Experiments on Recognizing Faces in Hyperspectral Images

Zhihong Pan, Glenn Healey, Manish Prasad, Bruce Tromberg

Department of Electrical and Computer Engineering

Beckman Laser Institute

University of California

Irvine, CA 92697

Abstract

We examine the utility of using near-infrared hyperspectral images for the recognition of human subjects over a database of 137 subjects. Hyperspectral images were collected using a CCD camera equipped with a liquid crystal tunable filter and calibrated to spectral reflectance. The face recognition algorithm exploits spectral measurements for individual facial tissue types and combinations of facial tissue types. We demonstrate experimentally that hyperspectral images provide the opportunity to recognize faces independent of facial expression and face orientation.

1 Introduction

In remote sensing, researchers have shown that hyperspectral data can be used for material identification over a range of conditions using only the spectrum at a single pixel [4]. A similar approach has the potential to support human face recognition using local image regions over variation in face orientation, expression, and the illumination conditions [3]. Most current face recognition systems use primarily spatial information [1] [5] [10] [9] [8] and many have performed well on databases acquired under controlled lighting and geometry [7] [6]. However, these approaches exhibit performance degradation as the scene geometry and illumination change.

While the spectral characteristics of human tissue are significantly different from person to person, spectral variation for a single subject is often small over a range of geometries. Figure 1 presents an example of the spectral variability in human skin and hair using measurements taken from a hyperspectral image. In figure 1(a), near-infrared skin and hair reflectance functions are plotted for two subjects as acquired in a front-view hyperspectral image. In figure 1(b),
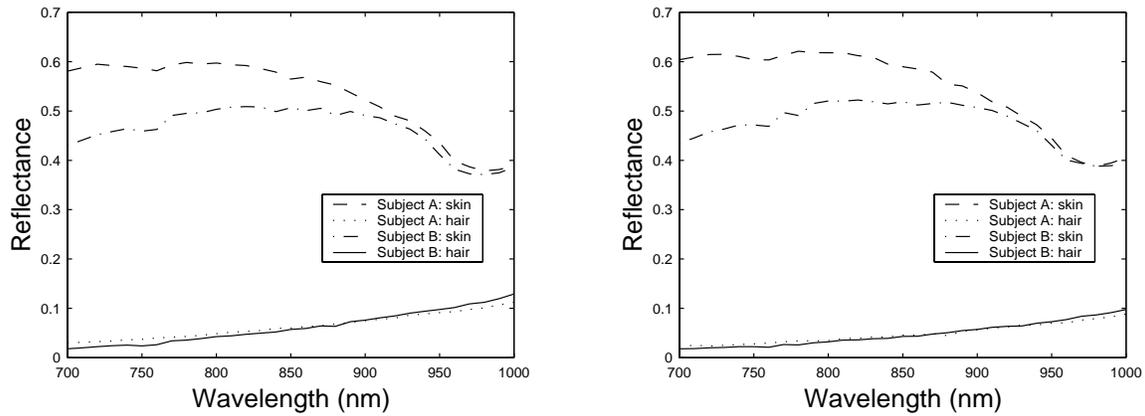
(a) Front view images (b) 90° side view images

Figure 1: Skin and hair reflectance spectra of two individuals

reflectance spectra for the same subjects are plotted as acquired in a side-view (profile) image. We see that there is significant spectral variability from one subject to the other. We also see that the spectral characteristics of the subjects remain stable over a large change in face orientation.

In this paper, we present experimental results on recognizing 137 human subjects using hyperspectral face images. For each subject, seven near-infrared images were acquired using a range of expression and geometric configurations. By utilizing spectral measurements over the near-infrared (NIR), we gain the ability to observe subsurface tissue structure that is difficult for a person to modify. Recognition is achieved by combining local spectral measurements for different tissue types. We show that subjects can be recognized over a broad range of conditions using only local spectral information.

