Three-dimensional facial recognition using passive long-wavelength infrared polarimetric imaging

Alex J. Yuffa,* Kristan P. Gurton, and Gorden Videen

Computational & Information Science Directorate, U.S. Army Research Laboratory, Adelphi, Maryland 20783, USA

*Corresponding author: ayuffa@gmail.com

Received 7 October 2014; revised 19 November 2014; accepted 21 November 2014;
posted 21 November 2014 (Doc. ID 224226); published 18 December 2014

We use a polarimetric camera to record the Stokes parameters and the degree of linear polarization of long-wavelength infrared radiation emitted by human faces. These Stokes images are combined with Fresnel relations to extract the surface normal at each pixel. Integrating over these surface normals yields a three-dimensional facial image. One major difficulty of this technique is that the normal vectors determined from the polarizations are not unique. We overcome this problem by introducing an additional boundary condition on the subject. The major sources of error in producing inversions are noise in the images caused by scattering of the background signal and the ambiguity in determining the surface normals from the Fresnel coefficients. © 2014 Optical Society of America

OCIS codes: (110.6820) Thermal imaging; (110.6880) Three-dimensional image acquisition; (260.5430) Polarization.

http://dx.doi.org/10.1364/AO.53.008514

1. Introduction

Long-wavelength infrared (LWIR) imaging has many practical applications. Since it is a measure of the signal emitted by an object, it is a passive tool that does not require an additional light source. One major disadvantage of LWIR imaging is the lack of contrast. Images of people tend to look washed out and ghostlike, and it is difficult to obtain information from them. Incorporating polarization information can provide additional discriminating information that can enhance the contrast of images, making their features more identifiable [1].

In this manuscript, we describe efforts to expand the capabilities of LWIR polarimetric imaging by extracting three-dimensional (3D) images. We demonstrate the capabilities by generating 3D images of human faces reconstructed from single Stokes-vector images. We use human faces because they have a complex morphology with both concave and convex regions that are especially challenging to analyze. In addition, they have applications in facial recognition. To avoid a possible source of confusion, we refer to an image that contains the depth and shape information as a "3D image." However, technically speaking, this image is a two-dimensional surface map $f(x,y)$ of the subject's face.

In the context of defense and security applications, e.g., nighttime/daytime surveillance and facial recognition, our method offers a number of advantages over methods that use visible light. Perhaps the oldest and the best-known of these methods is the shape-from-shading method [2,3]. In this method, a digital camera is used to capture an intensity image of the subject's face, and the gradual variation of shading in the image is used to obtain $f(x,y)$. Unfortunately, to use this method one has to know the reflectivity map (bidirectional reflectance distribution function) and the position of the light source (s). Moreover, it requires precise illumination conditions and may yield unreliable results if these conditions are not strictly satisfied [3,4]. Recently, the polarization state of visible light specularly and/or diffusely reflected from a smooth dielectric object has been used to obtain a 3D image of an object

1559-128X/14/368514-08$15.00/0

[5–9]. These shape-from-polarization methods require one or more strategically placed illumination sources and are sensitive to the illumination conditions. Furthermore, some of these methods [7,9] require two views of the object, which may be difficult to achieve in defense and security applications. Our method requires only a single view of the subject and does not require any illumination sources. Thus, it is a passive method that is better suited for surveillance applications.

We note that this method may also have medical applications. Since the end of the Cold War, the transfer of U.S. Military infrared imaging technology to the public medical sector has resulted in a number of new diagnostic techniques in medicine [10]. For example, infrared imaging has been used to treat and/or diagnose vascular disorders, breast cancer, dermatological disorders, and diabetes [10–14]. We speculate that the polarization state of emitted LWIR radiation contains additional information that may be of value to medical practitioners. For example, a rapid 3D modeling of body parts may be of use during surgery.

2. Image Acquisition

We record the polarization state of the *total* wave emanating from the face using a polarimetric camera equipped with a Stirling-cooled mercury cadmium telluride focal-plane array of 640 × 480 pixels. The array has a spectral response range of 7.5–11.1 μm [1]. A sequence of 32-bit images is recorded at a frame rate of 120 Hz, and a Fourier modulation technique [15] is applied to them to obtain the Stokes parameters (S_0, S_1, S_2, S_3) and the degree of linear polarization (DoLP) for each pixel of the array. From this image vector, we obtain the surface normal **N** on a pixel-by-pixel basis that we then integrate to obtain the surface $f(x,y)$. We note that the Stokes parameters cannot be measured directly and are inferred from six independent intensity measurements [1]. Therefore, at a frame rate of 120 Hz, the subject must remain stationary for approximately $\frac{6}{120} = 50$ ms. The recorded images contain both reflected and emitted waves. As a consequence of the Fresnel equations, the emitted waves are polarized parallel to the plane of emission, which is defined by the line of sight of the camera and the surface normal **N**. In contrast, reflected waves are polarized perpendicular to this plane. In other words, the emitted and the reflected polarization states are orthogonal. For human subjects in common indoor conditions, we find that the emitted radiation dominates, and the reflected radiation contribution to the polarization state may be disregarded. This is not unexpected, as less than 10% of the infrared radiation is specularly reflected by the skin [14]. Furthermore, we assume that the subject's face is sufficiently warmer than the ambient environment, and that the environment is devoid of strong LWIR radiation sources.

