSI) feature pair or features defined from color imagery].

## 5. TEST IMAGE DATA

Image data used in evaluating the performance of the (NDGRI, NDSI) feature space requires the availability of near-infrared spectrum. In this article a hyperspectral imager is used to collect the imagery such that the required spectral information is available. The imager is described and specifics regarding scene content and number of skin and background pixels are provided. Image preparation is important in evaluating the (NDGRI, NDSI) features and comparing them against color-image based features. This preprocessing includes atmospheric correction, band selection to generate the features, and conversion of the visible portion of the spectrum to RGB image data.

### A. Hyperspectral Image Data

Data for this test were collected with the SpecTIR HyperSpecTIR Version 3 (HST3) Hyperspectral Imager [48]. The HST3 is a mirror-scanning instrument that collects 227 spectral channels (400–2500 nm) in one line in the horizontal direction for each of the 250 pixels in the vertical direction, simultaneously. The HST3 nominally acquires 1024 lines (horizontal-scan direction). The number of lines acquired by the HST3 varies slightly due to control limitations for the scanning mirror. Image dimensions are summarized in Table 2. Each line is assembled in sequence to generate the resulting image cube (250 × 1024 × 227 elements). The spectral sampling

**Table 2. Hyperspectral Image Dimensions for Each Image in Fig. 5**

Image	Pixels	Lines	Spectral Channels
Fig. 5(a)	250	1023	227
Fig. 5(b)	250	1027	227
Fig. 5(c)	250	1022	227
Fig. 5(d)	250	1022	227
Fig. 5(e)	250	1022	227
Fig. 5(f)	250	1022	227
Fig. 5(g)	250	1027	227
Fig. 5(h)	250	1023	227

interval is nominally 12 nm in the visible portion of the spectrum and 8 nm in the near-infrared portion of the spectrum.

Eight hyperspectral images in Fig. 5 were collected on an overcast day in March 2009 in Dayton, OH (United States) and contains subjects of various Fitzpatrick skin types located between 3–40 m from the imager. (The implication of a cloudy day is added noise in the spectra used to compute the NDSI.) The images are arranged based on the following considerations. Figure 5(a) is the clearest reference image for subject labeling;



**Fig. 5.** Eight test images acquired in a suburban environment at different stand-off distances: (a) 3 m, (b) 20 m (subjects 7–9 are at 40 m visible on the left), (c) 3 m, and (d) 30 m, (e) 3 m (f) 7 m (g) 3 m, and (h) 3 m. Subjects (1,2,4,6–9,15,16) are Type I/II skin; subjects (3,5,13,14) are Type III/IV skin; and subjects (10–12) are Type V/VI skin. (a) Subjects 1–14 are labeled, (b) subject 15, and (g) subject 16. False-alarm sources include (left-to-right): a branch from a conifer, flesh-colored shirt, color photograph of a person, wood, brown leather boot, cardboard, red brick, stick lying horizontally, flesh-colored doll, and a leather glove all sitting on top of a gray plastic container. Spectralon calibration panels appear below the brown leather boot (partially visible, dark gray) and below the flesh-colored doll (white). Image was acquired by the authors using the SpecTIR HyperSpecTIR Version 3 (HST3) Hyperspectral Imager [48]. All faces are blurred in all images to protect the identity of the volunteers.

Fig. 5(b) is used for a qualitative assessment since it has the most diversity in distance from subjects to the imager; Figs. 5(c), 5(e), and 5(g) are close in images with the subjects grouped together; and Figs. 5(d), 5(f), and 5(h) have subjects at various distances from the imager.

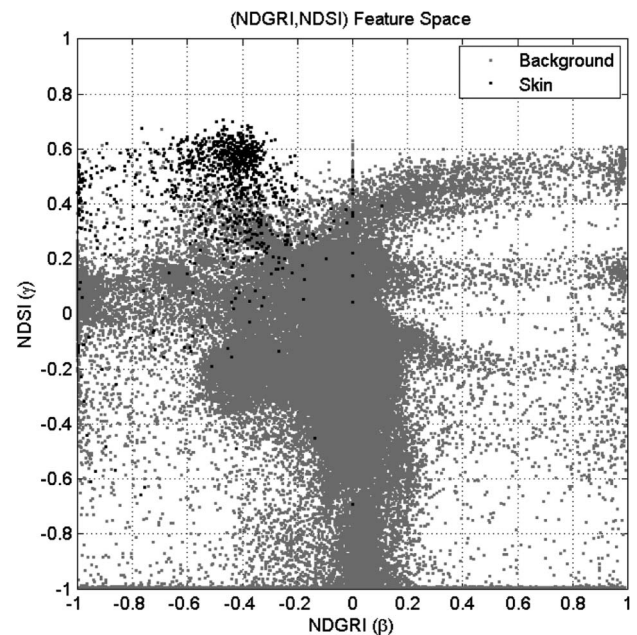
The images contain, in combination, 16 labeled subjects. Subjects (1,2,4,6–9,15,16) are Type I/II skin, subjects (3,5,13,14) are Type III/IV skin, and subjects (10–12) are Type V/VI skin. Subjects 1–14 are labeled in Fig. 5(a), subject 15 in Fig. 5(b), and subject 16 in Fig. 5(g).