2   Data Collection and Camera Calibration

Our data collection utilizes a hyperspectral camera from Opto-Knowledge Systems, Inc. (OKSI) that is based on a liquid crystal tunable filter [2] made by Cambridge Research Instruments. All images were captured with 31 bands over the near-infrared (0.7μm-1.0μm) with 468×494 spatial resolution using integration times between 30 and 60 ms for each band. The spectral channels have unknown gains due to filter transmission and CCD response and unknown offsets due to dark current and stray light. Therefore, we devised a method to convert the raw images acquired
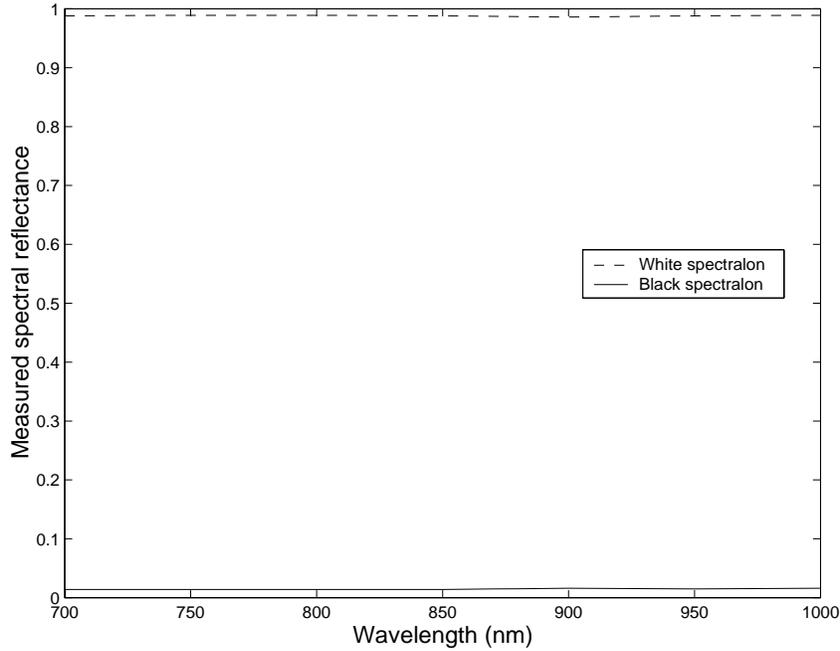
Figure 2: Spectral reflectance for spectralon panels

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 measured spectral reflectance functions for these panels are plotted in figure 2.

The raw measurement obtained by the hyperspectral imaging system at spatial coordinate (x, y) and wavelength λ_k is given by

$$I(x, y, \lambda_k) = L(x, y, \lambda_k)S(x, y, \lambda_k)R(x, y, \lambda_k) + O(x, y, \lambda_k) \tag{1}$$

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. For images of white and black spectralon, we have:

$$I_W(x, y, \lambda_k) = L(x, y, \lambda_k)S(x, y, \lambda_k)R_W(\lambda_k) + O(x, y, \lambda_k) \tag{2}$$

$$I_B(x, y, \lambda_k) = L(x, y, \lambda_k)S(x, y, \lambda_k)R_B(\lambda_k) + O(x, y, \lambda_k) \tag{3}$$

where $R_W(\lambda_k)$ and $R_B(\lambda_k)$ are the reflectance functions for the spectralon panels.

We average 10 images each of the white and black spectralon panels to obtain estimates of
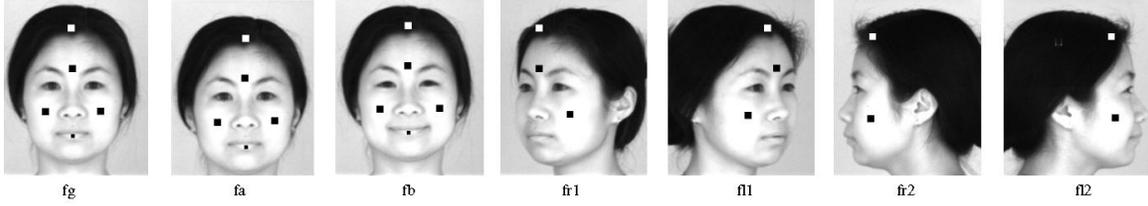
Figure 3: Examples of images with different expressions and rotations

$I_W(x, y, \lambda_k)$ and $I_B(x, y, \lambda_k)$. These estimates are used together with (2) and (3) to estimate $L(x, y, \lambda_k)S(x, y, \lambda_k)$ according to

$$L(x, y, \lambda_k)S(x, y, \lambda_k) = \frac{I_W(x, y, \lambda_k) - I_B(x, y, \lambda_k)}{R_W(\lambda_k) - R_B(\lambda_k)} \qquad (4)$$

and this estimate can be substituted into (3) to obtain an estimate for $O(x, y, \lambda_k)$. With these estimates, (1) can be solved for reflectance to give

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

We performed this calibration step at the beginning of each imaging session.

