

Analysing Musical Pieces Using harmony-analyser.org Tools

Ladislav Maršík

Dept. of Software Engineering, Faculty of Mathematics and Physics
Charles University, Malostranské nám. 25, 118 00 Prague 1, Czech Republic
`marsik@ksi.mff.cuni.cz`

Abstract. The tools provided under *harmony-analyser.org* are capable of recognizing harmonies, extracting the high-level harmony features, and plotting the harmony structure of the audio. They focus on the classical tonal analysis, as well as the distances between the harmonies to allow for the creation of novel descriptors. In the light of the recent expansion of the music retrieval techniques, the concepts of chord distances or chroma vector distances were still not studied to the full extent. With the presented tools we aim to provide an easy-to-use system for anyone interested in extracting these features, as well as an open-source framework written in Java for the developers interested in researching the concepts further. In this short paper, we offer the walk-through of *harmony-analyser.org* tools with the manual for the correct usage. We also summarize the results achieved using our system and we set the focus for the next development and research.

Keywords: harmony-analyser.org, tonal analysis, chord distance, chroma vector distance, music information retrieval

1 Introduction

The focus in Music Information Retrieval (MIR) is recently shifting to the large-scale approaches and techniques for a fingerprint extraction in the way that the most relevant audio features are retained [2]. The research teams are using fingerprints based on the spectrogram analysis [16], music theory [6], and most recently also experimenting with deep learning techniques to learn and distinguish musically relevant features [14].

In the search for the best features and fingerprints, it becomes increasingly difficult to know what features are already available. There are many proposed methods of extraction, as we can clearly see on benchmarking challenges such as MIREX¹ with over 15 distinct tasks, each requiring a different set of audio features, and the features changing every year since the first benchmarking in 2005.

¹ http://www.music-ir.org/mirex/wiki/MIREX_HOME

On the other hand, MIR is an interdisciplinary field and there are not many institutions worldwide having their own MIR team or laboratory. Therefore it can be challenging, especially for the young researchers, to join the common effort and be a part of the MIR project, unless a similar project is hosted by the researcher’s academic institution. To fulfil the need of onboarding the new researchers, popularizing the MIR field, giving an overview of the common techniques, and facilitating an open-source system, we have started the *harmony-analyser.org* project in 2016 [9].

Analysing harmonies is the main, but not the only aim of the project. The analysis output (harmony features) can be used for further retrieval easily, e.g. by employing Dynamic Time Warping (DTW) techniques [13]. We chose to focus on harmony to honour the fact that a musical piece usually contains multiple instruments played simultaneously, and the resulting harmony is one of the main features used for retrieval [6]. But the project is open for analysing melodies, rhythm, or beat tracking in the future, as well as using the machine learning approaches instead of the traditional feature extraction.

We continue by introducing the reader to the concepts and related work in Section 2. The step-by-step manual for the tools with screenshots is presented in Section 3. The first results obtained by our techniques are summarized in Section 4 and our future work is discussed in Section 5.

2 Harmony Features and Related Work

Chroma features is a common name for a series of 12-dimensional vectors of floating-point numbers, capturing the presence of each tone in a short music moment. They became popular after the works of Fujishima [5] and Bartsch and Wakefield [1]. Obtained directly from the Discrete-Time Fourier Transform output by grouping frequencies that belong together in one frequency bin, the resulting chroma vector has the form:

$$\langle c_A, c_{A\#}, c_B, c_C, c_{C\#}, c_D, c_{D\#}, c_E, c_F, c_{F\#}, c_G, c_{G\#} \rangle$$

where $c_A \in \mathbb{R}$ represents the presence of the A tone, $c_{A\#} \in \mathbb{R}$ represents the presence of $A\#$ tone, etc. The value distribution of $c_A, c_{A\#}, \dots$ depends on the algorithm used, but it is a common practice to normalize to $[0, 1]$ interval, where the value represents the loudness of the frequency bin. We refer the reader to Bartsch and Wakefield [1] for a detailed definition. One of the motivations for our work is, that chroma vectors have not yet been studied in terms of distances, even though the distances in between the chords have long been proposed by the works in music cognition [7].

Chord progression (a sequence of chord labels) is a familiar concept for musicians, who often use it to play together in an unrehearsed situation. The idea of using chord progression itself as a fingerprint for large-scale music retrieval was proposed by Khadkevich and Omologo [6], improving the state-of-the-art cover song identification results in 2013. The progression can be represented as

a sequence of strings (C , $F6$, $Gmaj7$, ...), or boolean vectors similar to chroma vectors.

