A Novel Serial Multimodal Biometrics Framework Based on Semisupervised Learning Techniques

Qing Zhang, Yilong Yin, De-Chuan Zhan, and Jingliang Peng

***Abstract*— We propose in this paper a novel framework for serial multimodal biometric systems based on semisupervised learning techniques. The proposed framework addresses the inherent issues of user inconvenience and system inefficiency in parallel multimodal biometric systems. Further, it advances the serial multimodal biometric systems by promoting the discriminating power of the weaker but more user convenient trait(s) and saving the use of the stronger but less user convenient trait(s) whenever possible. This is in contrast to other existing serial multimodal biometric systems that suggest optimized orderings of the traits deployed and parameterizations of the corresponding matchers but ignore the most important requirements of common applications. In terms of methodology, we propose to use semisupervised learning techniques to strengthen the matcher(s) on the weaker trait(s), utilizing the coupling relationship between the weaker and the stronger traits. A dimensionality reduction method for the weaker trait(s) based on dependence maximization is proposed to achieve this purpose. Experiments on two prototype systems clearly demonstrate the advantages of the proposed framework and methodology.**

***Index Terms*— Serial multimodal biometrics, user convenience, semi-supervised learning, dimensionality reduction.**

I. Introduction

BIOMETRIC recognition, or biometrics, has been in high demand for many purposes including criminal identification, secure access control, forensics and so forth. Correspondingly, biometrics has been intensively researched and widely applied in the last decade.

In early years, monomodal biometric systems were used, which usually suffer from problems such as noisy data, unacceptable error rate and non-universality (*e.g.*, 4% people have difficult fingers). In order to overcome those limitations, multimodal biometric systems were proposed, which use multiple biometric traits to complete the recognition task. Previous research has demonstrated that a multimodal biometric system usually yields more reliable performance than a monomodal one due to the presence of multiple, (fairly) independent pieces of evidence [1].

There currently exist two modes of multimodal biometric systems, *i.e.*, parallel fusion mode and serial fusion mode. The former fuses the information of all traits in the system simultaneously while the latter uses traits in the system one by one in sequence. By comparison, the serial fusion mode usually provides more flexibility in ordering the traits and parameterizing the corresponding matchers in the chain. Further, it is more user convenient since traits later in the chain will not be used if an earlier one already identifies the user with high confidence.

Due to its configuration flexibility and user convenience, we focus on the serial fusion mode of multimodal biometric systems in this work. Specifically, we make major contributions in the following aspects:

- **A novel framework of serial multimodal biometric systems**. Compared with the currently existent serial multimodal biometric systems, our proposed framework is novel in that:
  - we for the first time propose to always use more (less) user convenient traits earlier (later) in the chain and save the use of less user convenient traits whenever possible;
  - we for the first time propose to enhance the weaker traits' distinguishing capabilities for promoted user convenience and recognition rates at the same time.
- **Effective methodology to strengthen the weaker traits**. We propose to use Semi-Supervised Learning (SSL) techniques to strengthen the weaker traits, utilizing the tight coupling relationship between the weaker and the stronger traits.
- **Promoted face-fingerprint and gait-fingerprint biometric systems**. The proposed framework and methodology are implemented and applied to two prototype biometrics systems, leading to superior performance in both user convenience and recognition accuracy.

The rest of this paper is organized as follows. Section II gives a brief review of the related work. Sections III and IV describe the proposed framework and the weaker trait enhancement method, respectively. Experimental results of the proposed framework are given in Sections V and VI for two exemplar biometric systems. Finally, we conclude this work in Section VII.

Manuscript received June 10, 2014; accepted August 1, 2014. Date of publication August 8, 2014; date of current version September 11, 2014. This work was supported in part by the National Natural Science Foundation of China under Grant 61070097, Grant 61173069, and Grant 61105043 and in part by the Shandong Natural Science Funds for Distinguished Young Scholar under Grant JQ201316. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Sviatoslav S. Voloshynovskiy. *(Corresponding author: Yilong Yin.)*

Q. Zhang and Y. Yin are with the Machine Learning and Data Mining Laboratory, Shandong University, Jinan 250101, China (e-mail: zhangqing2008@sdu.edu.cn; ylyin@sdu.edu.cn).

D.-C. Zhan is with the National Key Laboratory for Novel Software Technology, Lamda Group, Nanjing University, Nanjing 210093, China (e-mail: zhandc@nju.edu.cn).

J. Peng is with the School of Computer Science and Technology, Shandong University, Jinan 250101, China (e-mail: jingliap@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2014.2346703

II. Related Work

A. Parallel Fusion

The parallel fusion mode has been investigated more intensively and for longer than the serial fusion mode. As early as in 1998, Hong and Jain [2] proposed the parallel fusion mode by using fingerprint and face simultaneously for identification. Since then, many papers concerning parallel fusion have been published. Most of them have focused on the study of fusion methods. The currently available fusion methods can be divided into three major classes according to the level where the fusion is conducted. Some methods [3]–[5] combine features of all the biometric traits into a new feature, which are categorized as feature level fusion methods. Some other methods [6], [7] study how to make a final decision based on the recognition results given by all the biometric traits, which are categorized as decision level fusion methods. It has been noticed that to fuse at the feature level is sometimes tricky and even infeasible because of the incompatibility of features; also, to fuse at the decision level would inevitably lose useful detailed information. Consequently, as a compromise, the majority researches have focused on the score level fusion which combines match scores of all the biometric traits to from a final match score [8]–[12]. Further, the score level fusion methods can be roughly classified into transformation based methods where scores are normalized into a common domain and then combined [8], [9], classifier based methods where scores are treated as a feature vector and a classifier is constructed to discriminate genuine and impostor scores [10], [11], and density based methods which are based on the estimation of genuine and impostor match score densities [12].

In the investigation of the fusion methods, several special issues were particularly emphasized. Some works took the diversity of users into consideration and emphasized that user-dependent methods should be applied for better performance. Jain *et al.* [13] attested that setting the fusion weight and the decision threshold according to user-dependent information can promote the performance of the multimodal biometric systems. Uludag *et al.* [14] then proposed a user-dependent score normalization method and a user-dependent weighting method. Several other works [15]–[18] concerning classifier based score level fusion were proposed to train different classifiers such as PM [15], Bayesian [16] and SVM [17], [18] for different users. In [15], [16], [18], the user-dependent and user-independent information were treated as local and global information, respectively. Based on this, methods that adaptively combine local and global information were used to achieve satisfactory performance. Besides the emphasis of user diversity, some other works concentrated on the quality of captured biometric traits in the investigation of fusion methods. Julian *et al.* [18] set the fusion weight according to the quality of the corresponding trait. Some other works [19], [20] took the use of quantified quality information directly in classifier training. Additionally, some works [7] and [21] were proposed to choose fusion methods adaptively according to the performance requirement of the application.

Besides, some works investigated how to deal with the generally existing intra-class variance problem which results in a performance decline in the multimodal biometric systems. Roli *et al.* [22], [23] proposed the template co-update method. It uses the mutual help of two biometric matchers to update the template of each trait on-line based on the concept of a semi-supervised learning method called co-training. Afterwards, Didaci *et al.* [24] extended this work to more than two biometric traits. In works [25]–[27], the authors analyzed and testified the effectiveness of the template co-update method empirically.

Lately, some works [28], [29] investigated the feature level fusion by exploiting the technique of Multiple Kernel Learning (MKL). Yang *et al.* [28] proposed a novel supervised local-preserving canonical correlation analysis method to combine fingerprint and finger-vein features at the feature level. Shekhar *et al.* [29] proposed a joint sparse representation of multimodal biometrics by the techniques of MKL and Sparse Representation (SR), which addressed the difficulties in feature fusion and achieved recognition robustness.

Besides fusing multiple main biometric traits, Jain *et al.* [30] proposed to combine auxiliary information such as gender, ethnicity, height and weight with the main biometric traits in a parallel mode to improve the performance. These auxiliary information are called "soft biometrics". Many recent works took use of the soft biometrics and obtained promising results in various applications such as face recognition [31], gait recognition [32] and new born recognition [33].

B. Serial Fusion

Works have been published which investigated how to use biometric traits sequentially for recognition. Zhou *et al.* [34] designed a serial fusion system in which a subset of candidate identities is provided by a gait matcher at first, and a face matcher is then used to pick the recognized identity from the candidate subset. Marcialis *et al.* [35] proposed a framework for the serial fusion of face and fingerprint traits in which the acceptance and rejection thresholds for the corresponding matchers are set according to the zero FAR (False Accept Rate) and the zero FRR (False Reject Rate) values. Further, some works studied how to arrange the processing chain of biometric traits from several different points of views. One earliest work was made by Takabashi *et al.* [36], which applied the sequential probability ratio test to a three-stage biometric verification system (face, iris, voice). Later, Marcialis *et al.* [37] proposed a model to find the processing chain of two traits allowing a trade-off between the recognition accuracy and the matching time. They extended this model to systems with more than two traits in [38]. Allano *et al.* [39] proposed a method to set the processing chain balancing between the user cost and the recognition performance. Presently, Akhtar *et al.* [40] studied the robustness of the system under spoofing attack. They found evidence that serial fusion multimodal systems may be more robust than parallel ones.

III. The Proposed Framework

The proposed serial multimodal biometric framework always places more user convenient (but weaker) traits at

earlier stages than less user convenient (but stronger) traits in the serial fusion chain; in addition, it proposes to enhance the discriminating power of the weaker traits in order to increase the success rates of the earlier stages and reduce the use of the later stages.

