Template Matching Approach for Pose Problem in Face Verification

Anil Kumar Sao and B. Yegnanaarayana

Speech and Vision Laboratory, Department of Computer Science and Engineering,
Indian Institute of Technology Madras, Chennai-600036, India
{anil, yegna}@cs.iitm.ernet.in

Abstract. In this paper we propose a template matching approach to address the pose problem in face verification, which neither synthesizes the face image, nor builds a model of the face image. Template matching is performed using edginess-based representation of face images. The edginess-based representation of face images is computed using one-dimensional (1-D) processing of images. An approach is proposed based on autoassociative neural network (AANN) models to verify the identity of a person using score obtained from template matching.

1 Introduction

It is important for a face recognition system to be able to deal with faces of varying poses, since the test face image is not likely to have the same pose as that of the reference face image. Other important aspects of the test face image that need to be considered for face recognition are illumination, expression, and background [1]. In this paper we focus on the pose problem.

The problem of pose variation has been addressed using two different approaches. In the first approach, a model for each person is developed using face images at different poses of that person. The resultant model is used to verify a given test face image. Such methods are discussed in [1,2]. But, the resultant models may average out some of the information of the face image which is unique for that person. In the second approach, the pose information of the given test face image is extracted, which is then used to synthesize a face image in a predefined orientation (pose) using a 3-D face model. The resultant synthesized face image is matched with a reference face image in the predefined pose. Such methods are discussed in [3,4]. In these cases some artifacts may be introduced, or some unique information may be lost while synthesizing the face image, which in turn can degrade the performance of face recognition system.

In this paper we have proposed a template matching based approach, which neither synthesizes the face image in a predefined pose, nor derives a model for a person's face. We have used face images (of different poses) separately for template matching. The template matching is performed using the edginess-based [5] representation of a face image. The scores obtained from separate template matching are combined in a selective way. The combined scores are used with an autoassociative neural network (AANN) [6] model for classification.

B. Gunsel et al. (Eds.): MRCS 2006, LNCS 4105, pp. 191–198, 2006.

The performance of the proposed approach is evaluated on the FacePix database [7,8]. The FacePix database consists of two sets of face images: a set with pose angle variation, and a set with illumination angle variation. We have used pose angle variation set in our experiments. This set consists of 30 subjects, each having 181 images (representing angles from -90° to 90° at 1° interval), with varying pose. In this paper we denote these images by $I^1, \ldots, I^{181}$.

The organization of the paper is as follows: Section 2 explains the template matching using the edginess-based representation of a face image, and the computation of the edginess image using 1-D processing. The scores of matching with several templates are combined in a selective way as explained in Section 3. An approach is proposed in Section 4 to classify based on the combined scores using AANN models. Experimental results are given in Section 5, a summary of the work in Section 6.

2 Template Matching

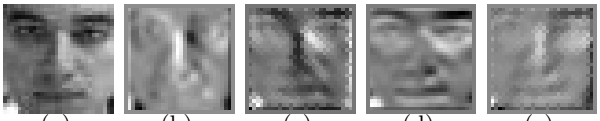
Template matching is performed using a correlation based technique [9]. The correlation between reference face image $r(x, y)$ and test face image $i(x, y)$ is computed as follows:

$$\begin{aligned} c(\tau_x, \tau_y) &= i(x, y) \odot r(x, y) \\ &= \int\int i(x, y) r(x + \tau_x, y + \tau_y) dx dy \\ &= \int\int I^*(u, v) R(u, v) \exp(j2\pi(u\tau_x + v\tau_y)) du dv \end{aligned} \tag{1}$$

where $I(u, v)$ and $R(u, v)$ are the Fourier transforms of $i(x, y)$ and $r(x, y)$, respectively, and $\odot$ denotes the correlation operator. The correlation output $(c(\tau_x, \tau_y))$ can be quantified using Peak-to-Sidelobe Ratio (PSR) measure [9]. The PSR should be high when the test and reference face images are similar. On the other hand PSR should be low.