Chord distance is a concept based on the acknowledged music cognition findings: the listeners perceive the differences in chords in a way that can be predicted by a formal tonal harmony model. Fred Lerdahl's Tonal Pitch Space (TPS) model [7] was proposed and backed up by the empirical studies. This concept was further studied by several MIR authors [4] [12] [15], combining the cognitive and computational chord distances. A thorough review of the available chord distances was assembled by Rocher et al. [15]. Notably, the TPS distance performed the best in the studies for the chord estimation or cover song identification tasks [12] [15].

3 Usage of Harmony Analyser Tools

In *harmony-analyser.org* project, we provide GUI tools published as executable JAR archives, to allow for a custom harmony analysis of WAV or MIDI input. The tools itself are using the *JHarmonyAnalyser* Java library, which we describe in details in the more technical report [9]. To achieve a high variety of analysis, we also incorporated GPL-licensed Vamp plugins² to the GUI tools. The advanced users can customize their analysis by downloading additional plugins or creating their own.

In this section we focus on a simple use case of running the tools to get a simple analysis of the MIDI keyboard input and WAV files. We also describe the differences from the other systems and possible usages for the research along the way.

3.1 Chord Transition Tool

When the application starts, the default tool selected is the *Chord Transition Tool* (see Figure 1). The user can either use the MIDI keyboard plugged in via the USB port, or use a text input field, to specify two chords. The added value compared to other common MIDI software is a list of functions and chord distances, based on the tonal analysis (described in more details in [10]). The fact that the chord can have multiple functions in music is commonly accounted for in the works on musicology, but less frequently in the MIR works. This is one of the many examples of a gap between MIR and musicology, which should be addressed, as pointed out by Lewis [8]. Chord Transition Tool shows the chord and all of its tonal functions, and the user can observe various chord distances (Chord Complexity Distance [10], or TPS Distance [7]) as seen on Figure 1, which gives him a good overview for developing advanced tonal features.

3.2 Visualization Tool

After the user is familiar with chordal analysis described in the previous section, the next step is to observe the chords, chord distances, or chroma vector distances

