

Context-Sensitive Self-Updating for Adaptive Face Recognition

C. Pagano, E. Granger, R. Sabourin , P. Tuveri, G.-L. Marcialis and F. Roli

Abstract Performance of state-of-the-art face recognition (FR) systems is known to be significantly affected by variations in facial appearance, caused mainly by changes in capture conditions and physiology. While individuals are often enrolled to a FR system using a limited number of reference face captures, adapting facial models through re-enrollment, or through self-updating with highly confident operational captures, has been shown to maintain or improve performance. However, frequent re-enrollment and updating can become very costly, and facial models may be corrupted if misclassified face captures are used for self-updating. This chapter presents an overview of adaptive FR systems that perform self-updating of facial models using operational (unlabelled) data. Adaptive template matching systems are first revised, with a particular focus on system complexity control using template management techniques. A new *context-sensitive* self-updating approach is proposed to self-update only when highly-confident operational data depict new capture conditions. This allows to enhance the modelling of intra-class variations while mitigating the growth of the system by filtering out redundant information, thus reducing the need to use costly template management techniques during operations. A particular implementation is proposed, where highly-confident templates are added according to variations in illumination conditions detected using a global luminance distortion measures. Experimental results using three publicly-available FR databases indicate that this approach enables to maintain a level of classification performance comparable to standard self-updating template matching systems, while significantly reducing the memory and computational complexity over time.

C. Pagano, E. Granger, R. Sabourin
Lab. d'imagerie, de vision et d'intelligence artificielle, École de technologie supérieure,
Université du Québec, Montreal, Canada, e-mail: cpagano@livia.etsmtl.ca, {eric.granger,
robert.sabourin}@etsmtl.ca

P. Tuveri, G.-L. Marcialis, F. Roli,
Pattern Recognition and Applications Group, Dept. of Electrical and Electronic Engineering, Uni-
versity of Cagliari, Cagliari, Italy e-mail: {pierluigi.tuveri, marcialis, roli}@diee.unica.it

1 Introduction

Automated face recognition (FR) has become an important function in a wide range of security and surveillance applications, involving computer networks, smart-phones, tablets, IP cameras, etc. Capturing faces in still images or videos allows to perform non-intrusive authentication in applications where the user's cooperation is either impossible (video-surveillance in crowded environments) or to be limited (continuous authentication). For example, in the context of controlled access to critical information on computer network systems, the face modality may allow for a continuous, non-intrusive authentication [18]. After initial log-in, a FR system may enroll the authenticated user using facial images captured from the computer's built-in camera, and design a facial model¹. The user's identity may then be periodically validated using facial images captured over time without requiring active co-operation (i.e. password prompt).

However, limited user co-operation as well as uncontrolled observation environments often make FR a challenging task. It is well known that the performance of state-of-the-art FR systems may be severely affected by changes in capture conditions (e.g., variation in illumination, pose and scale), as well as individual physiology [20, 31]. Moreover, such systems are usually initialized with a limited number of high-quality reference face captures, which may generate non-representative facial models (not modelling all possible variations) [19].

To account for such intra-class variations, several solutions have been investigated in the literature over the past decade. They can be organized into the following two categories:

1. Development of discriminative features that are robust to environmental changes [4, 36]. These techniques usually aim to develop facial descriptors insensitive to changes in capture conditions, to mitigate their effects on the recognition process.
2. Storage (or synthetic generation) of multiple reference images to cover the different capture conditions that could be encountered during operations [11, 13].

However, these approaches assume that FR is a stationary process, as they only rely on information available during enrolment sessions. In addition, depending on the application, a single enrolment session is often considered as multiple ones are not always possible [22]. This prevents to integrate new concepts² that may emerge during operations as capture conditions and individuals physiology evolve over time (for example due to natural lighting conditions and ageing).

To address this limitation, adaptive biometric systems have been proposed in the literature [28], inspired by semi-supervised learning techniques for pattern recognition [17]. These systems are able to adapt facial models (sets of templates or classifier parameters) by exploiting (either on-line or off-line) faces captured during sys-

¹ Depending on the classification system, a facial model may be defined as either a set of one or more reference face captures (template matching), or a statistical model estimated from reference captures (statistical classification).

² A *concept* can be defined as the underlying data distribution of the problem under specific operating conditions [1].

tem operations. Common approaches in adaptive biometrics fall under *self-updating* and *co-updating*, depending on whether they rely on a single or multiple modalities. They usually either: 1) add novel captures to individual specific galleries [29], or 2), fuse new input data into common templates referred to as *super-templates*, containing all information [12, 30] for each modality (for example, virtual facial captures constructed with patches from operational data).

This chapter focuses on *self-updating* techniques with template matching systems for FR. These methods update template galleries using faces captured during operations that are considered highly-confident, i.e. that produce very high matching scores (surpassing a self-updating threshold) [23]. Advantages and drawbacks of self-updating have been widely investigated [15, 23]. While these methods have been show to significantly improve the performance of biometric systems over time, an updating strategy only relying on matching score values may add redundant template to the galleries. This can significantly increase system complexity over time with information that do not necessarily improve performance, and also eventually reduce its response time during operations. To bound this complexity, template management methods (e.g. pruning) have been proposed in literature [8, 15, 23]. While clustering-based methods showed the most promising results, they remain computationally complex and thus not suited for seamless operations, if self-updating is performed frequently.

In this chapter, a survey of state-of-the-art techniques for adaptive FR using self-updating is presented, along with the key challenges facing these systems. An experimental protocol involving three real-life facial datasets (DIEE [26], FIA [9] and FRGC [21]) is proposed to evaluate the benefits and drawbacks of a self-updating methodology applied to a template matching system, with a particular focus on the management of system complexity. To address this challenge, a *context-sensitive* self-updating technique is proposed for template matching systems, combining a standard self-updating procedure and a change detection module. With this technique, only operational faces that were captured under different capture conditions are added to an individual's template gallery. More precisely, the addition of a new capture into the galleries depends on two conditions: 1) it's matching score is above the self-updating threshold (highly confident capture), and 2), the capture contains new information w.r.t. the samples already present in the gallery (i.e. captured under different conditions). This strategy allows to benefit from contextual information available in operational captures to limit the growth in system complexity. With this technique, one can avoid frequent uses of costly template management schemes, while still enhancing intra-class variation in facial models with relevant templates. A particular implementation of this proposed technique is considered for a basic template matching system, where changes are detected in illumination conditions.

The rest of this chapter is organized as follows. Section 2 provides a general survey of self-updating algorithms in the context of adaptive biometric systems. Then, section 3 introduces the new *context-sensitive* self-updating technique based on the detection of changes in capture conditions, and Section 4 presents the proposed experimental methodology. Finally, experimental results are presented and discussed in Section 5.

2 Self-Updating for Face Recognition

2.1 A General Face Recognition System

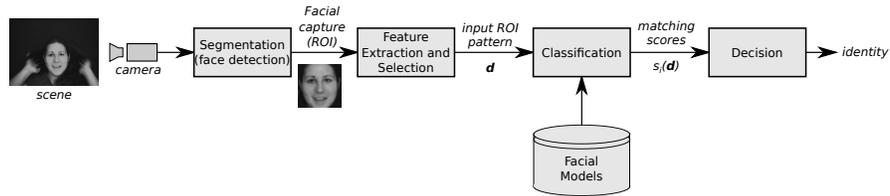


Fig. 1 General FR system trained for N individuals.

Figure 1 presents a generic system for the recognition of faces in images (stills or video frames) captured from a camera. It is composed by four modules: segmentation, feature extraction, classification and decision. In addition, facial models of the N enrolled individuals are stored into the system, to be used by the classification module to produce matching scores for each individual.

During operations, faces are isolated in the image using the segmentation module, which produces the regions of interest (ROIs). Then, discriminant features are extracted from each ROI (e.g. eigenfaces [32] or local binary patterns [2]) to produce the corresponding pattern $\mathbf{d} = (d[1], \dots, d[F])$ (with F the dimensionality of the feature space). This pattern is then compared to the facial model of each enrolled individual i by the classifier, which produces the corresponding matching scores $s_i(\mathbf{d})$, ($i = 1, \dots, N$).

The facial models are usually designed a priori using one or several reference patterns, from which the same features have been extracted, and their nature depends on the type of classifier used in the system. For example, with a template matcher, a facial model of an individual i can be a gallery of one or several reference patterns $\mathbf{r}_{i,j}$ ($j = 1, \dots, J$), in which case matching scores for each operational pattern \mathbf{d} would be computed from distance measures to these patterns. Classification may also be performed using neural networks (e.g. multi-layer perceptrons [27] and ARTMAP neural networks [3]) or statistical classifiers (e.g. naive Bayes classification [7]), in which case the facial models would consist of parameters estimated during their training using the reference patterns $\mathbf{r}_{i,j}$ (e.g. neural networks weights, statistical distribution parameters, etc.).

