Remote Biometrics for Robust Persistent Authentication

Mads I. Ingwar(✉) and Christian D. Jensen

Department of Applied Mathematics and Computer Science,
Technical University of Denmark, Kongens Lyngby, Denmark
{ming,cdje}@dtu.dk

Abstract. This paper examines the problem of providing a robust non-invasive authentication service for mobile users in a smart environment. We base our work on the persistent authentication model (*PAISE*), which relies on available sensors to track principals from the location where they authenticate, e.g., through a smart card based access control system, to the location where the authentication is required by a location-based service. The *PAISE* model is extended with remote biometrics to prevent the decay of authentication confidence when authenticated users encounter and interact with other users in the environment. The result is a calm approach to authentication, where mobile users are transparently authenticated towards the system, which allows the provision of location-based services. The output of the remote biometrics are fused using error-rate-based fusion to solve a common problem that occurs in score level fusion, i.e., the scores of each biometric system are usually incompatible, as they have different score ranges as well as different probability distributions.

We have integrated remote biometrics with the *PAISE* prototype and the experimental results on a publicly available dataset, show that fusion of two remote biometric modalities, facial recognition and appearance analysis, gives a significant improvement over each of the individual experts. Furthermore, the experimental results show that using remote biometrics increases the performance of tracking in persistent authentication, by identifying principals who are difficult to track due to occlusions in crowded scenes.

1 Introduction

What is in a face? Judging by children's drawings, two circles for the eyes, a line for the mouth, and perhaps a dot for the nose makes a face. While seemingly simple, these archetypical features distil faces down to their basic forms and resemble Haar-like features, which are used in face detection methods to find faces in real-time with robust results [1,2].

Christian D. Jensen—The research leading to these results has received funding from the [European Union] [European Atomic Energy Community] Seventh Framework Programme ([FP7/2007-2013] [FP7/2007-2011]) under grant agreement n [242497].

J. Garcia-Alfaro et al. (Eds.): DPM 2013 and SETOP 2013, LNCS 8247, pp. 250–267, 2014.
DOI: 10.1007/978-3-642-54568-9_16,

Faces are what allow us to differentiate people in a group. It might be a child identifying family members in an old photograph, or security personnel identifying people from their passport photos in the airport. Our faces are the most visible characteristic we have, and together with traits such as fingerprints, palm prints, DNA, and iris patterns possess a high discriminative power. In contrast, hair colour, skin colour, gait, height, and weight all have low discriminative power.

The discriminative power of these traits must be considered in security sensitive biometric applications where the performance of the biometric system is important, for instance in some airports, where holders of biometric passports can go through automated gates that authenticate them using facial-recognition. These security sensitive applications of biometric authentication requires robust and accurate results, but, at the same time they must satisfy user demands of a non-invasive and user friendly authentication process.

In his vision of ubiquitous computing, Mark Weiser states that technology must be *calm* [3,4] in order to allow users to focus on their primary tasks. This implies that any authentication technology should require minimal attention from the users, which excludes the use of many authentication techniques, such as passwords or fingerprints. This lends itself to the use of remote biometrics, that is, biometric characteristics that are measurable from a distance without user interaction, such as facial recognition, appearance or gait analysis.

In this paper we extend our Persistent Authentication model (*PAISE*) [5,6] with continuous authentications using remote biometrics. The *PAISE* model combines traditional authentication mechanisms with location information and tracking of principals. The goal in persistent authentication is to translate user authentication from a single event to a lasting session. The model uses strategically placed authentication points to establish an initial authentication session and principals are then tracked throughout the environment. In this paper we explore the addition of remote biometrics, which are regularly measured to prevent the decay of authentication confidence when authenticated users encounter and interact with other users in the environment. This multi-factor approach gives robust results by utilising the strengths of an interaction-based authentication system with the continuous evaluation of an unobtrusive biometric system.

One of the key applications of persistent authentication is to allow secure provision of location-based services, through calm authentication of mobile users in the smart environment. Indoor location systems have seen an increase in popularity in recent years. In particular, tracking of inhabitants in indoor environments have become vital in hospitals to locate and page staff, in homes for elderly people, and in industry for applications in logistics, warehousing and automation. Persistent authentication extends these applications by utilising the credentials associated with each principal's authentication session. This allows persistent authentication to act as the context manager in a sensor enhanced access control system [7], providing a fine-grained and flexible access control mechanism. The *PAISE* model makes it possible to take informed decisions based on the user's credentials, for instance, detecting that the cleaning personnel are accessing a restricted area, or that the carrier delivering goods is entering the premise through the loading area.

The users credentials are captured by the access control mechanism and provided to the persistent authentication system, which tracks the users and, as needed, verifies the identity of the users based on their biometric characteristics. To do so, a specialised algorithm, known as a biometric expert, processes a sample of the characteristic, referred to as the modality. The expert extracts a small amount of data containing the minutia features of the characteristic, called the biometric signature, which represents the unique aspect of the modality. The biometric signature is compared to a reference database, called the template, which links the true identity of the person to the previously captured biometric samples for that person. A match score is generated between the sample and the template, reflecting the expert's confidence in the identity of the person. Alternatively, the expert can be used for identification purposes, in which the persons signature is compared to all templates and the best match returned, however, in this paper we focus on biometric verification.

