MIREX CHORD RECOGNITION SYSTEM
SYSTEM 1 : MAJOR AND MINOR CHORDS

Laurent Oudre[1]**, Yves Grenier**[1]**, Cédric Févotte**[2]
[1]Institut TELECOM ; TELECOM ParisTech ; CNRS LTCI
[2]CNRS LTCI ; TELECOM ParisTech
37-39 rue Dareau, 75014 Paris, France
{oudre,grenier,fevotte}@telecom-paristech.fr

ABSTRACT

This paper describes a fast and efficient template-based chord recognition method. We introduce three chord models taking into account one or more harmonics for the notes of the chord. After extracting a chromagram from the signal, the detected chord over a frame is the one minimizing a measure of fit between the chromagram frame and the chord templates. Several popular measures in the probability and signal processing field are considered for our task. In order to take into account the time persistence, we perform a post-processing filtering over the recognition criteria.

1. GENERAL IDEA

Given N successive chroma vectors $\{\mathbf{c}_n\}_n$, K chord templates $\{\mathbf{p}_k\}_k$ and a measure of fit D, we define :

$$d_{k,n} = D\left(h_{k,n}\,\mathbf{c}_n;\mathbf{p}_k\right). \tag{1}$$

$h_{k,n}$ is a scale parameter whose role is to fit the chroma vector $\mathbf{c}_n$ with the chord template $\mathbf{p}_k$ according to the measure of fit used. In practice, $h_{k,n}$ is calculated such as :

$$h_{k,n} = \operatorname*{argmin}_h D\left(h\,\mathbf{c}_n;\mathbf{p}_k\right). \tag{2}$$

The detected chord $\hat{k}_n$ for frame n is then the one minimizing the set $\{d_{k,n}\}_k$:

$$\hat{k}_n = \operatorname*{argmin}_k \{d_{k,n}\}. \tag{3}$$

In our system, the chroma vectors are calculated from the music signal with the same method as Bello & Pickens [1]. The frame length is set to 753 ms and the hop size is set to 93 ms. We use the code kindly provided by these authors.

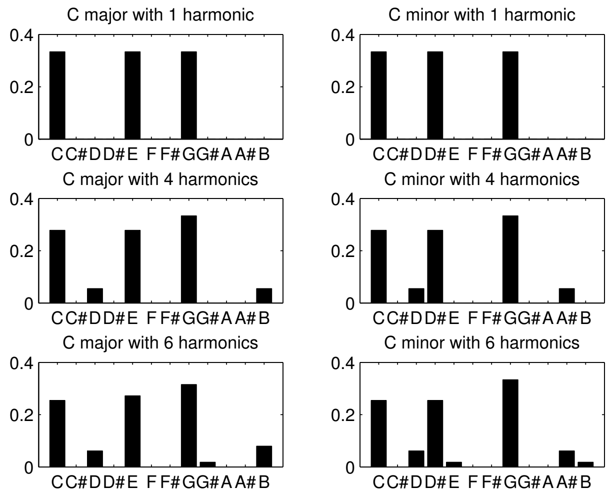

Figure 1. Chord templates for C major / C minor with 1, 4 or 6 harmonics.

2. CHORD MODELS

The intuitive chord model is a simple binary mask constituted of 1's for the chromas present in the chord and 0's for the other chromas [2], [3].

Yet, the information contained in a chromagram captures not only the intensity of every note but a blend of intensities for the harmonics of every note. Like Gomez [4] and Papadopoulos [5], we assume an exponentially decreasing spectral profile for the amplitudes of the partials. An amplitude of 0.6^{i-1} is added for the i^{th} harmonic of every note in the chord.

In our system three chord models are defined, corresponding to 1, 4 or 6 harmonics. Examples for C major and C minor chords are displayed on Figure 1.

From these three chord models we can build chord templates for all types of chords (major, minor, dominant seventh, diminished, augmented,...). By convention in our system, the chord templates are normalized so that the sum of the amplitudes is 1.

3. MEASURES OF FIT

We consider for our recognition task several measures of fit, popular in the field of signal processing.

The **Euclidean distance** (*EUC*) defined by

$$D_{EUC}\left(\mathbf{x}|\mathbf{y}\right)=\sqrt{\sum_i\left(x_i-y_i\right)^2} \tag{4}$$

has notably already been used by Fujishima [2] for the chord recognition task.

