Person Identity Verification Based on Multimodal Face-Gait Fusion

Emdad Hossain and Girija Chetty

University of Canberra, University Drive, ACT, Australia

Summary: In this paper we propose a novel approach for ascertaining human identity based on fusion of profile face and gait biometric cues The identification approach based on feature learning in PCA-LDA subspace, and classification using multivariate Bayesian classifiers allows significant improvement in recognition accuracy for low resolution surveillance video scenarios. The experimental evaluation of the proposed identification scheme on a publicly available database [2] showed that the fusion of face and gait cues in joint PCA-LDA space turns out to be a powerful method for capturing the inherent multimodality in walking gait patterns, and at the same time discriminating the person identity.

Keywords: Biometrics; gait recognition; PCA; LDA; Eigenface, Fisherface, Multivariate Gaussian Classifier,

1. Introduction

Person identity verification from arbitrary views in low resolution surveillance video is a very challenging problem, especially when one is walking at a distance. Over the last few years, recognizing identity from gait patterns has become a popular area of research in biometrics and computer vision, and one of the most successful applications of image analysis and understanding. Gait recognition is one of the new and important biometric technologies based on behavioural characteristics, and it involves identifying individuals by their walking patterns. Gait can be captured at a distance by using low resolution devices, while other biometrics need higher resolution. Also, gait is difficult to disguise, and can be performed at a distance or at low resolution, and requires no body-invading equipment to capture gait information. Gait recognition can hence be considered as a next-generation identity verification technology, with applicability to many civilian and high security environments such as airports, banks, military bases, car parks, railway stations etc. Further, gait is an inherently multimodal biometric as proposed by Murray et. al in[1], suggesting that there are 24 different components to human gait, and involves not just the lower body or legs, but also the upper body in terms of motion associated with the torso, the head and the hands. If all gait movements from full body images can be captured, it can be a truly a unique biometric.

In this paper we propose a multimodal fusion technique by combining face and gait features in learning subspaces based on principal component analysis (PCA) and linear discriminant analysis (LDA). Further, by processing the fusion features with multivariate Gaussian classifiers, it is possible to capture several inherent multimodal components present in human Gait. Extensive experiments conducted on a publicly available gait database [2] suggest that to obtain optimal performance, a integrated face, body and gait cues obtained from video sequences and processed appropriately with learning approaches mentioned above, can result in a simple, practical and robust identity verification technique in spite of poor quality data from surveillance video with significant degradations in operating environments.

Rest of the paper is organized as follows. Next Section discusses the background and the previous work, followed by our motivation for the proposed scheme in Section 3. In Section 4 we describe the identification scheme and the details of the experimental work carried out is discussed in Section 5. The paper concludes in Section 6 with conclusions and plan for further work.

2. Background

The recognition of people is of great importance, since it allows us to have a greater control about when a person has access to certain information, area or simply to identify if the person is the one who claims to be [3,4]. And one natural tool to identify a person is the biometric trait. Automated face recognition technology [3,4,5] first captured the public attention from the media reaction to a trial implementation at the January 2001 super bowl, which captured surveillance image and compared them to a database mug shots [5]. From 1960s till now vast number of research works have been conducted on biometric person authentication. Several research articles have been reported involving the use of signatures, fingerprints, face and voice information [6]. For face recognition systems, the performance of 2D face matching systems depends on capability of being insensitive of critical factors such as facial expression, makeup and aging, but also relies upon extrinsic factors such as illumination difference, camera viewpoint, and

Manuscript received June 5, 2011
Manuscript revised June 20, 2011

scene geometry [7]. Further, the 2D face recognition systems are vulnerable to pose, and illumination variations. Use of 3D face can make systems robust to pose and illumination variations. The state of the art 3D face recognition technique using isogeodesic stripes was proposed in [7], However, progress in 3D identification approaches has been slow as it suffers from higher infrastructure costs and limited availability of databases [8]. Hence, alternate biometric traits and combination of different types of biometric traits was explored by the biometric security community, Subsequently, due to increase in global demand for automated security and surveillance products, there was a proliferation of research works on identity verification based on different biometric modalities [9], [10], and [11]. Several research works have also reported importance of using multiple modalities instead of single biometric trait in order to enhance the accuracy and robustness [10], [11] and [12]. However, most of the systems have been tested in controlled laboratory environments, and it is a huge challenge to achieve similar accuracies and robustness in real world public surveillance applications. Further, the current generation of identity authentication systems are based on modalities based on fingerprint, palm print, face, iris, ear biometric traits [13], [14] and [15]. These modalities are of limited use for deployment in public surveillance scenarios or performing authentication at a distance.

Lately, an increasing need for surveillance in public places and facilities from a distance has been felt due to terrorist attacks or attack on public assets. And the automated video surveillance systems serve as the first line of defence for protecting assets and people for different types of operating scenarios and applications – be it a civilian public space for access control to a facility, or financial and transaction oriented applications, or the high security immigration and border control check points. It has become an enabler of trust, integrity and security in the new Digital Economy [16], [17] and [18].