The main challenge in biometric verification is that the process is not reliable: an expert may reject a genuine user, or conversely, an expert may accept an impostor. A biometric expert may have insufficient discriminative power, especially within a large group [8], or adverse environmental conditions, such as dust or poor luminosity, can affect the quality of biometric acquisition. These factors are further compounded when using remote biometrics as the quality and resolution of the biometric acquisition is significantly lower due to the uncontrolled acquisition process.

The reliability of remote biometrics can be improved by employing multiple biometric experts and fusing their outputs. In this paper we use *Error-Rate-based Fusion* [9], a novel fusion strategy that transforms individual scores into objective evidences and combines them using Bayesian inference. In more details, let us assume that an expert generates a match score y_i and the expert takes a decision that the claimant is genuine. The false acceptance rate (FAR) at the decision threshold y_i represents the probability that the claimant is an impostor. Similarly, the false rejection rate (FRR) at the threshold y_i represents the intrinsic probability of incorrectly rejecting a genuine user. Bayesian inference is used to combine the false acceptance and false rejection rates of different scores, calculated by different experts, and generate a confidence value representing the probability that claimant is genuine.

We evaluate the performance of our error-rate-based fusion strategy using two biometric experts, facial recognition and appearance analysis, on the publicly available CAVIAR dataset [10]. Our experimental results show a significant improvement in the error rate compared to the performance of each individual expert. In addition, we evaluate the increased tracking accuracy and persistence gained by including remote biometrics in the persistent authentication system. Our results show that including remote biometrics significantly improves tracking by identifying principals who are difficult to track due to environmental factors or occlusions in crowded scenes.

The rest of this paper is organised in the following way: an overview of the remote biometrics used in this paper is given in Sect. 2. Fusion of biometric

experts and a quick overview of error-rate-based fusion is presented in Sect. 3. Persistent authentication and the PAISE model are presented in Sect. 4. Our experimental results are presented in Sect. 5 and related work is examined in Sect. 6. Finally, Sect. 7 presents our conclusion and outlines the directions for future work.

2 Remote Biometrics

Compared to their intrusive counterparts, remote biometrics have a lower discriminative power and a higher error rate [11], but they are non-invasive and allow continuous authentication. This ensures a *calm* authentication process without user interaction. The two biometric characteristics we focus on in this paper are facial recognition and appearance analysis based on colour profiles. Our faces possess a high discriminative power, whereas our appearance, in terms of hair and skin colour and the clothes we wear, have a low discriminative power.

For facial recognition we use a linear subspace technique to project high-dimensional data into a lower dimensional subspace by linearly combining features. *Principal Component Analysis* (PCA) [12] and *Linear Discriminant Analysis* (LDA) [13] are well established linear subspace techniques and are considered the most robust methods for face recognition [14].

Consider a set of N facial images $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N$ with values in an n-dimensional image space. A linear transformation maps this n-dimensional image space into a lower m-dimensional feature space $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_N$ such that $\mathbf{y}_k$ represents $\mathbf{x}_k$ by introducing a transformation vector W such that:

$$\mathbf{y}_k = W^T \mathbf{x}_k \qquad k = 1, 2, ..., N$$

For the transformation to accurately represent the original data, it is important to retain the highest possible variation, thus the objective is to find a subspace in which the variance is maximised. Let the total scatter matrix S_T be defined as:

$$S_T = \sum_{k=1}^{N} (\mathbf{x}_k - \mu)(\mathbf{x}_k - \mu)^T$$

Where μ is the mean of all the images. The output is a set of n-dimensional eigenvectors $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m$ corresponding to the m largest eigenvalues, which account for the most variance in the training set. Since these eigenvectors have the same dimension as the original images, they are referred to as Eigenfaces [12].

In PCA, classification can be performed in this reduced feature space, for instance using a nearest neighbour classifier. However, a drawback of this approach is, that much of the variation we seek to maximise is caused by illumination changes [15], thus with images of faces under changing illumination the projected feature space will contain variation due lighting and not necessarily due to class separability. Consequently, the points in the projected space will not be well clustered. A better approach is to use Linear Discriminant Analysis,

where classification is performed by selecting W in such a way that the ratio of the between-class scatter S_B and the within-class scatter S_W is maximised. With the between-class scatter matrix defined as

$$S_B = \sum_{i=1}^{c} N_i(\mu_i - \mu)(\mu_i - \mu)^T$$

and the within-class scatter matrix defined as

$$S_W = \sum_{i=1}^{c} \sum_{\mathbf{x}_k \in X_i} (\mathbf{x}_k - \mu_i)(\mathbf{x}_k - \mu_i)^T$$

where μ_i is the mean image of class X_i, and N_i is the number of samples in class X_i. A projection, W_{opt} is then found, that maximises the class separability criterion

$$W_{opt} = \arg\max_W = \frac{|W^T S_B W|}{|W^T S_W W|}$$

For appearance analysis we use colour profiles, calculated using histogram comparison. Colour histograms are widely used for content-based image retrieval [16] as they are fast to compute, and despite their simplicity, have attractive properties. Since they contain no spatial information they are largely invariant to rotation and translation of objects in the image. Additionally, colour histograms are robust against partial occlusions and changes in camera viewpoint [17].

Colour histograms are typically represented in the RGB colorspace, and the difference between two histograms h_1, h_2 are expressed by the chi-squared distance:

$$\chi^2(h_1, h_2) = \frac{1}{2} \sum_k \frac{(h_{1k} - h_{2k})^2}{h_{1k} + h_{2k}}$$

To reduce the error rate of the remote biometric experts, the output of each of these experts are fused, which increases the robustness of the evaluation.

