

A collection and statistical analysis of skin reflectance signatures for inherent variability over the 250 nm to 2500 nm spectral range

Catherine C. Cooksey, Benjamin K. Tsai, and David W. Allen
Sensor Science Division
National Institute of Standards and Technology
Gaithersburg, MD, USA 20899

ABSTRACT

The spectral reflectance signature of human skin provides opportunities to advance observations ranging from medical treatment to security applications. In this study 28 volunteers participated in a skin reflectance measurement of the inside of the right forearm. The reflectance measurements were made over the 250 nm to 2500 nm spectral range. The analysis included estimates of the variability attributed to the instrument, variability of the same subject, and variability among subjects. This allowed for determining measures of similarity and differences that indicate the inherent separability within the distribution. While this sample size may not fully represent a full diverse-population, it does provide a provisional reference point for modeling and simulation.

Keywords: Skin, spectral, reflectance, traceable data, reference data, signatures, variability

1. INTRODUCTION

Spectral imaging can provide a means of identifying substances based on their spectrally distinct signatures. For instance, human skin has a number of distinct spectral features, which are fostering a wide range of current and emerging imaging applications [1,2,3]. Some examples include medical imaging, security, biometrics, and safety. For continued progress of these applications, some knowledge of what is considered a “typical” signature for the spectral reflectance of skin is required. The challenge is establishing what might be considered “typical” and how is it described. Given the diversity of the human population, it may be more appropriate to establish the range and distribution of spectral signatures in a population rather than define an average spectral signature. Knowledge of the variability of skin reflectance spectra and its scale is essential for all of the applications mentioned above.

Biological variability is routinely encountered when making physical measurements of organisms. The total variability for any given biological quantity, can be expressed as the sum of the variability due to the measurement error (i.e., instrument), the method (e.g., positioning), and the nature of inherent biological sample. Ideally, an assessment of the biological quantity is designed so that the variability due to the measurement and method are easily quantified and separable from its unknown biological variability.

The optical properties of human skin depend on its many components, such as cells, fibers, and chromophores, as well as its surface features, which include hair, freckles, wrinkles, and contours [4,5]. Thus, it is expected that the spectral reflectance of skin varies significantly, not only between individuals but within the same individual due to variation across different regions of the body, level of activity, time of day, time of year, and many other possible parameters. While this study is limited to the examination of one area, the middle of the inside of the forearm, for simplicity, other parameters, such as age, the presence of pigmentation features (freckles, moles, etc.), and the use of lotions, are not controlled in order to evaluate the biological variability of human skin.

The reflectance properties of human skin and its variability have been investigated previously [6]. This paper expands on that work by increasing the sample size and analysis methods. As with the previous study, the method of collection

is not driven by any particular use or application of the resulting reflectance spectra. Consequently, the spectral range, 250 nm to 2500 nm, is intentionally broad to encompass any number of applications that might fall within this region. The measurements are directly tied to the NIST scale of spectral reflectance, thus providing a known reference set with stated measurement uncertainties.

Various analysis methods are employed to explore the resulting data set for general trends, including spectral angle, principal component analysis, Mahalanobis Distance, and the coordinates for the International Commission on Illumination (CIE) color space. More thorough analyses would be expected for cases where a specific application is known. Although this sample size, 28 subjects, precludes drawing any conclusions on a global scale, it does, provide insight into the variable nature of skin reflectance spectra.

2. METHOD

2.1 Image collection and reflectance measurement of human subjects

The method for collecting skin reflectance measurements of human subjects has been previously established and is described in detail in Reference 6. Briefly, 12 subjects volunteered to participate in the initial round of this study and an additional 16 subjects volunteered to participate in the second round.¹ All of the subjects were federal employees. There was no attempt to select subjects based on age, gender, or ethnicity as might be related to skin tone. No subjects were excluded for the use of sun screen, body lotion, or medication, or for the presence of freckles, moles, tattoos, or skin conditions or disorders.