The images contain typical false-alarm sources for color-based skin-detection approaches, which include [left-to-right in Fig. 5(a)]: a branch from a conifer (from the yew family) as it has a high NDSI value, flesh-colored shirt, color photograph of a person, wood, brown leather boot, cardboard, red brick, stick lying horizontally, flesh-colored doll, and a leather glove. Below the brown leather boot is a partially visible dark gray Spectralon panel and below the flesh-colored doll is a white Spectralon panel. The scene is a suburban environment with houses (brick face and vinyl siding), asphalt streets, concrete sidewalks, trees, grass, bushes, cars, and children's toys.

All skin pixels are labeled by hand using three observers. Pixels are labeled as skin if two of the three observer's declared the same pixel as skin. All other pixels are considered background. There are a total of 35,198 skin pixels and 2,011,802 background pixels.

## B. Observance of Feature Separability

Figure 6 shows the 2D (NDGRI, NDSI) feature space where black dots are skin pixels and gray dots are background pixels from Fig. 5(b). Note that the skin pixels form a well-defined cluster in the upper left quadrant of the feature space while background pixels form an elongated cluster in the  $y$ -direction centered along the  $x$ -axis. These two clusters appear visually



**Fig. 6.** Display of the (NDGRI, NDSI) feature space computed from the HST3 Hyperspectral Image shown in Fig. 5(b). Black dots are skin features and gray dots are background features.

well separated although some overlap does exist. Factors contributing to the overlap include the effects of non-spectrally pure skin pixels (skin and hair, skin and clothing, and skin and background), skin pixels with low SNR due to heavy shadow (eyes, under the chin), cloud cover reducing the SNR at the wavelengths used in computing the NDSI, pixels with low SNR (open garage door on the right portion of the images), and legitimate false-alarm sources. However, a preponderance of the skin pixels shows good separation from the background, and the features defined in this article should result in good detection and false-alarm suppression results.

### C. Image Pre-processing

Images collected by the HST3 require post-acquisition processing before additional image or signal-processing tasks can take place. Proprietary software developed by SpecTIR incorporates spectral and radiometric calibration information to correct for aberrations or anomalies due to the optical components in the image pathway, to include the scan mirror, slit, optics, and a two-focal plane array (one visible/near infrared and one short-wave infrared). Once the imagery is corrected, it can undergo atmospheric correction to transform the radiance imagery into estimated reflectance, which is the image representation used in the definition of our (NDGRI, NDSI) feature space.

#### 1. Estimated Reflectance

Images are transformed into estimated reflectance through a linear regression known as the empirical line method (ELM) [49] where reflectance is the percentage of incident light on an object that is reflected off of that object. It is widely used in the remote sensing community as it has several desirable properties [49]. ELM removes linear attenuation and scattering affects caused by the Earth's atmosphere while removing errors due to viewing geometry and residual imager calibration artifacts. Any limitations resulting from the inability to remove non-linear effects are outweighed by its ease of implementation and good performance.

Although ELM is often implemented as a least-squares solution using several in-scene and laboratory measurements from both a bright and a dark reference panel, it is easily explained assuming only a single measurement pair exists for both reference panels. ELM estimates the wavelength dependent reflectance at  $\lambda_k$  as

$$\hat{\rho}_{\lambda_k}^i = G_{\lambda_k} L_{\lambda_k}^i + O_{\lambda_k}, \quad (5)$$

where  $\hat{\rho}_{\lambda_k}^i$  is the estimated reflectance for the  $i$ th pixel,  $G_{\lambda_k}$  is the gain (slope),  $O_{\lambda_k}$  is the offset ( $y$ -intercept), and  $L_{\lambda_k}^i$  is the radiance (digital number) recorded by the imager for the  $i$ th pixel.

The wavelength-dependent gain is defined at  $\lambda_k$  as

$$G_{\lambda_k} = \frac{\rho_{B,\lambda_k} - \rho_{D,\lambda_k}}{L_{B,\lambda_k} - L_{D,\lambda_k}}, \quad (6)$$

where  $\rho_{B,\lambda_k}$  and  $\rho_{D,\lambda_k}$  are the known reflectance measurements of the bright and dark reference panels (respectively), and  $L_{B,\lambda_k}$  and  $L_{D,\lambda_k}$  are the image values of the bright and dark reference panels (respectively). The wavelength-dependent offset is defined at  $\lambda_k$  as

$$O_{\lambda_k} = \frac{L_{B,\lambda_k} \rho_{D,\lambda_k} - L_{D,\lambda_k} \rho_{B,\lambda_k}}{L_{B,\lambda_k} - L_{D,\lambda_k}}. \quad (7)$$

**Table 3. Hyperspectral Image Bands Used to Generate the (NDGRI, NDSI) Feature Pair**

Target $\lambda$	Previous Image Band	Closest Image Band	Next Image Band
540 nm	531.70 nm	542.74 nm	554.08 nm
660 nm	648.68 nm	660.27 nm	672.00 nm
1080 nm	1069.91 nm	1078.06 nm	1086.29 nm
1580 nm	1570.83 nm	1579.03 nm	1587.29 nm

#### 2. Band Selection Used to Generate Features

Due to the noise inherent in the system/environment and the fact that the bands selected for the (NDGRI, NDSI) features do not line up with the HST3 band centers, the NDSI and NDGRI features are generated with a slightly different set of image bands to accommodate the available spectra. The algorithms are implemented with the mean of the estimated reflectance of the three HST3 bands closest to the band centers defined for the (NDGRI, NDSI) pair. The mean is used to help average out some of the noise caused by viewing geometry. For example, the estimated reflectance at 540 nm used for the NDGRI algorithm is implemented using the mean of the estimated reflectance at 531.37, 542.74, and 554.08 nm. Band selection for the four spectral regions used in this study is provided in Table 3.