3 Fusion of Biometric Experts

The main challenge in biometric fusion is that different biometric experts generate matching scores in different domains, and that these domains usually follow different probability distributions. Therefore, score normalisation and transformation are required to make the scores compatible, which are error prone processes. Moreover, the existing parametric models assume a certain distribution of scores which also introduces errors in the fusion process.

In our fusion strategy, error-rate-based fusion, we use measures of false acceptances and false rejections, which have the same definitions across different experts, and therefore do not need any normalisation. We work in a non-parametric model, namely we estimate false acceptance and false rejection rates

for certain discrete levels of thresholds. Further, the fused output is a confidence measure, which is a continuous probability value; therefore, the decision errors associated with a binary decision do not occur. In the following, we give a brief overview of error-rate-based fusion, and we refer interested readers to the complete algorithm presented in Ingwar et al. [9].

For biometric verification we consider two class labels, A and $\bar{A}$, where A is assigned when the expert concludes that a claimant is genuine, and $\bar{A}$ is assigned if the authentication status of the claimant is unknown. If the claimant is A but the expert wrongly labels him $\bar{A}$ then this event is called a false rejection (FR). Similarly, if the claimant is not A and an expert wrongly labels him A then this event is called a false acceptance (FA). The false acceptance and the false rejection rates (FAR and FRR) correspond to the fractions of FA and FR events taken over all genuine and impostor access.

Let us consider N biometrics experts. The output of the i-th expert is a match score, $y_i \in \mathbb{R}$, where $1 \leq i \leq N$.

For a decision threshold Δ_i, the decision function is defined as follows:

$$decision(\Delta_i, y_i) = \begin{cases} accept & \text{if } y_i \geq \Delta_i \\ reject & \text{otherwise} \end{cases}$$

With the match score y_i, let the functions $FAR(y_i)$ and $FRR(y_i)$ be the false acceptance rate and false rejection rates of the i-th expert with $\Delta_i = y_i$. Since $y_i \in \mathbb{R}$, these functions are continuous, such that $FAR(y_i) \in \mathbb{R}$ and $FRR(y_i) \in \mathbb{R}$. For precise evaluation of $FAR(y_i)$ and $FRR(y_i)$, we use a non-parametric approach, and model them as step functions, in which Δ_i can only take m different values: $\Delta_i \in \{\delta_i^1, \ldots, \delta_i^m\}$, where $\delta_i^1 < \cdots < \delta_i^m$. We call these values of Δ_i error decision thresholds (EDTs). This means that $FAR(y_i)$ and $FRR(y_i)$ are defined over a set of m EDTs.

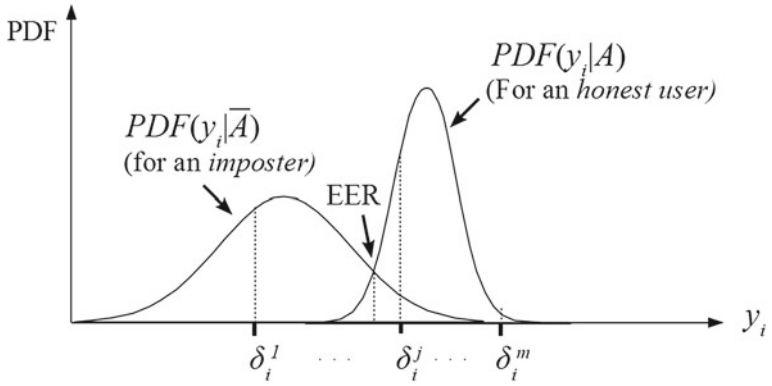

Fig. 1. Error Decision Thresholds (EDTs). Plot of the probability density functions of typical expert scores with the point of Equal Error Rate (EER) shown.

The different values of Δ_i are illustrated in Fig. 1, with a typical plot of the probability density functions (PDF) of expert scores. The figure illustrates that the match score for a genuine user is distributed on larger values as compared to that of an impostor. The figure also shows the point of equal error rate, where the false acceptances and false rejections have the same values.

To illustrate an error-rate-based fusion system, consider a verification system that contains N biometric experts. When biometric data of a claimant is available from the sensors, the system invokes the experts with the claimed identity A. Each expert extracts the relevant biometric signature from the data and compares the extracted signature with the signature templates of A. Each expert then generates a match score y_i, and we compute $FAR(y_i)$ and $FRR(y_i)$ and fuse the match score based on Bayesian inference.

The system has an a-priori confidence that the claimant is A, which is represented as the probability measure, *Pr(A)*. The complementary confidence that the claimant is not A is *1 - Pr(A)*. We compute a-posteriori confidence, $Pr(A|y_i \geq \Delta_i)$, i.e., the probability that the claimant is A after receiving the evidence y_i that meets the decision threshold Δ_i. For brevity, we do not include the decision threshold in the probability expressions, and therefore we write $Pr(A|y_i \geq \Delta_i)$ as $Pr(A|y_i)$. The value of the a-posteriori confidence is computed as follows:

$$Pr(A|y_i) = \frac{TAR(y_i)Pr(A)}{TAR(y_i)Pr(A) + FAR(y_i)(1 - Pr(A))} \tag{1}$$

With TAR being the *true acceptance rate*, i.e. $1 - FRR(y_i)$. Equation 1 allows us to fuse the outputs of N experts, by taking into account the prior confidence level. To get an intuitive feeling of Eq. 1, let us consider a traditional verification expert, which is assumed to be error free, e.g., a password-based authentication of claimant, A, on a computer terminal. If the password is correct then the computer has full confidence that the claimant is A. For such an expert, the values of FAR and FRR are assumed to be zero. As expected, the confidence is evaluated to *1* in Eq. 1 independent of a-priori confidence. In fact, any expert for which the value of FAR is zero will generate the confidence value of *1*, which is consistent with the fact that with zero false acceptances no impostor can ever be accepted by the expert.

