Face Recognition Using Convolution Filters and Neural Networks

V. Rihani
Head, Dept. of E&E,PEC
Sec-12, Chandigarh – 160012
vrihani@yahoo.com

Amit Bhandari
Department of CSE & IT, PEC
Sec-12, Chandigarh – 160012
amit.bhandari@yahoo.com

C.P. Singh
Physics Department, CFSL,
Sec-36, Chandigarh - 160036
cpureisingh@yahoo.com

Abstract — *Human Face Recognition has become a potential method of biometric authentication because of its most non-intrusive and user-friendly nature. Automatic face recognition poses various challenges due to: (a) inherent variability of face due to age, gender and race; (b) different facial expressions and orientations of same person's face; and (c) images containing faces have high degree of variability in size, texture, background and illumination. The proposed model uses a two pass method in which the input image is first processed using conventional image processing filters to enhance it and then the edges are detected in the image using convolution mask; which is further followed by face recognition using neural network. This helps the model to achieve more accuracy and the whole process to be more efficient than simply applying neural network model for face recognition.*

Keywords — Neural network, face recognition, classification, convolution filters.

1 Introduction

Face Recognition has been recently the most favored subject due to its most non-intrusive nature and wide range of applications. Its applications are in crime investigations, security (biometric authentication), etc. But the whole process depends on the images, so, the image quality play an important role and is influenced by some fundamental characteristics like lighting, camera position/angle, shade, face view (front/side), and emotional expressions.

2 Techniques and dataset used

2.1 Convolution Filter

In this approach, three different masks are used to detect edges named after their inventors, Sobel, Prewitt, and Kirsh. In each case, a horizontal version of the filter is applied to one bitmap, vertical version to another, and (1) is used to merge both horizontal and vertical version together.

pixel = sqrt(pixel1 *pixel1 + pixel2 * pixel2) … (1)

The convolution masks of 3x3 matrices are represented as below:

1	2	1
0	0	0
-1	-2	-1

Fig. 1. Sobel convolution mask for edge detection

1	1	1
0	0	0
-1	-1	-1

Fig. 2. Prewitt convolution mask for edge detection

5	5	5
-3	-3	-3
-3	-3	-3

Fig. 3. Kirsh convolution mask for edge detection

These filters perform the horizontal edge detection and rotating them through 90^{o} gives us the vertical, and then the merging takes place.

2.2 Face Detection

The proposed model uses a multi-layer feed forward backpropagation-algorithm based neural network which has one input layer, one hidden layer and one output layer.

The neural network examines small windows of an image, and decides the presence or absence of face in each window. The face detection component accepts a 20x20 pixel region of the image and output a value 1 or -1 signifying the presence of face in the image. The algorithm used for face detection proposed by H.A. Rowley, S. Baluja, and T. Kanade [4] is as shown below:

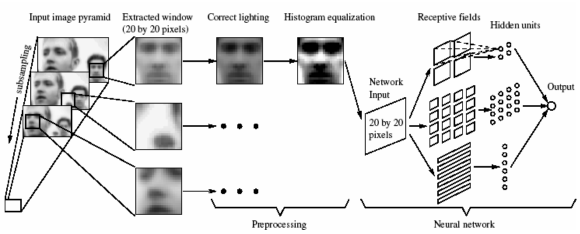

Fig. 4. System diagram of Rowley's method [5].

To detect faces in the input, the filter is applied at every location of the image. The input image can be sub sampled to reduce the size if the size of face in image is larger than the window size.

2.3 Dataset

The dataset for the testing purposes has been selected from various sources, which include the face database published by UMIST, University of Essex, and ORL.

The UMIST Face Database consists of 564 images of 20 people. Each face covers a range of poses from profile to frontal views.

The University of Essex has published the database in years 1994, 1995 and 1996. The backgrounds of images is plain color, however, there is small head scale variation, head turn, tilt and slant. Also, there is major expression variation, however, very little image lighting variation.

The ORL face database was originally published by Cambridge University. The sample of dataset is shown below.

Fig. 5. Sample of ORL Face Database.

3 Neural Network

The proposed neural network model uses sigmoid activation function and stochastic gradient backpropagation learning algorithm.

The sigmoid function is given as:

$$f(x) = \frac{1}{1+e^{-\beta x}}, \beta > 0 \qquad \ldots (2)$$

The sigmoid function is used to set the output value of any given neuron by processing the sum of weighted input values and any bias applied.

The learning algorithm used is backpropagation algorithm. The error backpropagation algorithm is a gradient descent algorithm where weights and biases are adjusted to minimize a cost function equal to the mean square error in the network.

The proposed model uses the stochastic version of the algorithm wherein the error is back-propagated after every learning case. The previous version works on global error.