Finally, the decision module produces a final response according to the application. For example, an identification system for surveillance may predict the identity of the observed individual with a maximum rule, selecting the enrolled individual with the highest matching score, while a verification system for access control usually confirms the claimed identity by comparing the corresponding matching score to a decision threshold.

2.2 Adaptive Biometrics

As mentioned earlier, the performance of FR systems can be severely affected by changes in capture conditions. Intra-class variations can be observed in the input data, as a consequence of changes in capture conditions (scene illumination, facial pose angle w.r.t. the camera, etc.) or individuals physiology (facial hair, ageing, etc.). Such diversity is difficult to represent using the limited amount of reference captures used for initial facial model design. To address this limitation, adaptive biometric systems have been proposed in the literature, providing the option for continuous adaptation of the facial models using the operational data [22, 23].

Adaptation can be either supervised or unsupervised, depending on the labelling process of the operational data. In semi-supervised learning [37], the facial model of each individual enrolled to the system is updated using operational data labelled as the same individual by the classification system. For example, a gallery \mathcal{G}_i of reference patterns may be augmented with highly-confident operational input patterns \mathbf{d} matched to the facial model of individual i . While this enables to refine facial models, the performance of such systems is strongly dependent on their initial classification performance. In addition, the integration of mislabelled captures could corrupt facial models, thus affecting the accuracy of the system for the corresponding individuals [23, 26]

An adaptive biometric system can also perform supervised adaptation, where the operating samples used to update the system are manually labelled, or obtained through some re-enrolment process [23]. While supervised adaptation may represent an ideal scenario with an error-free labelling process, human intervention is often costly or not feasible. Depending on the application, the ability to perform semi-supervised adaptation may be the only viable solution, which has led to the development of various strategies to increase the robustness of such systems.

These techniques can be categorized as *self-update* [12, 30] and *co-update* techniques [23, 24], depending on whether a single or multiple modalities are considered for the update of facial models with highly-confident patterns. This chapter focuses on *self-updating* methods for FR, where facial models are defined by galleries of reference patterns.

2.3 Self-Updating Methods

In the context of FR systems, self-updating methods update the facial models using only highly-confident operational captures, i.e. with matching scores surpassing a very high threshold, to prevent possible corruptions due to misclassification.

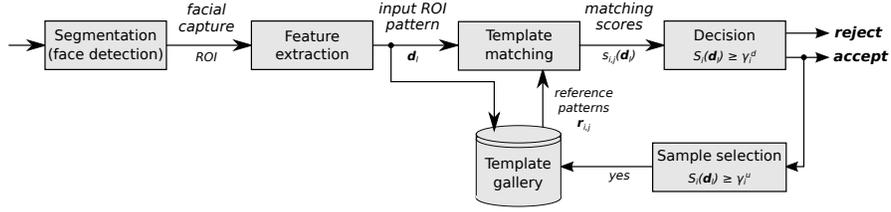


Fig. 2 A FR system based on template matching that allows for self-update.

2.3.1 General Presentation

To illustrate this principle, it is applied to a template matching system, presented in Figure 2. In this system, inspired by [25], the facial model of each individual i is designed by storing initial reference patterns from a labelled dataset into a gallery $\mathcal{G}_i = \{\mathbf{r}_{i,1}, \mathbf{r}_{i,2}, \dots\}$ (in this case, the terms *pattern* and *template* are used indiscriminately). To simplify the notation, the remaining of this section will omit the subscript i and only consider one individual, as this methodology can be extended to many with individual specific galleries and thresholds.

Algorithm 1 Self-update algorithm for adapting template galleries.

Inputs: - $\mathcal{G} = \{\mathbf{r}_1, \dots, \mathbf{r}_J\}$ // Gallery with initial templates
 - $\mathcal{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_L\}$ // Unlabelled adaptation set

Outputs: - $\mathcal{G}' = \{\mathbf{r}_1, \dots, \mathbf{r}_{J'}\}$, $J' \geq J$ // Updated Gallery

```

1: Estimate updating threshold  $\gamma^u \geq \gamma^d$  from  $\mathcal{G}$ 
2:  $\mathcal{G} \leftarrow \mathcal{G}'$  // Initialization with previous state
3: for  $l = 1, \dots, L$  do // For all samples  $\mathbf{d}_l \in \mathcal{D}$ 
4:   for  $j = 1, \dots, J$  do // For all references  $\mathbf{r}_j \in \mathcal{G}$ 
5:      $s_j(\mathbf{d}_l) \leftarrow \text{similarity\_measure}(\mathbf{d}_l, \mathbf{r}_j)$  // Prediction score for each reference
6:   end for
7:    $S(\mathbf{d}_l) \leftarrow \max_{j \in \{1, \dots, J\}} \{s_j(\mathbf{d}_l)\}$  // Maximum fusion of scores
8:   if  $S(\mathbf{d}_l) \geq \gamma_d$  then
9:     Output positive prediction
10:    if  $S(\mathbf{d}_l) \geq \gamma^u$  then
11:       $\mathcal{G}' \leftarrow \mathcal{G}' \cup \mathbf{d}_l$  // Add the sample which similarity surpasses  $\gamma^u$  to the gallery
12:    end if
13:  end if
14: end for

```

Alg. 1 presents a generic algorithm for self-updating a template gallery \mathcal{G} with several reference patterns \mathbf{r}_j ($j = 1, \dots, J$). During operations, the system is presented with an unlabelled data set \mathcal{D} of L facial captures. For each sample \mathbf{d}_l , similarity measures to each reference \mathbf{r}_j in the gallery are used to compute the set of matching scores $s_j(\mathbf{d}_l)$ ($j = 1, \dots, J$). Then, the final score $S(\mathbf{d}_l)$ is computed as a

combination of $s_j(\mathbf{d}_l)$ (e.g. the maximum fusion rule), and positive prediction is output if it surpasses the decision threshold γ^d . Finally, the sample selection module relies on a stricter updating threshold γ^u (usually $\gamma^u \geq \gamma^d$), updating the gallery \mathcal{G} with \mathbf{d}_l if $S(\mathbf{d}_l) \geq \gamma^u$, i.e. if the prediction has a high degree of confidence.

2.3.2 Challenges

While self-updating methods have been shown to improve system accuracy over time, the adaptation of the facial models using operational data might be detrimental, and the selection of the updating threshold is critical [23]. To prevent a decline in classification performance, the use of a strict updating threshold may enable to reduce the probability of updating the facial models with misclassified patterns [14, 29, 30]. However, it has been argued that updating with only highly confident patterns may result in the addition of redundant information in the galleries, and thus a marginal gain in performance at the expense of a considerable increase in system complexity [23].

In addition, operational samples with more drastic changes are less likely to generate classification scores surpassing the updating threshold, preventing the classification system to assimilate this new information. To address this limitation, *co-updating* methods have been proposed to benefit from complementary biometric systems [23, 24]. Each system is initialized with reference templates from a different source (or different features extracted from the same source), and performs classification of operational input data. In the same way as self-updating techniques, each system selects highly-confident samples based on an updating threshold, but this information is also shared with other systems. If the classification score of one system surpasses its updating threshold, the others will also consider the corresponding samples as highly confident, and perform adaptation. This enables to increase to probability of updating with different but genuine operational data, by relying on the supposition that a drastic change on one source is not necessarily observed on others. A recent model has been proposed to estimate optimal amounts of samples and iterations to improve system's performance under specific updating constraints [5]. This model has shown to be effective under the stringent hypothesis of 0% false alarm rate for the updating threshold of both systems. While *co-updating* is usually applied with multiple biometric traits, it could also be applied in, for example, a FR scenario involving multiple cameras. In this situation, relying on multiple point of views could mitigate the effect of disruptions such as motion blur that would be less likely to affect every camera at the same time.

Finally, system complexity is a critical issue for template matching systems in live FR. The ability to operate seamlessly depends on the computational complexity of the recognition operation, which is usually directly related to gallery sizes. Several template management strategies have been proposed to limit complexity in self-updating systems. In [8], template replacement strategies have been experimented to perform self-update in a constrained environment, where the maximum number of templates in a gallery is fixed by the user. When the maximum size is reached,

several criteria have been experimented to determine which obsolete template can be replaced, such as FIFO, LFU and clustering algorithms. Among them, the clustering algorithm MDIST showed the most promising results, reducing the number of impostors samples by maintaining a gallery with very close samples. While these methods enable to compromise between system performance and complexity, they remain computationally costly, and may interfere with seamless long-term operations. Once the maximum gallery size is reached, such process would have to be performed for each new highly-confident template, thus increasing system response time. To reduce these occurrences, operational data containing redundant information should be filtered out during operations. This would limit the self-updating process to only operational templates with relevant information, i.e. templates improving intra-class variability in facial models.

