Multi-view Multi-modal Gait Based Human Identity Recognition from Surveillance Videos

Emdad Hossain, Girija Chetty, and Roland Goecke

Faculty of ISE,
University of Canberra, Australia
emdad.hossain@canberra.edu.au

Abstract. In this paper we propose a novel human-identification scheme from long range gait profiles in surveillance videos. We investigate the role of multi view gait images acquired from multiple cameras, the importance of infrared and visible range images in ascertaining identity, the impact of multimodal fusion, efficient subspace features and classifier methods, and the role of soft/secondary biometric (walking style) in enhancing the accuracy and robustness of the identification systems, Experimental evaluation of several subspace based gait feature extraction approaches (PCA/LDA) and learning classifier methods (NB/MLP/SVM/SMO) on different datasets from a publicly available gait database CASIA, show significant improvement in recognition accuracies with multimodal fusion of multi-view gait images from visible and infrared cameras acquired from video surveillance scenarios.

Keywords: multimodal, multiview, PCA, LDA, MLP, identification, SMO, feature fusion.

1 Introduction

Automatic human identification from arbitrary views 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. Also, gait recognition is being considered as a next-generation recognition technology, with applicability to many civilian and high security environments such as airports, banks, military bases, car parks, railway stations etc. For these application scenarios, it is not possible to capture the frontal face, and even if it can be captured, it is of low resolution. Hence most of traditional approaches used for face recognition fail. However, several physiological and biomechanical studies have shown that human gait is inherently multimodal, and is based on kinematic interaction between several motion articulators, such as lower and upper limbs and other biomechanics of joints. It is person specific based on body weight, height, joint mobility in the limbs, and other behavioural nuances. If we can model these inherently multimodal traits , it is possible to identify human from a distance from their gait or from the way they walk. Even if frontal face is not visible, it is possible to establish the

F. Schwenker, S. Scherer, and L.-P. Morency (Eds.): MPRSS 2012, LNAI 7742, pp. 88–99, 2013.

identity of the person using certain static and dynamic multimodal cues from frontal and profile face, ear and head shape, walking style and speed, hand motion during walking etc. If automatic identification systems can be built based on this concept, it will be a great contribution to surveillance and security area Further. this will make a significant contribution to better understanding of gait abnormalities, and development of human computer interfaces. However, each of these cues or traits captured from long range low resolution surveillance videos on its own are not powerful enough for ascertaining identity, A combination or fusion of each of them, along with an automatic processing technique can result in robust recognition. In this paper, we propose usage of full profile silhouettes of persons from visible range and infrared range camera footage for capturing inherent multi-modal cues available from of the gait patterns of the walking human. This also addresses the need to establish identity from low resolution surveillance video images. In addition, user cooperation is not mandatory upon data collection making it suitable for surveillance and law enforcement scenarios. Further, capture of long range gait biometric data from surveillance videos contains several biometric traits such as side face, ear, body shape, and gait motion, which are a combination of physiological and behavioural biometrics. Automatic schemes that can process this rich multimodal information can result in robust human identification approaches.

In this paper, we propose the use of a principled approach involving feature extraction techniques based on multivariate statistical techniques, such as principle component analysis (PCA) and linear discriminant analysis (LDA), and efficient learning classifier approaches based on support vector machines and Bayesian classifiers. Further, we propose that the feature level fusion of multi-view multispectral images (from visible range cameras and infrared cameras) can enhance the performance of identification scheme as compared to single mode image features. Fusing features at the feature level is more effective than fusion at later stages, as the inherent multi-modality can be preserved at early stages of processing as compared to late fusion [2]. The experimental evaluation of the proposed approach with a publicly available CASIA [1] gait database shows a significant improvement in recognition performance as compared to other methods proposed in the literature. Rest of the paper is organised as follows. Next Section describes the background and motivation for proposed work, followed by the proposed multiview multimodal identification scheme in Section 3. The details of the experiments performed is described in Section 4, and conclusions and plans for further work is described in Section 5.