The specific use scenario of the proposed framework is described as follows. In the enrollment process, we acquire from each genuine user a set of samples of all the traits used in the serial fusion chain. These samples are stored as templates. At the run time, a user goes through the serial chain stage by stage. At each stage, the user's trait is sampled and matched against the pre-sampled templates. If she or he passes the authentication at a stage, there is no need to use the later stage(s) in the chain; otherwise, she or he has to be sampled and authenticated at later stage(s). It is expected that, in the majority of cases, a genuine user will pass the authentication at the first stage. If a user has to be sampled in multiple stages and finally passes the authentication, her or his dynamically sampled traits together with the enrolled traits are used to further strengthen the weaker trait(s) that is(are) deployed earlier in the chain.

Detailed problem analysis that leads to the proposed framework is given in the following subsections.

A. Parallel vs. Serial Fusion

Though the parallel fusion mode has been more intensively researched than the serial one, it has inherent disadvantages in convenience and efficiency for practical use, limiting its scope of application.

The parallel fusion mode demands that all types of required traits be always captured for each user in both the enrollment and the recognition stages. This will inevitably result in an unwelcome burden to the users. It is usually acceptable that the users spend time and effort capturing all the required traits in the one-time enrollment process. However, after the deployment of the system, the recognition process will be repetitively conducted and therefore capturing all the traits for each use of the system will cause much inconvenience for the users. For instance, with a multimodal biometric system set at the main entrance of a school library, it is apparently unacceptable if users are required to stop and spend time providing all the biometric traits every time they enter the building.

The parallel fusion mode aims at enhanced system reliability by congregating the power of all involved traits which, however, leads to efficiency issues of the system. The use of multiple biometric traits originates in compensating problems that are hard to solve with just one biometric trait. However, difficulties in mono-modal biometric systems usually exist in a small percentage of cases where parallel fusion helps the most; for the other cases where the user identity can easily be determined with one trait, parallel fusion will become inefficient by redundant capturing and matching of all the traits. For instance, in a multimodal biometric system with fingerprint and face matchers, assuming that the fingerprint matcher can achieve a recognition accuracy of 95%, we only need the face matcher to compensate in the 5% hard cases. However, the parallel fusion mode requires that both the fingerprint and the face traits be captured and matched for each recognition task, meaning that the face capturing and matching is literally wasted for 95% of the cases.

Based on the above analysis, we see that the parallel fusion will be best suitable for special applications (*e.g.*, control of access to confidential military information) for which reliability is most important but user convenience and system efficiency can be sacrificed. For common biometric applications, however, parallel fusion may not be appropriate because of its inconvenience and inefficiency for use. The research on parallel fusion so far has focused on improving the reliability of the biometric system, but mostly ignored other factors that are equally or more important for practical deployment of the system. As a result, parallel fusion multimodal biometric systems have rarely been successfully used. As pointed out by Wayman [41], the added user interface efforts and time has limited the deployment of the multimodal biometric systems, and the multimodal biometric systems have failed to achieve the promise despite about thirty years of research.

In order to overcome the shortcomings of the parallel fusion mode, serial fusion mode has been proposed and researched in the last decade. In the serial fusion mode, the user goes through the authentication process stage by stage. At each stage, a certain type of trait is sampled and matched against a template library. Once she or he passes the authentication at a certain stage, all the later stage(s) will be bypassed. Generally speaking, most users do not have to go through the whole chain of stages for the authentication. As a result, user time and effort will be significantly saved and system efficiency significantly improved which makes the serial fusion mode more promising in most real applications than the parallel fusion mode.

B. Proposed Design Philosophy

Regarding the design of the serial fusion chain, currently existing methods [37]–[39] have mainly concentrated on the balance among several factors (*e.g.*, recognition accuracy, response time and user cost). These methods have paid little attention to the issue of user convenience which, however, is frequently the most important factor for a system to get widely applied. It is worth noting that user convenience was considered as a factor by one work [39], in which user cost was measured by the number of biometric traits used. However, it ignores the disparate characteristics of each separate trait and therefore can not accurately model the user convenience. As a result, it may frequently happen that the "optimal" designs of chain made by these methods do not in fact meet the users' preferences.

Following the above analysis, we set our novel philosophy in designing the chain of biometric traits: more user convenient trait(s) should always be set earlier in the chain and any further optimization to the performance of the chain should be based on this ordering. As a result, it leads to more humanized designs and avoids the tricky trait ordering and matcher parametrization process that is typical of many existing serial frameworks.

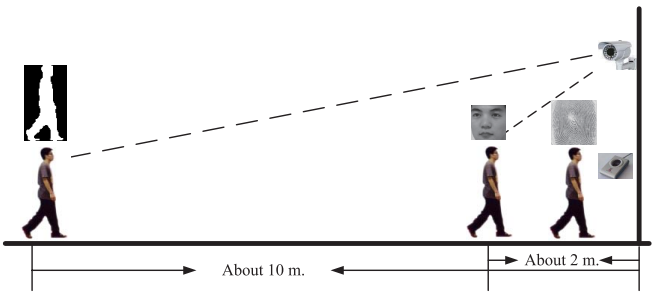


Fig. 1. Desirable scenario for an examplar serial fusion biometric system.

Take a multimodal biometric system of three most used biometric traits – fingerprint, face and gait as an example. Gait is the most convenient trait. The camera to capture the gait sequences is usually set at a distance, and users are recognized by gait while approaching without any further cooperation. Face recognition is the next convenient one. Generally speaking, face images should be captured at a relatively short distance from the objection where the cameras are set. If face images cannot be successfully captured during the walk, users have to stop and give more cooperation. Fingerprint is the most inconvenient trait. In most situations, users have to contact directly with a special capturing equipment to let the fingerprints captured. Obviously, as shown in Fig. 1, the most desirable way is that when a user is walking towards the objection from a distance (*e.g.*, 10 meters), her or his gait sequence is captured and matched first. If the user passes the gait recognition with high confidence, her or his face image and fingerprint will no more be needed; otherwise, the user then walks to a nearer position (*e.g.*, 1 or 2 meters) and has the face image captured and used for recognition. Only if the face matcher still cannot give a reliable result will the fingerprint be used. In this way, users can be whenever possible recognized by a more convenient trait, and inconvenient traits will only be used for compensation in necessary situations. Apparently, this serial fusion mode provides maximal convenience to the users.

C. Key Design Challenge

In biometrics, it is often the case that more (less) convenient traits are easier (harder) to capture but have less (more) discriminating power. For instance, the gait and the face traits are easier to capture than the fingerprint trait but the gait and the face matchers generally have lower recognition performance than the fingerprint matcher. There are mainly two reasons for this phenomenon. One is that inconvenient traits are often acquired at short distances from or even by direct contact with the capturing devices while convenient traits are often acquired under non-contacting environments or even at long distances from the devices. As a result, inconvenient traits often contain clearer and more detailed data, while convenient traits often contain discriminating data more implicitly. The other reason is that less controllable capturing environments of more convenient traits lead to more intra-class variance and, correspondingly, the performance of such traits will drop with the use of the system when the initial templates become poorly representative [22]–[27].

If we put more user convenient traits earlier in the chain, it may turn out that we still have to heavily use inconvenient traits for most users due to the weak performance of the earlier matchers. This will discount the advantage in user convenience which is the key goal of the proposed design. In other words, the contradiction between the convenience and the performance of biometric traits is the key obstacle to achieving user convenience and system performance simultaneously. This introduces the key challenge of our design: how to effectively enhance the discriminating power of the more user convenient but weaker traits for boosted user convenience and recognition performance at the same time.

IV. Weaker Traits Enhancement

Observing that stronger traits give more reliable recognition results, we consider learning from the stronger traits to enhance the weaker traits' recognition capability. For this purpose, we introduce SSL techniques into our framework. The reason to propose SSL techniques is that there are always limited labeled samples captured in the enrollment stage (due to the costly user interaction) but abundant unlabeled samples acquired with the use of the system. Specifically, our SSL method enhances the discriminating power of one weaker trait with the help of another stronger trait. For a chain with multiple traits, we may combinatorially group the traits into stronger and weaker pairs, and work on the selected pairs one by one. In the following, we focus on only one pair of traits, one stronger and the other weaker, for the simplicity of description.

Based on the analysis that the weaker trait may contain discriminating information more implicitly compared with the stronger trait, the SSL method attempts to extract discriminating information from the weaker trait with the help of the stronger one. A dimension reduction method is proposed to achieve this purpose. Dependence maximization is the central technique and we call it the DMDR (Dependence Maximization Dimensionality Reduction) method. Besides, a novel SSL scenario labeling strategy is proposed to assign pseudo labels to unlabeled samples which are then used to enhance the template library for future recognition tasks.

In the following subsections, we give an overview of the SSL based enhancement techniques, introduce the DMDR and the labeling methods, and make an analytical comparison with previous SSL based methods.

A. Overview of the Enhancement Techniques

In order to give a clear picture of the whole SSL based techniques, we show in Fig. 2 the flowchart of the enhancement process, specifics of which are given as follow.

