A Hybrid Neural Network Approach for Face Recognition

[1]Ashwini Kanade, [2]Akhil Khare
[1, 2] Department Of Information Technology, Bharati Vidyapeeth Deemed University
College Of Engineering, Pune-46

ABSTRACT: *The face recognition algorithm that is presented here is a memory based face recognition system. The memory-based technique for view-based frontal face recognition can outperform more than sophisticated algorithms that use Principal Components Analysis (PCA) and neural networks. The goal of this report is to write about the most common methods that have been used till now for face recognition. Analyse these methods and give a general idea of the background of the algorithm, ARENA. The capability of the face recognition is to find the exact mach of a face image from an image database System. The algorithm that is used in order to achieve that is called AREN.*

1. INTRODUCTION

The objective of our system is to recognise and identify faces, not previously presented to or in some way processed by the system. There are many datasets involved in this system. Some of them are the ORL, MIT database which consisting of a large set of images of different people. The database has many variations in pose, scale, facial expression and details. Some of the images are used for training the system and some for testing. The test set is not involved in any part of training or configuration of the system, except for the weighted committees.

The algorithm used for the face recognition, known as ARENA. Similar to several other approaches to face recognition and identification, which use Principal Component Analysis (PCA) as pre-processing, dimensionality reduction and feature extraction, of the input images. One of the main parts of the system is a neural network. The use of a neural network makes the algorithm perform better.

The purpose of face recognition algorithm is to examine a set of images and try to find the exact match of a given image. An advanced system would be a neural network face recognition algorithm. The system examines small windows of the image in order to calculate the distances of given points. That would be done from any algorithm but in a system where the system use neural networks the system arbitrates between multiple networks in order to improve performance over a single network.

The goal of the system is to formulate paradigms for detection and recognition of human faces, and especially develop an algorithm, which is going to have high performance in complex backgrounds. One of the applications would be towards adding face-oriented queries to our image database.

The fundamental principle, which we are exploiting for our face recognition algorithm, is Principal Component Analysis. Thought the algorithm is much simpler. One of the aims is to run tests in order to compare the algorithm with two PCA algorithm and also show that the calculation between two given point with the ARENA algorithm is efficient.

1.1 Face recognition

Face recognition is a part of a wide area of pattern recognition technology. Recognition and especially face recognition covers a range of activities from many walks of life. Face recognition is something that humans are particularly good at and science and technology have brought many similar tasks to us. Face recognition in general and the recognition of moving people in natural scenes in particular, require a set of visual tasks to be performed robustly. That process includes mainly three-task acquisition, normalisation and recognition. By the term acquisition mean the detection and tracking of face-like image patches in a dynamic scene. Normalisation is the segmentation, alignment and normalisation of the face images, and finally recognition that is the representation and modelling of face images as identities, and the association of novel face images with known models.

Given the requirement for determining people's identity, the obvious question is what technology is best suited to supply this information? The are many ways that humans can identify each other, and so is for machines. There are many different identification technologies available, many of which have been in commercial use for years. The most common person verification and identification methods today are Password/PIN known as Personal Identification Number, systems. The problem with that or other similar techniques is that they are not unique, and is possible for somebody to forget loose or even have it stolen for somebody else. In order to overcome these problems there has developed considerable interest in "biometrics" identification systems, which use pattern recognition techniques to identify people using their characteristics. Some of those methods are fingerprints and retina and iris recognition. Though these techniques are not easy to use. For example in bank transactions and entry into secure areas, such technologies have the disadvantage that they are intrusive both physically and socially. The user must position the body relative to the sensor, and then pause for a second to declare himself or herself. That doesn't mean that face recognition doesn't need specific positioning.

