

Summary: "Eigenfaces for Recognition" (M. Turk, A. Pentland)

Ed Lawson

1. Introduction

"Eigenfaces for Recognition" seeks to implement a system capable of efficient, simple, and accurate face recognition in a constrained environment (such as a household or an office). The system does not depend on 3-D models or intuitive knowledge of the structure of the face (eyes, nose, mouth). Classification is instead performed using a linear combination of characteristic features (eigenfaces).

Previous works cited by Turk et al. fall into three major categories: feature based face recognition, connectionist based face recognition and geometric face recognition. Feature based recognition uses the position, size and relationship of facial features (eyes, nose, mouth) to perform face recognition. The connectionist approach recognizes faces using a general 2-D pattern of the face. Geometric recognition models the 3-D image of the face for recognition.

2. Eigenfaces

The motivation behind Eigenfaces is that the previous work ignores the question of which features are important for classification, and which are not. Eigenfaces seeks to answer this by using principal component analysis of the images of the faces. This analysis reduces the dimensionality of the training set, leaving only those features that are critical for face recognition.

The system is initialized by first acquiring the training set (ideally a number of examples of each subject with varied lighting and expression). Eigenvectors and eigenvalues are computed on the covariance matrix of the training images. The M highest eigenvectors are kept. Finally, the known individuals are projected into the face space, and their weights are stored. This process is repeated as necessary.

Figure 1a shows 16 faces used for training. These images are 8-bit intensity values (0 to 255) of a 256 by 256 image. These images are converted from 256 by 256 images into a single dimension vector of size 65,536. This conversion is necessary because we need a 2-D square matrix to compute eigenvectors.

The mean of the training images ($\Gamma_1, \Gamma_2, \dots, \Gamma_M$) is the "average face" $\Psi = \frac{1}{M} \sum_{n=1}^M \Gamma_n$. Each training image differs from the average face by $\Phi_i = \Gamma_i - \Psi$.

The vectors u_k and scalars λ_k are the eigenvectors and eigenvalues of the covariance matrix of the face images Φ_i .

$$C = \frac{1}{M} \sum_{n=1}^M \phi_n^T \phi_n = AA^T$$

where $A = [\Phi_1, \Phi_2, \dots, \Phi_M]$.

The eigenvalues, λ are selected such that

$$\lambda_k = \frac{1}{M} \sum_{n=1}^M (u_k^T \phi_n)^2$$

is maximum, where $u_l^T u_k = \delta_{lk} = 1$ if $l = k$, 0 otherwise. These formulas attempt to capture the source of the variance, which will later be used for classification purposes.

The vectors u_k are referred to as the "eigenfaces", since they are eigenvectors and appear face-like in appearance. The major difficulty with this is that this computation will produce a large number of eigenvectors (N^2 by N^2). Since M is far less than N^2 , we expect at most $M - 1$ useful (nonzero) eigenvectors. A smaller matrix $L = A^T A$ will yield this smaller number of eigenvectors.

The eigenfaces are $u_t = \sum_{k=1}^M v_{tk} \phi_k$.

Figure 2 shows 7 eigenfaces from the training set.

3. Recognition using Eigenfaces

The new image Γ is projected into the face space using $\omega_k = u_k^T(\Gamma - \Psi)$. The weights form a vector $\Omega^T = [\omega_1, \omega_2, \dots, \omega_M]$.

The euclidean distance : $\epsilon_k^2 = \|\Omega - \Omega_k\|^2$ measures the distance between the new image and a class of faces k. Note that if there are more than one examples of the face, these weights are averaged among all of the examples. If the distance measure, ϵ_k is less than a threshold Θ_ϵ , the face is assigned to recognized, and assigned to class k. This threshold is assigned empirically.

The distance function previously assumed that new image Γ is a face. To determine the validity of this assumption, project the image onto the face space, and examine difference between the projected image and Γ . The image is projected by computing $\Phi = \Gamma - \Psi$ and projecting onto $\Phi_f = \sum_{i=1}^{M'} \omega_i u_i$. The distance $\epsilon^2 = \|\Phi - \Phi_f\|^2$ determines the distance between the face space.

There are one of four possibilities.

1. $\epsilon_k < \Theta_\epsilon$ and $\epsilon < \Theta_\epsilon$. In this case, the image is close to the face space, and close to a known class. It is classified as belonging to that class.
2. $\epsilon_k < \Theta_\epsilon$ and $\epsilon \geq \Theta_\epsilon$. This case is most likely noise and should be rejected.
3. $\epsilon_k \geq \Theta_\epsilon$ and $\epsilon < \Theta_\epsilon$. In this case, the image is close to the face space, but not close to a known class. This is a face, but the classification is unknown. Store this for future assignment.
4. $\epsilon_k \geq \Theta_\epsilon$ and $\epsilon \geq \Theta_\epsilon$. This is not a face image.

4. Location and Detection

The previous sections assume a centered face image that is the same size as the faces in the training images. The system cannot operate correctly if the faces are not in the same location with approximately the same size.

Turk et al. use frame differencing to track motion against a static background. The filtered image is thresholded to produce a number of motion blobs. Heuristics such as the small blob of the large blob is a face and the face must all move contiguously can be used to determine the location of the head. The size of this 'blob' can be used to scale estimate the size of the face, to perhaps scale it to fit the size of the faces in the face space. Alternatively the size can be used in a multi-scale eigenfaces approach.

If there are a number of 'blobs' to select from, the face can be located by examining a fixed sized subregion and determining the 'faceness' of that region. If the distance ϵ between the region and the face space can be used to analyze the image for faces. However, this type of brute force method can be quite time consuming, and may not be practical for any real time application. The authors present a derivation to suggest how the faceness might be computed in a less computationally expensive way, by calculating some of the fixed terms ahead of time.

5. Other Issues

A significantly different background will adversely affect recognition, as the algorithm cannot distinguish between face and background. To reduce this problem, the authors use a 2 dimensional Gaussian centered at the face to reduce the background intensity.

The size of the face may also play a major role in the recognition rate. The tracking algorithm gives an idea of the face size, but may not always be correct. The solution may be to train using different scale faces, then use the eigenfaces at various scale to estimate size.

Head tilt will also affect the recognition rate, face symmetry measures are used to determine tilt and to rotate to a standardized orientation.

6. Experiments

Experiments were conducted using recognition under various lighting, scale and orientation. The experimental database is over 2500 face images of the 16 subjects using all combination of 3 head orientations, 3 head sizes, and 3 lighting conditions. A six level Gaussian pyramid was constructed resulting in resolutions from 512 x 512 to 16 x 16.

The first experiment sets an infinite threshold. The system is trained using 1 image per class at a fixed orientation, size and lighting. All of the remaining images are assigned to different classes. This experiment shows 96% recognition over varied lighting, 85% over varied orientation and 64% over varied size.

The next experiment varied the value for the accept / reject threshold Θ_ϵ . While the experiments showed that it is possible to select a threshold such that the recognition rate is 100%, the number of 'unknown' images can be large, depending on lighting, orientation and size variations. Adjusting the threshold to achieve this resulted in 19% unknown for lighting, 39% unknown for orientation and 60% for size.

References

- [1] M. Turk, A. Pentland. "Eigenfaces for Recognition". Journal of Cognitive Neuroscience. Vol 3, No. 1. 71-86, 1991.