Face Color Constancy: Towards Illumination Invariant Face Recognition

Rajkiran Gottumukkal and Vijayan K. Asari; Department of Electrical and Computer Engineering; Old Dominion University; Norflok, VA-23529, USA

Abstract

We propose a novel technique to compensate for variations in face images due to changes in the illumination conditions. Other than color of the faces, shadows and specular reflections on the faces also change with respect to lighting environment. These variations result in false recognitions even in the best face recognition algorithms. We tackle this problem by processing the faces using a color constancy model specific to face images, which we named face color constancy.

The face color constancy algorithm is trained with face images in different lighting environments. The color changes observed are used to learn the color mapping from one lighting environment to the other. These mappings are represented in a low dimensional subspace to obtain basis vector fields. Using these basis vector fields we can model the nonlinear color changes to transform a new probe face to a reference face in the gallery. Using face color constancy preprocessing, we show that for fixed pose and expression we require a single training image per subject to perform accurate face recognition.

Introduction

The appearance of a face with fixed expression and pose is determined by the illumination conditions. Here we define illumination conditions as including all those parameters which affect the final RGB values recorded by the camera given that pose and expression are fixed. These parameters are determined by the intensity of the light, frequency of the light, position of the light with respect to the face, shadows over the face, reflectance of the regions in the neighborhood of the face, and camera characteristics. The appearance of the same face with fixed pose and expression under different illumination conditions can change significantly. The accuracy of a face recognition algorithm is going to be severely affected due to these large variations. In general, we know that appearance-based face recognition methods like Eigenfaces [2] need a number of training images for each subject in order to cope with illumination variations. We aim to deal with these variations by automatically transforming the faces images towards a reference lighting condition. When the pose and expression are fixed we require only a single training image per subject for accurate face recognition.

We test our approach on PIE database [4] since the illumination variations are more representative of the real world situations than other standard face databases currently available. The illumination variations in PIE database were generated due to presence or absence of ambient light combined with changing position of a point source with respect to the face. The changes in colors, shadows and specular reflections on faces due to the changing illuminations are highly complex. Also we do not have any physical parameters of the lighting, surface reflectances or camera sensitivities. Hence our face color constancy (FCC) approach needs to be completely unsupervised. We realize this by using the color changes observed on a face for different illuminations to drive the model. By training on few faces of different ethnicities, our model is able to generalize the color mapping for any unseen ethnicity from the PIE database. After the FCC algorithm is trained we obtain basis vector fields, which can be used to explain the observed color changes. Using these basis vectors we can transform a face image towards a reference illumination condition. In a face recognition system with a single image per subject, the basis vectors can be used to transform a probe face towards its reference in the gallery.

Previous work

Many interesting approaches have been used to address the problem of illumination invariant face recognition. Most of the work has been done in building models that try to explicitly model varying illumination. The human face is treated as a Lambertian surface and mathematical models are created which can describe all images it can produce under all possible illuminations. In [1], using spherical harmonics it has been 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. A shape from shading method was attempted in [9], but suffered from drawbacks of shape form shading. Another well known model that is widely used is the illumination cone model [3]. It was also shown that illumination cones of human faces can be approximated well by low-dimensional linear subspaces [8]. These appearance-based methods need training images of subjects under a number of different illumination conditions. This drawback is overcome in [5] and [6] by reconstructing 3D face information for each subject in the training set.

Color constancy algorithms are very appealing to use for compensating face image variations taken under different illumination conditions. A common approach in the color constancy literature is to use linear models of reflectance and illuminant spectra to estimate surface reflectances and various physical parameters. Most of the methods make strong assumptions about the distribution of reflectances. Gray world algorithms [11] assume that the average reflectance of all the surfaces in a scene is gray. While the white world algorithms [12] assume that the brightest pixel corresponds to a scene point with maximal reflectance. The assumptions make these algorithms impractical for use on real world images. Some of the algorithms require prior knowledge of the illuminant spectra, surface reflectances and camera sensitivities during training to achieve color constancy [10, 15]. There exist very few color constancy algorithms [13, 14] that work on real images, but only for limited conditions. Hence for the purpose of FCC on real world images we develop a data driven approach that can be applied directly to real world images. Our approach is influenced by the color eigenflows method [7], which was used to "flow" an image to a target image using the eigenflows. Their method was trained on a wide range of colors while our method is specific to faces hence more appropriate for FCC.