2.1 Edginess-Based Representation of Face Image

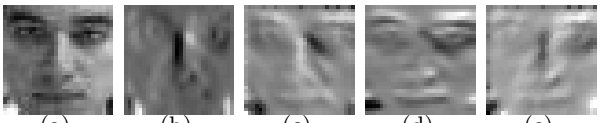
The edginess-based representation has advantage that it is less sensitive to illumination [5]. This representation is computed using one-dimensional (1-D) processing of the image, which gives multiple partial evidences of the same face image [5]. In the 1-D processing of a given image, the smoothing operator is applied along one direction, and the derivative operator is applied along the orthogonal direction. By repeating this procedure of smoothing followed by differential operation along the orthogonal direction, two edge gradients are obtained, which together represent the intensity gradient of the image. We call the edge gradient obtained by applying the derivative operator along θ direction with respect to horizontal scan line as i^g_θ. For different values of θ, the computed edge gradients are shown in Fig. 1. One of the problem with the edge gradient representation for correlation is that most of the values in the edge gradient image



(a) (b) (c) (d) (e)

Fig. 1. (a) Gray level face image. Edge gradient (i_θ^g) of the face image obtained for (b) $\theta = 0°$, (c) $\theta = 45°$, (d) $\theta = 90°$, and (e) $\theta = 135°$.

are very small, close to 0, except near the edges. Because of this, even a small deviation in the edge contour for the same face image can significantly reduce the value of the PSR. The edge information can be spread by using potential field representation [5] derived from the edge gradient, which brings interaction among the gradient values at different points in the image. Let u_θ be the potential field representation derived from the edge gradient (i_θ^g) by an approach described in [5]. Fig. 2 shows the potential fields obtained for different values of θ. The edge gradients (i_θ^g) for different directions (θ) give different information of



(a) (b) (c) (d) (e)

Fig. 2. (a) Gray level face image. Potential field (u_θ) developed from the edge gradient of the face image (a) for (b) $\theta = 0°$, (c) $\theta = 45°$, (d) $\theta = 90°$, (e) $\theta = 135°$.

the face image. Hence, we have performed correlation between partial evidence (u_θ) of given test and reference face images. Let c_θ be the correlation output obtained when the correlation between the partial evidence along θ direction of the test and reference face images is performed. The resultant correlation output is used to compute the PSR (P_θ). Ideally P_θ should be high if the given test face image is similar to the reference image. In our experiment we have computed the partial evidence (u_θ) along four directions (θ= 0°, 45°, 90°, and 180°). Hence, for a given test face image a four dimensional feature vector (four PSR values) is obtained. Fig. 3(a) shows the scatter plot obtained from the PSR vectors of the true class and false class face images for a person. For visualization we show the plot using three (θ= 0°, 45°, and 90°) of the four dimensions. In this example we have used I^1 as a reference face image of a person. The remaining 180 face images of the given person form examples of true class, and the corresponding PSR values are denoted by diamond ($\diamond$) symbol in the plot. For the false class, $29 \times 181 = 5249$ face images are available, and the PSR values are shown by point ($\cdot$) symbol in the scatter plot. Though the separation between the true and false class face images is not decisive, one can observe from the plot that high scores are given by the face images I^2, I^3, I^4, I^5, and I^6 of the true class.

These face images have pose that is close to the pose of the reference face image. One can also see that none of the face images of false class gives high scores and that these scores are clustered near the origin. It means that the chances of matching face images of two different persons even with the same pose, is less. Similar observations can be made from the scatter plot shown in Fig. 3 (b), which is obtained using I^{46} as the reference face image. This behavior is utilized to address the pose problem in face verification.

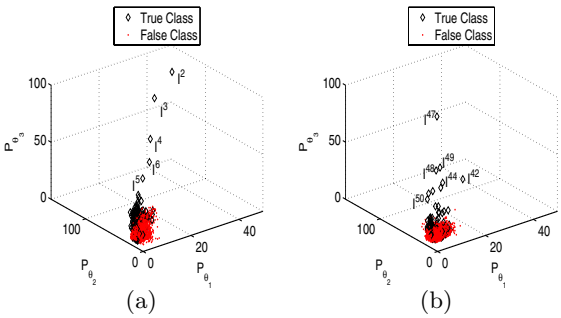


Fig. 3. Scatter plot of a person using potential field representation, obtained from θ_1=0°, θ_2=45°, θ_3=90°, and using reference face image as I^1 in (a), and as I^{46} in (b)

