BMVC 2002

Face Recognition using Shape-from-shading

W.A.P. Smith and E.R. Hancock
Dept of Computer Science, University of York
will@smith.net,erh@cs.york.ac.uk

Abstract

This paper explores how curvature attributes delivered by the Worthington and Hancock shape-from-shading algorithm can be used for the purposes of face recognition. We show that histograms of mean and Gaussian curvature can be used for pose invariant recognition. Recognition rates of up to 94% may be achieved.

1 Introduction

The use of shading and illumination information for face analysis has been a long-standing goal in computational vision. The topic has attracted considerable interest in both the psychology and computer vision literature. In the psychology literature, several authors have investigated the effects of orientation, albedo and shading and have established their importance for face perception and recognition. For instance, Enns and Shore [12] studied the influences of orientation and lighting on the inverted-face effect. Several authors including Kemp, Pike, White and Musselman [13] and Braje, Kersten, Tarr and Troje [4] have studied the role of illumination effects in face recognition. In the computer vision area, Troje and Bulthoff [15] investigated the role of facial shape and texture in the recognition of shaded faces. Yuille, Ferraro and Zhang [19] have shown how shading information can be used to control the warping of a prototype shape onto image data. Atick, Griffin and Redlich [1] show how principal component analysis can be used to reconstruct 3D facial surfaces from 2D images.

The aim in this paper is to investigate whether shape-from-shading can be used for the purposes of face recognition. Because of its perceived fragility, the use of surface orientation or needle map information has received little interest. One of the reasons for this is that shape-from-shading has proved notoriously difficult to regulate and the recovered surface orientation information is of questionable acuity. In particular, the recovered needle-maps are dominated by the smoothness constraints needed to solve the image irradiance equation. As a result, the image irradiance equation is only weakly satisfied and the recovered orientation information is a poor representation of reality. For this reason, much of the face-analysis literature has focused on recovering shape from multiple light source images or by developing more sophisticated appearance based models to account for the observed shading variations. For example, Belhumeur and Kriegman [2] have developed a photometric stereo method which relies on matrix factorisation

BMVC 2002 doi:10.5244/C.16.58

to extract shape and albedo information from multiple light source images of the human face. Jacobs, Belhumeur and Basri [7] have shown how to use this representation to compare faces under different illumination conditions. More recently, Belhumeur, and Hespanha and Kriegman [3] have shown how to use a class specific linear projection to recognise different faces under variations in illumination. Georghiades, Kriegman and Belhumeur [8] have a more physically based method which allows the illumination cone to be used to effect recognition under variable lighting conditions when a Lambertian reflectance model applies. Zhao and Chellappa [20] have recently descibed a symmetric shape-from-shahding method which is tailored to face analysis. The method has produced promising recognition results.

Worthington and Hancock have recently reported a new framework for shape-from-shading [17] which holds out the possibility of performing shaded face analysis using a single image. Their main contribution has been to develop an efficient geometric means of iteratively adjusting the needle-map orientation, which ensures that the image irradiance equation can be satisfied as a hard constraint. Moreover, they have developed alternatives to the bland assumption of quadratic smoothness which allow gradient consistency and curvature consistency constraints to be imposed on the recovered needle-maps. The net effect is to recover orientation information, which is a more faithful representation of ground-truth. The surface orientation and curvature information delivered by the method hase been successfully used for a number of tasks including 3D object recognition from 2D views [18], view synthesis and facial pose estimation.

The aim in this paper is to investigate whether the curvature information delivered by the method can be used for histogram-based face recognition. Histograms have proved to be simple and powerful attribute summaries which can be used to great effect in the recognition of objects from large image data-base. The idea was originally popularised by Swain and Ballard who used colour histograms [16]. There have since been several developments of the idea. For instance Gimelfarb and Jain [9] have used texture histograms for 2D object recognition, Dorai and Jain [6] have used shape-index histograms for range image recognition and Huet and Hancock have used relational histograms for line-pattern recognition [11].