While the pause and present interaction are useful in high-security, they are exactly the opposite of what is required when building a store that recognise its best customers, or an information kiosk that remembers you, or a house that knows the people who live there. Face recognition from video and voice recognition have a natural place in these next generation smart environments, they are unobtrusive, are usually passive, do not restrict user movement, and are now both low power and inexpensive. Perhaps most important, however, is that humans identify other people by their face and voice, therefore are likely to be comfortable with systems that use face and voice recognition.

2. FACE RECOGNITION

There are many algorithms that can be used for face recognition. Most of them are based on the same techniques and methods. Some of the most popular are Principal component analysis and the use of eigenfaces.

2.1 Principal Component Analysis

On the field of face recognition most of the common methods employ Principal Component Analysis. Principal Component Analysis is based on the Karhunen-Loeve (K-L), or Hostelling Transform, which is the optimal linear method for reducing redundancy, in the least mean squared reconstruction error sense. 1. PCA became popular for face recognition with the success of eigenfaces. The idea of principal component analysis is based on the identification of linear transformation of the co-ordinates of a system. "The three axes of the new co-ordinate system coincide with the directions of the three largest spreads of the point distributions." In the new co-ordinate system that we have now the data is uncorrected with the data we had in the first co-ordinate system. [2] For face recognition, given dataset of *N* training images, we create *N* d-dimensional vectors, where each pixel is a unique dimension. The principal components of this set of vectors is computed in order to obtain a *d x m* projection matrix, *W*. The image of the i^{th} vector may be represented as weights:

$$\vec{\theta}i = (\theta i1, \theta i2, \ldots, \theta im)^T \qquad (1)$$

Such that

$$\vec{xi} = \vec{\mu} + W\vec{\theta} \qquad (2)$$

Approximates the original image where μ is the mean, of the χ_i and the reconstruction is perfect when $m = d$. P1

As mentioned before the ARENA algorithm is going to be tested and its performance is going to be compared with other algorithms. For the comparison we are going to use two different PCA algorithms. The first algorithm is computing and storing the weight of vectors for each person's image in the training set, so the actual training data is not necessary. In the second algorithm each weight of each image is stored individually, is a memory-based algorithm. For that we need more storing space but the performance is better.

In order to implement the Principal component analysis in MATLAB we simply have to use the command *prepca*. The syntax of the command is

ptrans,transMat = prepca(P,min_frac)

Prepca pre-processes the network input training set by applying a principal component analysis. This analysis transforms the input data so that the elements of the input vector set will be uncorrected. In addition, the size of the input vectors may be reduced by retaining only those components, which contribute more than a specified fraction (min_frac) of the total variation in the data set. Prepca takes these inputs the matrix of centred input (column) vectors, the minimum fraction variance component to keep and as result returns the transformed data set and the transformation matrix.

A) Alorithm

Principal component analysis uses singular value decomposition to compute the principal components. A matrix whose rows consist of the eigenvectors of the input covariance matrix multiplies the input vectors. This produces transformed input vectors whose components are uncorrected and ordered according to the magnitude of their variance.

Those components, which contribute only a small amount to the total variance in the data set, are eliminated. It is assumed that the input data set has already been normalised so that it has a zero mean.

In our test we are going to use two different "versions' of PCA. In the first one the centroid of the weight vectors for each person's images in the training set is computed and stored. On the other hand in PCA-2 a memory based variant of PCA, each of the weight vectors in individually computed and stored.

B) Eigenfaces

Human face recognition is a very difficult and practical problem in the field of pattern recognition. On the foundation of the analysis of the present methods on human face recognition, a new technique of image feature extraction is presented. And combined with the artificial neural network, a new method on human face recognition is brought up. By extraction the sample pattern's algebraic feature, the human face image's eigenvalues, the neural network classifier is trained for recognition. The Kohonen network we adopted can adaptively modify its bottom up weights in the course of learning. Experimental results show that this method not only utilises the feature aspect of eigenvalues but also has the learning ability of neural network. It has better discriminate ability compared with the nearest classifier. The method this paper focused on has wide application area. The adaptive neural network classifier can be used in other tasks of pattern recognition.