Each subject participated in only one measurement session, which consisted of collecting a photographic image of the test area on the subject's forearm and acquisition of 3 reflectance measurements of the test area. The photographic image of the test area was intended to provide a means to document the uniformity of the general test region and to aid in the researchers' understanding of any significant variability that might be encountered. A commercial ultraviolet/visible/near-infrared spectrophotometer was used to acquire the 8°/h spectral reflectance factor of the subject's forearm over the wavelength region of 250 nm to 2500 nm at a wavelength interval of 3 nm. Three scans were acquired for each subject. Each scan lasted approximately 3 minutes, and the subject was provided approximately three minutes to relax their arm between scans. A post was used as a grip to provide nominally equivalent repositioning.



Figure 1. On the left, the imaging setup showing the exposed sample area of the skin. The white and grey scale references surround the aperture of the mask. The camera is positioned to view through the ring illuminator. On the right, the spectrophotometer is shown with the cover in the open position to illustrate the location of the integrating sphere in relation the arm.

¹ This human subject study, "Reflectance Measurements of Human Skin" Protocol #382, was initially approved by the NIST Institutional Review Board on May 29, 2012. Re-approval for continued collection of skin reflectance measurements was granted on September 19, 2013.

2.2 Calculation of reflectance factors and associated uncertainties

The reflectance values acquired using this spectrophotometer were determined by relative measurement, requiring comparison to a reflectance standard. The standard was sintered polytetrafluoroethylene (PTFE) and its reflectance scale was determined prior to the measurement sessions through comparison measurements of sintered PTFE and pressed PTFE. The scale for spectral reflectance factor was established for pressed PTFE in the NIST Spectral Tri-function Automated Reference Reflectometer (STARR) facility [7] using the absolute method of Van den Akker [8]. The wavelength scale of the spectrophotometer was validated using the spectrophotometer's internal atomic emission lamps.

The 8°/h spectral reflectance factor R at each wavelength λ of the item was calculated from

$$R(\lambda) = \frac{S(\lambda) - S_d(\lambda)}{S_s(\lambda) - S_d(\lambda)} \cdot R_s(\lambda), \quad (1)$$

where S is the average signal from the scan of the item, S_s is the average signal from the scan of the standard, S_d is the dark signal, and R_s is the 8°/h spectral reflectance factor of the sintered PTFE standard. Dark signals were acquired once, prior to measurement sessions with subject. The final 8°/h spectral reflectance factors were obtained by averaging the values from the three scans.

The estimated measurement uncertainties for the reflectance measurements are calculated according to the procedures outlined in [9]. Sources of uncertainty are the 8°/h spectral reflectance factor of the sintered PTFE standard, the sphere geometry, the wavelength, and random effects. The uncertainty due to the standard was evaluated during the scale transfer from pressed PTFE, and determined to be 0.0045. The uncertainty due to the difference in sphere geometries of the commercial spectrophotometer used in the skin reflectance study and NIST STARR, which was used to establish the reflectance scale for the sintered PTFE standard, was evaluated by comparing reflectance factors of a NIST-owned sintered PTFE working standard measured using both instruments. The uncertainty caused by wavelength is evaluated from the derivative of the representative spectral reflectance factors of the mean. The repeatability was determined from the standard deviation of repeat measurements of the sintered PTFE standard. The expanded uncertainty ($k = 2$) is the combined (root-sum-square) uncertainty from all contributions due to systematic and random effects multiplied by a coverage factor of two [9]. The evaluated contributions of the sources of uncertainty and the expanded uncertainty ($k = 2$) of the instrument are given in Table 1. These uncertainties are representative of the manner in which the instrument was used for this study (henceforth referred to as the instrument uncertainty) and may not represent the best performance.

Table 1. Measurement uncertainties

Source of Uncertainty	Standard Uncertainty	Uncertainty Contribution
Reflectance Standard	0.0045	0.0045
Geometry	0.001	0.001
Wavelength	0.3 nm	0.0008
Repeatability	0.0003	0.0003
		Expanded Uncertainty ($k = 2$)
		0.0094

3. RESULTS

3.1 Images of test area

The images of the subject's test area are shown in Figure 2. Inspection of the images reveals that each subject's skin is generally uniform within the test area with only a minimal presence of veins and hair. The images of the test areas are provided for a qualitative analysis only.



