A Classification-Selection Approach for Self Updating of Face Verification Systems Under Stringent Storage and Computational Requirements

Pierluigi Tuveri, Valerio Mura(✉), Gian Luca Marcialis, and Fabio Roli

Department of Electrical and Electronic Engineering,
University of Cagliari, Cagliari, Italy
{pierluigi.tuveri,valerio.mura,marcialis,roli}@diee.unica.it

Abstract. Nowadays face recognition systems have many application fields. Unfortunately, lighting variations and ageing effects are still open issues. Moreover, face changes over time due to ageing. A further problem is due to occlusions, for example the glass presence. Re-enrolling user's face is time-consuming and does not solve above problems. Therefore, unsupervised template update has been proposed, and named self update. Basically, this algorithm adapts/modifies templates or face models by collecting samples during system operations. The most effective variant of self update is based on the collection of multiple templates. However, this approach has been evaluated and tested in conditions under which the possible number of collectable templates is uncostrained. Actually, available resources are limited in memory and computational power, thus it is likely that it is not possible to have more than a pre-set number of templates. In this paper, we propose a classification-selection approach, based on the combination of self update and C-means algorithms, which keeps constant the number of templates and improve the ratio between intra-class variations and inter-class variations for each user. Experimental results show the effectiveness of this method with respect to standard self update.

1 Introduction

Face recognition is not a time-invariant process, because traits change over time, for instances due to expressions, illumination environment, and so on. In order to recognize a face the system needs to have a set of templates for each user to cover all the possible variations. A periodical re-enrollment phase can only partially follows these variations, and is time-consuming. The idea that the system can be automatically updated without the users cooperation has been addressed in the self update algorithm and its variants [9]. Basically, the system adds a new template if the match score is over a pre-calculate "updating" threshold t^*. This is a fundamental parameter, that biases the performance of the system. A low t^* allows the collection of more templates, but there is a high probability that

V. Murino and E. Puppo (Eds.): ICIAP 2015, Part II, LNCS 9280, pp. 540–550, 2015.
DOI: 10.1007/978-3-319-23234-8_50

the system associates the template to a wrong user. When the threshold is more selective it does not allow adding a face template with high variation respect to stored ones.

Self update has been always applied under the hypothesis that the potential number of possible templates is infinite [4]. This allows to point out at which extent self update can be useful; on the other hand, it is not a realistic hypothesis. Moreover, during collection, noisy and redundant samples can be added to the system. Therefore, we believe that self update should be analyzed under more realistic conditions. For example under limited storage and computational resources. The comparison between a large amount of templates makes slow the identification module. Another problem is the management of the memory, that is a precious and critical resource.

Actually, state of the art is aware of points above. With regard to the limited storage and the computational resources, it has been proposed to combine more templates [3], or to use replacement algorithms when the maximum number of templates is reached [4]. With regard to the problem of noisy, redundant or wrong samples, a dual stage method that collect and select the best templates has been proposed [5].

In this paper, we propose a classification-selection method dealing with both issues. First stage the system acts as standard self update, by collecting a set of possible templates under a stringent updating threshold. Second stage adopts the C-means algorithm, under the assumption that the templates of the users are probabilistically and geometry paritional in mono-modal clusters. Centroids of these clusters are assumed to be the centroid of the population associated to a given user. Experimental results on a standard benchmark data set show the effectiveness of the proposed method.

The paper is organized as follow. After a brief review of the self update approach and its variants (Section 2), we present our method (Section 3). Experiments are described in Section 4, and conclusions are reported in Section 5.

2 State of the Art of Self-Update

Let $T = \{t_{1,1}, t_{1,2}, ..., t_{l,j}, ..., t_{k,p}\}$ be a set of enrolled users in the system, named gallery. Where the pair l, j indicates $l - th$ user and $j - th$ template, p is the maximum allowed number of face templates per user, and k the number of enrolled users which are listed in the set $u = \{u_1, ..., u_j, ..., u_k\}$.

Let $b_i = \{x_{i_{1,1}}, x_{i_{1,2}}, ..., x_{i_{l,j}}, ..., x_{i_{k,n_k}}\}$ be a batch of input faces collected during system's operations, at a certain time. We added the i pedix in order to mean that b_i is collected before b_{i+1}.