Initially, a small set of samples labeled with the users' identities is acquired for each trait during the enrollment stage. Features of each trait are extracted by the stronger and the weaker trait feature extractors, respectively. Labeled samples of the stronger and the weaker traits are combined into pairs, each containing traits of the same user, and stored together with their features. These labeled sample pairs form an initial library of templates for recognition. For each use of

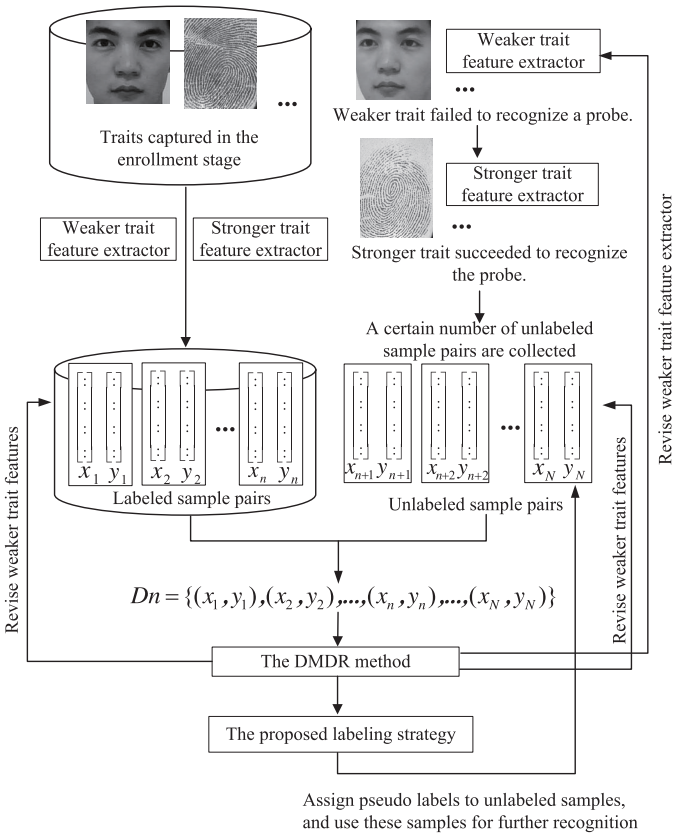


Fig. 2. The flowchart of the proposed SSL based enhancement techniques. The enhancement process is repeated with the accumulation of more unlabeled sample pairs.

this system, the user's weaker trait feature is first extracted and matched against the templates. It is expected that, in most cases, a genuine user will pass the authentication by the weaker trait at the first stage. However, if the weaker trait does not reliably pass a genuine user, she or he will have to get sampled and matched on the stronger trait. For all the users who have been sampled on two traits and finally recognized as genuine, we combine their dynamically sampled weaker and stronger traits into pairs, each containing traits of the same user, and store them as unlabeled sample pairs. When a certain number of unlabeled sample pairs are accumulated or the system runs for a certain period of time, based on all the labeled and the collected unlabeled sample pairs, we employ the DMDR method to obtain an optimal feature projection matrix $\boldsymbol{P}^*$ for the weaker trait. Thereafter, the weaker trait feature extractor is revised by $\boldsymbol{P}^*$ and new features are extracted from all the weaker trait samples including the labeled and the unlabeled ones. Unlabeled samples are each assigned a pseudo label of user identity by the proposed labeling strategy. The newly extracted weaker features and the assigned pseudo labels are used to update and extend the template library for later recognition tasks. The above-described process is repeated with the use of the system until an acceptable high performance of the weaker trait matcher is obtained.

It should be noted that the use of accumulated unlabeled sample pairs for weaker trait promotion is well supported. On one hand, accumulation of unlabeled sample pairs to a certain level signifies unsatisfactory performance of the weaker trait since an unlabeled sample pair may be captured only when the weaker trait fails to recognize a user; on the other hand, the accumulated unlabeled sample pairs contain the hardest cases encountered so far for the weaker trait and it makes the most sense to promote the weaker trait on those cases.

B. DMDR Method

Let $\mathcal{X}$ and $\mathcal{Y}$ denote the original feature space of the weaker trait and the stronger trait, respectively. In the enrollment stage, a sample of the weaker trait and a sample of the stronger trait with the same user identity (label) can be naturally combined into a pair to form a tight coupling relationship between the two traits. Labeled sample pairs in the templates are presented by $\{(\boldsymbol{x}_1, \boldsymbol{y}_1, l_1), (\boldsymbol{x}_1, \boldsymbol{y}_2, l_2), \ldots, (\boldsymbol{x}_n, \boldsymbol{y}_n, l_n)\}$, where $\boldsymbol{x}_i$ denotes a sample of the weaker trait, and $\boldsymbol{y}_i$ denotes a sample of the stronger trait, l_i is the corresponding label of the user identity, $i = 1, 2, \ldots, n$. For the ease of description, $\boldsymbol{x}_i$ ($\boldsymbol{y}_i$) may be interchangeably used for both a weaker (stronger) trait sample and its feature vector in the following text.

In the recognition stage, if the weaker trait is not able to recognize a user reliably, the stronger trait of the user should be used and therefore both the weaker and the stronger samples of the user are captured. The two samples are tightly related because they are captured form the same user. We formulate this tight coupling relationship by combining them into a pair. If a user has been sampled on both traits and finally passed the authentication, we keep her or his weaker and stronger trait samples as an unlabeled sample pair. When a certain number of unlabeled sample pairs are accumulated or when the system runs for a certain period of time, the DMDR method is applied to enhance the discriminating power of the weaker trait. Unlabeled sample pairs are presented by $\{(\boldsymbol{x}_{n+1}, \boldsymbol{y}_{n+1}), (\boldsymbol{x}_{n+2}, \boldsymbol{y}_{n+2}), \ldots, (\boldsymbol{x}_N, \boldsymbol{y}_N)\}$, where n is the number of labeled sample pairs, and N is the total number of labeled and unlabeled sample pairs. The total dataset of both labeled and unlabeled sample pairs is presented by $Dn = \{(\boldsymbol{x}_1, \boldsymbol{y}_1), (\boldsymbol{x}_2, \boldsymbol{y}_2), \ldots, (\boldsymbol{x}_n, \boldsymbol{y}_n), \ldots, (\boldsymbol{x}_N, \boldsymbol{y}_N)\}$.

By assuming that the weaker trait contains discrimination information implicitly, we need to extract these discriminations with the supervision of the stronger trait. More specifically, we attempt to find a lower-dimensional feature space for the weaker trait features in which the dependence between the information of the weaker trait and the information of the stronger trait is maximized. By denoting the projection vector of the weaker trait features as $\boldsymbol{p}$, a sample $\boldsymbol{x}$ is projected into a new space $\mathcal{F}$ by $\phi(\boldsymbol{x}) = \boldsymbol{p}^\top \boldsymbol{x}$ and the deduced kernel function is $\kappa(\boldsymbol{x}_i, \boldsymbol{x}_j) = \langle \phi(\boldsymbol{x}_i), \phi(\boldsymbol{x}_j) \rangle = \langle \boldsymbol{p}^\top \boldsymbol{x}_i, \boldsymbol{p}^\top \boldsymbol{x}_j \rangle$. For samples of the stronger trait, we define the kernel function $\ell(\boldsymbol{y}_i, \boldsymbol{y}_j) = \langle \boldsymbol{y}_i, \boldsymbol{y}_j \rangle$. Given the dataset Dn with joint distribution $P_{\boldsymbol{xy}}$, we define the kernel matrix for the weaker trait and the stronger trait as $\mathbf{K} = [\kappa_{ij}]_{N \times N}, \kappa_{ij} = \kappa(\boldsymbol{x}_i, \boldsymbol{x}_j)$ and $\mathbf{L} = [\ell_{ij}]_{N \times N}, \ell_{ij} = \ell(\boldsymbol{y}_i, \boldsymbol{y}_j)$, respectively. In realization, $\ell(\boldsymbol{y}_i, \boldsymbol{y}_j)$ can be assigned the match score of $\boldsymbol{y}_i$ and $\boldsymbol{y}_j$

by the stronger trait matcher. Then, we try to maximize the dependence between the information of the weaker trait in the projected feature space $\mathcal{F}$ and the information of the stronger trait.

In order to facilitate the dependence maximization and make full use of the potential non-linearities in both traits, we define a dependence between the kernels of different traits as

$$\mathfrak{D}(\mathbf{K}, \mathbf{L}) = tr(\mathbf{KL}) \tag{1}$$

by assuming that both $\mathbf{K}$ and $\mathbf{L}$ are centralized and normalized. For general kernels of data, if we define $\mathbf{H} = \mathbf{I} - \frac{1}{N} \times \mathbf{e}\mathbf{e}^\top$, where $\mathbf{I}$ is an identity matrix and $\mathbf{e}$ is an all-one column vector, the equation above becomes

$$\mathfrak{D}(\mathbf{K}, \mathbf{L}) = tr(\mathbf{HKHL}) \tag{2}$$

where $\mathbf{H}$ can be regarded as a centralized operator to eliminate the effect of sample position on the kernel matrix $\mathbf{K}$. Our dependence criterion is closely related to a kind of independence criterion called Hilbert-Schmidt Independence Criterion [42]. The dependence criterion computes the square of the norm of the cross-covariance operator over the domain $\mathcal{X} \times \mathcal{Y}$ in Hilbert Space. Due to the neat theoretical properties, we maximize the dependence of the information of both traits, *i.e.*,

$$\max \quad \mathfrak{D}(\mathbf{K}, \mathbf{L}) \quad = \quad \max \quad tr(\mathbf{HKHL}) \tag{3}$$

