There is Hope for Face Recogniation

Mahendra Pratap Panigrahy*
Neeraj Kumar**

Abstract

The temporal lobe of the brain (forward part of the brain) is partly responsible for our ability to recognize faces. Some neurons in the temporal lobe respond to particular features of faces. That particular lobe is also important for the processing of semantics in both speech and vision. The temporal lobe contains the hippocampus and plays a key role in the formation of long-term memory. Some people who suffer damage to the temporal lobe lose their ability to recognize and identify familiar faces. This disorder is called prosopagnosia. It has been estimated that at least 2% of the general population suffer from this infliction and it isn't necessarily because of an injury to the right occipitotemporal cortex. The occipital lobe is the visual processing center of the human brain and is at the back of the brain. During the recent years due to its many applications in different fields such as law enforcement, security scenarios or video indexing. Face recognition is a very challenging problem and up to date, there is no evident technique that provides a proper solution to all situations and different applications that face recognition may encounter. The context of this paper will address the question: is there any hope for face recognition? In a general context, I believe that face recognition in complex scenarios will remain unsolved for the next years. However there might be hope for specific contexts and applications if some techniques are further studied developed and combined.

Keywords: Hippocampus, Prosopagnosia, Occipitotemporal Cortex

1. Introduction

Designing a system for automatic image content recognition is a non-trivial task that has been studied for a variety of applications. Computer recognition of specific objects in digital images has been put to use in manufacturing industries, intelligence and surveillance, and image database cataloging to name a few. In this project, a prototype algorithm for automating the detection of human faces in digital photographs was developed and can serve as an introduction for future work in detecting people in images.

Prosopagnosia has long been known to occur as a result of stroke or other brain damage. But scientists now recognize that it can occur even in the absence of any disease or injury. This congenital or developmental form of face blindness can run in families, and occurs with varying degrees of severity in an estimated 2 per cent of the population. Thus, in Canada alone there may be over 650,000 individuals who are seriously impaired in their ability to tell one face from another. And many more may find remembering faces a challenge, even if the problems they experience are not as severe.

Recent research suggests that this more holistic processing –

*Department of ECE, Institute of Technology Roorkee (ITR), Roorkee, India
**Department of CSE, Institute of Technology Roorkee (ITR), Roorkee, India

seeing the forest instead of a bunch of trees, so to speak - is particularly important when considering individual differences in face-recognition abilities. A growing number of studies have revealed that face blindness is associated with a deficit in holistic processing. And superior face perception and recognition in a more general population has also been shown to be correlated with how strongly an individual displays holistic face processing.

Fortunately there may be hope for those who struggle to figure out if they've met someone before. Researchers have used computerized training over the course of several days to enhance individuals' attention to the overall configuration of facial features. Results suggest that such training in individuals with congenital face blindness can subsequently improve face perception, and enhance the co-ordination of activity between the occipital and fusiform areas.

2. Face Detection using Eigenfaces

The information theory approach of encoding and decoding face images extracts the relevant information in a face image, encode it as efficiently as possible and compare it with database of similarly encoded faces. The encoding is done using features which may be different or independent than the distinctly perceived features like eyes, ears, nose, lips, and hair.

Mathematically, *Principal Component Analysis (PCA)* approach will treat every image of the training set as a vector in a very high dimensional space. The eigenvectors of the covariance matrix of these vectors would incorporate the variation amongst the face images. Now each image in the training set would have its contribution to the eigenvectors (variations). This can be displayed as an *'eigenface'* representing its contribution in the variation between the images. These eigenfaces look like ghostly images and some of them are shown in figure 2.In each eigenface some sort of facial variation can be seen which deviates from the original image.

The high dimensional space with all the eigenfaces is called the image space (feature space). Also, each image is actually a linear combination of the eigenfaces. The amount of overall variation that one eigenface counts for, is actually known by the eigenvalue associated with the corresponding eigenvector. If the eigenface with small eigenvalues are neglected, then an image can be a linear combination of reduced no of these eigenfaces. For example, if there are M images in the training set, we would get M eigenfaces. Out of these, only M' eigenfaces are selected such that they are associated with the largest eigenvalues. These would span the M'-dimensional subspace 'face space' out of all the possible images (image space).

