HAL open science

Mistral : Plate-forme open source d'authentification biométrique

Eric Charton, Teva Merlin, Christophe Lévy, Anthony Larcher, Sylvain Meignier, Jean-François Bonastre, Laurent Besacier, Jérôme Farinas, Bertrand Ravera

To cite this version:

Eric Charton, Teva Merlin, Christophe Lévy, Anthony Larcher, Sylvain Meignier, et al.. Mistral : Plate-forme open source d'authentification biométrique. XXVIIe Journées d'étude sur la parole (JEP 2008), AFCP, Jun 2008, Avignon, France. pp.1654. hal-01451536

HAL Id: hal-01451536

https://hal.science/hal-01451536v1

Submitted on 22 Mar 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Mistral*: Plate-forme open source d'authentification biométrique

Éric Charton (LIA) *Teva Merlin (LIUM)* *Christophe Lévy (LIA)*
Anthony Larcher (LIA) *Sylvain Meigner (LIUM)* *Jean-François Bonastre (LIA)*
Laurent Besacier (LIG) *Jérôme Farinas (IRIT)* *Bertrand Ravera (Thales)*

{teva.merlin, sylvain.meigner}@lium.univ-lemans.fr
{eric.charton, christophe.levy, anthony.larcher, jean-françois.bonastre}@univ-avignon.fr
laurent.besacier@imag.fr jerome.farinas@irit.fr bertrand.ravera@fr.thalesgroup.com
http ://mistral.univ-avignon.fr

ABSTRACT

Mistral is an open source software for biometrics applications. This software, based on the well-known UBM/GMM approach, also includes the latest developments in speaker recognition such as latent factor analysis and unsupervised adaptation, or SVM supervectors. The software performance is demonstrated within the framework of the NIST evaluation campaigns. Several applications, like speaker diarization, embedded speaker recognition, language recognition, and pathological voice assessment are also presented.
Keywords: reconnaissance du locuteur, reconnaissance de la langue, segmentation, toolkit, application

1. Introduction

La prolifération des terminaux d'accès à l'information ainsi que l'usage croissant d'applications mettant en œuvre des transferts de données personnelles imposent de disposer de moyens d'identification performants.

Le besoin des administrations et des industriels de disposer de méthodes et d'outils fiables pour ces applications sensibles a fait émerger un champ de recherche scientifique dense et actif, encadré par des campagnes d'évaluation internationales rigoureuses et sélectives (telles que celles organisées par le NIST). Cette sélectivité rend aujourd'hui très difficile pour une équipe isolée de concevoir et développer une application biométrique performante et à l'« état de l'art ». De ce constat est né le projet Mistral qui propose d'offrir une plate-forme biométrique open-source libre (proposée sous licence LGPL).

L'originalité majeure du projet Mistral est de mettre en place un moteur de reconnaissance unique pour différentes modalités, utilisable pour des tâches variées de biométrie. Mistral peut traiter aussi bien données audio, que visages ou empreintes, sous de multiples formes (reconnaissance de genre, de langue, segmentation, multimodalité).

2. Architecture globale

Les applications de biométrie sont par nature complexes à déployer et configurer. Elle font appel à des moteurs de reconnaissance statistiques (modèles multigaussiens — GMM) et à de nombreux outils de préparation des données parfois difficiles à appréhender.

2.1. Environnement de travail

L'un des enjeux de Mistral est d'offrir à sa communauté d'utilisateurs un environnement de travail convivial et simple. Cette démarche de simplicité ne doit pas pour autant être mise en œuvre au détriment des utilisateurs plus expérimentés (chercheurs et ingénieurs) qui souhaitent pouvoir agir finement sur le choix des paramètres qu'ils appliquent à leurs expériences et applications.

Pour marier ces deux exigences contradictoires, Mistral est un outil de construction d'expérimentation totalement configurable. Il est utilisable soit en ligne de commande (tous les paramètres statistiques, de normalisation, de choix de formats de données peuvent être réglés), soit via une interface graphique plus conviviale développée en Java (figure 1).