3. Properties and Parameters of the Skin

To a first-order approximation, a macroscopic body at temperature T emits electromagnetic radiation in a manner proportional to Planck's radiation law. Namely, the distribution of the emitted energy density in an infinitesimal interval $d\omega$ depends on temperature via

$$u(\omega, T)d\omega = \frac{\hbar\omega^3}{\pi^2 c^3\left[\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1\right]}d\omega, \tag{1}$$

where ω is the angular frequency, and c, $\hbar$, k_B are known constants. Humans, being homeotherms, can maintain *core* body temperature at a nearly constant 37°C. However, the skin surface temperature is not constant, and it depends on the surrounding environment as well as the blood flow underneath the outer layer of the skin. For example, the temperature of buttock skin exposed to direct sunlight is ≈40°C [16], and the surface temperature of forehead and toes under room conditions is ≈34°C and ≈27°C, respectively [10]. From Eq. (1) and the above temperature data, it follows that ≈90% of the emitted radiation by humans is in the 6–14 μm wavelength range with a mean peak between 9 and 10 μm [10,11]. Thus, our polarimetric camera with a spectral response range of 7.5–11.1 μm is well suited for measuring the radiation emitted by human skin.

To utilize the polarizing LWIR properties to extract 3D information, we require a numerical value for the complex index of refraction n of the skin. Unfortunately, this value is not constant, and it tends to have a spatial and an environmental dependence [10,11,14,17,18]. In order to establish a reasonable value for n, we briefly review the composition of human skin.

Human skin consists of an outer layer called an epidermis and a dermis layer on which the epidermis layer rests. The epidermis layer can be further partitioned into a thin (≈10–20 μm) nonliving layer and a thick (≈100–200 μm) living layer [13,14,19]. The thickness of the living layer is nonuniform, and the interface between the epidermis layer and the dermis layer contains many nonplanar folds (see Fig. 1(a) in [13]). The dermis layer contains a network of blood capillaries, and these capillaries lose thermal energy to the epidermis layer, which in turn radiates the thermal energy to the surrounding environment [10,19]. From the above discussion it follows that human skin is highly variable and complex. For our purposes, we consider its characterization through an effective refractive index that we consider to be constant over the body. The real part of this effective n is between 1.4 and 1.6, and the imaginary part is between 0.01 and 0.03 [12,14,20–22].

4. From Polarization State to 3D Image

When thermal radiation is emitted by a hot body having an optically smooth surface, the emitted waves become partially linearly polarized [23–25].

Let us consider thermal radiation emanating from a large planar surface with directional spectral emissivity ε_p, absorptivity α_p, and reflectivity r_p. From the Kirchhoff radiation law, it is known that these three quantities are related via

$$\varepsilon_p(\theta,\phi) = \alpha_p(\theta,\phi) = 1 - r_p(\theta,\phi), \tag{2}$$

where θ is the angle between $\mathbf{N}$ and the camera's line of sight, ϕ is the azimuthal angle, and the polarization state with respect to the plane of emission is denoted by the subscript $p = \perp$ or $\parallel$ ($\perp$ for perpendicular and $\parallel$ for parallel). For brevity, in Eq. (2) we have suppressed the functional dependence of ε_p, α_p, and r_p on temperature and wavelength. Substituting Eq. (2) into the definition of DoLP,

$$P(\theta) = \frac{\varepsilon_\parallel(\theta) - \varepsilon_\perp(\theta)}{\varepsilon_\parallel(\theta) + \varepsilon_\perp(\theta)}, \tag{3}$$

and using the Fresnel reflection coefficients for perpendicular and parallel polarization states, we obtain the DoLP as a function of θ [23], i.e.,

$$P(\theta) = \frac{[1 - |r_\parallel(\theta)|^2] - [1 - |r_\perp(\theta)|^2]}{[1 - |r_\parallel(\theta)|^2] + [1 - |r_\perp(\theta)|^2]}, \tag{4a}$$

where

$$r_\parallel(\theta) = \frac{n^2\cos\theta - \sqrt{n^2 - \sin^2\theta}}{n^2\cos\theta + \sqrt{n^2 - \sin^2\theta}}, \tag{4b}$$

$$r_\perp(\theta) = \frac{\cos\theta - \sqrt{n^2 - \sin^2\theta}}{\cos\theta + \sqrt{n^2 - \sin^2\theta}}. \tag{4c}$$

Notice that $P(\theta)$ does not explicitly depend on the azimuthal angle ϕ because the polarization state of the emitted radiation has been decomposed into its parallel and perpendicular components. If we partition the human face surface into planar facets and neglect possible multiple reflections of the emitted radiation, then we can obtain θ. Equating Eq. (4) to the experimentally measured DoLP and then numerically solving for θ yields a value for θ on a pixel-by-pixel basis. Notice that by neglecting possible multiple reflections of the emitted radiation we are essentially treating $f(x,y)$ as a convex surface, which, clearly, a human face is not. In other words, the above method may produce inaccurate θs for *concave* portions of the face, e.g., eye sockets.

