# KEY DETECTION EVALUATION BY USING VARIABLE WINDOW CASTELLS-GALIN TRANSFORM METHOD (VCGT)

**David Castells-Rufas**
Universitat Autonoma de Barcelona
david.castells@uab.cat

**Adria Galin**
Universitat Autonoma de Barcelona
adria.galin@e-campus.uab.cat

## ABSTRACT

In this submission we present a method for automatically estimating the key of a musical piece in a digital audio format.

The proposed method relies on three main aspects, first, the generation of a spectrogram using the VCG method; second, the generation of chroma features at different conditions. Finally, the classification of the chroma features using 3 different methods and their combination with some heuristics to produce a key estimation.

## 1. INTRODUCTION

This work represents some progress in the scope of the previous efforts from the members of the team to implement an Automated Piano Music Transcription System [1] [2] [3]. Some of the technical decisions must be understood in this context.

The present contribution corresponds to the very first contribution on the MIREX contest by the authors.

The automated Key Detection Algorithm has been developed and implemented in Java.

## 2. PROPOSED ALGORITHM

The proposed algorithm is the combination of different techniques as shown in the following diagram (see Figure 1).

First, a Spectrogram is generated using a technique similar to the Constant Q Transform. We select 97 frequency bins with a logarithmic spacing starting with the A0 frequency and covering all the standard piano frequencies up to A8.

For every frequency we compute a cross-correlation matching filter. This is similar to what is done on a DFT or FFT. The difference here is that we do it for logarithmically spaced frequencies instead of equally spaced frequencies.

The other difference is that we do not use the same integration window for all the frequencies. Instead, we use an integration window which is multiple of the period of the frequency of the filter.

With this approach, the different filters would produce signals at different sampling rates. To unify the results in the same sampling rate we upsample the resulting signal to the same sampling rate equal as the input signal.

We name this method as VCGT (Variable Window Castells-Galin Transform), which is a flavor of constant Q transform.

The resulting spectrogram is further processed to obtain chroma descriptions and key detection. Some assumptions are made to improve the detection rate, especially in Piano and classical music.

1. We expect that the key will be dominant at the beginning and the end of the music themes.

2. We expect the lower frequencies to have more key information, and higher frequencies to have more frequencies harmonics, pass-notes, trills and mordents which could not be part of the key.

With these assumptions we provide some spectrogram filtering at both the time and frequency domain.

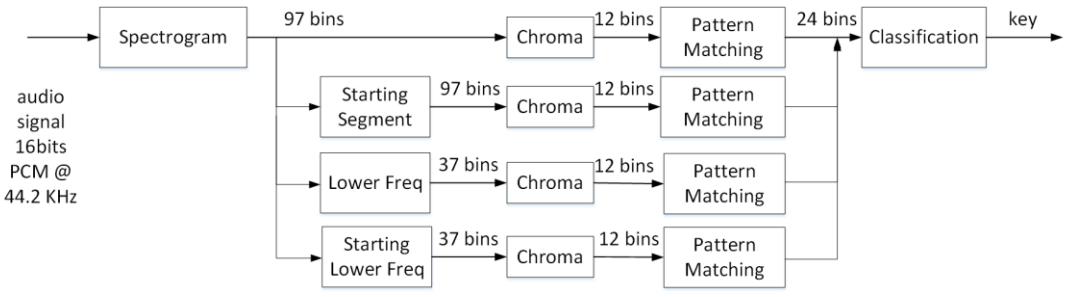


Figure 1 Logic diagram of the key detection process

As a result not a single chroma vector is produce but four. Moreover, several post-processing algorithms are used on resulting chroma vectors. The goal is to test different approaches and obtaining a confidence value for everyone.

We use three approaches:

1. Template based chroma analysis, similarly to the method indicated in [5] by Temperley, D.
2. Chord Recognition
3. Scale Matching (with some similarities to [6]).

In Chord Recognition we only look for the best chroma elements that match all possible major and minor chords.

In Scale Matching, we first try to identify the key signature and later decide between their major and minor possible keys.

Finally all results are combined and some heuristics are used to give an estimation of the present key.

## 3. RESULTS

The preliminary results in classical music are above 90%, achieving a 89.8 MIREX score from 96 pieces.

The observed results in electronic dance music from GiantSteps database [7] is much lower, around 30%. This can be explained by the low harmonic information present in the first 30 seconds of many pieces of the database. Many of them start with a long percusive loop and audio effects, having no melodic lines or chord progressions.

## 4. REFERENCES


[1] Galin, A. and Castells-Rufas, D. “Design, Implementation and Validation of a Novel Real-Time Automatic Piano Transcription System" Universitat Autonoma de Barcelona, 2017.

[2] Galin, A. and Morell-Perez, A. “Automatic Music Transcription" Universitat Autonoma de Barcelona, 2015.

[3] Galin, A. and Zennou, E. “Instantaneous Frequency Estimation Analysis" Institute Superieur de l'Aeronautique et de l'Espace, 2014.

[4] Gomez-Gutierrez, E. “Tonal Description of Music Audio Signals" Universitat Pompeu Fabra, 2006.

[5] Temperley, D. "Whats Key for Key? The Krumhansl-Schuckler Key-Finding Algorithm Reconsidered" Music Perception: An interdisciplinary Journal, 1999.

[6] Zhu, Yongwei, Mohan S. Kankanhalli, and Sheng Gao. "Music key detection for musical audio." *Multimedia Modelling Conference, 2005. MMM 2005. Proceedings of the 11th International*. IEEE, 2005.

[7] Knees, Peter, et al. "Two Data Sets for Tempo Estimation and Key Detection in Electronic Dance Music Annotated from User Corrections." *ISMIR*. 2015.