When the face image to be recognized (known or unknown), is projected on this face space Figure-1, we get the weights associated with the eigenfaces, that linearly approximate the face or can be used to reconstruct the face. Now these weights are compared with the weights of the known face images so that it can be recognized as a known face in used in the training set. In simpler words, the Euclidean distance between the image projection and known projections is calculated; the face image is then classified as one of the faces with minimum Euclidean distance.

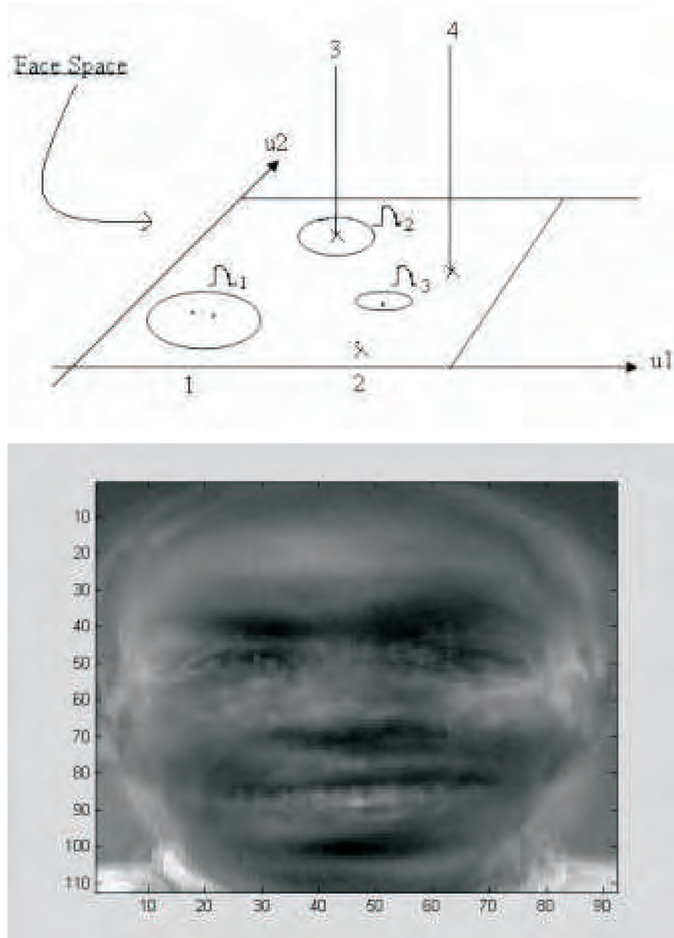

Fig. 1:
- **a. The face space and the three projected images on it. Here u1 and u2 are the eigenfaces**
- **b. The projected face from the training database**

2.1 Mathematical Formulations for Eigenfaces

Let a face image Γ (x, y) be a two-dimensional *N* by *N* array of intensity values. An image may also be considered as a vector of dimension N^2, so that a typical image of size 256 x 256 becomes a vector of dimension 65,536, or equivalently, a point in 65,536-dimensional space. An ensemble of images, then, maps to a collection of points in this huge space.

Images of faces, being similar in overall configuration, will not be randomly distributed in this huge image space and thus can be described by a relatively low dimensional subspace. The main idea of the principal component analysis is to find the vector that best account for the distribution of face images within the entire image space. These vectors define the subspace of face images, which we call "face space". Each vector is of length N^2, describes an *N* by *N* image, and is a linear combination of the original face images. Because these vectors are the eigenvectors of the covariance matrix corresponding to the original face images, and because they are face-like in appearance, they are referred to as "eigenfaces".

Let the training set of face images be $\Gamma_1, \Gamma_2, \Gamma_3, \ldots, \Gamma_M$. The average face of the set if defined by $\Psi = \frac{1}{M}\sum_{n=1}^{M}\Gamma_n$. Each face differs from the average by the vector $\Phi_n = \Gamma_n - \Psi$. An example training set is shown in Figure 1a, with the average face Ψ shown in Figure 1b. This set of very large vectors is then subject to principal component analysis, which seeks a set of M orthonormal vectors, μ_n which best describes the distribution of the data. The kth vector, μ_k is chosen such that

$$\lambda_k = \frac{1}{M}\sum_{n=1}^{M}(\mu_k^T\Phi_n)^2 \quad (1)$$

is a maximum, subject to

$$\mu_l^T\mu_k = \begin{cases} 1, l = k \\ 0, otherwise \end{cases} \quad (2)$$

The vectors μ_k and scalars λ_k are the eigenvectors and eigenvalues, respectively, of the covariance matrix

$$C = \frac{1}{M}\sum_{n=1}^{M}\Phi_n\Phi_n^T = AA^T \quad (3)$$

where the matrix $A = [\Phi_1\Phi_2...\Phi_M]$. The matrix C, however, is N^2 by N^2, and determining the N^2 eigenvectors and eigenvalues is an intractable task for typical image sizes. A computationally feasible method is needed to find these eigenvectors.