By representing the instances in $\mathcal{X}$ as $\phi(\boldsymbol{x})$, we can rewrite the target function in eq. 3 as

$$\boldsymbol{p}^* = \arg\max_{\boldsymbol{p}} \quad tr(\mathbf{H}\mathbf{X}^\top \boldsymbol{p}\boldsymbol{p}^\top \mathbf{X}\mathbf{H}\mathbf{L}) \tag{4}$$

To avoid the scaling problem, we add the constraint that the l_2-norm of $\boldsymbol{p}$ should be bounded. Therefore, we reformulate the optimization problem as

$$\begin{aligned} \boldsymbol{p}^* = \arg\max_{\boldsymbol{p}} \quad & tr(\mathbf{H}\mathbf{X}^\top \boldsymbol{p}\boldsymbol{p}^\top \mathbf{X}\mathbf{H}\mathbf{L}) \\ s.t. \qquad & \boldsymbol{p}^\top \boldsymbol{p} = 1 \end{aligned}$$

Note that

$$tr(\mathbf{H}\mathbf{X}^\top \boldsymbol{p}\boldsymbol{p}^\top \mathbf{X}\mathbf{H}\mathbf{L}) = \boldsymbol{p}^\top (\mathbf{X}\mathbf{H}\mathbf{L}\mathbf{H}\mathbf{X}^\top)\boldsymbol{p} \tag{5}$$

Since $\mathbf{X}\mathbf{H}\mathbf{L}\mathbf{H}\mathbf{X}^\top$ is symmetric, the eigenvalues are all real. Without any loss of generality, we can assume that the eigenvalues of $\mathbf{X}\mathbf{H}\mathbf{L}\mathbf{H}\mathbf{X}^\top$ are sorted as $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_D$. Thus, if d is the dimensionality of the new feature space, the optimal projection matrix $\boldsymbol{P}^*$ can be defined as $\boldsymbol{P}^* = [\boldsymbol{p}_1^*, \boldsymbol{p}_2^*, \ldots, \boldsymbol{p}_d^*]$, where $\boldsymbol{p}_i^*$ is the normalized eigenvector corresponding to the i-th largest eigenvalue λ_i, $i = 1, \ldots, d$ ($d \ll D$). Since the eigenvalues reflect the contribution of the corresponding dimensions, we can control d by setting a threshold $thr (0 \leq thr \leq 1)$ and then choose the first d eigenvectors such that

$$\sum_{i=1}^{d} \lambda_i \geq thr \times \sum_{i=1}^{D} \lambda_i \tag{6}$$

With different setting of thr, the outcome of the DMDR method varies, which results in different recognition performance of the weaker trait. Experimental results indicate that a larger thr usually yields better recognition performance, and when thr goes beyond 0.9, the performance becomes quite stable.

It should be noted that our proposed DMDR method is different from the MKL-based ones [28], [29]. They propose joint representations of the weaker and the stronger trait features, which usually change the feature space of each separate trait. We assume that the stronger trait already exhibits excellent discriminating capability. Therefore, we choose to maintain the original feature space of the stronger trait but, under the supervision of the stronger trait, find an optimal feature space for the weaker trait where its discriminating power is better expressed.

C. Labeling Strategy

By the DMDR process, we update all the weaker trait features in the templates by their projections on the newly found feature space. Further, we assign a pseudo label for each unlabeled sample pair. The updated weaker trait features and the assigned pseudo labels together lead to an updated and extended library of templates.

For an unlabeled sample pair, $(\boldsymbol{x}_u, \boldsymbol{y}_u) \in Dn$, we compute its match score with each labeled sample pair, $(\boldsymbol{x}_l, \boldsymbol{y}_l) \in Dn$, as follows. Assuming that the match score between $\boldsymbol{x}_u$ and $\boldsymbol{x}_l$ is obtained as $\boldsymbol{sx}_{ul}$, and the match score between $\boldsymbol{y}_u$ and $\boldsymbol{y}_l$ is obtained as $\boldsymbol{sy}_{ul}$, a final match score between the two sample pairs is computed as $\boldsymbol{s}_{ul} = \boldsymbol{w}_x \times \boldsymbol{sx}_{ul} + \boldsymbol{w}_y \times \boldsymbol{sy}_{ul}$, where $\boldsymbol{w}_x \in [0, 1]$, $\boldsymbol{w}_y \in [0, 1]$ and $\boldsymbol{w}_x + \boldsymbol{w}_y = 1$. In this way, both the weaker and the stronger traits take part in the labeling process with different weights. A reasonable way to determine the values of $\boldsymbol{w}_x$ and $\boldsymbol{w}_y$ is to set the weights according to the performance (measured by accuracy in the experiments) of the weaker and the stronger matchers. To an unlabeled sample pair, the label of the labeled sample pair with the highest match score is assigned as the pseudo label.

D. Proposed Method vs. Previous SSL Based Methods

There are mainly three SSL based methods proposed for the biometric systems which are named self-training based method [43], [44], graph based method [45], [46] and co-training based method [22], [23], respectively. The previous SSL based methods mainly concentrate on updating the library of templates with unlabeled samples to make the templates more representative. The self training based method adapts itself to the confidently labeled samples. The graph based method which is also called mincut method organizes samples using a graph-based structure, and chooses the optimum labeling of the unlabeled samples by partitioning the graph into two sub-graphs using the max-flow/min-cut algorithm.

The most related work is the co-training based template co-update method proposed for multimodal biometric systems. This work uses the mutual help of two biometric matchers to update the template of each trait. Our method is different form the template co-update method mainly in two respects. Firstly, in the proposed method, traits are not treated equally. We concentrate on promoting the weaker trait to address the

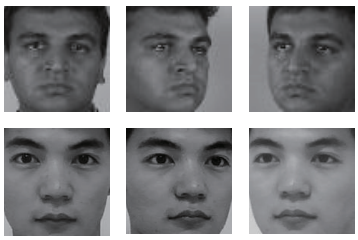


Fig. 3. Representative face images in the combined face database. Images in the first row are from the FacePix database, and images in the second row are from the database constructed by SDUMLA.

contradiction between the convenience and the performance. Secondly, the promotion of the weaker trait is not solely relied on labeling unlabeled samples. The core is a dimensionality reduction method by which information of the stronger trait guides the feature transformation of the weaker trait.

V. Experiments on Face and Fingerprint

In this section, the proposed framework with the SSL based method is applied to a use case of a serial multimodal biometric system in which a face matcher is deployed before a fingerprint matcher. Experiments testify the effectiveness of the proposed SSL based method in promoting the face matcher. The performance of the system in boosting the user convenience and the recognition accuracy is also verified.

A. Databases

The face database is obtained by combining a subset of the public FacePix database [47] and a subset of the database collected by the Group of Machine Learning and Applications, Shandong University (SDUMLA) in September, 2011. The special choice of the database is due to the requirement of the compared dimensionality reduction method–LDA method, which is strongly dependent on the number of classes (c), the number of samples (n), and the original space dimensionality (d). There should be $d + c$ samples at least to have a nonsingular within-class scatter matrix. Consequently, we combine two databases to guarantee abundant samples. FacePix database contains 30 users with 61 face images per user. In the database collected by SDUMLA, there are 32 users with 120 face images per user. We choose 30 users with 60 images per user from both databases. Totally, the combined database includes 60 users with 60 face images per user. The selected face images vary in pose, which can roughly be categorized into frontal images, profile towards left and profile towards right. All face images are normalized to have 32×32 pixels. Fig. 3 shows some representative samples of the combined face database.

We construct the fingerprint database from the database collected by SDUMLA in September, 2011. In this database, 32 users' fingerprints are collected including fingerprints of thumb, index and middle fingers from both hands. FPR620 optical fingerprint scanner developed by Zhong zheng Inc. is used to capture the fingerprints. We choose 30 users with 60 thumb fingerprints per user as half of the database. We use the same 30 users with 60 fingerprints of the index finger per user as the other half of the database.

In the constructed face and fingerprint databases, 60 virtual users are formed by assigning the same identity to two users in the face and the fingerprint databases, respectively. For each virtual user, there are 60 face samples and 60 fingerprint samples in the databases. The samples are then divided into three parts as follows.

Part I: Training set. We choose 3 face images and 3 fingerprints per user as the labeled samples in the template library. The labeled face images are frontal images to simulate the controllable capturing process in the enrollment stage.

Part II: Testing set. For each user, 30 face images and 30 randomly selected fingerprints are used as the testing set. Face images in the testing set include 10 frontal images, 10 profile towards left and 10 profile towards right for each user.

Part III: Unlabeled sample set. The rest 27 face images and 27 randomly selected fingerprints of each user are used to simulate the unlabeled samples collected with the use of the system.

It should be noted that, in a real scenario, unlabeled samples are gathered gradually with the running of the system. In general, a large pool of input samples should be needed to accumulate sufficient unlabeled samples for weaker trait promotion. However, it is hard for us to find such a large pool of input samples for experiment. Therefore, without loss of generality, we assume that a set of unlabeled samples (*i.e.* Part III) has already been accumulated from the previous running of the system, which will be directly applied to promote the weaker trait matcher in our experiments.