The normal to the surface is given by

$$\mathbf{N} = -\frac{\partial}{\partial x}f(x,y)\hat{\mathbf{x}} - \frac{\partial}{\partial y}f(x,y)\hat{\mathbf{y}} + \hat{\mathbf{z}}, \tag{5a}$$

where from simple geometric considerations it follows that

$$-\frac{\partial}{\partial x}f(x,y) = \tan\theta\cos\phi \tag{5b}$$

and

$$-\frac{\partial}{\partial y}f(x,y) = \tan\theta\sin\phi. \tag{5c}$$

The unknown angle ϕ in Eq. (5) can be obtained for each facet directly from the experimentally measured Stokes parameters [26], i.e.,

$$\phi = \psi + (0 \quad \text{or} \quad \pi), \tag{6a}$$

where

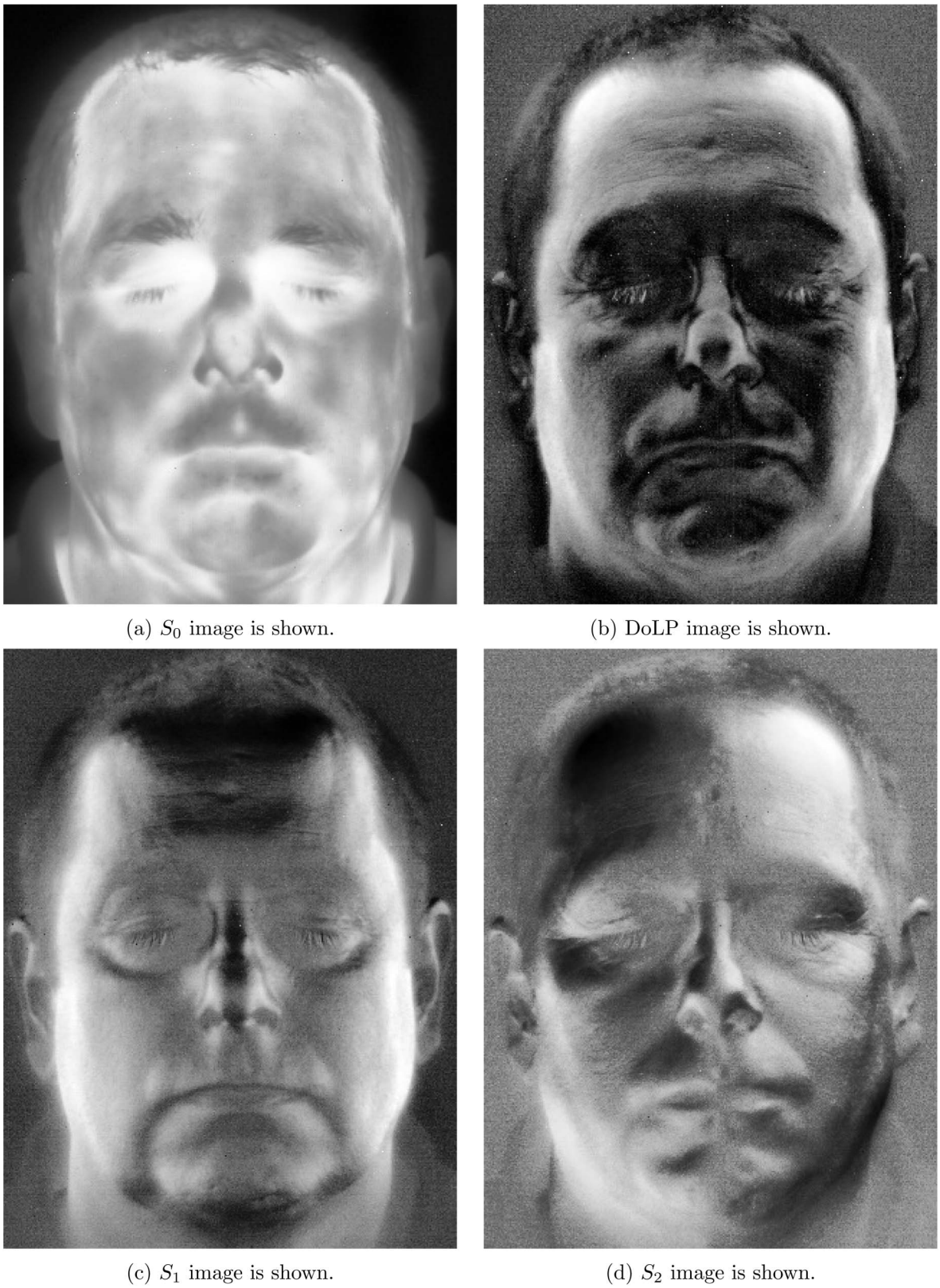
$$\psi = \frac{1}{2}\begin{cases} \arctan\left(\frac{S_2}{S_1}\right) + \frac{\pi}{2}, & S_1 \le 0, \\ \arctan\left(\frac{S_2}{S_1}\right) + \pi, & S_1 > 0 \quad \text{and} \quad S_2 < S_1, \\ \arctan\left(\frac{S_2}{S_1}\right) + 0, & S_1 > 0 \quad \text{and} \quad S_2 \ge S_1. \end{cases} \tag{6b}$$

