Multimodal Biometrics-based Student Attendance Measurement in Learning Management Systems

Elisardo González Agulla\textsuperscript{1} Enrique Argones Rúa\textsuperscript{1} José Luis Alba Castro\textsuperscript{1} 
\textsuperscript{1}Signal Technologies Group, Signal Theory and Communications Department, University of Vigo 36310, Vigo, Spain \{eli,eargones,jalba,carmen\}@gts.tsc.uvigo.es

Daniel González Jiménez\textsuperscript{2} Luis Anido Rifón\textsuperscript{3} 
\textsuperscript{2}Galician R&D Center in Advanced Telecommunications 36310, Vigo, Spain dgonzalez@gradiant.org
\textsuperscript{3}Telematics Engineering Department, University of Vigo 36310, Vigo, Spain lanido@det.uvigo.es

Abstract

In this paper we present a solution to obtain useful and reliable user logs in a Learning Management System (LMS). Current LMS logs are combined with biometric-based logs that show the student behaviour. Our system models the student behaviour, allowing to know exactly how much time the student spends in front of the computer examining the contents of the LMS. Besides, user verification and face tracking are also integrated, what guarantees that the student is the person actually interacting with the system. The presented multimodal solution for user tracking and user verification combines face tracking, face verification, speaker verification and fingerprint verification. Face tracking and face verification are performed in a non-collaborative fashion. Fingerprint or speaker verification is performed on demand, with the aim of avoiding a negative influence of adverse environmental or behavioural human factors in the reliability of the user logs generated by the system. These circumstances can thwart the non collaborative face verification performance involved in the tracking process. The presented solution solves the problem of user tracking and authentication even in adverse environments for face verification.

1. Introduction

The use of online learning applications is nowadays widely spread and still growing. There exist thousands of courses using e-learning environments currently available. Nevertheless there are some limitations in current LMSs that do not allow to provide virtual educational environments with full online learning capabilities.

One of the most important drawbacks in current LMSs is the lack of adequate tools to properly track the behaviour of the users in the system. Traditional solutions, such as simple session logs where the interactions between user and system are recorded, can not actually guarantee that the user is who he/she claims to be, and even worse, they can not guarantee whether the user is in front of the computer or not. This limitation hardly restricts the use of e-learning solutions for examinations. At least, online exams will provide much less confidence than face-to-face exams do, since it is easy to cheat the online system. Furthermore, even though we would not mind the authenticity of the user identity, the information provided by the traditional logs is far from actually reporting the time a user spends in front of the computer attending the online course contents. An user could log into the LMS, an five minutes later he/she could go out for a coffee, what obviously would not appear reflected in the log.

These are similar situations to the one described above, that are of special relevance in courses where the actual time
spent by the student has to be certified. This is the case for professional certifications, some life-long-learning courses, European ECTS, etc. Current LMS session logs store information about when a user gets a e-learning resource. However there is not any reliable mechanism to obtain accurate information of the user attendance, such as how much time he/she spends attending a given resource. In other domains biometrics have been successfully applied to check the identity of the users. In [5], for instance, the system uses webcams for employee surveillance, replacing the classic punch-card reader with facial recognition software. This made the old canard of getting someone else to clock in for you useless. In [10], detailed session logs from the LMS are combined jointly with a biometric log obtained without the collaboration of the user. The detailed LMS session logs include information about when a resource is being used and about browser events, providing information about when it is in foreground or background. Biometric support provided in this work allows to track the user activity and to assess the user’s identity. This biometric log is performed using non collaborative face tracking and recognition. Nevertheless, face biometrics’ technologies are nowadays far from perfect; there are yet some problems to deal with, such as bad light conditions in the student room, that can lower the verification performance. Furthermore, human factors can seriously degrade the performance of such a system: some users will sit looking at the screen almost all the time, but many others could adopt a non convenient pose for face verification almost all the time, occluding the face, looking down to read or write notes, etc. An alternate solution to use face biometrics is presented in [8], where the author addresses the use of fingerprint verification to support web-based course examinations. This system uses a collaborative approach to solve the user authentication problem in e-learning. However, there are some issues that remain unresolved. For instance, once the user is verified using fingerprint, it would be very easy to cheat the system if no user tracking is performed.