² <http://vamp-plugins.org>

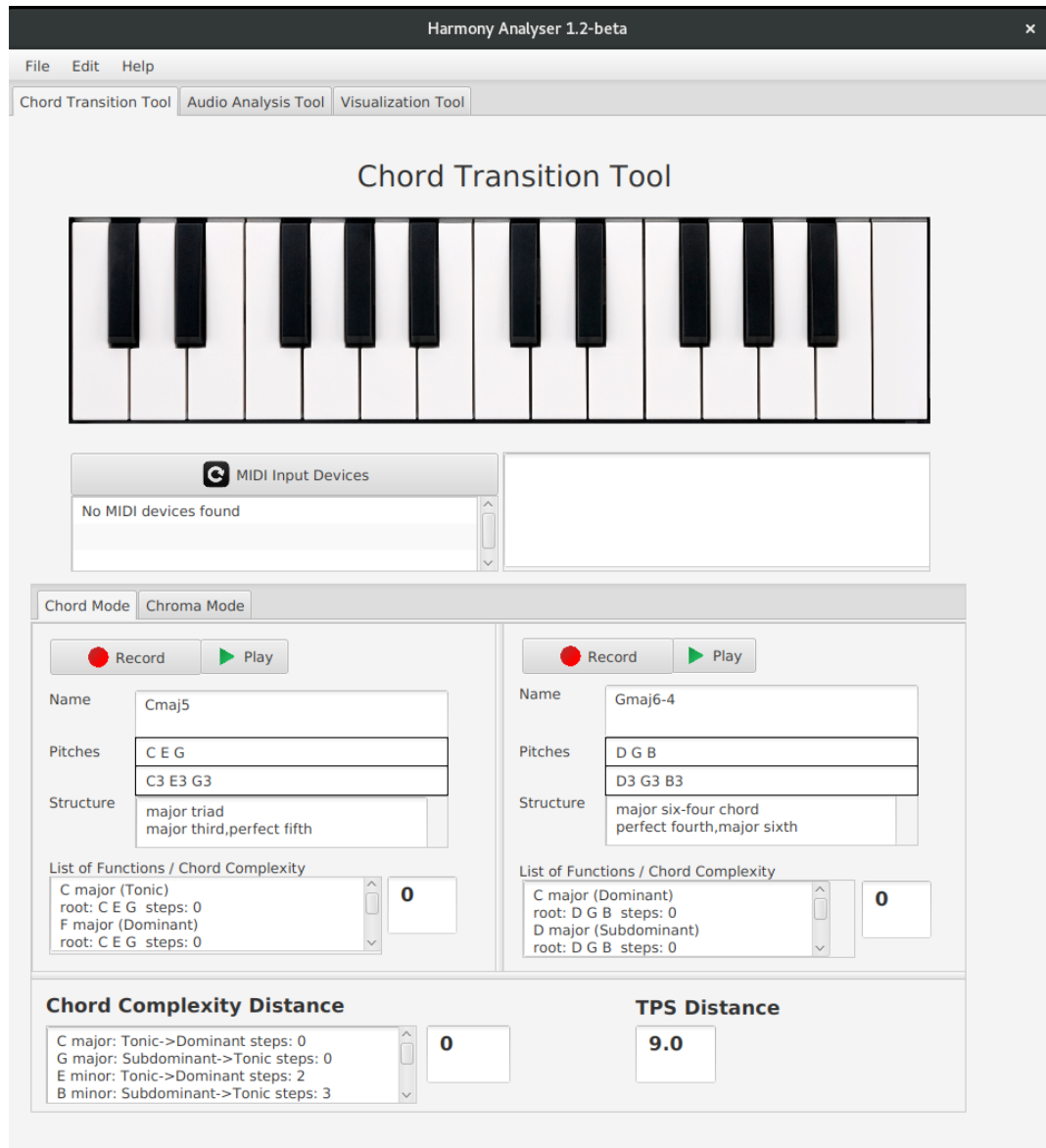


Fig. 1. Chord Transition Tool: capturing the MIDI input and outputting the chord labels, functions and the chord distances. *C major* and *G major* chords are analysed.

extracted from the real audio. We offer the Visualization Tool (see Figure 2) to visually understand how the labels and distances can help analysing a musical piece. In the musical piece analysis on Figure 2 (Hallelujah by Bastian Baker) we have time in seconds on the x axis, and chord distance values of *each pair of the subsequent chords* on the y axis. This is one of the song fingerprints that we experimentally studied. In the given analysis, the chord distance time series represents a typical curvature of the harmony movement in the piece. The local peaks around 30th or 80th second represent the transition between *Ami* and *F* chords. The peaks after the 150th second represent the same transition with the singer performing vocal ornaments in the last verse, yielding a higher (more complex) chord distance value, since the voice is accounted for in the chord estimation.

The same chart visualization can be shown for each plugin that extracts the values in the form of a time series (e.g. chroma vector distances), or labels with a timestamp (chord or key detection). Some plugins will output column charts, such as Average Chord Complexity Distance [11] on Figure 2. We have shown how these averaged features improve the music genre detection in one of our previous studies [10].

3.3 Audio Analysis Tool

The last step of the analysis after understanding the harmony features thoroughly, is to apply the chosen analysis on a folder with WAV files. This can be achieved by the Audio Analysis Tool (Figure 3). The plugins are categorized in the plugin groups (*Vamp plugins*, *Chord analyser*, *Chroma analyser*) and the details and parameters of the selected plugin are shown. After hitting the *Analyse* button, the tool creates text files with the analysis results in the selected folder. These can be used as an input for another analysis plugin, or an input for a retrieval technique. There is also an additional *Post Processing* tab that serves various purposes, such as applying a smoothing filter to a time series. The additional tabs can also be helpful for importing or exporting other file types, so that the application can be used for various projects. As an example, The Million Song Dataset from Bertin-Mahieux et al. [3] uses *HDF5* files, and by providing a conversion to text files the dataset can be used easily with Audio Analysis Tool.

4 Summary of Results

The tools from *harmony-analyser.org* have already been tested on various MIR tasks. We have gathered an average chord complexity distance and used this average for the genre detection, as one of the features for the neural network method. The usage of the feature yielded to 4% precision improvement for the dataset of 100 musical pieces [10].