### D. Hyperspectral Image to RGB Image Conversion

In order to compare detection performance of the (NDGRI, NDSI) feature space with color spaces derived from three-channel RGB image data, the hyperspectral data must be converted to RGB space. This is accomplished by first converting the data to the color space defined by the International Commission on Illumination (CIE) in 1931 [50]. The values are defined as an XYZ tuple through a transformation [50] where  $Y$  is the luminance component,  $X$  is a linear combination of cone-response curves while  $Z$  is modeled after blue-cone stimulation. Although conversion to the CIE color space is sufficient for transforming into other color spaces, it is transformed into RGB in this paper for both visual display and as a basis for defining color spaces used in the comparison of feature spaces for skin detection. The transformation of hyperspectral to the XYZ color space is accomplished by integrating the received spectra with the response functions shown in Fig. 7 [50].

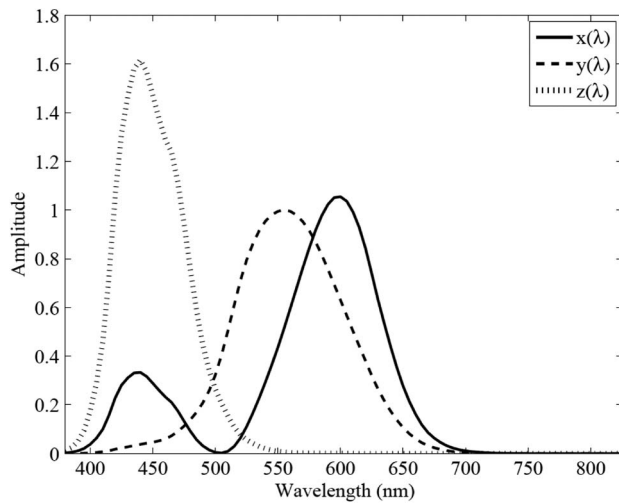
Once the hyperspectral data are converted to the XYZ space, it is converted to the RGB space (assuming a spectrally flat illumination source) through the following linear transform [51]:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 2.3710 & -0.9000 & -0.4710 \\ -0.5140 & 1.4250 & 0.0890 \\ 0.0005 & -0.0150 & 1.0090 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}. \quad (8)$$

### 6. COLOR-BASED FEATURE SPACES USED FOR PERFORMANCE COMPARISON TO THE (NDGRI, NDSI) FEATURE SPACE

There are numerous feature spaces definable from RGB imagery and many that appear in the skin-detection literature. The following sections describe two feature spaces that consistently





**Fig. 7.** Response curves for wavelength dependent conversion factors used to transform hyperspectral spectra to CIE XYZ form [50].

show best performance for skin detection: luminance and red and blue chrominance ( $YC_bC_r$ ) and hue, saturation, value (HSV).

#### A. Luminance, Blue Chrominance, and Red Chrominance

The  $YC_bC_r$  color-space projects RGB pixels into luminosity and chromaticity components [16]. The luminosity component ( $Y$ ) is a weighted sum of RGB values, and the blue ( $C_b$ ) and red ( $C_r$ ) chrominance values are the differences between luminosity and the blue and red components from the RGB data:

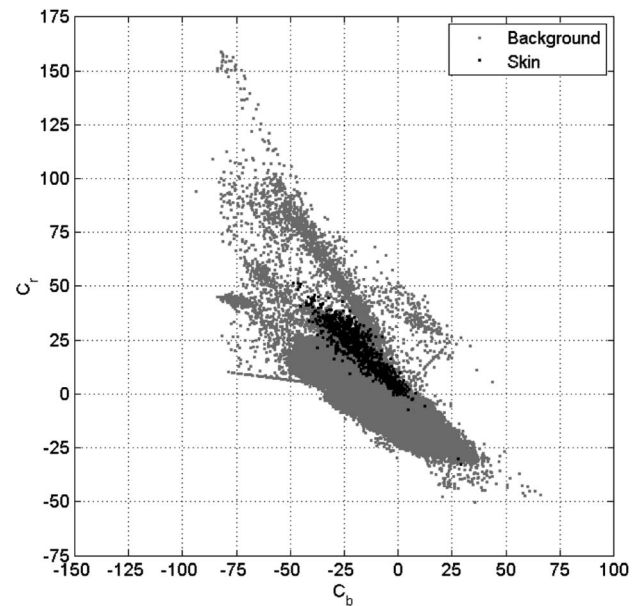
$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B, \\ C_b &= B - Y, \\ C_r &= R - Y. \end{aligned} \quad (9)$$

In addition to being computationally efficient, the  $(C_b, C_r)$  components are a robust representation of skin color under different lighting conditions [16,25] (Fig. 8). Skin remains clustered in an ellipse in  $(C_b, C_r)$ -space when taken from both shadowed and well-lit areas of an image. The authors in [22] use the feature space as a precursor to face detection where skin pixels are identified using an ellipse-shaped boundary.

