

Face Recognition Under Variable Lighting using Harmonic Image Exemplars

Lei Zhang Dimitris Samaras
Department of Computer Science
SUNY at Stony Brook, NY, 11790
{lzhang, samaras}@cs.sunysb.edu

Abstract

We propose a new approach for face recognition under arbitrary illumination conditions, which requires only one training image per subject (if there is no pose variation) and no 3D shape information. Our method is based on the recent result [1] which demonstrated that the set of images of a convex Lambertian object obtained under a wide variety of lighting conditions can be approximated accurately by a low-dimensional linear subspace. In this paper, we show that we can recover basis images spanning this space from just one image taken under arbitrary illumination conditions. First, using a bootstrap set consisting of 3D face models, we compute a statistical model for each basis image. During training, given a novel face image under arbitrary illumination, we recover a set of images for this face. We prove that these images are the set of basis images with maximum probability. During testing, we recognize the face for which there exists a weighted combination of basis images that is the closest to the test face image. We provide a series of experiments that achieve high recognition rates, under a wide range of illumination conditions, including multiple sources of illumination. Our method achieves comparable levels of accuracy with methods that have much more onerous training data requirements.

1. Introduction

Face recognition has recently received extensive attention as one of the most significant applications of image understanding. Although rapid progress has been made in this area during the last few years, the general task of recognition remains unsolved. In general, face appearance does not depend solely on identity. It is also influenced by illumination and viewpoint. Changes in pose and illumination will cause large changes in the appearance of a face. In this paper we demonstrate a new method to recognize face images under a wide range of illumination conditions using spherical harmonic images of the face. When the viewpoint is

fixed, our method requires only a single training image per subject.

In the past few years, there have been attempts to address image variation produced by changes in illumination [6][7][8]. Sim et al. [9] propose a model and exemplar-based approach for recognition. Georgiades et al. [3] present a new method using the illumination cone. The above two methods need to reconstruct 3D face information for each subject in the training set. Romdhani et al. [10] recover the shape and texture parameters of a 3D Morphable Model in an analysis-by-synthesis fashion. Both this method and ours compute a statistical model during a bootstrap phase which encapsulate texture and shape information. The difference between the two methods is that, during training, in [10] the 3D Morphable Model is fitted to a subject's face image explicitly to recover texture and shape parameters, whereas our method recovers a set of spherical harmonic images for the training subject instead. In general, appearance-based methods like Eigenfaces [14] and SLAM [15] need a number of training images for each subject, in order to cope with illumination variability. Previous research suggests that illumination variability in face images is low-dimensional e.g. [1][2][11][12][13][20]. Using spherical harmonics and signal-processing techniques, Basri et al. [1] and Ramamoorthi [2] have shown that the set of images of a convex Lambertian object obtained under a wide variety of lighting conditions can be approximated accurately by a 9 dimensional linear subspace. Furthermore, a simple scheme for face recognition with excellent results was described in [1]. However, to use this recognition scheme, the basis images spanning the illumination space for each face are required. These images can be rendered from a 3D scan of the face or can be estimated by applying PCA to a number of images of the same subject under different illuminations [2]. An effective approximation of this basis by 9 single light source images of a face was reported in [4]. The above mentioned methods need a number of images and/or 3D scans of the subjects in the database, thus requiring specialized equipment and procedures for the capture of the training set, thus limiting their

applicability. A promising earlier attempt by [5] used symmetric shape from shading but suffered from the drawbacks of SFS.

In this paper we propose a method that recovers the 9 basis images of the illumination space from just one image taken under arbitrary illumination conditions when the pose is fixed. Since the first nine spherical harmonic images capture more than 99 percent of the Lambertian reflectance energy [1], the weighted combination of the 9 basis images will be very close to the real image. Our method consists of three steps: statistical model computation, training and testing. Initially, we use a *bootstrap* set consisting of 3D face models. We compute a statistical model for each of the spherical harmonics. Our statistical models aim to learn the probability density function (pdf) for the basis images. The heaviest data capture requirements for our method are placed in the construction of the bootstrap set. This set and the resulting pdf's for the basis images need be constructed only once. No further modifications are necessary when new subjects are added to the database, allowing for increased scalability of the recognition system.

