# Face Verification Using Correlation Filters and Autoassocoative Neural Networks

**Anil Kumar Sao and B. Yegnanarayana**
Speech and Vision Laboratory
Department of Computer Science and Engineering
Indian Institute of Technology Madras, Chennai-600 036, India
{anil, yegna}@cs.iitm.ernet.in

**Abstract**

Face verification is the process of accepting or rejecting the identity claim of a person using information from his/her face. Representation of the face is an important issue in face verification. This paper propose edge gradient-based representation of face, for correlation-based face verification. The edge gradient based representation of face is obtained using one-dimensional (1-D) processing of the image, which has the advantage of providing multiple partial evidences for a given image. This representation of face is used to recognize the faces, which is performed by a specific type of correlation filter called Minimum Average Correlation Energy (MACE). Separate correlation filters are employed for each partial evidence. A method is proposed to combine the output of the filter using an Auto-Associative Neural Network (AANN) model to arrive at a decision to accept or reject the claim.

## 1. INTRODUCTION

Automatic recognition of faces is one of the challenging problems in human-computer interaction [1]. A number of methods have been proposed in the literature which include eigenface [2], template matching [3], correlation filter [4]. This paper proposes an edge gradient representation of faces for correlation-based face verification. The correlation is performed by Minimum Average Correlation Energy (MACE) filter [5]. This filter was proposed for face verification using the gray level values of the face image [4]. The edge gradients are robust to changes in illuminations [6], which is a major problem in using gray level values of 3-D objects like faces. The edge gradients are computed using one-dimensional (1-D) processing of the image [7], which has the advantage of providing multiple partial evidences for a given image. Separate correlation filters are employed for each partial evidence, and a method is proposed for combining the outputs of the filters using an Auto-Associative Neural Network (AANN) model. The performance of face verification system using proposed representation is evaluated on Facial Expressions Database, collected at Advanced Multimedia Processing (AMP) Lab at the Electrical and Computer Engineering Department of Carnegie Mellon University (CMU) [8]. The database consists of 13 subjects, each having 75 images with varying facial expressions. The size of the image was scaled down to 30×30 in all the experiments.

The organization of the paper is as follows: Section 2 provides background of the MACE filter. Section 3 explains the edge gradient based representation of face. This representation is used to recognize face using MACE filter and is explained in Section 4. The sections 5 gives results for the proposed representation, along with the results using gray scale values for comparison. Section 6 summarizes the work and discuses scope for future work.

## 2. MINIMUM AVERAGE CORRELATION ENERGY (MACE) FILTER

Cross-correlation is perhaps the most popular technique for object matching, because of its mathematical tractability, robustness to noise and shift, low computational complexity and simple hardware implementation. Face verification is performed by cross-correlating the input face image with the derived MACE filter, and processing the resultant correlation output. Fig. 1 shows the cross-correlation obtained using Discrete Fourier Transform (DFT) implementation. The filter output should give high peak value at the origin of the correlation plane when the input face image is from true class. Here the origin is the center of the cross-correlation output. On the other hand, if the input face image is from false class, then the filter output should give a low peak at the origin. The relative heights of these peaks are used to determine whether the input face is from the true or false class. The MACE filter minimizes the Average Correlation

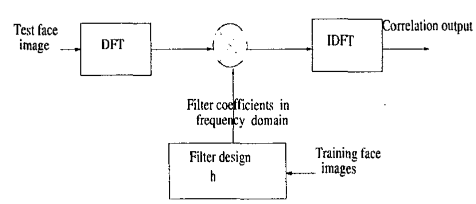



**Figure 1. Block diagram of correlation process**

Energy (ACE) of the correlation output due to the training face images, while maintaining the correlation peak constraint. The correlation peak constraint is that the height of the peak at the origin of the correlation plane due to different training face images should be same. The resulting MACE filter in vector form is as follows [5]:

$$\mathbf{h} = \mathbf{D}_{\mathbf{x}}^{-1}\mathbf{X}(\mathbf{X}^{+}\mathbf{D}^{-1}\mathbf{X}^{-1})\mathbf{c}. \qquad (1)$$

The various symbols are explained below: Suppose that we have N training images from the authentic class with each image having R×C pixels in it. We perform the R×C point 2D-DFT on these partial evidences, and convert the 2D-DFT array into a 1-D column vector by lexicographic order. These vectors form the columns of the d ×N matrix **X** in Eq. (2). Where d= R×C. The column vector **c** with N elements contains the prespecified correlation peak values. The d×d diagonal matrix **D** contains, along its diagonal, the average power spectrum of the partial evidences of the training images (i.e., average of the magnitude of squares of the columns of **X**). The synthesized filter **h** is a column vector with d elements, and the 2-D correlation filter is obtained by reordering the column vector back to a 2-D array. The superscript '+' symbol represents the complex conjugate transpose.

