A SURVEY OF FACE RECOGNITION TECHNIQUES

[1]*Omidiora E.O., [2]Fakolujo O.A., [3]Ayeni R.O. and T.M. Ajila[1]

[2]Department of Electrical & Electronics Engineering,
University of Ibadan, Nigeria
[1]Department of Computer Science & Engineering
[3]Department of Pure & Applied Mathematics
Ladoke Akintola University of Technology, Ogbomoso, Nigeria

ABSTRACT

A review of face recognition techniques has been carried out. Face recognition has been an attractive field in the society of both biological and computer vision of research. It exhibits the characteristics of being natural and low-intrusive. In this paper, an updated survey of techniques for face recognition is made. Methods of face recognition, such as, geometric, statistical and neural networks approaches are presented and analyzed. The comparative performance of the various approaches is discussed.

Keywords: Face recognition, Multilayer Perceptron, Statistical Approaches, Neural Network Approaches, Geometric Approaches, Biometric methods

*Corresponding Author

INTRODUCTION

Face recognition has aroused the interest of researchers from security, psychology and image processing to computer vision fields. it is one of the biometric techniques that allows or would grant access by "who we are" and not by conventional methods that allow access by "what we have" such as ID cards keys, passwords, PIN nos. Biometric systems are systems that identify or verify human beings. Also, it has the merits of both high accuracy and low intrusiveness. Different biometrics being investigated include fingerprints [1], speech [2], signature dynamics [3] and face recognition [4] and so on. Among the various biometric ID methods enumerated above, the physiological methods (face and fingerprint) are more stable than the behavioral methods (signature dynamics and speech). The reason being the non-alterable nature (except by injury) of the physiological features but the behavioral methods has the advantage of being non-intrusive [5,6]. As a result of these good attributes, biometric system is very difficult to forge since they exhibit biological characteristics to identify.

Face recognition is useful in finding a person within a large database of faces e.g. in a police, school or library database. These systems usually return a list of the most likely people in the database.

- It is also useful in identifying particular people in real-time for example a security monitoring system, location tracking system etc.
- It is used in allowing access to a group of people and denying access to all others for example access to a building, computer etc.
- It creates an easier way of making information access faster in a large database.

REVIEW OF RELATED WORKS

Faces represent complex, multidimensional, meaningful visual stimuli and developing a computer model for face recognition is difficult [6]. Face recognition is one of the few biometric methods that possess the merits of both high accuracy and low intrusiveness. It also has the accuracy of a physiological approach without being intrusive. For this reason, since the early 70's [37], face recognition has drawn the attention of the researchers in fields from security, psychology and image processing to computer vision. Psychologists and neuroscientists have studied issues such as uniqueness of face, how infants perceive faces and organization of memory of faces while engineering scientists have designed and developed face recognition algorithms [4,8-25].

A general statement of the problem of face recognition can be given as follows: To identify one or more persons in a given still or video images of a scene using a stored database of faces.

The solution to the problem involves segmentation of face from cluttered scenes, extraction of features from face region, identification and matching. Face Recognition problems and techniques can be separated into two groups: dynamic (video) and static matching [9]. *Dynamic matching* is used when a Video sequence is available. The video images tend to be of low quality, the background is cluttered and often more than one face present in the picture. However, since a video sequence is available, one could use a motion as a strong cue for segmenting faces of moving persons. On the other hand, *static matching* uses image with typically reasonably controlled illumination, background, resolution and distance between camera and the person. Some of the images that arisė in this group can be acquired from a video camera.

Automatic face recognition by computer can be categorized into two approaches [4,10] namely, constituent based and face-based. In *constituent–based* approach, recognition is based on the relationship between human facial features such as eyes, mouth, nose and face boundary [11-12,26-27] while *face-based* approach [13-15,26,28] attempts to capture and define the face as a whole. In this approach, face is matched through identifying its underlying statistical regularities.

Sirovich and Kirby [16] first proposed Karhunen-loeve (KL) transform to represent human faces. In their method, faces are represented by a linear combination of weighted eigenvector, known as Eigenfaces.