Notice that Eq. (6a) only determines ϕ up to a π ambiguity because the Stokes parameters are invariant under a rotation of π. Such an ambiguity constitutes an additional unknown in the mathematical solution. To resolve this ambiguity, we must introduce an additional boundary condition to find a solution. The boundary condition we employ for human faces is that the surface normal must point away from the subject's head on the occluding boundary. In other words, we set ϕ to ψ or $\psi + \pi$ on the occluding boundary so that the normal $\mathbf{N}$ satisfies the above requirement. Such a boundary condition is valid for human faces. After ϕ is chosen on the occluding boundary, we choose ϕ for the neighboring pixels in such a way that the normal for these pixels points in the same direction as the normal on the occluding boundary. This process is iteratively continued for all remaining ambiguous ϕs for which the corresponding θ is larger than some threshold. The ϕs for which the corresponding θ is smaller than the threshold are left unchanged because these facets are only slightly tilted with respect to the line of sight of the camera. We note that this additional boundary condition may not be valid for all objects.

A number of methods are available to integrate $\mathbf{N}$ and obtain $f(x,y)$ up to an additive constant; e.g., see [27–32]. For its simplicity and robustness to noise, we chose to use the well-known Frankot–Chellapa method [27]. Recall that the Frankot–Chellapa method enforces the integrability condition, $\nabla \times \mathbf{N} = \mathbf{0}$, for all points on the support of $f(x,y)$, and, thus, it smoothes the surface normal. This is beneficial because the surface of a human face may have discontinuities.

5. Results

All of the measurements presented in this section were taken in an air-conditioned laboratory where the temperatures were significantly lower ($\lesssim$21°C) than the subjects under investigation. This is important because LWIR fields from the background can scatter off the subject. The resulting image measured by the camera is a superposition of this noise field



(a) S_0 image is shown.

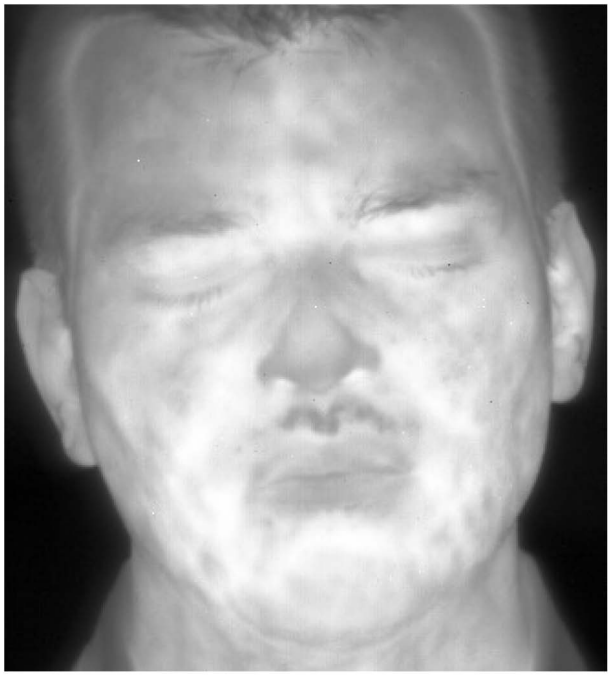
(b) DoLP image is shown.

(c) S_1 image is shown.

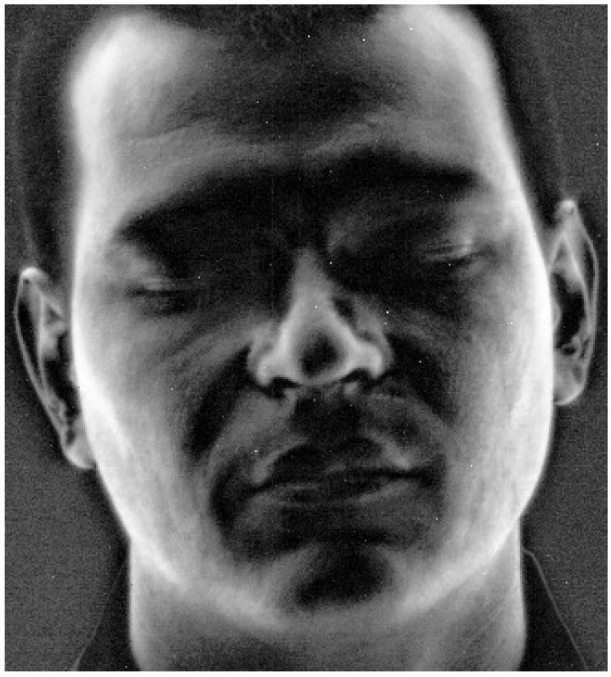
(d) S_2 image is shown.