#### B. Hue, Saturation, Value

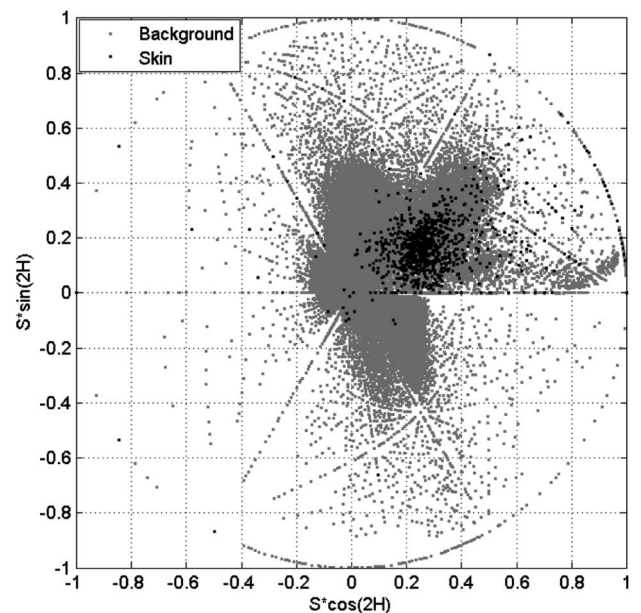
A popular color space for representing skin features is HSV. HSV is a cylindrical feature space that is often noted as having the best skin-detection performance in several studies [24] (and citations therein). The HSV representation can be derived from the RGB data with the following transformations:

$$\begin{aligned} H &= \arccos \frac{\frac{1}{2}((R - G) + (R - B))}{\sqrt{((R - G)^2 + (R - B)(G - B))}}, \\ S &= 1 - 3 \times \frac{\min\{R, G, B\}}{R + G + B}, \\ V &= \frac{1}{3}(R + G + B), \end{aligned} \quad (10)$$



**Fig. 8.** Display of the  $(C_b, C_r)$  feature space computed from the RGB conversion of the HST3 hyperspectral image shown in Fig. 5(b). Black dots are skin features and gray dots are background features.

where hue represents the dominant color and saturation is the proportion of dominant color to value, which corresponds to brightness. Dropping  $V$ , the  $(H, S)$ -space represents skin well, as skin has a characteristic hue near red, a low saturation level, and is invariant to  $V$ . Figure 9 shows the clustering of skin and non-skin pixels in the  $(H, S)$  feature space.



**Fig. 9.** Display of the  $(H, S)$  feature space computed from the RGB conversion of the HST3 hyperspectral image shown in Fig. 5(b). Black dots are skin features and gray dots are background features.



## 7. RESULTS

The three feature spaces described in this article are evaluated using real-image data described in Section 5.A, which was acquired with the HST3 hyperspectral imager also described in Section 5.A. The (NDGRI, NDSI) features are generated using the image bands outlined in Table 3 where the  $(C_b, C_r)$  and  $(H, S)$  features described in Section 6.A and Section 6.B, respectively, are generated using RGB images derived from the same hyperspectral image data used to generate the (NDGRI, NDSI) features. The hyperspectral imagery is converted to RGB imagery as described in Section 5.D. This is accomplished by first integrating the reflectance spectra with the three response functions shown in Fig. 7 and then multiplying the results by Eq. (8). The evaluation of the three feature spaces on data from the same system improves the fairness of the evaluation since the image acquisition and preprocessing up to the point of hyperspectral to RGB conversion and feature generation are identical.

Skin-detection results are a function of the feature space and the model used for the detector. The detector requires some number of mixture components for skin and background pixels and is specified in Section 7.A. Once the number of Gaussian mixtures is specified, experimental results are obtained using  $K$ -fold cross validation (for  $K = 10$  [52]). Stratified random sampling is used to generate the samples assigned to each of the  $K$  bins. The results presented in Section 7.B are evaluated using ROC curves, which characterize performance across the operating range of the detector. The EER criterion is used to select operating points that provide a balance between the miss probability and false-alarm probability. The images in Fig. 5(b) are evaluated at the EER in order to demonstrate the performance from a visual perspective. For each feature space evaluated, the AUC is computed to provide a composite score. EER and AUC results, which indicated performance improvement, are evaluated for normality and an appropriate statistical test is used to determine if statistical significance is achieved.

### A. Gaussian Mixture Models

A mixture of Gaussian distributions is used to model each of the 2D feature spaces. The number of Gaussian mixtures used in each model is determined by starting with a single Gaussian and then adding one mixture component up to the point before the converged model has an ill-conditioned covariance matrix. The number of mixtures for the skin pixels and background pixels is not necessarily the same. Table 4 provides the number of Gaussian mixtures used for each feature pair and for each of the two pixel classes (skin and background).

### B. Performance Comparison of Feature Spaces

The averaged ROC curves for the (NDGRI, NDSI),  $(C_b, C_r)$ , and  $(H, S)$  feature spaces are shown in Fig. 10. The curves show that the detector based on the (NDGRI, NDSI) feature pair dominates the detectors based on the  $(C_b, C_r)$  and  $(H, S)$  features. In the trials here, the detector based on the  $(C_b, C_r)$  feature space provides improved detection results compared to the  $(H, S)$  features space beyond a  $P_D \approx 0.68$  and a  $P_{FA} \approx 2.8 \times 10^{-2}$ .