3 Self-Updating Driven by Capture Conditions

This chapter introduces a new self-updating method that efficiently self-updates facial models based on capture conditions. This methodology is illustrated using a template matching system performing self-updating, as presented in [25]. As discussed in the previous sections, such methodology can significantly improve the overall classification performance through a better modelling of intra-class variations, specifically in applications exhibiting significant variations in capture conditions (e.g. continuous authentication using webcams). However, updating the galleries with only highly-confident inputs may not always provide new and beneficial information, as those samples are usually well-classified by the system, which could lead to an unnecessary increase in system complexity (e.g. the number of reference patterns stored in the galleries) [23]. While this complexity can be mitigated with template-management techniques [8], frequent gallery filtering may interfere with seamless operations over time.

To address this limitation, this section proposes a *context-sensitive* self-updating technique that integrates a template filtering process during operations. It is designed to ensure that only highly-confident data captured under novel conditions are added to template galleries, thus limiting the growth in memory complexity with redundant samples. In fact, in FR, intra-class variations in facial appearance are often related to changes in capture conditions (e.g. environmental illumination, facial pose, etc.) [20, 31], and such information can be detected during operations. Following this intuition, when a highly confident ROI pattern surpasses the updating threshold, non-discriminative information related to capture conditions are extracted to evaluate whether it has been captured under different conditions than of the reference templates already stored in the gallery. If not, the pattern is discarded, and the gallery is not augmented.

3.1 Framework For Context-Sensitive Self-Update

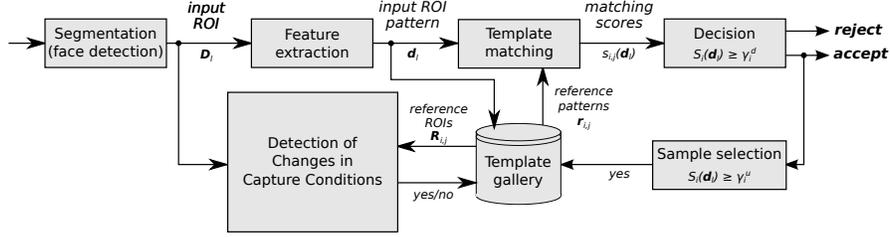


Fig. 3 A template matching system that integrates context-sensitive self-updating.

The diagram of a general template-matching system that employs the new *context-sensitive* technique is presented in Figure 3. It augments the system presented in Figure 2 with an additional decision module to detect changes in capture conditions.

In the same way than standard self-updating systems, when presented with a unlabelled data set $\mathcal{D} = \{\mathbf{d}_1, \dots, \mathbf{d}_L\}$, this system first selects highly-confident samples to perform adaptation of the template gallery \mathcal{G}_i , i.e. the set $\mathcal{D}' = \{\mathbf{d}_{i'} | S_i(\mathbf{d}_{i'}) \geq \gamma_i^d\}$. Then, an additional test is performed on these samples, only to select a final subset captured under novel capture conditions. To extract additional non-discriminative information, the individual galleries are augmented with the input ROIs $\mathbf{R}_{i,j}$ from which the reference patterns $\mathbf{r}_{i,j}$ are extracted. The augmented galleries are stored as $\mathcal{G}_i = \{\{\mathbf{R}_{i,1}, \mathbf{r}_{i,1}\}, \{\mathbf{R}_{i,2}, \mathbf{r}_{i,2}\}, \dots\}$. This additional measurement enables to maximize the intra-class variation of the galleries while mitigating their growth by rejecting redundant information. For example, contextual information such as environmental illumination or facial pose w.r.t. the camera can be measured on ROIs, to be compared with ROIs in the galleries.

3.2 A Specific Implementation

As a basic example of the framework presented in Figure 3, a particular implementation is proposed. It relies on the detection of changes in illumination conditions.

3.2.1 A Template Matching System

For classification, a standard template matching system is considered. For each individual i , a dedicated facial model is stored as a template gallery $\mathcal{G}_i = \{\{\mathbf{R}_{i,1}, \mathbf{r}_{i,1}\}, \{\mathbf{R}_{i,2}, \mathbf{r}_{i,2}\}, \dots, \{\mathbf{R}_{i,J_i}, \mathbf{r}_{i,J_i}\}\}$, as well as user-specific decision γ_i^d and updating γ_i^u thresholds.

For each input ROI isolated through segmentation, the corresponding pattern \mathbf{d}_l is extracted using a Multi-Bloc Local Binary Pattern (LBP) [2] algorithm. Features for block sizes of 3x3, 5x5 and 9x9 pixels are computed and concatenated with the grayscale pixel intensity values, and PCA is used to reduce the dimensionality to $F = 32^3$. The matching score for each individual i is then computed following:

$$S_i(\mathbf{d}_l) = \frac{1}{J_i} \cdot \sum_{j=1}^{J_i} s_{i,j}(\mathbf{d}_l) = \frac{1}{J_i} \sum_{j=1}^{J_i} \frac{[\sqrt{F} - d_{Eucl}(\mathbf{d}_l, \mathbf{r}_{i,j})]}{\sqrt{F}} \quad (1)$$

where $d_{Eucl}(\mathbf{d}_l, \mathbf{r}_{i,j})$ is the Euclidean distance between input pattern \mathbf{d}_l and template $\mathbf{r}_{i,j}$ (with $j = 1, \dots, J_i$) and J_i the total number of templates in \mathcal{G}_i . The matching scores $s_{i,j}(\mathbf{d}_l)$ are here computed as the normalized opposite to the distance $d_{Eucl}(\mathbf{d}_l, \mathbf{r}_{i,j})$ (a score of 1 is achieved for a null distance). The final matching score $S_i(\mathbf{d}_l)$ is obtained from the combination of these scores using the average fusion rule.

Finally, the system outputs a positive prediction for individual i if $S_i(\mathbf{d}_l) \geq \gamma_i^d$, and selects \mathbf{d}_l as a highly confident face capture for individual i if $S_i(\mathbf{d}_l) \geq \gamma_i^h$.

3.2.2 Detecting Changes in Capture Conditions

In Figure 3, for each individual i , the input ROIs \mathbf{D}_l corresponding to highly-confident operational captures are compared to the reference ROIs $\mathbf{R}_{i,j}$ ($j = 1, \dots, J_i$) stored in the galleries, and assess whether the capture conditions are novel enough to justify an increase in complexity. The universal image quality index Q [34] is considered to measure the distortion between \mathbf{D}_l and each reference ROI $\mathbf{R}_{i,j}$. This measure is a particular case of the Structural Similarity Index Measure (SSIM) presented in [35]. It can be written as a product of the three factors – loss of correlation, luminance distortion and contrast distortion:

$$Q(\mathbf{R}_{i,j}, \mathbf{D}_l) = \frac{\sigma_{\mathbf{R}_{i,j}, \mathbf{D}_l}}{\sigma_{\mathbf{R}_{i,j}} \cdot \sigma_{\mathbf{D}_l}} \cdot \frac{2\bar{\mathbf{R}}_{i,j} \cdot \bar{\mathbf{D}}_l}{\bar{\mathbf{R}}_{i,j}^2 + \bar{\mathbf{D}}_l^2} \cdot \frac{2\sigma_{\mathbf{R}_{i,j}} \cdot \sigma_{\mathbf{D}_l}}{\sigma_{\mathbf{R}_{i,j}}^2 + \sigma_{\mathbf{D}_l}^2} \quad (2)$$

where $\bar{\mathbf{R}}_{i,j}$ and $\bar{\mathbf{D}}_l$ are the average images, $\sigma_{\mathbf{R}_{i,j}}$ and $\sigma_{\mathbf{D}_l}$ their variances, and $\sigma_{\mathbf{R}_{i,j}, \mathbf{D}_l}$ the covariance.