The stochastic version is based on the fact that when the network weights approach a minimum solution, the gradient becomes small and step size diminishes too, giving very slow convergence. If a so called momentum factor is added to the weight update equations, the weights can be updated with some component of past updates. This can reduce the decay in learning updates and cause the learning to proceed through the weight space in fairly constant direction. The benefits of this, apart from faster convergence to the minimum, is that it may even be possible to sometimes escape a local minimum if there is enough momentum to travel through it.

The stochastic backpropagation changes the weight matrix for neuron i as follows:

$$W_{ij}(n+1) = W_{ij}(n) + \alpha.A_i.S_j + \gamma.(W_{ij}(n) - W_{ij}(n-1) \qquad \ldots (3)$$

The target output vector is given as

$$T_i(n+1) = T_i(n) - \alpha.A_i + \gamma.(T_i(n) - T_i(n-1)) \qquad \ldots (4)$$

where α and γ are positive valued scalar gain or learning rate constants, less than 1.

4 Face Detection

As shown in Fig. 4, the proposed model follows the approach. Before feeding the face as input to the neural network, however, the extracted face feature is preprocessed for intensity equalization. The process of intensity equalization is done automatically and is mandatory step for the whole detection process. The intensity equalization includes grayscale conversion, if required, followed by linear best fit adjustment and finally histogram equalization.

5 System Overview

The whole process of face recognition is divided into three modules, viz,

a Image Analysis Module
b Adjust if possible for orientation of face
c Classification based on class characteristics

The image analysis module analyzes the disputed image and figures out the profile of the face in image, camera angle, degree of illumination, resolution of image. The given image should exhibit the threshold values of all the parameters.

If the input image to the system contains some disorientation(s), then the image should be adjusted so that further classification/recognition operation can be applied. If the image can't be oriented then the image can be discarded.

The classification based measurement matrices will have a provision to compare and form a decision at this stage. Measured parameters need to be provided.

Also, the various factors that have been identified to affect the performance of the system are:

i *Pose*: The image of face vary due to the relative camera-face pose (frontal, 45^o, profile, upside down), and some special features such as an eye or the nose become partially or wholly occluded.

ii *Presence or absence of structural components*: Facial features such as beards, mustaches, and glasses may or may not be present and there is a great deal of variability among these components including shape, color and size.

iii *Facial expressions*: The appearance of face is directly affected by person's facial expressions.

iv *Image orientation*: Face images directly vary for different rotations about the camera's optical axis.

v *Occlusion*: Faces may be partially occluded by other objects. In an image with a group of people, some faces may partially occlude other faces.

vi *Imaging conditions*: When the image is formed factors such as lighting (spectra, source distribution and intensity) and camera characteristics (sensor response, lenses) affect the appearance of the face.

vii The objective of the system is to provide face recognition for forensic applications. The various aspects of the system are:

viii *Face localization*: The image is analyzed to determine the position and size of one or more face. (All other steps consider that only one face has been determined in the image)

ix *Eye localization*: The positions of eye centers in the face are determined.

x *Image quality check*: The quality of face image is checked to see whether it is sufficient for steps that follow.

xi *Normalization*: The face is extracted from image and is scaled and rotated such that an image of fixed size results, with eye centers at fixed positions in the image.

xii *Preprocessing*: The normalized image is preprocessed with standard techniques such as histogram equalization, intensity normalization, and others.

xiii *Feature extraction*: From the preprocessed image, those features are extracted that are relevant for distinguishing a person from another.

xiv *Construction of reference set (Training)*: During enrollment, the facial features of several images of a person are extracted and combined into reference set, also called Biometric Template.

xv *Recognition*: For recognition, the set of extracted features is compared with the reference set of the person who the person in the image processed is claimed to be.

6 System Design

The Fig. 7 shows the steps followed in feature extraction:

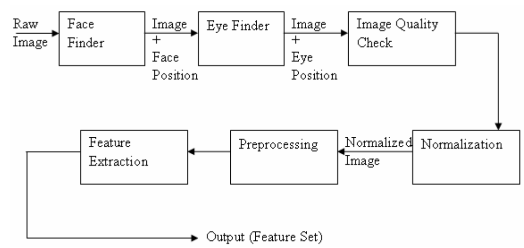

Fig. 6. Steps for creation of feature set

The following figures show the working of system in training and recognition phases:

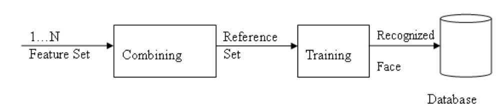

Fig. 7. Steps for training phase

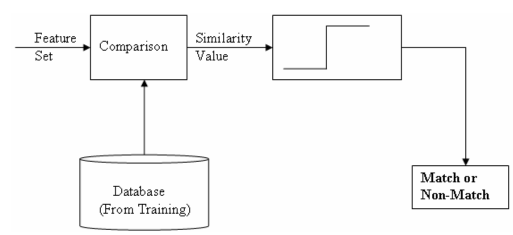

Fig. 8. Steps for recognition phase

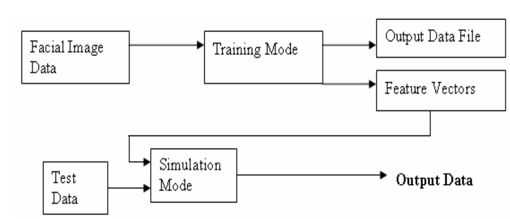

Fig. 9. Steps for both the processes

7 Implementation

The implementation of the system has been carried out in Microsoft Visual C#.NET. The system is fed with matching and non-matching samples including the mask for ignoring the background pixels.

8 Testing and Results

The images in the dataset have been categorized as non-faces and faces. This has been done manually so as to enhance the accuracy of results of the system. Practically, any image can be regarded as a non-face image, however, due to the similarity between the face image and non-face image, any face having more than 10^o of rotation is regarded as non-face.

The developed system is being tested with the selected dataset and is expected to give higher accuracy.

9 Conclusion

The proposed system can give more accurate results by varying the degree of rotation.

The proposed system works only with grayscale images; however, the system can be enhanced by providing functionality for color-based tracking and applying color segmentation. The face detector, in this case would be applied only to regions having skin color, thus speeding up the algorithm as well as eliminating much of the false detection.

The system has been designed by keeping forensic science application in mind. The enhancement of the system for more generalized applications of face recognition and optimizing the algorithm accordingly has been left as future work.

10 Acknowledgements

The work is supported by Central Forensic Science Labs, Chandigarh and Punjab Engineering College, Chandigarh. The authors would like to thank Daniel Graham for providing UMIST Face Database and Dr Libor Spacek, University of Essex, for providing collection of Facial Images.

11 References

[1] I. Aleksander and H. Morton, "An Introduction to Neural Computing," International Thomson Publishing, 1995.

[2] C. Stergiou, and D. Siganos, "Neural Networks," http://doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html.

[3] D. Patterson, "Artificial Neural Networks", Printce Hall India, 1996.

[4] V.V. Starovoitov, D. I. Samal, and D.V. Briliuk, "Three Approaches for Face Recognition," Proceeding of 6th International conference on Pattern Recognition and Image Analysis, Russia, 21-26 (10) 2002, pp. 707-711.

[5] H.A. Rowley, S. Baluja, and T. Kanade, "Human Face Detection in Visual Scenes," Technical Report, Department of Computer Science, Carnegie Mellon University, Pittsburgh, 1995, CMU-CS-95-158R.

[6] Yi-Qiong Xu, Bi-Cheng Li, Bo Wang, "Face Detection and Recognition using Neural Network and Hidden Markov Models," IEEE Conference on Neural Network and Signal Processing, 14-17 (12) (2003), vol. 1, pp. 228-231.

[7] S. Lawrence, C.L. Giles, A.C. Tsoi, and A.D. Back, "Face Recognition: A Convolutional Neural Network Approach," IEEE Transactions on Neural Network, 1997, vol. 8, No. 1, pp. 98-113.

[8] S.-H. Lin, S.Y. Kung, and M. Fang, "A Neural Network Approach to Face/Palm Recognition," *Proc.* 5th IEEE Wkshp. Neural Networks Signal Processing, Cambridge, MA, 1995, pp. 323-332.

[9] C. Peacock, A. Goode, and A. Brett, "Automatic forensic face recognition from digital images," Science & Justice, 2004, vol. 44, No. 1, pp. 29-34.

[10] Daniel B. Graham, and Nigel M. Allinson, "Face Recognition: From Theory to Applications," NATO ASI Series F, Computer and System Sciences, 1998, vol. 63, pp. 446-456.

[11] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, "From Few to Many: Illustration Cone Models for Face Recognition under Variable Lighting and Pose," IEEE Trans. Pattern Anal. Mach. Intelligence, 2001, vol. 23, No. 6, pp. 643-660.

[12] T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression (PIE) Database of Human Faces," Technical Report, The Robotics Institute, Carnegie Mellon University, Pittsburgh, CMU-R1-TR-01-02.

[13] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression (PIE) database," Proceedings of 5th IEEE International Conference on Face and Gesture Recognition, 2002.

[14] Jeong-Seong Park, You Hwa Oh, Sang Chul Ahn, and Seong-Whan Lee, "Glasses Removal from Facial Image Using Recursive Error Compensation," IEEE Transactions on Pattern Analysis and Machine Intelligence, May 2005, vol. 27, No. 5, pp. 805-811.