2 Background

Current state-of-the-art video surveillance systems, when used for recognizing the identity of the person in the scene, cannot perform very well due to low quality video or inappropriate processing techniques. Though much progress has been made in the past decade on visual based automatic person identification through utilizing different biometrics, including face recognition, iris and fingerprint recognition, each of these techniques work satisfactorily in highly controlled operating environments such as border control or immigration check points, under constrained illumination, pose and facial expression variations. To address the next generation security and surveillance requirements for not just high security environments, but also day-to-day civilian access control applications, we need a robust and invariant biometric trait [3] to identify a person for both controlled and uncontrolled operational environments. According to authors in [4], the expectations of next generation identity verification involve addressing issues related to application requirements, user concern and integration. Some of the suggestions made to address these issues were use of non-intrusive biometric traits, role of soft biometrics or dominant primary and non-dominant secondary identifiers and importance of novel automatic processing techniques. To conform to these recommendations; often there is a need to combine multiple physiological and behavioral biometric cues, leading to so called multimodal biometric identification system.

Each of the traits, physiological or behavioral have distinct advantages, for example; the behavioral biometrics can be collected non-obtrusively or even without the knowledge of the user. Behavioral data often does not require any special hardware (other than low cost off the shelf surveillance camera). While most behavioral biometrics are not unique enough to provide reliable human identification they have been proved to be sufficiently high accurate [5, 6]. Gait, is a powerful behavioral biometric, but as a single mode, on its own it cannot be considered as a strong biometric to identify a person. However, if we combine complementary gait information from another source, the multi-modal combination is expected to be powerful for human identification. Researchers have found that one of the most promising techniques is the use of multimodality or combination of differnt biometric traits or same biometric trait from multiple disparate sources. For example, researchers in [7, 8] have found that multi-modal scheme involving PCA on combined image of ear and face biometric results in significant improvement over either individual biometric. In addition, other recent attempts to improve the recognition accuracy include face, fingerprint and hand geometry [9]; face, fingerprint and speech [10]; face and iris [11]; face and ear [12]; and face and speech [13]. The fusion of complementary biometric information from disparate sources, however, did not attract much attention from the research community. This could be due to difficulty in acquiring the data, and processing and making sense out of them.

3 Multimodal Identification Scheme

For experimental evaluation of our proposed multimodal gait identification scheme, we used CASIA Gait Database collected by Institute of Automation, Chinese Academy of Sciences [1]. It is a large multi-view gait database, which is created in January 2005. There are more than 300 subjects. We used two different set of data known as dataset B and Dataset C. Dataset B was captured from 11 views with normal video camera, and 11 different views know as view angles. We used the data captured only in 90 degree view angle. The dataset C was captured with an infrared (thermal) camera. It takes into account four walking conditions: normal walking, slow walking, fast walking, and normal walking with a bag. The videos were all captured at night. Figure 1 shows the sample images in different view angles.

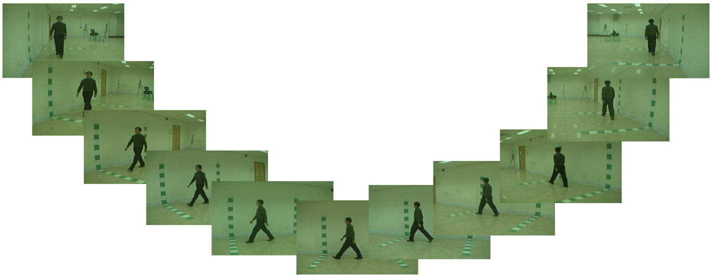

Fig. 1. Sample images from CASIA gait database

However, we used 50 subjects with a set of extracted silhouettes from Dataset B and another set of extracted silhouettes from Dataset C. Each subject consists of 16 images and in total 1600 images for 100 subjects (people). Figure 2 shows the extracted silhouettes from dataset B and C.

