Colour Eigenfaces

Graham D. Finlayson†, Janet Dueck⋆, Brian V. Funt⋆, and Mark S. Drew⋆

†Department of Computer Science, University of York, York YO1 5DD

⋆School of Computing Science, Simon Fraser University, Vancouver, Canada.

email graham@minster.cs.york.ac.uk, {janet,funt,mark}@cs.sfu.ca

Abstract

Images of the same face viewed under different lighting conditions look different. It is no surprise then that face recognition systems based on image comparisons can fail when the lighting conditions vary. In this paper we address this failure by designing a new lighting condition independent face matching technique. We begin by demonstrating that the colour image of a face viewed under any lighting conditions is a linear transform from the image of the same face viewed under complex (3 lights at 3 locations) conditions. Our new matching technique solves for the best linear transform relating pairs of face images prior to calculating the image difference. For a database of 15 (complexly illuminated) faces and 45 test face images the new matching method delivers perfect recognition. In comparison, matching without accounting for lighting conditions fails 25% of the time.

I. Introduction

One of the most successful and widely used technique for face recognition is the eigenface method of Turk and Pentland[9], [8]. The basic idea in that method is, that the **greyscale** images of the same face seen in different circumstances should be quite similar. Recognition takes place by comparing the image of an unknown face with face images stored in a database. The closest database image identifies the face. Because, in general, images are very large, image matching is very expensive. In order to reduce matching cost, Turk and Pentland approximated each face image as a linear combination of a small set of basis faces called *eigenfaces*.

Unfortunately, images of the same face viewed under different lighting conditions rarely look the same i.e. their shading fields will differ. This problem can be mitigated by viewing each face under a variety of lighting conditions and storing this variation in the face database[1], [3], [5]. The multiple image approach succeeds because each separate image encodes a certain amount of information about the shape of the face; that is, at an implicit level, the multiple image approach is concerned with matching shape. However, it is not clear how the notion of shape can be made explicit. We certainly do not want to solve for shape since although this can be done[10] highly specialized calibrated conditions are needed.

In this paper we show that shape information is easily obtained so long as face recognition is based on colour images. Specifically we show that: **the implicit notion of shape is explicitly captured in a single 3-band colour image**. This result follows from Petrov's[6] seminal work on the relationship between illumination, reflectance, shape and colour images in which he demonstrated that, so long as a Lambertian surface is viewed under a complex illumination field (at least 3 light spectrally distinct light sources at different locations), the rgb pixel triplets in an image are a linear transform from scene surface normals: **colour is a linear transform from shape**.

In our method each database face is created with respect to a complex illumination field. Face recognition simply involves matching the image of an unknown face to the face database. Each database face image is first transformed (by a linear transform) to best match the image colours of some unknown face. Thereafter, the residual difference is calculated. The database face with the smallest residual difference overall identifies the face. In line with Turk and Pentland the cost of matching is reduced by approximating face images using a small number of *colour eigenfaces*.

II. Face recognition using Eigenfaces

Let us represent an $n \times n$ greyscale image by the function I such that $I(x, y)$ denotes the grey-level at location x, y. Suppose we have a database $\mathcal{M}$ of m face images: $\mathcal{M} = \{I_1, I_2, \cdots, I_m\}$. Face recognition is all about finding the image I_c in $\mathcal{M}$ which is *closest* to some unknown face image I_u. Mathematically we might define a function Φ which takes I_u and $\mathcal{M}$ as parameters and returns the closest match I_c:

$$\Phi(I_u, \mathcal{M}) = I_c \;\; : \; I_c \in \mathcal{M} \; \& \; ||I_c - I_q||_d < ||I_i - I_q|| \;\; (i = 1, 2, \cdots, c-1, c+1, \cdots, m) \tag{1}$$

where $||.||_d$ is a distance measure (usually Euclidean) which quantifies the similarity of two images. To reduce computation a face image I can be represented (approximately) by a linear combination of basis faces (which

Turk and Pentland call *eigenfaces*).

