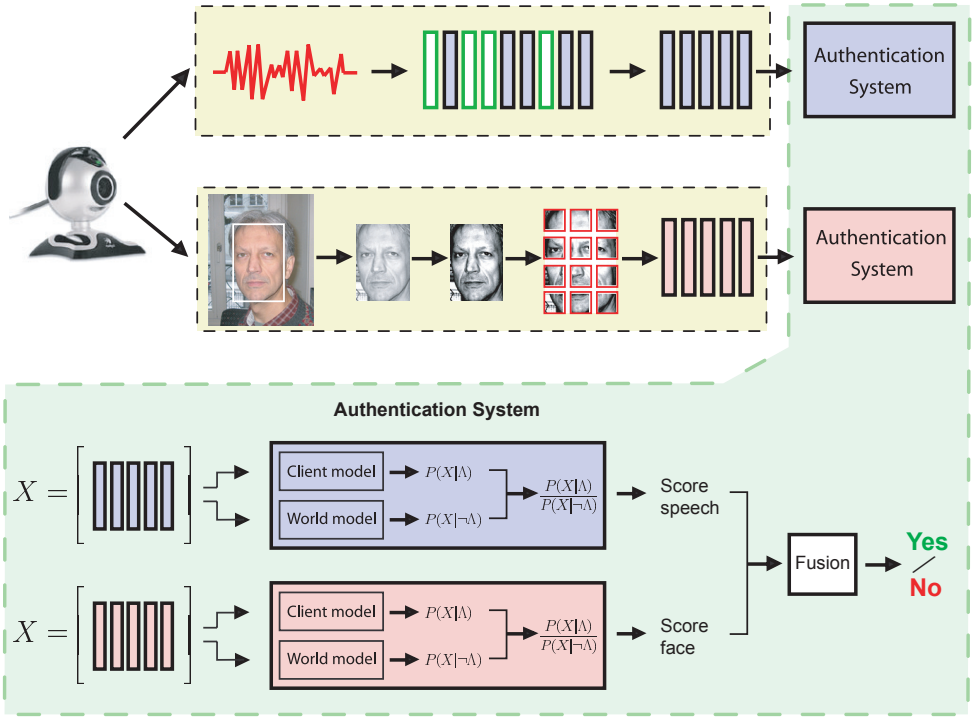
Identity authentication is a task that has many applications such as voice mail, secure teleworking, access control or transaction authentication (in telephone banking or remote credit card purchases for instance).

Biometric identity authentication systems are based on unique personal characteristics, such as a person's face, voice, fingerprint or signature. Identity authentication using face and speech information is a challenging research area that is currently very active, mainly because of its natural and non-intrusive interaction with authentication systems.

An identity authentication system has to deal with two kinds of events: either the person claiming a given identity is the one who he/she claims to be (in which case, he/she is called a **client**), or he/she is not (in which case, he/she is called an **impostor**). Moreover, the system may generally take two decisions: either **accept** the **client** or **reject** him/her and decide he/she is an **impostor**.



1. Face authentication

Most face authentication systems use as features the whole gray-scale face image, or its projection into a linear subspace, together with a discriminant classifier. Such methods are not robust to errors in face localization (imprecision in translation and scale) and their performance is degraded.

Our approach decomposes the image into blocks, computes an appropriate feature space, namely DCTmod2 [2], and models the distribution of these blocks across the image using Gaussian Mixture Models (GMMs) [1]. It has been shown that our method is more robust to errors in face localization than the above methods [3].

IDIAP Research Institute - Rue du Simplon 4 - P.O. Box 592 - CH-1920 Martigny
tel: +41 27 721 77 11 - fax +41 27 721 77 12 - info@idiap.ch - www.idiap.ch

2. Speaker authentication

The goal of a speaker authentication system is to decide whether a given speech utterance has been pronounced by a claimed client or by an impostor. Two different scenarios can be investigated. In text dependent applications, the system checks the lexical content of the authentication utterance while in text independent applications, there is no fixed phrase.

We address the problem of speaker authentication using a technique based on GMMs [1], and Maximum A Posteriori adaptation [3], that does not require any speech recognition techniques. The advantages of GMMs are their low complexity and that they can be used both in text dependent or in text independent applications.

3. Fusion

It has been shown that combining biometric authentication systems enables better performance than techniques based on only one biometric modality. More specifically, audio-visual biometric authentication systems (based on the face and the voice of an individual) have been successful. Fusion algorithms are methods whose goal is to merge the prediction of many algorithms in order to achieve better average performance than any of the individual methods. This fusion can be simple (maximum score, score average, ...), but it is often much better to train a fusion system using Machine Learning algorithms such as Artificial Neural Networks or Support Vector Machines.

4. Application

Our bi-modal (face and speech) authentication includes two applications:

- BioLogin User Manager (BLUM): Creates a new account *(Fig. 1)*,
- BioLogin: Login using your face and your voice *(Fig. 2)*.

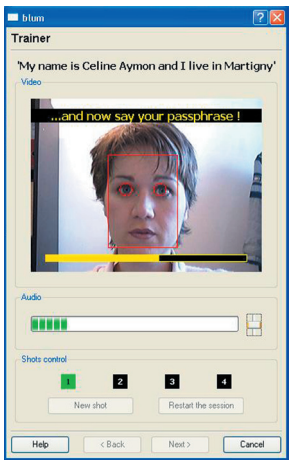


Fig. 1 - BLUM (Enrollment)

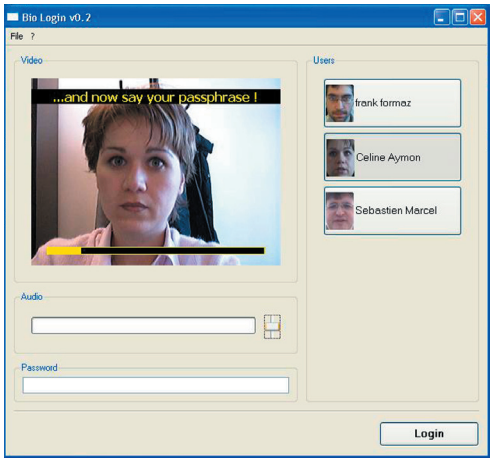


Fig. 2 - Bio Login (Test)

Contact : Dr. Sébastien Marcel - marcel@idiap.ch

References:
[1] Douglas A. Reynolds and Thomas F. Quatieri and Robert B. Dunn, «Speaker Verification Using Adapted Gaussian Mixture Models», Digital Signal Processing, vol 10, pp. 1--3, 2000

[2] C. Sanderson, and K. K. Paliwal, «Fast Features for Face Authentication Under Illumination Direction Changes», Pattern Recognition Letters, vol 24 (14), 2003

[3] F. Cardinaux and C. Sanderson and S. Marcel, «Comparison of MLP and GMM classifiers for face verification on XM2VTS», Proceedings of the International Conference on Audio- and Video-Based Biometric Person Authentication, Springer-Verlag, 2003

[4] J. Mariéthoz, and S. Bengio, «A Comparative Study of Adaptation Methods for Speaker Verification», International Conference on Spoken Language Processing, pp. 581--584, 2002