

**“ANALYSIS AND RECOGNITION OF AUDIO ORIENTED TO MUSIC APPLICATIONS”**

LUÍS GUSTAVO MARTINS<sup>1</sup> - FEUP<sup>2</sup> – INESC PORTO<sup>3</sup>

---

The advances that took place during the last decade in information technologies and the resultant profound implications in the way multimedia content is produced, shared and accessed on worldwide networks, have created great new challenges concerning the search and access to audio and music contents.

In order to search, access and process such media-rich contents it would be important to know some kind of descriptive information about the data itself, in a level that allows users to manage and process audio and music data in a semantic oriented, natural and efficient way.

Consequently, and in consonance with the recent MPEG-7 standardizing activities in the area of multimedia description formats, great interest has been put in the development of automatic systems capable of extracting semantically meaningful information from multimedia contents (*CBID – Content Based Identification Systems*), which can then be used for the semantic classification, retrieval and processing of audio and musical contents.

---

This PhD research proposal addresses the study, development and implementation of advanced solutions for the analysis and recognition of audio signals, with emphasis on applications for the semantic analysis of musical contents.

The research work comprises the study and development of new methodologies for the analysis of audio signals, such as the extraction of semantically relevant features and descriptors, and the implementation of robust classifiers for the automatic classification and description of audio contents.

Digital Signal Processing, Pattern Recognition and Machine Learning are the base disciplines of such a work, but the results from related research areas such as Psychoacoustics, Computational Auditory Scene Analysis, Music and even Musicology already present an increasingly importance in the attempt to understand and model the human sound and music perception mechanisms.

This PhD research proposal takes as a starting point the previous work in the area of multi-pitch estimation, developed during the MSc activities, as well as some previous developments of tools for the automatic analysis, segmentation and classification of audio signals (more detailed information about such results can be consulted at <http://utm.inescporto.pt/~lmartins>).

---

<sup>1</sup> [lmartins@inescporto.pt](mailto:lmartins@inescporto.pt) – <http://utm.inescporto.pt/~lmartins>

<sup>2</sup> <http://www.fe.up.pt>

<sup>3</sup> <http://www.inescporto.pt>

This PhD work is also in close relation with the activities in the *VISNET FP6 Network of Excellence* (<http://www.visnet-noe.org>), specifically in what concerns the work packages related to audio, speech and multimodal analysis (*WP4.1* and *WP4.3*).

This research proposal is part of a PhD programme at the School of Engineering of University of Porto (FEUP – <http://www.fe.up.pt>), and is being developed at INESC Porto (<http://www.inescporto.pt>), under the supervision of Prof. Aníbal Ferreira (<http://www.fe.up.pt/~ajf>). It officially started on December 2003 and has a foreseen duration of 3 to 4 years. This PhD program is funded by the *Fundação para a Ciência e Tecnologia* (FCT - <http://www.fct.mces.pt>).

Porto, 27th May 2004.