To accommodate spatial variations in image distortion, statistical features for Eq. 2 may be measured locally. A local quality index $Q(\mathbf{R}_{i,j}[k], \mathbf{D}_l[k])$ is thereby calculated, where $\mathbf{D}_l[k]$ ($\mathbf{R}_{i,j}[k]$) corresponds to window of \mathbf{D}_l ($\mathbf{R}_{i,j}$) sliding from the top-left corner to the bottom right corner for a total of K steps. These local measurements can then be combined into the global quality index GQ following:

$$GQ(\mathbf{R}_{i,j}, \mathbf{D}_l) = \frac{1}{K} \sum_{k=1}^K Q(\mathbf{R}_{i,j}[k], \mathbf{D}_l[k]) \quad (3)$$

³ This value has been determined experimentally as an optimal trade-off between accuracy and computational complexity using a nearest-neighbour classifier with Euclidean distance.

In this chapter, the proposed template filtering strategy is implemented through a detection of changes in ROI illumination conditions only. For that intent, the second term of the quality index Q (see Eq. 2) is considered, to compute the global luminance quality (GLQ) following:

$$GLQ(\mathbf{D}_l, \mathbf{R}_{i,j}) = \frac{1}{K} \sum_{k=1}^K LQ(\mathbf{R}_{i,j}[k], \mathbf{D}_l[k]) = \frac{1}{K} \sum_{k=1}^K \frac{2 \cdot \bar{\mathbf{D}}_l[k] \cdot \bar{\mathbf{R}}_{i,j}[k]}{\bar{\mathbf{D}}_l[k]^2 + \bar{\mathbf{R}}_{i,j}[k]^2} \quad (4)$$

where the local luminance quality measurements LQ measure the proximity of the average luminance between each window. Highly confident captures \mathbf{D}_l are then used to update the gallery \mathcal{G}_i if and only if

$$\frac{1}{J_i} \sum_{j=1}^{J_i} GLQ(\mathbf{D}_l, \mathbf{R}_{i,j}) \geq \gamma_i^c \quad (5)$$

with γ_i^c the capture condition threshold, computed as the average GLQ between all the references captures in \mathcal{G}_i .

4 Simulation Methodology

This section presents several experimental scenarios involving three real-world FR databases. The proposed simulations emulate realistic FR applications of different orders of complexity, with variations in capture conditions. The objective is to observe and compare the performance of new and reference self-updating techniques under different operation conditions, and within a basic template matching system described in Section 3.2.

4.1 Face Recognition Databases

Three publicly-available FR databases are considered for simulation. To standardize the experimental protocol, each database is separated into 6 different batches for all individuals. These scenarios are summarized at the end of Section 4.1, in Table 1.

4.1.1 Multi-Modal Dipartimento di Ingegneria Elettrica ed Elettronica

The multi-modal Dipartimento di Ingegneria Elettrica ed Elettronica⁴ (DIEE) dataset [26] regroups face and fingerprint captures of 49 individuals. In this study, only facial captures are considered. For each individual, 60 facial captures have been ac-

⁴ Department of Electrical and Electronic Engineering.

quired over 6 sessions at least three weeks apart, with 10 captures per session. The collection process spanned over a period of 1.5 years.

For simulations, the facial captures of each individuals are separated into 6 batches corresponding to the capture sessions. ROIs have been extracted with a semi-manual process [16]: an operator first selected the eyes in each frame, and the cropped region was then determined as the square of size $2d * 2d$ (d being the distance between the eyes), with the eyes located at the position $(d/2, d/4)$ and $(3 \cdot d/2, d/4)$. In this process, faces have been rotated to align the eyes to minimize intra-class variations [10], and then normalized to a size of 70x70 pixels.

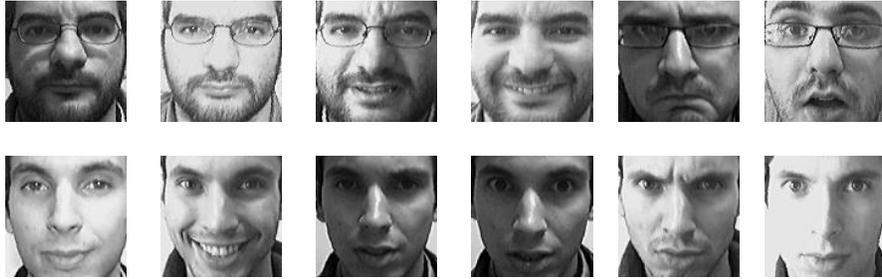


Fig. 4 DICE dataset. An example of randomly chosen facial captures for two individuals.

This dataset was explicitly collected to evaluate the performance of self-update and co-training algorithms. Over the 6 sessions, gradual changes can be observed in facial pose, orientation, and illumination (see examples in Fig. 4). While these changes generate visible differences in facial captures, the position of the individuals and their distance to the camera are controlled. For this reason, this dataset represents the easiest problem in this study, simulating an application of continuous authentication of individuals over a computer network.

4.1.2 CMU Faces in Action

The Carnegie Mellon University Faces In Action (FIA) dataset [9] contains a set of 20-second videos for 221 participants, mimicking a passport checking scenario in both indoor and outdoor environments. Videos have been captured in three separate sessions of 20 seconds at least one month apart, with 6 Dragonfly Sony ICX424 cameras (640x480 pixel resolution, 30 images per second). Cameras were positioned at 0.83m of the subjects, mounted on carts at three different horizontal angles (0° and $\pm 72.6^\circ$), with two focal lengths (4 and 8mm) each.

In this chapter, only ROIs captured during the indoor sessions, and using the frontal camera with 8mm focal length are considered. ROIs have been extracted using the OpenCV implementation of Viola-Jones face and eye detection algorithm [33]. In the same way than with DICE, faces have been rotated to align the eyes



Fig. 5 FIA dataset. An example of randomly chosen facial captures for two individuals.

[10], and normalized to a size of 70x70 pixels. For simulations, sequences from each session have been divided into two sub-sequences, in order to organize the facial captures into 6 batches.

This dataset simulates an open-set surveillance scenario as found in face re-identification applications. A restrained subset of 10 individuals of interest are monitored, but in an environment where a majority of ROIs are capture from non-target individuals. The 10 individuals of interest enrolled to the systems have been chosen with two experimental constraints: 1) the individuals must be present in all capture sessions, and 2), at least 30 ROIs per session have been extracted by the face detection algorithm.

Faces in this data set have been captured in semi-controlled capture conditions, where the individuals entered the scene and walked to stop at the same distance from the cameras, and talked while moving their head with natural movements until the end of the session. In addition to variations in illumination and facial expressions, ROIs also incorporate variations in pose, resolution (scaling), motion blur and ageing.

4.1.3 Face Recognition Grand Challenge

The Face Recognition Grand Challenge (FRGC) dataset as been collected at University Notre Dame [21]. In this chapter, the still face images of this dataset are considered. They were captured over an average of 16 sessions for 222 individuals for the training subset, and up to 22 sessions for the validation one, using a 4 Megapixels Canon camera. Each session contains four controlled and two uncontrolled captures, with significantly different illumination and expression.

Overall, 187 individuals have been selected for experiments, for which more than 100 ROIs are available (around 133 in average). In the same way than with the other datasets, 6 batches of the the sane size have been created for each individual, respecting the temporal relation between the capture sessions. ROIs have been extracted in the same way than with the DICE dataset [16], using the position of the eyes already available in the FRGC dataset.

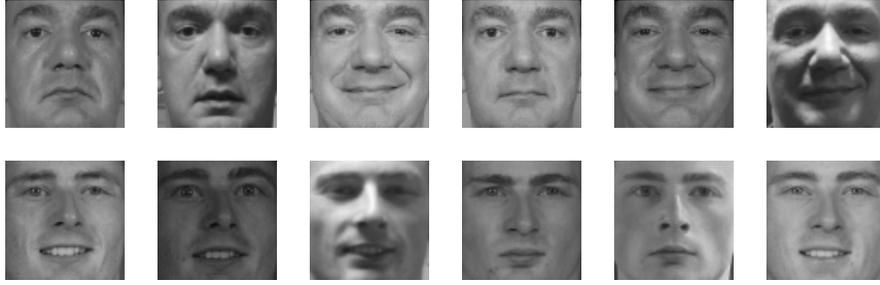


Fig. 6 FRGC dataset. An example of randomly chosen facial captures for two individuals

This dataset simulates a wide-range identification application, with multiple enrolment sessions where a very limited amount of reference templates are captured. Recurring and unpredictable changes in illumination and facial expression emerge in the operational environment in every capture session.

Dataset	Scenario	# enrolled individuals	# enrolment sessions	# ROIs per batch	Sources of variation
DIEE	continuous authentication	49	6	10	illumination, expression
FIA	video-surveillance	10	3	69	illumination, expression, pose, resolution, ageing, scaling, blur
FRGC	wide-range identification	187	16	22	illumination, expression, ageing

Table 1 Summary of the three experimental scenarios.