4 Persistent Authentication

The goal in persistent authentication is to translate authentication from a single event to a lasting session. We track principals from the point where they authenticate and throughout the environment. We use closed-circuit television (CCTV) cameras and image processing algorithms to provide the location data, and then employ filtering techniques to associate the location with target principals in consecutive frames.

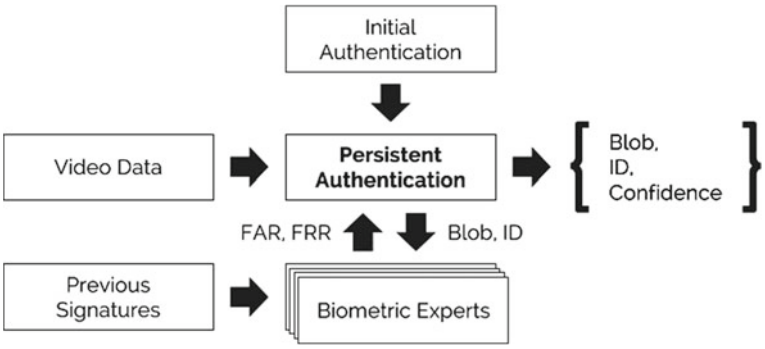

Fig. 2. Overview of the components in the persistent authentication model.

The core component of the *PAISE* model combines data from the authentication system, the smart environment and the biometrics experts, tracks authenticated principals and forwards this information to a location-based service. An overview of the components in the persistent authentication model are shown in Fig. 2. The figure shows how the three components, the authentication system, the smart environment and the biometric experts interface with the core *PAISE* component.

The authentication mechanism handles authentication of principals and provides the initial authentication of the principal. The operation of the authentication mechanism is external to the persistent authentication model, thus *state-of-the-art* solutions are supported, such as intrusive biometrics, smart-cards, wearable tokens, or a combination resulting in multimodal authentication [18].

The smart environment delivers the sensor data needed for tracking. In this paper we use a smart environment that consists of a camera-based location system. CCTV cameras are used to both track principals and to gather remote biometric samples.

The biometric experts process the modalities of the principals captured by the smart environment and returns an estimate of their identity. As mentioned in Sect. 2, the two remote biometrics explored in this paper are facial recognition and appearance analysis.

Finally, the persistent authentication component must: 1. Identify the principals and their locations from the video data and track them throughout the environment, 2. Associate the initial authentication sessions with the corresponding principals, and 3. Continuously provide the biometrics modalities of each tracked principal to the biometric experts and evaluate the feedback. The output is the location of each principal, the associated authentication sessions and the confidence in this assertion.

To identify principals in a video steam, we use image segmentation. In image segmentation objects that share certain characteristics are identified and labeled. In persistent authentication, this means assigning one label to the principals and

Fig. 3. Background segmentation. For each pixel in the image a label w is inferred denoting the absence or presence of a foreground object.

another label to their surroundings. The principals are then referred to as the image *foreground* and their surroundings as the image *background*.

A binary label $w_i \in \{0, 1\}$ is assigned to each pixel $\mathbf{x}_i$ in the image, indicating whether it is part of a known background ($w = 0$) or if it belongs to the foreground ($w = 1$), determined by the recent history of each pixel $\mathbf{x}_1, ..., \mathbf{x}_n$ modelled as a *Gaussian Mixture Model* [19,20]. The probability that a new pixel $\mathbf{x}$ belongs to the foreground is then given by:

$$Pr(\mathbf{x}|w = 1) = \sum_{k=1}^{K} \lambda_k \mathcal{N}(\mu_k, \Sigma_k)$$

where $\mu_{1...K}$ and $\Sigma_{1...K}$ are the means and covariances of the normal distributions and $\lambda_{1...K}$ are positive valued weights that sum to one. The combination of these normal distributions allows the Gaussian mixture model to describe complex multi-modal probability densities. The Gaussian mixture model is robust to noise and changes in illumination and it handles reflections and shadows well, making it particular suited for indoor surveillance applications. A typical result of the labelling process is shown in Fig. 3. The first image shows a complex scene, captured by a CCTV camera, containing five principals annotated with circles. The second image shows the output of the Gaussian mixture model, a black-and-white binary image. The white pixels in the binary image, also called the blobs, indicate the presence of a principal, and the figure shows that all five principals have been correctly identified.

With the foreground objects identified and labeled, we track them throughout the environment. The objective of the tracking is to associate the location of target principals in consecutive video frames. This association can be especially difficult when multiple users are in the environment, when users are occluded, or when the quality of the images are poor due to environmental conditions. In these situations the tracking system relies on the correlation of principals over time, either inferred from the physical properties of the environment or from a model which describes how the location of the target might change for different possible motions of the principals.