3 Combining Scores from Different Templates

One can conclude from the previous section that if a test face image of the true class has a pose that lies between poses of two reference face images, then the test image will give high scores with respect to both the reference face images. It is better to combine these score rather then use than separately for taking decision. One way to combine the scores is as follows. Let $P_\theta^{t,l}$ be the similarity score (PSR) obtained when the potential field representation along θ direction of the test face image I^t is correlated with the corresponding representation of reference image I^l. The combined similarity score for two reference images I^l and I^m is given by

$$P_\theta^{t,l,m} = \left[\frac{1}{2}\left[\left(P_\theta^{t,l}\right)^n + \left(P_\theta^{t,m}\right)^n\right]\right]^{\frac{1}{n}}, \tag{2}$$

where the parameter n decides the weights associated with the scores. For $n \leq 1$, $\min[P_\theta^{t,l}, P_\theta^{t,m}] \leq P_\theta^{t,l,m} \leq \frac{P_\theta^{t,l}+P_\theta^{t,m}}{2}$, and for $n \geq 1$, $\frac{P_\theta^{t,l}+P_\theta^{t,m}}{2} \leq P_\theta^{t,l,m} \leq \max[P_\theta^{t,l}, P_\theta^{t,m}]$. The value of the parameter n has to be chosen in such a way that $P_\theta^{t,l,m}$ should be small for false class face images and large for true class face images. We have found empirically that n=3 is a suitable choice. Fig. 4 (c) is the scatter plot obtained after combining the PSR scores in Figs. 4 (a) and (b) using (2), for n =3. Figs. 4 (a) and (b) are the same scatter plots as shown in Figs. 3 (a) and (b), respectively, but with a different view. In this example

we have shown only the scores obtained from the true class face images I^t for $2 \leq t \leq 45$. One can see that the separation between true and false class is better in Fig. 4 (c) as compared to Figs. 4 (a) and (b). Similarly, we can combine the scores obtained from other reference face images of adjacent poses.

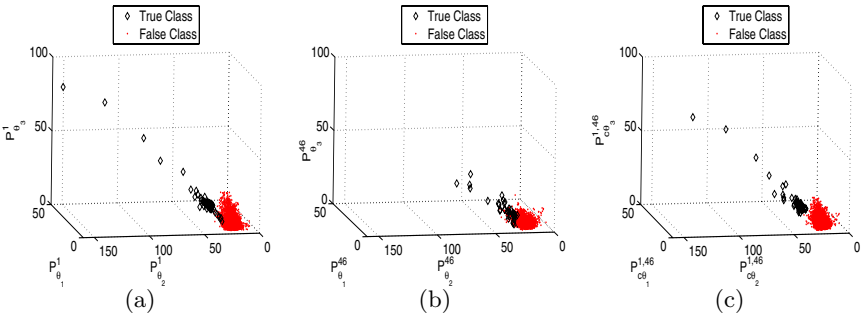


Fig. 4. Scatter plot of a person using potential field representation, obtained from θ_1=0°, θ_2=45°, θ_3=90°, and using reference face image (a) I^1, and (b) I^{46}. (c) Combining (a) and (b) using $n = 3$.

4 Autoassociative Neural Network Based Classifier

The next task is to classify a given test face image using the combined similarity score $P_\theta^{t,l,m}$. One can employ a classifier based on Multilayer Perceptron (MLP) [10] neural network model or Support Vector Machine (SVM) [10]. But, these models require samples from both true and false classes. Though one can have a large number of false class images for a given person, in practice it is not feasible to have that many face images of the true class. This problem can be overcome by exploiting the behavior of the false class face images in a scatter plot. In a scatter plot the points due to false class are more dense as compared to points due to true class. Hence, we have adopted the following strategy for classification. Take several false class face images and compute the combined similarity score ($P_\theta^{t,l,m}$) using θ= 0°, 45°, 90°, 135°. Now, capture the distribution of the four dimensional feature vector (four combined similarity scores) of false class face images. The distribution is captured using an auto associative neural network (AANN) [6] model. Thus using a suitable threshold for the output of the AANN model, a decision can be made weather to accept the claim of the test input or not.