The method proposed by Brunelli and Poggio [29] used a set of templates to detect the eye positions in a new image, by looking for the maximum absolute values of the normalized correlation of these templates at each point in the test image. Turk and Pentland [13] also developed a face recognition using Principal Component Analysis (PCA).

TECHNIQUES FOR SOLVING FACE RECOGNITION PROBLEMS

Based on the feature extraction and classification techniques used, the face recognition approaches are divided into three:

- Statistical Approaches
- Neural Network Approaches
- Geometric Approaches

STATISTICAL APPROACHES

In this section we focus on the statistical approaches to face recognition and more specifically on:

- Correlation Method
- Karhunen-loeve expansion based methods

Correlation Method

The most direct of the procedures used for face recognition is the matching between the test images and a set of training images based on measuring the correlation. The matching technique in this case is based on the computation of the normalized cross-correlation coefficient C_N, defined by [1]:

$$C_N = \frac{E\{I_T T\} - E\{I_T\}E\{T\}}{\sigma\{I_T\}\sigma\{T\}} \quad (1)$$

Where I_T is the image, which must be matched to the template T

$I_T T$ is the pixel-by-pixel product

E is the average operator

σ is the standard deviation over the area being matched.

This normalization rescales the templates and image energy distribution so that their average and variances match. However, correlation based methods are very dependent on illumination, rotation and scale. For reduction of the illumination variations, the intensity of the gradient $(|\delta_x I_T| + |\delta_y I_T|)$ was being used.

Correlation based recognition is possible at a good performance level using templates as small as 36 x 36 pixels.

Recognition using Correlation Methods

Brunelli and Poggio [29] described a correlation-based method for face recognition from frontal views. Their method is based on the matching of templates corresponding to facial features of relevant significance as the eyes, nose and mouth. The positions of these features are first detected so as to reduce the complexity of this approach.

Feature Extraction: After the facial features are detected, a set of template corresponding to these features in the test image is compared, in turn, with the corresponding features of all the images in the database, returning a vector of matching scores (one per feature) computed through normalized cross-correlation.

Classification: The similarity scores of different features can be integrated to obtain a global score.

The cumulative score an be computed in several ways:

- Choose the score of the most similar feature
- Sum the feature score
- Sum the feature scores, using constant weights
- Sum the features scores using person-dependent weights.

After computing the cumulative matching scores, a test face is assigned to the face class for which this score is maximized.

Recognition Performance: The recognition rate reported in [29] using correlation method for frontal faces is higher than 96%. The correlation method as described in this section requires a robust feature detection algorithm with respect to variations in scale, illumination and rotations in image plane and image depth.

Recognition under General Viewing Condition

Beymer [30] extended the correlation based approach to a view-based approach for recognizing faces under varying orientations, including rotations with axis perpendicular to the image plane (rotations in image depth). In the first stage, of this method, a pose estimation module detects the position of two facial features (the eyes and the nose) to determine the face pose. These features were being detected by looking for the maximum absolute values of the normalized correlation coefficients of the model templates at each point in the current image. After the pose has been determined, the task of recognition is reduced to the classical correlation method discussed above in which the facial feature templates are matched to the corresponding templates of the appropriate view-based models using the cross-correlation coefficient.

In this case, the computation complexity increases with the number of model views for each person in the database.

Recognition Performance: The recognition performances of the system were evaluated on a small database of 620 test images representing 62people. A recognition rate of 98.7% has been reported [30]. However, the method is computationally complex requiring 10-15mins to perform the classification of one face image.

Karhunen-Loeve Expansion-Based Methods
Recognition using Eigenfaces

As mentioned, one of the goals that the feature extraction routine wishes to achieve is to increase the efficiency. One simple way to achieve this goal is using alternative orthonormal bases other than the natural bases. One such basis is the Karhunen-Loeve Expansion (KL).

The "Eigenfaces" method proposed by Turk and Pentland [13,31] is based on the karhunen-loeve expansion and is motivated by the earlier work of Sirovitch and Kirby [16,28] for efficiently representing picture of faces. The eigenface method finds the Principal Components (KL) of the covariance matrix of the set of face images. These eigenvectors can be though as a set of features, which together characterize the variation between face images.

Let a face images I (x,y) be a vector of dimension n.