Our approach of using color constancy to achieve illumination invariant face recognition is unique in many ways compared to previous illumination invariant face recognition methods. We can deal with real world illumination environments where many light sources are active simultaneously, the position of the light sources are changing and the neighborhood scene is changing. Our model is also capable of FCC when camera processing such as auto-gain and camera color balancing functions are active. Given a few faces under different illumination environments in these real word situations, our model can learn the basis vectors needed to compensate for these variations in any new face image. Our model can also compensate for strong shadows and specular reflections to some extent.

Learning the joint color changes

We perform training on face images from the PIE database [4] to learn the color changes on faces for different lighting conditions. We use only frontal faces with normal expression. The database has two sets of illumination variations. In set 1 the ambient lights are on and the point source is changing its state and position. In set 2 the ambient lights are off and again the point source is changing its state and position. We use set 1 to learn the joint color changes.

Let the RGB color space be defined as, $C = \{(r, g, b)^T \in \Re^3 : 0 \le r \le 255, 0 \le g \le 255, 0 \le b \le 255\}$. This space defines all the possible color vectors observable in images. The color vector of an image pixel p is denoted as $c(p) \in C$. Let $1 \le i \le N$, where N is the number of subjects and let $1 \le j \le M$, where M is the number of illumination conditions under which each subject's image is taken. Also let $1 \le k \le P$, where P is the number of pixels in each face image. The mapping of colors under different illumination conditions is represented by difference of two corresponding pixels:

$$\mathrm{d}(I_{i0}^k, I_{ij}^k) = \mathrm{c}(I_{i0}^k) - \mathrm{c}(I_{ij}^k) \tag{1}$$

This difference vector tells us how a particular pixel's value changed from illumination condition of I_{i0} to illumination condition of I_{ij}. This difference vector is computed for each of the P pair of pixels to obtain a vector field that is defined at all points in C for which there are colors in image I_{ij}. The vector field is constructed by placing each vector difference at the point c(I_{ij}^k) in the color space C.

The vector field Φ' over C is defined as:

$$\Phi'(c(I_{ij}^k)) = d(I_{i0}^k, I_{ij}^k), \qquad 1 \le k \le P \tag{2}$$

This vector field is only defined at particular color points in C that happen to be in image I_{ij}^k, hence Φ' is called a partially observed color flow. We wish to obtain a full color flow from the partially observed color flow. A simple approach to obtain the full color flow is to follow an interpolation scheme as proposed in [7]. The color flow at a color point $(r, g, b)^T$ is obtained by a weighted proximity based average of nearby observed color flow vectors.

$$\Phi(r,g,b) = \frac{\sum_{k=1}^{P} e^{-\left\| c(I_{ij}^k) - (r,g,b)^T \right\|^2 / 2\sigma^2} \Phi'(c(I_{ij}^k))}{\sum_{k=1}^{P} e^{-\left\| c(p_1^k) - (r,g,b)^T \right\|^2 / 2\sigma^2}} \tag{3}$$

In [7] the full color flow is defined at every point in C, whereas our full color flow is defined only in a subset of C. The points where our Φ is defined depends on the colors present in I_{ij}. The variance term σ^2 controls the mixing of the observed flows to form the interpolated flow vectors. From our experiments we found that a variance of 16 gave us good results. Visualization of the color flow vector field for the face pair (I_{10}, I_{12}) is shown in figure 1.

The full color flows $\Phi(c(I_{ij}^k))$ are computed for all the training images; $1 \le i \le N$, $1 \le j \le M$. We used N=34 and M=23; i.e. we used images of 34 subjects taken under 23 different illumination conditions to obtain 782 color flows. The training images are of dimension 190 x 200; i.e. P = 38000. The training face images were created form the PIE database by interactively cropping and masking the faces from the images.