In order to calculate the eigenfaces and eigenvalues in MATLAB we have to use the command eig. The syntax of the command is

d = eig(A)
V,D = eig(A)
V,D = eig(A,'nobalance')
d = eig(A,B)
V,D = eig(A,B)

d = eig(A) returns a vector of the eigenvalues of matrix A. *V,D = eig(A)* produces matrices of eigenvalues (D) and eigenvectors (V) of matrix A, so that A*V = V*D. Matrix D is the canonical form of A, a diagonal matrix with A's eigenvalues on the main diagonal. Matrix V is the modal matrix, its columns are the eigenvectors of A. The eigenvectors are scaled so that the norm of each is 1.0. Then we use *W,D = eig(A'); W = W'* in order to compute the left eigenvectors, which satisfy *W*A = D*W*.

V,D = eig(A,'nobalance') finds eigenvalues and eigenvectors without a preliminary balancing step. Ordinarily, balancing improves the conditioning of the input matrix, enabling more accurate computation of the eigenvectors and eigenvalues. However, if a matrix contains small elements that are really due to round-off error, balancing may scale them up to make them as significant as the other elements of the original matrix, leading to incorrect eigenvectors. .

d = eig(A,B) returns a vector containing the generalised eigenvalues, if A and B are square matrices. *V,D = eig(A,B)* produces a diagonal matrix D of generalised eigenvalues and a full matrix V whose columns are the corresponding eigenvectors so that *A*V = B*V*D*.

C) Euclidean distance

One of the ideas on which face recognition is based is the distance measures, between to points. The problem of finding the distance between two or more point of a set is defined as the Euclidean distance. The Euclidean distance is usually referred to the closest distance between two or more points. So we can define the Euclidean distance d_{ij} between points x x_{ik} and x_{jk} as :

$$d_{ij} = \sum_{k=1}^{p} (x_{ik} - x_{jk})^2 \quad (3)$$

3. ARENA ALGORITHM

As mentioned before in the introduction, the algorithm that is used in the System is called ARENA. Is a memory-based technique for view-based frontal face recognition that can outperform more sophisticated algorithms that use Principal Components Analysis and neural networks. This method does not perform any complex feature extraction, nor does it incorporate any face-specific information. The ARENA algorithm technique is closely related to correlation templates. However, the use of novel distance metrics greatly improves the performance. Augmenting the memory base with additional, synthesised face images results in further improvements in performance.

The technique is going to be tested on standard face recognition databases, and direct comparisons with other techniques will show that our algorithm achieves comparable or superior results.

Arena algorithm has also a good asymptotic computation and storage behaviour, and is ideal for incremental training. The system has been integrated with a neural-network based face detection system into a real-word visitor identification system that has been operating successfully in an outdoor environment with uncontrolled lighting for several months.

A) The algorithm

Arena is, as mentioned, a memory based algorithm that employs reduced resolution images, like in Principal Component Analysis, 16x16 and the a parameter of similarity measure L_0^*. One of the most important parts of the system is to reduce the resolution of the image. That is achieved by averaging over non-overlapping rectangular regions in the image. The aim of the system is to find the exact mach of an image from the given datasets, so the distance from the query image to each of the datasets stored images is computed and the best much is returned.

The key point of the algorithm for its good performance is the L_p^* similarity measure. The measure that is used has a better performance than the Euclidean distance. Lp^* is defined as

$$Lp(\vec{a}) \equiv (\Sigma|a_i|^p)^{1/p} \qquad (4)$$

The Euclidean is defined for p=2 so that we have:

$$L_2(\vec{x} - \vec{y}). \qquad (5)$$

Because we are not interested in the actual distances, but only in the ordering we can say that equation 3 becomes:

$$Lp(\vec{a}) \equiv (\Sigma|a_i|^p)^{1/p} \qquad (6)$$

For each reduced resolution image we have is converted to a vector, $\vec{x}$, where each pixel in the image is represented as a component of the vector. So since the individual pixel intensities are noisy, we can define the similarity measure for p=0 as:

$$L_0(\vec{x} - \vec{y}) \equiv \sum_{|xi - yi| \geq \delta} 1 \qquad (7)$$

where δ is a threshold value, such that pixels whose intensities differ by less than δ are considered equivalent.

