

A Dual-staged Classification-Selection Approach for Automated Update of Biometric Templates

Ajita Rattani, Gian Luca Marcialis, Eric Granger* and Fabio Roli

Dept. of Electrical and Electronic Engineering, University of Cagliari, Italy

** École de technologie supérieure, Université du Québec, Montreal, Canada*

{ajita.rattani, marcialis, roli}@diee.unica.it, Eric.Granger@etsmtl.ca

Abstract

In the emerging field of adaptive biometrics, systems aim to adapt enrolled templates to variations in samples observed during operations. However, despite numerous advantages, few commercial vendors have adopted auto-update procedures in their products. This is due to limitations associated with existing adaptation schemes. This paper proposes a dual-staged template adaptation scheme that allows to capture ‘informative’ operational samples with significant variations but without increasing the vulnerability to impostor intrusion. This is achieved through a two staged classification-selection approach driven by the harmonic function and risk minimization technique, over a graph based representation of (enrolment and operational) samples. Experimental results on the D1EE fingerprint data set, explicitly collected for evaluating adaptive biometric systems, demonstrate that the proposed scheme results in 67% reduction in error over the baseline system (without adaptation), outperforming state-of-the-art methods.

1. Introduction

An intrinsic characteristic of the biometric verification technology is that a system’s error rate, e.g., the false accept rate (FAR), false reject rate (FRR) and equal error rate (EER) (the rate at which FAR is equal to FRR), simply cannot attain the absolute zero. The major cause of these errors is the compound effect of the scarcity of training samples during the enrollment phase as well as the presence of substantial sample variations (e.g., human sensor interaction, temporal variations etc.) during operations [6, 7]. An important consequence of these factors is that a biometric template (obtained during enrollment) cannot be expected

to fully represent a person’s identity.

Solutions in the form of adaptive biometrics have been introduced to address this issue of template representativeness [7]. Adaptive biometric systems attempt to update template galleries by integrating information captured in input operational samples.

For systems operating in semi-supervised mode, a scheme commonly adopted in adaptive biometrics is *self-update* [7, 3], where the system adapts itself to the variation of the highly confident input samples, to avoid impostor intrusion. Self-adaptive systems may operate in online mode (the adaptation process follows verification) or off-line mode (the adaptation process is carried out independently after some time). In either case, when a new batch of input samples becomes available, the update process is referred to as an *update cycle*. In [8], authors propose an off-line graph *min-cut* based scheme for adaptation, where input samples are assigned binary labels upon finding the minimum cut over the graph-based representation of the samples. These adaptive systems may also operate in supervised mode where the labels are assigned off-line by an operator or analyst.

Adaptive biometric systems have several advantages. Firstly, one no longer needs to collect a large number of templates during the enrollment process. Secondly, it is no longer necessary to re-enrol or re-train the system from the scratch in order to cope up with changing environments [7, 5]. Despite these advantages, few biometric vendors like *BIOsingle* (fingerprint) and *Recogsys* (hand geometry) have incorporated adaptation mechanism into their technologies. This is due in large part to the following limitations associated with the existing adaptive biometric systems:

1. **Inability to capture substantial amount of samples.** Commonly adopted self-update schemes can capture only limited amount of available samples [7]. This is due to the operation at the stringent threshold. As a consequence, large number of in-

formative input samples with significant variations remain unexploited and limited performance improvement is obtained over the baseline system.

2. **No distinction between informative, redundant and noisy samples.** Existing adaptation schemes (even supervised ones) cannot distinguish between informative, redundant and noisy input samples [7]. Input samples with significant variations can be both informative and noisy. For instance, fingerprint sample obtained through dirty sensor is an example of noisy sample with significant variation and should not be used for adaptation.
3. **Vulnerability to attacks.** Adaptive biometric systems are vulnerable to impostor attacks [7, 4] due to the intrinsic failure of the system (causing FAR). Adaptation using impostor samples may result in the counter-productive effect over time.