In this paper we propose to combine both collaborative and uncollaborative biometric modalities to obtain a single multimodal log. This biometrics-based log is based on the combination of face tracking, face verification and a collaborative biometric recognition, with the traditional LMS session log. In the current implementation, two different biometric modalities can be chosen for the collaborative process: speaker verification and fingerprint verification. Our system performs non-collaborative face-based tracking and identity verification. Whenever the system is unable to determine the identity of a user for a given period of time, for instance because the light conditions are bad or because the user is not in a frontal pose, a collaborative verification, based on fingerprint or voice, is performed. This will provide information about the user presence and identity even when the face verification is not possible. The incorporation of the collaborative verification will result in increasing the certainty, the reliability and the accuracy of the user tracking logs, even in an adverse scenario for face verification. Besides, multimodal biometric verification, based on face and fingerprint or speaker verification, is also incorporated into the student access control module, hence providing the system with improved access control security.

The rest of the paper is organized as follows: Section 2 describes the multimodal student access control module; Section 3 describes the student attendance module, detailing the tracking process and relevant events; the biometric authentication technology is described in Section 4; Section 5 describes the experimental framework used to test our system; experimental results are presented and discussed in Section 6; and the paper is drawn to conclusions in Section 7.

2. Multibiometric-based Student Access Control Module

The problem of access control using biometrics was addressed using the Biometrics for Web Authentication project [1, 14]. BioWebAuth is an open source Java framework intended to provide single sign-on web authentication based on BioAPI-compliant biometric software or devices. It uses JA-SIG Central Authentication Service architecture [2]. BioWebAuth provides an biometrics-based access control mechanism [11] that offers the LMS the guarantee that the user logging in the system is actually sat in front of the computer, thus avoiding unauthorized accesses without collaboration from a student. This module is mainly based on the work presented in [6].

BioWebAuth allows us to use a variety of combinations of biometric modalities for our access control module. The choice we finally is the combination of face and fingerprint or speaker verification. Several factors must be taken into account for this choice. For instance, fingerprint verification is not eligible if there is not any fingerprint capture device, and voice verification should not be chosen if the environment is noisy. The details concerning the biometric algorithms involved in the voice, fingerprint and face BSPs are described below, in Section 4.

3. Multibiometric-based Student Attendance Module

The student attendance module is designed as a continuous user monitoring system that stores the student behaviour during the whole e-learning session. This user monitoring system is basically composed by four main modules:

**Face tracker** The face tracker was developed using C++ programming language and the image processing li-
brary OpenCV [13]. It combines a face detector, a skin tracker and an eyes tracker to locate and track any new face appearing in the webcam field of view.

BioAPI-compliant face Biometric Service Provider
This element provides with face-based enrollment, verification and identification services to the system.

BioAPI-compliant voice Biometric Service Provider
This element provides with voice-based enrollment, verification and identification services to the system.

BioAPI-compliant fingerprint Biometric Service Provider
This element provides with fingerprint-based enrollment, verification and identification services to the system.

A detailed description of the face Biometric Service Provider (BSP), the voice BSP and the fingerprint BSP is provided below, in Section 4.

![Real-time visual output provided by the face tracking module. The colors of the squares and ellipses describe the authentication state of the detected face. Green stands for the logged student, red stands for an impostor, and orange stands for undetermined.](image1)

When a user enters the field of view of the webcam the system detects his/her face, and as soon as a frontal face with enough quality is detected, the system asks for a verification to the face BSP against the logged student identity. The detected face is then categorized as the true student or as an impostor. This face is continuously tracked afterwards until the tracker loses the face because it goes out of the webcam field of view or the face image is occluded, for instance because the user looks down hiding the face to the webcam. The face verification result is valid for a period of time, after which a new verification is needed. Figure 1 illustrates the visual output of the face tracking module, showing a face that has been verified as the face of the student, a face categorized as an impostor and another that has not been verified yet.