If the number of data points in the image space is less than the dimension of the space ($M < N^2$), there will be only $M-1$, rather than N^2, meaningful eigenvectors (the remaining eigenvectors will have associated eigenvalues of zero). Fortunately, we can solve for the N^2-dimensional eigenvectors in this case by first solving for the eigenvectors of and M by M matrix-e.g., solving a 16 x 16 matrix rather than a 16,384 x 16,384 matrix-and then taking appropriate linear combinations of the face images Φ_n. Consider the eigenvectors v_n of A^TA such that

$$A^T A v_n = \lambda_n v_n \quad (4)$$

Premultiplying both sides by A, we have

$$AA^T A v_n = \lambda_n A v_n \quad (5)$$

from which we see that are the eigenvectors of $C = AA^T$.

Following this analysis, we construct the M by M matrix $L = A^TA$, where $L_{mn} = \Phi_m^T\Phi_n$, and find the M eigenvectors v_n of L. These vectors determine linear combinations of the M training set face images to form the eigenfaces μ_n:

$$\mu_n = \sum_{k=1}^{M} v_{nk}\Phi_k = A v_n, n = 1,......, M \quad (6)$$

With this analysis the calculations are greatly reduced, from the order of the number of pixels in the images (N^2) to the order of the number of images in the training set (M). In practice, the training set of face images will be relatively small ($M < N^2$), and the calculations become quite manageable. The associated eigenvalues allow us to rank the eigenvectors according to their usefulness in characterizing the variation among the images.

2.2 Face Space

As the accurate reconstruction of the face is not required, we can now reduce the dimensionality to M′ instead of M. This is done by selecting the M′ eigenfaces which have the largest associated eigenvalues. These eigenfaces now span a M′-dimensional subspace instead of N^2.

2.3 Recognition

A new image T is transformed into its eigenface components (projected into 'face space') by a simple operation,

$$w_k = u_k^T(T - \psi) \quad (7)$$

here k = 1,2,….M′. The weights obtained as above form a vector $\Omega^T = [w_1, w_2, w_3, ... w_{M'}]$ that describes the contribution of each eigenface in representing the input face image. The vector may then be used in a standard pattern recognition algorithm to find out which of a number of predefined face class, if any, best describes the face. The face class can be calculated by averaging the weight vectors for the images of one individual. The face classes to be made depend on the classification to be made like a face class can be made of all the images where subject has the spectacles. With this face class, classification can be made if the subject has spectacles or not. The Euclidean distance of the weight vector of the new image from the face class weight vector can be calculated as follows,

$$\varepsilon_k = || \Omega - \Omega_k || \quad (8)$$

where Ω_k is a vector describing the kth face class. Euclidean distance formula can be found in [2]. The face is classified as belonging to class k when the distance ε_k is below some threshold value $\theta\varepsilon$. Otherwise the face is classified as unknown. Also it can be found whether an image is a face image or not by simply finding the squared distance between the mean adjusted input image and its projection onto the face space.

$$\varepsilon_2 = || \Phi - \Phi_f || \quad (9)$$

where Φ_f is the face space and $\Phi = T_i - \Psi$ is the mean adjusted input.

With this we can classify the image as known face image, unknown face image and not a face image.

2.4 Recognition Experiments

32 images were trained with four individuals having 8 images per individual. The 8 images had different lighting conditions, orientations and scaling. These images were recognized successfully with the accuracy of 100% for lighting variations, 90% for orientation variations, and 65% for size variations. The lighting conditions don't have any effect of the recognition because the correlation over the image doesn't change. The orientation conditions would affect more because of the image would have more hair into it than it had for training. Scaling affects the recognition significantly because the overall face data in the image changes. This is because the background is not subtracted for training. This effect can be minimized by background subtraction.