However, the surveillance systems designed to work in high security environments fail miserably when deployed for day-to-day civilian environments, due to unconstrained noisy and non-ideal operating conditions of civilian environments. Integrating multiple sources of information can solve some of the problems with these systems; and this integration could be at a lower level involving sensors, data and feature extractors, or at a higher level, at decision or at a score level.; or it could also be at an ancillary or side level, consisting of higher level context information Humans have long been using such multiple levels of information sources as cues to perform any identification tasks, particularly in difficult scenarios. This then suggests that there is an information fusion from multiple information sources which equips humans with better recognition capabilities as compared to any other species. Likewise for automatic person recognition systems, if multiple heterogeneous sources of information could be combined and used - for example, the side-views of the face, the facial gestures, limb movements, or the walking gait pattern, which may be least discriminative and on their own cannot be used for establishing the identity of the person, but can perhaps provide a unique form of contextual information - it could be possible to enhance the performance of automatic identity recognition task in video surveillance systems operating in unconstrained civilian operating environments. The non-dominant secondary information is normally captured without extra burden on the system, and is available as rich multimedia synchronized data from the same CCTV images as the primary facial. The proposed multimodal approach in this paper is based on side face and gait, and both can be extracted from low resolution imagery. And it is not necessary to have clear face or gait, making it suitable to collect data irrespective of user's disable, aged or any gender. The potential of gait as a powerful biometric has been explored in some of the recent works [18, 21], though inherent multimodal components present in the whole body during walking has not been much exploited by the research community. In this paper we investigate the potential of combining rich multimodal information available form face body (torso) and gait biometric cues, to ascertain the human identity in low resolution surveillance videos with unconstrained operating environments.

Face and gait based identity verification serves another important purpose – and that is of addressing sensitive privacy issues associated with capture and storage of biometric data. Some of the most important challenges for diffusion of biometrics in day-to-day civilian applications are issues related to invasion of privacy. In [19], an extensive study has shown that physiological biometrics as having no negative impact whatsoever on the privacy. That is an excellent motivation for us to investigate gender-specific face, body(torso) and gait cues during walking as a powerful biometric with inherent multimodality for ascertaining the identity of a person. Further, these video based cues can be captured remotely from a distance, and by using an appropriate biometric identification protocol as suggested by authors in [22], it can be ensured that sensitive privacy concerns are addressed as well. An appropriate protocol as in [22] can ensure that the identification system is not misused and that function creep (i.e. use for another purpose is prevented). This means in particular that a component should not be able to learn more information than what is really needed for a correct result. In fact our proposed fusion of face and gait cues captured from low resolution surveillance videos ("security check: pass") needs strong algorithms and processing techniques to be of any use for establishing identity, and of no use without them, and hence can safe-guard the privacy to a large extent automatically.

Next few sections describe out proposed multimodal fusion approach based on face and gait cues, the details of the publicly available gait database used for this research, and the results obtained from the experiments.

3. Motivation for Face-Gait Fusion

Since the performance of any classifier is more sensitive to some factors and relatively invariant to others, a recent trend has been to combine individual classifiers in order to integrate their complementary information and, therefore, create a system that is more robust than any individual classifier to variables that complicate the recognition task. In [22], researchers have shown that integrating multiple biometrics does indeed result in consistent performance improvement while in [23], authorss have empirically demonstrated that as the number of classifiers combined increases, so does the recognition accuracy. The encouraging results reported in the literature for such systems in conjunction with the conclusions reached by the previous studies mentioned in the background section of this paper, provide a strong basis for believing that a system constructed by combining various biometric characteristics is going to yield better recognition rates than the individual classifiers for those traits. Further, using multiple biometrics is a viable solution to real-world problems, such as non-universality of some biometric traits (e.g., some people's fingerprints cannot be reliably extracted because of the poor quality of the ridges also possible to do fingerprint alteration [16]), unavailability of data for a certain biometric (e.g., visual cues such as face, ear, etc. might be occluded in surveillance videos) and criminal activity (i.e., attempts to fool the single-biometric based system by duplicating the biometric trait or breaching the system).

In light of the above, some specific reasons for considering investigation of face and gait biometric fusion are as follows:

- The face is a short-range biometric, which can be used effectively for identification only when the subject is close enough to the camera for sufficient details of subject's facial features to be captured.
- Gait, on the other hand, is a medium to long-range biometric, which can be extracted reliably even from low-resolution imagery and is more invariant to slight changes in viewpoint. Researchers in [3] suggested finding invariant representation from inherently varying biometric signal (profile/side face and gait for example), by using an appropriate digital representation, such that the trait can be recognized despite changes in pore, illumination expression, aging and so on [3].
- Using these two biometric traits together would arguably make the system more robust to variations in subject to camera distance. Also, both face and gait are visual cues; both can be extracted from the same modality, (i.e., image sequences of people) precluding the need for separate or specialized equipment.
- Further, the face and gait biometrics make use of apparently independent personal characteristics: face recognition systems exploit the relatively detailed appearance of the facial surface, while gait recognition methods capture data from the coarse body shape as it changes over time. Consequently, some conditions that sharply degrade the performance of face recognition systems, such as large variations in illumination and facial expressions, affect gait to a much lesser extent or not at all. Similarly, some conditions that adversely affect the accuracy of gait recognition, such as clothing, footwear, and load, do not influence the performance of face recognition systems. Therefore, it is reasonable to believe that combining these complementary cues would improve the recognition accuracy.

4. Proposed Multimodal Scheme

For experimental evaluation of our proposed face and gait fusion scheme, we used a publicly available video database of human actions [2]. This video database contains six types of human actions (walking, jogging, running, boxing, hand waving and hand clapping) performed several times by 25 subjects in four different scenarios: outdoors s1, outdoors with scale variation s2, outdoors with different clothes s3 and indoors s4. Currently the database contains 2391 sequences.

