Student Research Abstract: Using Chord Distance Descriptors to Enhance Music Information Retrieval

Ladislav Maršík
Dept. of Software Engineering,
Faculty of Mathematics and Physics,
Charles University, Malostranské nám.
25, Prague, Czech Republic
marsik@ksi.mff.cuni.cz

ABSTRACT
Music Information Retrieval (MIR) is an established field that provides solutions to analyze, retrieve, classify, recommend, or visualize music. From many applications, however, only a fraction will provide results that are meaningful for musicians [4]. Scoring functions may use descriptors that are difficult to describe for the end-user, and visualizations are often bound to signal aspects of music such as displaying waveform. We have developed a music analysis system which is based on music theory and contains visualizations meaningful for those interested in harmony aspects of music. Music is first segmented to chords by known techniques, providing the basis that musicians understand. From there, distances between chords are evaluated by a novel approach and new descriptors are formed, based on recent music theory studies. End-user can visualize the musical piece and find interesting sequences in a color temperature graph. While having the music visualized, user can retrieve similar musical pieces, understanding the similarity between the chord progressions.

Keywords
chord distance; chord transcription; cover song identification

1. BACKGROUND AND RELATED WORK
The gap between music theory and recent MIR applications has been pointed out by multiple researchers, calling for more work on how music theory can help recent retrieval tasks [1][4]. Even if the application provides valid results (e.g. retrieves a correct cover song), users may have difficulties understanding, why the result was made such. Our work is motivated by this fact and employs music theory in the proposed application. The analysis does not assume to be complete in the context of all music properties, including melody or rhythm. We focus on music harmony, since it is one of the most important aspects to describe the song structure, and can also be used effectively for cover song identification, as shown by Ellis [2].

1.1 Chord transcription
To obtain the chord representation, a standard multiple-step process is followed. First, obtaining chroma vectors by a short-time Fourier transformation, where chroma vector is a representation of a short musical moment mapped to 12 tones of the piano keyboard [2]. The next step is chord segmentation and labeling, jointly referred to as chord transcription [5]. Segmentation finds an exact timestamp for every chord change, while labeling names the chord using a chord dictionary.

SAC 2017, April 03-07, 2017, Marrakesh, Morocco.
ACM 978-1-4503-4486-9/17/04.
http://dx.doi.org/10.1145/3019612.3019939

1.2 Chord distances
We propose chord distances to fulfil the goal of analyzing and retrieving music. This term is defined freely in the literature, and can suggest multiple definitions. In MIR, chords are usually represented either as 12-dimensional vectors similar to chroma vectors, or as a string of tones (e.g. *CEG*). If we want to obtain distance between 2 chords, we can therefore use vector or string distances. However, other chord distances based on music theory are being proposed, that make the decision of choosing one distance challenging. Musicology standard is to use “Circle of Fifths” [3] and more profound models from which the most acknowledged is Tonal Pitch Space (TPS) from Lerdahl [3]. In addition to the knowledge of the chord, many methods need to be context-aware - in particular, having the knowledge of the music key. The whole variety of possible chord distances was summarized in the work of Rocher et al. [6].

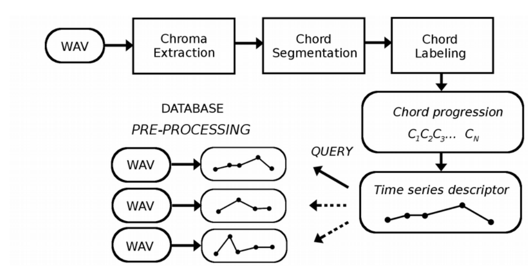

Figure 1: System outline for descriptor creation and retrieval query

2. APPROACH AND UNIQUENESS
Our approach starts with standard music processing tasks – obtaining chroma vectors and segmentation, as described in Section 2.1. We use methods proposed by Mauch and Dixon [5] for these tasks, allowing the use of other methods at the same time, as the analysis is independent from lower-level processing. Our approach differentiates from all known methods in the further steps: labeling (chord representation), key finding, chord distance and visualization, while common retrieval techniques can be used to evaluate the innovative approach. Figure 1 shows the outline of the whole system.

2.1 Chord representation

While most of the MIR works are using only a subset of all possible chords (chord dictionary), we think that the added value of dissonances can enhance the task. Therefore, we prefer to leave all sounding tones in the chord. As a result, even chords that are difficult to describe by music theory, containing dissonances, can appear in our analysis. There are advantages and disadvantages to this approach. Dissonant tones can be treated as invalid information towards retrieval, such as noise. However, non-chord tones can also play an important role in the character of the musical piece. An example can be a singer singing the voice above the chord accompaniment, causing a temporary dissonance.