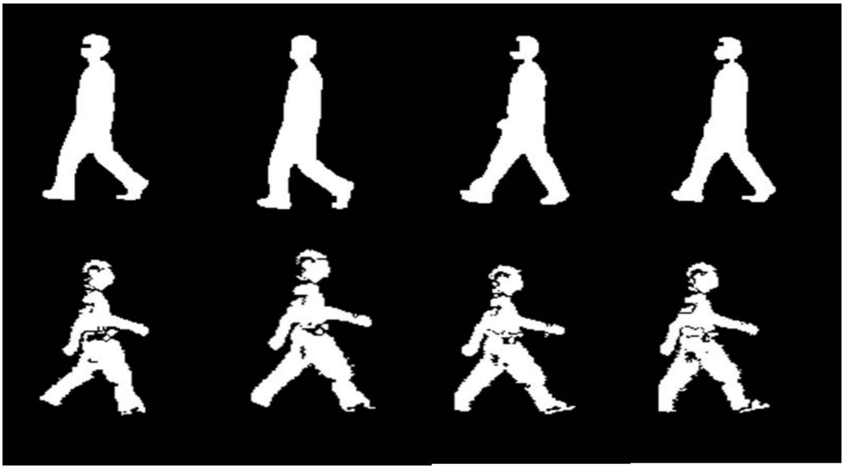

Fig. 2. Extracted silhouettes

We extracted the reduced dimensionality feature vector for each of the dataset separately by suing PCA (principal component analysis) and Linear Discriminant Analysis (LDA), and then have classified with different learning classifiers. Therefore our (cross camera feature level fusion) experiments involved evaluation of diiferent feature extraction and learning classifier combinations including PCA-MLP, LDA-MLP, PCA-SMO, and LDA-SMO.

3.1 Feature Extraction Using PCA-LDA Approach

Principle component analysis is a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. Since patterns in data can be hard to find in data of high dimension, where the luxury of graphical representation is not available, PCA is a powerful tool for analysing data. The other main advantage of PCA is that once we have found these patterns in the data, and we can compress the data, e.g. by reducing the number of dimensions, without much loss of information. Basically this technique used in image compression [14]. In the image analysis it works like;

$$X=(x1, x2, x3 \ldots\ldots N2) \tag{1}$$

where the rows of pixels in the image are placed one after the other to form a one dimensional image. Each image is N pixels high by N pixels wide. For each image it creates an image vector. And then it counts all the images together in one big image-matrix like;

$$\text{Matrix} = (v1, v2, v3 \ldots\ldots vN) \tag{2}$$

On the other hand, the LDA also closely related to principal component analysis (PCA) and factor analysis in that they both look for linear combinations of variables which best explain the data. LDA explicitly attempts to model the difference between the classes of data. PCA on the other hand does not take into account any difference in class, and factor analysis builds the feature combinations based on differences rather than similarities. Discriminant analysis is also different from factor analysis in that it is not an interdependence technique: a distinction between independent variables and dependent variables (also called criterion variables) must be made. LDA works when the measurements made on independent variables for each observation are continuous quantities. When dealing with categorical independent variables, the equivalent technique is discriminant correspondence analysis [15]. And in our experiment, LDA shows prominent than PCA. Next Section describes several classifiers we examined.

3.2 Naive Bayes and MLP Neural Network Classifier

Naive Bayes classifier can serve as a baseline classifier due to its simple probabilistic nature based on applying Bayes' theorem with strong (naive) independence assumptions. In other words, a naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature, given the class variable. Depending on the precise nature of the probability model, naive Bayes classifiers can be trained very efficiently in a supervised learning setting. In many practical applications, parameter estimation for naive Bayes models uses the method of maximum likelihood; in other words, one can work with the naive Bayes model without using any Bayesian methods [23]. In spite of their naive design and apparently over-simplified assumptions, naive Bayes classifiers have worked quite well in many complex real-world situations. Multi Layer perceptron (MLP) is a feedforward neural network with one or more layers between input and output layer. Feedforward implies that the data flows in on direction from input to output layer (forward). This type of network is trained with the backpropagation learning algorithm. MLPs are widely used for pattern classification, recognition, prediction and approximation. Multi Layer Perceptron can solve problems which are not linearly separable [16].

3.3 SVM and SMO Classifiers

Support Vector Machine (SVM) classifiers perform classification tasks by constructing hyperplanes in a multidimensional space that separates cases of different class labels. A support vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier. Whereas the original problem may be stated in a finite dimensional space, it often happens that the sets to discriminate are not linearly separable in that space. For this reason, in SVM, the original finite-dimensional space is mapped into a much higher-dimensional space, presumably making the separation easier in that space. To keep the computational load reasonable, the mappings used by SVM schemes are designed to ensure that dot products may be computed easily in terms of the variables in the original space, by defining them in terms of a kernel function selected to suit the problem.[24] The hyperplanes in the higher-dimensional space are defined as the set of points whose inner product with a vector in that space is constant.