$$I \approx \sum_{i=1}^{n} \beta_i B_i \tag{2}$$

here B_i is the ith (of n) eigenface and β_i are weighting coefficient chosen to minimize:

$$||I - \sum_{i=1}^{n} \beta_i B_i||_d \tag{3}$$

Clearly the error in the approximation defined in (3) depends on the set of eigenfaces used. In general the eigenfaces are selected to minimize the expected residual difference in (3). This is done using standard statistical techniques (e.g. principal component analysis[4]). However, eigenfaces based on other error criteria are sometimes used[7]. Turk and Pentland have shown that a small number of eigenfaces (just 7) renders the error in (2) reasonably small. Denoting eigenface approximations with the superscript $'$, the function Φ' is defined as:

$$\Phi'(I_u, \mathcal{M}) = I'_c \ : \ I'_c \in \mathcal{M}' \ \& \ ||I'_c - I'_q||_d < ||I'_i - I'_q|| \ \ (i = 1, 2, \cdots, c-1, c+1, \cdots, m) \tag{4}$$

Because each of I'_c and I'_q are defined by just n numbers (the coefficients β in (2)) it is staightforward to show that the cost of each image comparison is proportional to n. Usually $n << \#$ pixels in an image, so matching is very fast. Turk and Pentland[8] have shown that the function Φ' suffices for face recognition so long as illumination conditions are not allowed to vary too much.

III. Colour and shape

The light reflected from a surface depends on the spectral properties of the surface reflectance and of the illumination incident on the surface. In the case of Lambertian surfaces (these are the only kind we consider here), this light is simply the product of the spectral power distribution of the light source with the percent spectral reflectance of the surface. Illumination, surface reflection and sensor function, combine together in forming a sensor response:

$$\underline{\rho}^{\hat{x}} = \underline{e}.\underline{n}^x \int_w S^x(\lambda) E(\lambda) \underline{R}(\lambda) d\lambda \tag{5}$$

where λ is wavelength, $\underline{\rho}$ is the 3-vector of sensor responses (*rgb* pixel value) $\underline{R}$ is the 3-vector of response functions (red-, green and blue- sensitive), E (assumed constant across the scene) is the incident illumination and S^x is the surface reflectance function at location x on the surface which is projected onto location $\hat{x}$ on the sensor array. The relative orientation of surface and light is taken in account by the dot-product of the surface normal vector $\underline{n}^x$ with the light source direction $\underline{e}$ (both these vectors have unit length).

Let us denote $\int_w S^x(\lambda)E(\lambda)\underline{R}(\lambda)d\lambda$ as $\underline{q}^x$. It follows that (5) can be rewritten as:

$$\underline{\rho}^{\hat{x}} = \underline{q}\underline{e}^t \underline{n}^x \tag{6}$$

where ^{t} denotes vector transpose ($\underline{e}.\underline{n}^x = \underline{e}^t \underline{n}^x$). Now consider that a scene is illuminated by two spectrally distinct light sources at distinct locations. If we denote illumination dependence using the subscripts $_1$ and $_2$ then equation (6) becomes

$$\underline{\rho}^{\hat{x}} = \underline{q}_1 \underline{e}_1^t \underline{n}^x + \underline{q}_2 \underline{e}_2^t \underline{n}^x \tag{7}$$

Assuming k lights incident at x:

$$\underline{\rho}^{\hat{x}} = [\sum_{i=1}^{k} \underline{q}_i \underline{e}_i^t] \underline{n}^x \tag{8}$$

So long as $k \geq 3$ the term $[\sum_{i=1}^{k} \underline{q}_i \underline{e}_i^t]$ will define a 3×3 matrix of full rank. In this case there is a one to one correspondence between the colours in an image and the normal field of a scene. Shape and colour are inexorably intertwined.

It is important to note that the relationship between surface normal and camera response depends on the reflective properties of the observed surface and the particular set of illuminants incident at a point. Changing the reflectance or the illumination field changes the relationship between surface normal and image colour. Henceforth we will assume that faces are composed of a single colour and that faces are illuminated by a homogeneous illumination field and as such a single 3×3 matrix relates all surface normals and image colours.