2.2 Chord distance time series descriptor

After the chords are represented, we continue by evaluating transitions between every two subsequent chords and plotting them on the color temperature graph, such as on Figure 3. Using chord distances for obtaining time series is an innovative approach on its own – to the best of our knowledge, there are very few studies focusing on this aspect [1][6], despite the developed musicology background. Instead of using classic methods for chord distances such as TPS [3], we proceed by employing our new Chord Complexity Distance (CCD), based on a novel model, which is an alternative to TPS. Our innovation lies in taking every chord as a sentential form of a grammar-like system, and evaluating chord distance as the number of steps of derivation of one sentential form to another (see Figure 2). This way we can work with dissonant tones. Also, we can abstract from a separate key-finding method, as CCD model finds key and chord simultaneously, similar to the work of Rocher et al [7]. For a trained musicians, the proposed rules are understandable, since they are based on tonal harmony [3]. Once we extract time series descriptor, we can use this descriptor as a basis for a specific music retrieval task. We look for musical pieces with a similar series, as is depicted on the last steps on Figure 1. The descriptor is simple (most of the songs having less than 100 transitions), yet describing the whole chord progression, independent from the music key. As such it can be understood as a fingerprint of the chord progression and a way how to speed up music search.

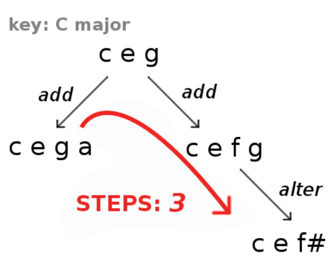

Figure 2: CCD model for distances based on adding/removing tones from the same key, and altering the tones outside the key. The key is evaluated for each chord tuple.

2.3 Visualization and important music parts

One of the main advantages of the time series descriptor introduced in previous section is the ease of visualization of a musical piece. As seen on Figure 3, both color temperature or line graph can be employed to show the data. This visualization can be easily used while playing the music in music player. When user retrieves a similar musical piece based on this descriptor, the visualization provides an understanding of why the piece was selected. Another advantage is, that user sees important chord sequences in the visualization easily, as the contrasting areas in the color temperature graph, or as the peaks in the line graph. These parts usually relate to interesting harmony movements, as are often in the bridge or before the start of the chorus.

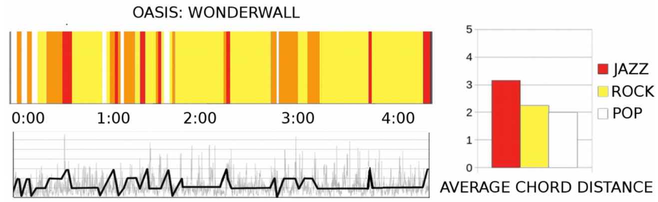

Figure 3: Visualization of the song Wonderwall by Oasis. On the top left, the chord progression is shown as a color temperature graph, on the bottom left as a line graph. On the right there are average chord complexity distances for 3 genres: Jazz, Rock and Pop made on a corpus of 120 songs.

3. RESULTS AND CONTRIBUTIONS

We have developed a system[1] capable of automatic analysis described in this abstract. The application provides visualizations which are easy to understand for musicians. On sample analysis on Figure 3, we can notice chord distance peaks around 0:40, 1:30 and 2:30, which correspond to the A5 chord followed by B7sus4, as well as a dissonance in the end caused by a guitar ornament. Preliminary results also show different average chord distances for different genres and promising results for music classification task (Figure 3 on the right). Such system can popularize MIR research among musicians and can be a useful plugin for today's music players. The system is distributed under GNU Public License, it fills in the gap in MIR studies and encourages further research on chord distances. As a future work, we focus on finding the best parameters for the model, as the effect of dissonances can sometimes cause inaccuracies (as at 3:40 on Figure 3). We also will test our descriptors for the cover song identification task. We think that the comparison of various chord distances and their impact on MIR tasks can be vital for the future research.

4. ACKNOWLEDGMENTS

This study was supported by the Charles University in Prague, project GA UK No. 708314 and by project SVV-2016-260331.

5. REFERENCES

[1] De Haas, W. B., Veltkamp, R. and Wiering, F. Tonal Pitch Step Distance: A Similarity Measure for Chord Progressions. *ISMIR* 2008

[2] Ellis, D. P. W. and Poliner, G. E. Identifying 'Cover Songs' with Chroma Features and Dynamic Programming Beat Tracking. *ICASSP* 2007

[3] Lerdahl, F. *Tonal Pitch Space.* Oxford University Press, Oxford, 2001

[4] Lewis, R. J., Fields, B. and Crawford, T. Addressing the Music Information Needs of Musicologists. *ISMIR* 2015

[5] Mauch, M. and Dixon, S. Approximate Note Transcription for the Improved Identification of Difficult Chords. *ISMIR* 2010

[6] Rocher , T., Robine , M., Hanna , P., and Desainte-Catherine , M. A Survey of Chord Distances With Comparison For Chord Analysis. *ICMC* 2010

[7] Rocher T., Robine M., Hanna P., and Oudre, L. Concurrent Estimation of Chords and Keys from Audio. *ISMIR* 2010

[1] http://harmony-analyser.org