## 3. EDGE GRADIENT REPRESENTATION

One of the problems in using gray scale values of images for object matching is that changes in illumination can significantly reduce the correlation between test and template images. The change due to illumination variation is difficult to predict without an a priori model of the object under consideration. It has been observed that, compared to gray scale values, edge maps are more robust to illumination variation [6]. Edge maps are obtained by computing the intensity gradients and thresholding the gradient values. The selection of the threshold value is a major problem in edge map representation. If the threshold value is low, then spurious edges come up in the edge map. On the other hand, a high threshold value can obliterate significant edges. Another problem in using edge map for correlation is that even a small deformation in the edge contours can significantly reduce the correlation.

We propose continuous edge gradient image to recognize the face. This representation has the advantage of not requiring any thresholding, being more spatially distributed compared to the edge map. The edge gradient computation requires two operators: A smoothing operator to decrease noise and a spatial differential operator to compute the gradient. We use a Gaussian function for smoothing and the derivative of Gaussian function for computing the edge gradient. A problem encountered when smoothing the image before computing the gradient is that genuine edges also tend to get smoothed [9]. The smearing of genuine edges can be reduced using one-dimensional (1-D) processing of the image [7].

### 3.1 Partial edge evidence

In 1-D processing of a given image, the smoothing operator is applied along one direction and the derivative operator is applied along the orthogonal direction. By repeating this procedure in the orthogonal direction, two partial edge gradients are obtained which together represent the intensity gradient of the image. As the smoothing is done along a direction orthogonal to that of the edge extraction, the smearing of edges is reduced.

1-D Gaussian function along a line which makes an angle $\theta$ degree with respect to the x-axis, and has a perpendicular distance of $\rho$ from the origin, is given by

$$f_{\rho,\theta}(x) = \frac{1}{\sqrt{2\pi}\sigma_1} \exp(-\frac{x^2 \sec^2\theta}{2\sigma_1^2}) \quad (2)$$

where $\sigma_1$ is the width of the Gaussian function. The derivative operator is obtained by taking first derivative of the 1-D Gaussian function and is given by

$$g_{\rho,\theta}(x) = \frac{-x\sec^2\theta}{\sqrt{2\pi}\sigma_2^3} \exp(-\frac{x^2 \sec^2\theta}{2\sigma_2^2}) \quad (3)$$

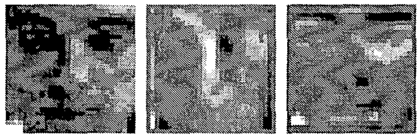
where $\sigma_2$ is the variance of the Gaussian function whose derivative is computed. For the smoothing operation 1-D convolution is performed along straight lines making an angle $\theta$+90 with x-axis in the image plane. The smoothed image $i_\theta^s(x,y)$ is obtained as

$$i_\theta^s(x, \frac{\rho + x\sin(\theta+90)}{\cos(\theta+90)}) = i(x, \frac{\rho + x\sin(\theta+90)}{\cos(\theta+90)}) * f_{\rho,\theta+90}(x) \quad (4)$$

where $*$ denotes the 1-D convolution operator and x is the independent variable. The convolution is performed for all values of $\rho$ to obtain the smoothed image. The partial edge gradient $i_\theta^g$ is computed by applying the derivative operator along the orthogonal direction, that is along the straight line which is at an angle $\theta$ with respect to x-axis in the image plane. The edge gradient is obtained from

$$i_\theta^g(x, \frac{\rho + x\sin\theta}{\cos\theta}) = i_\theta^s(x, \frac{\rho + x\sin\theta}{\cos\theta}) * g_{\rho,\theta}(x) \quad (5)$$

for different values of $\rho$ to get $i_\theta^g(x,y)$. By performing the operations in the orthogonal direction we get $i_{\theta+90}^g$, which along with $i_\theta^g$ represents the intensity gradient of the image. Fig. 2 shows an example of the gradient maps obtained for $\theta$=0$^o$ and $\theta$=90$^o$. These partial evidences of face image is used to recognize the face using MACE filter.



