

A New Approach for Speech Feature Extraction Based on Genetic Algorithms

C. Charbuillet, B. Gas, M. Chetouani, J. L. Zarader

**Laboratoire des Instruments et Systèmes d'Ile-de-France
Université Pierre et Marie Curie, Paris, FRANCE**

christophe.charbuillet@lis.jussieu.fr gas@ccr.jussieu.fr
mohamed.chetouani@lis.jussieu.fr zarader@ccr.jussieu.fr

Feature extraction plays a major role in a pattern recognition system. Its aim is to extract useful information in order to increase the classifiers' performances for a given task. Current speech feature extraction methods rely mainly on properties of speech production (LPC, LPCC) and perception (MFCC, PLP). State of the art methods in speech feature extraction don't take into consideration the specific information about the task to accomplish. However this field has recently been investigated using LVQ and MLP based nonlinear predictive methods [1] and giving interesting results.

We propose in this article to use genetic algorithms (GAs) to design a new feature extraction method adapted to a speaker recognition system.

Genetic algorithms were first proposed by Holland in 1975 [2] and have become widely used in various disciplines as a new mean of complex systems optimization. The basic idea of GAs is based on "natural selection", the principle of "survival of the fittest". A GA operates on a population of chromosomes, each generating a potential solution to the studied problem.

GAs most attractive quality is certainly their aptitude to avoid local minima. However, our study relies on another quality which is the capacity to optimize a complex system without needing any knowledge about it. This allows us to realize a feedback between the recognition system's output and the feature extractor to optimize.

In this article we present an application of our algorithm to the adaptation of the common MFCC extractor to the speaker clustering task. This adaptation consists in designing a filter bank, with optimized center frequencies and band-widths. In order to evaluate the performances of the developed algorithm, we test it on a sub-database from the ESTER [3] corpus of broadcasted radio emissions: two hours for the evolution stage and eight hours for the testing stage. Results showed that the obtained filter bank gives significant improvements compared to the Mel one, reducing the error rate from 7.68% to 5.52%.

More details about the experiments will be given in the final article and results will be discussed.

- [1] *"Non-linear Speech Feature Extraction for Phoneme Classification and Speaker Recognition"* M. Chetouani, M. Faundez, B. Gas and J.L. Zarader. in "Nonlinear speech processing : Algorithms and Analysis". Eds. G. Chollet, A. Esposito, M. Faundez, M. Marinaro. Springer Verlag (2005).
- [2] *"Adaptation in Natural and Artificial Systems"* J. H. Holland, University of Michigan Press, 1975, Ann Arbor.
- [3] *"The Ester Phase II Evaluation Campaign for the Rich Transcription of French Broadcast News"* S. Galliano, E. Geoffrois, D. Mostefa, K. Choukri, J.F. Bonastre, G. Gravier , Proc.Eurospeech/Interspeech, 2005.