4. COMPLEXITY

One of the most important aspects of an algorithm is the computational complexity and the storage. Testing the ARENA algorithm, and also different versions of principal component analysis, ARENA can be trained and tested faster, and also has a better storage. The advantage of ARENA compare to other algorithms can be for three main reasons.

Firstly the training time for arena scales linearly with the number of images that we use for training, in comparison with PCA methods. The training times of ARENA and two PCA methods that are used from MATLAB are as follows.

Method	Training Time (Computational Complexity)
PCA-1	$O(N^3 + N^2 d)$
PCA-2	$O(N^3 + N^2 d)$
ARENA	O(Nd)

Where N is the total number of the images that we use and d is the dimension of each image.

For the classification time of the algorithms that PCA-2 and ARENA are slower than PCA-1. But ARENA still faster than PCA-2

Method	Classification Time (Computational Complexity)
PCA-1	O(cm + dm)
PCA-2	O(Nm + dm)
ARENA	O(Nm + d)

Though ARENA has an advantage compare to the PCA-1 algorithm. ARENA is not computing the dimension of the reduced representation, m for each d. c is the number of people on the images.

From these two comparisons of the three algorithms but also from the tests show that ARENA requires less storage space than the other two face recognition methods. PCA-1 requires more storage because needs to store all the vectors of size d, Apart from that ARENA is performing all the computations in the reduced dimensional space. The only disadvantage of ARENA to the other algorithms is that requires more storage if the number of images that we use for training or testing is very large.

Method	**Storage Space**
PCA-1	O(cm + dm)
PCA-2	O(Nm + dm)
ARENA	O(Nm)

5. CONCLUSION

Face recognition is one of the several techniques for recognising people. There are several methods that can be used for that purpose. Some of the most common are using PCA or eigenfaces. Thought there are other new techniques more simple to understand use and implement but also with very good performance. The ARENA algorithm is one of those algorithms. As we show ARENA has very good performance and is a very accurate especially if we use a feedforward neural network.

Face recognition technology has come a long way in the last twenty years. Today, machines are able to automatically verify identity information for secure transactions, for surveillance and security tasks, and for access control to buildings. These applications usually work in controlled environments and recognition algorithms that can take advantage of the environmental constraints to obtain high recognition accuracy. However, next generation face recognition systems are going to have widespread application in smart environments, where computers and machines are more like helpful assistants. A major factor of that evolution is the use of neural networks in face recognition. A different filed of science that also is very fast becoming more and more efficient, popular and helpful to other applications.

The combination of these two fields of science manage to achieve the goal of computers to be able to reliably identify nearby people in a manner that fits naturally within the pattern of normal human interactions. “They must not require special interactions and must conform to human intuitions about when recognition is likely.” This implies that future smart environments should use the same modalities as humans, and have approximately the same limitations. “These goals now appear in reach however, substantial research remains to be done in making person recognition technology work reliably, in widely varying conditions using information from single or multiple modalities.”

References

1) High Performance Memory Based Face Recognition for Visitor Identification, Terence Sim, Shumeet Baluja, Rahul Sukthankar, Mathew D. Mullin

2) Face Recognition: A Convolutional Neural Network Approach, Steve Lawrence, C. Lee Giles, Ah Chung Tsoi, Andrew D. Back, IEEE Transactions on Neural Networks, Volume 8, Number 1, pp. 98-113, 1997.

3) Face Recognition with Multi-Layer Perceptrons,Erik Hjelmås and Jørn Wroldsen , Department of Electrical Engineering and Science Gjøvik College.