(a) (b) (c)

**Figure 2. (a) Gray level face image. Partial edge evidence ($i_\theta^g$) of face image (shown in (a)) obtained using $\sigma_1$=0.9, $\sigma_2$=1, and $\theta$= (b) 0°, (c) 90°**

## 4. PARTIAL EDGE EVIDENCE AS INPUT TO MACE FILTER

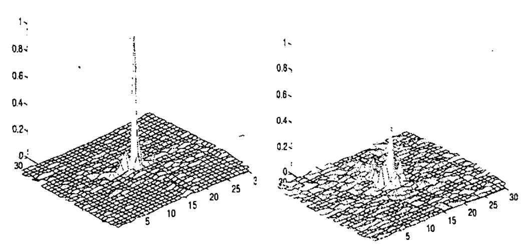
As the derivative operator is in one-dimension, each gradient map $i_\theta^g$ gives only partial evidence about the image. To obtain the complete information about intensity gradient, we must compute $i_\theta^g$ for different values of $\theta$. In our implementation, we computed gradients along four directions ($\theta$=0, 45, 90 and 135). As each of these direction give a different view of the image, it seems logical to build a separate filter for each direction. Thus for each of the 13 subjects in the database, we have a bank of four filters, one for each value of $\theta$. Fig. 3 shows the correlation output for horizontal ($\theta$=0°) edge gradient for a true class face image and a false class face image. One way to take decision from the correlation outputs of the MACE filter is by using the Peak to Sidelobe Ratio (PSR) [4]. The PSR measures the sharpness of the highest peak in the correlation output, the reason being that for a valid claimant's face image the peak will be high and sharp, and in case of a false claim the peak will be low and blunt, The PSR is denoted by $P_\theta$, and is defined as

$$P_\theta = \frac{p - \mu}{\sigma}. \tag{6}$$

where p is the value of maximum peak in the correlation output, $\mu$ is the mean of the correlation output of a window (of size $19 \times 19$ pixels) around the peak, and $\sigma$ is the standard deviation of the values in this window. In actual practice, we leave out a region of size $7 \times 7$ in the center while computing $\mu$ and $\sigma$.

The values of $\sigma_1$ and $\sigma_2$ are taken as 0.9 and 1, respectively. For low values of $\sigma_2$, the correlation between partial edge evidence is difficult, leading to higher intra-class variations. On the other hand for large values of $\sigma_2$, the inter-class discrimination is less due to smearing of edges.

Another way of controlling the degree of spreading the edge information in the image plane is by eliminating the Discrete Fourier Transform (DFT) coefficients of the image. Elimination of high frequency DFT coefficients leads to blurring of the image. Thus edges can be smeared by eliminating the high frequency DFT coefficients. The effect of removing the high frequency DFT coefficients is same as increasing $\sigma_2$. We have found that the elimination of 10 highest frequency DFT coefficients along both frequency axes is optimal for partial edge evidence.



(a) (b)

**Figure 3. Correlation output obtained if $i_\theta^g$ is used as input to the MACE filter for (a) True class (b) False class.**

### 4.1 Combining the partial evidences

It is difficult to verify the person using only one partial edge evidence, since it captures only partial information. One approach to combine is to form a four dimensional feature vector for a given image from the PSR values obtained from the filter bank, and classify these vector into true and false class. Fig. 4 shows the scatter plots for the PSR vectors of true and false class images for person 4. In order to visualize the result, we plotted only three dimensions instead of all the four ($\theta$=0, 45, 90, 135). We can see the separation between the clusters of true and false classes. A Multi-Layer Perceptron (MLP) could be used to perform the classification, but for training it needs large number of true class images. This will go against the design requirement of using as few training images as possible [4].

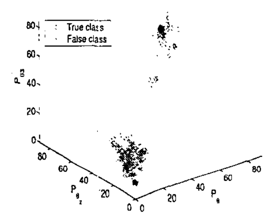



**Figure 4. Scatter plot for a person when partial edge evidence is used as feature for $\sigma_1$=0.9, $\sigma_2$=1, $\theta_1$=0°, $\theta_2$=45°, $\theta_3$=90°, number of training images=5, and 10 high frequency DFT coefficients along both the frequency axes are eliminated.**

An interesting observation one can make from Fig. 4 is that the vectors for the false class are tightly clustered near the origin in contrast to the scattered vectors of the true class image. This characteristic was verified to be present for all persons. An alternative to the MLP-based classification is to capture the distribution of false class vectors. For given a test vector, the classifier can determine whether it belongs to the false-class distribution or not. The false class distribution is captured using an Auto-Associative Neural Network (AANN) [10] model. A separate AANN model is derived for each person. Given a test vector and the claimed identity, we compare the error in associating the vector to the AANN model corresponding to the claimed identity. If the error is above a threshold the vector is said to belong to true class, else the false class.