During the second step, *training*, when presented with an image of a novel face, we use the statistical model learned during the bootstrap phase to recover the corresponding basis images. Given a face, we first estimate the weight coefficients for the basis images. Then we recover the basis images by computing the *maximum a posteriori* (MAP) estimate. We show that the set of images we recover is the set of basis images with maximum probability. For each subject in the training set, we recover the basis images that span the illumination space for this subject. The faces in the training set need not belong to the bootstrap set. The only requirement is that the bootstrap data capture the statistical characteristics of the training set. Finally, during the *testing* step, we use the recognition scheme proposed by Basri et al [1]. We return the face from the training set for which there exists a weighted combination of basis images that is the closest to the test face image.

We use a collection of 3D face scans supplied by USF [18] as the bootstrap set and the Yale Face Database B [19] for training and testing. We provide a series of experiments that show that our method has high recognition accuracy even though none of the subjects in the training set belongs to the bootstrap set. Our method achieves comparable levels of accuracy with methods that have much more onerous training data requirements. In our experiments, as predicted by the theory, the basis images recovered during training are noticeably insensitive to the particular illumination of the training image, which indicates that our method should perform well on larger databases. Finally, we provide experiments that show that our method remains accurate across different poses. This suggests that our method can be easily incorporated to generalized face recognition methods that

address multiple recognition issues.

This paper is organized as follows. In the next section, we will show how to acquire basis images from 3D models and how to estimate the statistical models for the basis images. In Sec. 3, we explain how to recover basis images from a novel face image. In Sec. 4, we describe the recognition process that uses recovered basis images. In Sec. 5, we describe our experiments and their results. The final Section presents the conclusions and future work directions.

2. Statistical Models of Basis Images

2.1. Spherical Harmonic Images

In this section, we will show how we render basis images from 3D models using the results of [1]. Let p_i denote the i th object point. Let λ denote the vector of the object's albedos, that is, λ_i is the albedo of p_i . Similarly, let n_x, n_y, n_z denote three vectors of the same length that contain the x, y and z components of the surface normals. Further, let n_{x^2} denote a vector such that $n_{x^2,i} = n_{x,i}n_{x,i}$. We define $n_{y^2}, n_{z^2}, n_{xz}, n_{yz}, n_{xy}$ similarly. We use $\lambda \cdot * \mathbf{v}$ to denote the component-wise product of λ with any vector \mathbf{v} . Using this notation, the first nine harmonic images of the objects in this specific pose are:

$$\begin{aligned} b_{00} &= \frac{1}{\sqrt{4\pi}}\lambda; & b_{11}^e &= \sqrt{\frac{3}{4\pi}}\lambda \cdot * n_x \\ b_{11}^o &= \sqrt{\frac{3}{4\pi}}\lambda \cdot * n_y; & b_{10} &= \sqrt{\frac{3}{4\pi}}\lambda \cdot * n_z \\ b_{20} &= \frac{1}{2}\sqrt{\frac{5}{4\pi}}\lambda \cdot * (2n_{z^2} - n_{x^2} - n_{y^2}) \\ b_{21}^e &= 3\sqrt{\frac{5}{12\pi}}\lambda \cdot * n_{xz}; & b_{21}^o &= 3\sqrt{\frac{5}{12\pi}}\lambda \cdot * n_{yz} \\ b_{22}^e &= \frac{3}{2}\sqrt{\frac{5}{12\pi}}\lambda \cdot * (n_{x^2} - n_{y^2}); & b_{22}^o &= 3\sqrt{\frac{5}{12\pi}}\lambda \cdot * n_{xy} \end{aligned} \quad (1)$$

The first row of Fig.1 shows an example of the 9 basis images of a face, rendered from a 3D face scan. Lighter gray indicates positive values and darker gray indicates negative values. More informative, high-resolution versions of the basis images can be found at: www.cs.sunysb.edu/~lzhang/hrbasis.jpg.