5 Experimental Results

Here, we give a brief summary of our experiments. The block diagram of the training phase is shown in Fig. 5(a). In this block diagram we have shown training with five reference face images I^1, I^{46}, I^{91}, I^{136}, I^{181}. This process can be

generalized for any number of reference face images. The reference face images are chosen, in such a way that their poses are uniformly spaced over the span of $0° - 180°$. Several false class images are chosen and their potential field representations (u_θ) are computed along four directions (θ=0°, 45°, 90°, and 135°). These representations are correlated with the corresponding representation of each reference face image. The resultant correlation output is used to compute the similarity score. Hence, five sets of four dimensional feature vectors (4 PSR values) are obtained for each false class image. The similarity scores obtained from the reference face images which have adjacent pose are combined using (2), as shown in Fig. 5(a). These combined scores are presented to an AANN model for training. Let $AANN^{1,46}$ denote the AANN model trained with the combined similarity scores ($P_\theta^{t,1,46}$) obtained using the reference face images I^1 and I^{46}. The structure of the AANN model is $4L\,8N\,2N\,8N\,4L$, where L denotes a linear unit, and N denotes a nonlinear unit. The AANN model is trained using back-propagation algorithm for 3000 epochs. Similarly we have designed $AANN^{46,91}$, $AANN^{91,136}$ and $AANN^{136,181}$ using the same false class face images. The block diagram of testing phase of face verification system is shown in Fig. 5(b). For a given test face image, the potential field representation u_θ is computed along four directions (θ=0°, 45°, 90°, and 135°). These representations are correlated with the corresponding representation of each reference face image of claimed identity. The resultant similarity scores are combined as in the training phase, and are presented to AANN models as shown in Fig. 5(b). The combined similarity score (4 dimensional feature vector) is used to compute the error in associating the vector with the AANN models corresponding to the reference face images. If the error is above a threshold in any one of the AANN models, the claim is accepted. Here, the threshold value for each AANN model could be different.

False acceptance (FA) and false rejection (FR) are two error metrics that are used to evaluate a face verification system. The trade off between FA and FR is a function of the decision threshold. Equal Error Rate (EER) is the value for which the error rates FA and FR are equal. Here, we will explain the computation of EER for a single person. We have used I^1, I^{46}, I^{91}, I^{136}, and I^{181} as reference face images. The remaining $176 = ((181 - 5))$ face images form the examples of the true class. For false class, $5259 = (29 \times 181)$ face images are available. Out of these, 3000 face images are used to train the $AANN^{1,46}$, $AANN^{46,91}$, $AANN^{91,136}$, and $AANN^{136,181}$ models. The remaining $2249 = (5349 - 3000)$ false class face images are used for testing. But the true class sample will be different for each AANN models, because the objective of AANN is to verify if the true class face image has a pose in a specific range. Thus for $AANN^{1,46}$, the true class samples will be the images between I^1 to I^{46}. By varying the threshold value of $AANN^{1,46}$ the Receiver Operating Characteristics (ROC) curve is obtained. In ROC curve the intersection point of FA and FR gives the EER for this model. Similarly we compute the EER for other AANN models, and the average of all them is the resultant EER for that person. The experiment is repeated for all the persons with this set of reference images, and the average of all EER values is E_{avg} is computed. The performance ((1-E_{avg})*100) [9] of

