2011 International Conference on Biomedical Engineering and Technology
IPCBEE vol.11 (2011) © (2011) IACSIT Press, Singapore

Next Generation Identity Verification Based on Face-Gait Biometrics

S.M.E. Hossain , G. Chetty

Faculty of information Science & Engineering, University of Canberra, Australia

Abstract. Biometric authentication of a person is highly challenging and complex problem. A significant research effort has gone into this areas and a number of research works were published, but still there is an immense shortage of accurate and robust methods and techniques. In this paper we survey several important research works published in this area and report our work in progress on next generation identity verification technologies based on face and gait biometrics.

Keywords: robust-method, next-generation, identity-verification.

1. Introduction

Biometric person identification is a common technological tool for identity verification. It is carries significant importance for national or international security. All most each and every part of human body is unique; some of the significant ones have been used for developing automate identity verification systems. Fingerprint, palm print, face, iris, ear [1, 2] etc. have been used immensely for current generation of person authentication technologies. There are still challenges in this area, and need for better biometric modalities, development of novel approaches and techniques are being an ongoing process. Video surveillance in public places and facilities has become omnipresent, and has become 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. The need for non-intrusive biometric modalities enjoys significant user acceptability. Though any one biometric modality on its own cannot address all the challenges, and importance of combining the information from multiple biometric modalities holds significant promise. Next Section reviews some of the current approaches using different biometric traits.

2. Background

Person authentication using fingerprint or voice biometric traits has increasingly being deployed for day-to day security and surveillance applications. However, one of most acceptable non-intrusive physiological attribute to authenticate is "face". Automated face recognition technology [3] 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 mugshots [3]. From 1960s till now vast number of research works has been conducted on biometric person authentication. Several research articles have been reported in use of signature, fingerprint, face and voice biometrics [4]. 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 scene geometry [5]. In fact, none of the methods result in acceptable false error rates. . Practically, it is not possible to obtain zero error rates in realistic operating environments, however the most

S.M.Emdad Hossain, Tel: +61-02-62012329
Fax: +61-02-62015231, Email: emdad.hossain@canberra.edu.au

of the research focussed on attempts to achieve acceptable false error rates (around 1 - 5 %), if not perfect error rates.

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 [5], 3D face recognition from single image using single reference face shape was proposed in [2], where researchers proposed a novel method for 3D shape recovery of faces that exploits the similarity of faces. It also should mention that a number of limitations of 3D identification are high costs, limited availability of databases [6].

There have been several works reporting use of fingerprints for authenticating identity. A fingerprint is made of a number of ridges and valleys on the surface of the finger [8]. The uniqueness of a fingerprint can be determined by the pattern of ridges and furrows as well as the minutiae points. There are five basic fingerprint patterns: arch, tented arch, left loop, right loop and whorl. Loops make up 60% of all fingerprints, whorls account for 30%, and arches for 10%. Fingerprints are usually considered to be unique, with no two fingers having the exact same dermal ridge characteristics [8]. In fact, there has been a debate on how stable is the uniqueness of fingerprints? Further, due to increasing use of fingerprints for criminal identification, there have been cases of abuse. Hundreds of asylum seekers in Sweden and French tried to cut or burn their fingertips to evade identification by "Eurodac", and EU fingerprint ID for asylum seekers [13], likewise, a Chinese women arrested for illegal entry had altered her fingerprints through surgery (Dec 8) [13]. According to most researchers, Iris and retina are not changeable, but still not out of limitation. The fail to enrol (FTE) rate brings up another important problem. Not all users can use any given biometric system. People without hands cannot use fingerprint or hand-based systems. Visually impaired people have difficulties using iris or retina based techniques. As not all users are able to use a specific biometric system, the authentication system must be extended to handle users falling into the FTE category. This can make the resulting system more complicated, less secure or more expensive [7]. The authors in [7] clearly indentified undeniable limitations for biometric person authentication using fingerprint, iris and retina. Same might goes to person authentication using signature, some systems may also compare visual images of signatures, but the core of a signature biometric system is behavioral, i.e. how it is signed rather than visual, i.e. the image of the signature [7]. It means it has limitations for usage with persons with disability, and it can't be applied to authenticate for large population due to behavioral nature of the trait.