2.2. Learning the statistical models

The key equation of our method is:

$$i(x) = b(x)^T \alpha + e(x, \alpha) \quad (2)$$

which states that at pixel position x , the pixel intensity $i(x)$, is the weighted combination of the basis images plus an error term. More precisely, let I be a d -dimensional image

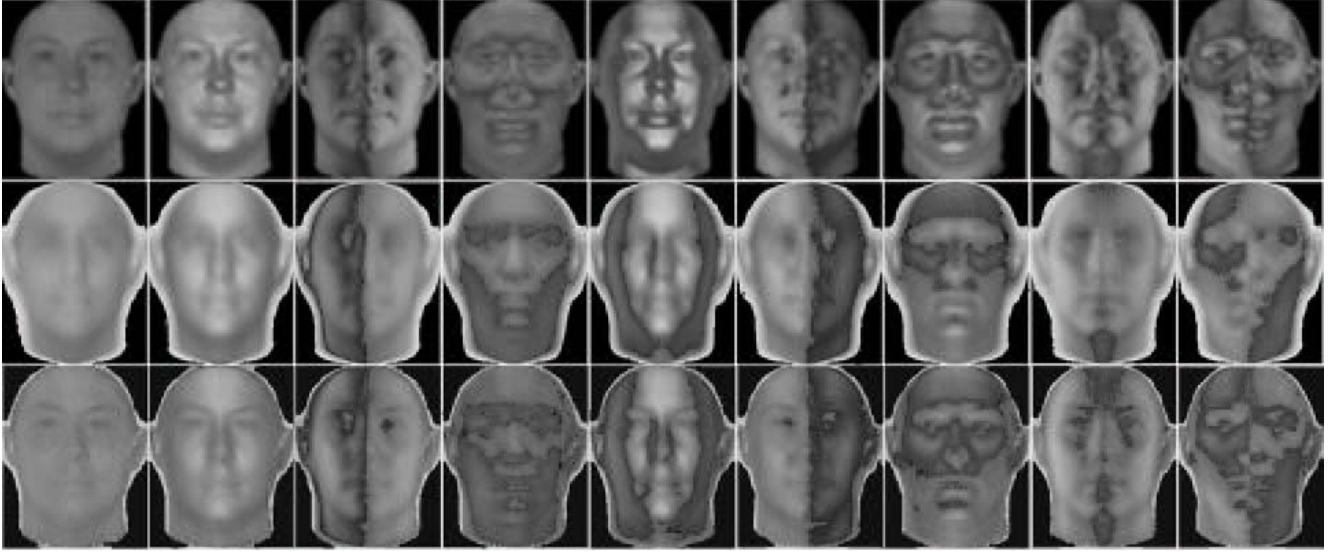


Figure 1. The first row is the first nine spherical harmonic images rendered from this 3D face scan. The second row is the statistical models computed from the rendered basis images and the third row is the recovered basis images. Lighter gray indicates positive values and darker gray indicates negative values (for positive values, the range of pixel intensities is from 140 to 240 and for negative values, the range is from 40 to 140).

and let B be the $9 * d$ matrix of basis images, the columns of which are the vectors $\{b(x)\}_{x=1}^d$. Also, let α be a 9-dimensional vector denoting the coefficients of $b(x)$ and let E be a d -dimensional vector denoting the error term. Thus we get: $I = B^T \alpha + E$. Similar to [9], our statistical models aim to learn the probability density function (pdf) for B and E . For B , we assume that the pdf's are Gaussian distributions of unknown means and covariances which we can estimate from the basis images rendered by 3D scans. In our experiments we had 50 3D face scans with texture information in our bootstrap set. We render 9 basis images per face model using Equation set (1). From these basis images, we compute the sample mean vector $\mu_b(x)$ and sample covariance matrix $C_b(x)$ for the basis image $b(x)$ of each face. The second row of Fig. 1 displays the statistical models of the basis images computed from our bootstrap set.

To estimate the error term, we need real images with known harmonic coefficients. Unfortunately, in our current implementation we do not have both 3D scans and real images for the subjects in our data set. Instead, we synthesized images and estimated their coefficients. Since we have 3D scans with texture information, we illuminated synthetically each 3D model, obtaining face images under a wide variety of illumination conditions. Then we applied a least-squares method to estimate the coefficients α . Hence, we can compute the sample mean $\mu_e(x, \alpha)$ and sample variance $\sigma_e^2(x, \alpha)$ of the error term by $E = I - B^T \alpha$. This error term models the deviations from Lambertian reflectance and the errors of the low-dimensional approximation which cause the estimates of the coefficients to diverge from the

true values.