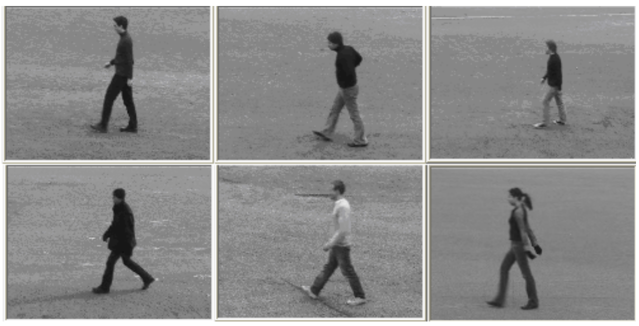

Fig. 1. Sample images from human action database for walking sequences [2].

Fig. 2 Sample images from multiple frames of a single person's walking sequence [2].

All sequences were taken over homogeneous backgrounds with a static camera with 25fps frame rate. The sequences were down-sampled to the spatial resolution of 160 × 120 pixels and have a length of four seconds in average. We used only the walking sequences for our experiments. Figure 1 shows some of the sample images from the walking video sequences., and Figrue 2 shows multiple frames of the sequences for a person walking in the video clip.

For all our experiments we used 100 video sequences for 25 people. There were 19 males and 6 females in the entire walking dataset. We performed some image pre-processing steps corresponding to cropping, filtering and histogram equalization and then extracted features based on PCA (principal component analysis) and LDA (linear discriminant analysis). We used separate set for performing training and testing. The low dimensional PCA and LDA features were then classified by a Bayesian classifier. We examined four different classifiers, the nearest neigbour (*k-NN*), the Bayesian linear and the Bayesian quadratic classifiers, and the Mahalanobis classifier. The combination of the low dimensional, discriminative PCA and LDA features along with powerful multivariate Bayesian linear/quadratic classifiers allow us to achieve significant improvement in recognition accuracy as compared to conventional Euclidean distance based methods reported predominantly in previous works.

This is because Bayesian classifiers have the flexibility to incorporate prior information, and can predict how a system's performance will change when there is a mismatch in train and test conditions. [24 - 27]. And *k-NN* is very effective simple classifier with noise reduction capabilities [24 -27]. The schematic for the proposed multimodal identification scheme is shown in Figure 3. A brief description of PCA and LDA feature processing technique is given next.

4.1 PCA Features

PCA is a useful statistical technique that has found application in fields such as face recognition and image compression, and is a common technique for finding patterns in data of high dimension [24-27].

Let $\{\mathbf{x}1, \mathbf{x}2, \ldots, \mathbf{x}n\}$, $\mathbf{x}k \in RN$, be n random vector, where N is the dimensionality of the vector obtained by concatenation of an image row-by-row. The covariance matrix is defined as $\sum\mathbf{x}=E([\mathbf{x}-E(\mathbf{x})][\mathbf{x}-E(\mathbf{x})]T)$, where $E(\cdot)$ is the expectation operator and T denotes the transpose operation. The covariance matrix $\sum\mathbf{x}$ can be factorized into the following form:

$$\sum x=\Phi\wedge\Phi \text{ .... (1)}$$

where $\Phi = [\Phi\ 1,\ \Phi\ 2\ \ldots\ \Phi\ N] \in RN\times N$ is the orthogonal eigenvector matrix of $\sum x$; $\wedge=\{\wedge 1\ \wedge 2\ \ldots\ \wedge N\} \in RN\times N$ is the diagonal Eigen value matrix of $\sum x$ with diagonal elements in descending order [24-27].

One important property of PCA is its optimal signal reconstruction in the sense of minimum mean square error (MSE) when only a subset of principal components are used to represent the original signal. An immediate application of this property is the dimensionality reduction [24-27]:

From Video Sequences

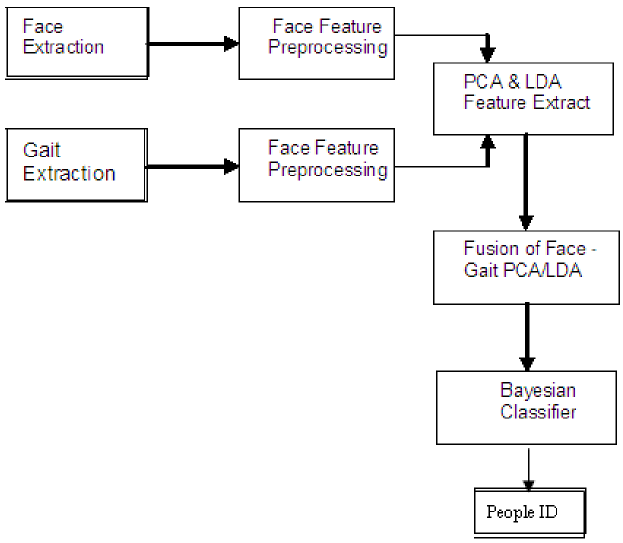

Fig. 3. Schematic for proposed multimodal identification scheme based on fusion of side face and gait cues extracted from low-resolution video.

$$\mathbf{y}k = \mathbf{P}T\, pca[\mathbf{x}k - E(\mathbf{x})],\ k= 1, 2, \ldots n, \text{ .... (2)}$$

where $\mathbf{P}pca = [\Phi\ 1, \Phi 2\ \ldots\ \Phi\ m], m \le N$. The lower dimensional vector $\mathbf{y}k \in Rm$ captures the most expressive features of the original data $\mathbf{x}k$ [19, 25, 26].

If **f** ∈ RN1 and **g** ∈ RN2 represent the PCA vectors corresponding to a person video data, where N1 and N2 are the dimensionality of the face and the gait feature spaces, respectively.
We obtain low dimensional feature vectors,

f′ = **M**f **f** and **g′** = **M**g**g**,

by using the PCA method as in Eq. (2).

