Illumination-invariant face recognition in hyperspectral images

Zhihong Pan, Glenn Healey, Manish Prasad, Bruce Tromberg[a]

Department of Electrical Engineering and Computer Science
[a]Beckman Laser Institute
University of California, Irvine, CA

ABSTRACT

We examine the performance of illumination-invariant face recognition in hyperspectral images on a database of 200 subjects. The images are acquired over the near-infrared spectral range of 0.7-1.0 microns. Each subject is imaged over a range of facial orientations and expressions. Faces are represented by local spectral information for several tissue types. Illumination variation is modeled by low-dimensional linear subspaces of reflected radiance spectra. One hundred outdoor illumination spectra measured at Boulder, Colorado are used to synthesize the radiance spectra for the face tissue types. Weighted invariant subspace projection over multiple tissue types is used for recognition. Illumination-invariant face recognition is tested for various face rotations as well as different facial expressions.

1. INTRODUCTION

Biometric identification techniques, especially face recognition, have been accepted as methods which can contribute to homeland security. Ideally, a face recognition system should identify human subjects under unconstrained conditions, such as indoor or outdoor illumination and arbitrary face orientation and occlusion. The ability to operate with a large distance from the camera to the subject is also highly desirable. Current face recognition systems use primarily spatial discriminants that are based on geometric facial features.[1-5] Many of these systems have performed well on databases acquired under controlled conditions.[6,7] However, these approaches often exhibit significant performance degradation in the presence of changes in face orientation. One study,[8] for example, showed that there is significant degradation in recognition performance for images of faces that are rotated more than 32° from a frontal image that is used to train the system. A more recent study,[9] which uses the light-fields model for pose-invariant face recognition, showed promising results on faces rotated over 60° at the cost of a large computational requirement. Algorithms that use geometric features can also perform poorly when subjects are imaged at different times. For example, recognition performance can degrade by as much as 20% when imaging sessions are separated by a two week interval.[8] Face orientation is another obstacle for traditional 2D face recognition. A 3D morphable face model has been used for face identification across different poses.[10] This approach has provided promising performance on a 68 subject dataset. At the current time, however, this system is computationally intensive and requires considerable manual intervention. Partial face occlusion also brings poor performance. A method[11] which divides the face into regions for isolated analysis can tolerate up to 1/6 face occlusion without losing accuracy. Geometric feature based algorithms have performed surprisingly well for face recognition under different indoor illumination conditions.[7] Under unknown outdoor illumination, however, the verification rate of the best systems can degrade by as much as 40%.[7] There is further performance degradation if the head pose variation and different illumination conditions are combined.[8]

Several of the above limitations for current face recognition systems can be overcome by using spectral information of human face tissue. Spectroscopy has been used in remote sensing applications[12] for several years. After hyperspectral cameras have become economically accessible, computational methods have been studied for applications like face recognition.[13] Face recognition in hyperspectral images[13] has been based on

---

Further author information: (Send correspondence to G. Healey)
G. Healey: E-mail: healey@ece.uci.edu; Z. Pan: E-mail: zpan@ece.uci.edu

near-infread (NIR) spectral properties of human tissue that are different from person to person while remaining reasonably constant for different face orientations and over time. The NIR spectral measurements can be extracted from different locations in the hyperspectral face images. Based on experiments with 200 subjects, this method is able to identify accurately human subjects regardless of facial expression and pose. It also has potential to identify human subjects from hyperspectral images taken on different days.

Linear models have been used to represent spectral properties of illumination as well as reflectance by physics-based image analysis algorithms. The seminal study of Judd, MacAdam, and Wyszecki[14] was the first in a series of analyses that considered the use of linear models for outdoor illumination spectra. The study concluded that a set of 622 spectra measured in different locations could be well approximated by a three-dimensional linear model. Subsequent studies[15–17] that analyzed sets of spectra measured in different locations or synthesized using atmospheric models[18] consistently concluded that low-dimensional models provide an accurate fit for outdoor illumination spectra. The largest set of outdoor illumination spectra that has been analyzed to date consists of 7258 spectra collected in Boulder, Colorado over the 0.35-2.2 μm spectral range in 1997.[19] This analysis showed that a low-dimensional linear model is able to represent accurately a large set of outdoor illumination spectra over both the visible and 0.4-2.2 μm spectral ranges. Related ten-dimensional linear models are used to represent the reflected radiance spectra for each of 233 materials. The reflected radiance spectrum of an unknown material can be compared with these linear subspaces and the study showed that accurate material identification is achieved in over 99% of all tests. A similar technique can be used for face recognition under unknown illumination.