Fig. 1. Intensity and polarization imagery of Subject A.

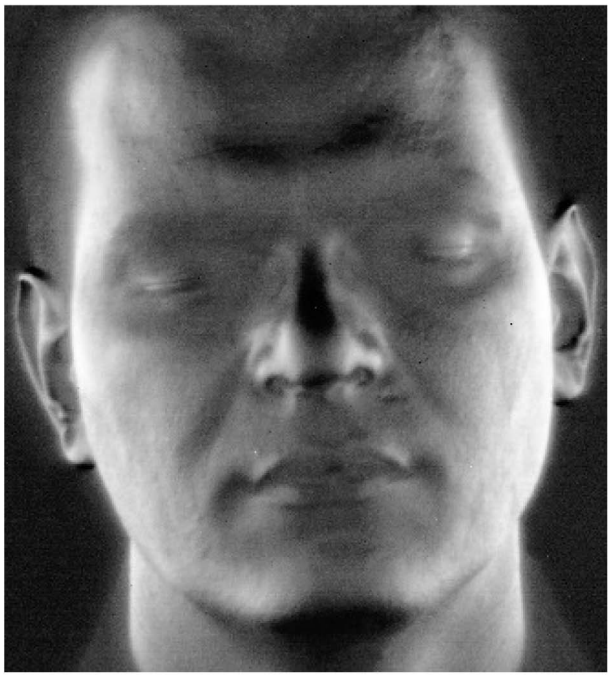
scattered from the background with the signal field emitted by the subject. In general, the scattered radiation has a different polarization state than the emission, resulting in errors in subsequent inversions. The Stokes-vector images (S_0, S_1, S_2) and the DoLP image recorded by the polarimetric camera of two subjects are shown in Figs. 1 and 2. The intensity S_0 images correlate to the temperature of the object and also to their emissivities. Perhaps most significant in remote-sensing applications is how little information can be obtained directly from these images; i.e., the facial features of the different subjects have different behaviors. For instance, the eye sockets of Subject A are significantly brighter and warmer than the rest of the face, but this characteristic is not present in Subject B. The ears of Subject A appear colder than the rest of the face, whereas they are relatively warm in Subject B. Because of



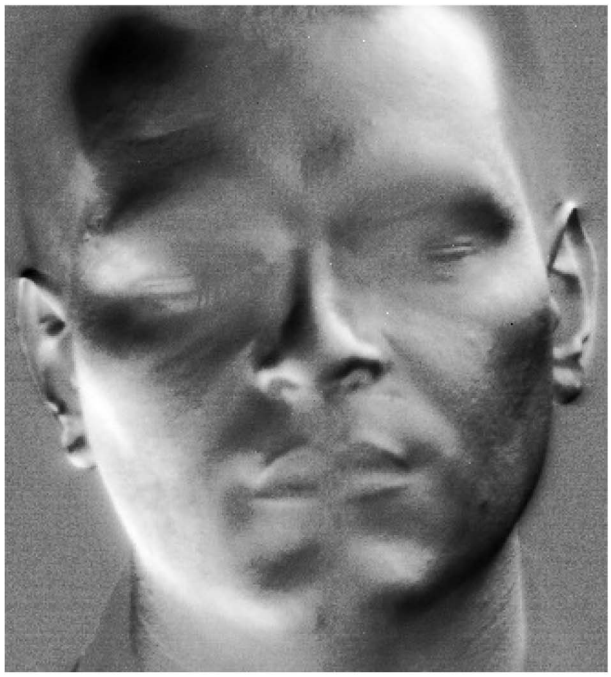
(a) S_0 image is shown.



(b) DoLP image is shown.



(c) S_1 image is shown.



(d) S_2 image is shown.

Fig. 2. Intensity and polarization imagery of Subject B.