2 Shape-from-Shading

The shape-from-shading algorithm of Worthington and Hancock has been demonstrated to deliver needle-maps which preserve fine surface detail [17]. The observation underpinning the method is that for Lambertian reflectance from a matte surface, the image irradiance equation defines a cone of possible surface normal directions. The axis of this cone points in the light-source direction and the opening angle is determined by the measured brightness. If the recovered needle-map is to satisfy the image irradiance equation as a hard constraint, then the surface normals must each fall on their respective reflectance cones. Initially, the surface normals are positioned so that their projections onto the image plane point in the direction of the image gradient. Subsequently there is iterative adjustment of the surface normal directions so as to improve the consistency of the needle-map. In other words, each surface normal is free to rotate about its reflectance cone in such a way as to improve its consistency with its neighbours. This rotation is a two-step process. First, we apply a smoothing process to the current surface normal estimates. This may be done in a number of ways. The simplest is local averaging. More sophisticated alternatives include robust smoothing with outlier reject and, smoothing with curvature or image gradient consistency constraints. This results in an off-cone direction for the surface normal. The hard data-closeness constraint of the image irradiance equation is restored by projecting the smoothed off-cone surface normal back onto the nearest position on the reflectance cone.

To be more formal let $\mathbf{s}$ be a unit vector in the light source direction and let $E_{i,j}$ be the brightness at the image location (i,j). Further, suppose that $\mathbf{n}^k(i,j)$ is the corresponding estimate of the surface normal at iteration k of the algorithm. The image irradiance equation is

$$E(i,j) = \mathbf{n}^k_{i,j}.\mathbf{s} \tag{1}$$

As a result, the reflectance cone has opening angle $cos^{-1}E(i,j)$. After local smoothing, the off-cone surface normal is $\bar{\mathbf{n}}^k_{i,j}$. The updated on-cone surface normal which satisfies the image irradiance equation as a hard constraint is obtained via the rotation

$$\mathbf{n}^{k+1}_{i,j} = \Phi\bar{\mathbf{n}}^k_{i,j} \tag{2}$$

The matrix Φ rotates the smoothed off-cone surface normal estimate by the angle difference between the apex angle of the cone, and the angle subtended between the off-cone normal and the light source direction. This angle is equal to

$$\theta = \cos^{-1} E - \cos^{-1} \frac{\bar{\mathbf{n}}^k_{i,j} \cdot \mathbf{s}}{||\bar{\mathbf{n}}^k_{i,j}|| \cdot ||\mathbf{s}||} \tag{3}$$

This rotation takes place about the axis whose direction is given by the vector

$$(u,v,w)^T = \bar{\mathbf{n}}^k_{i,j} \times \mathbf{s} \tag{4}$$

This rotation axis is perpendicular to both the light source direction and the off-cone normal. Hence, the rotation matrix is

$$\Phi = \begin{pmatrix} c + u^2c' & -ws + uvc' & vs + uwc' \\ ws + uvc' & c + v^2c' & -us + vwc' \\ -vs + uwc' & us + vwc' & c + w^2c' \end{pmatrix}$$

where $c = \cos\theta$, $c' = 1 - c$, and $s = \sin\theta$.

For the purposes of this paper, the smoothing constraint used was a simple Gaussian smoothing filter.

3 Attribute Histograms

The aim in this paper is to investigate whether histograms of mean and Gaussian curvature can be used for the purposes of pose invariant face recognition. Gordon [10] has listed three potential of curvature features offer over intensity-based features for face recognition. Specifically, curvature features 1) have the potential

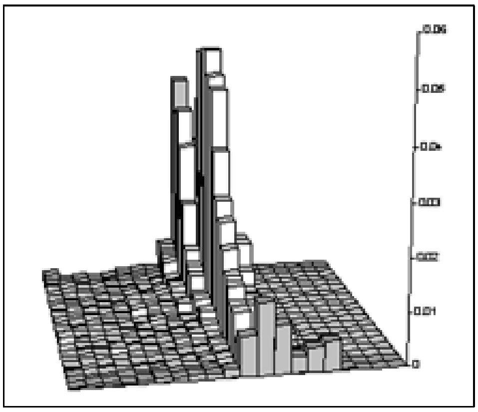

Figure 1: Attribute Histogram of mean and Gaussian Curvature

for higher accuracy in describing surface based events, 2) are better suited to describe properties of the face in areas such as the cheeks, forehead, and chin, and 3) are viewpoint invariant. Although there are several alternatives, Gordon [10] found mean and Gaussian curvature to be effective in the recognition of faces from range imagery. Moreover, the 3D object recognition experiments of Worthington and Hancock revealed that the mean and Gaussian curvatures extracted from the needle-maps delivered by shape-from-shading gave good results when used in conjunction with a simple histogram comparison method.