Fig. 2: The first row is some of the images used for training while the second row shows the eigenfaces with significant eigenvectors

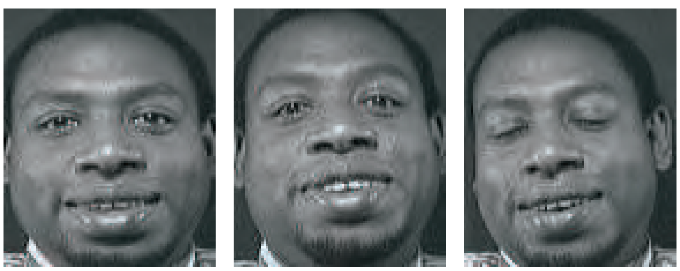

Fig. 3: The first image are the frontal image, whiles the second and third are the images of the same subject with different scaling and orientation

3. Face Recognition Using Line Edge Map

This algorithm describes a new technique based on line edge maps (LEM) to accomplish face recognition. In addition, it proposes a line matching technique to make this task possible. In opposition with other algorithms, LEM uses physiologic features from human faces to solve the problem; it mainly uses mouth, nose and eyes as the most characteristic ones. In order to measure the similarity of human faces the face images are firstly converted into gray-level pictures. The images are encoded into binary edge maps using Sobel edge detection algorithm. This system is very similar to the way human beings perceive other people faces as it was stated in many psychological studies. The main advantage of line edge maps is the low sensitiveness to illumination changes, because it is an intermediate-level image representation derived from low-level edge map representation. The algorithm has another important improvement, it is the low memory requirements because the kind of data used.

4. Conclusion

An eigenfaces-based face recognition approach was implemented in MatLab. This method represents a face by projecting original images onto a low-dimensional linear subspace-'face space', defined by eigenfaces. A new face is compared to known face classes by computing the distance between their projections onto face space. This approach was tested on a number of face images downloaded from airily good recognition results were obtained.

One of the major advantages of eigenfaces recognition approach is the ease of implementation. Futhermore, no knowledge of geometry or specific feature of the face is required; and only a small amount of work is needed regarding preprocessing for any type of face images.

However, a few limitations are demonstrated as well. First, the algorithm is sensitive to head scale. Second, it is applicable only to front views. Third, as is addressed in and many other face recognition related literatures, it demonstrates good performance only under controlled background, and may fail in natural scenes.

5. References

1. *Matthew A. Turk, Alex P. Pentland, "Face Recognition Using Eigenfaces," Proc. IEEE Conference on Computer Vision and Pattern Recognition: 586–591. 1991.*
2. *"Face recognition for smart environments", A. Pentland and T. Choudhury, Computer, Vol.33 Iss.2, Feb. 2000.*
3. *"Automatic recognition and analysis of human faces and facial expressions: A survey", A. Samal and P. A. Iyengar, Pattern Recognition, 25(1): 65-77, 1992.*
4. *L.I.Smith. "A tutorial on principal component analysis". Feb 2002.*
5. *"Low dimensional procedure for the characterization of human faces", Sirovich, L. and Kirby, M, Journal of the Optical Society of America A, 4(3), 519-524.*
6. *Ryan Johnson, Kevin Bonsor, "How Facial Recognition Systems Work," How Stuff Works, 2007.*
7. *L.C. Jain, "Intelligent biometric techniques in fingerprint and face recognition", Boca Raton: CRC Press, 1999.*
8. *Aurélio Campilho, Mohamed Kamel, "Image analysis and recognition: International Conference, ICIAR 2004, Porto, Portugal, September 29-October 1, 2004". Berlin; New York: Springer, 2004.*
9. *Yongsheng Gao; Leung, M.K.H., "Face recognition using line edge map", Pattern Analysis and Machine Intelligence, IEEE Transactions on, Volume: 24 Issue: 6, June 2002, Page(s): 764 - 779.*
10. *W. Zhao, R. Chellappa, A. Rosenfeld, and J. Phillips, "Face Recognition: A Literature Survey". ACM Computing Surveys, Vol. 35, No. 4, December 2003, pp. 399–458.*
11. *Face recognition home page: http://www.face-rec.org/.*
12. *Keren, D., Osadchy, M., Gotsman, C.: Antifaces: A novel fast method for image detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(7) (2001) 747–761.*
13. *Dollar, P., Tu, Z., Tao, H., Belongie, S.: Feature mining for image classification. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. (2007).*
14. *Amit, Y., Geman, D.: A computational model for visual selection. Neural Computation 11(7) (1999) 1691–1715.*