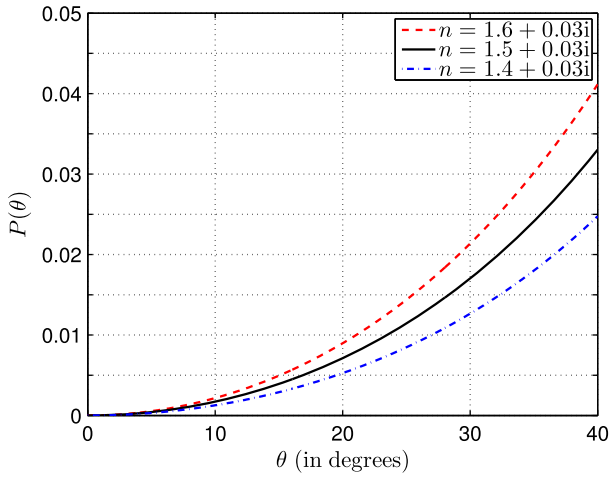
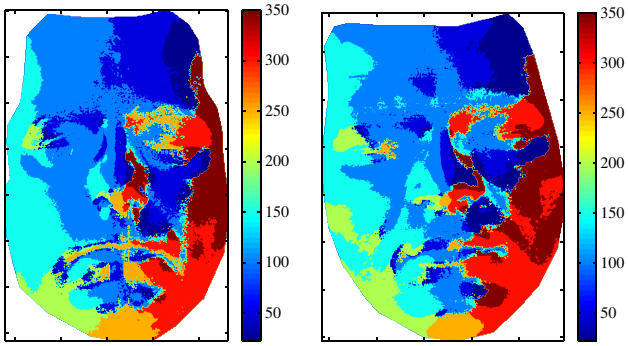


Fig. 3. Degree of linear polarization as a function of θ is shown for three refractive indices.

these inherent behavioral differences it is difficult to identify specific features of a face. Both intensity S_0 images appear splotchy and washed out, lacking in definition and texture that is seen in the DoLP image. This is not surprising because S_0 images are a measure of the total signal and do not contain polarimetric information that is sensitive to subtle changes in the skin texture and especially the orientation [1].

Clearly, the value of the θ threshold (see Section 4) depends on the index of refraction of the skin. This value may be estimated by plotting $P(\theta)$ for a number of different refractive indices of the skin. From Fig. 3, we see that $P(\theta \lesssim 10°) \approx 0$ and the $P(\theta)$ curves do not differ appreciably for $\theta \lesssim 10°$. Furthermore, from Fig. 3 we see that the curved parts of the face will tend to stretch as n decreases. Fortunately, this effect is not very pronounced because the range of the index of refraction of the skin is rather narrow (see Section 3). For the results presented below, we used $\theta_{\text{threshold}} = 11°$ and $n = 1.5 + 0.03i$.

In Fig. 4 contour plots of ϕ derived from the S_1 and S_2 images are shown. Although there are some parts near the eyes and the mouth where the algorithm described in Section 4 does not properly resolve the



(a) Subject A. (b) Subject B.

Fig. 4. Contour plots of ϕ (in degrees).

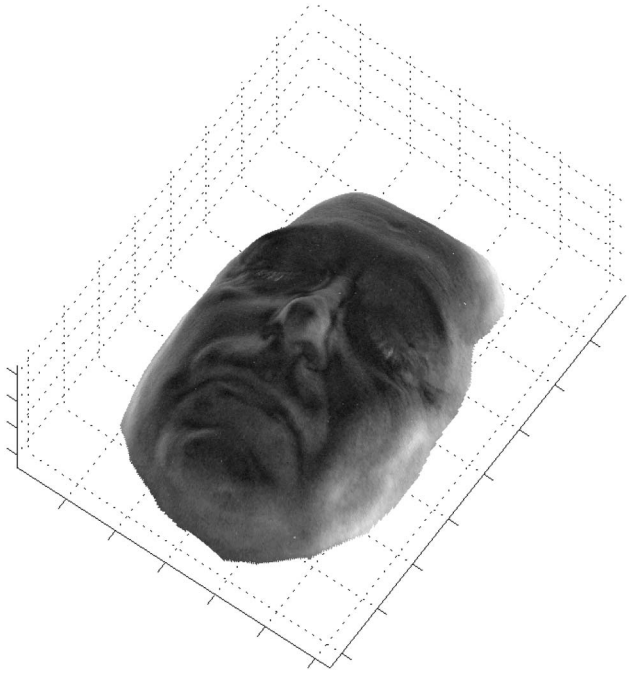


Fig. 5. 3D image of Subject A (see Media 1).

π ambiguity, overall, the algorithm performed very well. Finally, 3D images of the subjects are shown in Figs. 5 and 6 (also, see Media 1 and Media 2). To enhance the visual perception of the surface in these figures, we use the grayscale DoLP image as a texture map for the surface. From the figures we see that the main features of the face are faithfully reconstructed. However, the eyebrow region and some aspects of the eye sockets appear to be somewhat deformed. This is expected as the eyebrow hair