B. Methods for Comparison and Experiment Configuration

To testify the effectiveness of the proposed SSL based method in promoting the face matcher, three other SSL based methods - the self-training based method, the mincut based method and the template co-update method are used for comparison. In the following, we call the three methods the Self-update, Mincut and TCU (Template Co-Update) methods for convenience. In the proposed SSL based method, the initial feature extractor for the weaker trait can be any existing feature extraction method, we use the PCA [48] method in the experiments. In the novel labeling strategy, the values of the weights are obtained according to the ratio of the weights, and the constraint that the weights should be sum to one. The ratio of the weights is determined by the ratio of accuracies of the face and the fingerprint matchers where accuracies are evaluated by three-fold cross-validation on the labeled samples since there are three labeled samples of each enrolled user for each trait. The value k of the k-NN sub graph for the Mincut method is set to 3, which is a desirable setting according to [46]. In the previous SSL based methods, five commonly used dimensionality reduction methods for face recognition are used for comparison, including PCA, LDA [49], ICA [50] and other two LPP [51] methods named LPP1 and LPP2, respectively. The adjacency Matrix G is constructed differently in LPP1 and LPP2. In LPP1, $G_{ij} = 1$ if the i-th and the

j-th samples have the same label; $G_{ij} = 0$ otherwise. In LPP2, G is constructed using the similarities between face samples where similarities are measured by Euclidean distance. A typical minutiae-based fingerprint matching algorithm [52] is used for the fingerprint matcher.

The performance of the multimodal biometric system under the proposed framework with the SSL method is compared with a parallel fusion system and serial fusion systems using the proposed serial fusion strategy but different SSL based methods (*i.e.*, Self-update, Mincut and TCU) for face matcher promotion. These systems are compared mainly on two aspects, user convenience and recognition accuracy.

In order to measure the user convenience of a system, we first quantify the user inconvenience of each trait. For that purpose, we design a questionnaire that asks the subjects to choose an integral score of inconvenience for each of the three traits – fingerprint, face and gait. The integral score of inconvenience ranges from 1 (the least inconvenient) to 5 (the most inconvenient). Sixty-five subjects of various ages and genders responded to this questionnaire. Based on their responses, the average scores of inconvenience are 2.8 for fingerprint, 1.9 for face and 1.2 for gait.

TABLE I
NUMBERS AND CATEGORIES OF FACE SAMPLES IMPOSED FOR EACH USER AT EACH STEP

Step	Number of imposed face samples			
	Frontal	Profile to right	Protile to left	Total
1	0	3	4	7
2	7	2	1	10
3	0	5	5	10

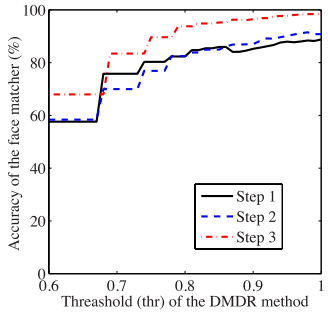


Fig. 4. Accuracy of the face matcher under the proposed SSL based method with different setting of thr for the DMDR method. Results at each step of the promotion are shown.

C. Face Matcher Enhancement

This subsection shows the effectiveness of the proposed work in promoting the performance of the face matcher (weaker matcher), and the comparison with previous works.

Using the training set as the template library, we test the recognition performance of the fingerprint matcher and the face matchers (with different dimensionality reduction methods) using the testing set. Because the testing set is a closed dataset, *i.e.*, imposters of a user are other users in the dataset, we can evaluate the performance of each matcher by the measurement of accuracy. The accuracy of the fingerprint matcher reaches 99.5% while the highest accuracy of the face matchers is 53% when LPP1 is used. Clearly, the fingerprint matcher is stronger than the face matchers when they are trained with only the labeled samples.

In our experiments, to simulate the enhancement process of the weaker trait, 27 unlabeled face and fingerprint pairs per user are imposed as collected unlabeled sample pairs in three sequential batches at three steps, respectively. At each step, all the labeled and the unlabeled samples imposed so far are used by the proposed SSL method to promote the face matcher. To simulate the uncontrollable recognition scenario, different kinds of face samples are imposed at each step randomly without any regularity. The numbers and categories of face samples imposed for each user at each step are listed in Table I.

For the DMDR method, different settings of the threshold thr lead to different numbers of retained dimensions in the transformed feature space and, correspondingly, different performance of the face matcher. We investigate the performance of the face matcher at each step of the promotion with different settings of thr for the DMDR method. Results shown in Fig. 4 indicate that larger thr values usually yield better performance and, when thr is above 0.9, the performance becomes stable. In other words, the DMDR method becomes robust when thr is big enough, and a thr value above 0.9 often yields satisfactory performance.

In the following experiments, the value of thr that leads to the best performance is always chosen for each running of the proposed DMDR method. For a fair comparison we set the threshold parameter of each compared dimensionality reduction method to the value that maximizes the corresponding face matcher's performance.

At each promotion step, we test the performance of the face matchers using the proposed method and the compared methods on the testing set. The mostly used zero FAR and 1% FAR thresholds are set for the Self-update method and the TCU method to determine the unlabeled samples that are used to augment the template library. For the TCU method, LDA can be used only at the 2nd and the 3rd steps of the promotion when enough samples are imposed, while the fingerprint matcher is relied on at the 1st step to augment the template library. The result of LDA can not be investigated in the the Self-update and the Mincut methods since the help of the fingerprint matcher is missing. Fig. 5 to Fig. 7 show the results. The accuracy of the face matcher in the proposed method gets improved from around 50% to higher than 98%, which means that the face matcher is able to successfully recognize most genuine users through the promotion.

The superiority of the proposed method is most apparent when compared with the Self-update and the Mincut methods. It is mainly because the Self-update and the Mincut methods depend solely on the face trait itself for the promotion so that tend to augment the template library with only easy samples that can be reliably recognized by the face matcher on the initial templates. When the imposed unlabeled samples contain

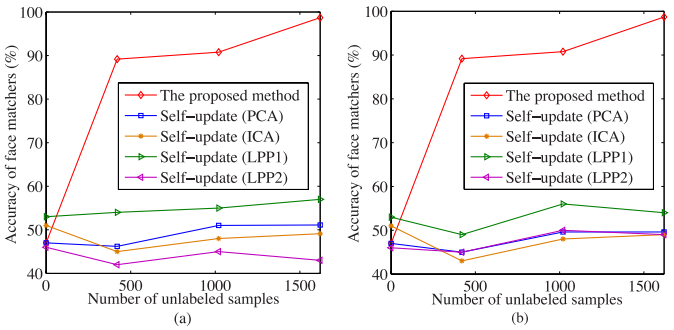


Fig. 5. Accuracy of face matchers using the proposed method and the Self-update methods, respectively, with the increment of imposed unlabeled samples. The threshold for the Self-update method to accept unlabeled samples is set to zero FAR and 1% FAR respectively in (a) and (b).

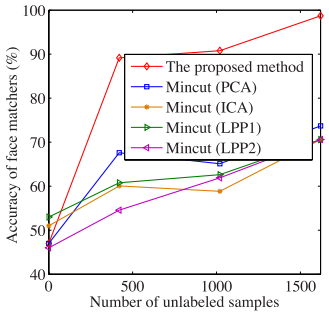


Fig. 6. Accuracy of face matchers using the proposed method and the Mincut methods, respectively, with the increment of imposed unlabeled samples.

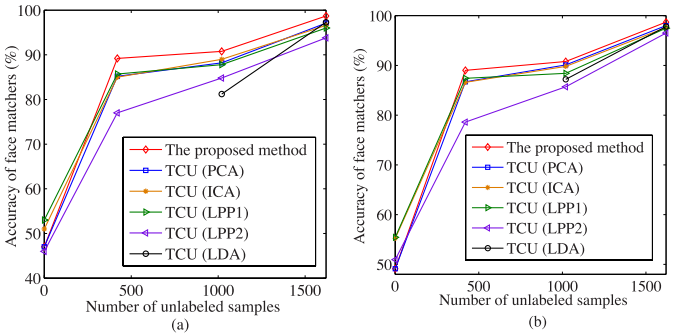


Fig. 7. Accuracy of face matchers using the proposed method and the TCU methods, respectively, with the increment of imposed unlabeled samples. The threshold for the TCU method to accept unlabeled samples is set to zero FAR and 1% FAR respectively in (a) and (b).

high intra-class variance, as is the case of this experiment, the promotion of the two methods is not significant. The Self-update method with some dimensionality reduction methods even fails the promotion. The Mincut method performs better than the Self-update method in admitting more variable template samples as has been verified in [46].

Different from the Self-update and the Mincut methods, the TCU method enhances the face matcher with the help of the fingerprint matcher. Hard samples for the face matcher are labeled by the fingerprint matcher and used to augment the template library. Still, our proposed method outperforms the TCU method. When 27 unlabeled sample pairs of each user are imposed, all the face matchers reach the highest performance. For the TCU method, the highest accuracy of the face matcher is 97.3 % (when LDA is used) and 98% (when PCA is used) on the zero and 1% FAR threshold settings, respectively. The accuracy of the face matcher for the proposed method reaches 98.7% at the highest point, which is 1.4% and 0.7% higher than the best matchers, respectively, with the TCU method.

D. Performance of the System

In this subsection, we testify the superiority of the proposed system (*i.e.*, the system with the proposed framework and SSL based method) by evaluating and comparing the performance of various systems in user convenience and recognition accuracy.