SMO, on the other hand is an SVM classifier with learning based on Sequential Minimal Optimization (SMO). SMO decomposes the overall QP problem into QP sub-problems, using Osuna's theorem to ensure convergence [16]. Unlike the other methods, SMO chooses to solve the smallest possible optimization problem at every step. The advantage of SMO lies in the fact that solving for multi instance multipliers can be done analytically. In addition, SMO requires no extra matrix storage at all. There are two components to SMO: an analytic method for solving for the two Lagrange multipliers, and a heuristic for choosing which multipliers to optimize [17].

$$y1 \;/\!= y2 \Rightarrow \alpha 1 - \alpha 2 = k \tag{1}$$

$$y1 = y2 \Rightarrow \alpha 1 + \alpha 2 = k \tag{2}$$

However, the multi instance multipliers must fulfil all of the constraints of the full problem. The linear equality constraint causes them to lie on a diagonal line. Therefore, one step of SMO must find an optimum of the objective function on a diagonal line segment [17].

4 Experiments and Results

Different sets of experiments were performed on two datasets in CASIA database-Dataset B containing visible normal images of walking humans, and Dataset C consisting of infrared images. By using PCA and LDA techniques, we extracted the feature vector for both datasets, training different learning classifiers and performed identification experiments with multiple fold cross validation in single mode and multimodal fusion mode. We used different combinations of features (for example PCA-Dataset B, PCA-Dataset C, LDA-Dataset B, LDA-Dataset C and the feature level fusion of visible and infrared gait images from Dataset B and Dataset C. Table 1 to Table 5 show the recognition performance for each set of experiments in terms of recognition accuracy and several statistically significant performance measures such as true positive rate (TPR), false positive rate (FPR), precision, recall and Fmeasure.

All experiments involved either 5 or 10 fold cross validation. Cross-validation is a technique for assessing how the results of a statistical analysis will generalize to an independent data set. One fold of cross-validation involves partitioning a sample of data into complementary subsets (training and testing subsets), performing the analysis on one subset (called the training set), and validating the analysis on the other subset (called the validation set or testing set). To reduce variability, multiple folds of cross-validation are performed using different partitions, and the validation results are averaged over the folds. We examined 5 fold and 10 fold cross-validation for each set of experiments.

Table 1. Classifier Performance for Dataset B (Visible range dataset) with PCA features with 50 dimensions. (NB – naïve Bayes; MLP – Multilayer Perceptron; TPR-true positive rate; FPR – false positive rate).

Classifier	Folds	Accuracy(%)	TPR	FPR	Precision	Recall	F-measure
NB	10	48.63	0.49	0.01	0.49	0.49	0.48
NB	5	47.68	0.48	0.01	0.49	0.48	0.48
MLP	10	79.5	0.8	0	0.8	0.8	0.79
MLP	5	75.13	0.75	0.01	0.76	0.75	0.75

The first set of experiments involve Dataset B (visible range dataset) with 50 dimensional PCA features. As can be seen in Table 1, The recognition accuracy for naïve Bayes classifier with different 10-fold and 5 fold cross-validation is low, with 48.63 % for 10 folds and 47.68 for 5 folds. Using MLP neural net classifier (with backpropagation learning) results in better accuracy with 79.5% for 10 folds and 75.13% for 5 folds. However, the MLP classifier is computational intensive with long train and test times. This could be due to inability of PCA features to discriminate multiple classes (50 classes here) with the available data size or the structure of the neural network used.

The second set of experiments involved use of linear discriminant analysis features and use of support vector machine classifier. As can be seen in Table 2, the naïve Bayes classifier with 50 dimensional LDA features results in significant improvement in performance with 92.5% recognition accuracy as compared to 48.6 % with PCA features for 10 fold cross-validation (CV). With 5 fold CV, the LDA features result in an accuracy of 92.25% as compared to 47.68% for PCA features. Due to computational intensive nature of neural net classifiers, we examined SVM classifier

Table 2. Classifier Performance for Dataset B (Visible range dataset) with LDA features with 50 dimensions. (NB – naïve Bayes; MLP – Multilayer Perceptron; SVM-L(Support Vector Machine-Linear Kernel); SVM-RBF (Radial Basis Function Kernel); SVM-poly (Polynomial Kernel); SVM-Signmoid (Signmoidal kernel).