Figure 2. Images of each subject's test area are shown for a qualitative assessment of the uniformity.

The photographic images enable a visual assessment of the uniformity that might be attributed to texture, hairs, and pigmentation, as examples. A cursory estimation of the uniformity was made using the green band from the collected images. The standard deviation of all of the pixels in the image was divided by the mean of all of the pixels. The estimated uniformity is provided in Table 2. A significant amount of the non-uniformity occurred near the edge of the image suggesting an edge effect related to the mask and/or illumination. Consequently, these estimated values should be considered from a conservative perspective.

Table 2. The estimated uniformity of all subjects based on the green band from the images.

Estimated Uniformity (%)	
Minimum	2.2
Maximum	6.8
Mean	3.6

3.2 Reflectance spectra of human subjects

Selected reflectance spectra from the set of spectra collected from the 28 human subjects in this study are shown in Figure 3. The spectrum depicted by the black dashed curve is representative of the mean spectral reflectance values of all subjects participating in this study. The representative spectrum was selected from among the set of reflectance spectra based on its similarity to the mean spectrum using the following equation:

$$\theta_i = \cos^{-1} \left(\frac{s_m^T s_i}{\|s_m\| \|s_i\|} \right) \quad (2)$$

S_m is the mean spectrum of reflectance factors for the full set of scans acquired (84 scans) and S_i is an individual spectrum from the set of scans. The difference between the spectra, θ_i , is reported in radians. The spectrum with the smallest resulting angle is considered to be the closest match to the mean spectrum, and is selected as the representative of the mean. Selecting a representative spectrum from the overall set prevented the loss of spectral features that would have resulted from averaging the small shifts inherent in the spectral variability.

The population variability observed for the set of subjects was calculated using the standard deviation of the full set of scans. This variability is depicted in Figure 3 by the grey shaded area about the representative of the mean. For wavelengths less than 1125 nm, there is considerable variation in reflectance factors among the set of human subjects. This is further captured by the spectra (gray solid) provided in Figure 3, which represent the total range of observed reflectance factors in the ultraviolet (UV), visible (Vis), and near infrared (NIR) spectral regions.

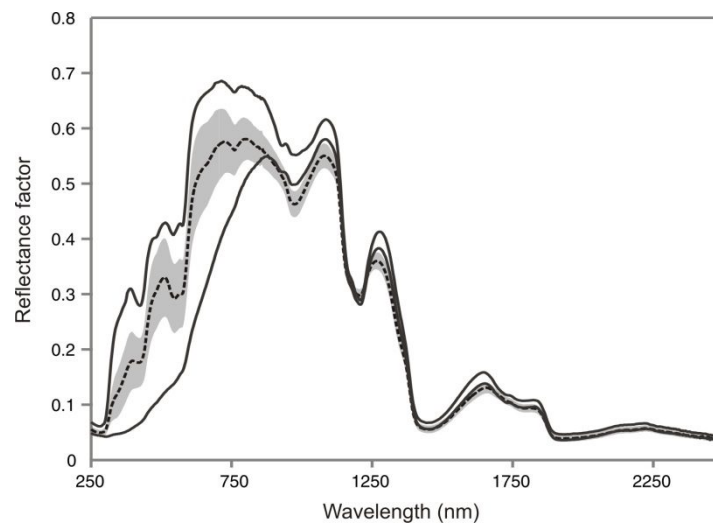


Figure 3. The reflectance spectrum of the representative of the mean (black dashed) with grey shaded area representing the population variability for all subjects. Representative spectra showing the range of variation in reflectance factors in the visible region (grey solid) are also depicted.