The standard self-update algorithm estimates t^* through T. Usually the threshold is set such that only a small percentage of impostors can be wrongly classified (false accepptance rate or FAR). This operating point estimates how many genuines users and impostors can pass the verification process. If a certain input sample exhibits a distance from the template $d < t^*$, it is added to the gallery of the claimed user. Obviously if t^* is such that $FAR(t^*) > 0$, a small percentage of impostors may be included in the user's gallery, thus dropping the system's performance over time. In other words, self update algorithm is as follows:

estimate the update threshold t^* using T
for all e element in u_i **do**
  if $distance(x \in u_i, T) < t^*$ **then**
    $T_{new} = T \cup x$
  end if
end for

This algorithm iterates at each collected batch of samples.

Rattani et al. [5] introduced a system composed by two stages. Firstly, samples are classified on the basis of a minimum energy function. Secondly, accepted samples are selected on the basis of a pre-defined and minimized risk function. This two-stages approach is finalized at limiting the number of impostors in the user's gallery, under the usual condition that $p \rightarrow inf$ [5].

In [4] the authors take into account the problem that p is finite by replacing the less frequently used samples in order to maintain p samples in each user's gallery. Unfortunately, using these protocols we risk replacing the useful template, in this case we don't care about the intra-class variation, because we have a rules about the time of face template [4].

3 The Proposed Method

Our method is two-staged. The first stage classifies samples of u_i for inclusion in the user's gallery as usually by standard self update, and the second one selects the best p templates among them. Selection is based on unsupervised clustering by the C-means method.

Main hypothesis behind our method is that users are clusterized in overlapping subsets as follows:

$$p(x) = \sum_{i=1}^{k} p(x|x \in c_i) \cdot P(x \in c_i) \tag{1}$$

Where c_i is the cluster associated to a well-defined user u_j. We hypothesize that the samples distribution $p(x|x \in c_i)$ is mono-modal, thus only one user at a time can be associated to a certain cluster. In order to find these clusters, we use the C-means algorithm aimed to maximize the intra-cluster/inter-clusters ratio variability. Cluster c_i is associated to a user u_j if this user exhibits the maximum number of samples in c_i.

The overall algorithm is as follows:

First stage
estimate the update threshold t^* using T
for all e element in u_i **do**
  if $distance(x \in u_i, T) < t^*$ **then**
    $T_{new} = T \cup x$
  end if
end for
Second stage
generate k clusters by C-means algorithm on T_{new}, where $C = k$ being k the number of enrolled users
for each i in $[1, .., k]$ **do**
  let c_i be the $i - th$ cluster generated by C-means (with $C = k$)
  let u_j the user with the highest number of samples in c_i
  select the p nearest samples to the centroid of c_i and update the gallery of u_j accordingly
end for

Even in this case, the algorithm is repeated once a novel batch u_{i+1} is collected.

Our system is supposed to work by accepting that hypothesis about users distributions is true. In case it is false, we may have two problems: (1) more than one cluster is associated to the same user; (2) one cluster can be associated to more than one user. However, if the matcher is working at a very stringent operational point, namely, zeroFAR, it is expected the classified samples are very close each other. Therefore, our basic hypothesis may hold in this case.

4 Experimental Result

4.1 Dataset

The dataset used to simulate our method is the Multimodal-DIEE. It is made up of 49 users with 60 images per user. The period of the acquisition is about one year and half per user. The dataset presents variations of lighting, face pose, expression and occlusions (glasses). An example is in Figure 1. Each face image is rotated in order to align eyes, then cropped and normalized at the size 128x128 pixels in order to keep the same interocular distance.

4.2 Experimental Protocol

The dataset is randomly subdivided into seven parts at incremental period of time. The first part is the initial gallery T, and the last part is the test set used to measure the performance of the system at each update cycle. In each update cycle, the intermediate five parts are submitted to the system. An independent

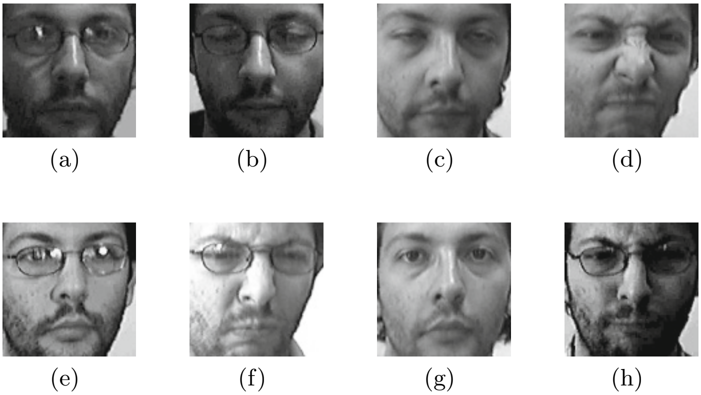


Fig. 1. Example of variation of a face in the Multimodal DIEE.

test-set allows us to appreciate the benefits of self update and proposed variant in order to have a precise reference and the same for all cycles.