3. Recovering the Basis Images

Given a training image of a novel face (possibly a different subject from those in the bootstrap) under arbitrary illumination, first we need to estimate the coefficients α in Eq. (2) and to update the error term according to the coefficients. Then we recover $b(x)$, the corresponding basis images, at each x by computing the *maximum a posteriori* (MAP) estimate,

$$b_{MAP}(x) = \operatorname{argmax}_{b(x)} (P(b(x)|i(x)))$$

This is explained in detail in the following subsections.

3.1. Estimating α and the error term

To estimate the unknown spherical harmonic coefficients α for the given novel face image, we use kernel regression. Notice that since in the bootstrap set we have images with known coefficients, we can regard coefficient estimation as a continuous-valued classification problem. The approach is similar to the recovery of unknown illumination which is a well-studied problem [17][9]. We first store all the K bootstrap images, $\{J_k\}_{k=1}^K$ along with their spherical harmonic coefficients $\{\alpha_k\}_{k=1}^K$. Similar to [9], given a new image, i_{tra} , we can estimate the coefficients α_{tra} as follows:

$$\alpha_{tra} = \frac{\sum_{k=1}^K w_k \alpha_k}{\left(\sum_{k=1}^K w_k\right)} \quad (3)$$

where $w_k = \exp[-\frac{1}{2}(D(i, J_k)/\sigma_k)^2]$ and $D(i, J_k) = \|i - a_k\|_2$ the L_2 norm, σ_k is the width of the k -th Gaussian kernel which controls the influence of J_k on the estimation of coefficient α_{tra} . We can pre-compute all $\{\sigma_k\}_{k=1}^K$ in a way such that ten percent of the bootstrap images are within $1 \times \sigma_k$ at each σ_k .

We treat the statistics of the error term E similarly. In Section 2.2, we have shown how to get the statistical model for the error term. The mean and the variance of the error estimation $e_{tra}(x, \alpha_{tra})$ at α_{tra} will be interpolated from the known mean and variance of the error term at coefficients $\{\alpha_k\}_{k=1}^K$ which have been calculated in Section 2.2. The equation used here is similar to that in Eq. (3).

To check the accuracy of our estimation of the coefficients, we synthesized images using the basis images with various sets of coefficients $\{\alpha_j\}_{j=1}^J$. From the synthesized images, we estimated coefficients $\{\alpha'_j\}_{j=1}^J$ using Eq. (3). Ideally, the two sets of α_j and α'_j should be identical. Experimental results show that the estimated values are very close to the actual values.

3.2. Computing the basis images

After estimating the coefficients α , we can recover the basis images for a novel face image. As we have shown already: $b_{MAP}(x) = \operatorname{argmax}_{b(x)}(P(b(x)|i(x)))$. It is hard to calculate $\operatorname{argmax}_{b(x)}P(b(x)|i(x))$ directly. Similar to [16], using Bayes' rule, $P(b(x)|i(x))P(i(x)) = P(i(x)|b(x))P(b(x))$. Since we have already estimated the spherical harmonic coefficients α , according to Eq. (2), we know that given an image $i(x)$ and known coefficients α , $P(i(x))$ is a constant. Thus,

$$b_{MAP}(x) = \operatorname{argmax}_{b(x)}(P(i(x)|b(x))P(b(x))) \quad (4)$$

In this term, $P(b(x))$ is the Gaussian probability density function we have learned previously and $P(i(x)|b(x))$ can be computed from Eq. (2). Thus, $i(x)$ is a random variable with Gaussian pdf of mean $b(x)^T \alpha + \mu_e(x, \alpha)$ and variance $\sigma_e^2(x, \alpha)$. Thus,

$$\operatorname{argmax}_{b(x)}(Gauss(b(x)^T \alpha + \mu_e, \sigma_e^2) \times Gauss(\mu_{b(x)}, C_{b(x)})) \quad (5)$$

Using log probability, and ignoring constant terms, Eq. (5) becomes:

$$\operatorname{argmax}_{b(x)}(-\frac{1}{2}(\frac{i - b^T \alpha - \mu_e}{\sigma_e})^2 - \frac{1}{2}(b - \mu_b)^T C_b^{-1}(b - \mu_b)) \quad (6)$$

Then we set the derivatives of Eq. (6) (w.r.t $b(x)$) to 0 in order to get the maximum probability, we get:

$$-\frac{2}{\sigma_e^2}(i - b^T \alpha - \mu_e)\alpha + 2C_b^{-1}(b - \mu_b) = 0 \quad (7)$$

Rearranging, we can get the following linear equation: $A * b_{MAP} = T$ where $A = \frac{1}{\sigma_e^2}\alpha\alpha^T + C_b^{-1}$ and $T = \frac{(i - \mu_e)}{\sigma_e^2}\alpha + C_b^{-1}\mu_b$.