Figure 4 plots the population variability of this set of human subjects along with subject variability and instrument uncertainty. The subject variability was calculated using the standard deviation of the three scans of each subject. It represents the dynamic nature of human skin observed for each subject. The instrument uncertainty is described above in Section 2.2. Overall, the population variability is the most significant source of uncertainty for the skin's spectral signature over the UV-Vis-NIR region. In the shortwave infrared (SWIR), the instrument uncertainty is the dominant source of uncertainty.

The first derivative of each subject's mean reflectance spectrum is shown in Figure 5. This plot aids in defining the most significant absorption features. When comparing the full set of our subject's spectra, it is often difficult to distinguish the common spectral features among the varying levels of reflectance factor. The first derivative easily enables viewing of the peak locations, and provides a quantitative method for grouping spectra with similar spectral features. As noted in the lower panel of Figure 5, some of the features are not as easily identifiable for some subjects.

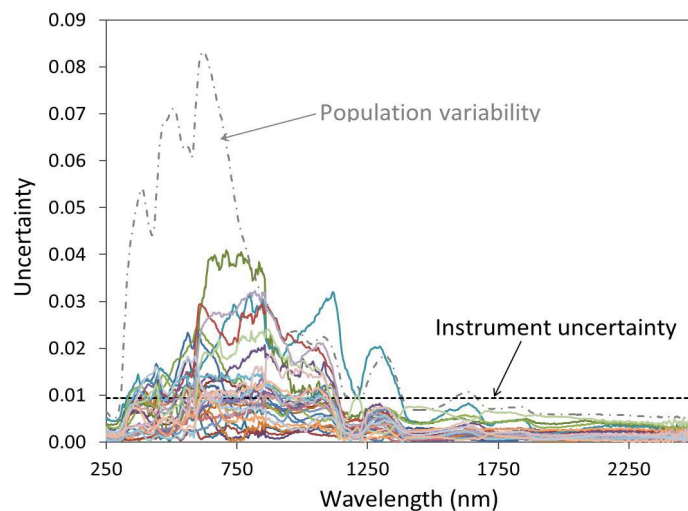


Figure 4. The instrument uncertainty (black dashed line), subject variability for each subject (solid colored lines) and population variability for all subjects (grey dot-dashed line).

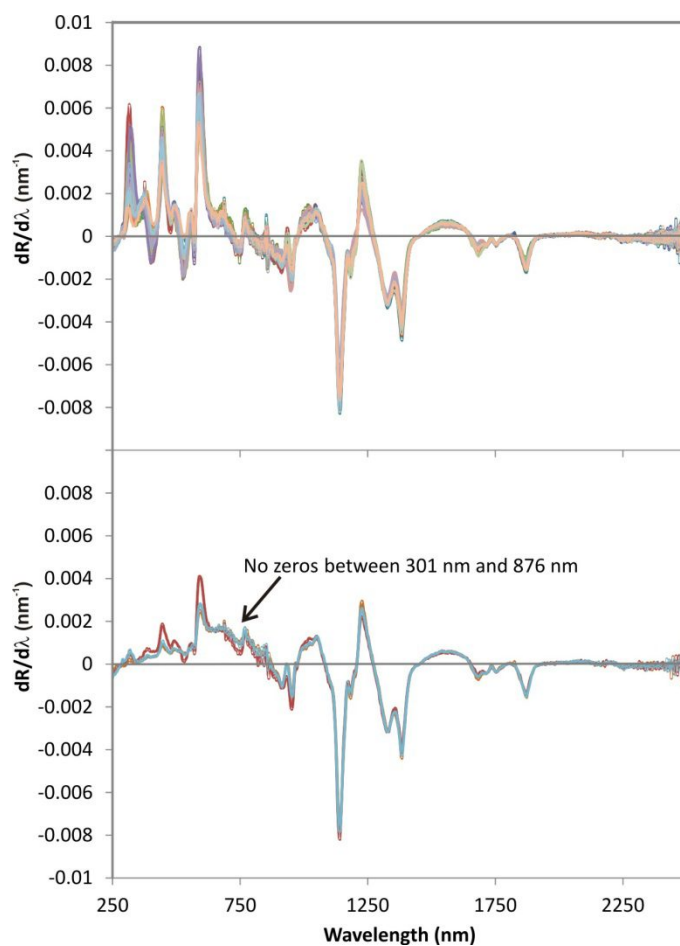


Figure 5. The first derivative of each subject's mean spectrum showing the reflectance features. Not all subjects displayed prominent spectral features in the visible and near-infrared regions.