Another possible biometric trait is use of hand geometry. In large populations, hand geometry is not suitable for so-called one-to-many applications, in which a user is identified from his biometric without any other identification [10]. Some extreme biometric traits have also been proposed such as use of ear canal. Researchers found that one of the most promising techniques is use of multimodality or combination of biometric traits. Using PCA on combined image of ear and face, researchers in [5, 11] have found that multi-modal recognition results in significant improvement over either individual biometric. Since we have reviewed some of the prominent biometric traits for person authentication, we now look at some of the desirable characteristics of biometric traits as proposed by [1]. Next Section discusses some of the desirable characteristics for good biometric traits.

2.1 Comparison of Various Biometric Technologies

The choice of a particular human characteristic to be used as a biometric trait depends on the following criteria [12]:

- **Uniqueness** is how well the biometric separates individually from another.
- **Permanence** measures how well a biometric resists aging.
- **Collectability** ease of acquisition for measurement.
- **Performance** accuracy, speed, and robustness of technology used.
- **Acceptability** degree of approval of a technology.
- **Circumvention** ease of use of a substitute.

The following table shows a comparison of existing biometric systems in terms of those parameters. A low ranking indicates poor performance in the evaluation criterion whereas a high ranking indicates a very good performance [12].

Comparison of various biometric technologies, according to A. K. Jain (H=High, M=Medium, L=Low)

Biometrics:	Universality	Uniqueness	Permanence	Collectability	Performance	Acceptability	Circumvention*
Face	H	L	M	H	L	H	L
Fingerprint	M	H	H	M	H	M	H
Hand geometry	M	M	M	H	M	M	M
Keystrokes	L	L	L	M	L	M	M
Hand veins	M	M	M	M	M	M	H
Iris	H	H	H	M	H	L	H
Retinal scan	H	H	M	L	H	L	H
Signature	L	L	L	H	L	H	L
Voice	M	L	L	M	L	H	L
Facial thermograph	H	H	L	H	M	H	H
Odor	H	H	H	L	L	M	L
DNA	H	H	H	L	H	L	L
Gait	M	L	L	H	L	H	M
Ear Canal	M	M	H	M	M	H	M

Circumvent ability listed with reversed colours because low is desirable here instead of high [12]

As can be seen in this Table, each and every individual technology has limitation either in universality, uniqueness, permanence, collectability, or performance, acceptability, circumvention. Due to these limitations, no single biometric can provide a desired performance and the usage of multimodal biometric traits sounds promising.. Exploiting information from multiple biometric sources or features improves the performance and also robustness of person authentication [14].One of most widely reported multimodal biometric authentication is combination of speech and signature features. Research shows that they result in good performance, but limited applications. Perhaps they didn't collect the data from practical environment [14]. So, that's still far from public applicability. Another popular multimodal trait is combined authentication of "face and iris", First of all face alone is not good enough to identify a person that has been proved few times. Now, in the case of iris, there would be problem for disabled people The research work reported by authors in [15] suggest usage of iris and face biometrics for robust identification and verification. They specifically applied 2-D discrete wavelet transform to extract the feature sets of low dimensionality from iris and face [15]. One interesting aspect of human iris is, that a person iris might change if he/she undergoes a medical surgery on eye. Research shows it is possible to have colour surgery on human iris [17]. There has also been some work reported on fusion of face and ear biometric. However, the result obtained under controlled environment is about 4% FRR, and authors in [16] are working on improving the performance in uncontrolled operating environments. Next Section describes some of the futures directions in biometric identification technologies.

3. Next Generation Biometric Technologies

Having reviewed the capabilities and limitations of present current generation biometric identification technologies [1], [7], [9], [14], [15] and [16], we now discuss some of the next generation biometric technologies that could play a major role in security and authentication applications. According to authors in [1], 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 fusion protocols. We report here some of the work in progress in our laboratory in this direction, where we are investigating face and gait biometrics as potential candidates. Use of visual biometric traits such as face and gait appears promising as they require no user involvement actively in collecting the data, with camera sensors collecting the information automatically. However, git

being a weak biometric, is more behavioural and on its own cannot be a powerful biometric trait. The fusion of dominant physiological biometric - *the facial image patterns*, and a weak behavioral biometric – *the gait patterns*, can however, be a powerful combination. Both are non-intrusive, inexpensive to deploy, require no cooperation from the subjects, and provide abundant data for analysis. However, there is a lack of efficient algorithms and fusion models for processing the combination of near range face image patterns and medium/long range gait patterns, and make a sensible decision on the identity of the individual - as a civilian or criminal, client or impostor - and detect their actions as benign or harmful. We are trying to address this problem by proposing several new computational algorithms and fusion models for processing the two biometric modalities. Some of the algorithms and fusion models we are currently investigating are described in next Section.

