Proceedings of KFIS Autumn Conference 2007 Vol. 17, No. 2.

얼굴과 발걸음을 결합한 인식

Fusion algorithm for Integrated Face and Gait Identification

Imran Fareed Nizami, Sugjun Hong, Heesung Lee, Toh kar Ann, Euntai Kim and Mignon Park

School of Electrical and Electronic Engineering, Yonsei University, Seoul, Korea
Biometrics Engineering Research Center

Abstract

Identification of humans from multiple view points is an important task for surveillance and security purposes. For optimal performance the system should use the maximum information available from sensors. Multimodal biometric systems are capable of utilizing more than one physiological or behavioral characteristic for enrollment, verification, or identification. Since gait alone is not yet established as a very distinctive feature, this paper presents an approach to fuse face and gait for identification. In this paper we will use the single camera case i.e. both the face and gait recognition is done using the same set of images captured by a single camera. The aim of this paper is to improve the performance of the system by utilizing the maximum amount of information available in the images. Fusion is considered at decision level. The proposed algorithm is tested on the NLPR database.

Key Words : gait recognition, face recognition, fusion method, RM.

1. INTRODUCTION

Identification of persons has become very important from surveillance and security point of view. Different modalities can be used for identification based on the given situation. If a person is far away from the camera then the style of walking best represents the biometric trait of the person. However the inclusion of available face information along with gait can act as a powerful cue to improve the performance of the system.

Recognition based on gait offers many advantages, as gait is a non-invasive biometric. Gait offers potential for recognition at low resolution. Gait does not require close contact between the subject and the recording probe. The two approaches to gait recognition are model based and model free. Model based approaches explicitly try to model the shape of human body or the motion. The parameters for these models are determined by processing of whole gait sequences. The advantages of model based approaches are that they are scale invariant, but they are computationally very expensive. In contrast the holistic or model free approaches operate directly on the gait sequences without assuming any specific model for the gait sequence, so they are computationally very efficient as compared to the model based approaches.

The disadvantage of traditional holistic approaches is that they ignore the temporal component of gait [1]. In contrast the model based approaches aim to model accurately how a subject walks by analysis of the motion of the legs, but the drawback with this approach is the computational complexity. One of the aims of this paper is to introduce a model free approach which considers the temporal information. This is done successfully by using Moving Motion Silhouette Image (MMSI). MMSI is an instantiated version of MSI. However it was found that MMSI performs better than MSI. It embeds the spatial and temporal information of the gray scale binarized silhouette images.

The main contribution of this paper is the fusion of face and gait using MMSI images and RM classifier. RM is a classifier based on Reduced Multivariate Polynomial. It offers the advantage of being computationally very efficient. The rest of this paper is organized as follows. In section 2 we explain the face detection algorithm and feature extraction using ICA. Recognition algorithm based on multiple-views is discussed in section 3. Section 4 displays the experimental results. Lastly conclusions and future works are given in section 5.

2. FACE DETECTION AND FEATURE EXTRACTION

2.1 Face Detection

In the proposed algorithm, face is detected from the images before recognition is performed. An intuitive method for Face detection is given here. Initially background subtraction is used to get the gray silhouette images, Bhanu et al. suggested that human head is 16% of the human body [2], so the top 16% of the image is taken as head and cut off. Next left and right coordinates for the head region are calculated from the binarized silhouette images. The binarized silhouette images are obtained using normalization and background subtraction. The face images are then resized to 32*32 dimensions. The process for face detection is given in Fig 1.

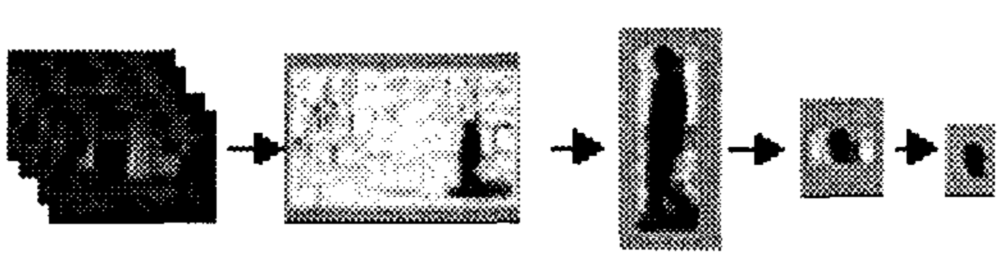

Figure 1. Face Detection Algorithm

2.2 Moving Motion Silhouette Image

An effective representation for gait was introduced in [3] called Moving Silhouette Image (MSI). MSI is a gray level image which retains the critical spatiotemporal information. The intensity of each pixel in an MSI is a function of its temporal history. MSI has high discriminating power because it represents the spatial as well as the temporal information present in a gait sequence. MSI is given by eq (1).