We use a combination of spatial-temporal coherence, filtering and flow techniques to ensure consistent tracking in consecutive frames. The spatial-temporal coherence uses the physical reality of the world to infer correlation. Spatial coherence describes the correlation between signals at different points in space, while temporal coherence describes the correlation between signals observed at different moments in time. In tracking this is used to infer correlation based on the speed and trajectory of the principals, which must be consistent with the physical restrictions of the environment.

In addition to the spatial-temporal coherence, we filter the output of the image segmentation process with the Kalman filter [21,22], which in essence is a sensor fusion algorithm that uses the system dynamics model to form an estimate of the system's state, which improves tracking under rapidly changing environmental conditions.

Finally dense optical flow [23,24] is used as a global approach, that is not affected by labelling ambiguities to ensure consistent tracking even in noisy situations. In optical flow it is assumed that when a pixel moves from one frame to another, its intensity or colour does not change. This is a combination of a number of assumptions about the reflectance properties and illumination of the scene and is known as the *brightness constancy*. Solving the brightness constancy results in the magnitude and direction of motion for each pixel in the image. By comparing the displacement to the Kalman filter estimate, tracking becomes possible even when principals are partially occluded, as their direction in the environment helps to differentiate them. Additionally, as the optical flow analysis is applied directly to the image it helps ensure that errors in the labelling process does not carry over into the tracking process.

Tracked principals continuously have their remote biometrics measured and compared to a signature database. This database contains all previous matching signatures and, optionally, high quality enrolment signatures. For each biometric characteristic a set of false acceptance and false rejection rates (FAR and FRR) values are generated. These values are fused, using error-rate-based fusion, which helps reduce the impact of the high error rate of remote biometrics. The result is a biometric confidence score in the identity of the principal, i.e., the confidence on the assertion that a blob has a certain identity. This score is matched with the trackers current confidence score, which turns the confidence into a dynamic value based on positive biometric signatures.

A dynamic score allows the system to take occlusions and other noisy measurements into consideration when determining the confidence in identity, such that, when principal moves through the environment, the confidence in his identity changes based on the quality of the tracking. An example is shown in Fig. 4. The figure shows two paths, a solid line that corresponds to the motion of a principal A and a dashed lined that corresponds the motion of a principal who is not A, denoted, $\bar{A}$. Events on the paths have timestamps, and the time t_0 corresponds to the initial authentication, where A is authenticated using an interactive authentication mechanism, giving an initial confidence of 1.

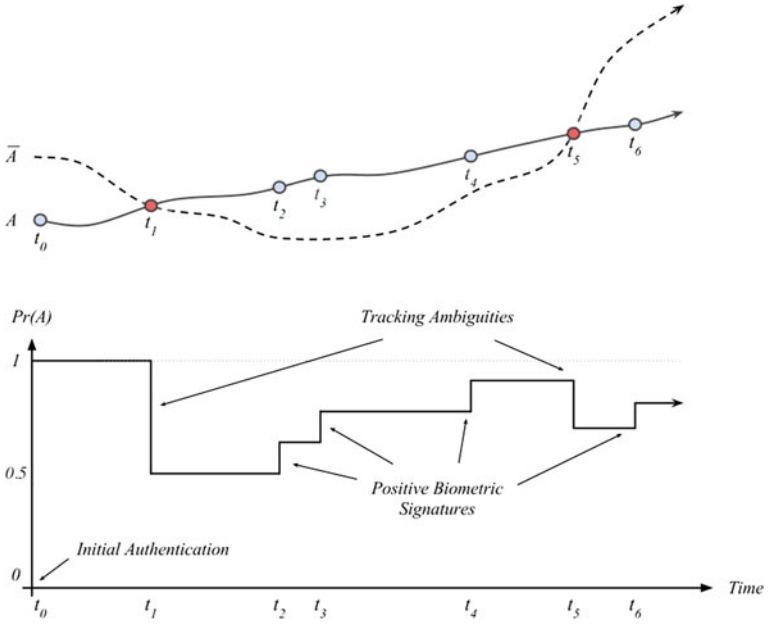

Fig. 4. Confidence in the identity of A. The confidence in A's identity decrease when the paths of A and $\bar{A}$ *intersect and increase with positive biometric signatures.*

The principals are reliably tracked from the point of initial authentication until the time t_1, where occlusions causes ambiguities in the labelling process, which, in turn, causes ambiguity in which of the paths the tracked principal A is following. As a result, the confidence in the identity of A is lowered. How much the confidence is lowered depends on the output of the tracking algorithm, but for the sake of the example, we assume that there is an equal chance of A following either path.

The remote biometrics of A are continuously measured and at time t_2, t_3 and t_4 a positive signature is captured. The resulting biometric confidence score is used to increase the confidence in the identity of A. As A can only follow either the solid line or on the dashed line, the confidence for A on both lines must sum to *1*. Therefore, an increase in the confidence on the solid line automatically decrease the confidence that A is following dashed line. The increase in confidence depends on the quality of the biometric sample and the output of the biometric expert. This cycle of decreasing confidence due to noise or occlusions and increasing confidence with positive signatures continues as long as A is in the environment.