3.1 Face-Gait Fusion models

We are currently investigating some novel algorithms and fusion models to integrate face, a physiological biometric, with gait, a behavioural biometric at low level and high level,.Some of the work in progress is discussed here:

- **Face and gait decision fusion model:** For this approach, we are lookng at Hidden Markov Models and Fisher faces method for gait and face classification, respectively. And then, the results obtained from the two classifiers will be utilized and integrated at match score level. The proposed face-gait fusion approach will be tested on video sequences of several individuals collected from different directions. The results of fusion of face and gait will be tested for robustness and better recognition performance compared with face only or gait-only method.

- **Static and dynamic body biometric decision fusion model**: For this approach, a new human recognition algorithm by combining static and dynamic body biometrics is being investigated. For each sequence involving a walking human, temporal pose changes of the segmented moving silhouettes will be represented as an associated sequence of complex vector configurations and will then be analysed using the Procreates shape analysis method to obtain a compact appearance representation, called static information of body. In addition, a model-based approach under a condensation framework will be explored, which will track the walker and recover joint-angle trajectories of lower limbs, called dynamic information of gait. Both static and dynamic cues obtained from walking person video footage will be independently used for recognition using the nearest exemplar classifier. They will then be fused at the decision level using different combinations of rules and will be tested for improvement in performance for both identification and verification tasks. Experimental evaluation with a video surveillance dataset with several subjects (at least 20 – 30) will be done to demonstrate the feasibility of the proposed algorithm.

- **Multi camera cross-modal fusion model:** For this approach, the face and gait cues will be derived from multiple simultaneous camera views, and we propose a visual hull algorithm for the fusion to create imagery in canonical pose prior to recognition. These view-normalized sequences, containing frontal images of face and profile silhouettes, will be separately used for face and gait recognition, and the results will be combined using a range of strategies. We will explore the concept of cross-modal correlation and score transformations for different modalities, with probabilistic settings for the cross-modal fusion. The effectiveness of various strategies will be evaluated on a data set with several subjects. We envisage that this novel fusion model will be useful in developing further statistical framework for multi-modal recognition.

- **Face and gait feature fusion model:** This new fusion approach will allow recognition of non-cooperating individuals at a distance in video, who expose side views to the camera. Information from two biometric sources, side face and gait, will be utilized and fused at feature level. For face, a high-resolution side face image will be constructed from multiple video frames. For gait, we propose a Gait Energy Image (GEI), a spatio-temporal compact representation of gait in video, to characterize human walking properties. Face features and gait features will be obtained separately using Principal Component Analysis (PCA) and Multiple Discriminate Analysis (MDA) from the high-resolution side face image and Gait Energy Image (GEI), respectively. The system will be tested on a database of video sequences corresponding to several people. It is expected that this face-gait fusion approach will carry more discriminating power as compared to any individual biometric.

- **Holistic and Hierarchical fusion protocols:** For this fusion approach, we plan to investigate the important of a new fusion protocol, by integrating face and gait cues for the single camera case. We will employ a view invariant gait recognition algorithm for gait recognition. A sequential importance sampling based algorithm will be used for probabilistic face recognition from video. We will employ decision fusion to combine the results of our gait recognition algorithm and the face recognition algorithm. We then consider two new fusion protocols: hierarchical and holistic. The first protocol will involve using the gait recognition algorithm as a filter to pass on a smaller set of candidates to the face recognition algorithm. The second protocol will involve combining the similarity scores obtained individually from the face and gait recognition algorithms Simple rules like the SUM, MIN and PRODUCT will be used for combining the scores. The results of the fusion will be tested on a face-gait database which has outdoor gait and face data of several subjects.

- **Adaptive face-gait fusion model:** For this fusion approach we plan to investigate adaptive fusion of face-gait patterns. Most work on information fusion for human identification is normally based on static fusion rules which cannot respond to the changes of the environment and the individual users. The adaptive fusion, which dynamically adjusts the fusion rules to suit the real-time external conditions. Two factors that may affect the relationship between gait and face in the fusion will be considered, i.e., the view angle and the subject-to-camera distance. Together they can determine the way gait and face are fused at an arbitrary time. Experimental evaluation will be carried out to assess the performance of adaptive fusion as compared to not only single biometric traits, but also those widely adopted static fusion rules including SUM, PRODUCT, MIN and MAX.