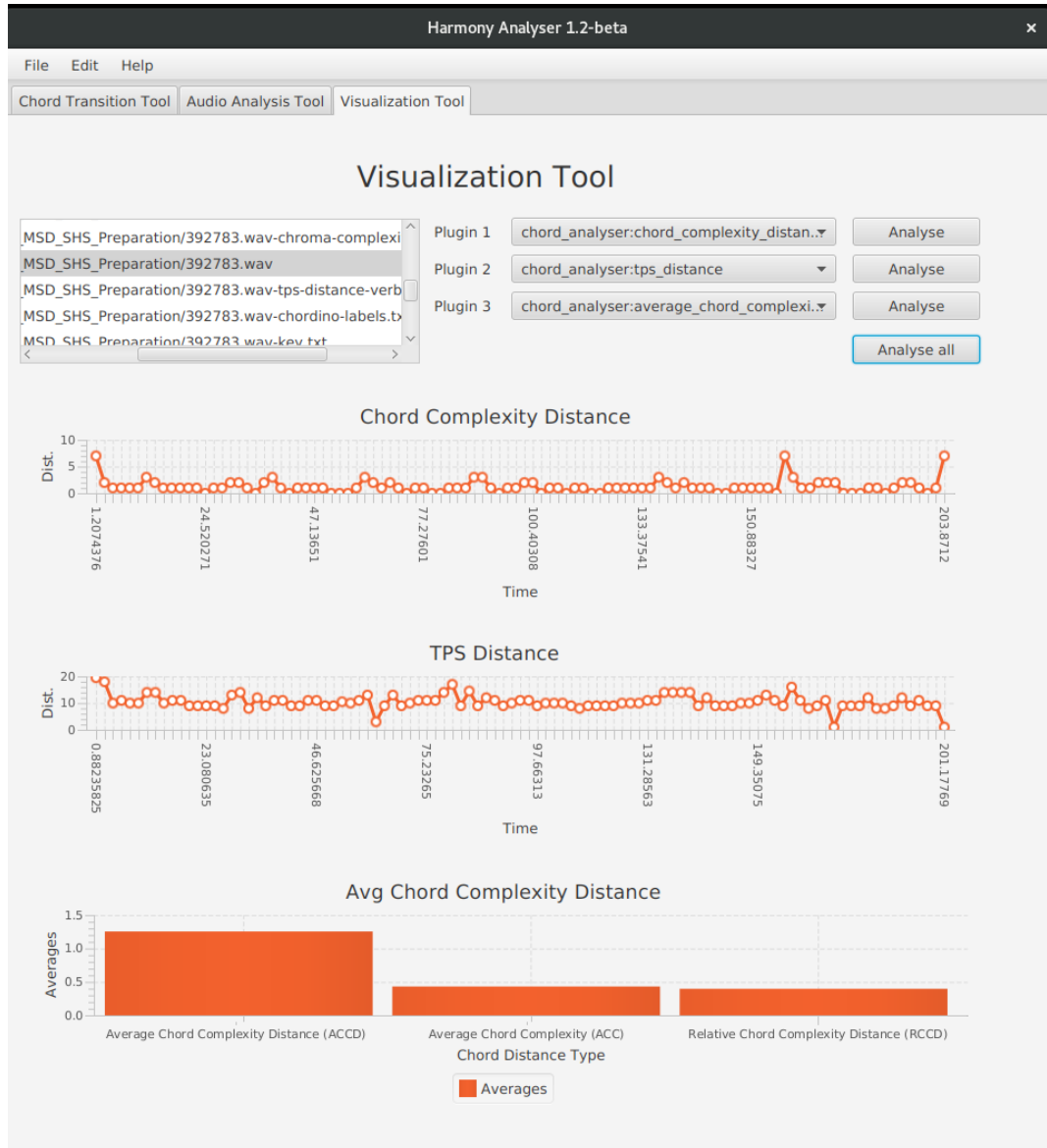


Fig. 2. Visualization Tool: Analysis of Hallelujah by Bastian Baker is shown, containing results for Chord Complexity Distance, TPS Distance, and three types of averages for Chord Complexity Distance [11].