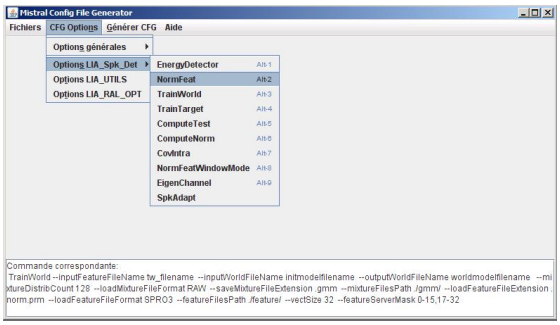

Fig. 1: Le gestionnaire de configuration de Mistral

Cette interface offre à l'utilisateur un ensemble de choix pré-définis pour configurer son système de reconnaissance biométrique et génère automatiquement ses plans d'expériences (scripts, configuration des répertoires, explication des paramètres de configuration).

L'ensemble des ces logiciels est accessible depuis un site internet (`http://mistral.univ-avignon.fr`) utilisé en tant qu'instrument de dissémination, mais aussi comme plate-forme d'ingénierie logicielle. Ce site joue également un rôle fédérateur en accueillant les publications de la communauté de chercheurs qui utilisent Mistral. Cette communauté est actuellement composée de nombreux utilisateurs issus de laboratoires scientifiques ou industriels d'une

*Mistral est un projet financé par l'ANR (programme technologies logicielles 2006)

dizaine de pays.

2.2. Accès aux applications

Ce modèle d'organisation, qui a fait ses preuves dans de nombreux projets open-source, offre l'avantage de regrouper dans un environnement unique développeurs et utilisateurs.

Le portail internet est architecturé selon les standards des grands projets open-source et fournit l'ensemble des informations et des outils indispensables à la mise en œuvre d'une expérience biométrique.

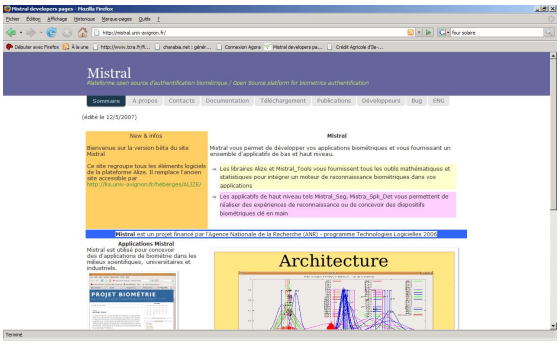

Fig. 2: Le portail collaboratif de la plate-forme Mistral

Le site de Mistral regroupe un ensemble de documentations de haut et bas niveau pour assister les utilisateurs dans leur prise en main de la plate-forme.

Les applications qui composent Mistral sont mises à la disposition du public sous diverses formes afin de s'adapter à différents types d'utilisateurs. Les utilisateurs qui le souhaitent trouveront des versions compilées et prêtes à l'emploi compatibles avec Linux Ubuntu et Microsoft Windows[1].

Pour des utilisateurs plus expérimentés, des versions sources « stables » (c'est-à-dire figées lors du développement et considérées comme fiables) sont proposées. Ces versions donnent la possibilité d'adapter les outils de Mistral à des besoins spécifiques. Elles ont été par exemple mises en œuvre dans le cadre de travaux étudiants[2] ou d'applications spécialisées dans la reconnaissance de pathologies vocales [4].

2.3. Serveur Subversion

Un accès au contenu des dépôts Subversion[3] est proposé, complété par le téléchargement de distributions générées régulièrement[4]. Ces moyens d'accès permettent aux utilisateurs les plus expérimentés de bénéficier des dernières innovations ajoutées à Mistral (résultats de recherches, nouveaux modules, implémentation de nouvelles possibilités en fonction des avancées scientifiques).

[1] `http://mistral.univ-avignon.fr/download.html`

[2] Voir `http://mistral.univ-avignon.fr/applications.html`

[3] Outil logiciel de développement collaboratif utilisé par les contributeurs du projet

[4] Consulter la section réservée aux développeurs à l'adresse `http://mistral.univ-avignon.fr/mistral_dev/`

3. Fonctionnalités de Mistral

Mistral vise à intégrer des applications et des bibliothèques de haut niveau permettant de remplir des tâches comme la reconnaissance du locuteur, du visage et de la langue, ou la segmentation et classification en locuteurs.

3.1. Les modèles statistiques de Mistral

Mistral est composé d'une bibliothèque d'outils et d'un système applicatif complet reposant sur deux bibliothèques logicielles : le système de reconnaissance ALIZÉ et les API avancées du projet Mistral.

Les fonctions de bases proposées sont dédiées à la modélisation des données (locuteur, visage, langue, empreintes) à l'aide de mixtures de gaussiennes (figure 3) entièrement paramétrables. Les modèles GMM peuvent ensuite exploités dans le cadre d'expériences de détection et de classification.