The differential structure of a surface is captured by the Hessian matrix, which may be written in terms of the derivatives of the surface normals as

$$\mathcal{H} = \begin{pmatrix} -\left(\frac{\partial \mathbf{n}}{\partial x}\right)_x & -\left(\frac{\partial \mathbf{n}}{\partial x}\right)_y \\ -\left(\frac{\partial \mathbf{n}}{\partial y}\right)_x & -\left(\frac{\partial \mathbf{n}}{\partial y}\right)_y \end{pmatrix} \tag{5}$$

where $(\cdots)_x$ and $(\cdots)_y$ denote the x and y components of the parenthesized vector respectively. The eigenvalues of the Hessian matrix, found by solving the equation $|\mathcal{H} - \kappa\mathbf{I}| = 0$, are the principal curvatures of the surface, denoted by $\kappa_{1,2}$. The mean curvature is related to the trace of the Hessian matrix and is given by $H = \frac{1}{2}(\kappa_1 + \kappa_2)$. The Gaussian curvature is equal to the determinant of the Hessian and is given by $K = \kappa_1 \kappa_2$.

From the mean and Gaussian curvatures, we compute a 2D histogram. The histograms contained 20×20 bins, an example of which is shown in Figure 1.

Our aim is to perform face recognition by comparing histograms of mean and Gaussian curvature. The choice of distance metric for measuring the similarity between histograms may affect recognition performance [11]. The L1 and L2 norms are both commonly used in histogram comparison. However, Huet and Hancock [11] found that the Bhattacharyya distance offers significant advantages over the L1 and L2 norms within the context of histogram-based retrieval from large image databases. Suppose we have a query histogram P_Q and a data-histogram P_M. The Bhattacharyya distance between them is defined as

$$d_{bhat}\left(P_Q, P_M\right) = -\log\left(\sum_{i=1}^{x}\sum_{j=1}^{y}\sqrt{P_Q\left(i,j\right) \times P_M\left(i,j\right)}\right), \tag{6}$$

where $P_Q\left(i,j\right)$ denotes bin $\left(i,j\right)$ of histogram P_Q.

4 Experiments

Our experimental evaluation of the face recognition method is divided into two parts. We commence by discussing the data used in our experiments and showing some of the needle-maps. In the second part, we report the recognition results obtained. The experiments are conducted on both synthetic and real world data is shown.

4.1 The Data

The method was tested on two sets of data. The first was taken from a publicly available database of pre-rendered synthetic face images from the Max Planck Institute of Biological Cybernetics, Germany. This database contains images of 200 laser-scanned heads. There are 100 male subjects and 100 female subjects. Each head was scanned using a Cyberware laser range imager. What makes this database particularly useful is that there is registered range and intensity data. This means that by texture mapping the intensity image onto the 3D model recovered using the range imager, visually accurate reconstructions for arbitrary poses may be made. Moreover, these synthetic views contain realistic surface variations. Hence the data is not only more realistic than purely synthetic data that contains no surface reflectance variations, but it also furnishes accurate pose information for the purposes of ground truthing the method. The database contains 7 views of each head: a frontal image and six yaw rotations of 0 degrees, +30 degrees, -30 degrees, +60 degrees, -60 degrees , +90 degrees and -90 degrees. The last two poses hence correspond to profile views. Each subject wore a black swimming hat during the range and intensity data acquisition process. This removes the majority of the hair from the recovered model. The images are rendered using a uniform black background.

However, this database was not prepared with shape from shading in mind. Hence, the publicly available version does not contain illumination conditions that would be ideally suited to analysis using shape from shading. For this reason, for 40 of the faces (20 male and 20 female) three poses were re-rendered using just a single point light source situated at the camera viewpoint. The 3 poses were frontal, +30 degrees and -30 degrees.

For the purposes of testing the approach on real world data, we constructed a database of 12 subjects in 3 poses. The rotations were not accurately controlled but were intended to approximate those present in the database of synthetic images. This database includes many real-world imperfections such as regions of brightness saturation, albedo changes and departures from the Lambertian reflectance assumption. Figure 6 shows the three poses for a subject from each of