Two values are drawn from the ROC curves. First is the EER or the point along the ROC that appears closest to the point  $(P_D = 1, P_{FA} = 0)$ . It is determined by the first

**Table 4. Number of Gaussian Mixtures used for the LRT**

Features	Skin Mixtures	B/G Mixtures
(NDGRI, NDSI)	2	4
$(C_b, C_r)$	1	1
$(H, S)$	1	2

45 deg line originating from  $(P_D = 1, P_{FA} = 0)$  corner moving toward the  $(P_D = 0, P_{FA} = 1)$  corner that intersects the ROC curve tangentially. That intersection can be computed as follows:

$$P_{FA} = P_{Miss},$$

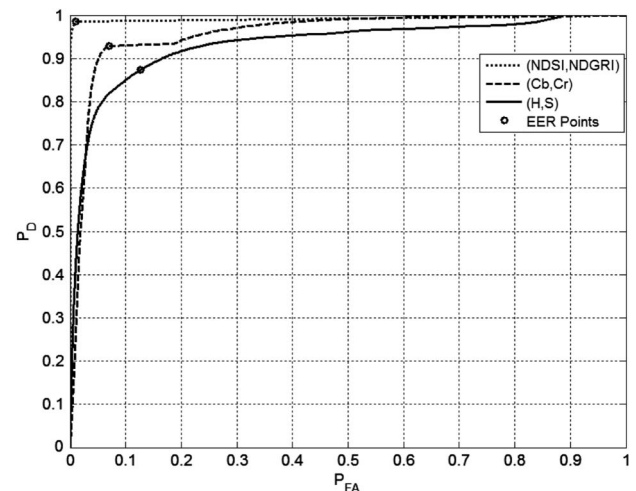
$$P_{FA} = (1 - P_D),$$

$$P_{FA} + P_D = 1. \quad (11)$$

The EER results for the evaluated features are presented in Table 5 and include the  $(P_D, P_{FA})$  for each of the  $K = 10$  trials. Included are the mean and standard deviation for these EERs. The results in Table 5 show that the proposed (NDGRI, NDSI) features provide better performance at that suggested operating point (a single threshold). Overall, detection rates are improved by 5.6%–11.2% while false-alarm rates are reduced by 6.0%–11.6% compared to well performing color-feature spaces. The EER shows a clear ranking between the performance of the features used in skin detection with the LRT. That rank ordering is the (NDGRI, NDSI) features followed by the  $(C_b, C_r)$  features followed by the  $(H, S)$  features.

Figures 11(a)–11(c) show detection images for the  $(H, S)$ ,  $(C_b, C_r)$ , and (NDGRI, NDSI) features (respectively) for the test scene shown in Fig. 5(b). The LRT models and thresholds used to obtain the EER values reported in Table 5 are used to generate the detection images. Each detection image is the result of a six out of ten majority vote. The outcome is a skin-detection for pixel  $i$  if six of the ten detection models declare it as skin, otherwise pixel  $i$  is declared as background. Skin pixels are represented as white pixels in the detection image while background pixels are represented as black pixels.

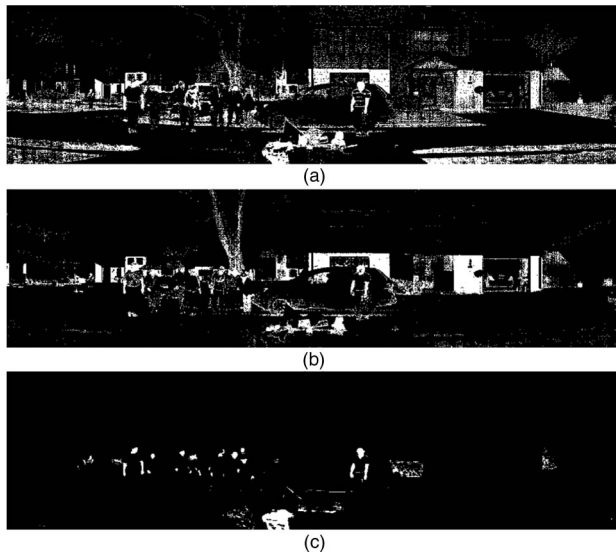
Consistent with the results shown in Table 5, the  $P_D/P_{FA}$  values for  $(H, S)$ ,  $(C_b, C_r)$ , and (NDGRI, NDSI) feature



**Fig. 10.** Skin detection ROC curves averaged over  $K = 10$  cross validation trials.

**Table 5. Equal Error Rate Detection and False Alarm Values and Statistics for Each Set of the Three Detection-Features Conducted over  $K = 10$  Trials<sup>a</sup>**

Features		Detection Results and Statistics Reported in %.											
		$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$T_8$	$T_9$	$T_{10}$	$\mu$	$\sigma$
(NDGRI, NDSI)	$P_D$	99.00	98.18	99.00	98.52	98.52	98.18	98.69	98.35	98.27	98.89	98.56	0.32
	$P_{FA}$	1.32	0.92	1.28	0.88	0.89	0.90	1.33	0.93	0.95	1.20	1.06	0.19
$(C_b, C_r)$	$P_D$	93.15	92.55	92.12	93.09	92.92	93.12	93.06	93.00	93.43	92.06	92.85	0.45
	$P_{FA}$	6.89	7.15	7.08	6.77	6.97	6.92	6.97	7.03	6.59	7.12	6.95	0.17
$(H, S)$	$P_D$	87.49	87.00	87.03	87.57	87.69	87.06	87.12	87.88	87.86	86.55	87.32	0.44
	$P_{FA}$	12.49	12.99	13.00	12.44	12.28	12.94	12.89	12.12	12.15	12.49	12.68	0.45

<sup>a</sup>Trials are denoted  $T_k$ ,  $k = 1 \dots K$ .**Fig. 11.** (a) Detection results for the test seen, evaluated at the thresholds used to determine the EERs listed in Table 5 for (b)  $(H, S)$  features, (c)  $(C_b, C_r)$  features, and (d) (NDGRI, NDSI) features. Skin detections are based on an agreement with six of the ten models and are shown as white pixels in (b)–(d). Black pixels are background pixels.

spaces is 87.46%/12.54%, 92.16%/7.84%, and 98.75%/1.25%, respectively. Note that the color-based features have false alarms associated with vegetation and the skin confusers at the bottom of Figs. 5(a)–5(h). The (NDGRI, NDSI) features remove many of the false detections keyed in on the color image-based features; however, there are still some challenges with respect to the flesh-colored shirt and vegetation from the

yew-family. From a visual perspective, the features proposed in this work produce much cleaner (less cluttered) detection results.