Firstly, we compare with a representative parallel fusion system that adopts the sum rule score level fusion method. For each use of the parallel fusion system, one face sample and one fingerprint sample from the testing set are input and checked by the face matcher and the fingerprint matcher, respectively. The final match score is the sum of the match scores from the two matchers. If the final match score is above a pre-defined threshold, the user is recognized as genuine; otherwise, an imposter. LPP1 (the best method for the face matcher on the template library) and the minutiae-based method in [52] are used for the face and the fingerprint matcher, respectively, in the parallel fusion system. In addition, we compare with the serial fusion systems using the proposed serial fusion strategy but different SSL based methods (*i.e.*, Self-update, Mincut and TCU) for the face matcher promotion. For the serial fusion systems using the Self-update and the TCU methods, respectively, zero FAR thresholds are set to admit unlabeled samples for face matcher promotion. For all the serial fusion systems, zero FAR thresholds are set for both the face and the fingerprint matchers to verify genuine users, which is an extremely strict setting for security.

In the same way as in Section V-C, the face matcher is promoted in three steps for each serial fusion system. After each step, we evaluate the performance of each serial fusion system with the testing set. For the parallel fusion system, we evaluate the performance just once with the testing set. There are 30 face samples and 30 fingerprints per user for 60 users in the testing set. Consequently, we can simulate 30 uses of the system per user, which sum up to 1800 uses in total.

Table II shows the experimental results on user convenience. We obtain the inconvenience score for a system by summing up the inconvenience scores of all uses of the traits. As introduced in Section V-B, each use of the face (fingerprint) trait has an inconvenience score of 1.9 (2.8). The baseline system is the parallel fusion system in which both matchers are needed in each use of the system. Therefore, it has the maximum inconvenience score of 8,460. With a serial fusion system, the ideal minimum inconvenience score is 3,420 when only the face matcher is used for every use of the system. As shown in Table II, the system with the proposed framework and SSL method is apparently superior to all the other systems. It should be noted that, for the serial fusion systems, the inconvenience scores are obtained with extremely strict thresholds (*i.e.*, zero FAR thresholds). In real

TABLE II
USER INCONVENIENCE OF DIFFERENT SYSTEMS AT EACH PROMOTION STEP OF THE FACE MATCHER

Systems		Number of uses of the face matcher			Number of uses of the fingerprint matcher			User inconvenience		
		step 1	step 2	step 3	step 1	step 2	step 3	step1	step2	step3
Parallel fusion system		1800			1800			8460.0		
The proposed serial fusion strategy	The proposed SSL method	1800	1800	1800	917	762	397	**5987.6**	**5553.6**	**4531.6**
	Self-update (PCA)	1800	1800	1800	1260	1054	1079	6948.0	6371.2	6441.2
	Self-update (ICA)	1800	1800	1800	1313	1129	1125	7096.4	6581.2	6570.0
	Self-update (LPP1)	1800	1800	1800	1341	1126	1103	7174.8	6572.8	6508.4
	Self-update (LPP2)	1800	1800	1800	1375	1129	1144	7270.0	6581.2	6623.2
	Mincut (PCA)	1800	1800	1800	1164	889	877	6679.2	5909.2	5875.6
	Mincut (ICA)	1800	1800	1800	1229	966	967	6861.2	6124.8	6127.6
	Mincut (LPP1)	1800	1800	1800	1145	953	955	6626.0	6088.4	6094.0
	Mincut (LPP2)	1800	1800	1800	1238	1048	946	6886.4	6354.4	6068.8
	TCU (PCA)	1800	1800	1800	1191	840	635	6754.8	5772.0	5198.0
	TCU (ICA)	1800	1800	1800	998	897	607	6214.4	5931.6	5119.6
	TCU (LPP1)	1800	1800	1800	991	894	644	6194.8	5923.2	5223.2
	TCU (LPP2)	1800	1800	1800	1217	943	653	6827.6	6060.4	5248.4
	TCU (LDA)	-	1800	1800	-	950	613	-	6080.0	5136.4

TABLE III
ACCURACY OF DIFFERENT SYSTEMS AT EACH PROMOTION STEP OF THE FACE MATCHER

Systems		Recognition accuracy (%)		
		step 1	step 2	step 3
Parallel fusion system		99.6		
The proposed serial fusion strategy	The proposed SSL method	**99.3**	**99.7**	**99.7**
	Self-update (PCA)	97.6	97.9	97.9
	Self-update (ICA)	97.5	97.6	97.9
	Self-update (LPP1)	97.8	98.0	98.0
	Self-update (LPP2)	97.5	97.5	97.6
	Mincut (PCA)	98.2	98.2	98.6
	Mincut (ICA)	97.9	98.0	98.4
	Mincut (LPP1)	98.0	98.2	98.4
	Mincut (LPP2)	97.5	97.9	98.4
	TCU (PCA)	98.8	99.1	99.4
	TCU (ICA)	99.1	99.4	99.6
	TCU (LPP1)	99.0	99.6	**99.7**
	TCU (LPP2)	98.9	99.6	99.6
	TCU (LDA)	-	98.9	99.6

applications, those systems will have lower user inconvenience since more users can be recognized by the face matcher as genuine users under looser thresholds.

Table III shows the experimental results on recognition accuracy. Samples that are wrongly recognized by the face or the fingerprint matcher are treated as errors. Samples that can not be recognized reliably at last are also treated as errors because the testing set is a closed dataset with no samples from outside users. The results indicate that all the systems yield fairly high accuracy. That is mainly because of the high performance of the fingerprint matcher. The capability of the face matcher also influences the accuracy. The accuracy of the system with the proposed framework and SSL method is higher than the other systems in most cases. Further, it is interesting to observe from Table III that the system with the proposed framework and SSL method even outperforms the parallel fusion system, when the face matcher is promoted in the last two steps.

One most concerned issue is often the computational complexity of a system. It should be noted that, under the proposed framework, the weaker trait enhancement process can run off-line, *i.e.*, when the system is not used for recognition, and therefore does not interfere with normal system use. The on-line computational efficiency of the system is dependent on the efficiency of concrete matchers deployed in the system. It in general is not a problem since most prevalent biometric matchers can respond in real time. Therefore, instead of a strict analysis of computational complexity, our proposed user convenience metric can be used to evaluate the system efficiency from the aspect of user cost. Another most concerned issue is the robustness of a system. Similarly, it depends on the robustness of the concrete matchers deployed in the system. Since our focus of experiment is on corroborating the effectiveness of the proposed framework and methodology in general, we do not investigate the robustness of specific matchers in this work.

VI. EXPERIMENTS ON GAIT AND FINGERPRINT

In this section, the proposed framework with the SSL based method is applied to a use case of a serial multimodal

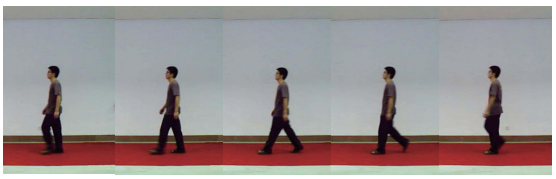


Fig. 8. Representative gait sequence in the database presented by 5 frames.

biometric system in which a gait matcher is deployed before a fingerprint matcher. The content and the configuration of the experiments are the same as in the former use case described in section V, except that LDA is not included in this use case for the limited samples.

A. Databases

The gait database is collected by SDUMLA in September, 2011. In the database, gait sequences of 25 users are collected with 40 gait sequences per user. For each gait sequence, a user is asked to walk for 10 meters, and a walking video of the user is captured by a digital camera set in a distance. We choose the first 50 frames in each video to represent a gait sequence. A representative gait sequence is shown in Fig. 8 with 5 frames chosen from the sequence. For each of the 25 users, 38 gait sequences are randomly chosen for the gait database. Contour features of gait sequences are extracted as in [53] to construct the original feature space for the gait trait. The fingerprint database is constructed by randomly choosing 25 users with 38 thumb fingerprints per user from the database collected by SDUMLA in September, 2011.

Finally, 25 virtual users are constructed by assigning the same identity to two users in the gait and the fingerprint databases, respectively. For each user, there are 38 gait samples and 38 fingerprints in the database. The samples are then divided into three parts as follows.

Part I: Training set. For each user, 3 gait sequences and 3 fingerprints are randomly selected as the labeled samples in the template library.

Part II: Testing set. 20 gait sequences and 20 fingerprints per user are randomly selected to form the testing set.

Part III: Unlabeled sample set. The rest 15 gait sequences and 15 fingerprints of each user are used as unlabeled samples collected with the use of the system.

B. Gait Matcher Enhancement

The fingerprint and gait matchers are trained with the training set and tested on the testing set. The accuracy of the fingerprint matcher reaches 99.2% while the accuracy of the best gait matcher is 69.2% when LPP1 is used. The fingerprint matcher is obviously stronger than the gait matchers when they are trained on the labeled samples.

To simulate the progressive weaker trait enhancement process with the use of the system, we impose 15 unlabeled gait and fingerprint pairs per user as the collected unlabeled sample pairs in three sequential batches at three steps, respectively. At each step, 5 gait and fingerprint pairs per user are selected randomly and imposed to simulate the uncontrollable recognition scenario.

The influence of thr in the DMDR method on the performance of the gait matcher is investigated. Results indicate the same trend as in the former use case, *i.e.*, larger thr values usually yield better performance, and when thr is above 0.9, the performance becomes stable.