Images of 137 human subjects have been acquired in sets of up to seven images per subject. Figure 3 shows all 7 images of one subject. Two front-view images were taken with neutral expression (fg and fa). fg is used to represent the subject in the gallery. Another front-view image was taken with a different expression, which is referred to as the fb image. Four other images were taken with face orientations of -90°, -45°, 45°, and 90° as shown in figure 3. These images are referred to as fr2, fr1, fl1 and fl2 respectively.

3 Face Recognition

Our preliminary face recognition algorithm utilizes spectral reflectance functions recovered from small facial regions. Squares overlayed on the images in figure 3 indicate the size and location of the facial regions that are used for reflectance recovery for a subject. For the frontal images fg, fa, and fb, five facial regions corresponding to forehead, left cheek, right cheek, hair, and lip are used. For images acquired at other facial orientations, the subset of the same facial regions

that are visible are used as shown in figure 3. The forehead, for example, is not visible for a facial orientation of $90°$.

The spectral reflectance for each facial region is estimated by averaging over the N pixel squares shown in figure 3 according to

$$R_{ts}(\lambda_k) = \frac{1}{N} \sum_{x,y} R(x, y, \lambda_k) \tag{6}$$

where the x, y sum is over the N pixels in the square and ts is one of the following tissue types: f(forehead), lc(left cheek), rc(right cheek), h(hair), l(lip). The normalized spectral reflectance is defined by

$$\overline{R}_{ts}(\lambda_k) = \frac{R_{ts}(\lambda_k)}{\sqrt{\sum_k R_{ts}(\lambda_k)^2}} \tag{7}$$

and the distance between face i and face j for tissue type ts is defined by

$$D_{ts}(i, j) = \sqrt{\sum_k (\overline{R}_{ts}(i, \lambda_k) - \overline{R}_{ts}(j, \lambda_k))^2} \tag{8}$$

which is the Euclidean distance between normalized spectral reflectance functions.

Consider a recognition experiment with C subjects. For each subject, the front-view image fg is put in the gallery. Each image in the gallery is called a target face. The other images are used for recognition tests and are called probes. For each probe i, the target face which corresponds to the same human subject is denoted as t_i. For a probe i, we compute $D_{ts}(i, j)$ for all target faces j in the gallery. Probe i is correctly recognized using tissue type ts if $D_{ts}(i, t_i)$ is the smallest of the C distances. Given a set of probes, the total number of correctly recognized probes is denoted as M_1. Similarly, M_n is the number of probes for which $D_{ts}(i, t_i)$ is one of the n smallest of the C distances. The recognition rate $R_n = M_n/P$ will be calculated for each experiment described in section 4 where P is the total number of probes in the experiment.

For a given probe, recognition performance can be enhanced by utilizing all visible tissue types. Thus, the distance between probe i and target j is defined as the weighted sum over all tissue types that are visible in the probe

$$D(i, j) = \omega_f D_f(i, j) + \omega_{lc} D_{lc}(i, j) + \omega_{rc} D_{rc}(i, j) + \omega_h D_h(i, j) + \omega_l D_l(i, j) \tag{9}$$

The weights ω_{ts} are derived from the statistics of the tissue type reflectance spectra.