The possible changes of a pixel's color on the face surface due to changes in illumination conditions are compact. While in principle its possible for a change in illumination condition to map any color from a Lambertian surface to any other color independently of all other colors, we know from experience that many such joint maps are not observed in real world situations. Hence there is significant structure in the space of color flows. The image pair shown in figure 1 has a nonlinear variation; some regions of the face become more bright than other regions due to position of the point source. As can be seen from the visualization, this kind of nonlinear variations can be captured by modeling the space of color flows.

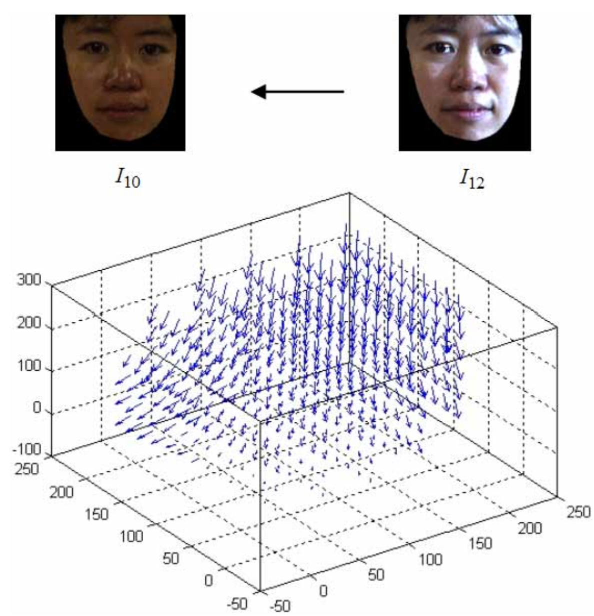

***Figure 1**. Visualization of the full color flow vector field for the image pair (I_{10}, I_{12}).*

Given a large number of color flows, we wish to model their distribution. We chose to use Principal Component Analysis (PCA) due to the following reasons, which were originally given in [7]. The flows are well represented by a small number of principle components and finding the optimal description of a difference image in terms of color flows is computationally efficient using this representation.

There are ~16 million points in the color space C, hence to represent the color flows we quantize it at Q distinct points. Therefore the color flow Φ can be represented as a collection of $3Q$ coordinates. We chose Q = 4096 distinct and equally spaced points in the color space for our experiments. Hence the full color flow is a vector of 3 * 4096 components. Using a higher value of Q would give us a more accurate color flow field, but due to computational speed and memory limitations we settle for $Q=16^3$. We compute principal components of the color flow covariance matrix. These principal components were named as eigenflows in [7]. Figure 2 shows the eigenvectors associated with the first 50 eigenflows. As expected the curve drops off rapidly indicating that most of the variance in the color flow distribution is represented by the first few eigenflows.

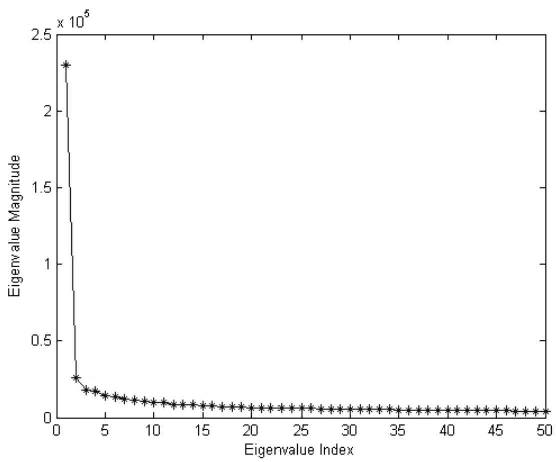

***Figure 2**. Magnitude of the eigenvalues vs. eigenvalue index.*

Face color constancy from basis vectors

Using the eigenflows, we can transform a face image taken under arbitrary illumination conditions towards its reference illumination condition. During the training procedure we used the face image taken with ambient light on and all point sources off as the reference, hence using the eigenflows we can obtain the face closest to that illumination condition given a face taken under an arbitrary illumination condition.