We testify the performance of the gait matcher of the proposed method and the compared methods on the testing set at each promotion step. The accuracy of the gait matcher in the proposed method gets improved from 65.2% to 87.8%. As is analyzed in section V-C, when the imposed unlabeled samples have severe intra-class variance problem, the superiority of the proposed method is most apparent compared with the Self-update method and the Mincut method. The TCU method outperforms the other two compared methods and the proposed method outperforms the TCU method. When 15 unlabeled sample pairs per user are imposed, all the gait matchers get the highest performance. For the TCU method, the highest accuracy is 84.2% (when LPP1 is used) and 85.2% (when ICA is used) on the zero and 1% FAR thresholds, respectively. The accuracy of the gait matcher for the proposed method reaches 87.8%, which is 3.6% and 2.6% higher than the best matchers, respectively, with the TCU method.

C. Performance of the System

The content and configuration of the experiments in this section is the same as is described in section V-D. There are 20 gait samples and 20 fingerprints per user for 25 users in the testing set. Consequently, we can simulate 20 uses of the system per user, which sum up to 500 uses in total.

Table IV shows the experimental results on user convenience. As introduced in Section V-B, each use of the gait (fingerprint) trait has an inconvenience score of 1.2 (2.8). The baseline system is the parallel fusion system with the maximum inconvenience score of 2,000. The ideal minimum inconvenience score is 600 when only the gait matcher is used for every use of the system. As shown in Table IV, the system with the proposed framework and SSL method is apparently superior to all the other systems.

Table V shows the experimental results on recognition accuracy. The results indicate that all the systems yield fairly high accuracy. That is mainly because of the high performance of the fingerprint matcher. The capability of the gait matcher also influences the accuracy. The accuracy of the system with the proposed framework and SSL method is higher than the other serial fusion systems and is even not inferior to the parallel fusion system.

VII. Conclusion and Future Work

In this work we have proposed a novel framework for serial multimodal biometrics and introduced semi-supervised learning techniques into this framework, resulting in highly boosted user convenience and system performance at the same time.

In general, the serial multimodal biometric system suits common applications better since it does not force all traits to be captured and matched for each use of the system. In other words, whenever the user is recognized on a certain

TABLE IV
USER INCONVENIENCE OF DIFFERENT SYSTEMS AT EACH PROMOTION STEP OF THE GAIT MATCHER

Systems		Number of usages of the gait matcher			Number of usages of the fingerprint matcher			User inconvenience		
		step 1	step 2	step 3	step 1	step 2	step 3	step1	step2	step3
Parallel fusion system		500			500			2000		
The proposed serial fusion strategy	The proposed SSL method	500	500	500	355	227	166	**1594.0**	**1235.6**	**1064.8**
	Self-update (PCA)	500	500	500	454	453	454	1871.2	1868.4	1871.2
	Self-update (ICA)	500	500	500	420	451	416	1776.0	1862.8	1764.8
	Self-update (LPP1)	500	500	500	431	435	417	1806.8	1818.0	1767.6
	Self-update (LPP2)	500	500	500	487	467	483	1963.6	1907.6	1952.4
	Mincut (PCA)	500	500	500	427	376	427	1795.6	1652.8	1795.6
	Mincut (ICA)	500	500	500	395	340	323	1706.0	1552.0	1504.4
	Mincut (LPP1)	500	500	500	414	336	342	1759.2	1540.8	1557.6
	Mincut (LPP2)	500	500	500	464	402	429	1899.2	1725.6	1801.2
	TCU (PCA)	500	500	500	404	348	226	1731.2	1574.4	1232.8
	TCU (ICA)	500	500	500	383	299	207	1672.4	1437.2	1179.6
	TCU (LPP1)	500	500	500	405	329	221	1734.0	1521.2	1218.8
	TCU (LPP2)	500	500	500	439	374	219	1829.2	1647.2	1213.2

TABLE V
ACCURACY OF DIFFERENT SYSTEMS AT EACH PROMOTION STEP OF THE GAIT MATCHER

Systems		Recognition accuracy (%)		
		step 1	step 2	step 3
Parallel fusion system		**99.0**		
The proposed serial fusion strategy	The proposed SSL method	**96.4**	**98.6**	**99.0**
	Self-update (PCA)	95.2	96.4	96.4
	Self-update (ICA)	95.4	96.4	96.6
	Self-update (LPP1)	94.4	95.6	96.4
	Self-update (LPP2)	93.8	94.2	94.2
	Mincut (PCA)	95.4	96.2	96.6
	Mincut (ICA)	95.2	95.8	96.4
	Mincut (LPP1)	95.9	96.6	97.0
	Mincut (LPP2)	93.6	94.2	94.0
	TCU (PCA)	95.4	97.4	98.2
	TCU (ICA)	96.2	97.0	97.6
	TCU (LPP1)	96.0	97.2	97.7
	TCU (LPP2)	93.8	94.2	95.2

trait, the rest traits do not have to be captured and matched. Compared with the other existent serial multimodal biometric systems, the proposed one is novel in that it for the first time proposes to always order the traits based on their extents of user convenience, with more user convenient traits positioned earlier in the chain. By doing this, the user convenience is maximized since users prefer to use convenient traits whenever possible.

More user convenient traits are, however, usually weaker ones that yield lower recognition accuracy. Therefore, for the proposed framework to work successfully, the recognition performance of the weaker traits set earlier in the chain must be improved. An SSL based method has been proposed for the promotion. The core methodology in the SSL method is to improve a weaker trait's discriminating capability by learning from a tightly coupling stronger trait in the same chain with the utility of both labeled and unlabeled data. Specifically, we have proposed to transform the weaker trait's features to a lower dimensional space such that the information dependence between the transformed weaker trait features and the stronger trait features is maximized. The dimensionality reduction method, *i.e.*, the DMDR method, is the central method that we have used to achieve this purpose. In addition, a novel labeling strategy is also included in the SSL method to assign pseudo labels to unlabeled samples, which in turn facilitate future recognition tasks.

The proposed framework and methodology have been implemented in two prototype multimodal biometric systems – one with a face and a fingerprint matchers and the other with a gait and a fingerprint matchers. Experimental results verified the effectiveness and the superiority of the proposed SSL based method in promoting the weaker trait matcher. The accuracy of the face matcher is promoted from 49.2% to 98.7% in the former system, and the accuracy of the gait matcher is promoted form 65.2% to 87.8% in the latter. Further, experimental results indicate that the system with the proposed framework and methodology is superior in user convenience and recognition accuracy, when compared with the parallel fusion system and other serial fusion systems using different SSL based methods.

We currently focus on biometric systems using two traits. In the future, we plan to extend the proposed framework and SSL based method to process more biometric traits in a chain. Further, it is interesting to investigate other advanced dimen-

sionality reduction techniques in the proposed SSL process, and even other advanced machine learning techniques to promote the weaker trait's recognition capability.

REFERENCES

[1] A. A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of Multibiometrics* (International Series on Biometrics). New York, NY, USA: Springer-Verlag, 2006.

[2] L. Hong and A. Jain, "Integrating faces and fingerprints for personal identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 12, pp. 1295–1307, Dec. 1998.

[3] K. Chang, K. W. Bowyer, S. Sarkar, and B. Victor, "Comparison and combination of ear and face images in appearance-based biometrics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1160–1165, Sep. 2003.

[4] T. Zhang, X. Li, D. Tao, and J. Yang, "Multimodal biometrics using geometry preserving projections," *Pattern Recognit.*, vol. 41, no. 3, pp. 805–813, Mar. 2008.

[5] A. A. Ross and R. Govindarajan, "Feature level fusion using hand and face biometrics," *Proc. SPIE*, vol. 5779, pp. 196–204, Mar. 2005.

[6] V. Chatzis, A. G. Bors, and I. Pitas, "Multimodal decision-level fusion for person authentication," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 29, no. 6, pp. 674–680, Nov. 1999.

[7] K. Veeramachaneni, L. A. Osadciw, and P. K. Varshney, "An adaptive multimodal biometric management algorithm," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 35, no. 3, pp. 344–356, Aug. 2005.

[8] A. Ross and A. Jain, "Information fusion in biometrics," *Pattern Recognit. Lett.*, vol. 24, no. 13, pp. 2115–2125, Sep. 2003.

[9] S. Ribaric and I. Fratric, "A biometric identification system based on eigenpalm and eigenfinger features," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 11, pp. 1598–1709, Nov. 2005.

[10] B. Gutschoven and P. Verlind, "Multi-modal identity verification using support vector machines (SVM)," in *Proc. ICIF*, Paris, France, Jul. 2000, pp. THB3/3–THB3/8.

[11] K.-A. Toh, J. Kim, and S. Lee, "Biometric scores fusion based on total error rate minimization," *Pattern Recognit.*, vol. 41, no. 3, pp. 1066–1082, Mar. 2008.

[12] K. Nandakumar, Y. Chen, S. C. Dass, and A. K. Jain, "Likelihood ratio-based biometric score fusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 342–247, Feb. 2008.

[13] A. K. Jain and A. Ross, "Learning user-specific parameters in a multibiometric system," in *Proc. ICIP*, Sep. 2002, pp. 57–60.

[14] R. Snelick, U. Uludag, A. Mink, M. Indovina, and A. Jain, "Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 450–455, Mar. 2005.

[15] K.-A. Toh, X. Jiang, and W.-Y. Yau, "Exploiting global and local decisions for multimodal biometrics verification," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 3059–3072, Oct. 2004.

[16] J. Fierrez-Aguilar, D. Garcia-Romero, J. Ortega-Garcia, and J. Gonzalez-Rodriguez, "Bayesian adaptation for user-dependent multimodal biometric authentication," *Pattern Recognit.*, vol. 38, no. 8, pp. 1317–1319, Aug. 2005.

