Mistral : open source biometric platform.

*

Eric Charton
Laboratoire Informatique d'Avignon
339, chemin des Meinajaries
Avignon, France
eric.charton@univ-avignon.fr

Anthony Larcher
Laboratoire Informatique d'Avignon
339, chemin des Meinajaries
Avignon, France
anthony.larcher@univ-avignon.fr

Christophe Levy
Laboratoire Informatique d'Avignon
339, chemin des Meinajaries
Avignon, France
christophe.levy@univ-avignon.fr

Jean-Francois Bonastre†
Laboratoire Informatique d'Avignon
339, chemin des Meinajaries
Avignon, France
jean-francois.bonastre@univ-avignon.fr

ABSTRACT

Mistral is an open source software for biometrics applications. This software, based on the well-known UBM/GMM approach includes also the latest speaker recognition developments such as latent factor analysis, unsupervised adaptation or SVM supervectors. The software performance is highlighted in the framework of the NIST evaluation campaigns.
Several high level applications, like speaker diarization, embedded speaker recognition, language recognition, face recognition and pathological voice assessment are also present.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

General Terms

Biometric platform

*Mistral Project is funded by the French National Research Agency (ANR) - Programme Technologies Logicielles 2006

†Official coordinator of the Mistral Consortium.

SAC'10 March 22-26, 2010, Sierre, Switzerland.

Keywords

GMM, Biometric, Speaker recognition, Face recognition

1. INTRODUCTION

Recent proliferation of computer terminal gives accesses to information in multiple ways. The emergence of personal mobility tools, associated with software applications related to personal data, increases the need of efficient and easy new ways for user identification. According to this, many administrations and industrials express their need to offer to their users or customers new methods and tools for accessing in a comfortable and secure way to such sensitive applications.

Biometrics recognition methods are one of the answers of the scientific community to allow a comfortable and easy way for a user to access to its personal data, applications or computer terminals. The scientific activity is wide and active in the biometrics field of research. It is also evaluated through rigorous and selective international evaluation campaigns like NIST-SRE for speaker recognition [4], FRGC for face recognition [9] or MBGC for face and iris recognition for example [11].

As the scientific competition is becoming intense in the biometric field, an isolated team of researcher would have difficulties to elaborate a biometric tool at a state-of-the-art level. The Mistral project issues from this observation. Mistral is an high quality open-source, free biometric platform (distributed under LGPL licence), supported by an efficient and active scientific community.

The main originality of the Mistral project is to use a unique statistical engine for various modalities of biometric applications. With a unique modeling tool, Mistral is adapted for voice, gender, image and language recognition. Mistral also includes medium level applications like segmentation tools (speech, speaker, ...), and high level applications like friendly graphic user interface.

2. ARCHITECTURE OF MISTRAL

Biometric applications are complex systems, usually difficult to configure and deploy. They use multiple modules, each dedicated to achieve a specific step of the biometric task. A complete biometric application often includes a statistical engine (*e.g.* GMM models) and a wide variety of pre/post processing tools like those dedicated to data preparation and score calculation. This software complexity is difficult to manage for an entrant user, even if this user has an important theorical background.

The Mistral platform offers a complete biometric tool performing at the state-of-the-art level in highly competitive tasks such as international evaluation campaigns like the NIST-SRE [7]. Furthermore, a special attention is paid to the ease of use of Mistral tools.

Work environment

Mistral is a fully configurable scientific tool which can be used through a graphic user interface offering ergonomic facilities.

Each module of Mistral can be used from a command line or processed in a full automated biometric task running from models training to score generation and performance calculation. In this configuration, users still have access to all the configuration parameters (GMM configuration, file format, normalisation, statistical engine parameters) through standardised configuration files.

In the perspective of Mistral use by student or novice users, a Java interface is also available. First visual interface, called *Mistral Config* (*cf.* Figure 1), allows beginners to use a set of pre-defined parameters. This aims to decrease the risk of bad settings and to give a textual explanation of each parameter. This Java interface allows also automatic generation of experiments sets (scripts, folder configuration).