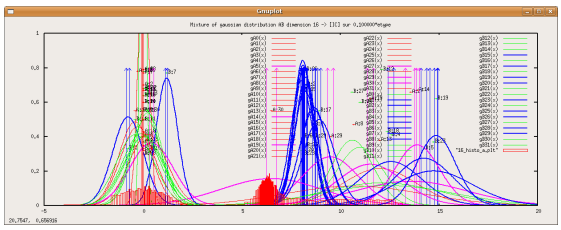

Fig. 3: Modélisation par GMM de Mistral

Au système de base reposant sur les GMM s'ajoutent des API dont le rôle est d'implémenter des fonctions évoluées de classification et de normalisation telles que SVM ou *Symmetrical Latent Factor Analysis* (SFA).

3.2. Aperçu d'ALIZÉ

Le moteur statistique et de gestion des serveurs de données de Mistral est la bibliothèque ALIZÉ. Il s'agit d'une bibliothèque orientée objet, développée en langage C++ et mise à disposition sous licence LGPL. L'architecture générale suit un découpage en plusieurs serveurs proposant des traitements élémentaires. Les serveurs principaux sont :

- le serveur de paramètres acoustiques qui gère l'accès aux données acoustiques ;
- le serveur de modèles (stockage, modification, partage de paramètres, . . .) ;
- le serveur de statistiques qui propose des calculs élémentaires — et le stockage des résultats correspondants — pour l'apprentissage de modèles, ainsi que le calcul de vraisemblance (pour l'évaluation d'un modèle, la normalisation de score ou un décodage Viterbi par exemple).

3.3. Détection et classification

La détection et la classification peuvent être réalisées selon les processus classiques des systèmes de modélisation multigaussienne à l'état de l'art. Il est possible d'appliquer une modélisation GMM pour chaque classe et de confronter un échantillon à ces modèles.

Ce test repose sur les hypothèses suivantes :

- L'hypothèse $H0$ postule que l'échantillon Y appartient au modèle L.
- L'hypothèse $H1$ postule que l'échantillon Y n'appartient pas au modèle L.

Mistral permet de considérer que l'échantillon Y représente toute forme de données (signal de parole numérisé, empreintes digitales ou visages, signal de parole correspondant à un genre ou à une langue). Le test optimal de décision est donné par le rapport entre les deux hypothèses comparées à l'estimateur de vraisemblance :

$$\frac{p(Y|H0)}{p(Y|H1)} \begin{cases} \geq \theta & \text{accepter l'hypothèse H0} \\ < \theta & \text{rejeter l'hypothèse H0} \end{cases}$$

Il est possible d'utiliser le modèle GMM-UBM pour l'hypothèse $H1$. Ce dernier exploite un modèle du monde (*Universal Background Model*) produit d'après des données représentant le modèle universel des données à classer. Il est associé à des modèles GMM pour chaque classe, qui peuvent être adaptés d'après ce modèle du monde [5].

Reconnaissance de locuteur — SpkDet : Les expériences de reconnaissance automatique du locuteur (RAL) sont réalisées avec l'ensemble SpkDet de Mistral. Ce système applicatif clé en main est composé d'un ensemble d'outils logiciels dédiés à chaque étape d'une expérience de RAL. Il est complété par des utilitaires d'analyse de cette chaîne de traitement (génération d'histogrammes, fusion de scores, visualisation de modèles GMM, etc). *EnergyDetector* est un outil de détection de l'existence de « parole » et de « non-parole » dans des trames acoustiques, basé sur le calcul de l'énergie des trames. *NormFeat* est utilisé pour normaliser la distribution des trames. Les applications *TrainWorld* et *TrainTarget* sont utilisées pour modéliser des données et produire une mixture de gaussiennes. *TrainWorld* est une application générique capable d'apprendre un GMM modèle du monde en utilisant l'algorithme EM. *TrainTarget* est utilisé pour modéliser des données en utilisant les méthodes MAP et EM. *ComputeTest* fournit un score LLR pour un segment de test en fonction d'un modèle donné. Ce programme supporte les tests multiples et peut être déployé dans le cadre de campagnes d'évaluation telles que celles organisées par le NIST. L'ensemble de ces outils bénéficie des dernières possibilités de l'état de l'art : les API avancées de Mistral utilisées par SpkDet intègrent les SVM, Factor Analysis, ou encore NAP.