Initial gallery varies from five to seven samples per user. This number also defines the maximum number of storable samples for our method, namely, the value of p, whilst for standard self update $p \rightarrow inf$.

Finally, we briefly describe the adopted facial features. After a pre-processing based on DoG [7] in order to alleviate the illumination variations problem, the resulting image is divided into non-overlapped blocks of 7x7 pixels. A set of Binarized Statistically Independent Features (BSIFs) is extracted from each block and concatenated in order to form the final feature vector [8]. This matcher is created for European MAVEN Project.

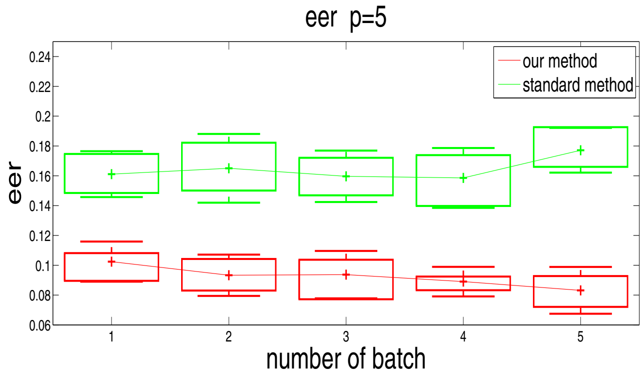
4.3 Results

In this Section, we show some experimental results about the application of standard self update and the proposed method under realistic constraints of keeping p samples in the users gallery. Updating threshold is set at the operational point such that $FAR = 0.1\%$.

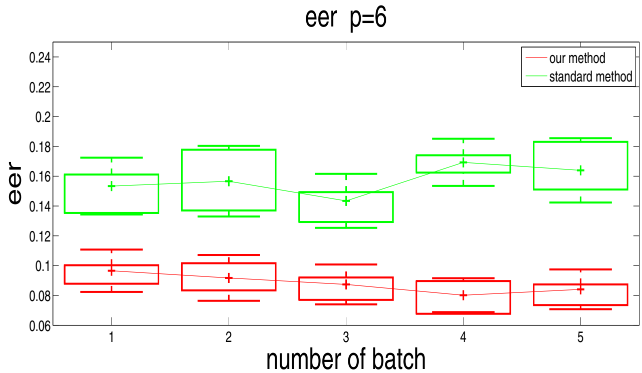
The EER plots 2(a), 2(b), 2(c) show that the standard self update start dropping the performance, with average and standard deviation per user, after some batches. This is due to the fact that the more the wrongly impostors inserted into the system gallery, the more the error rate. On the contrary, the selection stage strongly reduce the error rate (Figs. 4(a), 4(b), 4(c)).

We may see that our algorithm exhibits the same behaviour independently on the investigated values of p that we kept very low in order to meet the stringent requirement of our working hypothesis. EER is substantially reduced with respect to standard self update which uses the same number of p samples.

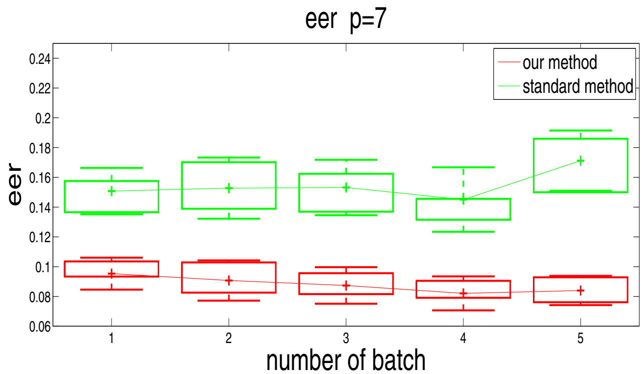
AUC values are shown in plots 3(a), 3(b), 3(c). The proposed method exhibits the best performance even in this case, thus suggesting a substantial improvement for all operational points.