Mf and **M**g are the PCA transformation matrices for face and gait, respectively. We choose a subset of principal components to derive the lower dimensional face and gait features, **f′**∈ Rm1 and **g′**∈ Rm2 ,where m1 and m2 are the dimensionality of the reduced face feature space and gait feature space, respectively.
On one hand, we hope to lose as little representative information of the original data as possible in the transformation from the high dimensional space to the low dimensional one. On the other hand, the eigenvectors corresponding to the small Eigen values are excluded from the reduced space so that we can obtain more robust PCA & LDA projection as well as reduce the problem of curse of dimensionality [24-27]. The Eigen value spectrum of the covariance matrix of the training data supplies useful information regarding the choice for the dimensionality of the feature space. Before face features and gait features are normalized to have their values lie within similar ranges. We use a linear method [24-27], which provides normalization via the respective estimates of the mean and variance. For the jth feature value in the ith feature vector w_{ij}, we have:

$$\hat{w}_{ij} = (w_{ij} - w'_j) / \delta_j,$$

$$i = 1, 2, \ldots, I, \quad j = 1, 2, \ldots, L, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

where $w_j = (1/I) \sum_{i=1}^{I} w_{ij}$ and S_B . I is the number of available
feature vectors and L is the number of features for each feature vector. The resulting normalized features have zero mean and unit variance. To take advantage of information for a walking person in video, we use all possible combinations of complete images, side face features and gait features to generate the maximum number of vectors **h**. specifically; we have two feature vectors of side face and two feature vectors of gait for one person from one video. Therefore, we have four concatenated features **h** for one person from one video. Generation of all possible low dimension vectors **h** from PCA analysis for side face and gait data helps to reduce the problem of curse of dimensionality for the subsequent LDA transformation [24-27].

4.2 LDA Transformation and Multimodal Fusion

Suppose that $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_c$ and $n_1, n_2, \ldots, n_c$ denote the classes and the number of concatenated feature vectors h within each class, respectively,
with **w** = **w**1 ∪ **w**2 ∪ · · · ∪ **w**c and n^ = n1 + n2 +· · ·+nc. Note that the value of n^ is two times of n. c is the number of classes. LDA seeks a transformation matrix **W** that maximizes the ratio of the between-class scatter matrix S_B to the within-class scatter matrix

$$\sum_{i=1}^{c} ni(Mi - M)(Mi - M)^T \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

S_W=J(W)|WTS$_B$W|/|WTS$_W$W|.

The within-class scatter matrix is

$$\mathbf{Sw} = \sum_{i=1}^{c} \sum_{h \varepsilon w_i} (h - Mi)(h - Mi)^T$$

and the between-class scatter matrix is

$$\mathbf{S_B} = \sum_{i=1}^{c} ni(Mi - M)(Mi - M)^T , \ldots\ldots\ldots\ldots\ldots (5)$$

where **M**i = $M_i = (1/ni)\sum_{h\varepsilon w_i}$
and $M = (1/n^\wedge)\sum_{h\varepsilon w} h$
are the means of the class i and the grand mean, respectively.

We use all possible combinations of side face features and gait features to generate the maximum number of concatenated feature vectors based on the characteristics of face and gait. Specifically, four concatenated features are constructed based on two face features and two gait features for one person from each video.

Let **V**i , i = 1, 2, . . . , c, the mean of the training synthetic features of class i, be the prototype of class i. The unknown person is classier to class K to whom the synthetic feature **p** is the nearest neighbour [19]:

$$\|p - VK\| = \min \|p - Vi| \ldots\ldots\ldots\ldots\ldots (6)$$

When multiple synthetic features are obtained for one person, Eq. (6) means that the unknown person is classified to the class which has the minimum distance out of all the distances corresponding to all the classes [19].

Instead of using traditional Euclidean distance based classifiers we use learning classifiers (Bayesian linear and quadratic classifiers). The Bayesian linear and quadratic discriminant classifier uses Bayesian decision rule for classifying a set of learned feature vectors to a class [27]. While the linear classifier fits a multivariate normal density to each group, with a pooled estimate of covariance, the quadratic discriminant classifier fits MVN (multivariate normal) densities with covariance estimates stratified by group. Both methods use likelihood ratios to assign observations to groups. Given a set of classes M characterized by a set of known parameters in model Ω a set of extracted feature vector X belongs to the class which has the highest probability. This is shown in Eq.(7)) and is known as Bayes decision rule.

$$X \in M_k \qquad P(M_k|X,\Omega) \geq P(M_l|X,\Omega) \qquad \forall l \neq k \qquad \ldots \ldots\ldots\ldots\ldots(7)$$

To calculate the *a*-posteriori probability shown, we used Bayes law of statistics which finally by assuming that features are distributed normally, leads to a quadratic classifier format known as Bayes Quadratic classifier [26]. The model Ω consists of the mean and the covariance of our training vectors, and likelihoods are calculated as stated above. The details of the experiments carried out in the next Section.

5. Experiments and Results

We performed different sets of experiments for examining the discriminating ability of proposed feature extraction /transformation and classifier techniques. Initial experiments involved creating separate gender-specific datasets and examining the performance of each experiment. However, the performance was relatively poor due to imbalance in data available in number of females in the data set. Hence they are reported in this paper. For each experiment we used datasets corresponding to face-only partial gait (lower body) and full gait (full images) for all 25 people for examining the performance of single mode and fusion of PCA and LDA features for different types of classifiers.