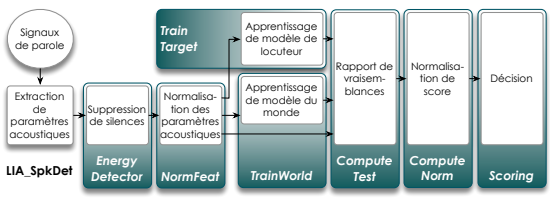

Fig. 4: Expérience de reconnaissance du locuteur avec Mistral

Vérification de la langue — LangDet : L'ensemble LangDet de Mistral est une boîte à outils dérivée de Spkdet, qui permet de réaliser un système complet de reconnaissance automatique de la langue parlée. Il a été testé lors de la campagne d'évaluation internationale NIST LRE 2007[5]. Deux systèmes ont été proposés :

- Un premier basique par l'IRIT : il permet de valider la modélisation acoustique par GMM [3]. Contrairement aux campagnes de vérification du locuteur, le modèle de chaque langue ou dialecte n'est pas appris par adaptation d'un modèle du monde, mais est consitué d'un GMM appris sur les données.
- Un second proposé par l'INRETS[6] utilise les outils de modélisation GMM de Mistral et leur adjoint des fonctions de modélisation prosodique à court et moyen terme [6].

Un outil de segmentation SegForwardBackward [1] a été implémenté dans Mistral et permet de produire une segmentation statistique du signal de parole.

Classification en locuteurs — hSeg : Les outils hSeg permettent de mettre en œuvre un système complet de segmentation et de classification en locuteurs, basé sur un algorithme de classification hiérarchique.

L'outil de segmentation en locuteurs permet de découper le signal en segments de petite taille. Généralement l'objectif est d'obtenir des segments suffisamment courts pour garantir qu'un seul locuteur s'exprime dans chaque segment, tout en étant suffisamment longs pour permettre l'apprentissage sur chaque segment d'un modèle de locuteur robuste.

L'outil de classification a ensuite pour but de regrouper ces segments en classes homogènes, correspondant chacune à un locuteur distinct. Les méthodes disponibles reposent sur des modèles de locuteurs représentés par une gaussienne pleine ou diagonale, utilisant diverses métriques.

4. Performances de Mistral

4.1. Vérification du locuteur

Mistral, dans la continuité du projet ALIZÉ est régulièrement évalué dans le cadre de campagne d'évaluation internationales (NIST). Les performances de Mistral ont été mesurées d'après la base de données NIST SRE 2006 [2]. Comparés aux résultats officiels de NIST SRE 2006, les scores obtenus par la plateforme biométrique Mistral, avec des résultats EER inclus dans la zone des 5 %, la situent au niveau des systèmes à l'état de l'art.

4.2. Segmentation et classification en locuteurs

Les expériences utilisent les émissions radiophoniques issues de la campagne d'évaluation ESTER 2005. La métrique d'évaluation correspond à celle proposée lors de la campagne d'évaluation ESTER et est identique

[5] Descriptif de la campagne : `http://www.nist.gov/speech/tests/lang/2007/`

[6] Laboratoire ayant utilisé la technologie Mistral `http://www.inrets.fr/`

à la métrique du NIST pour la tâche de *speaker diarization* des campagnes *Rich Transcription*. Le système de segmentation et de classification en locuteurs obtient 13 % d'erreur sur le corpus de test.

5. Applications de Mistral

Reconnaissance du locuteur en mode Client/Serveur — DistantServer : Ce module permet d'utiliser une partie des fonctionnalités de Mistral en mode client/serveur. Le client se connecte au serveur en utilisant une liaison TCP/IP. Un client peut ainsi envoyer du signal audio ou des paramètres acoustiques (qu'il a déjà extraits) à un serveur dédié. Le client indique ensuite au serveur le type de traitement qu'il veut y appliquer[7] (apprentissage de modèle, reconnaissance du locuteur, identification du locuteur, ...).

Démonstrateur Thales Communications (TCF) : Dans le cadre du projet Mistral, TCF est en charge du portage et de l'intégration sur cible mobile de modules de biométrie vocale. Le démonstrateur actuel, basé sur une plate forme UMPC (*Ultra Mobile PC*), est un outil de valorisation industrielle du projet qui démontre la simplicité d'utilisation de la plateforme Mistral.

Démonstrateur Calistel : Calistel est une société travaillant dans le domaine des serveurs vocaux interactifs. Dans le cadre du projet Mistral, un système de détection de genre a été développé avec la plateforme sur des données réelles d'exploitation. Les données sont celles du service audiotel 3635 de la SNCF. L'exploration d'autres applications de la plate-forme Mistral aux serveurs vocaux interactifs est prévue : reconnaissance automatique de mots clés courts par GMM et routage d'appels suivant l'affect du locuteur.