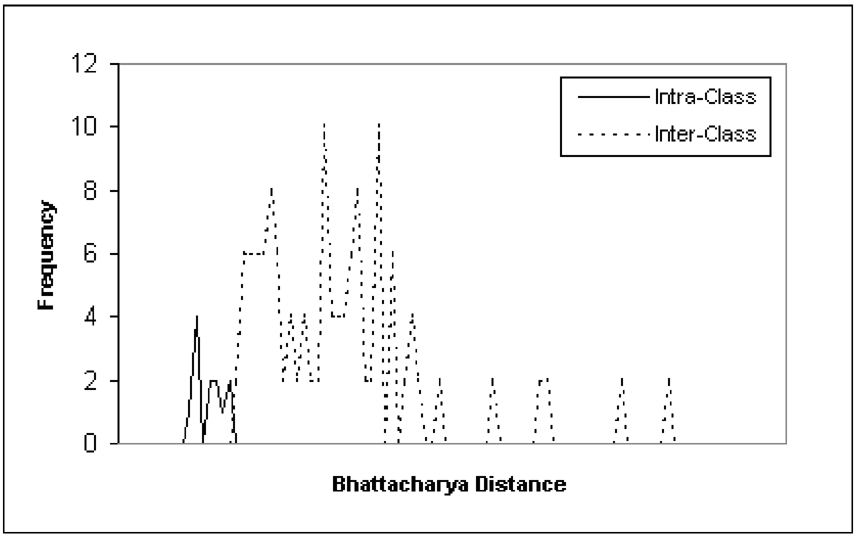

Figure 2: Distribution of Intra- and Inter-class distances for the real world images

the two sets. Figure 7 shows an example of a needle-map recovered using the shape-from-shading scheme on one of the real world images.

4.2 Results

We now turn our attention to the recognition performance for the data-base described above.

We commence in Figure 2 by showing the distributions of within and between class distances for the real world images. The solid curve shows the within class distance distribution and the dotted curve shows the between class distance distribution. The distances are computed for every pair of images. The within distribution is computed from the different poses of the same subject. The between class distribution is computed from all pairs of images of different subjects. The distributions show perfect seperation, i.e. the maximum intra-class distances is smaller than the minimum inter-class distribution.

Figure 3 shows the distributions of between and within class distances for the synthetic face images. For this dataset the distributions are not so well seperated. Approximately half the between-class distances lie within the within-class distance distribution, and would hence lead to mis-recognition. This is due to the increased size of the dataset, which increases the likelihood of misidentification.

A useful measure of performance is the number of correct identifications made when the nearest k-matches are considered. Figure 4 shows a bar chart of the number of correct matches for when the number of nearest neighbours considered is 1, 5 and 10. The plot also investigates the effect of the number of smoothing iterations used in the shape-from-shading scheme. In each bar of the plot, the different sub-bars show the fraction of correct matches obtained with the differ-

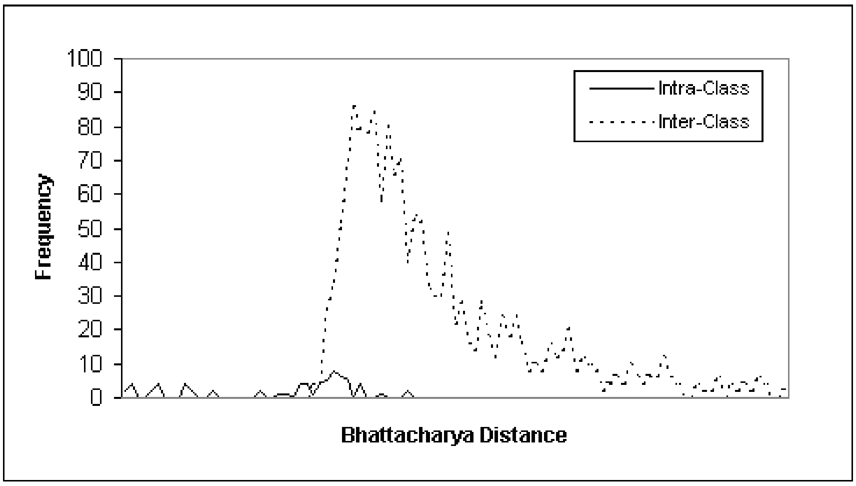

Figure 3: Distribution of Intra- and Inter-class distances for the synthetic images

ent number of nearest neighbours. The different bars are for different numbers of smoothing iterations. The plot shows results for the real-world data. Each image was matched against the entire database, with the exception of the image itself. This means that any correct matches were with different poses. There are a number of features which deserve comment. First, the optimal number of smoothing iterations is about 15. When fewer iterations are used then the roughness of the needle map limits performance. When more iterations are used then the curvature detail is oversmoothed. For this optimal number of smoothing iterations, nearly 50 percent of the first matches are correct. When it is considered that most of the rotations in the database exceed 30 degrees and that the majority of face recognition systems degrade rapidly up to rotations of 20 degrees this result is impressive.