Classifier	Folds	Accuracy(%)	TP	FP	Precision	Recall	F-measure
NB	10	92.5	0.93	0	0.93	0.93	0.93
NB	5	92.25	0.92	0	0.93	0.92	0.92
SVM -L	5	81.13	0.81	0	0.82	0.81	0.81
SVM -L	10	78.75	0.78	0	0.81	0.79	0.78
SVM -RBF	5	31.30%	0.3	0.02	0.74	0.3	0.39
SVM -poly	5	27.63%	0.28	0.02	0.75	0.28	0.36
SVM -sigmoid	5	29.13%	0.29	0.02	0.74	0.29	0.38

for this set of experiments, as SVMs are known to have better generalization ability, are less computation intensive, and are based on sound theory, unlike neural networks whose development has followed a more heuristic path. Other advantages of SVM over neural networks are - whilst ANNs can suffer from multiple local minima, the solution to an SVM is global and unique, and SVMs have a simple geometric interpretation and give a sparse solution. Unlike ANNs, the computational complexity of SVMs does not depend on the dimensionality of the input space. ANNs use empirical risk minimization, whilst SVMs use structural risk minimization. SVMs outperform ANNs often, as they are less prone to overfitting [17]. However, the performance depends on the kernel used and other SVM parameters. As can be in Table 2, different types of kernels - linear kernel (SVM-L), radial basis function kernel (SVM-RBF), polynomial kernel (SVM-poly) and sigmoidal kernel (SVMsigmoid), result in different recognition accuracies. The SVM with linear kernel performs best with 81.3% recognition accuracy for 5 fold CV, and has a 78.75% for 10 fold CV. Also, for both naïve Bayes and SVM classifier with linear kernel, the performance with 5 fold cross-validation partition was almost similar to 10 fold cross validation. Hence, for rest of the experiments, we used 5 fold CV partition.

Table 3. Classifier Performance for Dataset C (Infrared range dataset) with LDA features with 5 folds. (NB – naïve Bayes; MLP – Multilayer Perceptron; SVM-L(Support Vector Machine-Linear Kernel); SMO(Poly)-Sequential Minimum Optimization-Polynomial Kernel.

Classifier	Features	Dim	Accuracy(%)	TPR	FPR	Precision	Recall	F-measure
NB	PCA	50	56.63%	0.57	0.01	0.59	0.57	0.57
SVM-L	PCA	50	79.88	0.8	0	0.81	799	799
SVM-L	LDA	50	86.25%	0.86	0	0.88	0.86	0.87
NB	LDA	50	93.75%	0.94	0	0.94	0.94	0.94
NB	LDA	25	93.5	0.94	0	0.94	0.94	0.94
SVM-L	LDA	25	83.25%	0.83	0	0.85	0.83	0.84
SMO -poly	LDA	25	94	0.94	0	0.95	0.94	0.94

For the third set of experiments, we examined Dataset C, the infrared camera gait image dataset, with 5 fold cross validation. As can be seen in Table 3. infrared image dataset performs better than visible range dataset for both PCA and LDA features. The recognition accuracy achieved with 50 dimensional PCA features results is 56.3% for naïve Bayes classifier for Dataset C as compared to 47.68% for Dataset B (Table 1). A similar improvement in performance was achieved with 50-dimensional LDA features resulting in a recognition accuracy of 93.75% for Dataset C as compared to 92.25% for Dataset B. Further, we also examined reduced dimensional LDA features, as LDA features seem to model the identities better, even with large number of classes (50 classes/subjects). As can be seen in Table 3, there is no significant loss of accuracy with reduced dimensional feature vectors. With 25 dimensional LDA feature vector, the recognition accuracy achieved was 93.5 % for naïve Bayes classifier (as compared to 93.75% for 50 dimensions) and the accuracies were 83.25% for SVM with linear kernel (86.25%). This has a significant advantage as the reduced dimensional feature vector results in improvement in computational speed. In addition, for this set of experiments, we examined a different version of SVM classifier – SMO, the SVM with Sequential minimal optimization(SMO). SMO classifier uses an efficient algorithm for solving the optimization problem needed for training of support vector machines, and is known to result in a better performance than a traditional SVM which uses much more complex quadratic optimization problem during training. As can be seen in Table 3, the recognition accuracy achieved with SMO classifier with polynomial kernel is 94% as compared to 93.25 % achieved with SVM classifier with linear kernel.