Let I_{test} be a face under arbitrary illumination condition and let $I_{test\,0}$ be its reference. We compute the difference image as

$$D = I_{test0} - I_{test} \tag{4}$$

The difference image basis vectors for the test image and a set of E eigenflows Ψ_i, $1 \le i \le E$, can be represented as

$$D_i = I_{test}(\Psi_i) \tag{5}$$

Here the operator $I_{test}(\cdot)$ takes the pixel values at the location [x, y] and generates a difference image basis vector by placing at each [x, y] the closest eigenflow. The closest eigenflow is determined based on the distance in the color space from the color vector at [x, y]. The transformed image is obtained as

$$I_T = I_{test} + \sum_{i=1}^{E} \gamma_i D_i \tag{6}$$

where γ_i are scalar multipliers. We can directly solve for γ_i's by solving the system

$$\gamma_i = D_i^{\dagger} D \tag{7}$$

Here D is the difference image defined in eq. (4), and $D_i^{\dagger}$ is the pseudo-inverse of the difference image basis vectors defined in eq. (5).

We applied this procedure on 10 faces of arbitrary illumination conditions and transformed them to their reference illumination condition. The mean RMS error of the original faces with respect to the reference face was 70.51. We then computed the mean RMS errors of the transformed faces obtained using different number of eigenflows from the reference face (see figure 3). Using only 20 eigenflows the mean RMS error reduced to 7.65.

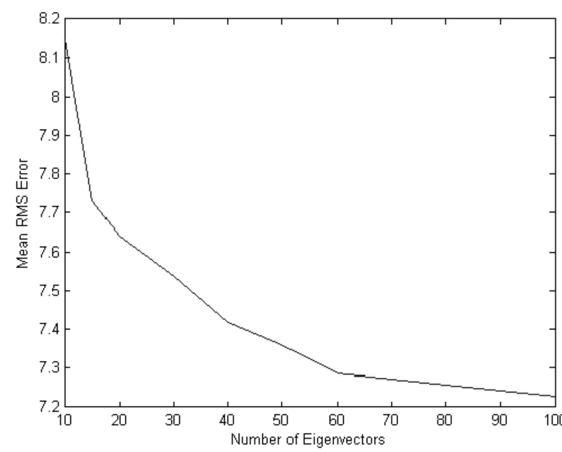

***Figure 3**. Mean RMS error using different number of eigenflows. Mean RMS error of the original images was 70.51.*

Figure 4 shows a few probe faces transformed towards the reference illumination condition. Notice that the effects of shadows and specular reflections are reduced and the color of the face is consistent with respect to the reference face. The subject shown in figure 4 was not used during training to compute the eigenflows. This shows that our FCC algorithm is capable of generalizing color constancy for unseen faces.

Face recognition using FCC

We describe a procedure to perform face recognition based on the RMS error between the transformed probe face and the face in the gallery. Let N be the number of images in the gallery, each image corresponding to a different subject. Also let all the gallery images be taken under the same illumination condition. Then we transform a probe image taken under arbitrary illumination condition towards all the N gallery images. We obtain N RMS errors between the N transformed images and their corresponding references in the gallery. Then the reference face in the gallery with minimum RMS error is considered to be the recognized face.

Using this procedure we performed face recognition with 10 images in the gallery, a single image used for each of the 10 subjects. We used 50 probe images, 10 images each of 5 subjects, the 5 subjects in probe set are present in the gallery set. The recognition results obtained are shown in table 1. We compared the face recognition accuracy based on the RMS error for direct image matching without processing the probe images and after processing the probe images with FCC.

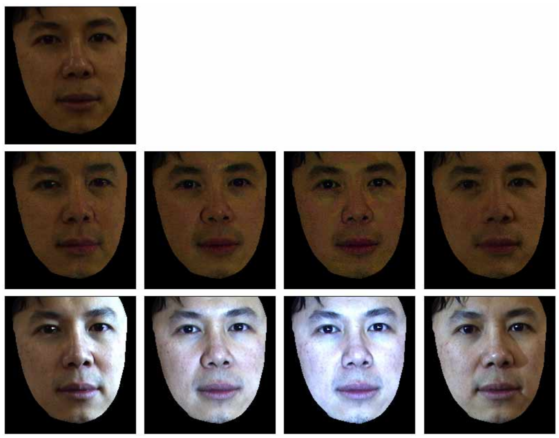

***Figure 4**. Probe faces (third row) are transformed (second row) to match the illumination condition of the reference face (first row).*

Comparisons of face recognition results: Table 1

Face recognition approach	% True positive	% False positive
Direct	20	80
FCC	94	6.0

We aim to transform the probe face to its correct reference face as well as possible, but at the same time we expect our FCC model is not endowed with enough capacity to transform a probe face to a wrong reference accurately. An ideal FCC model would thus transform a probe face to the correct reference with zero RMS error, and have large RMS errors for faces transformed with wrong references. From the face recognition results we see that our FCC model is not over-parameterized.