This part of the student attendance module remains the same than the student attendance module presented in [10]. However there is a problem in this approach. If the face tracker is unable to find a frontal face, or if all the detected frontal faces are below the minimum required quality threshold then the system can not ask for a face-based user authentication, and hence the identity of the user remains undetermined. This problem is specially difficult for students that work in an environment without an appropriate illumination. This problem can be solved by using any collaborative biometric authentication method, including for instance a collaborative face recognition. However, the possible problem of bad light conditions would not be solved, what impeled us to adopt a multimodal approach to cope with this problem. When a new face is found by the face tracker module, if there is no chance for a face-based user authentication for a given period of time, the system asks for a collaborative verification. This alternate user authentication allows us to always determine the authenticity of the identity of the users tracked by the system, thus eliminating the uncertainty of the identity of any face detected by the system, whether it is possible to perform a face verification or not.

4. Biometric Service Providers

4.1. Face BSP

The face verification engine is based on multi-scale and multi-orientation Gabor features, which have been shown to provide accurate results for face recognition [17, 16], due to biological reasons and because of the optimal resolution in both frequency and spatial domains [9]. A recent survey on the use of Gabor filters for face-based biometrics [19] reveals the huge number of papers that had adopted such features for facial processing, including [20, 12, 16].

More specifically, the face recognition system relies upon extraction of local Gabor responses (jets) at each of the nodes from a $10 \times 10$ rectangular grid overimposed on the (geometrically and photometrically corrected) face region. The similarity between two jets is given by their normalized dot product, and the final score between two faces combines the 100 local similarities using the median rule. Given that we are working with video sequences, we may use multiple shots (frames) when verifying students’ identities. To provide a general framework, let us assume that we have two videos $V_1$ and $V_2$, where $V_1$ is used for enrollment, and $V_2$ comprises the set of “test” frames that are available whenever verification is needed. The strategy to compare $V_1$ and $V_2$, is to perform pairwise comparisons between all frames in $V_1$ and $V_2$. Hence, a matrix of similarity values $S$ is obtained:
\[ S = \begin{pmatrix}
  s_{1,1} & s_{1,2} & \cdots & s_{1,N_{f_2}} \\
  s_{2,1} & s_{2,2} & \cdots & s_{2,N_{f_2}} \\
  \vdots & \vdots & \ddots & \vdots \\
  s_{N_{f_1},1} & s_{N_{f_1},2} & \cdots & s_{N_{f_1},N_{f_2}}
\end{pmatrix} \]

where \( N_{f_1} \) and \( N_{f_2} \) stand for the number of processed frames in \( V_1 \) and \( V_2 \) respectively. In Equation (1), \( s_{i,j} \) represents the similarity value between the \( i \)-th face from \( V_1 \) and the \( j \)-th image from \( V_2 \). The final score between \( V_1 \) and \( V_2 \) is calculated as follows:

\[ s = \text{median} \{ \max_j \{ s_{i,j} \} \} \quad (2) \]

We use the max operator in order to soften the influence of expression and pose changes that naturally may occur during a student session. In fact, since each face from \( V_2 \) is compared to all faces from \( V_1 \), taking the maximum similarity aims to select the face from \( V_2 \) that shows the most similar expression and pose with an enrolled face within \( V_1 \).

### 4.2. Voice BSP

The speaker verification system is a Gaussian Mixture Model-Universal Background Model (GMM-UBM) system [18] where both UBM and user models are 256 mixtures GMMs. UBM is generated by separately training and merging two gender-specific 128 mixtures GMMs. User model is obtained by means of the MAP adaptation of the UBM to the enrollment features. The audio waveform is processed with a Wiener filter to mitigate the influence of noise. 12 MFCC, the logarithmic energy and their delta and acceleration coefficients are extracted from the denoised audio waveform. Silence frames are then removed by an energy-based Voice Activity Detector (VAD). Given a sequence of speech feature vectors \( S = \{ s_1, \ldots, s_{N_S} \} \), verification score for identity \( u \) is provided by the following logarithmic likelihood ratio:

\[ s_{u,S} = \sum_{i=1}^{N_S} \log_{10} \left( \frac{f_u (s_i)}{f_{UBM} (s_i)} \right), \quad (3) \]

where \( f_u (s_i) \) and \( f_{UBM} (s_i) \) are the user \( u \) and UBM probability density functions respectively.