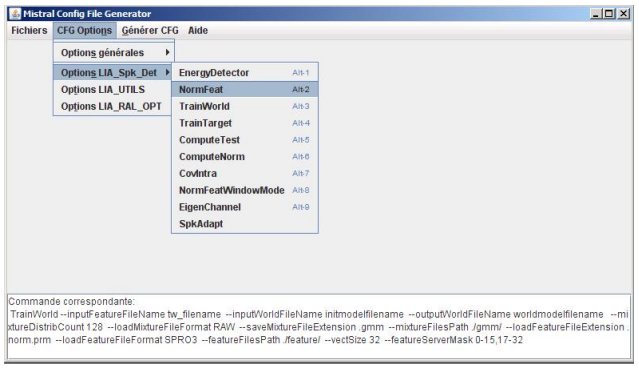

Figure 1: The configuration manager of Mistral

For demonstration purposes and/or student experiments, a second interface called *MistralIHM*, is available. This software is a graphical easy-to-use test tool allowing voice signal acquisition through computer audio interface, automated construction of speakers models and fast test-computation. This interface avoid the need of complex configuration and command line use of Mistral modules.

3. INTERNET PORTAL TO SUPPORT COMMUNITY

All Mistral applications and tools can be accessed through an Internet portal[1] (*cf.* Figure 2). This portal is also used as a dissemination tool and as a collaborative workspace. To federate community of Mistral, this website collects and publishes all the scientific publications and industrial realisations related to Mistral. This portal has a wide visibility with 2000 to 4000 visitors a month.

It includes also a set of engineering software applications to help developers community (UML tools for modelisation, Doxygen for documentation, SVN for dissemination, ...).

Another tool is available for the Mistral community: an open wiki system. It is provided to Mistral users and developers. Moreover, it is used as documentation and support service [2].

All those elements federate a community of users from numerous scientific laboratories, faculties and companies around the world (mostly from Europe, North America, China).

The organisation of the Mistral open-source community follows the organisation of well known existing open-source projects and gives access to a wide collection of useful informations and tools to deploy a biometric application. This includes access to multiple levels of source code (night builds, stable versions), executable and installable code ready to use for unexperienced users. Moreover, multiples plateforms are available: windows, mac-os, linux Ubuntu. A full Doxygen documentation, describing all the biometrics API of Mistral as well as the statistical engine Alize is daily updated.

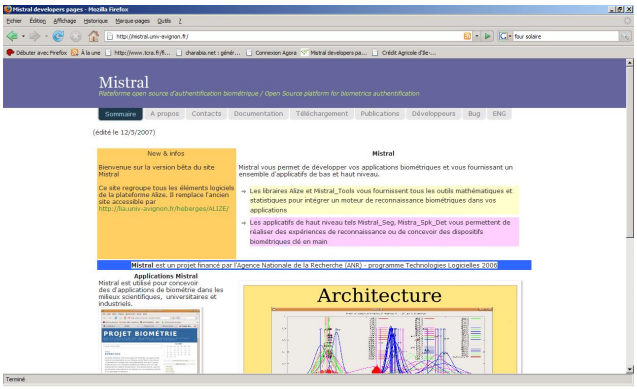

Figure 2: Mistral community is supported by an Internet portal

Subversion Server and code standard

Software architecture of Mistral components is based on UML modeling and strict code conventions. This development strategy allows numerous developers from various universities and companies involved in the code maintenance to share their work through a versioning repository software. The well known Subversion server is used. This server is fully opened for code downloading, allowing all Mistral-users to access to the latest versions of the Mistral code, branch variations and beta testing of new scientific discovery implemented.

[1] `http://mistral.univ-avignon.fr`
[2] `http://mistral.univ-avignon.fr/wiki/`

4. MISTRAL FUNCTIONALITIES

Objective of Mistral is to give high level libraries (Alize and Mistral_tools) and ready-to-use modules, for any user wishing to manage a task of speaker identification and verfification, speaker diarization, face and language recognition. The architecture includes a set of tools, based on two software libraries: the Alize library dedicated to feature and model management, and the Mistral_tools library providing usefull functions mainly for post-processing.

Statistical models of Mistral

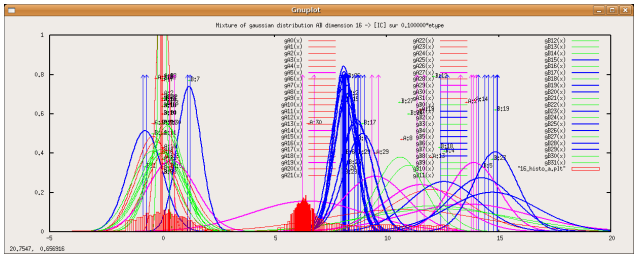

Figure 3: Mistral GMM modelisation

For basic applications, Mistral proposes a complete GMM statistical engine (Gaussian Mixtures) fully configurable by user. The models can then be completed by the two main sets of functionalities included in libraries:

The Alize library allows data reading and synchronization, algorithms (EM/ML/MAP) for models training, score computation, discriminant classifiers, session variability modeling.