(a)



(b)



(c)

Fig. 2. Equal Error Rate comparison between state of the art system and the new proposed method by varying p.

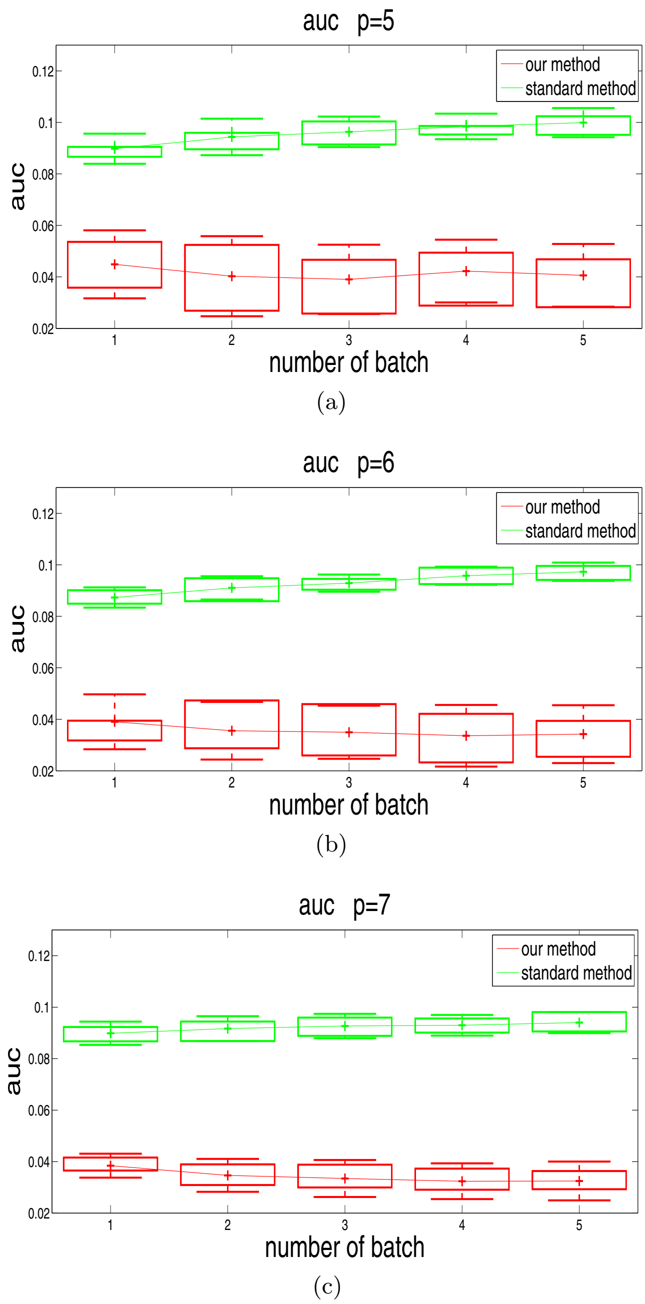
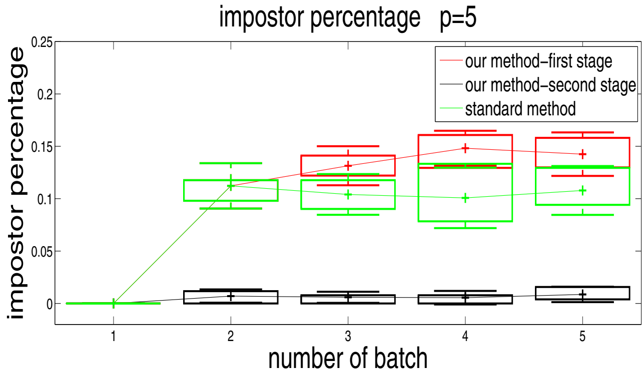
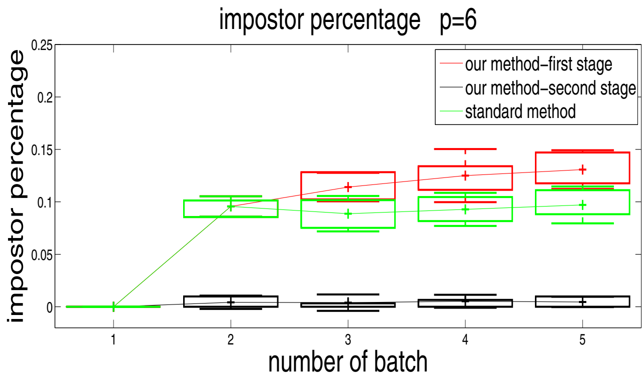