category	fg	fa	fb	fr1	fl1	fr2	fl2
# of images	137	125	130	131	130	124	129

Table 1: Number of images in each category
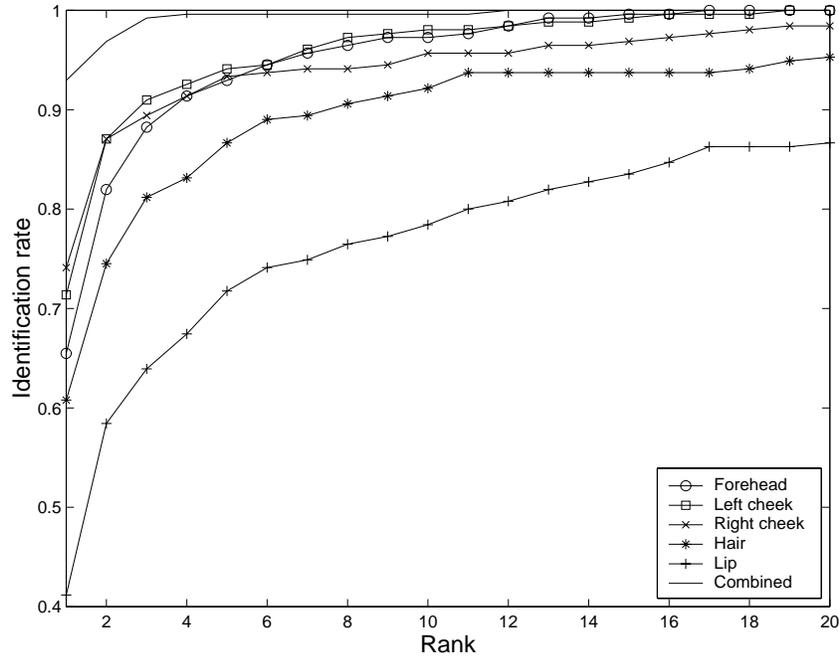
Figure 4: Identification performance using fa and fb probes

4 Experimental Results

We conducted a series of recognition experiments using a database of images obtained for 137 subjects. The total number of images in each category is shown in Table 1. For every probe image, there is an image of the same subject in the gallery. In Figure 4, the recognition performance using fa and fb probes is presented. The recognition rates using $D_{ts}(i, j)$ for individual tissue types as well as using $D(i, j)$ for all tissue types are shown. We see that skin is the most useful tissue type for recognition while hair and lips are less useful. The top curve in Figure 4 shows that the best performance is achieved by utilizing all of the tissue types.

Figure 5 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
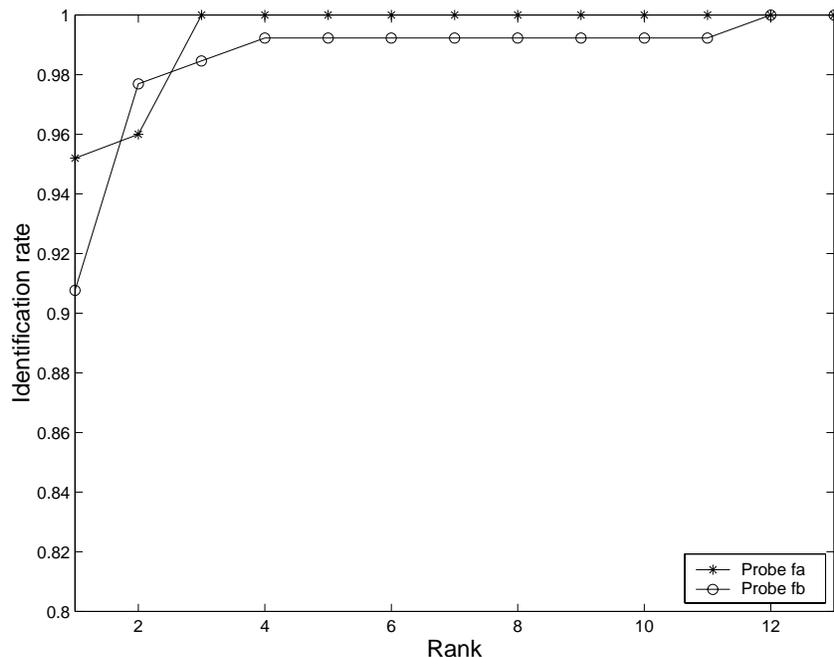
Figure 5: Performance comparison of probe fa and fb

gallery images while the fb images have another expression. We see that performance is similar in both cases which suggests that recognition using hyperspectral discriminants is not impacted significantly by changes in facial expression.