3.2 Primary/Secondary identifier extraction

From the same face-gait video surveillance footage, high level contextual information or secondary identifiers such as gender, age, aggression and emotion will be extracted which can then be used to automatically enhance the confidence level and the reliability of the decision taken by human identification stage. The approach for gender recognition is described here. We propose a fusion model based on canonical correlation analysis (CCA) technique. The canonical correlation analysis (CCA) is a powerful multivariate statistical analysis tool, well suited for relating two sets of measurements, by fusing the two modalities at the feature level. Experiments on large datasets will be carried out to examine the gender recognition capability of face-gait fusion approach as compared to individual face and gait patterns. Figure 1 shows the proposed fusion approach fro gender (secondary identifier) extraction.

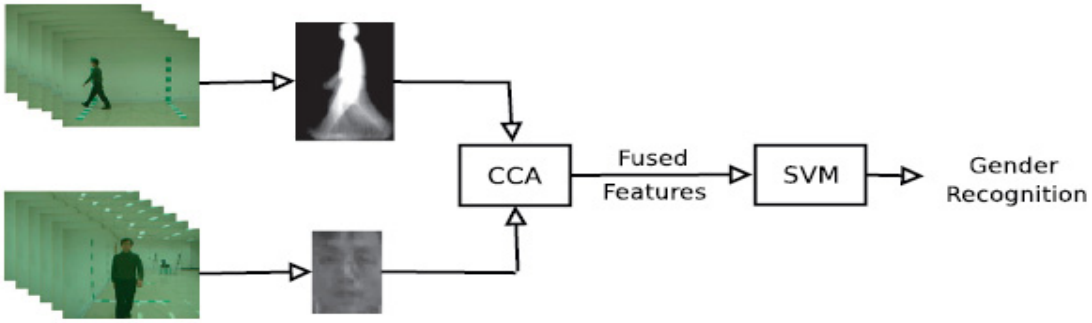

Figure 1: Face-gait fusion for gender (secondary identifier) extraction

Using gait for determining gender is a novel approach, and has not been explored before. Most of the existing work attempts to classify gender from human faces. In our work, we would like to investigate structural features and dynamic features of gaits for gender recognition, by adopting Gait Energy Image (GEI), a novel spatiotemporal compact representation of gaits. GEI has been demonstrated to be effective for representing gaits in the human identification problem. Using background subtraction techniques, the walking subjects can be extracted from the original image sequences to derive binary silhouette image sequences. To make the gait representation insensitive to the distance between the camera and the subject, we can perform silhouette pre-processing procedures including size normalization and horizontal alignment. Some examples of normalized and aligned silhouette images are shown in Figure 2. The entire human gait sequence can be divided into cycles as human walking repeats at a stable frequency. We decide the gait cycles by counting the number of foreground pixels in the bottom half of the silhouette and the two

consecutive strides in the variation of the number constitute a gait cycle. Given the pre-processed binary silhouette image $B_t(x, y)$ at time t in a sequence, the GEI is defined as follows:

$$G(x,y) = \frac{1}{N}\sum_{t=1}^{N} B_t(x,y) \qquad (1)$$

where N is the number of frames in the complete cycle(s) of a silhouette sequence, t is the frame number of the sequence, and x and y are values in the 2D image coordinate (see Figure 3 for an example of GEI). GEI reflects shapes of silhouette and their changes over the gait cycle, and it is not sensitive to incidental silhouette errors in individual frames.

Figure 2: Examples of normalized and aligned silhouette images. The rightmost image is the corresponding GEI.