4. ANALYSIS

Spectral angle is a simple metric that quantifies the difference between spectra when treated as vectors [10]. It is sensitive to differences in spectral shape but insensitive to overall lightness or darkness. The spectral angle for all reflectance spectra was calculated according to Equation 2 for three arbitrary spectral regions (roughly corresponding to common detector responses) and the overall spectral range. The resulting spectral angles had the largest divergence in the UV, followed by the VNIR and the SWIR, respectively. The full range is approximately equal to averaging the three individual ranges.

Table 3. The divergence of spectral angles, in radians, with respect to the mean spectrum of the full data set. The smaller the spectral angle the closer the match to the mean spectrum.

Wavelength Range (nm)	Spectral Angle (radians)			
	Minimum	Maximum	Mean	Standard Deviation
UV: 250 to 403	0.013	0.362	0.090	0.079
VNIR: 403 to 1002	0.007	0.260	0.053	0.059
SWIR: 1002 to 2500	0.008	0.110	0.028	0.015
Full: 250 to 2500	0.015	0.243	0.059	0.054

In the field of hyperspectral imaging, principal component analysis (PCA) is often used for dimensionality reduction, but can also indicate the inherent variance of a given data set [11]. This statistical procedure transforms a data set of possibly correlated variables into a set of linearly uncorrelated variables, or principal components, such that the first component has the largest possible variance and each succeeding (orthogonal) component has less variance than the previous. The analysis is applied to the full spectral range of skin reflectance data set. The resulting eigenvalues are plotted versus their respective PCA band in Figure 6. The cumulative variance of each band is also plotted. This shows that nearly all of the variance (99.5 %) within the skin reflectance data set occurs within the first 7 PCA bands. Conservatively, this suggests that there are at least two spectrally distinct groups and possibly more.

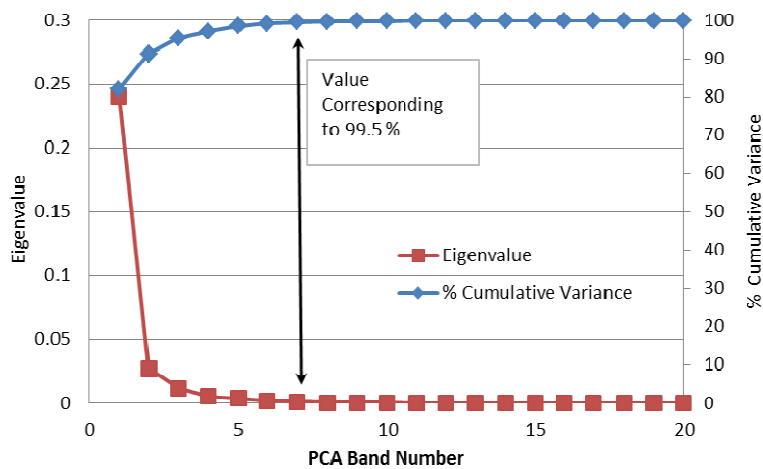


Figure 6. A plot of Eigenvalue versus PCA band number (left vertical axis) and cumulative variance versus PCA band number (right vertical axis). The double arrow indicates the point at which 99.5 % of all variance is included, corresponding to PCA band 7.

The Mahalanobis Distance (MD), which is often used in target detection algorithms, provides another way of measuring the difference, or distance, between spectra in hyperspace [12]. While the spectral angle method considers each spectrum as a discrete point in n-dimensional space, MD includes the covariance matrix from the data set as a means of

describing the distribution inherent in a specific data set. (This method is the basis for the matched filter widely used in hyperspectral imagery analysis.) The Mahalanobis Distance Δ is calculated using

$$\Delta^2 = (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Gamma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \quad (3)$$

where \mathbf{x} is a reflectance spectrum treated as an n-dimensional vector, $\boldsymbol{\mu}$ is the mean vector of the data set, and $\boldsymbol{\Gamma}$ is the covariance matrix.