In this paper, we present experimental results on recognizing 200 human subjects under unknown illumination in hyperspectral face images. For each subject, several near-infrared images were acquired under different poses and expressions. Spectral reflectances of different tissue types were estimated. One hundred randomly selected outdoor illumination spectra from the Boulder dataset were used to synthesize reflected radiance spectra as well as to generate a low-dimensional linear model for each tissue type for each subject. Recognition under unknown illumination is achieved by projecting reflected radiance spectra of different tissue types onto the linear models for each subject.

2. LINEAR MODEL OF GLOBAL IRRADIANCE SPECTRA

The global spectral irradiance $L(\lambda)$ includes all of the sunlight and skylight that is incident on a surface. The function $L(\lambda)$ will also be called an illumination function. Suppose that $L_1(\lambda), L_2(\lambda), \ldots, L_M(\lambda)$ is a set of illumination functions where the subscript denotes a particular set of conditions. If the illumination functions are sampled at W wavelengths, then each illumination function $L_i(\lambda)$ is represented by the W-dimensional vector $\mathbf{L}_i = (L_i(\lambda_1), L_i(\lambda_2), \ldots, L_i(\lambda_W))^T$. Following the work of Judd[14] and subsequent researchers, we can approximate $\mathbf{L}_i$ over a range of wavelengths using

$$\mathbf{L}_i \approx \sum_{j=1}^{N} \sigma_{ij} \ell_j \tag{1}$$

where the W-dimensional vectors ℓ_j define a fixed basis for the M illumination vectors and the constants σ_{ij} are weighting coefficients. The approximation in (2) typically enables each of the M spectra to be represented accurately with N coefficients where $N \ll W$.

We define the quality of the approximation in (1) by the squared error

$$E_i = \|\mathbf{L}_i - \sum_{j=1}^{N} \sigma_{ij} \ell_j\|^2 \tag{2}$$

For a set $\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_M$ of illumination vectors corresponding to different conditions, the total squared error associated with a set of basis vectors is
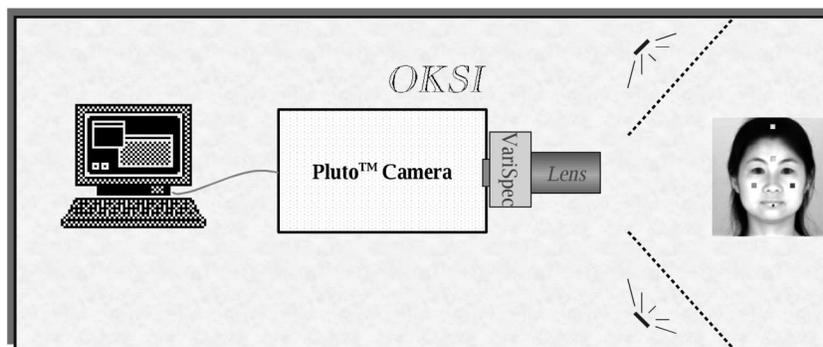
$$E_T = \sum_{i=1}^{M} E_i \tag{3}$$

Figure 1: Hyperspectral imaging system set-up

Figure 2: Thirty-one bands for a hyperspectral image of one subject

Given the set $\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_M$, the singular value decomposition (SVD)[20] can be used to generate an orthonormal set of basis vectors $\ell_1, \ell_2, \ldots, \ell_N$ that minimize E_T for any N if we let $\sigma_{ij} = \mathbf{L}_i \cdot \ell_j$.