[17] J. Fierrez-Aguilar, J. Ortega-Garcia, D. Garcia-Romero, and J. Gonzalez-Rodriguez, "A comparative evaluation of fusion strategies for multimodal biometric verification," in *Proc. 4th Int. Conf. AVBPA*, Jun. 2003, pp. 830–837.

[18] J. Fierrez-Aguilar, D. Garcia-Romero, J. Ortega-Garcia, and J. Gonzalez-Rodriguez, "Adapted user-dependent multimodal biometric authentication exploiting general information," *Pattern Recognit. Lett.*, vol. 26, no. 16, pp. 2628–2639, Dec. 2005.

[19] J. Fierrez-Aguilar, J. Ortega-Garcia, J. Gonzalez-Rodriguez, and J. Bigun, "Discriminative multimodal biometric authentication based on quality measures," *Pattern Recognit.*, vol. 38, no. 5, pp. 777–779, May 2005.

[20] D. E. Maurer and J. P. Baker, "Fusing multimodal biometrics with quality estimates via a Bayesian belief network," *Pattern Recognit.*, vol. 41, no. 3, pp. 821–832, Mar. 2008.

[21] A. Kumar, V. Kanhangad, and D. Zhang, "A new framework for adaptive multimodal biometrics management," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 1, pp. 92–102, Mar. 2010.

[22] F. Roli, L. Didaci, and G. L. Marcialis, "Template co-update in multimodal biometric systems," in *Proc. ICB*, Seoul, Korea, Aug. 2007, pp. 1194–1202.

[23] F. Roli, L. Didaci, and G. L. Marcialis, *Advances in Biometrics: Sensors, Algorithms and Systems*. London, U.K.: Springer-Verlag, 2008.

[24] L. Didaci, G. L. Marcialis, and F. Roli, "Semi-supervised co-update of multiple matchers," in *Proc. 8th Int. Workshop MCS*, Reykjavik, Iceland, Jun. 2009, pp. 152–160.

[25] A. Rattani, G. L. Marcialis, and F. Roli, "Capturing large intra-class variations of biometric data by template co-updating," in *Proc. IEEE CVPRW*, Anchorage, AK, USA, Jun. 2008, pp. 1–6.

[26] L. Didaci, G. L. Marcialis, and F. Roli, "A theoretical and experimental analysis of template co-update in biometric verification systems," in *Proc. S+SSPR*, Orlando, FL, USA, Dec. 2008, pp. 745–754.

[27] G. L. Marcialis, A. Rattani, and F. Roli, "Biometric template update: An experimental investigation on the relationship between update errors and performance degradation in face verification," in *Proc. S+SSPR*, Orlando, FL, USA, Dec. 2008, pp. 684–693.

[28] J. Yang and X. Zhang, "Feature-level fusion of fingerprint and finger-vein for personal identification," *Pattern Recognit. Lett.*, vol. 33, no. 5, pp. 623–628, Apr. 2012.

[29] S. Shekhar, V. M. Patel, N. M. Nasrabadi, and R. Chellappa, "Joint sparse representation for robust multimodal biometrics recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 1, pp. 113–126, Jan. 2014.

[30] A. K. Jain, S. C. Dass, and K. Nandakumar, "Soft biometric traits for personal recognition systems," in *Proc. ICBA*, Hong Kong, Jul. 2004, pp. 731–738.

[31] U. Park and A. K. Jain, "Face matching and retrieval using soft biometrics," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 3, pp. 406–415, Sep. 2010.

[32] K. Moustakas, D. Tzovaras, and G. Stavropoulos, "Gait recognition using geometric features and soft biometrics," *IEEE Signal Process. Lett.*, vol. 17, no. 4, pp. 367–370, Apr. 2010.

[33] S. Tiwari, A. Singh, and S. K. Singh, "Integrating faces and soft-biometrics for newborn recognition," *Int. J. Adv. Comput. Eng. Archit.*, vol. 2, no. 2, pp. 201–209, Jun. 2012.

[34] X. Zhou, B. Bhanu, and J. Han, "Human recognition at a distance in video by integrating face profile and gait," in *Proc. AVBPA*, New York, NY, USA, Jul. 2005, pp. 533–543.

[35] G. L. Marcialis and F. Roli, "Serial fusion of fingerprint and face matchers," in *Proc. 7th Int. Workshop MCS*, May 2007, pp. 151–160.

[36] K. Takabashi, M. Mimura, Y. Isobe, and Y. Seto, "A secure and user-friendly multimodal biometric system," *Proc. SPIE*, vol. 5404, pp. 12–19, Apr. 2004.

[37] G. L. Marcialis, F. Roli, and L. Didaci, "Personal identity verification by serial fusion of fingerprint and face matchers," *Pattern Recognit.*, vol. 42, no. 11, pp. 2807–2817, Nov. 2009.

[38] G. L. Marcialis, P. Mastinu, and F. Roli, "Serial fusion of multi-modal biometric systems," in *Proc. BIOMS*, Taranto, Italy, Sep. 2010, pp. 1–7.

[39] A. Lore, D. Bernadette, and G. Garcia-Sonia, "Tuning cost and performance in multi-biometric systems: A novel and consistent view of fusion strategies based on the sequential probability ratio test (SPRT)," *Pattern Recognit. Lett.*, vol. 31, no. 9, pp. 884–890, Jul. 2010.

[40] Z. Akhtar, G. Fumera, G. L. Marcialis, and F. Roli, "Evaluation of serial and parallel multibiometric systems under spoofing attacks," in *Proc. IEEE 5th BTAS*, Washington, DC, USA, Sep. 2012, pp. 283–288.

[41] J. L. Wayman, "A path forward for multi-biometrics," in *Proc. ICASSP*, Toulouse, France, May 2006, pp. V1069–V1072.

[42] A. Gretton, O. Bousquet, A. Smola, and B. Schölkopf, "Measuring statistical dependence with Hilbert–Schmidt norms," in *Proc. 16th Int. Conf. ALT*, Oct. 2005, pp. 63–77.

[43] X. Tan, S. Chen, Z.-H. Zhou, and F. Zhang, "Face recognition from a single image per person: A survey," *Pattern Recognit.*, vol. 39, no. 9, pp. 1725–1745, Sep. 2006.

[44] F. Roli and G. L. Marcialis, "Semi-supervised PCA-based face recognition using self-training," in *Proc. SSPR+SPR*, Hong Kong, Aug. 2006, pp. 560–568.

[45] A. Rattani, G. L. Marcialis, and F. Roli, "Biometric template update using the graph mincut algorithm : A case study in face verification," in *Proc. BSYM*, Tampa, FL, USA, Sep. 2008, pp. 23–28.

[46] A. Rattani, G. L. Marcialis, and F. Roli, "Biometric system adaptation by self-update and graph-based techniques," *J. Vis. Lang. Comput.*, vol. 24, no. 1, pp. 1–9, Feb. 2013.

[47] *FacePix Database*. [Online]. Available: http://www.facepix.org/, accessed Feb. 23, 2011.

[48] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognit. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.

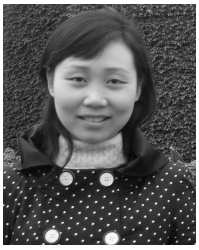
[49] P. N. Belhumeur, J. P. Hepanha, and D. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[50] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Trans. Neural Netw.*, vol. 13, no. 6, pp. 1450–1464, Nov. 2002.

[51] X. He and P. Niyogi, "Locality preserving projections," in *Advances in Neural Information Processing Systems*. Cambridge, MA, USA: MIT Press, Dec. 2003, pp. 153–160.

[52] X. D. Jiang and W.-Y. Yau, "Fingerprint minutiae matching based on the local and global structures," in *Proc. 15th ICPR*, Barcelona, Spain, Sep. 2000, pp. 1038–1041.

[53] L. Liu, Y. Yin, W. Qin, and Y. Li, "Gait recognition based on outermost contour," *Int. J. Comput. Intell. Syst.*, vol. 4, no. 5, pp. 1090–1099, Sep. 2011.



Qing Zhang received the B.S. and M.S. degrees from the School of Computer Science and Technology, Shandong University, Jinan, China, in 2005 and 2008, respectively, where she is currently pursuing the Ph.D. degree in computer application technology. Her main research interests are biometrics, machine learning, and data mining.

Yilong Yin is the Director of the Machine Learning and Applications Group and a Professor with Shandong University, Jinan, China. He received the Ph.D. degree from Jilin University, Changchun, China, in 2000. From 2000 to 2002, he was a Post-Doctoral Fellow with the Department of Electronic Science and Engineering, Nanjing University, Nanjing, China. His research interests include machine learning, data mining, and biometrics.

De-Chuan Zhan received the Ph.D. degree in computer science from Nanjing University, Nanjing, China, in 2010. He became a Faculty Member with the Department of Computer Science and Technology, Nanjing University, in 2010, where he is currently an Associate Professor. His research interests are mainly in machine learning, data mining, and mobile intelligence. He and other LAMDA members received the Grand Champion in the PAKDD 2006 Data Mining Competition.

Jingliang Peng received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 2006, and the B.S. and M.S. degrees in computer science from Peking University, Beijing, China, in 1997 and 2000, respectively. He is currently a Professor with the School of Computer Science and Technology, Shandong University, Jinan, China. His research interest mainly resides in digital geometry processing, 3-D animation, and image/video analysis.