A similar bar-chart is shown for the synthetic images in Figure 5. In this case, the numbers of nearest neighbours considered are 1, 5, 10 and 20. The results are poorer than obtained with the real world data. However some of the results are still encouraging. Taking the first 10 matches, 10 iterations yields nearly 70 percent correct matches. If the first 20 matches are considered, 10 iterations yields 90 percent correct matches. These results are quantitatively good.

One conclusion that certainly be drawn from these investigations is that the over smoothing of the needle map limits recognition performance. Since all faces are similar in shape, the small variations in surface detail are critical for recognition, but smoothing erodes them.

It is also important to note that rotations of over 30 degrees are quite severe and many face recognition systems would suffer dramatic performance degradation for such severe variation. For instance, the accuracy of Brunelli and Poggio's [5] template matching system dropped from 100 percent to 20 percent accuracy for a rotation of just 20 degrees. Additionally, the elastic bunch graph matching system of Wiskott et al [14] suffers performance degradation from 98 percent to 57 percent for half-rotated faces.

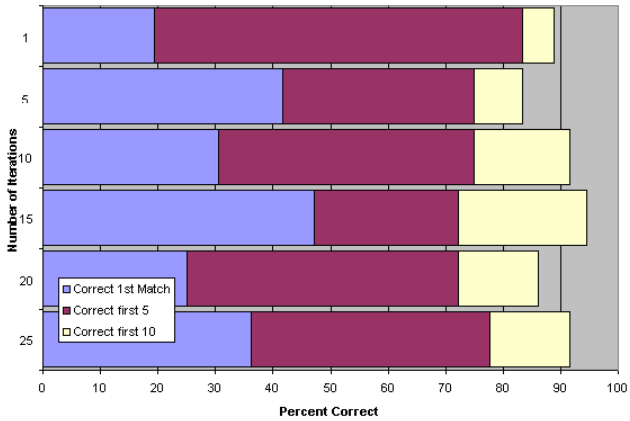

Figure 4: Best k-matches bar chart for the real world image database

5 Conclusions

In this paper, we have shown how the output of the Worthington and Hancock shape-from-shading scheme can be used for face recognition under variations in subject head pose. The recognition scheme uses histograms of mean and Gaussian curvature. Recognition is effected by comparing the comparing histograms using the Bhattahcharyya distance.

We demonstrate that the recognition results are senstive to the degree of smoothing applied to the needle-maps delivered by shape-from-shading. This is measured in terms of the number of smoothing iterations used. If too little smoothing is employed then "roughness' limits the recognition accuaracy. When too much smoothing is employed then curvature detail is eroded.

It is interesting that the performance results compare favourably with those obtained by Worthington and Hancock in the recognition of 3D objects from the COIL database. This is an important observation, since the COIL data-base contains objects of disimilar shape. Face recognition, on the other hand, is a problem of distinguishing between similarly shaped objects. Hence, that 1-NN classification rates as good as 50% can be achieved is promising. This recognition rate should also be qualified by noting that it is achieved under average differences of pose angle in excess of 30 degress.

There is clearly scope for further work in several directions. The first of these would be to investigate the use of orientation correction for the curvature attributes. Woodham has suggested a way in which this can be done, and has reported the advantages of the method in conjunction with photometric stereo. Second, histogram-based recognition is clearly a simplistic method. We are currently investigating the use of finer grained recognition stratagies.

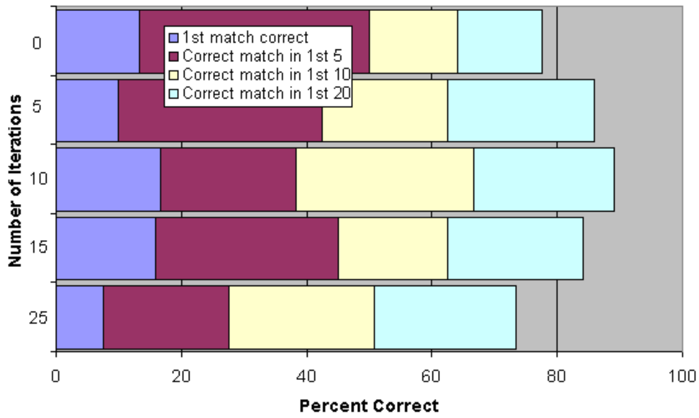

Figure 5: Best k-matches bar chart for the synthetic image database

References

[1] Joseph J. Atick, Paul A. Griffin, and A. Norman Redlich. Statistical approach to shape from shading: Reconstruction of three-dimensional face surfaces from single two-dimensional images. *Neural Computation*, 8(6):1321–1340, 1996.