6. Enseignement

6.1. Université du Maine

Un projet, réalisé par 4 étudiants de licence professionnelle Système Réseaux et Informatique, consiste à mettre en place un démonstrateur de serrure vocale pour la gestion de l'accès à des salles de cours. La partie reconnaissance du locuteur est effectuée utilisant le client/serveur distant disponible dans Mistral. L'interface Homme-Machine est hébergée sur un système embarqué ARCOM intégrant un écran tactile et une carte son.

6.2. Université Joseph Fourier (Grenoble)

La plate-forme Mistral est utilisée par les étudiants ingénieurs de Polytech' Grenoble (troisième année) qui suivent un cours de 12h intitulé « Tatouage et Biométrie ». Un TP sur 4 séances d'1h30 leur permet de manipuler des outils de modélisation multigaussienne pour la détection de genre et la reconnaissance automatique du locuteur[8].

6.3. Université d'Avignon

L'IUP de l'université d'Avignon utilise régulièrement Mistral en tant qu'outil support pour des travaux pratiques ou dirigés. Chaque année, des étudiants de Master 1 ou 2 réalisent leur projets avec Mistral et ALIZÉ. En 2008, un groupe de 4 étudiants du master (première année) informatique d'Avignon développe un projet de robotique intégrant la reconnaissance de la parole[9] et la reconnaissance du locuteur afin de commander un robot par la voix.

7. Conclusion

En mutualisant le développement d'un instrument scientifique de haut niveau, accessible grâce à la licence LGPL, le projet Mistral permet de maintenir et de développer une communauté de chercheurs et d'industriels intéressés par la biométrique.

Il aide à lever des verrous scientifiques par la coopération de chercheurs, et simplifie la participation des équipes à des campagnes internationales d'évaluation. Il favorise la transmission à des étudiants des avancées scientifiques les plus récentes dans le domaine de la biométrie.

À ce titre il aide les chercheurs et les enseignants à mobiliser leurs ressources disponibles pour leur activité principale. Mistral décharge ainsi ses utilisateurs des coûteuses étapes de mise au point d'un outil d'expérimentation, tout en leur permettant, par son ouverture, d'intégrer librement leurs propres recherches à la plate-forme.

Références

[1] R. André-Obrecht. A new statistical approach for automatic speech segmentation. *IEEE Transactions on Acoustic, Speech and Signal Processing*, 36(1) :29–40, janvier 1988.

[2] J.-F. Bonastre, N. Scheffer, D. Matrouf, C. Fredouille, A. Larcher, A. Preti, G. Pouchoulin, N. Evans, B. Fauve, and J. S. Mason. ALIZE/SpkDet : a state-of-the-art open source software for speaker recognition. In *Odyssey 2008 — The Speaker and Language Recognition Workshop*, 2008.

[3] F. Pellegrino. *Une approche phonétique en identification automatique des langues : la modélisation acoustique des systèmes vocaliques*. Thèse de doctorat, Université Paul Sabatier, Toulouse, France, décembre 1998.

[4] G. Pouchoulin, C. Fredouille, J.-F. Bonastre, A. Ghio, and A. Giovanni. Frequency study for the characterization of the dysphonic voices. In *Interspeech*, Antwerp, Belgium, septembre 2007.

[5] D.A. Reynolds, T.F. Quatieri, and R.B. Dunn. Speaker verification using adapted gaussian mixture models. *Digital Signal Processing (DSP), a review journal – Special issue on NIST 1999 speaker recognition workshop*, 10(1-3), janvier/avril/juillet 2000.

[6] J.-L. Rouas. *Caractérisation et identification automatique des langues*. Thèse de doctorat, Université Paul Sabatier, Toulouse, France, mars 2005.

[7] Une documentation détaillée de *DistantServer* est disponible sur `http://mistral.univ-avignon.fr/mistral_dev/pdf/mistral_standard.001.distantIP.pdf`

[8] La page web du TP est accessible à l'adresse : `http://www-clips.imag.fr/geod/User/laurent.besacier/NEW-TPs/TP-Biometrie/`

[9] Basé sur lia_stok — `http://lia.univ-avignon.fr/fileadmin/documents/Users/Intranet/chercheurs/linares/Softs/lia_stok.tgz`