This paper proposes a dual-staged template adaptation scheme using probabilistic semi-supervised learning [9]. Specifically, the input samples are assigned the probability of belonging to genuine class in the first (classification) stage, and the second (selection) stage chooses the most informative samples based on the risk minimization criteria. Compared to state-of-the-art methods, this scheme provides the advantage of mitigating the first two limitations by capturing substantial amount of *informative* samples through efficient labeling and sample selection criteria. These advantages result in a higher level of performance for the biometric system over existing adaptation techniques.

Section 2 explains the proposed dual-staged template adaptation scheme, and Section 3 presets proof-of-concept experiments and discussions.

2. A classification-selection technique for off-line template adaptation

The proposed dual-staged approach uses the labeling scheme based on probabilistic semi-supervised learning introduced in [9]. Specifically, soft probabilistic labels (f_u) are assigned to each batch of input samples (U), by finding the minimum energy (harmonic) function (f) on the graphical representation. This harmonic function (f) is *unique* and ensures that labels to input samples are assigned using both the enrolled and nearby input data. Then the genuinely classified samples (from each U) undergo the selection process based on the risk minimization process. The proposed approach follows the steps below, and is summarized in the Algorithm 1. In the algorithm 1, steps 1-3 and 4-5 are executed in the classification and selection stages, respectively.

A) Classification Stage:

1. Given L enrolled, U input batch of samples and $N = \{L \cup U\}$. An adjacency matrix W is computed where $w_{i,j} = \text{matchscore}(n_i, n_j)$.
2. The minimum harmonic energy function f , using W , is computed using the following matrix method [9]. Laplacian matrix, Δ , is obtained by $\Delta = D - W$. D is the diagonal matrix defined as $D = \text{diag}(d_i)$ whose entries $d_i = \sum_{i,j} w_{i,j}$ are the weighted degrees of the nodes in G . Matrix Δ is partitioned into blocks for enrolled (l) and input data samples (u) as follows:

$$\Delta = \begin{pmatrix} \Delta_{ll} & \Delta_{lu} \\ \Delta_{ul} & \Delta_{uu} \end{pmatrix}$$

The harmonic function, f_u , for the input samples can be given as:

$$f_u = \Delta_{uu}^{-1} \times \Delta_{ul} \times f_l \quad (1)$$

where f_l =labels of the enrolled templates. f_u lies in the range [0,1] and interpreted as the posterior probability of the samples being genuine.

3. Classification of the input samples, U , is done through Bayes decision rule [1]. Input samples, $u(i) \in U$, are considered positive and added to set P (i.e., genuine) in case $f_u(i) > 0.5$.

B) Selection Stage:

1. The informative samples (among P) are selected based on minimizing the estimated expected risk as follows. Firstly, the estimated Bayes risk [1] ($R(f)$) of the baseline classifier (enrolled on initial templates (L)) in classifying input samples (U) is computed as:

$$R(f) = \sum_{i=1}^n \min(f_i, 1 - f_i) \text{ for } n \in N \quad (2)$$

It is important to note that f_i for enrolled samples remain fixed integer values.

2. Then only those input samples, $x_k \in P$, are used whose addition to the enrolled templates, L , reduces the risk of the classifier denoted by $R(f^{+(x_k)})$ ($f^{+(x_k)}$ can be efficiently recomputed using matrix method [9]) as:

$$R(f^{+(x_k)}) = \sum_{i=1}^n \min(f_i^{+(x_k)}, 1 - f_i^{+(x_k)}) \quad (3)$$

Algorithm 1 Classification-Selection Technique for Off-line Template Adaptation.

- Given: L templates, U input set and $N = \{L \cup U\}$.
 1. Compute adjacency matrix W using N
 2. Compute harmonic function f_u for input samples U using (1)
 3. Classify the input samples $u_i \in U$ as positive, P , if $f_u(i) > 0.5$.
 4. Estimate the risk of the classifier using (2)
 5. While (risk can be minimized)
 - Find sample $x_k \in P$ using (4)
 - Add (x_k) to L , remove x_k from PEnd While
 - Output: the updated template set, L .
-

The iterative procedure of input sample, x_k , selection is followed until the risk can be minimized (stopping criteria) as:

$$k = \arg \min^k R(f^{+(x_k)}) \quad (4)$$

For a memory limited system, iterative procedure of input sample selection may be repeated only a fixed number of times. The proposed scheme mitigates the limitation of few sample capture through efficient labeling (under classification) scheme. Further the efficient sample selection criteria distinguishes between redundant, noisy and informative intra-class variations.

