KEY DETECTION EVALUATION BY USING CASTELLS - GALIN TRANSFORM (CGT)

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

David Castells-Rufas
Universitat Autonoma de Barcelona
David.Castells@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 an spectrogram using the VCG method; second, the generation of chroma features at different conditions. Finally, the classification of the chroma features.for key extraction.

1. INTRODUCTION

The Automated Key Detection Algorithm was developed following the research work on Instantaneous Frequency Estimation, submitted on Institute Supérieur de l'Aéronautique et de l'Espace (ISAE-SUPAERO) [3], followed by the improvements on the publications at Universitat Autonoma de Barcelona [1] [2] by the authors.

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

Automated Key Detection Algorithm has been developed and implemented with a MATLAB interface plugin for general use.

2. PROPOSED ALGORITHM

The proposed algorithm is the combination of different techniques.

First, a Spectrogram is generated using a technique similar to the Constant Q Transform. We select 97 frequency bins with 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 the FFT does, the difference here is that we do it for non-equally spaced frequencies, but logarithmically spaced frequencies.

The other difference is that we do use the same integration window for all the frequencies. Here is where Castells-Galin Trasform (CGT) takes place instead of the Variable Windown Castells-Galin Transform (VCGT) which is presented in a separated input. The difference between CGT and VCGT is that the former is based on a constant integrating window, so that the resulting transformed signal has a lower sampling rate which is the same for all the frequency bins.

The proposed algorithm is based on the Castells-Galin Transform (CGT), the presented transform aims to obtain a representation of the chroma from time domain input signal.

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

Once the Chroma is obtained from a given input signal, several post-processing calculations are performed in order to identify the proper key including SAD calculation similarly than the indicated in [5] by Temperley, D. as well as chord recognition and scale matching.

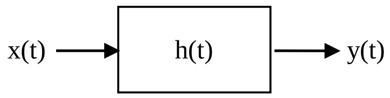
With the presented information Key estimation can be obtained.

3. INSTRUMENT CHARACTERIZATION

In order to be able to estimate with confidence the key of a given input signal, first it is required to detect with as much accuracy as possible the present notes in the input signal itself.

From this very reason we can take advantage of the input signal to try to characterize the instrument(s) being present.

From the submitted research work it is stated that there is a clear relationship relating the played notes and the generated output signal, which can be modeled as a filter given by:



where x(t) corresponds to the played notes, h(t) is the response of the given instrument(s) and y(t) corresponds to the overall time domain characteristics of the generated sound which corresponds to the input signal.

The relationship between x(t) and y(t) is as following:

y(t) = x(t)*h(t) = int x(t)·h(t-tau) dtau

Solving the above convolution equation can be time consuming, for this very reason one can switch to the frequency domain, where:

Y(f) = X(f)·H(f)

Now the relationship between y(t) and Y(f) is given by the Fourier transform:

Y(f) = int x(t) e^-i2·pi·ft dt

And same relationship applies to convert to h(t) and x(t) to H(f) and X(f) respectively. At this stage we can easily obtain X(f) as:

X(f) = Y(f)/H(f)

By mens of Inverse fourier transform it can be obtained the desired x(t) which corresponds to the played noted by the musician.

x(t)= int X(f) e^i2·pift df = Y(f)/H(f) e^i2·pift df

= (int y(t) e^-i2·pi·ftdt/int h(t) e^-i2·pi·ft dt) e^i2·pift df

4. AUDIO KEY FINDING

Once the input signal is analyzed, the extraction of the Chrome vector space from the audio signal is performed, which is related to its pitch class distribution. These features are averaged over a given segment and compared to a tonal model in order to find the key of the piece.

5. 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 11^{th} International. IEEE, 2005.