$$MSI(x,y,t) = \begin{cases} 255 & I(x,y,t) = 1 \\ \max(0, MSI(x,y,t-1) - 1) & otherwise \end{cases} \quad (1)$$

where I is the silhouette image, t is the current time, x and y are the horizontal and vertical coordinates of the input silhouette image respectively.

Figure 2. MSI Images for two persons

This paper introduces a variant of MSI called the Moving Motion Silhouette Image (MMSI). The basic difference between MSI and MMSI is that, the intensity value of the pixels in an MSI represents motion information of the whole gait sequence, whereas MMSI represents the motion information of only the images within the scope of the window. To generate MMSI images the window is moved forward one frame at a time until the window reaches the last frame in the gait sequence e.g. if the size of the window is chosen to be k and the total number of frames in the gait sequence are n then the first MMSI image includes the temporal information of frames from t=1,...,k, the second MMSI image has the temporal information of samples from t=2,...,(k+1) and so on, until the last MMSI image will have the temporal information from t=n-(k-1),...,n.

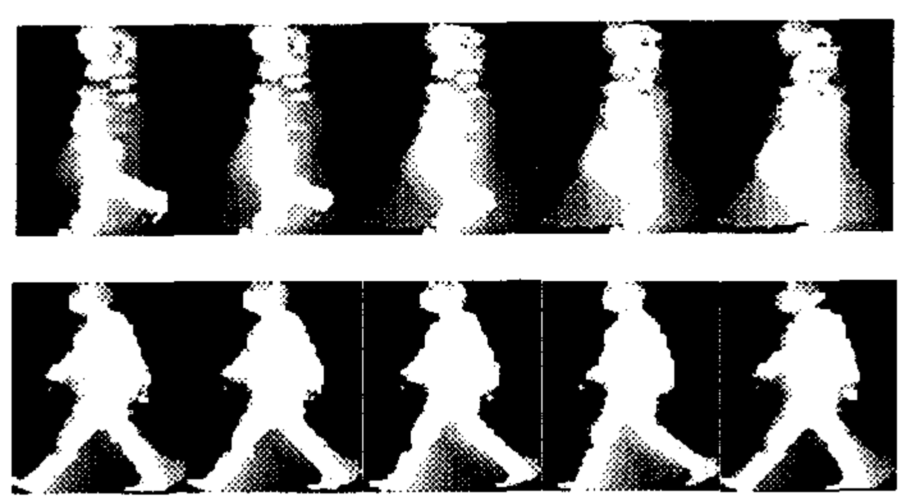

Figure 3. MMSI Images for 2 persons

2.3 Feature Extraction

Once face is detected and gait is represented by MMSI, the next step is dimension reduction and feature extraction.

PCA and ICA have been used extensively as the feature extraction and dimensionality reduction techniques for face recognition. Marian et al. and Jian et al. found in [4, 5] respectively that ICA significantly outperforms the standard PCA in face recognition when pose variation exists. This fact justifies the use of ICA for feature extraction in this paper. ICA is represented by the following liner input mixture model.

$$X = MS \tag{2}$$

where S is the m-dimensional random vector, X is a sequence of n-MMSI images and M is a full rank (n,m) mixing matrix. S and M are estimated by using X as the input. Once we know M we can compute its inverse A and get the feature vector as

$$S = AX \tag{3}$$

3. RECOGNITION ALGORITHM

The proposed algorithm can be divided into, three steps. (a) Face Recognition, (b) Gait Recognition, (c) Fusion of face and gait. Fig 4 shows the flow chart for the proposed algorithm.