5 Experimental Results

In this section we present and discuss our experimental results. We evaluate how remote biometrics, namely facial recognition and appearance analysis, perform

when implemented in a persistent authentication system. Both of these characteristics are measured from a distance, and authentication is performed continuously by sampling the modality recurrently.

The data used for the experiments are from the CAVIAR dataset [10]. The dataset comprises of a number of clips that show the frontal view of a corridor in a shopping centre. The clips include people walking alone, meeting with others, conversing, and window shopping. All the video clips are filmed in half-resolution PAL standard (384×288 pixels, 25 frames per second) and compressed using MPEG2.

We track each principal in the video and sample the modalities as they are available. As the setting is a corridor with principals walking in both directions, then principals are not always facing the camera and as a result, the facial expert is only able to extract modalities from a subset of the total principals. In contrast, the appearance expert is always available, though the area that is considered may contain little relevant information due to occlusions of the tracked principal. In the dataset 32 unique principals have been identified by both the facial and the appearance expert, on which we test the performance of our error-rate-based fusion technique. We measure the performance of the tracking by recording the number of frames each principal have been successfully tracked by the persistent authentication system and compare this to the ground truth. In addition, we run the experiments again, this time tracking the principals using only the filtering and flow techniques to evaluate the performance without the biometric experts.

The 32 unique subjects are tracked over multiple video clips, in varying poses and illumination. An example of the captured faces for three principals are shown in Fig. 5. The resolution of the video data is low, and as a result, the resolution of the facial images are very low at 50×50 pixels.

We use the first captured face to construct an initial training set, then for each subsequently captured face, we calculate the error rate using leave-one-out cross-validation, after which the new face is added to the training set. It may happen

Fig. 5. Example of the captured faces for three principals from the CAVIAR dataset.

Table 1. Error rates of the biometric experts

Biometric expert	Error rate
Facial expert	4.72 %
Apperance expert	5.01 %
Error-rate-based fusion	**1.44 %**

that a high number of biometric modalities are captured, thus we limit the size of the training set to the six most recently captured images. In a production system, we recommend using high quality enrolment signatures as the initial training set and augmenting the training set with good quality captured samples. The process is completed for both biometric experts and each step is monitored by a human expert who records the performance of the system and of each biometric experts. The resulting error rates are shown in Table 1.

The table shows that the overall performance of the system when using error-rate-based fusion is significantly lower than any of the individual experts. Our fusion technique has an error rate of 1.44 %, which is expected given the performance of the individual experts and the results are in line with the results published in earlier work [9]. The increase in performance is due to the fact that we weigh the decision given by each expert based on their FAR and FRR values as outlined in Eq. 1; a result, in our error-rate-based fusion strategy, the conflict between experts are more likely to be resolved in favour of the best performing expert.

The individual biometric experts have an error rate of 4.72 % for the facial expert and 5.01 % for the appearance expert. We conjecture that the relatively high error rate of the facial expert is caused by the very low resolution of the training images and the greatly varied poses of the principals. However, this shows that even in adverse conditions the LDA method gives robust results. The performance of the appearance expert is not as affected by the low resolution and thus the results are comparable to our previous studies.

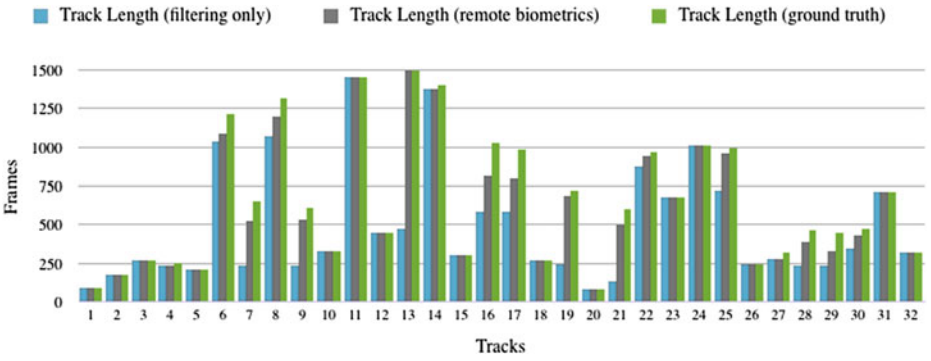

Fig. 6. Computed track lengths vs. ground truth for the CAVIAR dataset

To evaluate the impact of the remote biometrics we compare the performance of the tracking in persistent authentication to the performance without using biometric experts. We use the training data acquired from the CAVIAR dataset and track each of the 32 principals from the point they enter the scene. We measure the number of frames each principal is successfully tracked, with and without the biometric experts, and compare this to the ground truth.

The results are shown in Fig. 6, which charts the results for each the 32 tracks. The majority of the tracks have few or no occlusions and no drop-outs (principals leaving the scene completely), and in these situations both systems achieve near perfect tracking of the principals. The accuracy of the tracking drops when occlusions and drop-outs occur, for instance when principals enter a shop or when multiple principals crowd the scene. The system may completely lose track of a principal, in this case the remote biometrics are used to re-associate the session with the correct principal. As a result, the system using remote biometrics greatly outperforms the other system for a number of the tracks, which is most profound in the tracks 7, 13 and 21.

6 Related Work

In this section, we explore the state of the art related to continuous authentication.