The MD analysis is applied only to the spectral subset ranging from 400 nm to 1000 nm, which corresponds to 82 bands. The resulting MD for this subset can then be compared to the MD obtained for a model data set that has the same mean and variance, on a band-by-band basis, as the collection of skin spectra, but with a normal multivariate probability distribution. The comparison is shown in Figure 7 as the probability of exceedance. The distribution of the data set of skin reflectance spectra collected in this study departs from that of the model for MD values less than 9.0. This may be attributed to a bimodal distribution for the skin reflectance data set.

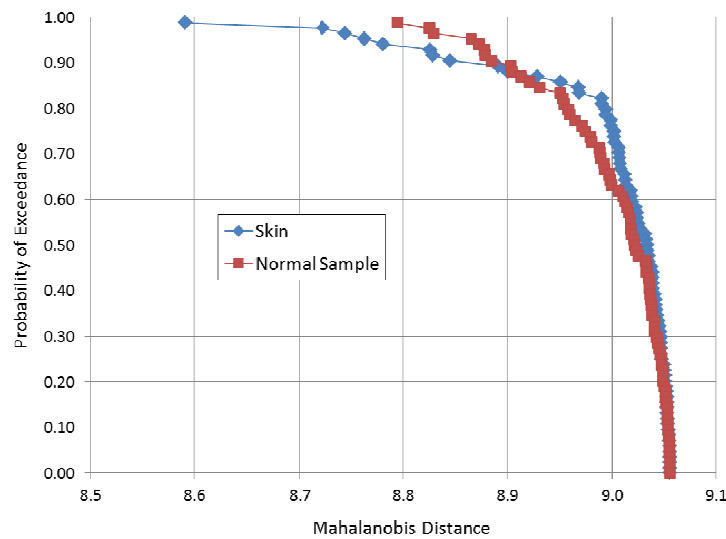


Figure 7. A plot of the probability of exceedance vs. the Mahalanobis distance (MD) of the spectral reflectance of skin. For comparison a normal sample distribution is also shown.

In relation to human vision, the spectral features of each subject in the visible range can be described using the CIE color space coordinates, commonly referred to as CIELAB or CIE $L^*a^*b^*$ [13]. These coordinates quantify visually discernable differences in color and lightness. To determine the coordinates, the reflectance factors, $R(\lambda)$, from the skin reflectance spectra for the 28 subjects are linearly interpolated to wavelengths ranging from 360 nm to 830 nm in increments of 5 nm. Then, the tristimulus values, X , Y , and Z , for a 2° field of view and D65 illuminant are calculated according to

$$X = k \int_{\lambda} R(\lambda) S(\lambda) \bar{x}(\lambda) d\lambda \quad (3)$$

$$Y = k \int_{\lambda} R(\lambda) S(\lambda) \bar{y}(\lambda) d\lambda \quad (4)$$

$$Z = k \int_{\lambda} R(\lambda) S(\lambda) \bar{z}(\lambda) d\lambda \quad (5)$$

where S is the spectral power distribution of the D₆₅ illuminant, the \bar{x} , \bar{y} , and \bar{z} are the red, green, and blue color-matching functions of the CIE 1931 Colorimetric scale, respectively, and k is the normalization factor. The corresponding tristimulus values for a white reference perfect diffuser are calculated separately as X_n , Y_n , and Z_n . When X/X_n , Y/Y_n , and $Z/Z_n > 0.008856$, the CIELAB coordinates are defined and calculated as

$$L^* = 116 \left(\frac{Y}{Y_n} \right)^{1/3} - 16 \quad (6)$$

$$a^* = 500 \left[\left(\frac{X}{X_n} \right)^{1/3} - \left(\frac{Y}{Y_n} \right)^{1/3} \right] \quad (7)$$

$$b^* = 200 \left[\left(\frac{Y}{Y_n} \right)^{1/3} - \left(\frac{Z}{Z_n} \right)^{1/3} \right] \quad (8)$$

The coordinate L^* represents the lightness of the color where $L^* = 0$ indicates black and $L^* = 100$ indicates diffuse white. The coordinates a^* and b^* represent color along the magenta-green and yellow-blue continuums, respectively. The resulting CIELAB coordinates for all 28 subjects are plotted in Figures 8 and 9.