### 4.3. Fingerprint BSP

Fingerprint authentication software used in this work was developed in two master’s thesis [7, 15], and it uses the Libfprint, which is part of the fprint project [3]. The fingerprint authentication kernel implemented in this software is based on the standard NIST’s NFIIS2, which is now known as NBIS [4]. The main modules of this software are:

- A minutiae detector called MINDTCT.
- A fingerprint image quality assessment algorithm called NFIQ.
- A minutiae-based fingerprint matching algorithm called BOZORTH3, which accepts the minutiae information from the MINDTCT module.

### 5. Experimental Framework

The goal of this study is to check the accuracy of the measurements obtained by our attendance module. The system has been evaluated in a test developed in a real e-learning scenario. Sixteen students of the Image Processing subject, of Telecommunications Engineering School at the University of Vigo, were tracked while they were using the LMS Moodle to solve a quiz test about subject contents. Tests were performed on the usual classroom of the Image Processing subject, where there is a computer with webcam per student. Figure 2 shows a general view of the classroom where the tests are performed.

![Figure 2. View of the classroom where the test is performed. Each student sits in front of a computer with a webcam](image)

The designed test can be briefly divided in three parts: the student fills his/her quiz, the student fills another student’s quiz, and the student consults his/her score in the LMS. More precisely:

1. The student checks out that the webcam is located on the top of the computer and that his/her face is properly framed in the video flow.
2. The templates for the different modalities are created by access control module. Face template is created us-
ing a video of 20s at maximum. Voice template is created from a voice recording of 20s at maximum. One fingerprint from each of the 10 fingers is taken for the fingerprint template.

3. The student, in the role of client, performs a short quiz of 7 questions using Moodle, spending around 8 minutes.

4. The student moves to the computer at his/her left and then, assuming the role of an impostor, performs other short quiz of 7 questions using Moodle, spending around 8 minutes.

5. The student comes back to his/her original computer and checks out the scores obtained in the test, in the role of client.


In order to provide a numerical comparison of the work presented in this paper with the work presented in [10], collaborative verification events are not performed during the test. They are generated offline, as a result of postprocessing the biometric log. When a tracked face is not properly verified as a client or an impostor (undetermined identity) for a period of time higher than a given expiry time, an offline collaborative test is simulated. These tests are performed using the fingerprint modality in this experiment, and the expiry time for undetermined tracked faces was fixed to 30s. The result of this offline verification is to label those undetermined identities as impostors or as clients, depending on each case.

6. Experimental Results

Figure 3 shows us a typical biometric log obtained using the non-collaborative scheme for tracking. As we can see, the system labels some intervals as intervals where the student’s presence is detected and verified, and some other intervals as intervals where there is a person in front of the camera who is verified as an impostor. Figure 4 shows another example of a biometric trace. In this trace we can observe that there is a large interval where the system was unable to properly determine whether the user is the true student or an impostor. This could be because the system did not get any frontal face or because the captured frontal faces do not exceed the minimum required quality for a face verification. The system can overcome this situations by means of a collaborative verification. The biometric log when a collaborative authentication is performed is shown in Figure 5. In this figure we can see that only orange intervals smaller than the expiry time for undetermined tracked faces (30s in our experiments) remain in the log.

Several statistics have been extracted from the biometric logs in order to compare the system performance with and without the aid of the collaborative authentication for undetermined tracked faces. The Total Session Time (TST) is the length of the full session; the Any Presence Detected Time (APDT) is total time that at least one face is being tracked by the system; the Any Presence Detected Time Ratio (APDTR) is the ratio between APDT and TST; the Student Presence Detected Time (SPDT) is the total time that the system tracks a face verified as belonging to the true student, whichever method for verification (collaborative or non collaborative) has been used; and the Student Presence Detected Time Ratio (SPDTR) is the ratio between SPDT and APDT. Results shown are the mean between all the videos recorded and the standard deviation. The statistics TST, APDT and APDTR, which are the same for both

Figure 3. Simplified Biometric Log of a typical student. The most usual situation is that collaborative authentication is not required.