3. DATA COLLECTION AND REFLECTANCE SPECTRAL CALIBRATION

The hyperspectral imaging system utilized for data collection is set up as in figure 1. The hyperspectral camera from Opto-Knowledge Systems, Inc. (OKSI) is based on a liquid crystal tunable filter[21] made by Cambridge Research Instruments (CRI). All images were captured with 31 bands sampled every 0.01μm over the near-infrared (0.7μm-1.0μm) with 468×494 spatial resolution. Figure 2 displays one example of all 31 bands for one subject. The 31 bands are shown in ascending order from left to right and from top to bottom.

In order to convert the raw images acquired by the hyperspectral camera to spectral reflectance images for analysis, two spectralon panels were used during calibration. A panel with approximately 99% reflectance is referred to as white spectralon and a panel with approximately 2% reflectance is referred to as black spectralon. The raw measurement obtained by the hyperspectral imaging system at spatial coordinate (x, y) and wavelength λ_k is given by

$$I(x, y, \lambda_k) = L(x, y, \lambda_k)S(x, y, \lambda_k)R(x, y, \lambda_k) + O(x, y, \lambda_k) \tag{4}$$

where $L(x, y, \lambda_k)$ is the illumination, $S(x, y, \lambda_k)$ is the system spectral response, $R(x, y, \lambda_k)$ is the reflectance of the viewed surface, and $O(x, y, \lambda_k)$ is the offset which includes dark current and stray light. To obtain the spectral reflectance image $R(x, y, \lambda_k)$, we take hyperspectral images of the white and black spectralon to get the
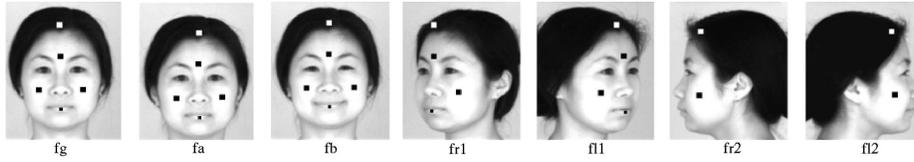
Figure 3: Examples of images with different expressions and rotations

raw measurements $I_W(x, y, \lambda_k)$ and $I_B(x, y, \lambda_k)$ respectively. Both measurements are averaged over 10 images. With the averaged I_W and I_B, we can calibrate to reflectance[13] as

$$R(x, y, \lambda_k) = \frac{(I(x, y, \lambda_k) - I_B(x, y, \lambda_k)) R_W(\lambda_k)}{I_W(x, y, \lambda_k) - I_B(x, y, \lambda_k)} + \frac{(I_W(x, y, \lambda_k) - I(x, y, \lambda_k)) R_B(\lambda_k)}{I_W(x, y, \lambda_k) - I_B(x, y, \lambda_k)} \quad (5)$$

We collected hyperspectral face images of 200 human subjects. Images of all human subjects were acquired in sets of seven images per subject. Figure 3 shows the seven images for one subject. Two front-view images were taken with neutral expression (fg and fa). Another front-view image fb was taken with a different expression. Four other images were taken with face orientations of -90°, -45°, 45°, and 90°, which are referred to as fr2, fr1, fl1, and fl2 respectively.

In order to test the feasibility of hyperspectral face recognition under unknown illumination, we represent each face image using spectral reflectance vectors that are extracted from small facial regions. Squares overlayed on the images in figure 3 indicate the size and location of the regions that are considered for each subject. Up to five facial regions corresponding to the forehead, left cheek, right cheek, hair, and lips are used in each of the seven images, depending on the visibility.

For each facial region, the spectral reflectance vector $\mathbf{R}_t = (R_t(\lambda_1), R_t(\lambda_2) \cdots, R_t(\lambda_B))^T$ is estimated by averaging over the $\mathbf{P}$-pixel squares shown in figure 3 according to

$$R_t(\lambda_k) = \frac{1}{\mathbf{P}} \sum_{x,y} R(x, y, \lambda_k) \qquad k = 1, 2, \ldots, B \quad (6)$$

where the sum is over the $\mathbf{P}$ pixels in the square, B is the number of spectral bands, and t is one of the following tissue types: f(forehead), lc(left cheek), rc(right cheek), h(hair), or l(lip). The normalized spectral reflectance vector $\overline{\mathbf{R}}_t$ is defined by