We give a brief summary of the complete process. During training, for each person, a few face images are chosen as training images. Edge gradient along four directions (0, 45, 90, 135) are computed for the images, and are used to derive the correlation filters for each person. Several false class images are given to the filter bank to get the four dimensional feature vectors of the PSR values. These feature vectors are used to train an AANN model, which captures their distribution. The structure of the AANN model is 4L 8N 2N 8N 4L, where L denotes a linear unit and N denotes the nonlinear unit. The AANN model is trained using 700 four dimensional feature vectors of the PSR of false classes. During testing, given a test image and the claimed identity, the edge gradients along different directions are computed and given to the filter bank to get a four dimensional PSR vector. The four dimensional feature vector is used to compute the error in associating the vector to AANN model corresponding to the claimed identity. If the error is above a threshold then claim is accepted.

## 5. RESULTS OF EXPERIMENTS

False acceptance (FA) and false rejection (FR) are the two errors that are used in evaluating a face verification system. The tradeoff between FA and FR is a function of decision threshold. Equal Error Rate (EER) is the value for which the error rates FA and FR are equal. Table 1 shows the Equal Error Rate (EER) obtained for all the 13 subjects, when gray values are given as input to the MACE filter, as reported in [4] . Table 2 show the EERs obtained for the proposed representation, for different sizes of the training set. In general, as the size of the training set decreases the performance degrades.

**Table 1. Results for gray level representation: EER for different number (N) of training face images.**

| Person | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| N=5 | 0 | 1.3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| N=3 | 0 | 0.9 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**Table 2. Results for the edge gradient representation: EER for different number (N) of training face images for each person using AANN for combining partial evidences from four directions ($\theta$=0°,45°,90°,135°).**

| Person | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| N=5 | 0 | 2.6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| N=3 | 0 | 12.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

## 6. SUMMARY

This paper discussed an edge gradient based representation of face images for face verification using MACE filter. The edge gradient is obtained using 1-D processing of the image. We have also proposed a classification technique based on AANN for combining the partial evidence. Experimental results indicate that the performance of the proposed representation is equivalent to the gray level except for one person. The reason is using edge gradient, it is difficult to correlate the partial evidence of two face images. But, the proposed representation has the advantage that it is less sensitive to illumination. The performance of the proposed representation can be improved by spreading the edge gradient information with potential field [11].

## REFERENCES


[1] R. Chellpa, C. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," Proc. IEEE, vol. 83, pp. 705–740, May. 1995.

[2] M. Turk and A. Pentland, "Eigenfaces for recognition," J. Cognitive Neuroscience, vol. 3, pp. 71–86, 1991.

[3] R. Brunelli and T. Poggio, "Face recognition: features versus templates," IEEE Trans. Pattern Anal. Machine Intell., vol. 15, no. 10, pp. 1042–1052, Oct. 1993.

[4] B. V. K. V. Kumar, M. Savvides, K. Venkataramani, and C. Xie, "Spatial frequency domain image processing for biometric recognition," IEEE Int. Conf. Image Processing, New York, pp. 53–56, Sep. 2002.

[5] A. Mahalanobis, B. V. K. V. Kumar, and D. Casasent, "Minimum average correlation energy filters," Appl. Opt., vol. 26, pp. 3633–3640, Sept. 1987.

[6] Y. Gao and M. K. H. Leung, "Face recognition using line edge map," IEEE Trans. Pattern Anal. Machine Intell., vol. 24, pp. 765–779, Jun. 2002.

[7] P. Kiran Kumar, Sukhendu Das, and B. Yegnanarayana, "One-Dimensional processing of images," in Proc. Int. Conf. Multimedia Processing and Systems, IIT Madras, Aug. 2000.

[8] http://amp.ece.cmu.edu, "Advanced mulitimedia procesing lab web page at electrical and computer engineering department at CMU,"

[9] P. P. Raghu and B. Yegnanarayana, "Segmentation of gabor filtered textures using deterministic relaxation," IEEE Trans. on Image Processing, vol. 5, pp. 1625–1636, Dec. 1996.

[10] B. Yegnanarayana and S. Kishore, "AANN: An alternative to GMM for pattern recognition," Neural Networks, vol. 15, pp. 459–469, Apr. 2002.

[11] R. Courant and D. Hilbert, "Methods of Mathematical Physics," Interscience, New Yor, vol. 1, 1953.