4.2 Protocol

The following three template matching systems are experimentally compared in this chapter:

1. **baseline system**, performing template matching in the same way as in Figure 3, but without any adaptation of the template galleries \mathcal{G}_i . User-specific decision thresholds γ_i^d are stored for decision.
2. standard **self-updating** system, updating the template galleries \mathcal{G}_i with highly confident ROI patterns, which scores surpass user-specific updating thresholds γ_i^u , and decision thresholds γ_i^d .

3. proposed **context-sensitive self-updating** system, only updating the template galleries \mathcal{G}_i with highly confident samples that also passed the concept change test (Eq. 5), using user-specific updating γ_i^u , capture condition γ_i^c and decision thresholds γ_i^d .

4.2.1 Simulation Scenario

The scenario described below is considered for each database. At each time step $t = 1, \dots, 6$, and for each individual $i = 1, \dots, N$, the performance of the baseline and the two self-updating systems updated with batch $b_i[t-1]$ is evaluated on batch $b_i[t]$. The self-updating systems are updated, and then tested with batch $b_i[t+1]$, and so on. A pseudo-code of the simulation process is presented in Alg. 2.

Algorithm 2 Protocol for simulations.

```

1: for  $i = 1, \dots, N$  do                                // Initialization of the galleries  $\mathcal{G}_i$  for each individual
2:    $\mathcal{G}_i \leftarrow$  first 2 patterns of  $b_i[1]$ 
3: end for
4: for  $i = 1, \dots, N$  do                                // Initialization of thresholds for each individual
5:   Evaluate update and decision thresholds  $\gamma_i^u$  and  $\gamma_i^d$  using negative distribution estimation
6:   Initialize change detection threshold  $\gamma_i^c$  as the average GLQ measure between each ROI in  $\mathcal{G}_i$ 
7: end for
8: for  $i = 1, \dots, N$  do                                // Processing of remaining samples from  $b_i[1]$ 
9:   Estimate genuine scores using remaining samples from  $b_i[1]$ 
10:  Estimate impostor samples using a random selection of impostor samples
11:  Update gallery
12:  Update thresholds
13: end for
14: for  $t = 2, \dots, 6$  do                                // Remaining data blocks
15:   for  $i = 1, \dots, N$  do                                // Each individual
16:    Estimate genuine scores using remaining samples from  $b_i[t]$ 
17:    Estimate impostor samples using a random selection of impostor samples
18:    Update gallery
19:    Update thresholds
20:   end for
21: end for

```

For each system, the individual galleries \mathcal{G}_i are initialized with the two first samples of the corresponding initial batches $b_i[1]$. For *context-sensitive* self-updating, corresponding ROIs are also stored to compute GLQ measures during operations (see Eq. 4). Then, the initial values of the decision thresholds γ_i^d are computed using negative distribution estimation: each gallery \mathcal{G}_i is compared to every other gallery to generate negative scores, and a threshold γ_i^d is chosen as the highest possible value respecting an operational false alarm constraint. For the self-updating variants, the updating threshold γ_i^u is initialized in the same way, and for the *context-sensitive*

self-updating system, γ_i^c is computed as the average GLQ measure between each ROI in \mathcal{G}_i .

Then, for each system, performance is evaluated using the remaining patterns from $b_i[t]$ to compute genuine scores, and a random selection of impostor patterns for the impostor scores. For the DIEE and FRGC datasets, impostor patterns for each individual are randomly selected among batches from other individuals. In the case of the FIA dataset, impostor patterns are selected from the non-target dataset individuals during the same session. To avoid any bias in performance evaluation, the same amount of impostor and genuine patterns are considered.

Finally, using genuine and impostor patterns, the self-updating systems galleries are updated according to their updating strategies, and the thresholds are re-estimated using the same methodology. This scenario is then reproduced for the remaining 5 batches.

4.2.2 Performance Measures

For each system, performance is measured with average true positive rate (tpr) and false positive rate (fpr) for each individual. These are respectively the proportion of genuine patterns correctly classified over the total number of genuine patterns (tpr), and the proportion of impostor patterns classified as genuine over the total number of negative patterns (fpr). These measures depend on the decision thresholds γ_i^d , computed during update to respect a given fpr constraint.

System complexity is also presented, as the average number of templates in the galleries. In addition, facial model corruption due to the addition of misclassified templates in the galleries is presented as the ratio of impostor over genuine templates. Following Doddington's classification [6], only the 10 galleries with the highest ratio are presented, to focus on lamb-type individuals which are easy to imitate.

Finally, a constraint of $fpr = 5\%$ has been chosen to compute the decision thresholds γ_i^d . In addition, for each scenario, the updating thresholds γ_i^u correspond to an ideal $fpr = 0\%$ and a laxer $fpr = 1\%$. For each performance measure, results are presented as the average and standard deviation values for every enrolled individual, computed using a Student distribution and a confidence interval of 10%.

5 Simulation Results

5.1 Continuous User Authentication with DIEE Data

Figure 7 presents the average performance results of the baseline, self-updating and *context-sensitive* self-updating techniques within the template matching system described in Section 3.2. Results are presented for the ideal $fpr = 0\%$ updating thresholds for the self-updating techniques.

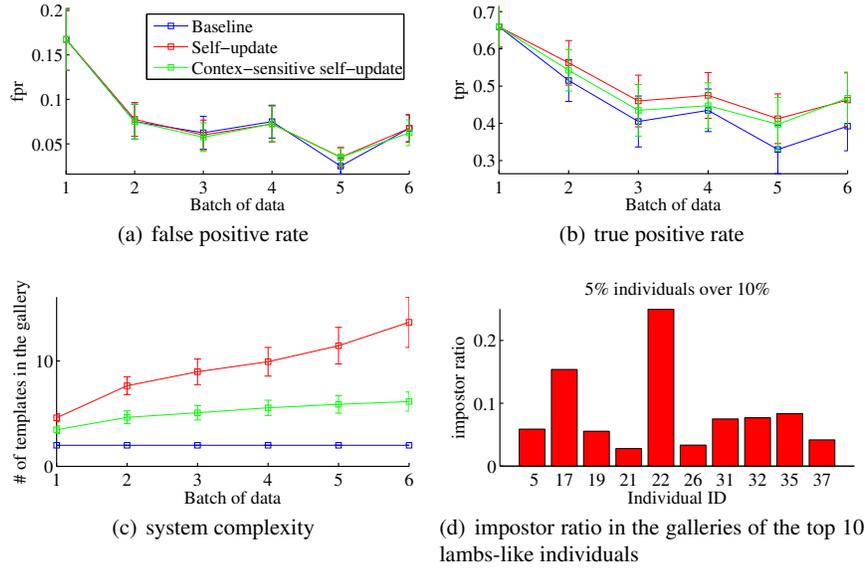


Fig. 7 Simulation results with DLEE dataset where the updating threshold is selected for $fpr = 0\%$.

While all 3 systems present similar fpr (between 7 – 17%) in Fig. 7 (a), a significant differentiation can be observed in the tpr with batches 5 and 6 (Fig. 7 (b)). In fact, the introduction of batch 5 generates a decline in tpr performance for the baseline system (from $43.5 \pm 5.7\%$ down to $33.0 \pm 6.5\%$), that ends at $tpr = 39.3 \pm 6.6\%$ at batch 6. On the other hand, the self-updating and *context-sensitive* self-updating systems exhibit a moderate decline (respectively from $47.5 \pm 6.1\%$ to $41.3 \pm 6.7\%$), and end at a higher performance of $tpr = 46.3 \pm 7.3\%$.

Even with a $fpr = 0\%$ updating threshold, it can be observed that this FR scenario benefits from a self-updating strategy, as the addition of up to an average 13.7 ± 2.4 templates in the galleries (see Fig. 7 (c)) enabled to increase the system’s performance. In addition, despite the limited amount of captures (10 per session), the filtering of the *context-sensitive* self-updating system enabled to maintain a comparable level of performance with a significantly lower amount of templates in the gallery, ending at an average of 6.1 ± 0.9 templates.

Despite the relative simplicity of this scenario and the restrictive updating threshold, impostor templates have been incorrectly added to the galleries during the updating process. Following Doddington’s analysis, the ratio of impostor over genuine templates in the galleries of the top 10 lamb individuals (i.e. the individuals with the highest ratio) are presented in Fig. 7 (d). While 95% of the galleries contain under 10% of impostor samples, two lamb-like individuals (ID 17 and 22) stand out with over 10% and 20% impostor samples in their galleries.