Let the training set of images be $I_1, I_2, ---, I_N$. The average face image of the set is defined by:

$$\Psi = \frac{1}{N} \sum_{i=1}^{N} Ii$$

Each face differ from the average by the vector

$$\Phi = \frac{1}{N} - \Psi$$

This set of very large vectors is subject to Principal Component Analysis which seeks a set of K orhonomal vectors V_K, K=1,2,---,K and their associated eigenvalues λ_k which best described the distribution of data.

The vectors V_K and scalars λ_k are the eigenvectors and eigenvaluesof the covariance matrix:

$$C = \frac{1}{N} \sum_{i=1}^{N} \Phi \Phi i = AA^T \quad (2)$$

Where the matrix

$A = [\Phi 1, \Phi 2, - - -, \Phi]$. Finding the eigenvectors of matrix $C_{n \times n}$ is computationally intensive. However, the eigenvectors of C can be determined by first finding the eigenvectors of a much smaller matrix of size N X N and then taking a linear combination of the resulting vectors [14].

Feature Extraction: the space spanned by the eigenvectors V_K, K=1,2,---,K, corresponding to the largest K eigenvalues of the covariance matrix C is called the "face space". The eigenvectors of matrix C, which are called the eigenfaces form a basis set for the face images. A new face image

Γ is transformed into its eigenface components (projected into the space) by:

$$\omega = \langle V_K(\Gamma - \Phi) \rangle = V_K(\Gamma - \Phi) \qquad (3)$$

For K =1, 2, ---, k

The projection W_K forms the feature vector $\Omega = [W_1, W_2, ---, WK]$, which describes the contribution of each eigenface in representing the input image.

Classification: Given a set of face classes E_q and the corresponding feature vectors Ω_q, the simplest method for determining which face class provides the best description of an input face image Γ is to find the face class j that minimizes the Euclidean distance in the feature space:

$$\xi_q = \|\Omega - \Omega_q\| \qquad (4)$$

A face is classified as belonging to class E_q when the minimum E_q is below some threshold θ_ε

$$E_q = \text{argmin}_q \{\xi_q\} \qquad (5)$$

Otherwise, the face is classified as unknown.

Recognition Performance: The Eigenface method was tested on a large database of 2500 face images digitized under controlled conditions. Various groups of sixteen images corresponding to sixteen different subjects were selected and used as the training set. The recognition performance reported are 96% correct classification over lighting variations, 85% correct classification over orientation variation and 64% over size variation. It can be seen that the approach is fairly robust to changes in lighting conditions [5], but degrades quickly as the scale changes. The recognition takes about 350msec running on a sparcstation 1, using face images of size 128x128 pixels.

Recognition Under General Viewing Conditions

The Parametric Approach

H. Murase and S. Nayar [8] and S. Nayar et al [40] extended the capabilities of the eigenface method to general 3D object recognition under different illumination and viewing conditions. Given N object images taken under P views and L different illumination conditions, a universal image set is built which contains all the available data. In this way, a single "Parametric Space" describes the object identity as well as the viewing or illumination conditions. The eigenface decomposition of this space was used for feature extraction and classification.

The View-Based Approach

Due to the eigenface decomposition, Pentland and others [8] developed a "view-based" eigenspace approach for human face recognition under general viewing conditions. Given N individuals under Pdifferent views, recognition is performed over Pseparate eigenspaces, each capturing the variation of the individuals in a common view.

The "view-based" approach is essentially an extension of the eigenface technique to multiple sets of eigenvectors, one for each face orientation. To deal with multiple views, in the first stage of this approach, the orientation of the test face is determined and the eigenspace which best describes the input image is selected. This is achieved by calculating the residual description error (Distance From Feature Space: DFFS) for each view space.Once the proper view is determined, the image is projected onto the appropriate view space and then recognized.

Recognition Performance: The recognition performance of the view-based and parametric approaches was evaluated on a database of 189 images consisting of nine views of 21 people [8]. The nine views of each person were evenly spaced from $-90°$ to $+90°$. Performance of the algorithms was tested by training on a subset of the available views, $\pm 90°, \pm 45°, 0°$ and testing on the intermediate views $\pm 68°, \pm 23°$ (interpolation performances). The average recognition rate reported were 90% for the view-based and 88% for the parametric eigenspace methods. A second series of experiments tested the extrapolation performance by training on a range of views (e.g. 90^0 to $+45°$) and testing on novel views outside the training range (e.g. $+68°$ and $+90°$) for testing views separated by $\pm 23°$ from the training range the average recognition rates were 83% for the view-based and 78% for the parametric eigenspace method.