$$\overline{\mathbf{R}}_t = \mathbf{R}_t / \|\mathbf{R_t}\| \quad (7)$$

4. SPECTRAL METRIC FOR FACE RECOGNITION

For the reflectance vector $\overline{\mathbf{R}}_t(i)$ of one frontal face image of subject i and a set of illumination vectors $\mathbf{L}_1, \mathbf{L}_2, \ldots, \mathbf{L}_M$, we can generate M reflected radiance vectors $\Re_t(i, 1), \Re_t(i, 2), \ldots, \Re_t(i, M)$ where the value of vector $\Re_t(i, m)$ at λ_k is the product of the values of $\mathbf{L}_m$ and $\overline{\mathbf{R}}_t(i)$ at λ_k respectively. These vectors can be represented compactly by an orthonormal set of N basis vectors $\mathbf{b}_{t1}(i), \mathbf{b}_{t2}(i), \ldots, \mathbf{b}_{tN}(i)$ that minimizes the total squared error for any N as in section 2. We can approximate a vector $\overline{\Re}_t(j, m)$, which is a normalized vector computed from $\Re_t(j, m)$ as in (7), using the N-dimensional linear subspace for subject i as

$$\overline{\Re}_t(j, m) \approx \sum_{n=1}^{N} (\overline{\Re}_t(j, m) \cdot \mathbf{b}_{tn}(i)) \mathbf{b}_{tn}(i) = \hat{\Re}_t(i, j, m) \quad (8)$$

where $\hat{\Re}_t(i, j, m)$ is the projection of $\overline{\Re}_t(j, m)$ onto the linear subspace for subject i. The distance from $\overline{\Re}_t(j, m)$ to face image i for tissue type t under illumination $\mathbf{L}_m$ is defined by the square of the Mahalanobis distance[22]

$$D_t(i, j, m) = \left(\overline{\Re}_t(j, m) - \hat{\Re}_t(i, j, m) \right)^T \Sigma_t^{-1} \left(\overline{\Re}_t(j, m) - \hat{\Re}_t(i, j, m) \right) \quad (9)$$

where Σ_t is the covariance matrix for the distribution of projection error between the reflected radiance vector $\overline{\Re}_t$ for a subject i and $\hat{\Re}_t(i, i, m)$. Note that we use a single Σ_t to represent variability for tissue type t over the entire database of subjects. We also approximate Σ_t as a diagonal matrix for efficiency. For every face image of subject i and a given illumination $\mathbf{L}_m$, we compute the projection error vector $\overline{\Re}_t(i, m) - \hat{\Re}_t(i, i, m)$. The value of this vector at each λ_k is squared to get a variance vector $\mathbf{V}_t(i, m)$. By averaging the variance vectors $\mathbf{V}_t(i, m)$ for all images of subject i that includes tissue t over the M illumination spectra, we get the variance vector $\overline{\mathbf{V}}_t(i)$ for subject i. The diagonal elements of Σ_t are approximated by averaging the elements of $\overline{\mathbf{V}}_t(i)$ over all subjects i.

Recognition performance can be enhanced by utilizing all visible tissue types. Thus, the distance between a frontal face image i and a test face image j is defined as

$$D(i, j, m) = \omega_f D_f(i, j, m) + \omega_{lc} D_{lc}(i, j, m) + \omega_{rc} D_{rc}(i, j, m) + \omega_h D_h(i, j, m) + \omega_l D_l(i, j, m) \tag{10}$$

where ω_t is 1 if tissue type t is visible in the test image and 0 otherwise.