Figure 6 examines the impact of changes in face orientation on recognition performance. Large changes in face orientation are a significant source of difficulty for many face recognition systems. We see that for probes that are rotated 45° to the left or right from the gallery, 80% of the probes are recognized correctly and 95% of the probes have the correct match ranked in the top 5. For the difficult case of probes that are rotated 90°, about 80% of the probes have the correct match ranked in the top 10.

5   Summary

We have collected a database of hyperspectral face images over the near-infrared for 137 subjects and conducted a series of face recognition experiments. The images have been carefully calibrated to spectral reflectance. Face recognition is based on matching local spectral measurements for different tissue types. We showed that face recognition using hyperspectral images
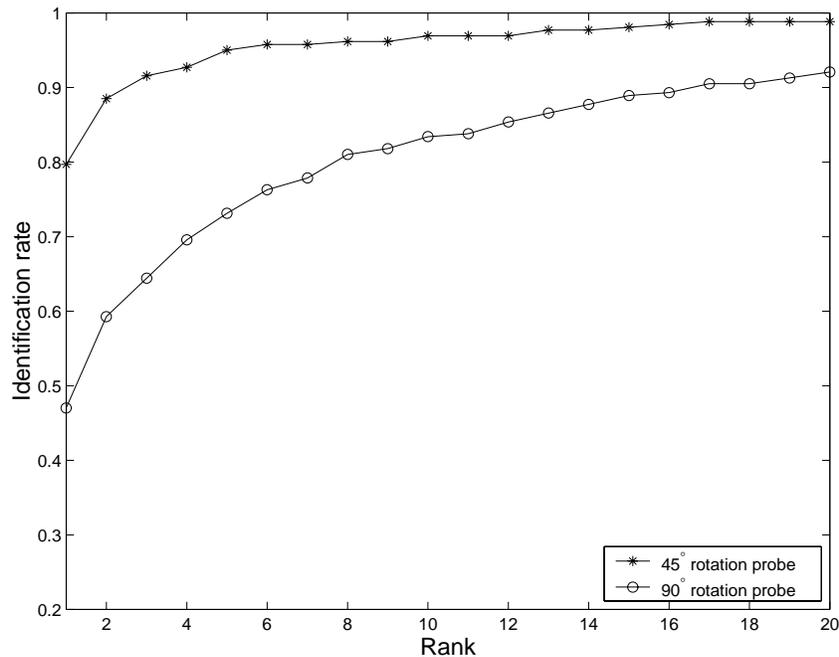
Figure 6: Identification performance of rotated face images

performs well even in the presence of changes in facial expression and face orientation. We also quantified the value of different tissue types for face recognition in hyperspectral images.

References

[1] K. Etemad and R. Chellappa. Discriminant analysis for recognition of human face images. *J. Opt. Soc. Am. A*, 14:1,724–1,733, 1997.

[2] N. Gat. Imaging spectroscopy using tunable filters: a review. In *SPIE Conference on Algorithms for Multispectral and Hyperspectral Imagery VI*, Orlando, April 2000.

[3] G. Healey, Z. Pan, and B. Tromberg. Models for recognizing faces in hyperspectral images. In *SPIE Proceedings Vol. 4381, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VII*, pages 317–326, April 2001.

[4] 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(6):2706–2717, November 1999.

[5] B. Moghaddam and A. Pentland. Probabalistic visual recognition for object recognition. *IEEE Trans. Pattern Anal. Machine Intell.*, 19(7):696–710, July 1997.

[6] P.J. Phillips, H. Moon, A. Rizvi, and P. Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE Trans. Pattern Anal. Machine Intell.*, 22(10):1090–1104, October 2000.

[7] P.J. Phillips, H. Wechsler, J. Huang, and P. Rauss. The FERET database and evaluation procedure for face recognition algorithms. *Image and Vision Computing*, 16(5):295–306, 1998.

[8] D. Swets and J. Weng. Using discriminant eigenfeatures for image retrieval. *IEEE Trans. Pattern Anal. Machine Intell.*, 18:831–836, August 1996.

[9] J. Wilder. Face recognition using transform coding of grayscale projection and the neural tree network. In R.J. Mammone, editor, *Artifical neural networks with applications in speech and vision*, pages 520–536. Chapman Hall, 1994.

[10] 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(7):775–779, July 1997.