The second value drawn from the ROC curve is the AUC. For each of the  $K = 10$  trials, the AUC is computed as well as the mean and standard deviation over the set of trials (Table 6). The AUC for the proposed (NDGRI, NDSI) features increases by 0.034–0.173 compared to the  $(C_b, C_r)$  and  $(H, S)$  features. The AUC shows a clear ranking between the performance of the features used in skin detection with the LRT. That rank ordering is the (NDGRI, NDSI) features followed by the  $(C_b, C_r)$  features followed by the  $(H, S)$  features.

The EER  $P_D$ , EER  $P_{FA}$ , and AUC results are evaluated for statistically significant improvement between the (NDGRI, NDSI) and color features. Results are first evaluated for normality using the Shapiro–Wilk test. Although several normality tests are available, the Shapiro–Wilk test is shown to work well for small sample sizes (less than 20) [53]. All results, with the exception of the false-alarm probability results for the (NDGRI, NDSI) feature pair, are shown to be normally distributed at the  $\alpha = 0.05$  level and the paired t-test is used to evaluate whether or not the population pairs come from a distribution with the same mean and different variances or from separate distributions all together. When comparing non-normally distributed data to normally distributed data, the Wilcoxon signed ranks test [54] is used. When comparing the three feature combinations ( $\{(NDGRI, NDSI), (C_b, C_r)\}$ ,  $\{(NDGRI, NDSI), (H, S)\}$ , and  $\{(C_b, C_r), (H, S)\}$ ), the feature pair on the left performs better than the feature pair on the right (using the LRT detector) at the  $\alpha = 0.05$  level for normally and non-normally distributed pairs. The results in this article show that the (NDGRI, NDSI) feature space provides statistically significant improvement in detection and false-alarm suppression when using the LRT.

**Table 6. Area under the Receiver Operating Characteristic Curve and Statistics for Each Set of the Three Detection-Features Conducted over  $K = 10$  Trials<sup>a</sup>**

Features	Area under the Curve Results Scaled by $\times 10^2$											
	$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$T_8$	$T_9$	$T_{10}$	$\mu$	$\sigma$
(NDGRI, NDSI)	99.46	99.05	99.47	99.23	98.23	99.06	99.30	99.14	99.10	99.41	99.25	0.16
$(C_b, C_r)$	95.88	95.77	95.62	95.88	95.82	96.00	96.00	95.81	96.15	95.59	95.85	0.17
$(H, S)$	82.13	81.56	81.73	82.36	82.16	81.80	81.71	82.42	82.31	81.33	81.95	0.38

<sup>a</sup>Trials are denoted  $T_k$ ,  $k = 1 \dots K$ .

## 8. CONCLUSIONS

This article presented a new 2D feature space used in skin detection and false-alarm suppression. The NDSI is based on observations from spectral measurements of human tissue and refined based on known physical properties of the tissue. The result is improved detection in complex backgrounds while the NDGRI inverts known color properties of human tissue in order to suppress false alarms. The (NDGRI, NDSI) feature space is bounded  $-1 \leq (\beta, \gamma) \leq 1$ .

The (NDGRI, NDSI) feature space proposed in this article produces statistically significant improved results compared to good performing color image-based feature spaces [specifically  $(C_b, C_r)$  and  $(H, S)$  features]. The ROC curve based on the (NDGRI, NDSI) features dominates the ROC curves for both the  $(C_b, C_r)$  and  $(H, S)$  feature spaces. Furthermore, the EER shows statistically significant improvements in  $P_D$  of 5.6%–11.2% and statistically significant decrease in  $P_{FA}$  of 6.0%–11.6% compared to well performing color feature spaces. Finally, the AUC shows a statistically significant increase in total area under the curve by 0.034–0.173.

The results for the color features deviates slightly from published literature. For the data evaluated in this article the  $(C_b, C_r)$  feature space shows improved performance over the  $(H, S)$  feature space. Others have shown that cylindrical feature spaces generally perform better for the skin-detection task. These results may be associated with the data set used for analysis in this article.

There are several detection and classification schemes that could be employed in the skin-detection problem. The LRT is used in this article due to its effectiveness and its wide acceptance. It is possible that other approaches would produce improved results using the features defined in this article. Investigating some of the more prominent approaches is an area of future work.

Due to the narrow spectral channels acquired by hyperspectral imagers, data acquisition is highly affected by illumination conditions such as solar position due to time of day or time of year, aerosols due to humidity and pollution, and color temperature due to haze and cloud cover. Additional work is needed to understand the sensitivity of the features generated in the article to any non-linear effects created by these different illumination conditions.