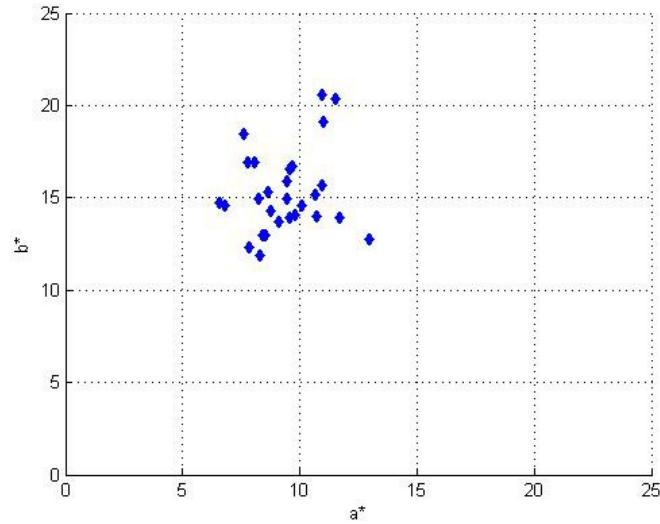


Figure 8. A two dimensional scatter plot of CIE a^* and b^* values for all subjects.

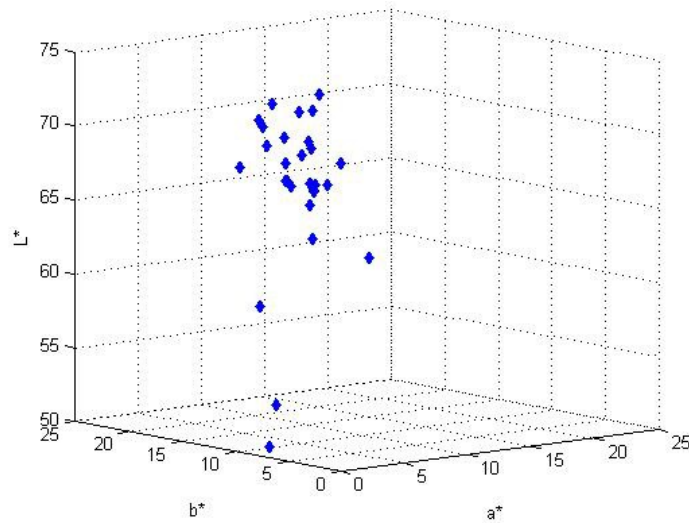


Figure 9. A three dimensional plot of the CIE $L^*a^*b^*$ values for all subjects.

In the two-dimensional plot of the a^* and b^* coordinates, all subjects are clustered in the same region of the positive a^*b^* quadrant (magenta-yellow, respectively) with the possible separation of several subjects from the main cluster.

However, the three-dimensional plot of $L^*a^*b^*$ shows more distinctly separation between the main cluster and several subjects. These results only provide a glimpse of the volume of color space that might be filled by a larger sample of the population.

5. SUMMARY

This work addresses the need for establishing skin reflectance spectral signatures and their inherent variability. The method presented here provides one model for sampling and evaluating the differences in skin reflectance spectra. The variability values attributed to measurement error and method are significantly smaller than the variability between any two individuals. This suggests that the overall variability can primarily be attributed to biological variability. Although the sample size is small, there are distinct trends that should be noted for applications where skin reflectance spectra are key features of interest. Visual inspection of the complete set of reflectance spectra reveals that the greatest variability in the reflectance of human skin occurs in the VNIR. In the SWIR, the variability across the data set is on the order of, or less than, the instrument variability. Additionally, not all of the spectral features are as easily identifiable for some subjects as they are for others. The variability of the reflectance signatures for human skin can be described using the results of several analysis methods. Spectral angle revealed that the UV region is the most variable and the SWIR is the least. The remaining analysis methods, PCA, MD, and CIELAB, point to the possible existence of several spectrally distinct groups of reflectance signatures.

