DEPARTAMENTO DE INFORMÁTICA

Computational Audition

Multimodal Systems Team

CHi CENTER FOR INFORMATICS AND INFORMATION TECHNOLOGIES

Sofia Cavaco

Prof. Auxiliar

Ph.D. in Computer Science, Carnegie Mellon University

M.Sc. In Computer Science, Carnegie Mellon University

M.Sc. In Artificial Intelligence, University of Edinburgh

http://ctp.di.fct.unl.pt/~sc

Objectives

My research focuses on the analysis of natural and musical sounds (Fig. 1). We develop methods to model, classify and synthesize the sound structure of these sounds. The goal of this research is to identifying the intrinsic acoustic dimensions that govern the structure of the sounds. I am interested in identifying which of these dimensions are relevant to the auditory perception of natural sounds. The identification of such dimensions provides information about the features that must be preserved in audio compression, as well as about the features that must be manipulated in audio synthesis.

In addition, we use these methods to develop software applications, which include tools and educational games for the blind, and tools for multimedia indexing.

Fig. 1 – Waveform.

Methodology

We use audio signal processing techniques, data-driven methods and multivariate decomposition techniques to learn the intrinsic structures that characterize the sounds (Fig. 3).

Fig. 2 – SonarX – from color to sound.

Current Work

Video annotation with audiovisual information – we are developing methods that:
• combine audio and visual information to annotate video (work developed in collaboration with Duvideo, a multimedia content provider company), and
• that extract information from the transients in the sounds to identify and recognize sound events (work developed in collaboration with LabROSA, Columbia University).

Unveiling the world of color for the blind – we are developing:
• a software tool which converts color information from images and video into sound (Fig. 2) and
• educative games for blind children.
(Work developed in collaboration with CESEM.FCSH, Buffalo State College and Biblioteca Nacional de Portugal).

Sound synthesis – we are exploring statistical methods for modeling and synthesizing sounds with both sinusoidal and attack transient components.

Sound classification – which includes the analysis and classification of environmental sounds, classification of music genres (Fig. 4), modeling and classification of harmonic and percussion instruments, modeling and classification of transients.
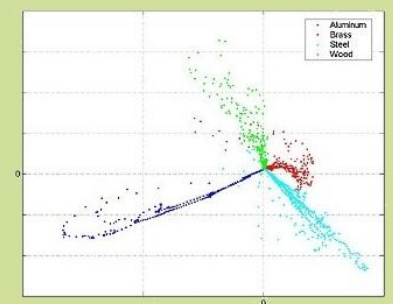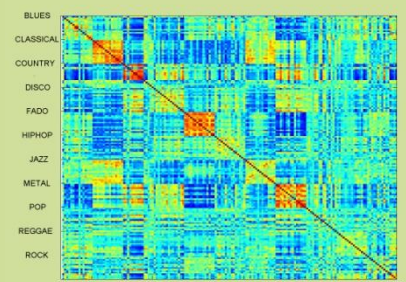

Fig. 3 – Principal components of impacts on rods.

Fig. 4 – Similarity matrix for music genre.