Due to the nature of SAR, the ability to perform skin detection in real time is important. The features defined in this article are computationally efficient to generate (requiring four additions and two divisions) and the LRT detection space can be implemented efficiently as a look up table. The complexity of feature computation is comparable to generating the  $(C_b, C_r)$  features and significantly less than generating the  $(H, S)$  features. However, the computational workload associated with atmospheric correction can add considerable complexity. A mapping of the detection space (defined from e.g., the EER) to the current imaging conditions could significantly reduce that workload, and research in this area is warranted.

The acquisition of the spectra required to generate the (NDGRI, NDSI) features comes at an increased financial and processing cost if using a hyperspectral imager. Such systems are

not as common as color cameras and can only be acquired through a limited number of manufacturers. Even though the features are efficiently computed and the results shown to be statistically significant, the performance gains may not justify the incurred cost, acquisition, and post-processing speed limitations. However, an application-specific optical system could meet the needs of real-time applications such as SAR with a dramatically reduced cost compared to a hyperspectral solution. Work in [55] demonstrates the feasibility of an application-specific solution. Although the solution in [55] is an order of magnitude less expensive than a hyperspectral imager, it is still costly given today's technology. As technology advances occur, the cost is likely to reduce making such a solution more viable for certain applications.

**Funding.** Air Force Research Laboratory (AFRL).

**Acknowledgment.** The authors would like to thank Christina Schutte and the late Dr. Devert Wicker of the Air Force Research Laboratory Sensors Directorate for sponsoring this work. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

## REFERENCES

1. D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.* **19**(1), 29–43 (2002).
2. S. Subramanian and N. Gat, "Subpixel object detection using hyperspectral imaging for search and rescue operations," *Proc. SPIE* **3371**, 216–225 (1998).
3. M. Topping, J. Pfeiffer, A. Sparks, K. Jim, and D. Yoon, "Advanced airborne hyperspectral imaging system (AAHIS)," *Proc. SPIE* **4816**, 1–11 (2002).
4. J.-R. Simard, P. Mathieu, G. Fournier, and V. Larochelle, "A range-gated intensified spectrographic imager: an instrument for active hyperspectral imaging," *Proc. SPIE* **4035**, 180–191 (2000).
5. B. Stevenson, R. O'Connor, W. Kendall, A. Stocker, W. Schaff, R. Holasek, D. Even, D. Alexa, J. Salvador, M. Eismann, R. Mack, P. Kee, S. Harris, B. Karch, and J. Kershenstein, "The civil air patrol archer hyperspectral sensor system," *Proc. SPIE* **5787**, 17–28 (2005).
6. C. Leonard, D. Michael, J. Gradie, J. Iokepa, and C. Stalder, "Performance of an EO/IR Sensor system in marine search and rescue," *Proc. SPIE* **5787**, 122–133 (2005).
7. C. Simi, A. Hill, and H. Kling, "Airborne remote spectrometry support to rescue personnel at Ground Zero after the World Trade Center attack on September 11, 2001," *Proc. SPIE* **4816**, 23–32 (2002).
8. Z. Pan, G. Healey, M. Prasad, and B. Tromberg, "Face recognition in hyperspectral images," *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 1552–1560 (2003).
9. M. Yang, D. Kriegman, and N. Ahuja, "Detecting faces in images: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.* **24**, 34–58 (2002).
10. R. Sanchez-Reillo, C. Sanchez-Avila, and A. Gonzalez-Marcos, "Biometric identification through hand geometry measurements," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 1168–1171 (2000).
11. C. Li and K. Kitani, "Pixel-level hand detection in ego-centric videos," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2013), pp. 3570–3577.
12. J. Daugman, "Face and gesture recognition: Overview," *IEEE Trans. Pattern Anal. Mach. Intell.* **19**, 675–676 (1997).
13. A. S. Nunez and M. J. Mendenhall, "Detection of human skin in near infrared hyperspectral imagery," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (2008), Vol. **2**, pp. 621–624.