For $\pm 45°$ testing views, the average recognition rates were 50% for view-based and 43% for parametric.

NEURAL NETWORK APPROACH

An artificial neutral network (ANN) is a computation system inspired by the functioning of the human brain. The system is made up of the highly interconnected processing elements (artificial neurons) [9,32,40]. The artificial neuron is a simplified mathematical representation of the biological neuron, which executes the sum of its weighted inputs through the weights, associated to those inputs (synaptic weights) and then applies a function to that result in order to generate the output. The function, which is applied to the result, is called the activation function, and is normally a non-linear function, a feature that enables the ANN to represent more complex problems. An input layer, output layer, and one or more hidden layer form neural network architecture. ANN knowledge is stored in the synaptic weights of each link between two processors. The ANN modifies the initial weight values so as to assimilate the desired mapping between the input and the output by means of a learning algorithm.

Recognition Using Neural Network

Multilayer Perceptron (MLP) Neural network is a good tool for classification purposes [9,32,33,40]. It can approximate almost any regularity between its inputs and outputs. The ANN weights are adjusted by supervised training procedure called back propagation.

MLP contains three layers: the inputs unit which is equal to the number of pixels in the image hidden unit and output units which is equal to the number of persons to be recognized. Every output unit is associated with one person. The neural network [NN] is trained to respond "-1" on output unit corresponding to recognized person and "+1" on other outputs [34]. Using discrete cosines transform coefficients (DCT), the sample size was reduced and significantly speedup the training process. A gradient map allows achieving partial invariance to lightning conditions.

Starovoitov et al [34] made use of a threshold rule, which improved recognition performance by considering all outputs of NN. This threshold rule is called 'sqr' rule, which calculates the Euclidean distance between the perfect and real outputs for recognized person. When this distance is greater than the threshold, the person is being rejected otherwise accepted.

Recognition Performance: -The NN method was tested on a database of 70 images under controlled conditions. The recognition performance reported is 94% for correct classification [34]. The recognition takes about 34min 46sec using face images of size 92 x 112pixels, quantized to 256 gray levels.

GEOMETRIC APPROACH

The first historical way to recognize people was based on face geometry, which is to extract the features form facial images. The discrete features such as the eyes, mouth, nose and chin is the important cues for discrimination and recognition of faces.

For different facial contours, different models are used to extract them from the original portrait. Because one shape of eyes and mouth are similar to some geometric figures, they are being extracted on terms of the deformable template model [35]. The other facial feature such as eyebrows, nose and face are called active contour model [36,37].

Effective Feature selection

Features are the basic elements for object recognition. Because the variance of each feature associated with the face recognition process is relatively large, the features are classified into three major types [18,23]:

Firstly-Order features Values: Discrete features such as eyebrows, mouth and nose which, have been found to be important [38] in face identifies and are specified without reference to other facial features, are called first-order features.

Second-Order features Values: are set of features, which characterize the special relationships between the positions of the first order features and information about the shape of the face.

Higher Order feature Values: features whose values depend on a complex set of feature values. For instance, age might be a function of hair coverage, hair colour, presence of wrinkles etc.

Feature Extraction Using Deformable template: The deformable template are specified by a set of parameters which uses a priori knowledge about the expected shape of the features to guide the acts on three representations of the image, as well as on the image itself [18]. The first two representations are the peak and villages in the image intensity changes quickly.

The templates are flexible enough to change their size and other parameter values, so as to match themselves to the data. The final values of these parameters can be used to describe the feature.

Feature Extraction using Active contour mode (snake): The active contour or snake is an energy minimizing spline guided by external constraint forces and influenced by image force that pull it toward features such as lines and edges [18]. This approach differs from deformable approaches, which detect edges, and links them. In the active contour model, image and external forces together with the connectivity of the contours and the presence corners affect the energy function and the locally option contour.