Figure 8 presents the average performance results for the $fpr = 1\%$ updating thresholds for the self-updating techniques. An overall performance increase is

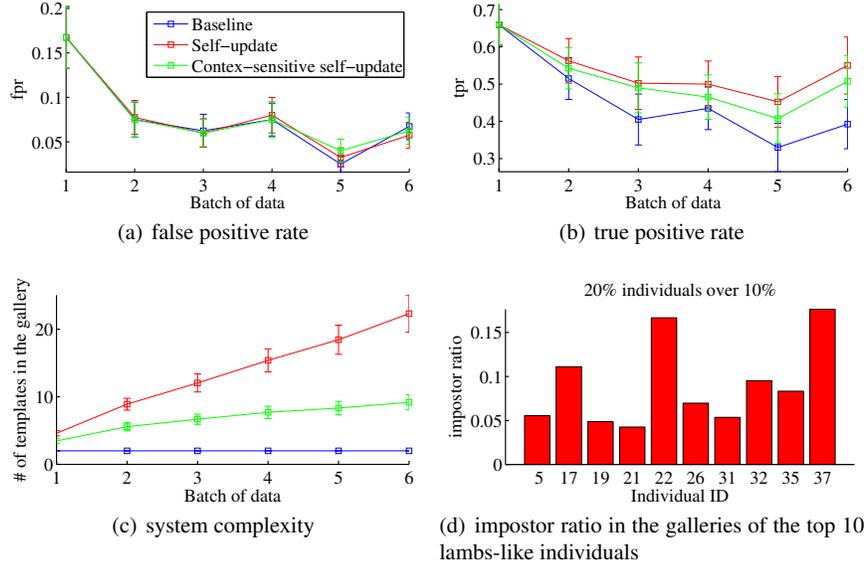


Fig. 8 Simulation results with DICE dataset where the updating threshold is selected for $fpr = 1\%$.

shown for the self-updating methods. A higher tpr is observed throughout the entire simulation, ending at $tpr = 55.0 \pm 7.7\%$ for self-updating, and $tpr = 50.8 \pm 7.0\%$ for *context-sensitive* self-updating (see Fig. 8 (b)).

While results with self-updating are higher in this application, it is important to note that improvements come at the expense of a doubled average gallery size (see Fig. 8 (c)), as well as an increase in the impostor ratio (see Fig. 8 (d), 20% of the galleries are composed by more than 10% impostor templates). Comparing these ratios with the previous ones (in Fig. 7), it is apparent that this increase is not connected to specific lamb-type individuals, but to all the enrolled individuals. This underlines the importance of updating thresholds, specifically for long-term operations where the impostor ratio would be likely to grow exponentially as the facial models become corrupted.

5.2 Video Surveillance with FIA Data

Figure 9 presents the average performance results for the $fpr = 0\%$ updating thresholds for the self-updating techniques. In this scenario involving more sources of variations in capture conditions than the DICE dataset (see Table 1), the benefits of a self-updating strategy are more significant, as the self-updating systems exhibit a significantly higher tpr during the entire simulation (see Fig. 9 (b)). From batch 2 to

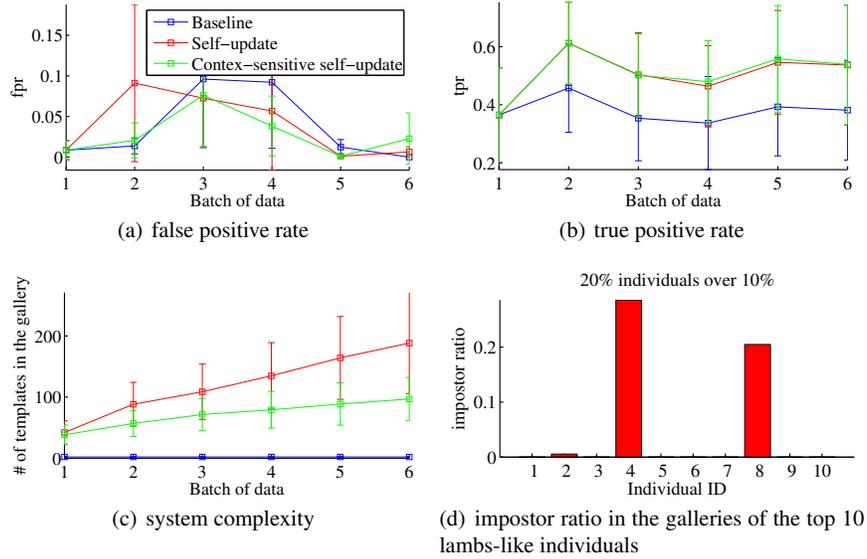


Fig. 9 Simulation results with FIA dataset where the updating threshold is selected for $fpr = 0\%$.

6, the self-updating systems are stable close to $tpr = 60\%$ (both ending at $53 \pm 20\%$), while the baseline system remains close to $tpr = 40\%$ (ending at $38.1 \pm 17.2\%$).

As a consequence of the more complex nature of a semi-controlled surveillance environment as well as the higher number of facial captures, performance improvements come at the expense of significantly larger galleries than with the DICE dataset (see Fig. 9 (c)), ending at an average of 188 ± 83 templates for self-update, and 97 ± 35 templates for *context-sensitive* self-update. It can still be noted that the filtering strategy of the *context-sensitive* self-update technique enables to maintain a comparable level of performance, for gallery sizes approximately two times smaller.

Among the 10 individuals of interest, 2 lamb-like individuals (ID 4 and 8) can be identified, with an impostor ratio over 20% (see Fig. 9 (d)). Despite the added complexity of a semi-constrained environment, the higher number of faces captured in video streams enables a better definition of facial models of target individuals during the first batch. This explains that impostor templates have only been added to two difficult lamb-type individuals, and not all the galleries.

In Figure 10 (b), it can be observed that a more relaxed $fpr = 1\%$ constraint for the updating threshold didn't have a significant impact on the performance of self-updating systems. However, the average gallery size of the self-updating technique increased to end at 268 ± 71 templates, while the *context-sensitive* self-updating technique enabled to remain at a lower size of 109 ± 38 templates (see Fig. 10 (c)), comparable to the $fpr = 0\%$ threshold results (see Fig. 9 (c)). This observation reveals that a majority of the new templates added with the $fpr = 1\%$ thresholds contained redundant information, that was already present in the galleries. This under-

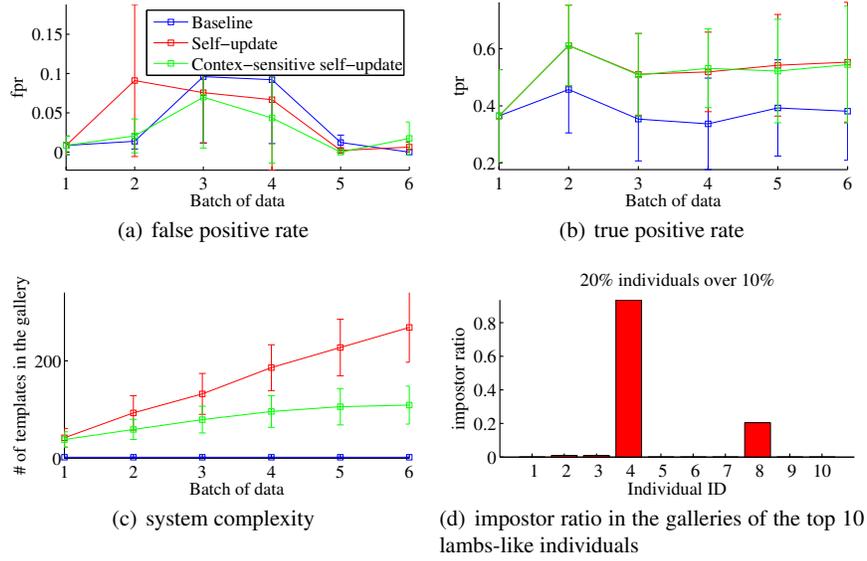


Fig. 10 Simulation results with FIA dataset where the updating threshold is selected for $fpr = 1\%$.

scores the benefits of the *context-sensitive* self-updating technique when operating with videos, where higher quantities of templates may be selected for self-updating. By reducing the number of updates, this technique enables to mitigate the growth in computational complexity of the prediction process as well as the need to use a costly template-management system, without impacting system performance.

Impostor ratios in Fig. 10 (d) show a significant increase for individual ID 8, which ends at 80%. This confirms the rapid addition of impostor templates to the galleries in long term operations. In this video-surveillance scenario where more facial captures are presented to the system (compared to the DIEEE scenario), the gallery of lamb-like individual 4 is updated with a larger amount of impostor templates at the beginning of the simulation. This gallery then keeps attracting impostor templates over time, which reduces the pertinence of the facial model.