Table 4. Classifier Performance for fusion of visible and infrared gait images (Dataset B + Dataset C) with equal weights a and with LDA features with 5 fold cross validation. (NB - naïve Bayes; SVM-L(Support Vector Machine-Linear Kernel; SMO(Poly)-Sequential Minimum Optimization- Polynomial Kernel.

Classifier	Dim	Accuracy(%)	TPR	FPR	Precision	Recall	F-measure
NB	50	98.38%	0.98	0	0.99	0.98	0.98
SVM-L	50	74.88%	0.79	0.01	0.75	0.75	0.97
SMO-poly	50	98.25%	0.98	0	0.98	0.98	0.98
NB	25	98.50%	0.99	0	0.99	0.99	1
SVM-L	25	72.50%	0.73	0.01	0.77	0.73	0.73
SMO-Poly	25	97.75%	0.98	0	0.98	0.98	0.98

The fourth set of experiments involved the feature level fusion of visible and infrared images from Dataset B and Dataset C. As we found the LDA features to be more discriminatory as compared to PCA, we used LDA features for all fusion experiments. As can be seen in Table 4, the fusion of normal visible camera and infra red camera images is synergistic, resulting in improvement in recognition performance as compared to single mode images. For naïve Bayes classifier, 50-dimensional LDA features result in 98.38% accuracy and 25-dimensional LDA features result in 98.5%. The recognition accuracy achieved with SVM-L (linear kernel) for 50-dim LDA features is 74.88% and 72.5% for 25-dim LDA vector. The SMO version of SVM classifier with polynomial kernel results in 98.25% accuracy for 50-dim LDA vector, and for 25 dimensional LDA features, the accuracy is 97.75%. Once again for fusion mode, SMO with polynomial kernel performs better than traditional SVM with linear kernel. An interesting observation was that the multimodal fusion (feature level) performs a more dominant role as compared to the type of classifier or the type of features, as irrespective of classifier used (naïve Bayes or SVM), the recognition accuracy is significantly higher with multimodal fusion (higher than 95 %).

The final set of experiments involved investigating the role of soft or secondary biometric information, in terms of walking style(fast walking and normal walking) for enhancing the recognition accuracy. The walking style data was available for visible camera images only for all 50 subjects (persons). We used the data for each person walking in two (2) different styles - fast and normal walking. In this final set of experiments, we examined three different approaches. First, we applied LDA-MLP separately to (1) normal walking data, (2) the fast walking data and (3) combined the data corresponding to slow and fast walking information into a single dataset. This represents a challenging scenario with both dominant identity specific gait information (primary biometric) and non-dominant secondary/soft biometric information (walking style) modeled by LDA/MLP approach.

Table 5. Result in fast walking and normal walking

No	Method	Dataset	Accuracy
1	LDA-MLP	Normal Walking	95.5%
2	LDA-MLP	Fast walking	94.5%
3	LDA-MLP	Combined	82.50%

As can be seen in Table 5, while individually fast and slow walking style information modeled by LDA/MLP technique results in good identification accuracy, with 95.5% for normal walking, and 94.5% for fast walking, the modeling of weak soft biometric information (walking style) along with strong biometric information (identity of each subject) weakens the overall identification accuracy (82.5%). However, this depicts more real world scenario, and development of appropriate high performance subspace features and efficient classifier methods can result in better identification performance. It should be noted that the fusion of primary and soft/secondary biometric features is not reported in Table 5 due to lack of space, but some of our preliminary experiments show that fusion of primary and secondary/soft biometric information (walking style) can result in synergistic fusion. Also, use of motion based static and dynamic features is currently being investigated.