IV. Face recognition using colour eigenfaces

Let us represent an $n \times n$ colour image by the vector function $\underline{I}$ such that $\underline{I}(x,y)$ denotes the (r,g,b) vector at location x, y and records how *r*ed, *g*reen and *b*lue a pixel appears. As before let us suppose we have a database $\mathcal{M}$ of m images: $\mathcal{M} = \{\underline{I}_1, \underline{I}_2, \cdots, \underline{I}_m\}$. Crucially, we assume that each database face image is created with respect to a complex illumination field and is thus a linear transform from the corresponding normal field. This relationship is made explicit in (9) where where $\underline{N}_i(x,y)$ is a vector function which returns the surface normal corresponding to $\underline{I}_i(x,y)$. The 3×3 matrix relating normal field to image colours is denoted $\mathcal{T}_i$.

$$\underline{I}_i(x,y) \;=\; \mathcal{T}_i \underline{N}_i(x,y) \;\;,\; (i = 1, 2, \cdots, m) \tag{9}$$

Suppose $\underline{I}_u$ denotes the image of an unknown face viewed under arbitrary lighting conditions. Clearly,

$$\underline{I}_u(x,y) \;=\; \mathcal{T}_u \underline{N}_u(x,y) \tag{10}$$

Suppose that $\underline{I}_j$ is an image of the same face (in $\mathcal{M}$). It is unlikely that $\mathcal{T}_j$ will equal $\mathcal{T}_u$. However, it is immediate from (9) and (10) that $\underline{I}_j$ must be a linear transform from $\underline{I}_u$:

$$\mathcal{T}_u \mathcal{T}_j^{-1} \underline{I}_j \;=\; \underline{I}_u \tag{11}$$

where $^{-1}$ denotes matrix inverse. It follows that a reasonable measure for the distance between a database image $\underline{I}_i$ and $\underline{I}_u$ can be calculated as:

$$||\mathcal{T}(\underline{I}_i, \underline{I}_u)\underline{I}_i \;-\; \underline{I}_u||_d \tag{12}$$

where $\mathcal{T}(\underline{I}_i, \underline{I}_u)$ is the 3×3 matrix which *best* maps $\underline{I}_i$ to $\underline{I}_u$. In the experiments reported later $\mathcal{T}()$ returns the matrix which minimizes the sum of squared errors and is readily computed using standard techniques[2]. Relative to (12) a closeness function Ψ for colour face images can be defined as:

$$\Psi(\underline{I}_u, \mathcal{M}) \;=\; \underline{I}_c \;\;:\;\; \begin{array}{l} \underline{I}_c \in \mathcal{M} \;\& \\ ||\mathcal{T}(\underline{I}_c, \underline{I}_u)\underline{I}_c - \underline{I}_u||_d < ||\mathcal{T}(\underline{I}_i, \underline{I}_u)\underline{I}_i - \underline{I}_u|| \;\; (i = 1, 2, \cdots, c-1, c+1, \cdots, m) \end{array} \tag{13}$$

To reduce computational cost of computing (13) we represent (in a similar way to the greyscale method) each band of a colour image as a linear combination of basis vectors:

$$\underline{I}^{\alpha} \;\approx\; \sum_{i=1}^{n} \beta_i^{\alpha} B_i \;\;,\; (\alpha = r, g, b) \tag{14}$$

where r, g and b denote the red, green and blue colour bands. the coefficients β_i^{α} $(\alpha = r, g, b)$ are chosen to minimize the approximation error.

To derive the eigenfaces to use in (14) a training set of colour face images is compiled. Each image is split into its 3 component band images and thereafter principal component analysis on the entire band image set. Denoting colour eigenface approximations with the superscript $'$, the function Ψ' is defined as:

$$\Psi'(\underline{I}'_u, \mathcal{M}) \;=\; \underline{I}'_c \;\;:\;\; \begin{array}{l} \underline{I}'_c \in \mathcal{M} \;\& \\ ||\mathcal{T}(\underline{I}'_c, \underline{I}'_u)\underline{I}'_c - \underline{I}'_u||_d < ||\mathcal{T}(\underline{I}'_i, \underline{I}'_u)\underline{I}'_i - \underline{I}'_u|| \;\; (i = 1, 2, \cdots, c-1, c+1, \cdots, m) \end{array} \tag{15}$$

It can be shown that the cost of calculating (15) is bound by the square of the number of eigenfaces used: matching costs $O(n^2)$ (instead of $O(n)$ for black and white faces).

V. Results

The colour images of 15 people (see Figure 1) viewed under 3 complex illuminations provide a training set for eigenface analysis. We found that 8 eigenfaces provide a reasonable basis set (the approximation in (14) is fairly good). The eigen approximations for the 15 faces viewed under one of the complex illuminations comprises the face database. A further 45 test images were taken (the same faces under 3 more illuminants) under non-complex illuminations.