The Mistral_tools provides ready to use functions for SVM classification, Latent Factor Analysis, features/scores normalisation.

4.1 Alize library

Alize library provides statistical engine, and feature server management of Mistral. This library is object oriented, programmed in C++ langage and provided under the LGPL licence. The global architecture of Alize divides the various tasks of the biometric modelisation and classification processes in servers, which provide complementary treatments:

The feature server manages access to acoustic or raw features.

The model server manages representations of mixtures and their parameters

The statistical server is in charge of calculations for models training, likelihood estimation, score normalisation. It includes complementary functions like Viterbi decoder or EM estimator.

4.2 Mistral_tools library

The Mistral_tools library includes Support Vector Machines (SVMs), the associated Nuisance Attribute Projection compensation (NAP) and Factor Analysis (FA).

The use of these differents techniques allows to improve significantly the performance of speaker recognition systems.

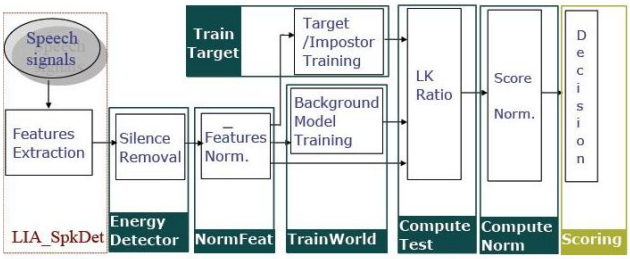

Figure 4: Mistral a complete experience

4.3 Mistral applications

Mistral provides ready to use applications to create a complete biometric experience, based on Alize and Mistral_tools libraries.

Speaker verification - SpkDet

Automatic speaker recognition experiments are achieved with the SpkDet tool. This complete application allows user to manage a complete experiment of speaker recognition, from acoustic features preparation to modelisation and score calculation.

NormFeat is used to normalise frame distribution ([1]).

TrainWorld allows to train a GMM World Model ([12]).

TrainTarget allows to adapt a model from another one. For example, this function is called to train a speaker model from a UBM.

ComputeTest returns a LLR score and is compatible with NIST-SRE evaluation campaign file format.

All those tools give and easy access to configuration options, allowing to select available options from the Alize and Mistral_tools libraries, including SVM configuration, Factor Analysis or NAP.

SpkDet includes some complementary tools like energy Detector (speech and non-speech detection in signal), GMM visualisation, histogram generation, and score fusion.

Segmentation and classification of speakers - hSeg:

Segmentation application, called hSeg, allows segmentation and hierarchical classification of speakers from a speech signal. Signal segments are low enough to allow classification even in difficult conditions, like multiple speakers intervention. The available methods include plain or diagonal Gaussian with various evaluation metrics.

Langage detection - LangDet:

The LangDet application is a set of tools derived from SpkDet. Those tools allow to build a complet language detection system, fully compatible with the NIST-LRE evaluation campaign datas.

Facial biometrics:

Several modules of Mistral are dedicated to facial biometrics. Local features approaches have been shown to be more

robust than holistic ones in terms of pose variations or translations due to face localisation [13]. The use of such parameters in the GMM/UBM paradigm has provided good performance [8]. Local Principal Component Analysis (LPCA), eigenfaces and 2D DCT extraction tools as well as some normalisation tools are part of the Mistral platform and could be easily combined to the statistic tools.

Dealing with imposture detection for audio-video biometrics task, Mistral holds a liveness detection module based on Co-Inertia Analysis [5].

User-customised password speaker recognition:

This module allows speaker recognition task to deal with embedded constraints which could involve resources limitations (memory, computational power) but also speech material quantity limitations due to ergonomic constraints.

A specific three stage hierarchical architecture is created to merge a state of the art GMM/UBM system and a semi-continuous HMM architecture, following the ideas developed in [2] and [6]. This architecture allows to reduce both memory and computational requirements.

Speaker recognition using Client/Server mode

This module allows to use a part of Mistral functionalities using C/S mode. Client connects to the serverthanks to the TCP/IP protocol. Clients are allowed to send speech signal or speech parameters to the dedicated server. They have then to specify the kind of treatment they want to apply to the sent data (model training, speaker recognition, speaker authentication, ...).

5. MISTRAL PERFORMANCES

Mistral performances are evaluated on regular basis throughout national and international scientific evaluation campaigns. During the last 3 years, Mistral has been deployed by Mistral Consortium members in NIST-SRE 2006-08 (speaker detection), ESTER 2 (speaker segmentation).