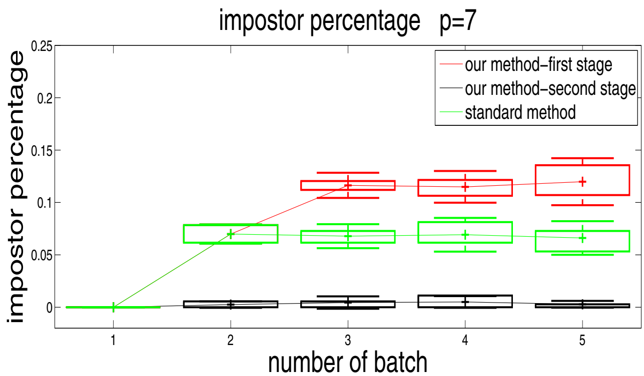
Fig. 3. AUC comparison between state of the art system and the new proposed method by varying p.



(a)

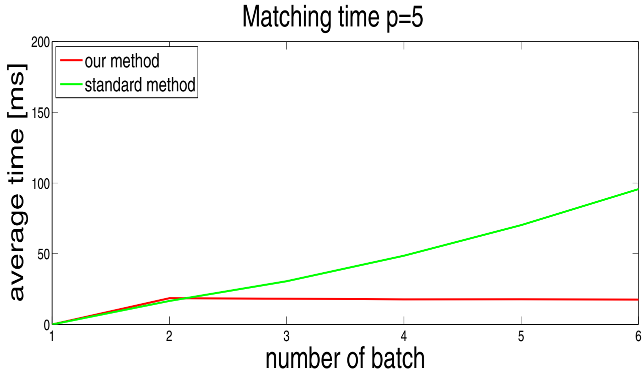


(b)

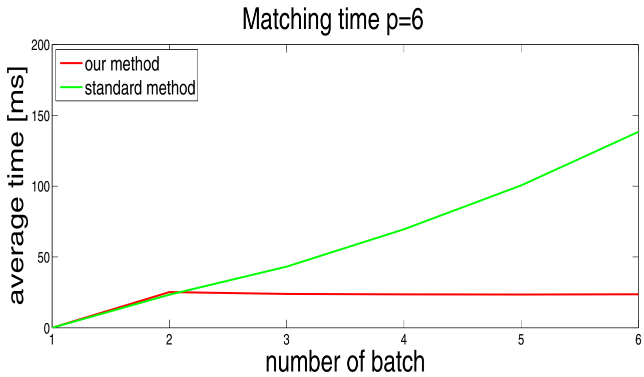


(c)

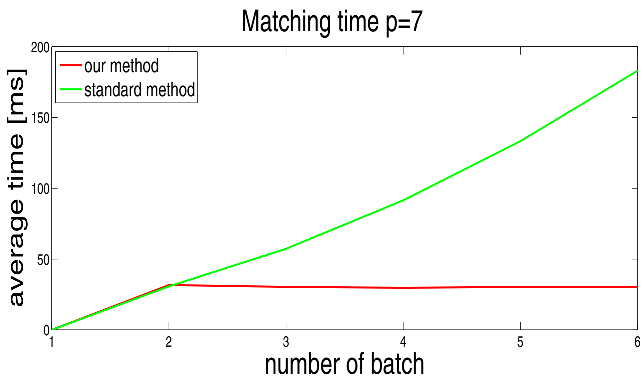
Fig. 4. Comparison between state of the art system and the new proposed method by varying p on rate of undetected impostors.



(a)



(b)



(c)

Fig. 5. Matching time comparison among the standard self-update system and the new proposed method by varying p.

The Figs. 4(a), 4(b), 4(c) show the average and standard deviation percentage of impostors in the users gallery. This percentage is always near to zero for all batches: the effectiveness of the selection approach can be appreciated by looking at the decrease of impostors number from first to second stage. This is an intrinsic confirmation about the hypothesis of users clusterization behind the rationale of our algorithm in the investigated data set.