For the analysis above, we have assumed that the calibrated normalized reflectance spectra are invariant to face orientation. This assumption may not be correct for large face rotations. Generally, the spectra of rotated faces tend to be flatter than the front view spectra. Therefore, we can compensate the spectra of faces rotated 45° or 90° using singular value decomposition (SVD)[20] techniques. For each of the C subjects, we have 3 normalized spectral reflectance vectors: $\overline{\mathbf{R}}_f^0(i)$ for the front view forehead spectrum, $\overline{\mathbf{R}}_f^1(i)$ and $\overline{\mathbf{R}}_f^2(i)$ for 45° right and left face rotation respectively. We can get $2C$ spectral variation vectors $\mathbf{U}(j)$ $j = 1, \ldots, 2C$ by subtracting $\overline{\mathbf{R}}_f^0(i)$ from $\overline{\mathbf{R}}_f^1(i)$ and $\overline{\mathbf{R}}_f^2(i)$ respectively. Applying the SVD to these $2C$ spectral variation vectors we can generate an orthonormal set of basis vectors $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_N$ which characterize the spectral variation vectors. If $\mathbf{b}_1$ accounts for most of the spectral variation, then we can approximate the spectral variation by

$$\mathbf{U}(j) = \sum_{b=1}^{N} \epsilon_b(j) \mathbf{b}_n \approx \epsilon_1(j) \mathbf{b}_1 \tag{11}$$

where the coefficients $\epsilon_b(j) = \mathbf{U}(j) \cdot \mathbf{b}_n$. To account for this variation, we adjust each 45° spectrum as

$$\hat{\mathbf{R}}_f^1(i) = \overline{\mathbf{R}}_f^1(i) - \epsilon_1 \mathbf{b}_1 \tag{12}$$

$$\hat{\mathbf{R}}_f^2(i) = \overline{\mathbf{R}}_f^2(i) - \epsilon_1 \mathbf{b}_1 \tag{13}$$

where ϵ_1 is the average of the $\epsilon_1(j)$. For other tissue types as well as other face rotations, similar adjustments are utilized.

5. EXPERIMENTAL RESULTS

We conducted a series of recognition experiments using an image database consisting of $C = 200$ subjects. At each imaging session, seven images of each subject were acquired as shown in figure 3. The images then were calibrated to generate the spectral reflectance images. Image fg is used to represent the subject in the *gallery* set which is the group of hyperspectral images of known identity.[6] The remaining images are used as *probes* to test the recognition algorithm. Thus, the experiments follow the *closed universe* model[6] where the subject in every image in the probe set is present in the gallery.

We used the outdoor illumination spectra from the Boulder dataset to synthesize the reflected radiance spectra. One hundred of the Boulder spectra were used for the experiments in this paper. Three example spectra are shown in figure 4. One spectrum has the maximum magnitude of the 100 spectra, the second spectrum has the minimum magnitude and the third spectrum has the median magnitude. For each of the 200 subjects, the front view image in the gallery was used to generate a 4-dimensional linear subspace for illumination variation for each tissue type for subject i.

The results of the experiments will be presented using cumulative match scores.[6] For a probe image j, the image in the gallery which corresponds to the same subject is denoted by T_j. Given a probe image j under
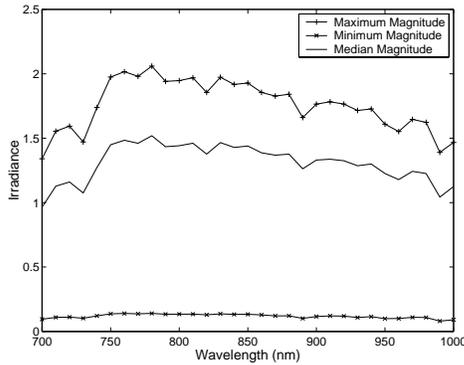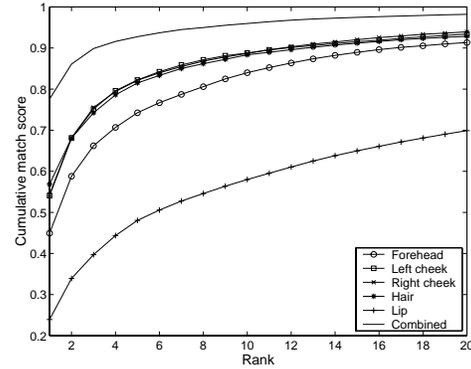
Figure 4. Examples of Boulder outdoor illumination spectra