The third row of Fig. 1 shows the recovered basis images compared to the bootstrap basis images (the first row).

4. Face Recognition

We perform recognition based on [1]. During testing, we recognize the face for which there exists a weighted combination of basis images that is the closest to the test face image. Here B is the set of basis images with size $r * d$, d is the number of points in the image and r the number of basis images used (9 is a natural choice as discussed in [1]). Every row of B contains one spherical harmonic image, and the rows of B form a basis for the linear subspace. We can simply apply QR decomposition to B to obtain an orthonormal basis. Thus, we compute the distance from the test image, I_{test} , and the space spanned by B as $\|QQ^T I - I\|$.

Recalling the error term that appears in Eq. (2), we alternatively perform recognition by minimizing $\|B^T \alpha + E - I_{test}\|$. Thus, for each set of basis images in the training set, we estimate the coefficients by solving the least square equation: $B^T \alpha = I_{test}$. Then, we update the error term E as in Sec. 2.2. After calculating both α and E for each set of basis images, we recognize the i -th face for which $B_i^T \alpha_i + E_i - I_{test}$ is the minimum.

Comparing the results of the two above schemes, we found that recognition rates were very close while the second scheme was much slower than the first scheme.

5. Experiments and Results

For training and testing, we used the Yale Face Database B [19] which contains faces that do not belong to our bootstrap set. Despite its relatively small size, this database provides images that sample sufficiently the whole illumination space and has therefore become a testing standard for variable illumination recognition methods. The Yale Database contains images of 10 people at 9 poses and 64 illuminations per pose. We used 45*10 face images for 10 subjects in a single pose with each subject having 45 face images taken under different directional light sources. Fig. 2 shows some of the images we used in our experiments.

Subsets	1	2	3	4
Number of Images	70	120	120	140
Illumination	0-12	13-25	26-50	51-77

Table 1: The separation of the Yale Database

Following [4], we group our data set in four subsets (Table 1). Each subset contains images illuminated from a specific range of directions. Subset 4 contains the extreme illu-



Figure 2. Some real images used in our experiments.

mination conditions. Please refer to [4] for more information about this grouping.

5.1. Experiments of Basis Image Recovery

Since our recognition method is based on the basis images recovered during training, it is very important that we have an accurate method to recover the basis images. Recovered basis images and bootstrap basis image for bootstrap subjects are compared in Fig. 1. Fig. 3 shows the basis images recovered from various face images under different illuminations for one subject belonging to the training set (but not the bootstrap).

We found that, except for the basis images recovered from images under extreme illumination (i.e. images in subset 4), the resulting basis images recovered from images under different illumination are very close. For each subject in the training set, we calculated the basis images using 20 images under different illumination (5 of them belonging to subset 4). The per person mean variance of the 20 resulting sets of basis images, was 14.66 intensity levels per pixel for the full set of images and 5.34 for the images belonging to Subsets 1-3. For comparison purposes, per person variance of the original training set images was 31.13 (24.77 for subsets 1-3) while the cross-person mean variance of the 200 total image basis sets for all subjects was 31.47 for the full set (16.33 for subsets 1-3). We see that the basis images we recover exhibit much greater invariance to illumination effects.

5.2. Recognition

We report two sets of experimental results in Table 2. For the first set, we use the basis images recovered from randomly selected training images, excluding extreme illumination conditions. For the second set, we perform recognition using the basis images recovered from the training im-

ages captured only under extreme illumination. We tested 400 face images and the error rates are shown in Table 2.

	Subset1	Subset2	Subset3	Subset4
Non-Ext	0.0	0.0	0.3	3.1
Extreme	8.0	7.5	7.5	6.8

Table 2: the first row of "Non-Ext" are the error rates achieved by using the set of basis images recovered from randomly selected training images (8 per person) from Subsets 1,2 and 3, and the second row ("Extreme") are the error rates by using basis images recovered from randomly selected training images (6 per person) from Subset 4. The columns denote the testing sets.