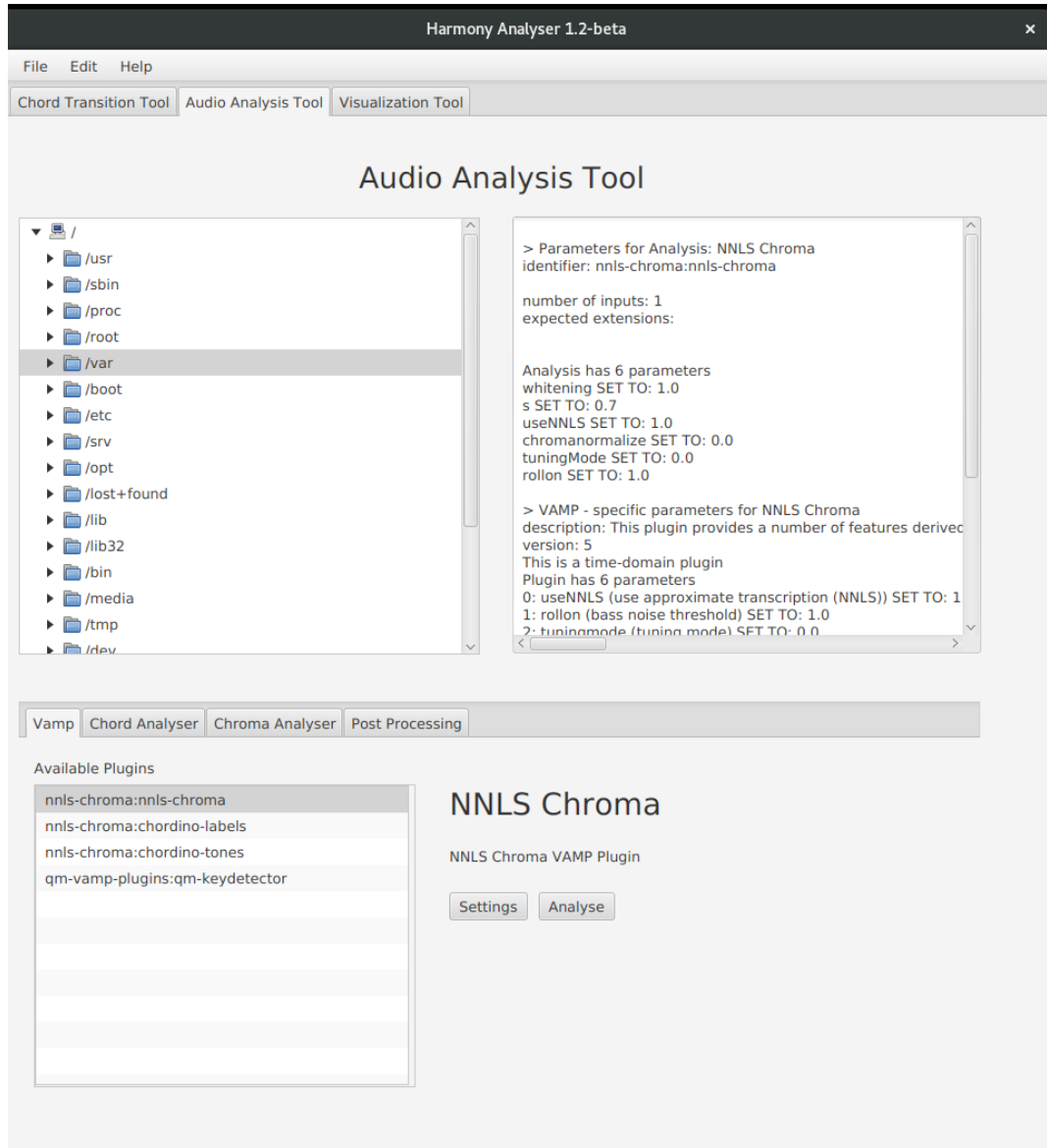


Fig. 3. Audio Analysis Tool: selecting a folder with WAV files and choosing a desired plugin for analysis.

The chord distance time series were tested on both covers80 dataset³ and a subset of SecondHandSongs dataset⁴ (999 songs), on a cover song identification task using DTW method [12]. The results show that TPS distance have outperformed Chord Complexity Distance in the MAP (Mean arithmetic of Average Precision) score. Overall, the usage of a chord distance time series means a loss in the MAP score compared to more low-level features: from 0.482 (full chroma features) to 0.198 (TPS distance) for covers80 dataset, but it comes with a more than a two thousand times faster performance (56s versus 25ms execution time for DTW matrix calculation of 80 songs).

The chroma vector distances were tested on the same datasets and task. The results were comparable to the results of a TPS chord distance (0.174 MAP score), which is promising for a first feature of this type.

These experiments show that the chord or chroma vector distance features do not provide enough information on their own for the retrieval, but if used properly in the combination with more low-level features, they can improve the performance.

5 Conclusion and Future Work

The *harmony-analyser.org* tools can be used for a musical piece analysis, feature extraction from audio files, or as a basis for further research and retrieval. They contain a variety of plugins for analysis, giving a thorough overview of what harmony features are currently available. The tools are also extensible in the way that new plugins can be downloaded or developed. We provided an overview of the usage of the main tools, and the summary of the achieved results, showing the ways to enhance the algorithms for MIR tasks.