4) Face Recognition in Dynamic Scenes, Stephen McKenna, Shaogang Gong and Yogesh Raja, Machine Vision Laboratory, Department of Computer Science Queen Mary and Westfield College.

5) Neural Network Based Face Recognition, Henry A. Rowley, Shumeet Baluja and Takeo Kanade, IEEE.

6) Face recognition: A hybrid neural network approach, S. Lawrence, C. Giles, A. Tsoi, A. Back, Technical report UMIACS-TR-96-16, University of Meriland, 1996

7) S. Arya and D.M. Mount. Algorithms for fast vector quantization. In J. A. Storer and M. Cohn, editors, *Proceedings of DCC 93: Data Compression Conference*, pages 381–390. IEEE Press, 1993.

8) 8 Hans-Ulrich Bauer and Klaus R. Pawelzik. Quantifying the neighborhood preservation of Self-Organizing Feature Maps. *IEEE Transactions on Neural Networks*, 3(4):570–579, 1992.

9) Yoshua Bengio, Y. Le Cun, and D. Henderson. Globally trained handwritten word recognizer using spatial representation, space displacement neural networks and hidden Markov models. In *Advances in Neural Information Processing Systems 6*, San Mateo CA, 1994. Morgan Kaufmann.

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

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

12) R. Brunelli and T. Poggio. Face recognition: Features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10):1042–1052, October 1993.

13) D. K. Burton. Text-dependent speaker verification using vector quantization source coding. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(2):133, 1987.

14) R. Chellappa, C.L. Wilson, and S. Sirohey. Human and machine recognition of faces: A survey. *Proceedings of the IEEE*, 83(5):705–740, 1995.

15) Ingemar J. Cox, Joumana Ghosn, and Peter N. Yianilos. Feature-based face recognition using mixture-distance. In *Computer Vision and Pattern Recognition*. IEEE Press, 1996.

16) David DeMers and G.W. Cottrell. Non-linear dimensionality reduction. In S.J. Hanson, J.D. Cowan, and C. Lee Giles, editors, *Advances in Neural Information Processing Systems 5*, pages 580–587, San Mateo, CA, 1993. Morgan Kaufmann Publishers.

17) H. Drucker, C. Cortes, L. Jackel, Y. Le Cun, and V.N. Vapnik. Boosting and other ensemble methods. *Neural Computation*, 6:1289–1301, 1994.

18) K. Fukunaga. *Introduction to Statistical Pattern Recognition, Second Edition*. Academic Press, Boston, MA, 1990.

19) S. Haykin. *Neural Networks, A Comprehensive Foundation*. Macmillan, New York, NY, 1994.

20) S. Haykin. Personal communication, 1996.

21) D.H. Hubel and T.N. Wiesel. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology (London)*, 160:106–154, 1962.

22) R.A. Jacobs. Methods for combining experts' probability assessments. *Neural Computation*, 7:867–888, 1995.

23) T. Kanade. *Picture Processing by Computer Complex and Recognition of Human Faces*. PhD thesis, Kyoto University, 1973.

24) Hajime Kita and Yoshikazu Nishikawa. Neural network model of tonotopic map formation based on the temporal theory of auditory sensation. In *Proceedings of the World Congress on Neural Networks, WCNN 93*, volume II, pages 413–418, Hillsdale, NJ, 1993. Lawrence Erlbaum.

25) T. Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78:1464–1480, 1990.

26) T. Kohonen. *Self-Organizing Maps*. Springer-Verlag, Berlin, Germany, 1995.

27) Martin Lades, Jan C. Vorbr¨uggen, Joachim Buhmann, J¨org Lange, Christoph von der Malsburg, Rolf P.W¨urtz, and Wolfgang Konen. Distortion invariant object recognition in the dynamic link architecture. *IEEE Transactions on Computers*, 42(3):300– 311, 1993.