In section 4, we proposed two recognition methods. Table 3 shows the different results of the two methods.

	Subset1	Subset2	Subset3	Subset4
Method1	0.0	0.0	0.3	3.5
Method2	0.0	0.0	0.3	3.1

Table 3: Since we only perform QR decomposition in method 1, recognition using method 1 is much faster than that using method 2, while error rates are very similar.

Clearly, the error rates using the two approaches are very close while the first method is much faster than the second method.

The comparison between our method and other methods [4] is shown in Table 4. In order to compare with other methods, we randomly select face images to recover the corresponding spherical harmonic images and perform recognition on 400 images. We repeated the process 10 times and calculated the average recognition rates.

All the methods in Table 4 which have an off-line training process require all the images of Subset1 and 2 for training while we require only one image per subject for training. The method of Nine Points of Light (9PL) [4] does not have off-line training, but still needs nine images per subject for training. With our method, training on one image randomly selected from subsets 1 and 2 and testing on images from Subsets 1-3, we obtain perfect results and even when we test on images from Subset 4, we still achieve high accuracy. As can be seen from Table 4, the results from our method are comparable with methods that require extensive training data per subject even though our method requires only one training image per pose.

Methods	Subset1,2	Subset3	Subset4
Correlation	0.0	23.3	73.6
Eigenfaces	0.0	25.8	75.7
Linear Subspace	0.0	0.0	15.0
Cones-attached	0.0	0.0	8.6
9PL	0.0	0.0	2.8
Cones-cast	0.0	0.0	0.0
Our method	0.0	0.3	3.1

Table 4: Recognition results using various previous methods and our method. Except for our method, the data were taken from [4]



Figure 3. The basis images recovered from images of the same subject under various single directional illuminants. The first column is the images we used for the recovery followed by the set of basis images.

5.3. Multiple Directional Illumination

It is interesting to study the performance of our method on images taken under multiple directional illumination sources. For comparison between the performance on images taken under a single directional illumination source and that on images taken under multiple directional illumination sources, we synthesized images by combining face images in our data set. We randomly selected 2-4 images from the training data set and combined them together with random weights to simulate face images under multiple directional illumination sources (12 synthetic images per person). Fig. 4 shows the recovered results of basis images from the synthesized images compared to the recovered results from a real image under a single illuminant.

We performed recognition experiments based on the basis images we recovered from training images under multiple illuminants. Error rates were very close to the single illuminant training set, e.g. 0.3 for test images under multiple illumination, 3.3 for test images from subset 4.

5.4. Pose Variance

So far, we have provided experiments on the images with frontal view images. We applied our method to two sets of images with poses -12 degrees and 24 degrees azimuth from the frontal view. We rotated the 3D scans in the bootstrap set by -12 degrees and 24 degrees respectively and rendered basis images in the same way described previously. For training and testing, from the Yale Face Database B we selected two sets of face images which are taken approximately -12 degrees and 24 degrees (8 images per person per angle) from the frontal view respectively and repeated the same

steps described previously in Sec. 3 and 4. Experimental results are similar to the experimental results of our method for frontal views. Our experiments imply a solution for face recognition in the presence of both illumination and pose variance. The basic idea would be to combine face pose estimation with our method. For example, [21] provides a learning-based pose estimation method using Support Vector Regression. Efficient ways to integrate pose estimation and our method will be explored as a future research direction.

6. Conclusions and Future Work

We have shown that by using statistical models, we can recover illumination space basis images from a single image taken under arbitrary illumination conditions. Experimental results indicate that our method's recognition rates are comparable to other methods for images under single illumination; however our method performs as well with multiple illuminants. During the training phase, we need only one image per subject taken under general illumination to recover the basis images. Thus the training set can be expanded easily with new subjects. Such scalability is desirable in Face Recognition Systems. In our experiments, as predicted by the theory, the basis images recovered during training are noticeably insensitive to the particular illumination of the training image, which indicates that our method should perform well on much larger databases, than the ones available to us now.