Our latest ideas were to utilize the concepts of chord and chroma vector distances differently. Rather than a stand-alone time series, we will be experimenting with using the distances in DTW calculation (comparison of two vectors done by the chord or chroma vector distance instead of the Euclidean distance). We also plan to include more types of chord distances in our tools to get a thorough comparison. Last but not least, we will continue to present the tools in the open-source community, to get more developers for the project, with the overall aim to make *harmony-analyser.org* an all-in-one music retrieval system.

Acknowledgments. The study was supported by the Charles University in Prague, project GA UK No. 1580317.

Bibliography

1. Bartsch, M.A., Wakefield, G.H.: To Catch a Chorus: Using Chroma-Based Representations for Audio Thumbnailing. In: *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. WASPAA 2001 (2001)

³ <https://labrosa.ee.columbia.edu/projects/coversongs/covers80>

⁴ <https://labrosa.ee.columbia.edu/millionsong/secondhand>

2. Bertin-Mahieux, T., Ellis, D.P.W.: Large-Scale Cover Song Recognition Using Hashed Chroma Landmarks. In: *IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics*. WASPAA 2011, IEEE (2011)
3. Bertin-Mahieux, T., Ellis, D.P., Whitman, B., Lamere, P.: The Million Song Dataset. In: *Proceedings of the 12th International Society for Music Information Retrieval Conference*. ISMIR 2011 (2011)
4. De Haas, W.B., Veltkamp, R., Wiering, F.: Tonal Pitch Step Distance: A Similarity Measure for Chord Progressions. In: *Proceedings of the 9th International Conference on Music Information Retrieval*. ISMIR 2008 (2008)
5. Fujishima, T.: Realtime Chord Recognition of Musical Sound: A System Using Common Lisp Music. In: *Proceedings of the International Computer Music Conference*. ICMC 1999 (1999)
6. Khadkevich, M., Omologo, M.: Large-Scale Cover Song Identification Using Chord Profiles. In: *Proceedings of the 14th International Society for Music Information Retrieval Conference*. ISMIR 2013 (2013)
7. Lerdahl, F.: Tonal Pitch Space. Oxford University Press, Oxford (2001)
8. Lewis, R.J., Fields, B., Crawford, T.: Addressing the Music Information Needs of Musicologists. In: *Proceedings of the 16th International Society for Music Information Retrieval Conference*. ISMIR 2015 (2015)
9. Marsik, L.: harmony-analyser.org - Java Library and Tools for Chordal Analysis. In: *Proceedings of 2016 Joint WOCMAT-IRCAM Forum Conference*. WOCMAT 2016, Kainan University, Taiwan (2016)
10. Marsik, L., Pokorný, J., Ilčík, M.: Improving Music Classification Using Harmonic Complexity. In: *Proceedings of the 14th conference Information Technologies - Applications and Theory (ITAT 2014)*. Ústav informatiky AV ČR (2014)
11. Marsik, L., Pokorný, J., Ilčík, M.: Towards a Harmonic Complexity of Musical Pieces. In: *Proceedings of the 14th Annual International Workshop on Databases, Texts, Specifications and Objects (DATESO '14)*. CEUR Workshop Proceedings, vol. 1139. CEUR-WS.org (2014)
12. Marsik, L., Rusek, M., Slaninová, K., Martinović, J., Pokorný, J.: Evaluation of Chord and Chroma Features and Dynamic Time Warping Scores on Cover Song Identification Task. In: *Proceedings of the 16th International Conference on Computer Information Systems and Industrial Management Applications*. CISIM 2017, Springer (2017)
13. Müller, M.: Information Retrieval for Music and Motion. Springer Berlin Heidelberg (2007)
14. Pons, J., Lidy, T., Serra, X.: Experimenting with Musically Motivated Convolutional Neural Networks. In: *14th International Workshop on Content-based Multimedia Indexing*. CBMI 2016, IEEE (2016)
15. Rocher, T., Robine, M., Hanna, P., Desainte-Catherine, M.: A Survey of Chord Distances With Comparison For Chord Analysis. In: *Proceedings of the International Computer Music Conference*. ICMC 2010 (2010)
16. Wang, A.L.: An Industrial-Strength Audio Search Algorithm. In: *Proceedings of the 4th International Society for Music Information Retrieval Conference*. ISMIR 2003 (2003)