5.1 Recognition Performance With PCA-Features

For the first set of experiments we applied PCA transformation and performed classification with Bayesian (linear/quadratic) and k-nearest neighbour classifiers. Table 1 shows the recognition accuracies achieved for PCA only features. For this experimental scenario, we received 80% recognition accuracy for Bayesian-linear classifier, 90% accuracy for Bayesian quadratic and 1-NN classifier. Though we expect a 100% accuracy for face-only mode, what we found was that quality of side face images was very poor, resulting in failure to recognize some poor quality faces. However, PCA was still able to model the low resolution side faces pretty good. This reduction of recognition accuracy is expected as PCA cannot capture the dynamic gait variations accurately no matter how efficient classifier is, The recognition performance gets worst for partial gait images (40 - 60%) and by including full gait images, it is slightly better (60-70%), perhaps because of inclusion of torso in full images. (The persons wore same clothes in train and test sessions).

Next, we performed experiments with fusion of face and gait (both partial and full gait images), and recognition accuracies achieved is shown in Table 2. An improvement in recognition performance was achieved with face-partial gait fusion resulting an accuracy of 70% for all three classifiers, and face-full gait fusion resulting in further improvement of accuracies (60 – 90%), with multivariate Bayesian Classifier performing the best with an accuracy of 90%.

Table 1: PCA with Bayesian Classifiers and 1-Nearest Neighbour Classifier

Name	**Face-Only**	**Partial-Gait**	**Full-Gait**
1-NN classify	90%	55%	70%
Bayesian-linear	80%	40%	60%
Bayesian-quadratic	90%	65%	65%

Table 2: PCA with face-gait fusion with Bayesian Classifiers and 1-Nearest Neighbour Classifier

Name	**Face-PartialGait**	**Face_FullGait**
1-NN classify	70%	85%
Bayesian-linear	70%	60%
Bayesian-quadratic	70%	90%

Table 3: LDA with Bayesian Classifiers and 1-Nearest Neighbour Classifier

Name	**Face-Only**	**Partial-Gait**	**Full-Gait**
1-NN classify	95%	80%	85%
Bayesian-linear	95%	90%	90%
Bayesian-quadratic	95%	95%	90%

Table4: LDA face - gait fusion with Bayesian Classifiers and 1-Nearest Neighbour Classifier

Name	**Partial-Gait**	**Full-Gait**
1-NN classify	95%	100%
Bayesian-linear	100%	100%
Bayesian-quadratic	100%	100%

5.2 Recognition Performance With PCA-LDA - Features

For this set of experiments, we transformed the PCA vectors in LDA space, and there was a significant improvement in recognition without fusion and with fusion. The results without fusion is shown in Table 3, and as can be seen from this table, even the gait only modes (both partial and full gaits) resulted in good accuracies (80% – 95%), with Bayesian quadratic classifier performing the best, perhaps due to its ability to model the nonlinear gait dynamics accurately.

When we performed the face and gait fusion of LDA transformed features, we got a remarkable improvement in accuracies with all three types of classifiers resulting in 100% accuracy. Thus a combination of PCA_LDA processing along with efficient classifiers, it was possible to identify a walking human from a distance even in low resolution video with poor backgrounds. Further, for all modes the multivariate classifier, particularly the quadratic one performs better as compared to 1-NN classifier used by several earlier reported studies. Also, we found the LDA has a remarkable capability to model the gait variations in the person and retain the identity specific information.

Figure 4 shows the first 8 most significant Eigen Images of faces, partial gaits and full gaits and Figure 5 shows that most significant Fisher Images of faces. partial gaits and full gaits.

6. Conclusions and Further Plan

In this paper we propose a novel multimodal identification approach based on fusion of face and gait biometric cues form low resolution surveillance videos. The proposed approach based on transforming the features in PCA-LDA subspace, and classification with Multivariate Guassian (linear and quadratic classifiers. The experimental evaluation of the proposed scheme on a publicly available database [2] showed that the combined PCA-LDA approach turns out to be a powerful method for capturing the inherent multimodality in walking gait patterns and at the same time discriminating the identity from low resolution video with noisy backgrounds. Further work involves carrying out experiments with person wearing different clothes and exploring novel methods for identity verification for unconstrained operating environments with less training data.

References

[1] Murray, M.P., Drought, A.B., and Kory, R.C , “Walking pattern of movement”, American Journal Medicine, vol.46, no. 1, pp.290-332, Jan. 1967.

[2] Schuldt Christian, Laptev Ivan and Caputo Barbara, “Recognizing Human Actions: A Local SVM Approach”. In Proceedings ICPR 2004. ftp://ftp.nada.kth.se/CVAP/users/laptev/icpr04actions.pdf (retrieved on 28th May 2011).

[3] J.N.Pato and L..I.Millett, Editor: Biometric Recognition: Challenge and opportunities, Research opportunities and the future of biometrics, National Research Council of the national Academics, Copyright 2010 by the national Academy of Science. All right reserved.

[4] D. Sanchez and P. Melin, Modular neural network with fuzzy integration and its optimization using genetic algorithms for human recognition based on iris, ear and voice

[5] J. C. Vazquez, M. Lopez and P. Melin, Real Time Face Identification Using a Neural Network approach, soft comp. for regcogn, based on biometric, SCI 312, pp. 155- 169, @ Springer-Verlag Berlin Heidelberg 2010

[6] F. Gaxiola, P. Melin and M. Lopez, Modular Neural Network for Person Authentication Using Counter Segmentation of the Human Iris Biometrics Measurement, Soft Comp. for Revogn, Based of Biometric, SCI 312, pp. 137-153, @ Springer-Verlag, Berlin Heidelberg 2010

[7] S. Berretti, A. Bimbo and P. Pala, 3D face recognaition using isogeodesic stripes, IEEE transaction on pattern analysis and machine intelligence, vol. 32, no. 12, December 2010

[8] Human Face Recognition, Advantages and disadvantages of 3D face recognition, http://www.tutorial.freehost7.com/human_face_recognition/biometrics_and_human_biometrics.htm

[9] A.K. Jain, Next Generation Biometrics, Department of Computer Science & Engineering, Michigan State University, Department of Brain & Cognitive Engineering, Korea University, December 10, 2009.