At this time, there exist relatively few publicly available sets of images of faces with multiple images taken under variable illumination, so we plan to continue validation of our method with a more complete database, both



Figure 4. The basis images recovered from an image taken under multiple illuminants (bottom) compared to those recovered from an image under single illumination (top).

in the number of subjects and in the type of light sources used, e.g. area sources. As the number of subjects increases, more sophisticated training methods should be explored (e.g. SVMs). Finally we plan to integrate our method with pose estimation methods.

Acknowledgments

We would like to thank Vivek Srivastava and Yang Wang for their help, and Sudeep Sarkar for the USF 3D face database.

References

- [1] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *Proc. ICCV*, (II:384-390), 2001.
- [2] R. Ramamoorthi, "Analytic PCA Construction for Theoretical Analysis of Lighting Variability in Images of a Lambertian Object" *IEEE Trans. on PAMI*, vol.24, no.10, Oct. 2002.
- [3] A.S. Georghiades, P.N. Belhumeur and D.J. Kriegman, "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose," *IEEE Trans. on PAMI*, vol.23, no.6, pp.643-660, June 2001.
- [4] Kuang-Chih Lee, Jeffrey Ho and D. Kriegman, "Nine Points of Light: Acquiring Subspaces for Face Recognition under Variable Lighting," *Proc. CVPR*, (I:519-526), 2001.
- [5] W.Y. Zhao, R. Chellappa, R. Rama, "Illumination-Insensitive Face Recognition using Symmetric Shape-from-Shading," *Proc. CVPR* (I: 286-293), 2000.
- [6] W. Zhao, R. Chellappa and A. Rosenfeld, "Face Recognition: A Literature Survey," *UMD CfAR Technical Report CAR-TR-948*, 2000.
- [7] A. Pentland, "Looking at people: Sensing for Ubiquitous and Wearable Computing," *IEEE Trans. on PAMI*, vol.22, no.1, pp.107-119, Jan. 2000.
- [8] R. Chellappa, C. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Proc. IEEE*, vol.83, no.5, pp.705-740, 1995.
- [9] Terence Sim, Takeo Kanade, "Combining Models and Exemplars for Face Recognition: An Illuminating Example," *Proc. CVPR Workshop on Models versus Exemplars in Comp. Vision*, December, 2001.
- [10] S. Romdhani, Volker Blanz and T. Vetter, "Face Identification by Fitting a 3D Morphable Model using Linear Shape and Texture Error Functions" *Proc. ECCV* (IV:3 ff.), 2002.
- [11] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions," *Proc. CVPR*, (995-999), 1994.
- [12] R. Epstein, P. Hallinan, and A. Yuille, "5+/-2 eigenimages suffice: An empirical investigation of low-dimensional lighting models," *PBMCV*, (SESSION 4), 1995.
- [13] P. Belhumeur and D. Kriegman, "What is the set of images of an object under all possible lighting conditions," *IJCV*, vol.28(3), pp.245-260, 1998.
- [14] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cognitive Neuroscience*, vol.3, no.1, pp.71-96, 1991.
- [15] H. Murase and S. Nayar, "Visual Learning and Recognition of 3D Objects from Appearance," *IJCV*, vol.14, pp.5-24, 1995.
- [16] T. Sim and T. Kanade, "Illuminating the Face," *Technical Report CMU-RI-TR-01-31*, Robotics Institute, CMU, 2001.
- [17] R. Zhang, P.S. Tsai, J.E. Cryer, and M. Shah, "Shape from Shading: A Survey," *IEEE Trans. on PAMI*, vol.21(8), Aug. 1999.
- [18] Volker Blanz and T. Vetter "A Morphable Model for the synthesis of 3D Faces," *Proc. SIGGRAPH*, (187-194), 1999.
- [19] P.N. Belhumeur and J.P. Hespanha and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection" *IEEE Trans. on PAMI*, vol.19, no.7, pp.711-720, Mon., 1997.
- [20] Y. Adini, Y. Moses and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," *IEEE Trans. on PAMI*, vol. 19, pp.721-732, July 1997.
- [21] Yongmin Li, Shaogang Gong and H. Liddell, "Support Vector Regression and Classification Based Multi-view Face Detection and Recognition" *Proc. AFGC*, (300-305), 2000.
- [22] H. Bischof, H. Wildenauer, A. Leonardis, "Illumination Insensitive Eigenspaces," *Proc. ICCV* (I: 233-238), 2001.