The **Itakura-Saito divergence** defined by

$$D_{IS}\left(\mathbf{x}|\mathbf{y}\right)=\sum_i\frac{x_i}{y_i}-\log\left(\frac{x_i}{y_i}\right)-1 \tag{5}$$

is a measure of fit between two spectra, popular in the speech community. Since it is not symmetrical, it is not a distance and it can therefore be calculated in two ways. $D_{IS}\left(h_{k,n}\,\mathbf{c}_n|\mathbf{p}_k\right)$ is referred as *IS1* and $D_{IS}\left(\mathbf{p}_k|h_{k,n}\,\mathbf{c}_n\right)$ as *IS2*.

The **Kullback-Leibler divergence** measures the dissimilarity between two probability distributions. In the present paper we use the generalized Kullback-Leibler divergence defined by

$$D_{KL}\left(\mathbf{x}|\mathbf{y}\right)=\sum_i x_i\log\left(\frac{x_i}{y_i}\right)-x_i+y_i. \tag{6}$$

Just like the Itakura-Saito divergence, the Kullback-Leibler divergence is not a distance so that we can build two measures of fit : $D_{KL}\left(h_{k,n}\,\mathbf{c}_n|\mathbf{p}_k\right)$ *(KL1)* and $D_{KL}\left(\mathbf{p}_k|h_{k,n}\,\mathbf{c}_n\right)$ *(KL2)*.

4. FILTERING METHODS

In order to take into account the time-persistence, we introduce some post processing filtering methods which work upstream on the calculated measures and not on the sequence of detected chords.

The new criterion $\tilde{d}_{k,n}$ is based on L successive values $\{d_{k,n'}\}_{n-\frac{L-1}{2}\leq n'\leq n+\frac{L-1}{2}}$ previously calculated. In our system two types of filtering are used.

The **low-pass filtering** takes the mean of the L values. It tends to smooth the output chord sequence and to reflect the long-term trend in the chord change.

The **median filtering** takes the median of the L values. It has been widely used in image processing and is particularly efficient to correct random errors.

In every case, the detected chord $\hat{k}_n$ on frame n is the one that minimizes the set of values $\left\{\tilde{d}_{k,n}\right\}_k$:

$$\hat{k}_n=\underset{k}{\operatorname{argmin}}\left\{\tilde{d}_{k,n}\right\} \tag{7}$$

5. MIREX SUBMISSION

Our system can work with many parameter sets [6] and types of chord considered [7]. Our MIREX submission detects the major and minor chords, with the Kullback-Leibler divergence *KL2*, the single harmonic chord model and the median filtering with $L=15$.

6. ACKNOWLEDGMENT

The authors would like to thank C. Harte for his very useful annotation files used during the development phase.

7. REFERENCES

[1] J.P. Bello and J. Pickens. A robust mid-level representation for harmonic content in music signals. In *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*, pages 304–311, London, UK, 2005.

[2] T. Fujishima. Realtime chord recognition of musical sound: a system using Common Lisp Music. In *Proceedings of the International Computer Music Conference (ICMC)*, pages 464–467, Beijing, China, 1999.

[3] C.A. Harte and M.B. Sandler. Automatic chord identification using a quantised chromagram. In *Proceedings of the Audio Engineering Society*, Barcelona, Spain, 2005.

[4] E. Gómez. Tonal description of polyphonic audio for music content processing. In *Proceedings of the INFORMS Computing Society Conference*, volume 18, pages 294–304, Annapolis, MD, 2006.

[5] H. Papadopoulos and G. Peeters. Large-scale study of chord estimation algorithms based on chroma representation and HMM. In *Proceedings of the International Workshop on Content-Based Multimedia Indexing*, pages 53–60, Bordeaux, France, 2007.

[6] L. Oudre, Y. Grenier, and C. Févotte. Chord recognition using measures of fit, chord templates and filtering methods. Accepted in *Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New York, USA, 2009.

[7] L. Oudre, Y. Grenier, and C. Févotte. Template-based chord recognition : influence of the chord types. Accepted in *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*, Kobe, Japan, 2009.