Corner and Noble [25–27] examine the problem of authentication when mobile devices are lost or users leave a workstation logged in. They define traditional authentication mechanisms as *persistent* because they rarely limit the duration that the authentication is valid, so a user may leave a computer logged in for several days. This means that anyone who steals a device that is logged in or gets physical access to the workstation may usurp the authentication of the original user. They define a *transient authentication* mechanism, where all data in the system is encrypted and a small *authentication token*, worn by the user, is needed to provide access to the encrypted data, thus ensuring that access can only be granted when the token is in close proximity to the system. The token stores the cryptographic keys and the proximity mechanism is based on short range wireless communication.

The definitions of persistent and transient authentication by Corner and Noble are device centric, authentication sticks to the device as long as the user is present, so restrictions are put on the users, e.g., they have to wear the authentication token. This creates problems when authentication tokens are forgotten, borrowed or lost. Our definition of persistent authentication is user centric, which means that authentication sticks to the user as long as the tracking is considered reliable. This means that any authentication mechanism, e.g., passwords, PIN or biometrics, can be used and that no additional requirements are placed on the user.

Bardram et al. [28] define a context-aware user authentication mechanism, where users need a smart card to identify themselves to the system and an RFID based tracking system that is used to authenticate the user. This adds complexity

for the users, by requiring them to carry two tokens, without offering significantly improved convenience, i.e., the user still has to insert the smart card into the system whenever authentication is required. In comparison, our method removes the need to perform repeated authentication actions.

Klosterman and Ganger [29] define a *continuous biometric-enhanced authentication* mechanism, which uses a biometric authentication module, based on face recognition, to periodically re-authenticate users who are logged in to the system. If, at some point, the biometrics of the user sitting in front of the monitor does not correspond to the biometrics of the authenticated user, re-authentication is required. This means that continuous authentication is achieved without additional requirements placed on the user, but their system authenticate a specific user at a specific location, whereas we propose to track the user so that his authentication may be reused in different locations.

Altinok and Turk [30] present an approach for temporal integration based on uncertainty propagation over time for a multimodal biometric system. Their method operates continuously by computing expected values as a function of time differences. The system generates continuous results in terms of confidence in the identity of the user, which makes it possible to adjust the security level accordingly in real time. Experimental results with simulated data of face, voice, and fingerprints have shown that the system can provide continuous authentication results which are consistently better than the individual components of the system. The authors conclude that comparing these preliminary results to a true multimodal database is very important for continued work in the field.

Sim et al. [31] develop a continuous authentication system based on multimodal remote biometrics in a Bayesian framework that combines both temporal and modality information holistically. This approach allows the system to evaluate the probability that the user is still present even when there is no observation. The authors are successful in integrating results from a fingerprint biometric classifier with a face classifier and develop a model that intuitively separates the uncertainty of the dynamic model from that of the sensor model. Muncaster and Turk [32] take similar approach as [31], but use a Dynamic Bayesian Network to achieve continuous authentication using multimodal biometrics. The advantage of a dynamic Bayesian network is its ability to account for more hidden variables and by modelling more hidden variables, the network is capable of modelling important contextual information. Both approaches focus on a controlled environment, such as a workstation, where an impostor hijacks a logged-in session. In comparison, persistent authentication operates in an uncontrolled and unconstrained environment, where the sessions are user centric, requiring an impostor to displace a legitimate user instead of hijacking an empty workstation.

Niinuma and Park [33] propose a framework for continuous authentication that uses soft biometrics traits, similar to the appearance analysis presented in this paper. The proposed framework automatically registers soft biometric traits every time the user login and fuses soft biometric matching with conventional authentication schemes, namely password and face biometric. The proposed scheme has high tolerance to the user's posture and the experimental

results show the effectiveness of the proposed method for continuous authentication. The authors make a number of assumptions about the pose of the users and the location of the body for appearance analysis, furthermore, occlusions are handled on a very ad hoc basis. In contrast, persistent authentication uses image segmentation to locate users which ensures that the regions of interest are correctly identified for appearance analysis, additionally, advance filtering algorithms are used to ensure occlusions does not revert the authentication session and require the user to start over.

7 Conclusion

In this paper we examined the problem of providing a robust non-invasive authentication service for mobile users in a smart environment. We used the persistent authentication model, *PAISE*, to track principals and employed continuous authentication, based on remote biometrics, to identify principals and re-associate lost authentication sessions. The result is a calm approach to authentication, where mobile users are transparently authenticated towards the system, which allows the provision of location-based services.

We used error-rate-based fusion to solve a common problem that occurs in score level fusion, i.e., the scores of individual experts are usually incompatible, as they have different score ranges as well as different probability distributions. In our fusion strategy, we use error rates (false acceptance and false rejection rates), which have the same definitions across different domains, and therefore does not require any normalisation.

We evaluated our error-rate-based fusion strategy on two remote biometric modalities, namely facial recognition and appearance analysis. Our experimental results on a publicly available dataset, show that our fusion strategy gives a significant improvement over each of the individual experts. This increase in accuracy is especially useful for security sensitive biometric applications where the performance of the biometric system is important. We further evaluated the performance of the persistent authentication system with regard to the accuracy of the tracking. Our results show that using remote biometrics help identify principals who are difficult to track due to occlusions in crowded scenes. In addition, remote biometrics allows the system to re-identify principals who drop out of view of the camera and re-enter at a later stage.

Finally, we conclude that the *PAISE* model provides a useful abstraction for authentication systems, which may greatly improve the usability of traditional user authentication.