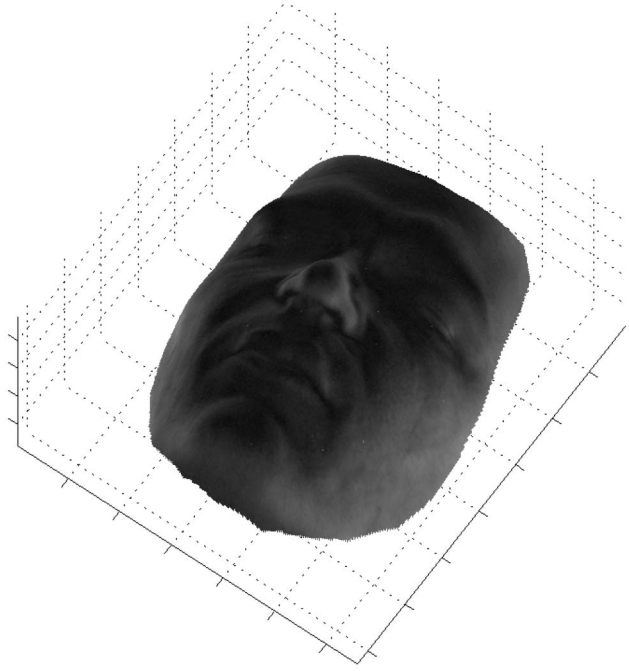


Fig. 6. 3D image of Subject B (see Media 2).

effectively causes the surface to be discontinuous. Furthermore, the indentations and arches near the eye region may produce multiple reflections of the emitted radiation, which we do not take into account in our method.

6. Conclusions

In this manuscript, we demonstrate that a 3D image of a human face may be obtained from a single Stokes-vector image of the emitted LWIR radiation. To do this, we analyze the polarization state of the emitted radiation to obtain the surface normal **N**. Then, by integrating **N**, we obtain the surface $f(x,y)$ up to an additive constant. For the integration, we use the well-known Frankot–Chellapa method [27], which requires unambiguous knowledge of the azimuthal angle ϕ. One shortcoming of this method is that the retrieved surface normals are nonunique. To address this problem, our method of producing an unambiguous ϕ relies on the occluding boundary. Unfortunately, this is precise only for a surface containing no concavities. The face represents a complex surface, and there are problematic regions in which this algorithm can be improved. It seems reasonable to conjecture that using a more sophisticated integration method, such as one used by Ecker *et al.* (2008) [29], which allows for a twofold ambiguity in ϕ, would further improve the fidelity of $f(x,y)$. We will consider this conjecture in future work.

While the methodology we have described should encounter no difficulties in a surface having no concavities, and has been demonstrated to be successful for complex surfaces that are smoothly varying but contain concavities, like a face, challenges may occur in other systems. Objects that have discontinuities in surface slopes due to sharp edges, for instance, can present a challenge that is not addressed by the current methodology.

This research was supported in part by an appointment to the U.S. Army Research Laboratory Postdoctoral Fellowship Program administered by the Oak Ridge Associated Universities through a cooperative agreement with the U.S. Army Research Laboratory.