5.3 Unconstrained Face Recognition with FRGC Data

Figure 11 presents the average performance results for the $fpr = 0\%$ updating thresholds for the self-updating techniques. It can be observed in Fig. 11 (b) that this scenario represents a significantly harder FR problem, as all three systems perform below $tpr = 23\%$ during the entire simulation. In addition, despite the increase in average gallery size up to respectively 18.8 ± 2.7 and 10.8 ± 1.5 templates for the self-update and *context-sensitive* self-update techniques (see Fig. 11 (c)), only

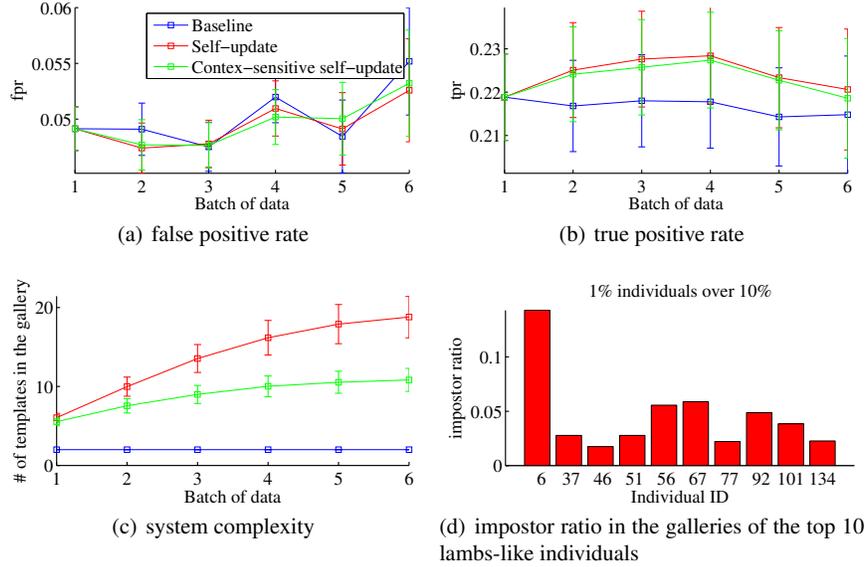


Fig. 11 Simulation results with FRGC dataset where the updating threshold is selected for $fpr = 0\%$.

a marginal performance gain can be observed. The two self-updating systems end at $tpr = 22.1 \pm 1.4\%$ and $tpr = 21.9 \pm 1.4\%$, while the baseline case exhibits a $tpr = 21.5 \pm 1.4\%$.

A bigger impact can be observed in Fig. 12 (b), presenting tpr performance of the three systems for the $fpr = 1\%$ updating threshold. From batch 2 to 6, the two self-updating cases present significantly higher tpr performance, both ending at $tpr = 25.6 \pm 1.5\%$. However, as in the previous scenarios, this performance gain comes at the expense of a significantly higher system complexity. Both systems with self-update end with respectively 82.4 ± 5.2 and 41.6 ± 2.0 templates in the galleries (see Fig. 12 (c)). The average impostor ratio also increased significantly, as 18% of the galleries contain more than 10% impostor templates (see Fig. 12 (d)), while only 1% of the galleries were in this situation with the $fpr = 0\%$ updating threshold.

Results are related to the nature of the scenario presented in Section 4.1.3. The multiple enrolment sessions (up to 16), where small numbers of ROI were captured (6 ROIs), favour the presence of genuine captures that are different enough to fail the updating threshold test. Fewer than 20 templates per individuals have been added to self-updating galleries with $far = 0\%$ self-updating threshold (see Fig. 11 (c)), despite the presence of more than 100 genuine samples in batches. In addition, the systems are initialized with the first capture session, where only 4 controlled stills are available to build the facial model before processing uncontrolled captures in future sessions. This prevents the generation of representative facial models, that

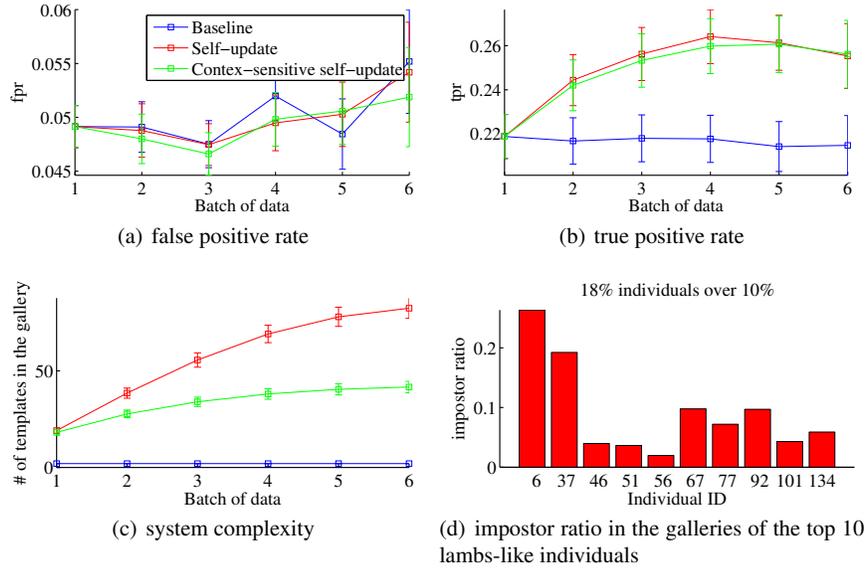


Fig. 12 Simulation results with FRGC dataset where the updating threshold is selected for $fpr = 1\%$.

either reject a majority of genuine templates, or accept a significant amount of impostor templates depending on the updating threshold (see Fig. 12 (d)).

Despite the improved performance achieved using self-updating techniques, this dataset raises the limitations of using a self-updating system relying on a two-threshold update strategy in complex environments, with limited reference data and uncontrolled variations in capture conditions.

5.4 Summary and Discussions

In all experimental results, the following general observations have emerged:

1. Both self-updating techniques generate a significant and stable performance boost over time.
2. The template filtering strategy of the proposed *context-sensitive* self-updating technique significantly reduces system complexity. The galleries are approximately 2 times smaller than a standard self-updating system, without impacting performance.
3. Using a less stringent constraint of $fpr = 1\%$ for the updating threshold does not always have an impact on the performance boost, but always increases system complexity as well as the number of impostor templates in the galleries.

While these observations remain valid for each scenario, a more precise analysis reveals potential limitations of these approaches depending on the represented application.

In a semi-controlled FR application with limited changes mainly caused by illumination and expression (DIEE dataset), benefits of a self-updating techniques are quite clear. In fact, despite the increase in the number of impostor samples in the gallery, a significant performance boost can be observed when a more relaxed updating-threshold is selected.

In the case of a video-surveillance scenario involving a higher amount of impostor individuals not modelled by the system (FIA dataset), a more relaxed updating threshold didn't show any performance improvement, despite a doubled average gallery size for the self-updating technique (while the *context-sensitive* self-update technique prevented any increase in average gallery size). While the overall performance wasn't lowered, the gallery of one specific individual was severely affected, ending with around 80% of impostor samples. In such scenario, involving multiple causes of variation (face angle, resolution, motion blur, etc.) as well as a greater amount of impostor individuals, manual intervention may be necessary at regular intervals, to ensure that the gallery of some specific individuals (lambs) are not getting corrupted over time.

Finally, in the more complex scenario represented by the FRGC dataset, the performance gain observed with the self-updating techniques was considerably lower, even with the less stringent updating threshold of $far = 1\%$. In this scenario, systems are presented with significantly different samples in early operations (after the 4th image), as opposed to the DIEE and FIA scenarios (with respectively 10 and around 30 samples for a first session). In such application, a manual intervention may be required at the early stages of operations, to ensure that the facial models are initialized with enough representative templates to be able to keep updating over time.

6 Conclusion

Despite the advances in feature extraction and classification techniques, face recognition in changing environments remains a challenging pattern recognition problem. Changes in capture condition or individuals physiology can have a significant impact on a system performance, where initial facial models are often designed with a limited amount of reference templates, and frequent re-enrolment sessions are not always possible. Adaptive classification techniques have been proposed in the past decade to address this challenge, relying on operational data to adapt the system over time. Among them, self-updating techniques have been proposed for automatic adaptation using highly-confident captures labelled by the system. While this enables to automatically benefit from a considerable source of new information without requiring a costly manual updating process, these systems are particularly sensitive to their internal parameters. A trade-off between assimilation of new

information and protection against the corruption of facial models with impostor templates has to be considered, as well as a limitation of system complexity over time. While template management techniques can be used to limit system complexity, they remain costly and may interfere with seamless operations.