It is not known what sample size or source of participants would best represent the population at-large. It can be expected that the variability is no smaller than the distribution presented here. In further exploring the inherent variability of skin reflectance spectra, any one study will likely be limited and biased to local demographics. In order to address this issue, the authors will seek to expand this work to include studies performed elsewhere as a meta-analysis. Differences in a meta-analysis might be explained by the sources of variability discussed in the introduction and similarities might suggest that a sufficiently large sample size has been achieved.

ACKNOWLEDGEMENTS

The authors would like to thank Nicholas Paulter and Ronald Resmini for support and guidance on this research. We would also like to acknowledge the volunteers who took the time to contribute and were essential to this study.

*Note: References are made to certain commercially available products in this paper to adequately specify the experimental procedures involved. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that these products are the best for the purpose specified.

REFERENCES

- [1] Cooksey, C. C., Neira, J. E., and Allen, D.W., "The evaluation of hyperspectral imaging for the detection of person-borne threat objects over the 400nm to 1700 nm spectral region," Proc. SPIE 8357, 83570O (2012).
- [2] Cooksey, C.C. and Allen, D. W., "Investigation of the potential use of hyperspectral imaging for stand-off detection of person-borne IEDs," Proc. SPIE. 8017, 80171W (2011).
- [3] Xu, R. X., Allen, D. W., Huang, J., Gnyawali, S., Melvin, J., Elgharably, H. Godillo, G., Huang, K. Bergdall, V., Litorja, M., Rice, J. P., Hwang, J., and Sen, C. K., "Developing digital tissue phantoms for hyperspectral imaging of ischemic wounds," Biomedical Optics Express, 3(6), 1433-1445 (2012).
- [4] Igarashi, T., Nishino, K., and Nayar, S. K., "The appearance of human skin," Columbia University Technical Report CUCS-024-05 (2005).
- [5] Vo-Dinh, T., [Biomedical Photonics: Handbook], CRC Press, Washington, D.C., (2003).

- [6] Cooksey, C. C. and Allen, D. W., "Reflectance measurements of human skin from the ultraviolet to the shortwave infrared (250 nm to 2500 nm)," *Proc. SPIE* 8734, 87340N (2013).
- [7] Barnes, P. Y., Early, E. A., and Parr, A. C., "NIST Measurement Services: Spectral Reflectance," NIST Special Publication 250-48 (1998).
- [8] Venable, W. H., Hsia, J. J., and Weidner, V. R., "Establishing a Scale of Directional-Hemispherical Reflectance Factor I: The Van den Akker Method," *J. Research of the Natl Bureau of Stnds*, 82, 29 (1977).
- [9] Taylor, B. N. and Kuyatt, C. E., "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results," NIST Technical Note 1297 (1994).
- [10] Kruse, F. A., Lefkoff, A. B., Boardman, J. B., Heidebrecht, K. B., Shapiro, A. T., Barloon, P. J., and Goetz, A. F. H., "The Spectral Image Processing System (SIPS) - Interactive Visualization and Analysis of Imaging spectrometer Data," *Remote Sensing of the Environment*, 44, 145-163 (1993).
- [11] Richards, J. A., [Remote sensing digital image analysis: An introduction], Springer-Verlag, Berlin, Germany (1999).
- [12] Manolakis, D., Marden, D., and Shaw, G. A., "Hyperspectral image processing for automatic target detection applications," *Lincoln Laboratory J.* 14, 79 (2003).
- [13] ASTM E308-13, "Standard Practice for Computing the Colors of Objects by Using the CIE System."