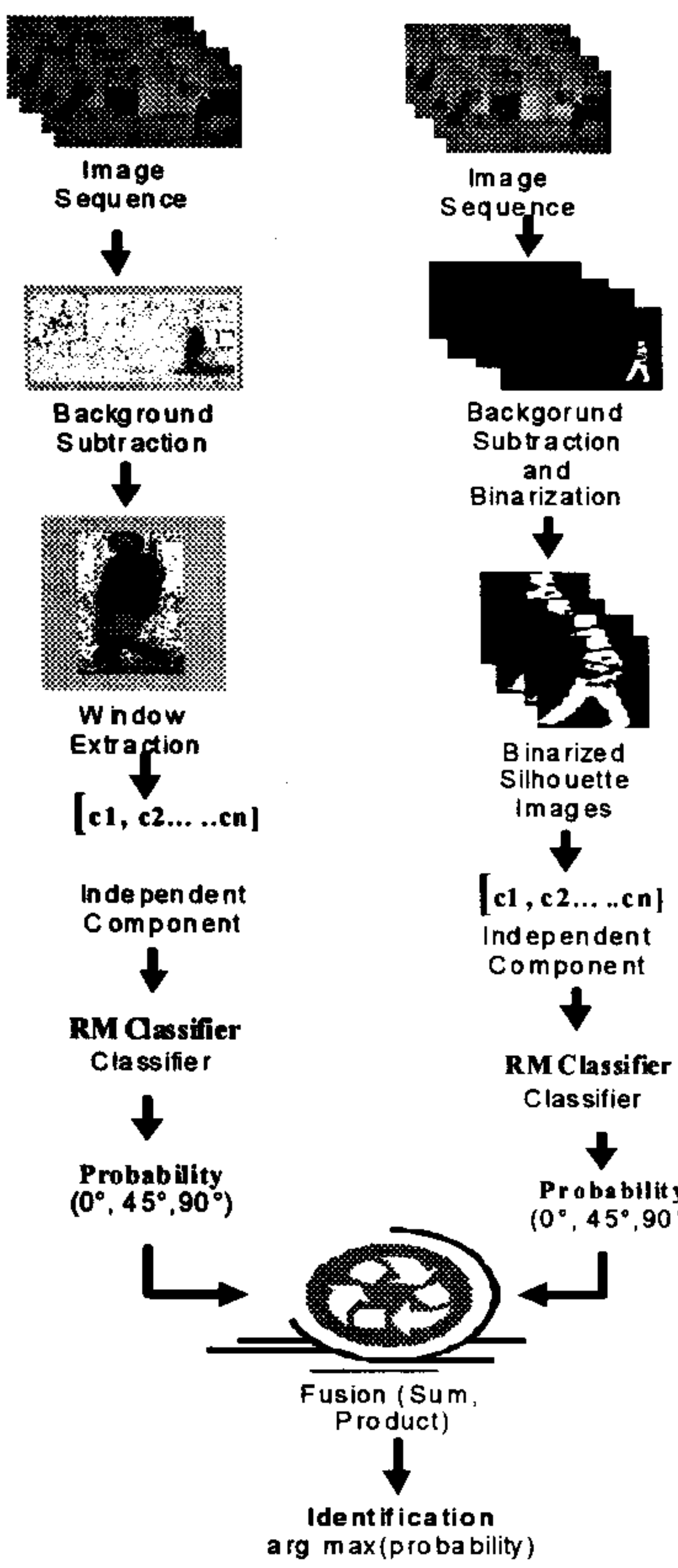

Figure 4. Proposed Algorithm

RM is based on a Reduced Multivariate Polynomial [7], the method is particularly suitable for problems with small no of features and large number of examples. The basic operation is as follows:

$$a = (P^t P + bI)^{-1} P^t y \tag{4}$$

where P denotes the Jacobean feature matrix for face or gait and y has the target values between [0, 1]. Value between 0 and less than 0.5 constitutes an imposter and a value greater then and equal to 0.5 and less than +1 constitutes the genuine user. Once we have α then testing can be done by

$$z = P_{test} a \tag{5}$$

Once we have the probability z then we can use this probability to identify the individual. We fuse the gait and face using product and sum rules. Let us suppose that the probability that z_{face} is the probability that the face test image has identity n^i and z_{gait} is the probability that the test MMSI image has identity n^i, where n^i is the identity of ith subject in the database. Then the identification based on product and sum rules can be given by eq (6) and (7) respectively.