Speaker verification

Mistral is regularly evaluated with participation to several national/international campaign like NIST-SRE. During the NIST-SRE 2006 campaign, Mistral achieved an EER of 5%, what placed this system close to the state-of-the-art systems (*cf.* [3]). During 2008 evaluations, Mistral was still to the state-of-the-art level (*cf.* [7]).

Segmentation and classification

Mistral had been deployed in the ESTER 2 campaign: speaker diarisation task. The context of the evaluation is the detection of speakers in broadcast news data. The evaluation rules follow those adopted in NIST *Speaker Diarisation*, from the *Rich Transcription* evaluation campaign. The segmentation module of Mistral obtained a 13 % error rate on the test corpus. *cf.* [10])

6. CONCLUSIONS

The aim of Mistral project is to developp a state-of-the-art biometric system. The performance of Mistral plateform is highlighted during several national and international campaigns.

Another main objective of the Mistral platform is to facilitate access to state-of-the-art biometric technologies for academics (research and education) and industrials. To achieve this objective, Mistral project provides an efficient, modular and easy to master platform, able to manage different biometrics samples and tasks in different environments. One of the expected impact of the project is to animate a strong community of users of the platform Mistral composed of academic and industrial players. The major originality of Mistral is to propose a single recognition engine for multiple modalities: primarily voice and face. Mistral is also available on different operating systems (Linux / Windows / Mac OS). For users of Mistral, adoption of a unique and free recognition engine allows to focus on research and development efforts on specific aspects of the studied modality, rather than on the development and monitoring of complex low level software engines. Nevertheless, developers are able to add their own part of code thanks to the LGPL licence used for the distribution of Mistral. With Mistral, it is becomes easier to integrate new methods based on the know-how acquired over those already studied. Furthermore, this approach simplifies the integration of biometrics in demonstrators or products.

7. REFERENCES

[1] F. Bimbot, J.-F. Bonastre, C. Fredouille, G. Gravier, I. Magrin-Chagnolleau, S. Meigner, T. Merlin, J. Ortega-Garcia, D. Petrovska-Delacretaz, and D. A. Reynolds. A tutorial on text-independent speaker verification. 4:430–451, April 2004.

[2] J.-F. Bonastre, P. Morin, and J.-C. Junqua. Gaussian Dynamic Warping (GDW) method applied to text-dependent speaker detection and verification. In *European Conference on Speech Communication and Technology (Eurospeech)*, Geneva (Switzerland), 2003.

[3] J.-F. Bonastre, N. Scheffer, D. Matrouf, C. Fredouille, A. Larcher, A. P. G. Pouchoulin, N. Evans, B. Fauve, and J. S. Mason. Alize/spkdet: a state-of-the-art open source software for speaker recognition. 2008.

[4] G. R. Doddington, M. A. Przybockib, A. F. Martin, and D. A. Reynolds. The NIST speaker recognition evaluation - Overview, Methodology, Systems, Results, Perspective. *Speech communication*, 31(2-3):225–254, 2000.

[5] N. Eveno and L. Besacier. A speaker independent "liveness" test for audio-visual biometrics. In *European Conference on Speech Communication and Technology (Eurospeech)*, Lisboa (Portugale), 2005.

[6] C. Lévy, G. Linares, and J.-F. Bonastre. Fast adaptation of gmm-based compact models. In *International Conference on Speech Communication and Technology*, Antwerpen, August 2007.

[7] D. Matrouf, J.-F. Bonastre, C. Fredouille, A. Larcher, S. Mezaache, M. McLaren, and F. Huenupan. LIA GMM-SVM system description: NIST SRE. In *NIST SRE*, Montreal (Canada), april 2008.

[8] C. McCool and S. Marcel. Parts-Based Face Verification Using Local Frequency Bands. *IEEE IAPR International Conference on Biometrics (ICB)*, page 72, 2009.

[9] NIST. Nist face recognition challenge, 2004-2006. http://face.nist.gov/frgc/.

[10] NIST. Dn4.2: Report on benchmark-based evaluations, 2008. Muscle, European Project no. FP6-507752.

[11] P. J. Phillips, A. Martin, C. Wilson, and M. Przybocki. An introduction to evaluating biometric systems. *Computer*, 33(2):56–63, 2000.

[12] D. A. Reynolds. Speaker identification and verification using Gaussian mixture speaker models. *Speech Communication*, 17(1-2):91–108, 1995.

[13] C. Sanderson, M. Saban, and Y. Gao. On local features for GMM based face verification. In *International Conference on Information Technology and Applications*, pages 638–643, 2005.