Conclusions and future work

The results show that our face color constancy (FCC) model is robust and can compensate for large illumination variations, shadows and specular reflections over faces. We also showed that our approach could have application in illumination invariant face recognition. The face recognition experiments presented are preliminary and much large experiments need to be performed to establish its utility for illumination invariant face recognition. We believe an important application of our FCC algorithm could be to use it in conjunction with other subspace based illumination invariant face recognition methods to make those methods more robust to illumination variations.

References

[1] R. Basri and D. Jacobs, "Lambertian Reflectance and Linear Subspaces," IEEE Trans. on PAMI, Vol. 25, No. 2, pp. 218-233, Feb. 2003

[2] M.Turk and A.Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, Vol. 3, No. 1, pp. 71-96, 1991.

[3] P. Belhumeur and D. Kriegman, "What is the set of images of an object under all possible lighting conditions?," Int'l J. Computer Vision, Vol. 28, No. 3, pp. 1-16, 1998.

[4] T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression Database," IEEE Trans. on PAMI, Vol. 25, No. 12, pp. 1615-1618, Dec. 2003.

[5] T. Sim and T. Kanade, "Combining Models and Exemplars for Face Recognition: An Illuminating Example," Proc. CVPR Workshop on Models versus Exemplars in Computer Vision, 2001.

[6] A. Georghiades, P. Belhumeur, and D. 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, Dec. 2001.

[7] E. Miller and K. Tieu, "Color Eigenflows: Statistical Modeling of Joint Color Changes," Proc. ICCV, Vol. 1, pp. 607-614, 2001.

[8] A. Georghiades, D. Kriegman, and P. Belhumeur, "Illumination cones for recognition under variable lighting: Faces," Proc. CVPR, 1998.

[9] W.Y. Zhao and R. Chellappa, "Illumination Insensitive Face Recognition using Symmetric Shape-from-Shading," Proc. CVPR, Vol. 1, pp. 286-293, 2001.

[10] D. Brainard and W. Freeman, "Bayesian Color Constancy," J. Opt. Soc. Am. A, Vol. 14, No. 7, pp. 1393-1411, July 1997.

[11] G. Buchsbaum, "A spatial processor model for object color perception," Journal of the Franklin Institute, 310, 1980.

[12] J. McCann, J. Hall, and E. Land, "Color mondrian experiments: The study of average spectral distributions," J. Opt. Soc. Am., 67, 1380, 1977.

[13] B. Buluswar and S. Draper, "Color recognition in outdoorimages," Proc. ICCV, January 1998.

[14] D. Forsyth, "A novel approach for color constancy," Int'l J. of Computer Vision, 5, pp. 5–36, 1990.

[15] B. Funt and V. Cardei, "Bootstrapping color constancy," Proc. SPIE: Electronic Imaging IV, Vol. 3644, Jan. 1999.

Author Biography

Rajkiran Gottumukkal received the BS degree in Electronics & Communication Engineering from Bangalore University, India, in 1999; and the MS degree in Electrical Engineering from Old Dominion University, in 2003. He is currently pursuing a PhD degree in Electrical and Computer Engineering at Old Dominion University.

Vijayan K. Asari joined the Department of Electrical and Computer Engineering, Old Dominion University, Virginia as an Associate Professor in August 2000. Dr. Asari received the B.Sc. (Engg.) degree in Electronics and Communication Engineering from the University of Kerala, M.Tech. and Ph.D. degrees in Electrical Engineering from the Indian Institute of Technology, Madras, India.. His research activities are in the areas of signal processing, image processing, neural networks, and digital system architectures for application specific integrated circuits. Dr. Asari is a senior member of the IEEE.