The proposed scheme differs from self-update [7] in the sense that the latter labels the input samples using only the enrolled templates. Thus self-update's capability to capture input samples depends on the representativeness of the initial templates [7]. The proposed scheme bears some similarity with min-cut based scheme [8] in assigning labels using both the enrolled and nearby input samples. Yet the min-cut based scheme performs classification on finding minimum cut on the graph (as opposed to harmonic function in the proposed scheme). Moreover, the obtained minimum cut may not always be a unique solution [8] and the binary labeling procedure does not allow us to invoke the selection process.

3. Experimental Validations

For proof-of-concept experiments, simulations were performed using the D1EE fingerprint [6] dataset. It is

comprised of 49 individuals with 90 samples per individuals collected over 1.5 years. These 90 samples are acquired in nine sessions with 10 samples per session (batch) in a gap of every three weeks on an average. Variations like rotation, changes in pressure, non-linear deformations are introduced in each batch of samples. This data set has been explicitly collected for evaluating adaptive biometric systems [6].

To simulate template adaptation in real systems, the first two images from the first session (batch b_1) are assigned as templates (L) and the rest other images from first session and sessions two to eight ($b_{2:8}$) are assigned as input batches (U), available over time. Performance of the adaptive system is evaluated after every update cycle. That is the system adapted using each batch b_i (U) is evaluated using the next batch b_{i+1} and EER_i (Equal Error Rate) computed. Impostor scores (for EER_i) are computed using the input batch b_{i+1} of the rest other users. The overall process resulted in 8 update cycles. Templates are updated by appending qualified input samples from U (using algorithm 1) to the template set. To evaluate the impact of impostor intrusion, 5 randomly selected impostor samples are added to each batch, b_i , used for adaptation. Threshold for self-update is set at 0.001% false acceptance rate (FAR) on impostor distribution obtained on comparing each initial template to the templates of rest other users. Next we present the obtained results.

Figure 1 shows the EER values obtained at different update cycles for the proposed scheme in comparison to baseline (enrolled on initial two templates), self-update [7], graph min-cut [8] and supervised scheme (using all the available genuine samples) to adaptation [7]. Average performance gain (EER reduction over 8 cycles) over the baseline is 2.5%, 9.7%, 64.1% and 67% for self-update, graph min-cut, supervised and the proposed scheme, respectively. It can be seen that the proposed adaptation scheme outperforms all the existing schemes by significant error reduction over the baseline.

Table 1 shows the percentage of genuine samples added, among 490 (392 in the 1st cycle), for all the update cycles. It can be seen that self-update captures the least amount of samples i.e., only 5.4% samples on an average, due to stringent threshold condition (i.e., highly confidently classified sample selection). Graph min-cut is able to capture more amount of samples over self-update i.e., 58.8% of the available samples. On contrary, the proposed scheme could capture maximum amount of samples i.e., 87.6% (in classification stage), mitigating the limitation of limited sample capture usually associated with self-update schemes. Then the selection stage chooses only 33.9% of the samples deemed informative over all the update cycles.

Table 1. % of genuine samples, among 490, added to the template set at each update cycle for all adaptation schemes.

Update Cycle	Self-update (%)	Graph min-cut (%)	Proposed (%)	
			Classification	Selection
1	26	66	68	58
2	6	65	85	29
3	1	64	86	32
4	2	63	93	29
5	2	54	92	28
6	2	51	92	32
7	2	55	90	32
8	2	52	95	31

Our conventional wisdom suggests that efficacy of the proposed scheme over graph min-cut [8] and supervised method [7] (see Figure 1) is due to its capability to select optimum set of informative samples. It can be verified from Table 1 that the selection stage of the proposed scheme chooses lesser amount of samples than graph min-cut and supervised scheme. Efficacy of our proposed scheme over supervised adaptation is also in agreement with the existing filter-based off-line template selection schemes [2]. These schemes [2] emphasize that systematic selection of the subset of templates (i.e., discarding noisy and redundant templates) is critical to the performance of the biometric system.