In this chapter, self-updating methods have been surveyed in the context of a face recognition application with template matching. A *context-sensitive* self-update technique has been presented to limit the growth in system complexity over time, by relying on additional information related to the capture conditions. With this technique, only highly-confident faces captured under new conditions are selected to update individual facial models, effectively filtering out redundant information. A specific implementation of a template matching system with *context-sensitive* self-update has been proposed, where changes are detected in illumination conditions. Proof-of concept experimental simulations using three publicly-available face databases showed that this technique enables to maintain the same level of performance than a regular self-updating template matching system, with a significant gain in terms of memory complexity. By using additional information available in the face captures during operations, this technique allows to reduce the size of template galleries by half, effectively mitigating the computational complexity of the recognition process over time. In applications where memory footprint has to be restricted, this strategy would also limit the need to use costly template management techniques during operations.

However, application-specific limitations have been observed during simulations. When faced with recognition environments with significant variations, and a limited pool of reference patterns for initial enrolment, self-updating systems can be very sensitive to the initialization of their template galleries, as well as the updating threshold. A stricter updating rule may be required to prevent updating with impostor samples, which can significantly reduce the benefits of a self-updating strategy that would never detect any highly confident samples. In addition, while the proposed *context-sensitive* self-updating techniques enabled to significantly reduce system complexity, it relies on the storage of input ROIs in addition to reference patterns in the galleries, as well as an additional measurement during operations.

While self-updating techniques can significantly improve the recognition performance of face recognition systems, their implementation should always be tailored to the specificities of the application as well as the recognition environment. While human intervention can be reduced with automatic strategies, it will still play a significant role in certain applications, especially when dealing with significant variations in capture conditions. In those cases, occasional manual confirmation should be considered, in order to maintain the system's performance by adapting to abrupt changes.

References

1. A. Narasimhamurthy, A., Kuncheva, L.I.: A framework for generating data to simulate changing environments. In: Proc. of the 25th IASTED International Multi-Conf.: artificial intelligence and applications, pp. 84–389 (2007)
2. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: application to face recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence* **28**(12), 2037–41 (2006)
3. Carpenter, G.A., Grossberg, S., Reynolds, J.H.: Artmap. supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks* **4**(5), 565–588 (1991)
4. De Marsico, M., Nappi, M., Riccio, D., Wechsler, H.: Robust face recognition for uncontrolled pose and illumination changes. *IEEE Transactions on Systems, Man, and Cybernetics* **43**(1), 149–163 (2012)
5. Didaci, L., Marcialis, G.L., Roli, F.: Analysis of unsupervised template update in biometric recognition systems. *Pattern Recognition Letters* **37**, 151–160 (2014)
6. Doddington, G., Liggett, W., Martin, A., Przybocki, M., Reynolds, D.: Sheep, goats, lambs and wolves: A statistical analysis of speaker performance in the nist 1998 speaker recognition evaluation. Tech. rep., DTIC Document (1998)
7. Duda, R.O., Hart, P.E.: *Pattern recognition and scene analysis* (1973)
8. Freni, B., Marcialis, G.L., Roli, F.: Replacement algorithms for fingerprint template update. In: *Image Analysis and Recognition*, pp. 884–893. Springer (2008)
9. Goh, R., Liu, L., Liu, X., Chen, T.: The cmu face in action (fia) database. In: *Analysis and Modelling of Faces and Gestures. Second International Workshop. AMFG 2005. Proceedings*, pp. 255–63. Springer-Verlag, Berlin, Germany (2005)
10. Gorodnichy, D.O.: Video-based framework for face recognition in video. In: Proc. of 2nd Canadian Conference on Computer and Robot Vision, pp. 330–338 (2005)
11. Jafri, R., Arabnia, H.: A survey of face recognition techniques. *Journal of Information Processing Systems* **5**(2), 41–68 (2009)
12. Jiang, X., Ser, W.: Online fingerprint template improvement. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **24**(8), 1121–1126 (2002)
13. Li, S.Z., Jain, A.K.: *Handbook of Face Recognition*, 2nd edn. Springer Publishing Company, Incorporated (2011)
14. Liu, X., Chen, T., Thornton, S.M.: Eigenspace updating for non-stationary process and its application to face recognition. *Pattern Recognition* **36**(9), 1945–1959 (2003)
15. Marcialis, G.L., Rattani, A., Roli, F.: Biometric template update: an experimental investigation on the relationship between update errors and performance degradation in face verification. In: *Structural, Syntactic, and Statistical Pattern Recognition*, pp. 684–693. Springer (2008)
16. Marcialis, G.L., Roli, F., Fadda, G.: A novel method for head pose estimation based on the "virtuvian man". *International Journal of Machine Learning and Cybernetics* **5**(11), 111–124 (2014)
17. Nagy, G.: Classifiers that improve with use. In: *Procs. Conference on Pattern Recognition and Multimedia*, vol. 103, pp. 79–86. IEICE, Tokyo, Japan (2004)
18. Niinuma, N., Park, U., Jain, A.: Soft biometric traits for continuous user authentication. *IEEE Trans. on Information Forensics and Security* **5**(4), 771–780 (2010)
19. Pagano, C., Granger, E., Sabourin, R., Gorodnichy, D.O.: Detector ensembles for face recognition in video surveillance. In: *Neural Networks (IJCNN), The 2012 International Joint Conference on*, pp. 1–8. IEEE (2012)
20. Pagano, C., Granger, E., Sabourin, R., Marcialis, G., Roli, F.: Adaptive ensembles for face recognition in changing video surveillance environments. *Information Sciences* **286**, 75–101 (2014)
21. Phillips, P.J., Flynn, P.J., Scruggs, T., Bowyer, K.W., Chang, J., Hoffman, K., Marques, J., Min, J., Worek, W.: Overview of the face recognition grand challenge. In: *Proc. of the 2005 Conference on Computer Vision and Pattern Recognition (CVPR'05)*, pp. 947–954. IEEE (2005)

22. Rattani, A.: Adaptive biometric system based on template update procedures. Ph.D. thesis, PhD Thesis, University of Cagliari, Italy (2010)
23. Rattani, A., Freni, B., Marcialis, G.L., Roli, F.: Template update methods in adaptive biometric systems: A critical review. In: *Lecture Notes in Computer Science (included Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 5558, pp. 847 – 856 (2009)
24. Rattani, A., Marcialis, G.L., Roli, F.: Capturing large intra-class variations of biometric data by template co-updating. In: *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on*, pp. 1–6. IEEE (2008)
25. Rattani, A., Marcialis, G.L., Roli, F.: Self adaptive systems: An experimental analysis of the performance over time. In: *Computational Intelligence in Biometrics and Identity Management (CIBIM), 2011 IEEE Workshop on*, pp. 36–43. IEEE (2011)
26. Rattani, A., Marcialis, G.L., Roli, F.: A multi-modal dataset, protocol and tools for adaptive biometric systems: a benchmarking study. *International Journal of Biometrics* **5**(4), 266 – 287 (2013)
27. Riedmiller, M.: Advanced supervised learning in multi-layer perceptrons from backpropagation to adaptive learning algorithms. *Computer Standards & Interfaces* **16**(3), 265 – 278 (1994)
28. Roli, F., Didaci, L., Marcialis, G.: Adaptive biometric systems that can improve with use. In: N. Ratha, V. Govindaraju (eds.) *Advances in Biometrics*, pp. 447–471. Springer London (2008)
29. Roli, F., Marcialis, G.L.: Semi-supervised pca-based face recognition using self-training. In: *Structural, Syntactic, and Statistical Pattern Recognition*, pp. 560–568. Springer (2006)
30. Ryu, C., Kim, H., Jain, A.K.: Template adaptation based fingerprint verification. In: *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 4, pp. 582–585. IEEE (2006)
31. De-la Torre, M., Granger, E., Radtke, P.V., Sabourin, R., Gorodnichy, D.O.: Partially-supervised learning from facial trajectories for face recognition in video surveillance. *Information Fusion* (2014)
32. Turk, M.A., Pentland, A.P.: Face recognition using eigenfaces. In: *Proc. of 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 586–91. IEEE (1991)
33. Viola, P., Jones, M.J.: Robust real-time face detection. *International Journal of Computer Vision* **57**, 137–154 (2004)
34. Wang, Z., Bovik, A.C.: A universal image quality index. *Signal Processing Letters, IEEE* **9**(3), 81–84 (2002)
35. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* **13**(4), 600–612 (2004)
36. Wright, J., Yang, A., Ganesh, A., Sastry, S.: Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **31**(2), 210–227 (2009)
37. Zhu, X.: Semi-supervised learning literature survey (2005)