[2] P. Belhumeur and D. Kriegman. What is the set of images of an object under all possible lighting conditions. *CVPR*, page pp. 270, 1996.

[3] Peter N. Belhumeur, Joao Hespanha, and David J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. In *ECCV (1)*, pages 45–58, 1996.

[4] Tarr M.J. Braje W.L., Kersten D. and Troje N.F. Illumination effects in face recognition. *Psychobiology*, (26):371–380, 1998.

[5] R. Brunelli and T. Poggio. Face recognition: Features versus templates. *IEEE Trans. PAMI*, 15(10):1042–1062, 1993.

[6] Chitra Dorai and Anil K. Jain. Shape spectrum based view grouping and matching of 3d free-form objects. *IEEE Trans. PAMI*, 19(10):1139–1146, 1997.

[7] P.N. Belhumeur D.W.Jacobs and R. Basri. Comparing images under variable illumination. In *Proc. Conf. on CVPR*, pages 610–617, 1998.

[8] A. Georghiades, D. Kriegman, and P. Belhumeur. Illumination cones for recognition under variable lighting: Faces. In *IEEE Conf. on CVPR*, 1998.

[9] G. Gimelfarb and A. Jain. On retrieving textured images from an image database. *Pattern Recognition*, 29(9):1461–1483, 1996.

[10] G.G. Gordon. Face recognition based on depth and curvature features. *Proc. IEEE Conf.*, 1992.

[11] B. Huet and E.R. Hancock. Line pattern retrieval using relational histograms. *IEEE Trans. PAMI*, 21(12):1363–1370, 1999.

[12] Enns J.T. and Shore D.I. Separate influences of orientation and lighting in the inverted-face effect. *Perception and Psychophysics*, (59):23–31, 1997.

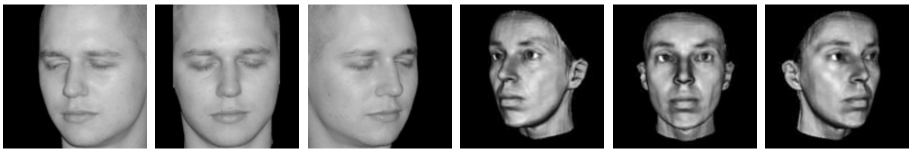

Figure 6: Two example image sets from the real and synthetic face image sets

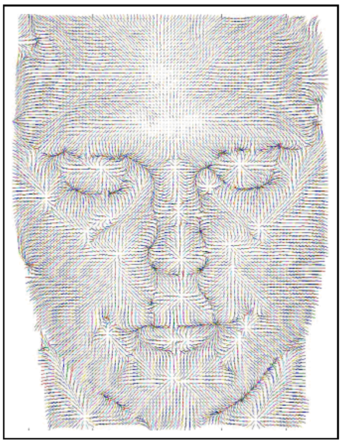

Figure 7: Needle map for real face image

[13] White P. Kemp R., Pike G. and Musselman A. Perception and recognition of normal and negative faces: The role of shape from shading and pigmentation cues. *Perception*, (25):37–52, 1996.

[14] N. Kruger C. von der Malsbnurg L. Wiskott, J. Fellous. Face recognition by elastic bunch graph matching. *IEEE Trans. PAMI*, 19(7), 1997.

[15] Troje N. and Bulthoff H. H. Face recognition under varying pose: The role of texture and shape. *Vision Research*, 36(12):1761–1771, 1996.

[16] Michael J. Swain and Dana H. Ballard. Color indexing. *International Journal of Computer Vision*, (7):11–32, 1991.

[17] P.L. Worthington and E.R. Hancock. New constraints on data-closeness and needle map consistency for shape-from-shading. *IEEE Trans. PAMI*, 21(12):1250–1267, 1999.

[18] P.L. Worthington and E.R. Hancock. Object recognition using shape-from-shading. *IEEE Trans. PAMI*, 23(5):535–542, 2001.

[19] Alan L. Yuille, Mario Ferraro, and Tony Zhang. Image warping for shape recovery and recognition. *Computer Vision and Image Understanding: CVIU*, 72(3):351–359, 1998.

[20] W. Zhao and R. Chellappa. Robust face recognition using symmetric shape-from-shading. Technical Report CARTR -919, Center for Automation Research, University of Maryland, College Park, 1999.