Figure 4. Simplified Biometric Log of one of the students’ videos without collaborative authentication. Large orange intervals in the left correspond to intervals where the system is unable to capture a good frontal face for verification.

Figure 5. Simplified Biometric Log of one of the students’ videos with collaborative authentication. Large orange intervals in the left have disappeared due to collaborative authentication events.
Table 1. Mean and standard deviation of the statistics extracted from the biometric logs without the aid of the collaborative authentication (NCA) and with the aid of collaborative authentication (CA) for undetermined tracked faces. SPDT is given in minutes:seconds

<table>
<thead>
<tr>
<th>System</th>
<th>SPDT</th>
<th>SPDTR%</th>
</tr>
</thead>
<tbody>
<tr>
<td>NCA</td>
<td>12:10 ± 3:54</td>
<td>54% ± 17%</td>
</tr>
<tr>
<td>CA</td>
<td>12:40 ± 3:22</td>
<td>57% ± 15%</td>
</tr>
</tbody>
</table>

systems, are the following:

- **TST** = 26:23 ± 02:49
- **APDT** = 22:33 ± 03:09
- **APDTR%** = 85.81% ± 9.95%

SPDT and SPDTR are different for both systems. The results obtained in our tests can be seen in Table 1. We can see that the system that performs a collaborative user authentication (CA) when the non collaborative approach is unable to determine the user’s identity obtains higher mean SPDT. This happens because the main long intervals where the system was unable to determine the identity of the tracked face were mainly located in the intervals where the true student was in front of the computer. Besides, standard deviation of the SPDT is lower for the CA system. This can be explained because some videos that had a low SPDT in the NCA system have more typical values in the CA system, decreasing the standard deviation of this statistic.

The improvements shown in Table 1 could be seen as not too much impressive. However, we must take into account that the collaborative part of the tracking system is thrown mainly due to human factors. These factors are not easy to measure or compare quantitatively. Some users can partially occlude his/her face, or can adopt a pose far from frontal, so inappropriate for face verification. We should pay a special attention to how well our system is reacting against these behaviours. An important output from our experiments is that only in 3 out of 16 videos the help of the collaborative authentication was required, so most of the users are not asked for a collaborative authentication during the whole test. The most of the users usually behave such that the non collaborative face tracking and verification is enough for an adequate tracking and identity assessment. However, for the users that are asked for a collaborative authentication, the periods of time that the system spent without reliable information about these user’s identity would be unacceptably long if non collaborative authentication was performed: 5:32, 1:50 and 1:25. And most important, only a 18.75% of the users where asked for a collaborative verification, so the degree of comfort remains the same for the majority of the users.

7. Conclusions

We have presented a multimodal system for control access and student attendance measurement system in LMS environments. The objective of the student attendance measurement system is to provide with biometric logs that allow for a reliable feedback of the students’ behaviour while logged into the LMS. The control access is based on multimodal fusion of face and other modality, namely voice or fingerprint. The student attendance measurement system is mainly based on face tracking and face verification. Collaborative authentication has been incorporated in the system in order to avoid long periods of time without reliable information about the identity of the tracked faces. One of the outcomes is that the reliability of the biometric logs is considerably reinforced. This has been achieved by incorporating a collaborative authentication method when the tracking system is unable to authenticate the student’s identity. A remarkable conclusion that we can extract from this work is that human factors can significantly degrade a non collaborative authentication scheme. This degradation in performance is mainly due to special behaviours of some of the users of the system. The effect of incorporating this collaborative authentication is to reduce the influence of such human factors in the performance of the system. Numerically noticeable improvements in terms of Student Presence Detected Time are achieved, and most important, these improvements due to the collaborative part of the system are achieved without being a nuisance, since only 3 out of 16 users are asked for a collaborative authentication.

References