Recognition Performance: In the [34], different features set were tested. The ORL face database was used with each photo being an image size of 92x112 pixels, and quantized to 256 gray levels. The method was tested on 70 images of 12persons and 28 features were selected. The recognition performance reported is 98.5% of correct classification.

COMPARISON OF METHODS

The recognition results of the approaches have been discussed in the previous session. The parameters of comparison among these methods are the recognition rates reported under variations in lightning and viewing conditions, facial expressions, no of classes used by the recognition system as well as computation complexity.

- The Correlation method performs with high accuracy if lightning and size normalization is applied, under variations in facial expression and pose.
- A more efficient approach to face recognition is the Eigenface method. Although the recognition performance is lower than the correlation method, the substantial reduction in computational complexity of the Eigenface method makes this method very attractive. The recognition rates increase with the number of Principal component used and in the limit as more principal component are used, performances approaches that of correlation.
- The recognition performances of the Eigenface and Correlation methods against large variations in illumination and pose were improved by using either a parametric approach or a view-based approach. The parametric approach has a reduced complexity but the view-based is more accurate.

The Neural Network and Geometric approaches perform with high accuracy and may be improved by utilizing any additional features. Using a corresponded face rotation and gesture geometric model may speed up the execution time are shown in Table 1

CONCLUSION

In this paper, a survey of comparison of methods of techniques for face recognition is given. Also, this paper shows that statistical, geometric and neural network approaches would perform well in any face recognition system with regards to multimedia information access, network security, content indexing and retrieval.

Most of the approaches have been considered, analyzed and quantitative percentage of accuracy of each of the methods has been given. Most of the methods under consideration have been found to be very accurate (greater than) under various viewing conditions. From the result of comparison in Table 1, one would easily conclude the wide acceptance of face recognition as one of the most important biometric methods in access or security monitoring control system.

Table 1:Comparison of Face Recognition Methods

Method	Training set	Testing set	Recognition results	Type of database	Complexity	Comparison with other methods
Correlation	Not specified	Not specified	Over 96%	Frontal faces, small variations in illumination, scale	High	NA
Correlation	62people 15images per person	62people 10images per person	98.7%	Strong rotations in depth, small variation in scale and illumination	10-15min on sparc2	NA
Eigenface	16	2500	96%	Lighting variations	350msec on sparc1	NA
			85%	Orientation variations		
			64%	Variations in size		
Eigenface parametric approach	128	7562	88%	Variations in pose (interpolation)	Higher than the	90%view-based approach
			78%	Variations in pose (extrapolation)		83%view-based approach
Eigenface view based approach	128	7562	90%	Variations in pose (interpolation)	Lower than the parametric approach	83% parametric approach
			83%	Variations in pose (extrapolation)		90% parametric approach
Multilayer Perceptron (MLP)	12people 70images per person		94%		34mins 46secs	NA
Geometric Approach	12peopl 70 images per person		98.5%		Complexity	NA

REFERENCES

Blue J., Candela. G., Grother, P., Chellappa, R. and Wilson, C. (1994), 'Evaluation of Pattern classifiers for fingerprint and OCR applications', *Pattern Recognition* 27(4), 485–501.

Bottou, L., Cortes, C., Denker, J., Drucker, H., Guyon, I., Jackel, L., Le Cun, Y., Muller, U.,Sackinger, E., Simard, P. and Vapnik, V. (1994), Comparison of classifier methods: A Case Study in handwritten digit recognition, *in* 'Proceedings of the International Conference on Pattern Recognition', IEEE Computer Society Press, Los Alamitos, CA

Qi, Y. and Hunt, B.(1994), "Signature verification using global and grid features", *Pattern Recognition* 27(12), 1621–1629.

Chellapa, R., Wilson, C. L., and Sirohey, S.(1995), "Human and Machine recognition of Faces: A Survey, Proceedings of IEEE, Vol.83, No 5, pp 705-740.

Lin, Shang-Hung (2000); "An Introduction to face Recognition Technologies-Part 2, Vol.3, No 1, pp 1 - 8.