5 Conclusions and Further Plan

In this paper we proposed a novel human-identification scheme from long range gait profiles in surveillance videos. We investigated the role of multi view gait images acquired from multiple cameras - infrared and normal visible images in ascertaining identity. We also examined the benefits achieved with multimodal fusion, the roles of efficient subspace features and classifier methods, and the importance of soft/secondary biometric (walking style) in enhancing the accuracy and robustness of gait based identification systems, Experimental evaluation of several subspace based gait feature extraction approaches (PCA/LDA) and classifier methods (NB/MLP/SVM/SMO) on different datasets from a publicly available CASIA gait database, showed a significant improvement in recognition accuracies with multimodal fusion of multiview gait images acquired from normal visible and infrared video surveillance scenarios.

References

1. Zheng, S.: CASIA Gait Database collected by Institute of Automation, Chinese Academy of Sciences, CASIA Gait Database, `http://www.sinobiometrics.com`
2. Huang, L.: Person Recognition By Feature Fusion, Dept. of Engineering Technology Metropolitan State College of Denver. IEEE, Denver (2011)
3. Bringer, J., Chabanne, H.: Biometric Identification Paradigm Towards Privacy and Confidentiality Protection. In: Nichols, E.R. (ed.) Biometric: Theory, Application and Issues, pp. 123–141 (2011)
4. Jain, A.K.: Next Generation Biometrics. Department of Computer Science & Engineering, Michigan State University, Department of Brain & Cognitive Engineering, Korea University (2009)
5. Yampolskiy, R.V., Govindaraja, V.: Taxonomy of Behavioral Biometrics. Behavioral Biometrics for Human Identification, 1–43 (2010)
6. Meraoumia, A., Chitroub, S., Bouridane, A.: Fusion of Finger-Knuckle-Print and Palmprint for an Efficient Multi-biometric System of Person Recognition. IEEE Communications Society Subject Matter Experts for Publication in the IEEE ICC (2011)
7. Berretti, S., Bimbo, A., Pala, P.: 3D face recognaition using isogeodesic stripes. IEEE Transaction on Pattern Analysis and Machine Intelligence 32(12) (2010)
8. Yuan, L., Mu, Z., Xu, Z.: Using Ear Biometrics for Personal Recognition,School of Information Engineering, Univ. of Science and Technology Beijing. Beijing 100083, `http://yuanli64hotmail.com`
9. Ross, A., Jain, A.K.: Information fusion in biometrics. Pattern Recognition Letters 24, 2115–2125 (2003)
10. Jain, A.K., Hong, L., Kulkarni, Y.: A multimodal biometric system using fingerprints, face and speech. In: 2nd. Int'l. Conf. AVBPA, vol. (10), pp. 182–187 (1999)
11. Wang, Y., Tan, T., Jain, A.K.: Combining Face and Iris Biometrics for Identity Verification. In: Kittler, J., Nixon, M.S. (eds.) AVBPA 2003. LNCS, vol. 2688, pp. 805–813. Springer, Heidelberg (2003)
12. Chang, K., et al.: Comparison and Combination of Ear and Face Images in Appearance-Based Biometrics. IEEE Trans. PAMI 25, 1160–1165 (2003)

13. Kittler, J., et al.: On combining classifiers. IEEE Trans. Pattern Anal. Mach. Intell. 20, 226–239 (1998)
14. Smith, L.I.: A tutorial on Principal Components Analysis
15. Linear discriminant analysis, Wikipedia, http://www.wikipedia.org
16. Multi Layer Perceptron, http://www.neoroph.sourceforge.net
17. Platt, J.C.: Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines, Microsoft Research. Technical Report MSR-TR-98-14 (17) (1998), http://jplatt@microsoft.com
18. Shlizerman, I.K., Basri, R.: 3D Face Reconstruction from a Single Image Using a Single Reference Face Shape. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(2) (2011)
19. Hossain, E., Chetty, G.: Multimodal Identity Verification Based on Learning Face and Gait Cues. In: Lu, B.-L., Zhang, L., Kwok, J. (eds.) ICONIP 2011, Part III. LNCS, vol. 7064, pp. 1–8. Springer, Heidelberg (2011)
20. Chin, Y.J., Ong, T.S., Teoh, A.B.J., Goh, M.K.O.: Multimodal Biometrics based Bit Extraction Method for Template Security. In: Faculty of Information Science and Technology, Multimedia University, Malaysia, School of Electrical and Electronic Engineering, Yonsei University, IEEE, Seoul (2011)
21. Multilayer Perceptron Neural Networks, The Multilayer Perceptron Neural Network Model, http://www.dtreg.com