[10] S. Bengio and J. Mariethoz, Biometric Person Authentication IS A Multiple Classifier Problem, Google Inc, Mountain View, CA, USA, bengio@google.com, IDIAP Research Institute, Martigny, Switzerland, marietho@idiap.ch

[11] G. Shakhnarovich T. Darrell, On Probabilistic Combination of Face and Gait Cues for Identification, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 200 Technology Square, Cambridge MA 02139, ƒgregory,trevorg@ai.mit.edu

[12] M. N. Eshwarappa and M. V. Latte, Bimodal Biometric Person Authentication System Using Speech and Signature Features, International Journal of Biometrics and Bioinformatics, (IJBB), Volume (4): Issue (4).

[13] B. Son and Y. Lee, Biometric Authentication System Using Reduced Joint Feature Vector of Iris and Face, Division of Computer and Information Engineering, Yonsei University, 134 Shinchon-dong, Seodaemoongu, Seoul 120-749, Korea, {sonjun,yblee}@csai.yonsei.ac.kr. T. Kanade, A. Jain, and N.K. Ratha (Eds.): AVBPA 2005, LNCS 3546, pp. 513–522, 2005.@ Springer-Verlag Berlin Heidelberg 005)

[14] N. B. Boodoo, R. Subramanian, Robust Multi-biometric recognition Using Face and Ear Images, (IJCSIS) International Journal of Computer Science and Information Security, Vol. 6, No. 2, 2009.

[15] I. K. Shlizerman, R. Basri, 3D Face Reconstruction from a Single Image Using a Single Reference Face Shape, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No.2 February 2011.

[16] Chetty, G. & White, M. (2010). “Multimedia Sensor Fusion for Retrieving Identity in Biometric Access

Control Systems", ACM Transaction on Multimedia Computing, Communications and Applications (Special Issue on Sensor Fusion), Nov. 2010.
[17] Chetty, G. & Wagner, M. (2008). Robust face-voice based speaker identity verification using multilevel fusion, Image and Vision Computing, 26, 1249-1260.
[18] S.M.E. Hossain & G. Chetty, "Next Generation Identity Verification Based on Face and Gait Biometric" International Conference on Biomedical Engineering and Technology" Kuala Lumpur 17-19 June 2011.
[19] R. Chakraborthy, H. Rengamani, P. Kumaraguru, R. Rao, "The UID Project": Lessons Learned from the Wast and Challenges Identified for India, Cyber Security, Cyber Crime and Cyber Forensics: Application and Perspective, Copyright © 2011, IGI Global.
[20] J. Feng, A.K. Jain, Fingerprint alteration, submitted to IEEE TIFS 2009.
[21] X. Zhou, B. Bhanu, Feature fusion of side face and gait for video-based human identification, © 2007 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.doi:10.1016/j.patcog.2007.06.019
[22] Bringer J., Chabanne H., Biometric Identification Paradigm Towards Privacy and Confidentiality Protection, Biometric: Theory, Application and Issues, Editor: Ellen R. Nichols, pp.123-141 © 2011 Nova Science Publishers, Inc.
[23] L. Hong, A. Jain, and S. Pankanti, "Can multi-biometrics improve performance?" in IEEE Workshop Identification of Advanced Tech., 1999, pp. 59–64.
[24] B. Schiele, "How many classifiers do I need?" in Int'l Conf. Pattern Recognition, vol. 2, 2002, pp. 176–179.
[25] Schuckers M.E., "Bayesian Method" Computational Methods in Biometric Authentication: Statistical Method for performance evolution, ©Springer-Verlag London Limited 2010.
[26] Cunningham P. and Delany S. J., k-Nearest Neighbour Classifiers, University College Dublin, Padraig.Cunningham@ucd.ie, Dublin Institute of Technology Sarahjane.Delany@comp.dit.ie , Technical Report UCD-CSI-2007-4, March 27, 2007.