3.3 Protocols for fusion of primary and secondary biometric identifiers

Once the primary identifiers and secondary identifiers are available, an appropriate protocol is needed to integrate the identifiers to address different user requirements based on the security level. The premise for this is that inherently the primary biometric identifiers for identifying the individual from the close-range face information and long-range gait information captured from video of a walking person, have several desirable properties under ideal and constrained operating environments, like universality, distinctiveness, permanence, collectability, acceptability, and resistance to spoofing. However, in reality, these systems operate in not so ideal environments. As a result, none of these traits can provide perfect recognition, and there is a need to improve the performance of these systems for day-to-day, civilian public access application scenarios. This is required for wide-spread diffusion and deployment of automated identification technologies based on non-intrusive, user friendly, biometric traits. Certain high level contextual information (soft characteristics) like gender, ethnicity, age, emotion/aggression, height, weight, and eye colour information could be extracted from the same video surveillance footage. Although, these soft characteristics are not unique and reliable, weak on their own, and are not capable of being decisive, they do provide important secondary level-additional demographic information about the user. They can certainly complement the identity information provided by the stronger biometric identifiers like face. The forensic identity recognition community have been using such soft characteristics for suspect and victim identification for a long time. The usage of secondary biometric identifiers – like the gender, age, ethnicity of the person, or the emotion- like aggression - on their own may not be useful to detect a criminal or identify an impostor. However, if used along with primary identifiers, it would be possible to enhance the robustness, reliability, and performance for different user and security requirements. Figure 3 shown below is the fusion structure for the combining primary and secondary identifier information

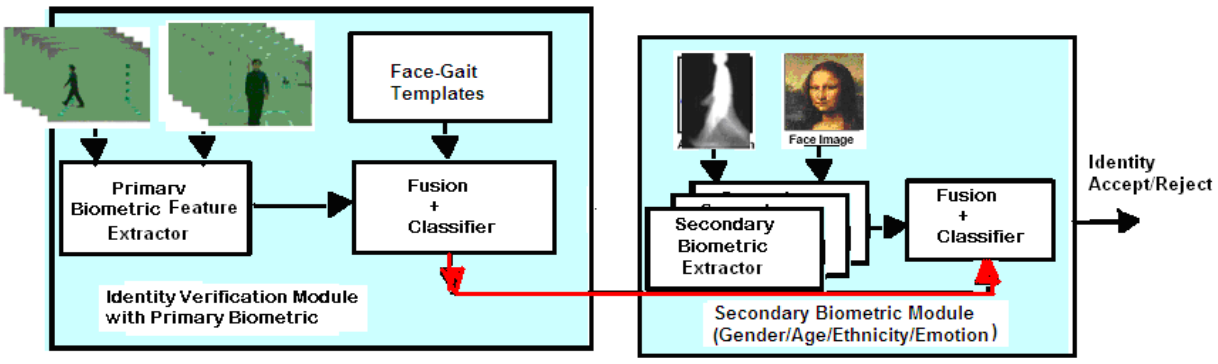

Figure 4: Fusion structure for primary and secondary identifiers using face and gait patterns.

If the evidence from the primary biometric modules is not sufficient to take a decision about the identity of the speaker, the secondary biometric modules supplement the evidence and enhance the confidence level

of the decision process. Such a setup for authentication process can provide several benefits. Modules in such authentication structures with primary biometric identifiers – gait image from long range camera 1, face image from short range camera2, followed by gender, age and emotion extraction (secondary identifiers) will spring into action, increasing the level of security to meet the requirements of increasing authentication accuracy.

4. Conclusions

In this paper we have presented a review of current biometric identification technologies and suggested the potential of face and gait biometric traits for next generation biometric technologies. Some of work in progress in relation to development of face-gait fusion models, the importance of primary and secondary biometric traits and the role of fusion protocols in addressing the requirements of next generation biometrics is discussed. The future work will involve evaluation of fusion models being developed for different face-gait databases in terms of false accept rates, false reject rates and equal error rates under controlled and uncontrolled environments.

5. References

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

[2] 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

[3] 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

[4] 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

[5] 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

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

[7] 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

[8] Fingerprint Identification Technology, Principles of fingerprint biometrics, www.biometricvision.com

[9] 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, fgregory,trevorg@ai.mit.edu

[10] Biometric technology, Hand Geometry Identification Technology, www.biometricvision.com

[11] L. Yuan, Z. Mu, and Z. Xu, Using Ear Biometrics for Personal Recognition, School of Information Engineering, Univ. of Science and Technology Beijing. Beijing 100083, yuanli64@hotmail.com

[12] Comparisons of Various Biometric Technologies, www.biometricvision.com

[13] J. Feng, A.K. Jain, Fingerprint alteration, submitted to IEEE TIFS 2009.

[14] 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)

[15] 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, Seodaemoon-gu, 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 2005)

[16] 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

[17] T. Bennett, New Iris Color Surgery, May 19, 2010, www.ehow.com

[18] F. Perronnin, J. C. Junqua, J. L. Dugelay, Biometrics Person Authentication: From Theory to Practice