$$ID^{pr}_{fusion} = \text{argmax}(z_{face}(w|n^i) \times z_{gait}(w|n^i)) \tag{6}$$

$$ID^{su}_{fusion} = \text{argmax}(z_{face}(w|n^i) + z_{gait}(w|n^i)) \tag{7}$$

4. EXPERIMENTAL RESULTS

The algorithm is evaluated on the well known NLPR gait database [8]. The database includes 20 subjects and 4 sequences for each subject. There are a total of 3 views, so a total of 240 sequences are present (20 x 4 x 3=240). The experimental setup is as follows. 10 fold cross validation was used to calculate the results i.e. 63 images corresponding to each person were chosen for training and 7 images were used for testing. For each person probability z is calculated. After the calculation of probability the identity with the maximum probability is assigned to the person as the recognized identity as explained in previous section. The recognition rate is calculated by taking the average of the individual recognition rates of all the 10 runs.

Table 1. Recognition Rate.

Regulating Factor (b)	RM Sum		RM Prod	
	Features used		Features used	
	31.75%	63.5%	31.8%	63.5%
0.1	84.71	94.86	84.3	94.86
0.01	83.57	94	83.14	93.57
0.001	72.57	80.43	71.57	81.3
0.0001	70	67	70.43	69.3

Table 1 shows that the proposed algorithm performs very well. By using only 64% of the feature space the proposed algorithm gives recognition rates of up to 95%. It can be observed that the performance of the system is improved as we use a larger portion of the feature space. The other parameter that affects the performance of the system is the regulating factor b. It is observed from Table 1 that these fine tuning parameters affect the performance of the system adversely. We use regulating factor b ranging from 0.1 to 0.0001. It was observed that a regulating factor of 0.1 gave the best results. We used two well known fusion rules for combining gait and face. Both rules i.e. product and sum rule give similar results.

5. CONCLUSION AND FUTURE WORK

The proposed algorithm performs well. The fusion of gait and face utilizes the maximum amount of information in the images and improves the performance from using either trait individually. But the proposed algorithm is tested on the NLPR database, which is a small. To check the practical application of the proposed algorithm it has to be checked on a larger database. Furthermore, fusing the three views i.e. 0°, 45° and 90° to get at decision level to get one decision may improve the performance.

ACKNOWLEDGMENT(S)

This work was supported by the Korea Science and Engineering Foundation (KOSEF) through the Biometrics Engineering Research Center (BERC) at Yonsei University.

REFERENCES

[1] J. P. Foster, M. S. Nixon, and A. Prugel-Bennet, "Automatic gait recognition using area-based metrics," *Pattern Recognition Letters*, vol. 24, pp. 2489 - 2497, 2003.

[2] X. Z. Bhanu, "Integrating Face and Gait for Human Recognition" *Computer Vision and Pattern Recognition Workshop*, pp. 55-55, Jun., 2006.

[3] T. Lam and R. Lee, "A New Representation for Human Gait Recognition: Motion Silhouette Image MSI," *Intl. Conf. on Biometrics 2006, LNCS*, pp. 612-618, 2005.

[4] C. Shi, H. Li, X. Lian, and X. Li, "Multi-Resolution Local Moment Feature for Gait Recognition", *Conf. 5th Machine Learning and Cybernetics*, pp. 3709-3314, Aug., 2006.

[5] M. Bartlett, J. Movellan and T. Sejnowski", Face Recognition by Independent Component Analysis", *IEEE Trans. Neural Networks*, Vol. 13, No. 6, pp. 1460-1464, Nov., 2002.

[6] F. Esposito, T, Scarabino, A. Hyvarinen, J. Himberg, E. Formisano, S. Comani, G. Tedeschi, R. Goebel, E. Seifritz, and F. Di Salle, "Independent Component Analysis of fmri Group Studies By Self-Organizing Clustering", *Neuroimage*, vol. 25, pp. 193-205, 2005.

[7] K. A. Toh, Q. L. Tran and D. Srinivasan, "Benchmarking a Reduced Multivariate Polynomial Pattern Classifier," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 740-755, Jun 2004.

[8] L. Wang, T. Tan, W. Hu, and H. Ning, "Automatic gait recognition based on statistical shape analysis", *IEEE Trans. Image Processing*, vol. 12, no. 9, 2003.