References

1. K. P. Gurton, A. J. Yuffa, and G. W. Videen, "Enhanced facial recognition for thermal imagery using polarimetric imaging," Opt. Lett. **39**, 3857–3859 (2014).
2. B. K. P. Horn and M. J. Brooks, eds., *Shape From Shading*, Artificial Intelligence (Massachusetts Institute of Technology, 1989).
3. R. Zhang, P.-S. Tsai, J. Cryer, and M. Shah, "Shape-from-shading: a survey," IEEE Trans. Pattern Anal. Mach. Intell. **21**, 690–706 (1999).
4. J.-D. Durou, M. Falcone, and M. Sagona, "Numerical methods for shape-from-shading: a new survey with benchmarks," Comput. Vis. Image Underst. **109**, 22–43 (2008).
5. D. Miyazaki, M. Saito, Y. Sato, and K. Ikeuchi, "Determining surface orientations of transparent objects based on polarization degrees in visible and infrared wavelengths," J. Opt. Soc. Am. A **19**, 687–694 (2002).
6. D. Miyazaki, R. T. Tan, K. Hara, and K. Ikeuchi, "Polarization-based inverse rendering from a single view," in *Proceedings of Ninth IEEE International Conference on Computer Vision* (IEEE, 2003), pp. 982–987.
7. D. Miyazaki, M. Kagesawa, and K. Ikeuchi, "Transparent surface modeling from a pair of polarization images," IEEE Trans. Pattern Anal. Mach. Intell. **26**, 73–82 (2004).
8. G. A. Atkinson and E. R. Hancock, "Recovery of surface orientation from diffuse polarization," IEEE Trans. Image Process. **15**, 1653–1664 (2006).
9. G. A. Atkinson and E. R. Hancock, "Shape estimation using polarization and shading from two views," IEEE Trans. Pattern Anal. Mach. Intell. **29**, 2001–2017 (2007).
10. B. F. Jones, "A reappraisal of the use of infrared thermal image analysis in medicine," IEEE Trans. Med. Imaging **17**, 1019–1027 (1998).
11. E.-K. Ng, "A review of thermography as promising non-invasive detection modality for breast tumor," Int. J. Therm. Sci. **48**, 849–859 (2009).
12. M. Milanič and B. Majaron, "Spectral filtering in pulsed photothermal temperature profiling of collagen tissue phantoms," J. Biomed. Opt. **14**, 064024 (2009).
13. S. Liakat, A. P. M. Michel, K. A. Bors, and C. F. Gmachl, "Mid-infrared (λ=8.4–9.9 μm) light scattering from porcine tissue," Appl. Phys. Lett. **101**, 093705 (2012).
14. A. P. M. Michel, S. Liakat, K. Bors, and C. F. Gmachl, "*In vivo* measurement of mid-infrared light scattering from human skin," Biomed. Opt. Express **4**, 520–530 (2013).
15. J. S. Tyo, D. L. Goldstein, D. B. Chenault, and J. A. Shaw, "Review of passive imaging polarimetry for remote sensing applications," Appl. Opt. **45**, 5453–5469 (2006).
16. S. Cho, M. H. Shin, Y. K. Kim, J.-E. Seo, Y. M. Lee, C.-H. Park, and J. H. Chung, "Effects of infrared radiation and heat on human skin aging in vivo," J. Investig. Dermatol. Symp. Proc. **14**, 15–19 (2009).
17. S. Nickell, M. Hermann, M. Essenpreis, T. J. Farrell, U. Krämer, and M. S. Patterson, "Anisotropy of light propagation in human skin," Phys. Med. Biol. **45**, 2873–2886 (2000).
18. T. L. Troy and S. N. Thennadil, "Optical properties of human skin in the near infrared wavelength range of 1000 to 2200 nm," J. Biomed. Opt. **6**, 167–176 (2001).
19. A. N. Bashkatov, E. A. Genina, V. I. Kochubey, and V. V. Tuchin, "Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000 nm," J. Phys. D **38**, 2543–2555 (2005).
20. B. Majaron and M. Milanič, "Effective infrared absorption coefficient for photothermal radiometric measurements in biological tissues," Phys. Med. Biol. **53**, 255–268 (2008).
21. J. Steketee, "Spectral emissivity of skin and pericardium," Phys. Med. Biol. **18**, 686–694 (1973).
22. J. A. Viator, B. Choi, G. M. Peavy, S. Kimel, and J. S. Nelson, "Spectra from 2.5–15 μm of tissue phantom materials, optical clearing agents and ex vivo human skin: implications for depth profiling of human skin," Phys. Med. Biol. **48**, N15–N24 (2003).
23. D. C. Bertilone, "Stokes parameters and partial polarization of far-field radiation emitted by hot bodies," J. Opt. Soc. Am. A **11**, 2298–2304 (1994).
24. K. P. Gurton, R. Dahmani, and G. Videen, "Measured degree of infrared polarization for a variety of thermal emitting surfaces," Tech. Rep. ARL-TR-3240 (U.S. Army Research Laboratory, 2004).
25. K. P. Gurton and R. Dahmani, "Effect of surface roughness and complex indices of refraction on polarized thermal emission," Appl. Opt. **44**, 5361–5367 (2005).
26. L. B. Wolff, "Polarization vision: a new sensory approach to image understanding," Image Vis. Comput. **15**, 81–93 (1997).
27. R. T. Frankot and R. Chellappa, "A method for enforcing integrability in shape from shading algorithms," IEEE Trans. Pattern Anal. Mach. Intell. **10**, 439–451 (1988).
28. J. Ho, J. Lim, M.-H. Yang, and D. Kriegman, "Integrating surface normal vectors using fast marching method," in *Computer Vision—ECCV 2006*, A. Leonardis, H. Bischof, and A. Pinz, eds., Vol. **3953** of Lecture Notes in Computer Science (Springer Science + Business Media, 2006), pp. 239–250.
29. A. Ecker, A. D. Jepson, and K. N. Kutulakos, "Semidefinite programming heuristics for surface reconstruction

ambiguities," in *Computer Vision—ECCV 2008*, D. Forsyth, P. Torr, and A. Zisserman, eds., Vol. **5302** of Lecture Notes in Computer Science (Springer, 2008), pp. 127–140.
30. P. Bhat, B. Curless, M. Cohen, and C. L. Zitnick, "Fourier analysis of the 2D screened Poisson equation for gradient domain problems," in *Computer Vision—ECCV 2008*, D. Forsyth, P. Torr, and A. Zisserman, eds., Vol. **5303** of Lecture Notes in Computer Science (Springer, 2008), pp. 114–128.
31. J.-D. Durou, J.-F. Aujol, and F. Courteille, "Integrating the normal field of a surface in the presence of discontinuities," in *Energy Minimization Methods in Computer Vision and Pattern Recognition*, D. Cremers, Y. Boykov, A. Blake, and F. R. Schmidt, eds., Vol. **5681** of Lecture Notes in Computer Science (Springer, 2009), pp. 261–273.
32. S. Galliani, M. Breuss, and Y. C. Ju, "Fast and robust surface normal integration by a discrete eikonal equation," in *Proceedings of the British Machine Vision Conference*, R. Bowden, J. Collomosse, and K. Mikolajczyk, eds. (British Machine Vision Association, 2012), pp. 106.1–106.11.