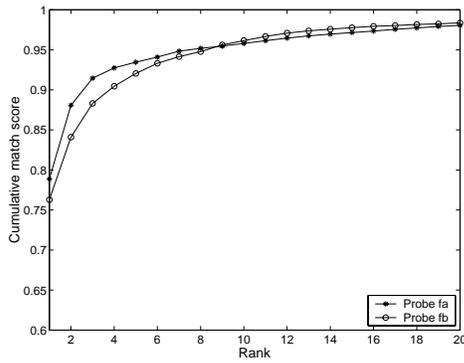**Figure 5.** Identification performance using fa and fb probes

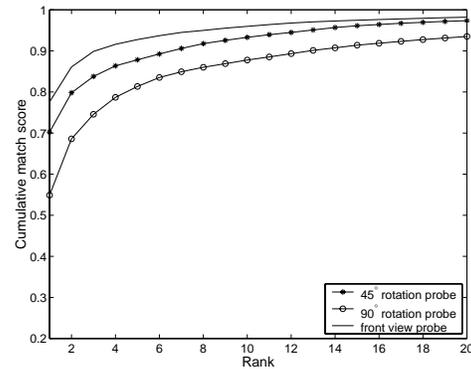**Figure 6.** Performance comparison of probe fa and fb

Figure 7. Identification performance of rotated face images

illumination k, we can compute $D(i, j, k)$ for each of the C images i in the gallery. Probe j is correctly recognized if $D(T_j, j, k)$ is the smallest of the C distances. Given a set of probes, the total number of correctly recognized probes is denoted as V_1. Similarly, V_n is the number of probes for which $D(T_j, j)$ is one of the n smallest of the C distances. The cumulative match score function for an experiment is defined by $R_n = V_n/(PM)$ where P is the total number of probes used in the experiment and M is the total number of illumination vectors used for each probe. We used the same set of 100 illumination spectra for probe images as those selected to construct gallery image linear subspaces.

We first consider the use of the frontal fa and fb probes to examine the utility of the various tissue types for hyperspectral face recognition. Figure 5 presents the cumulative match scores as a function of the rank n that are obtained when using $D_t(i, j, m)$ for each of the tissue types individually and $D(i, j, m)$ for the combination of all tissue types. We see that the cheeks and hair are the most useful tissue types for recognition while the forehead is somewhat less useful and the lip is the least successful. The recognition rate increases significantly when all tissue types are utilized. Over 90% of probes are correctly identified in the top 3 matches. The top curve in figure 6 compares recognition performance when using probes fa and fb separately with the algorithm that considers all tissue types. The fa images have the same facial expression as the gallery images while the fb images have different expressions. Accurate recognition is achieved in both cases which suggests that recognition using hyperspectral discriminants is not impacted substantially by changes in facial expression. Nevertheless, probes with different facial expressions are somewhat harder to identify at the top ranks.

Figure 7 examines the impact of changes in face orientation on recognition performance. Current 2D face recognition systems experience significant difficulty in recognizing probes that differ from a frontal gallery image by more than 32°, especially when combined with unknown illumination.[8] As expected, however, hyperspectral images can be used to achieve accurate recognition results for larger rotations. In figure 7 we see that for probes that are rotated 45° to the left or right from the frontal gallery image, over 93% of the probes have the correct match ranked in the top 10. For the difficult case of probes that are rotated 90°, about 88% of the probes have the correct match ranked in the top 10. These results utilize the distance function defined in terms of all visible tissue types.

6. CONCLUSION

We have tested the utilization of hyperspectral imaging for face recognition with changes in head pose and facial expression when the illumination is unknown. The experiments considered a database of near-infrared $(0.7\mu m$-$1.0\mu m)$ hyperspectral images for 200 subjects. A set of 100 global spectral irradiance functions measured at Boulder, Colorado were used to synthesize images of each subject. A four-dimensional linear model for each subject is used to test illumination invariant face recognition. The results show that the algorithm performs significantly better than current face recognition systems for identifying rotated faces under unknown illumination. The algorithm also provides accurate recognition performance in the presence of facial expression change.