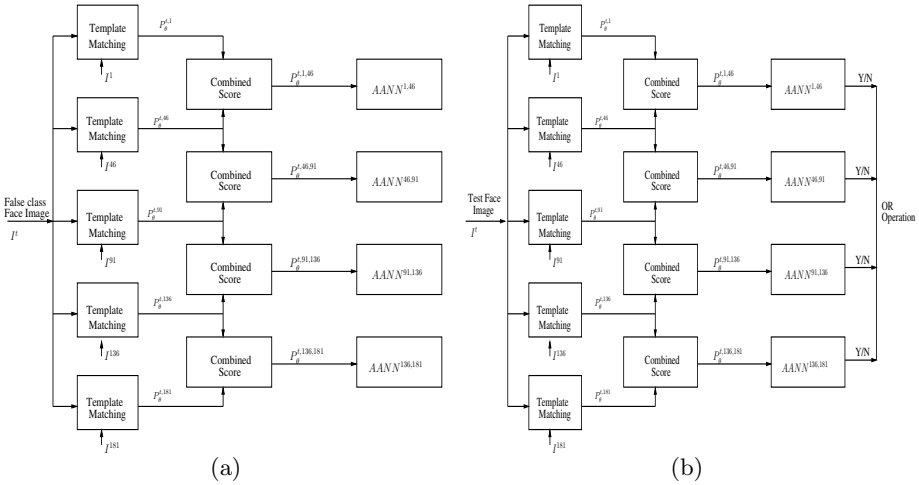


Fig. 5. Block diagram of Face verification system for (a) Training Phase, and (b) Testing Phase

the proposed method is shown in Table 1 for different sets of reference face images along with the performance obtained using existing approaches [8]. One can see from the table that proposed method performs better than the existing approaches. The reason could be that proposed method preserves the unique information of a person's face image in a given pose.

Table 1. Recognition rate (in %) for diferent sets of reference face images by the proposed method in comparison with the results for different method given in [8]

Set of reference face images	I^{91}	I^{1}, I^{91}, and I^{181}	I^{1}, I^{46}, I^{91}, I^{136}, and I^{181}
PCA	20.74	50.53	71.6
LDA	20.70	56.92	78.63
HMM	31.68	41.27	63.50
BIC	18.42	45.19	69.47
Proposed approach	**49**	**65.45**	**85.83**

6 Summary

We have proposed an approach to address the pose problem in face verification. This approach uses the given reference face images at different poses separately for template matching rather than building a model or synthesizing a face image. The template matching is performed using correlation based technique, and edge gradient based representation. The edge gradient representation cannot be used for correlation matching directly because of sparsity. This problem was overcome by spreading the edge information using potential field representation. The edge

representation was derived using 1-D processing of images, to obtain multiple partial evidences for a given image. An approach was proposed to combine the scores obtained by matching multiple partial evidences of different face images. The resultant combined scores were used to verify the identity of the person using AANN model based approach. The proposed approach for classification has the advantage that it does not require training images of the true class. Experimental results show that the proposed approach is a promising alternative to other existing approaches for dealing with the pose problem in face recognition.

References

1. Deniz, O., Castrilon, M., M, H.: Face recognition using independent component analysis and support vector machine. Pattern Recognition Letters **22** (2003) 2153–2157
2. Kim, K.I., Jung, K., Kim, K.: Face recognition using support vector machine and with local correlation kernels. Int'l Journal Pattern Recognition and Artifical Intelligence **16** (2002) 97–111
3. Sanderson, C., Bengio, S., Gao, Y.: On transforming statistical models for non-frontal face verification. Pattern Recognition **39** (2006) 288–302
4. Vetter, T., Blanz, V.: Estimating coloured 3D face models from single images: An example based approach. In: Proceedings of Conf. Computer Vision ECCV'98. (1998)
5. Sao, A.K.: Significance of image representation for face recognition. Master's thesis, Department of Computer Science and Engineering, Indian Institute of Technology, Madras (2003)
6. Yegnanarayana, B., Kishore, S.P.: AANN: an alternative to GMM for pattern recognition. Neural Networks **15** (2002) 459–469
7. Black, J., Gargesha, M., Kahol, K., Kuchi, P., Panchanathan, S.: A framework for performance evaluation of face recogniton algorithm. In: Internet Multimedia System, Boston (2002)
8. Little, G., Krishna, S., Black, J., Panchanathan, S.: A methodology for evaluating robustness of face recognition algoritms with respect to change in pose and illumination angle. In: ICASSP, Philadelhia (2005)
9. Kumar, B.V.K.V., Savvides, M., Venkataramani, K., Xie, C.: Spatial frequency domain image processing for biometric recognition. IEEE Int. Conf. Image Processing, New York (2002) 53–56
10. Haykin, S.: Neural Networks - A comprehensive foundation. Macmillan College Publishing Company Inc., New York (1994)