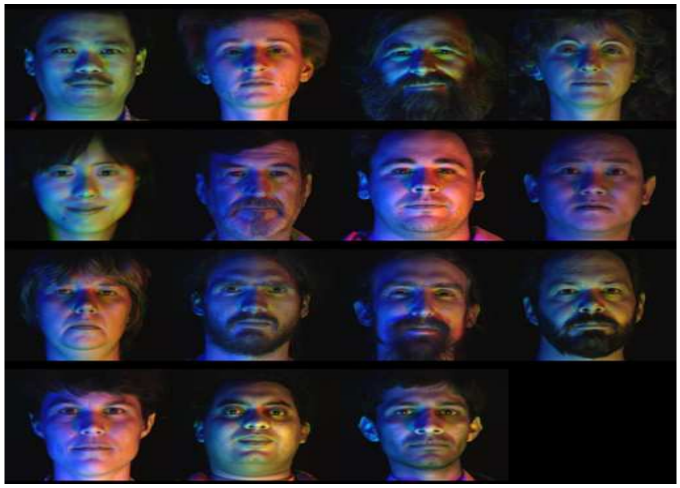

Fig. 1. Colour Face Images

Each test image was compared with each database image using equation (15). The closest database image defines the identity of the face in the test image. We found that all 45 faces (a 100% recognition rate) were correctly identified. Importantly we found that faces were matched with good confidence; on average the second closest database face was at least twice as far from the test image as the correct answer.

We reran the face matching experiment in greyscale using Turk and Pentland's original eigenface method. Greyscale images were created from the colour images (described above) by summing the colour bands together (greyscale=red+green+blue). We found that 7 eigenfaces is sufficient to approximate the training set. As before the face database comprises eigen approximations of each of the 15 faces viewed under a single illuminant. Test images were compared with the face database using (4). We found that only 32 of the faces were correctly identified (a recognition rate of 73%). This is quite poor given that the face database is quite small.

VI. Conclusion

Shape and colour in images are inexorably intertwined. A single coloured Lambertian surface viewed under complex illumination conditions is a linear transform from the surface normal field. It follows that the image of a face observed under any lighting conditions is a linear transform from the same face viewed under a complex illumination field. We use this result in a new system for face recognition.

Database faces are represented by colour images taken with respect to a complex illumination field. Matching takes place by finding the linear transforms which takes each database face as close as possible to a query image. The closest face overall identifies the face (in the query image). To speed computation all faces are represented as a linear combination of a small number eigenfaces. Experiments demonstrated that the colour eigenface method delivers excellent recognition. Importantly recognition performance, by construction, is unaffected by the lighting conditions under which faces are viewed. That this is so is quite significant since existing methods[9] require the lighting conditions to be held fixed (and fail when this requirement is not met).

References

[1] Russell Epstein, Peter J. Hallinan, and Alan L. Yuille. 5±2 eigenfaces suffice: an empirical investigation of low-dimensional lighting models. In *Workshop on Physics-Based Modelling, ICCV95*, pages 108–116, 1995.

[2] G.H. Golub and C.F. van Loan. *Matrix Computations.* John Hopkins U. Press, 1983.

[3] Peter W. Hallinan. A low-dimensional representation of human faces for arbitrary lightin conditions. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 995–99, 1994.

[4] I.T. Jolliffe. *Principal Component Analysis.* Springer-Verlag, 1986.

[5] Shree K. Nayar and Hiroshi Murase. Dimensionality of illumination manifolds in eigenspace. Technical report, Columbia University, 1995.

[6] A.P. Petrov. On obtaining shape from color shading. *COLOR research and application*, 18(4):236–240, 1993.

[7] D.J. Kriegman P.N. Belhumeur, J.P. Hespanha. Eigenfaces vs fisherfaces: recognition using class specific linear projection. In *The Fourth European Conference on Computer Vision (Vol I)*, pages 45–58. European Vision Society, 1996.

[8] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, March 1991.

[9] M. Turk and A. Pentland. Face recognition using eigenfaces. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 586–91, 1991.

[10] R.J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19:139–144, 1980.