MIREX 2011: AUDIO KEY DETECTION SYSTEM WITH STATISTICAL KEY PROFILES

Cristina de la Bandera, Lorenzo J. Tardón, Isabel Barbancho and Simone Sammartino.
Dept. Ingeniería de Comunicaciones, E.T.S. Ingeniería de Telecomunicación,
Universidad de Málaga, Campus Universitario de Teatinos s/n, 29071, Málaga, Spain
cdelabandera@ic.uma.es

ABSTRACT

This extended abstract details a submission to the Music Information Retrieval Evaluation eXchange (MIREX) 2011 for the Audio Key Detection task. The algorithm named "cbmirex2011_pdfs_modes" is a novel method that performs the key and mode detection. The main innovation is the use of two set of probability density functions (PDFs), computed by the experimental data, to generate the statistical profiles major and minor.

1. INTRODUCTION

The goal of this task is the recognition of the key from polyphonic music in audio format. There is a lot of research done in this area because the key of a music piece is a very important aspect to take into account in any analysis of tonal music: musical form analysis, structural segmentation, pitch detection, automatic transcription, chrod recognition, automatic harmonization, etc.

In recent years, many researchers have taken as a starting point the key profiles defined by Krumhansl *et al.* [6] and Temperley [7] to develop new tonal models and to optimize the key profiles. In [4], Izmirli compared the key estimation results obtained by analyzing spectral or chroma representation. He used three different chroma profiles, the ones by Krumhansl, the ones by Temperley and the flat diatonic profiles. Later, Izmirli presented in [5] two models for key detection, the first one was based on the template based correlation and the second one used a low-dimensional tonal representation.

Gómez [3] developed other similar method for key recognition based on a modified Pitch Class Profile (PCP) called *Harmonic Pitch Class Profile* (HPCP).The PCP was originally proposed by Fujishima [2]. In her method, Gómez studied machine learning methods for key estimation using HPCP. Due to the difficult task of key detection, other works were focused on audio tonality mode classification [1].

2. DESCRIPTION OF THE SYSTEM

The complete system developed can be divided into two stages. In the first stage, two sets of 12 probability density functions (PDFs) are calculated and their modes are used to replace the common key profiles. In the second stage, a music piece is analyzed and the main key is detected. In Figure 1, a general flowchart of the system is shown.

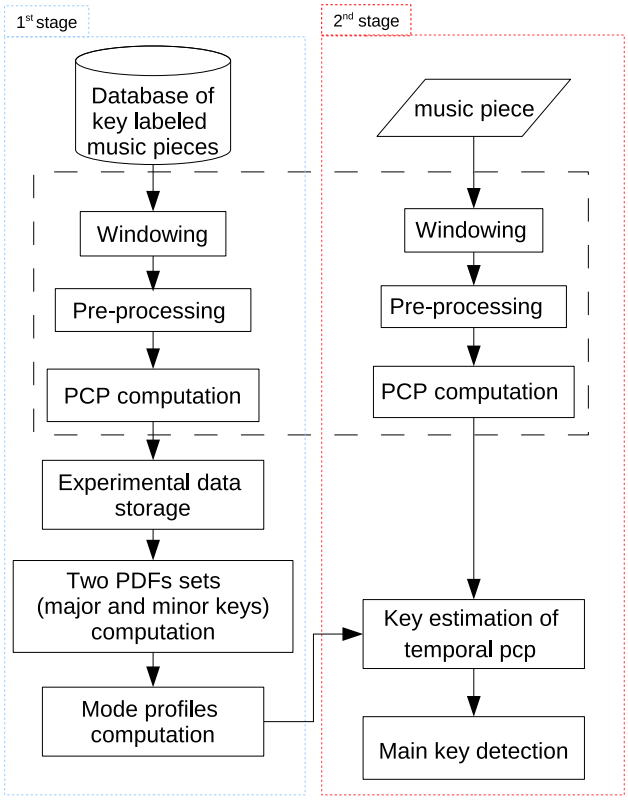

Figure 1. Flowchart of the of the key detection method developed.

The first stage is carried out only once, to fit the two set of PDFs to the experimental data extracted to the polyphonic music database and compute the statistical key profiles (Section 2.1). The second stage performs the key detection employing the profiles extracted from the PDFs (Section 2.2). This second stage is the algorithm called "cbmirex2011_pdfs_modes" and submitted to the "Audio key detection" task.

2.1 Key profiles computation

The first requirement is to have a database music pieces labeled with the corresponding key. These music pieces are pre-processed before the PCP computation. The first 30 seconds of the audio signal are cutted, in order to avoid possible modulations, then the windowing process starts and the spectrum is calculated and analyzed obtaining the PCP vector.

The key mode (major or minor) of the musical piece is extracted from the key label and determines if the PCP is stored in the PCP matrix of the major or minor keys. When the whole database has been analyzed, the data in each pitch class (each row of the matrix) are statistically analyzed in order to fit a probability density function (PDF) according to the specific data.

This process is carried out twice, for the major and minor keys, giving rise to two sets of 12 PDFs.. The mode of each PDF is calculated,to represent the contribution of each pitch class to the major and the minor keys. The results are two 12-dimensions key profiles.

2.2 Key detection

This section describes the key detection procedure submitted to the "Audio key detection" task in MIREX 2011. The main steps are represented in the $2^{nd} stage$ area of the Figure 1. The previous steps to calculate the PCP vector are the same as in the first stage (Section 2.1), windowing and preprocessing.

For each temporal window, the euclidean distance between the PCP of the temporal window and the key profiles is calculated. The PCP vector is shifted 12 times, to calculate the distances for the 12 major keys and the 12 minor keys, obtaining a 24-dimensional distance vector.

The key that give the minimum distance is stored in a vector and its mode is selected as the predominant key of the music piece.

3. REFERENCES

[1] Z. Duan, L. Lu and C. Zhang: "Audio tonality mode classification without tonic annotations" *Multimedia and Expo, 2008 IEEE International Conference on*, pp. 1361–1364, 2008.

[2] T. Fujishima : "Realtime Chord Recognition of Musical Sound : a System Using Common Lisp Music" *Proceedings of the International Computer Music Conference (ICMC)*, pp. 464–467, 1999.

[3] E. Gómez and P. Herrera: "Estimating The Tonality Of Polyphonic Audio Files: Cognitive Versus Machine Learning Modelling Strategies" *Proceedings of the 5th International Society for Music Information Retrieval Conference (ISMIR)*, Barcelona, Spain, 2004.

[4] Ö. Izmirli: "Template Base Key Finding From Audio" *Proceedings of the International Computer Music Conference (ICMC)*, pp. 211–214, Barcelona, Spain, 2005.

[5] Ö. Izmirli: "Audio Key Finding Using Low-Dimensional Spaces" *Proceedings of the 7th International Society for Music Information Retrieval Conference (ISMIR)*, pp. 127–132, Victoria, Canada, 2006.

[6] C. Krumhansl, E. J. Kessler : "Tracing the Dynamic Changes in Perceived Tonal Organization in a Spatial Representation of Musical Keys" *Psychological Review*, Vol. 89, No. 4, pp. 334–368, 1982.

[7] D. Temperley : "The Cognition of Basic Musical Structures" *The MIT Press*, Cambridge, 2001.