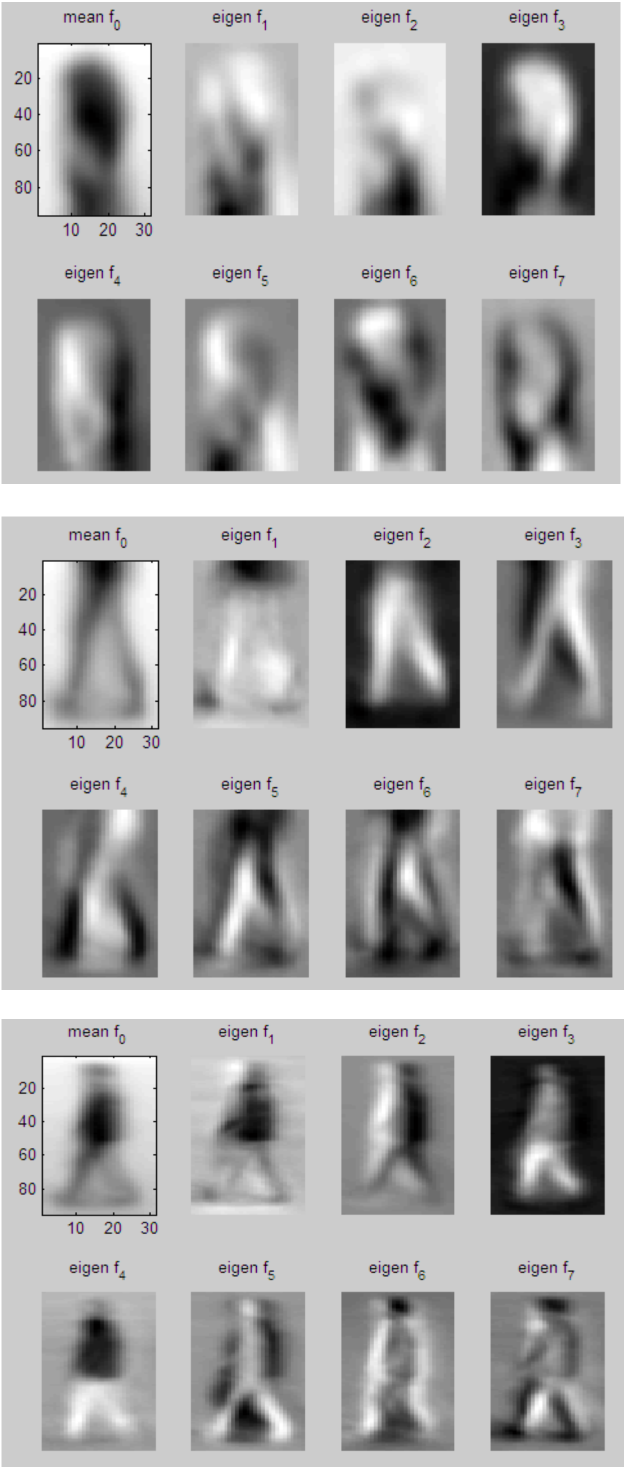

Fig. 4: Side Eigen-Faces and Eigen Gaits for face-only, partial gait and full-gait images.(from top to bottom)

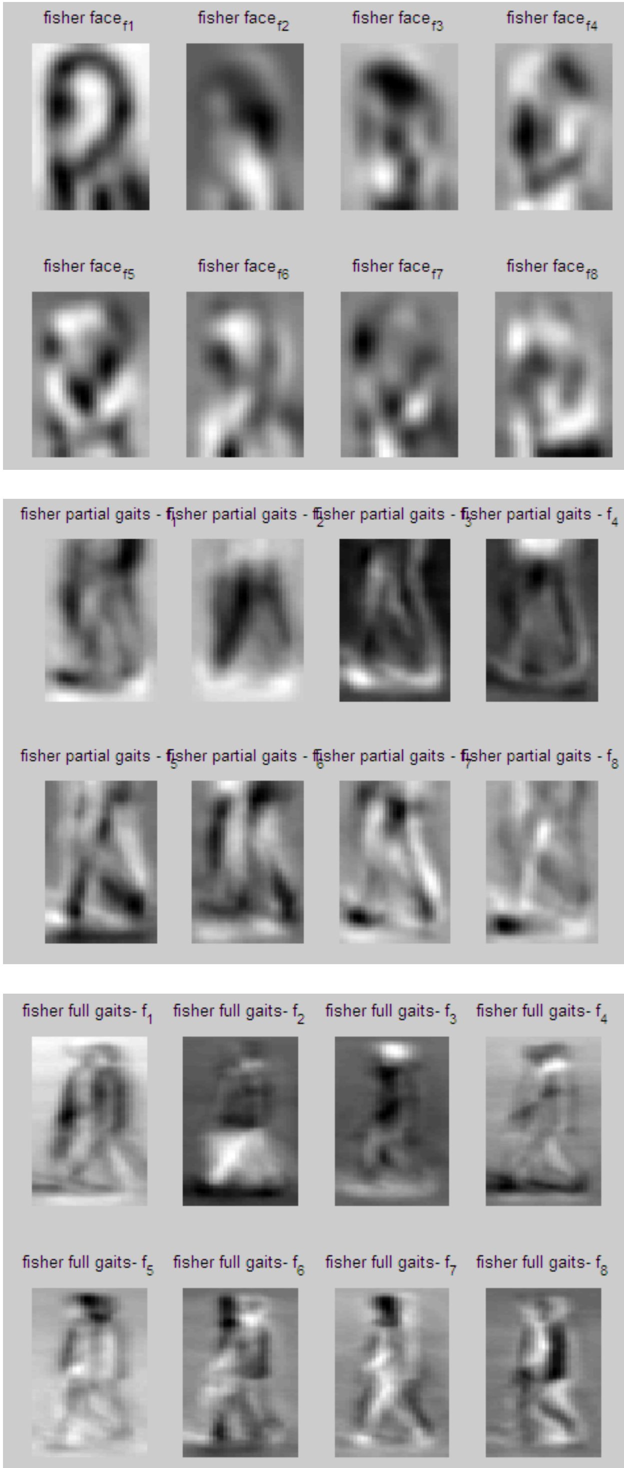

Fig. 5: Side Fisher-Faces and Fisher Gaits for face-only, partial gait and full-gait images.(from top to bottom)