References

1. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Computer Vision and Pattern Recognition (2001)
2. Jones, M., Viola, P.: Robust real-time object detection. In: Workshop on Statistical and Computational Theories of Vision (2001)

3. Weiser, M., Brown, J.: Designing calm technology. PowerGrid J. **1**, 1–5 (1996)
4. Weiser, M.: The computer for the 21st century. Scientific American **265**(3), 66–75 (1991)
5. Kirschmeyer, M., Hansen, M.S.: Persistent authentication in smart environments. IMM-THESIS: 2008–16, Technical University of Denmark (2008)
6. Ingwar, M.I., Jensen, C.D.: Towards secure intelligent buildings. In: Proceedings of the 5th Nordic Workshop on Dependability and Security (NODES'11) (2011)
7. Jensen, C.D., Geneser, K., Willemoes-Wissing, I.C.: Sensor enhanced access control: extending traditional access control models with context-awareness. In: Fernández-Gago, C., Martinelli, F., Pearson, S., Agudo, I. (eds.) IFIPTM 2013. IFIP AICT, vol. 401, pp. 177–192. Springer, Heidelberg (2013)
8. Cole, S.: More than Zero: accounting for error in latent fingerprint identification. J. Crim. Law Criminol. (1973-) **95**(3), 985–1078 (2005)
9. Ingwar, M.I., Ahmed, N., Jensen, C.D.: Error-rate-based fusion of biometric experts. In: PST2013 International Conference on Privacy, Security and Trust (PST) (2013)
10. Fisher, R.: CAVIAR Test Case Scenarios (2004)
11. Bhattacharyya, D.: Biometric authentication: a review. Int. J. u-and e-Service **2**(3), 13–28 (2009)
12. Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cogn. Neurosci. **3**(1), 71–86 (1991)
13. Fisher, R.: The use of multiple measurements in taxonomic problems. Ann. Hum. Genet. **7**(2), 179–188 (1936)
14. Belhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 711–720 (1997)
15. Adini, Y., Moses, Y., Ullman, S.: Face recognition: the problem of compensating for changes in illumination direction. Pattern Anal. Mach. Intell. **19**(7), 721–732 (1997)
16. Flickner, M., Sawhney, H., Niblack, W.: Query by image and video content: the QBIC system. Computer **28**(9), 23–32 (1995)
17. Kakumanu, P., Makrogiannis, S., Bourbakis, N.: A survey of skin-color modeling and detection methods. Pattern Recogn. **40**(3), 1106–1122 (2007)
18. O'Gorman, L.: Comparing passwords, tokens, and biometrics for user authentication. Proc. IEEE **91**(12), 2021–2040 (2003)
19. Stauffer, C., Grimson, W.: Adaptive background mixture models for real-time tracking. Comput. Vis. Pattern Recogn. **2**, 246–252 (1999)
20. Stauffer, C., Grimson, W.: Learning patterns of activity using real-time tracking. IEEE Trans. Pattern Anal. Mach. Intell. **22**(8), 747–757 (2000)
21. Kalman, R.: A new approach to linear filtering and prediction problems. J. Basic Eng. **82**, 35–45 (1960)
22. Welch, G., Bishop, G.: An introduction to the Kalman filter. Technical report, University of North Carolina (1995)
23. Farneback, G.: Fast and accurate motion estimation using orientation tensors and parametric motion models. In: Proceedings of 15th International Conference on Pattern Recognition (2000)
24. Farneback, G.: Very high accuracy velocity estimation using orientation tensors, parametric motion, and simultaneous segmentation of the motion field. In: Proceedings of the Eighth International Conference on Computer Vision 2001, ICCV 2001 (2001)

25. Corner, M., Noble, B.: Zero-interaction authentication. In: Proceedings of the 8th Annual International Conference on Mobile Computing and Networking, pp. 1–11 (2002)
26. Noble, B.D., Corner, M.D.: The case for transient authentication. In: Proceedings of the 10th Workshop on ACM SIGOPS European Workshop: Beyond the PC - EW10 , p. 24 (2002)
27. Corner, M., Noble, B.: Protecting applications with transient authentication. In: International Conference on Mobile Systems, Applications, and Services (MobiSys) (2003)
28. Bardram, J.E., Kjær, R.E., Pedersen, M.Ø.: Context-aware user authentication - supporting proximity-based login in pervasive computing. In: Dey, A.K., Schmidt, A., McCarthy, J.F. (eds.) UbiComp 2003. LNCS, vol. 2864, pp. 107–123. Springer, Heidelberg (2003)
29. Klosterman, A., Ganger, G.: Secure continuous biometric-enhanced authentication. Technical report, Parallel Data Laboratory (2000)
30. Altinok, A., Turk, M.: Temporal integration for continuous multimodal biometrics. In: Proceedings of the Workshop on Multimodal User Authentication (1) (2003)
31. Sim, T., Zhang, S.: Continuous verification using multimodal biometrics. Pattern Anal. Mach. Intell. **29**(4), 562–570 (2007)
32. Muncaster, J., Turk, M.: Continuous multimodal authentication using dynamic Bayesian networks. In: Proceedings of the 2nd Workshop of Multimodal User Authentication (2006)
33. Niinuma, K., Park, U., Jain, A.K.: Soft biometric traits for continuous user authentication. IEEE Trans. Inf. Forensics Secur. **5**(4), 771–780 (2010)