14. A. Kumar, "An empirical study of selection of the appropriate color space for skin detection," in *International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)* (2014), pp. 725–730.
15. R. Khan, A. Hanbury, and J. Stoetinger, "Skin detection: A random forest approach," in *17th IEEE International Conference on Image Processing (ICIP)* (2010), pp. 4613–4616.
16. V. Vezhnevets, V. Sazonov, and A. Andreeva, "A survey on pixel-based skin color detection techniques," in *Proceedings of GraphiCon* (2003), pp. 80–92.
17. C. Doukim, J. Dargham, A. Chekima, and S. Omatu, "Combining neural networks for skin detection," *Signal Image Process.* **1**, 1–11 (2010).
18. J. Brand and J. Mason, "A comparative assessment of three approaches to pixel-level human skin-detection," in *Proceedings of 15th International Conference on Pattern Recognition* (2000), Vol. **1**, pp. 1056–1059.
19. D. R. Rajesh, ed., *Combining Color Spaces for Human Skin Detection in Color Images using Skin Cluster Classifier* (ACEEE, 2013).
20. J. Kovac, P. Peer, and F. Solina, "Human skin color clustering for face detection," in *The IEEE Region 8 EUROCON 2003: Computer as a Tool* (2003), Vol. **2**, pp. 144–148.
21. M. M. Fleck, "Finding naked people," in *Proceedings of the 4th European Conference on Computer Vision* (1996), Vol. **1065**, pp. 593–602.
22. R. Hsu, M. Abdel-Mottaleb, and A. K. Jain, "Face detection in color images," in *IEEE International Conference on Image Processing (ICIP)* (2001), Vol. **1**, pp. 1046–1049.
23. G. Gomez, "On selecting colour components for skin detection," in *Proceedings of International Conference on Pattern Recognition* (2002), Vol. **2**, pp. 961–964.
24. J. C. SanMiguel and S. Suja, "Skin detection by dual maximization of detectors agreement for video monitoring," *Pattern Recogn. Lett.* **34**, 2102–2109 (2013).
25. A. Albiol, L. Torres, and E. Delp, "Optimum color spaces for skin detection," in *IEEE International Conference on Image Processing (ICIP)* (2001), Vol. **1**, pp. 122–124.
26. I. Pavlidis, P. Symosek, B. Fritz, M. Bazakos, and N. Papanikolopoulos, "Automatic detection of vehicle occupants: the imaging problem and its solution," *Mach. Vis. Appl.* **11**, 313–320 (2000).
27. J. Dowdall, I. Pavlidis, and G. Bebis, "Face detection in the near-IR spectrum," *Proc. SPIE* **5074**, 745–756 (2003).
28. G. A. Kilgore and P. R. Whillock, "Skin detection sensor," U.S. patent 7,446,316 (Nov 4, 2005).
29. J. Peers, *1,001 Logical Laws, Accurate Axioms, Profound Principles, Trusty Truisms, Homey Homilies, Colorful Corollaries, Quotable Quotes, and Rambunctious Rumi* (Doubleday, 1979).
30. P. Matts, P. Dykes, and R. Marks, "The distribution of melanin in skin determined in vivo," *Br. J. Dermatol.* **156**, 620–628 (2007).
31. A. Krishnaswamy and G. Baranoski, "A study of skin optics," Tech. Rep. CS-2004-01 (University of Waterloo, 2004).
32. S. L. Jacques, "Skin optics," in *Oregon Medical Laser Center News*, 1998, <http://omlc.org/news/index.html>.
33. S. Prahl, "Tabulated molar extinction coefficient for hemoglobin in water," 1998, <http://omlc.org/spectra/hemoglobin/summary.html>.
34. H. Buiteveld, J. Hakvoort, and M. Donze, "The optical properties of pure water," *Proc. SPIE* **2258**, 174–183 (1994).
35. K. F. Palmer and D. Williams, "Optical properties of water in the near infrared," *J. Opt. Soc. Am.* **64**, 1107–1110 (1974).
36. S. Jacques, "Optical properties of biological tissues: a review," *Phys. Med. Biol.* **58**, R37–R61 (2013).
37. S. Jacques, "Origins of tissue optical properties in the UVA, visible, and NIR regions," in *OSA TOPS on Advances in Optical Imaging and Photon Migration* (1996), Vol. **2**, pp. 364–369.
38. R. Anderson and J. Parrish, "The optics of human skin," *J. Invest. Dermatol.* **77**, 13–19 (1981).
39. N. Kollias, "The physical basis of skin color and its evaluation," *Clinics Dermatol.* **13**, 361–367 (1995).
40. E. Angelopoulou, "Understanding the color of human skin," *Hum. Vis. Electron. Imaging VI* **4299**, 243–251 (2001).
41. M. Storrington, T. Kocka, H. J. Andersen, and E. Granum, "Tracking regions of human skin through illumination changes," *Pattern Recogn. Lett.* **24**, 1715–1723 (2003).
42. M. J. Jones and J. M. Rehg, "Statistical color models with application to skin detection," *Int. J. Comput. Vis.* **46**, 81–96 (2002).
43. M. Abdel-Mottaleb and A. Elgammal, "Method for detecting a face in a digital image," U.S. patent 6,574,354 (June 3, 2001).
44. H. Wang and S.-F. Chang, "Rapid modeling of diffuse reflectance of light in turbid slabs," *J. Opt. Soc. Am.* **15**, 936–944 (1998).
45. J. Rouse, R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS," in *Third Earth Resources Technology Satellite-1 Symposium* (1973), pp. 309–317.
46. R. Clark, G. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, and S. Sutley, *USGS Digital Spectral Library Splib06a: U.S. Geological Survey, Digital Data Series 231*, 2007. Online <http://speclab.cr.usgs.gov/spectral.lib06>.
47. T. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Mag.* **13**(6), 47–60 (1996).
48. C. Jengo and J. LaVeigne, "Sensor performance comparison of HyperSpecTIR instruments 1 and 2," in *Proceedings of IEEE Aerospace Conference* (March 2004).
49. F. Kruse, K. Kierein-Young, and J. Boardman, "Mineral mapping at Cuprite, Nevada with a 63-channel imaging spectrometer," *Photogramm. Eng. Remote Sens.* **56**, 83–92 (1990).
50. International Commission on Illumination, "Selected colorimetric tables," 1931, <http://www.cie.co.at>.
51. G. Hoffmann, "CIE Color Space," 2000, <http://www.fho-enden.de/hoffmann/ciexyz29082000.pdf>.
52. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning* (Springer, 2001).
53. S. Shapiro and M. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika* **52**, 591–611 (1965).
54. F. Wilcoxon, "Individual comparisons by ranking methods," *Biom. Bull.* **16**, 80–83 (1945).
55. K. Poskosky, "Design of a monocular multi-spectral skin detection, melanin estimation, and false alarm suppression system," Master's thesis (Air Force Institute of Technology, 2010).