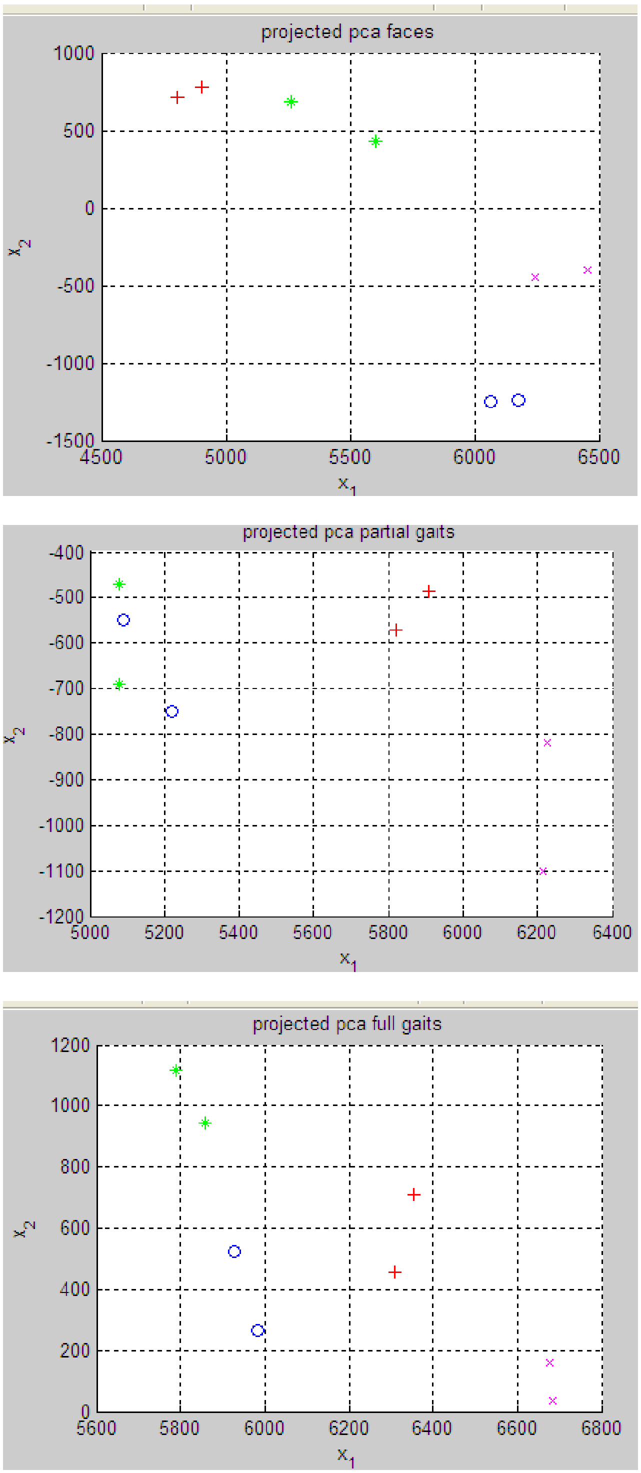

Fig. 6: Side Fisher-Faces and Fisher Gaits for face-only, partial gait and full-gait images.(from top to bottom)

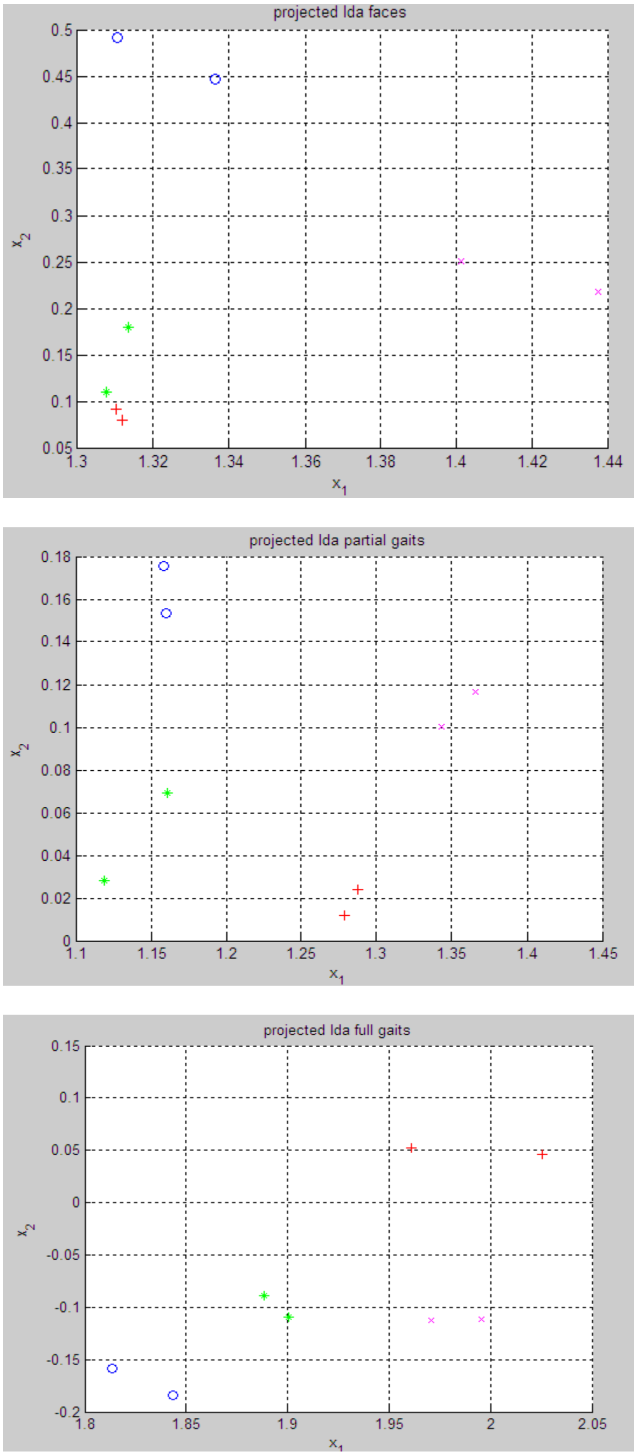

Fig7: Side Fisher-Faces and Fisher Gaits for face-only, partial gait and full-gait images.(from top to bottom)