Lawrence, S., Giles, L.C., Chung, T.A. and Andrew, D.B.(1997) "Face Recognition: A Convolutional Neural Network Approach" IEEE Transactions on Neural Network, Vol.8, **No 1 pp 98-132.**

Kelly M. D., "Visual Identification of People by Computer" (1976), Stanford AI Project, Stanford, CA, Techncal Report. Pp. 1-38.

H. Murase and S. Nayar (1995), "Visual Learning and Recognition of 3-Dimensional," International Journal of of Computer Vision, vol. 14, pp. 5-24.

Omidiora E. O., Fakolujo A. O., Ayeni R. O. and Olabiyisi S.O. (2005), "Generic Framework for Face Recognition and Detection, International Journal of Biological and Physical Sciences (Science Focus) (In Press)

Chow, G. and Li, X. (1993) "Towards a System for Automatic Facial Feature Detection, Pattern Recognition, Vol.26, No 12, 1739-1755.Lam. K.M. and Yan, H.(1996)

Locating and Extracting the eye in Human face images" Pattern Recognition Vol. 29, No 5, 771-779.

Valentin, D., Abdi. H..O'. Toole. A.J. and Cottrell, G.W. (1994)"Connectionist models of face processing: A survey, Pattern Recognition, Vol.27,1209-1230.

Turk M. and Pertland, A.(1991), "Eigenfaces for recognition" S. Cognitive Neuroscience, Vol.3, 71-86.

Swets, D.L. and Weng, J.J (1996) ""Using discriminant eigenfeatures for image retrieval" IEEE Trans. PAMI, Vol.18 No.8,836

Wickerhance, M.V. (1994), "Large-rank approximate component analysis with wavelets for signal feature discriminant and the inversion of complication maps"S. chemical Information and Computer Sciences. Vol. 34, No5, 1036-1046.

Sirovich, L. and Kirby, M. (1987), "Low-dimensional procedure for the charaterization of human-faces" J. OPT. Soc. Am. A, Vol. 4, No.3, 519-524.

Lin, S-H, Kung, S.Y. and Lin, L-J (1997), "Face Recognition/ Detection by Probabilistic Decision-Based Neural Network", IEEE Trans. Neural Networks, vol.8 no.1, pp.114-132.

Jia, X. and Nixon, M.S.(1990), "Extending the Feature Vector for Automatic Face Recognition", IEEE Trans. on PAMI, Vol. 17, No. 12,pp 1167-1176.

Kruizinga, P And Petkow, N(1995),"Person identification based on multiscale matching of cortical images",Proceeding of the In ternational Conference and exhibition on high- preference computing and networking, HPCN Europe'95,B Hertzberger and G. Srazzi Eds., 420-427.

Pentland, A.,Morgaddan, B and Starner (1994), T.,"View-Based and Modular Eigenspaces for Face Recognition", Proc. IEEE Conf. Computer Vision and Pattern Recognition, Seattle, pp 84-91.

Rowley, H. A.,Baluja, S. and Kanade, T. (1996) "Neural network Based Faced Detection", CVPR 19'96, pp 203-208.

Samal, A.(1996) "Minimum Resolution for Himan Face Detection and Identification" Proceedings of SPIE Human Vision, Visual Processing, and Digital Display, Vol. 1453, pp 203-208.

Sirovich, L. and SiroVich, C.(1989), "Low Dimensional Description of Complicated

Phenomena", Contemp. Mathematics, Vol 99, pp 277-305.

Tzanakou, E. M.,Uyeda, E.,Ray, R., Sharma, A.,Ramanujah, R. and Dong, J.,(1995) "Comparison of Neural Network Algorithms for Fcae Recgnition",Vol 64,pp 1,15 27.

Ilker Atalay (1996), "Face Recognition Using Eigenfaces", Unpublished M.Sc. Thesis, Department of Computer Engineering, Institute of Science and Technology, pp 24-

Goudail, F., Lange, E., Iwamoto, T. and Otsu, N. (1996)" Face Recognition System using Local Autocorrelations and Multiscale Integration", IEEE Trans. PAMI, Vol.18, No. 10 1024-1028.

Yuille, A.L., Hallinan, P.W and Colen, D.S,(1992) "Feature extraction from faces using deformable templates, Int.J. of Computer vission, Vol.8, No2, 99-111