Finally Figs. 5(a), 5(b), 5(c) show the matching time of the standard method, and the proposed one. We can appreciate that our method has a quite constant time at varying of p, because the variation of this parameter among the experiments is small. The standard self-update shows a linear matching time with number of templates. At varying of the variable p, the time increases for self-update, due to the fact that there is not a selection stage, which prunes the redundant templates. All the experiments were performed with a desktop PC with operating system Ubuntu 14.04 LTS 64bit,16 GB RAM, intel core i7-4790@3,60GHzx8, using MatLab v.R2013a.

The same arguments can be used to the memory issue, as matter of fact, our method has only p templates to every cycle. Instead in the standard self-update the number increase every time that there is a new batch of users, thus we have a linear proportionality between number of templates and occupied memory.

On the other hand, standard self update exhibit a higher number of impostors. This explains why both EER and AUC are worse than that of our method. It is important to notice that, potentially, the number of templates after first iteration is much superior than p. Despite this fact, the two-staged approach showed a more stable performance and a reduced number of impostors in the users galleries. As claimed in the introduction, this allows maintaining less complex the overall system architecture in order to be implemented even in systems with small storage and computational resources.

5 Conclusions

In this paper, we proposed an adaptive face recognition system based on two stages. First one is aimed to select a possible set of templates for each user's gallery. Second one is aimed to identify the user's cluster based on the hypothesis that the distribution of feature spaces is mono-modal for each user. Finally, the p samples nearest to each cluster's centroid are included into the clients gallery.

This approach has the ability of limiting the impostors introduction and keeping constant the number of samples in the user's gallery, by maximizing intra-class/inter-classes variations ratio.

Reported results are encouraging but experiments are limited to only one data set. Future works will be focused on extensive experiments and on a deep investigation on when and where the hypothesis behind our approach is verified.

Acknowledgement. The research leading to these results has received funding from the European Unions Seventh Framework Programme managed by REA - Research Executive Agency http://ec.europa.eu/research/rea (FP7/2007-2013) under Grant Agreement n 606058.

References

1. Yun, F., Guo, G., Huang, T.S.: Age Synthesis and Estimation via Faces: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence **32**(11), 1955–1976 (2010)
2. Uludag, U., Ross, A., Jain, A.: Biometric Template Selection: A Case Study in Fingerprints. In: Kittler, J., Nixon, M.S. (eds.) AVBPA 2003. LNCS, vol. 2688, pp. 335–342. Springer, Heidelberg (2003)
3. Jiang, X., Ser, W.: Online fingerprint template improvement. IEEE Transactions on Pattern Analysis and Machine Intelligence **24**(8), 1121–1126 (2002)
4. Freni, B., Marcialis, G.L., Roli, F.: Replacement Algorithms for Fingerprint Template Update. In: Campilho, A., Kamel, M.S. (eds.) ICIAR 2008. LNCS, vol. 5112, pp. 884–893. Springer, Heidelberg (2008)
5. Rattani, A., Marcialis, G.L., Granger, E., Roli, F.: A Dual-Staged Classification-Selection Approach for Automated Update of Biometric Templates. In: IAPR/IEEE 21th Int. Conf. on Pattern Recognition (ICPR 2012), Tsukuba, Japan, November 11–15, pp. 2972–2975 (2012) ISBN: 978-4-9906441-1-6
6. Ryu, C., Kim, H., Jain, A.K.: Template adaptation based fingerprint verification. In: 18th International Conference on Pattern Recognition, ICPR 2006, vol. 4, pp. 582–585. IEEE (2006)
7. Tan, X., Triggs, B.: Enhanced local texture sets for face recognition under difficult lighting conditions. IEEE Transactions on Image Processing **19**(6), 1635–1650 (2010)
8. Kannala, J., Rahtu, E.: BSIF: Binarized statistical image features. In: 2012 21st International Conference on Pattern Recognition (ICPR), November 11–15, pp. 1363–1366 (2012)
9. Roli, F., Didaci, L., Marcialis, G.L.: Adaptive biometric systems that can improve with use. In: Ratha, N., Govindaraju, V. (eds.) Advances in Biometrics: Sensors, Systems and Algorithms, pp. 447–471. Springer (2008) doi:10.1007/978-1-84628-921-7 ISBN 978-1-84628-920-0