It is also worth remarking that the number of users affected by impostor intrusion (i.e., 4.1%) are almost similar to other adaptation schemes. Impostor intrusion is a vital issue for adaptive biometric systems and needs a separate modeling [4, 7]. These experimental results validated the claims of the proposed template adaptation scheme mentioned in the outset.

4. Conclusion

This paper proposed a dual-staged off-line template adaptation scheme, derived through harmonic function and risk minimization criteria. Compared to the existing schemes, the proposed scheme results in higher error reduction (67%) over baseline possible through the capture of many samples followed by the systematic selection of informative ones in the classification and selection stages. Although we used fingerprint based biometrics, the proposed scheme may be applied to other biometric traits (such as face and iris) as well.

Acknowledgement: This work has been partially supported by a grant awarded to Ajita Rattani by Regione Autonoma della Sardegna, PO Sardegna FSE 2007-2013, L.R. 7/2007 “Promotion of the scientific research and technological innovation in Sardinia”, and by the Natural Sciences and Engineering Research Council of Canada.

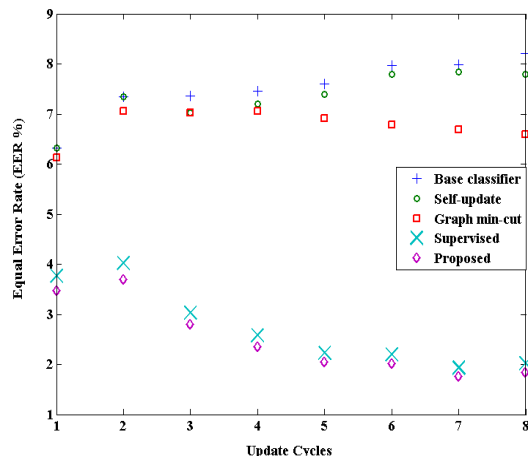


Figure 1. EER values obtained at each update cycle after adaptation using batch b_i , for all adaptation schemes.

References

- [1] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley-Interscience Publication, 2000.
- [2] A. Jain, U. Uludag, and A. Ross. Biometric template selection: a case study in fingerprints. In *Proc. of the 4th Int'l Conf. on Audio and Video based person authentication*, LNCS 2688, pages 335–342, Guildford, UK, 2003.
- [3] X. Jiang and W. Ser. Online fingerprint template improvement. *IEEE Tran. Pattern Analysis and Machine Intelligence*, 24(8):1121–1126, 2002.
- [4] 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. of Joint IAPR Int'l Workshop on SSPR and SPR*, LNCS 5342, pages 684–693, Florida, USA, 2008.
- [5] N. Poh, R. Wong, J. Kittler, and F. Roli. Challenges and research directions for adaptive biometric recognition systems. In *Proc. of the 3rd Int'l Conf. on Biometrics*, pages 753–764, Sardinia, Italy, 2009. LNCS,5558.
- [6] A. Rattani. *Adaptive Biometric System based on Template Update Procedures*. PhD thesis, University of Cagliari, Italy, 2010.
- [7] A. Rattani, B. Freni, G. L. Marcialis, and F. Roli. Template update methods in adaptive biometric systems: A critical review. In *Proc. of the 3rd Int'l Conf. on Biometrics*, LNCS 5558, pages 847–856, Sardinia, Italy, 2009.
- [8] A. Rattani, G. L. Marcialis, and F. Roli. Biometric template update using the graph mincut: a case study in face verification. In *Proc. of the 6th IEEE Biometrics Symposium*, pages 23–28, Tampa, USA, 2008.
- [9] X. Zhu, Z. Ghahramani, and J. D. Lafferty. Semi-supervised learning using gaussian fields and harmonic functions. In *Proc. of the Int'l Conf. on Machine Learning*, pages 912–919, Washington, USA, 2003.