ACKNOWLEDGMENTS

REFERENCES

1. K. Etemad and R. Chellappa, "Discriminant analysis for recognition of human face images," *J. Opt. Soc. Am. A* **14**, pp. 1,724–1,733, 1997.
2. B. Moghaddam and A. Pentland, "Probabalistic visual recognition for object recognition," *IEEE Trans. Pattern Anal. Machine Intell.* **19**, pp. 696–710, July 1997.
3. D. Swets and J. Weng, "Using discriminant eigenfeatures for image retrieval," *IEEE Trans. Pattern Anal. Machine Intell.* **18**, pp. 831–836, August 1996.
4. J. Wilder, "Face recognition using transform coding of grayscale projection and the neural tree network," in *Artifical neural networks with applications in speech and vision*, R. Mammone, ed., pp. 520–536, Chapman Hall, 1994.
5. L. Wiskott, J.-M. Fellous, N. Kruger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Trans. Pattern Anal. Machine Intell.* **19**, pp. 775–779, July 1997.
6. P. Phillips, H. Moon, A. Rizvi, and P. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Machine Intell.* **22**, pp. 1090–1104, October 2000.
7. P. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, and M. Bone, "Face recognition vendor test 2002: Overview and summary," tech. rep., Defense Advanced Research Projects Agency, March 2003.
8. R. Gross, J. Shi, and J. Cohn, "Quo vadis face recognition?," Tech. Rep. CMU-RI-TR-01-17, Robotics Institute, Carnegie-Mellon University, June 2001.
9. R. Gross, I. Matthews, and S. Baker, "Appearance based face recognition and light-fields," Tech. Rep. CMU-RI-TR-02-20, Robotics Institute, Carnegie-Mellon University, August 2002.
10. V. Blanz, S. Romdhani, and T. Vetter, "Face identification across different poses and illuminations with a 3D morphable model," in *Proc. IEEE Int. Conf. on Automat. Face and Gesture Recog.*, pp. 202–207, (Washington, DC), 2002.
11. A. Martínez, "Recognizing imprecisely localized, partially occluded and expression variant faces from a single sample per class," *IEEE Trans. Pattern Anal. Machine Intell.* **24**(6), pp. 748–763, 2002.

12. G. Healey and D. Slater, "Models and methods for automated material identification in hyperspectral imagery acquired under unknown illumination and atmospheric conditions," *IEEE Trans. Geosci. Remote Sensing* **37**, pp. 2706–2717, November 1999.

13. Z. Pan, M. Prasad, G. Healey, and B. Tromberg, "Recognizing faces in hyperspectral images," in *SPIE Proceedings Vol. 4725, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VIII*, pp. 168–176, April 2002.

14. D. Judd, D. MacAdam, and G. Wyszecki, "Spectral distribution of typical daylight as a function of correlated color temperature," *J. Opt. Soc. Am.* **54**, pp. 1031–1040, 1964.

15. S. Das and V. Sastri, "Spectral distribution and color of tropical daylight," *J. Opt. Soc. Am.* **55**, pp. 319–323, 1965.

16. J. Hernández-Andrés, J. Romero, and J. Nieves, "Color and spectral analysis of daylight in southern europe," *J. Opt. Soc. Am. A* **18**, pp. 1325–1335, June 2001.

17. G. Winch, M. Boshoff, C. Kok, and A. DuToit, "Spectroradiometric and colorimetric characteristics of daylight in Southern Hemisphere: Pretoria, South Africa," *J. Opt. Soc. Am.* **56**, pp. 456–464, 1966.

18. D. Slater and G. Healey, "Analyzing the spectral dimensionality of outdoor visible and near-infrared illumination functions," *J. Opt. Soc. Am.* **15**(11), pp. 2913–2920, 1998.

19. Z. Pan, G. Healey, and D. Slater, "Global spectral irradiance variability and material discrimination at Boulder, Colorado," *J. Opt. Soc. Am. A* **23**(3), pp. 513–521, 2003.

20. G. Golub and C. van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1983.

21. N. Gat, "Imaging spectroscopy using tunable filters: a review," in *SPIE Conference on Algorithms for Multispectral and Hyperspectral Imagery VI*, (Orlando), April 2000.

22. R. Duda, P. Hart, and D. Stork, *Pattern Classification*, Wiley-Interscience, New York, 2nd ed., 2001.