

QUANTIFYING AND VISUALIZING TONAL COMPLEXITY

Christof Weiß¹, Meinard Müller²

¹ *Fraunhofer Institute for Digital Media Technology Ilmenau, Germany*

² *International Audio Laboratories Erlangen, Germany*

Correspondence should be addressed to: christof.weiss@idmt.fraunhofer.de

Abstract: In Western classical music, the structure of a piece is reinforced by the contrasting harmonic nature of its sections. The structural parts are characterized by the presence of certain chords or chord changes. A section that is harmonically stable may be followed by a contrasting section that feels unstable or tense. In the sonata form, for example, the unstable development part is located between the stable exposition and recapitulation phases. In this paper, we try to measure this kind of harmonic stability and present visualizations for such analyses. To this end, we propose novel features for quantifying tonal complexity and discuss their musicological implications. The features are based on statistical measures calculated from chroma representations of the music recording.

The characteristics of tonal complexity apply to different time scales. To illustrate this time scale dependence for the proposed features, we use hierarchical visualizations based on previously introduced scape plot representations. On a fine temporal level, tonal complexity is related the character of chords or scales. For example, in a modulating transition phase, we usually find more complex chords than at the beginning of a piece. To analyze such differences, we study the feature values for isolated chords. Looking at a coarser level, the presence of modulations is an indication for a segment's complexity. In the sonata form, for example, the development usually contains several modulations. To account for this property, we calculate the complexity features based on a coarse resolution of the chroma features. For evaluation of this coarse-scale complexity, we analyze Beethoven's sonatas where we find higher complexity in the development parts.

1. INTRODUCTION

In Western art music, one major purpose of harmony is to emphasize musical structure. Typical harmonic phenomena are used to highlight pivotal moments of a composition. This observation applies to different time scales: Vertical structures such as intervals or chords show different characteristics with respect to harmonic stability, creating a feeling of either tension or resolution. Horizontal structures such as pendula, sequences, modulations, and cadences use chords of appropriate quality in order to form larger lines of development. Over the course of a work, different structural parts may differ significantly with respect to their tonal characteristics, thus creating an arc of tension of a musical piece.

Apart from such intra-work aspects, there is a related but more abstract quality describing the harmony of complete pieces or even a compositional style. The pitch class selection of Western music evolved from a diatonic scale to a fully chromatic set of equally relevant pitches in the atonal period [1]. The applied chords and chord progressions became more complex—on a rough scale—over the centuries. In this work, we try to quantify such characteristics related to tonal complexity on the basis of an audio recording. We present several complexity measures and investigate their behaviour for harmonic phenomena on different time scales.

There have been several attempts to approach similar concepts for symbolic data. In [2], ideas from pitch class set theory have been tested to measure degrees of tonality. Notions such as pitch entropy are used for style classification in [3]. Considering audio data, methods to quantify properties related to tonal complexity were presented in [4, 5] addressing sequential properties of tonality. We propose a different approach, accounting for the local pitch class distribution on various temporal scales of the audio data.

As main contribution of this paper, we discuss concepts such as harmonic tension, pitch entropy, and degree of tonality under the label “tonal complexity”. To obtain a more precise definition of

this notion, we compile in Section 2 a set of assumptions for several musicological tasks regarding various temporal scales. In Section 3, we propose novel features based on a chroma representation of audio data. The time-scale dependence of these features is visualized using scape plots. In Section 4, we test the introduced hypotheses on two different datasets.

2. MUSICOLOGICAL IMPLICATIONS

Assuming the existence of a musical dimension that can be referred to as “tonal complexity”, we want to approach the meaning of this quantity by considering several musicological questions. The smallest vertical—meaning simultaneously sounding—items regarding tonality are intervals and chords. The quality of these items plays an important role to create stabilizing and destabilizing musical moments. Consider the simple cadence of a G dominant seventh chord followed by a C major chord. The striving nature of the seventh chord with the dissonant tritone interval requires a resolution to a consonant chord. In a more advanced tonal context such as late romantic or jazz harmony, more complex resolution chords can arise. However, in that case the previous chord is often found to be even more dissonant. Thus, chord level complexity may be related to the grade of dissonance of the local tonal content. A major chord in root position suggests a more stable feeling to the listener than a diminished chord, a dominant seventh chord, or just this major chord while playing ornamental off-chord notes.

On a coarser scale, the change of chords and their tonal relationships are responsible for the complexity. This level refers to the scales representing the local pitch content, and the way these scales change. Chord changes within the pitch content of a diatonic scale do not sound very surprising, neither do chords from a neighboring key with only one or two “new” accidentals. However, a C major chord followed by an F \sharp major chord without harmonic progression will be perceived as an abrupt change. To quantify such relationships, ordering the pitches in distances of perfect fifth intervals has turned out useful. Just as in the circle of fifth, keys, chords, and pitches that differ by a small number of fifths show closer tonal relationships and, thus, constitute smoother harmonic changes. The amount and distance of modulations in the music may have an influence on the perceived tonal complexity.

Motivated by these considerations, we want to find a measure, say Γ , that expresses some kind of complexity of the tonal content of audio data on various musical scales. Intuitively, such a measure should support the following trends on different structural levels.

- **Chord level.** Different chords or scales should show distinct complexity:

$$\Gamma(\text{complex chord}) > \Gamma(\text{simple chord}) \quad (1)$$

- **Fine structure.** The subparts of a sonata exposition should be different in complexity:

$$\Gamma(\text{transition phase}) > \Gamma(\text{theme}) \quad (2)$$

- **Coarse structure.** The parts of a sonata form movement should show specific trends in complexity:

$$\Gamma(\text{development}) > \Gamma(\text{exposition}) \quad (3)$$

- **Cross-work.** Considering the oeuvre of one composer, we expect the late works to be more complex than the early ones:

$$\Gamma(\text{late sonata}) > \Gamma(\text{early sonata}) \quad (4)$$

- **Cross-composer.** On a cross-composer level, we assume stylistic trends. The historical periods may be characterized by different complexity:

$$\Gamma(\text{romantic}) > \Gamma(\text{classical}) \quad (5)$$

In the following, these hypotheses will be the basis for the discussion and evaluation of the proposed features. We are conscious of the limitations of these rather simplistic assumptions and use them only as a guiding principle for testing certain tendencies. Some of the hypotheses could be verified with perceptual studies and listening tests, for others, a closer look at the musical scores and a detailed view on musical styles may be necessary. For testing the different assumptions, we consider test scenarios based on two datasets. First, we use a set of recorded piano chords, second, we consider the first movements of the Beethoven sonatas.

3. FEATURES

3.1. Basic Features

For an appropriate description of tonality, we want the complexity features to be invariant against timbral variations. For example, an orchestra chord should obtain a similar value as the same chord played on a piano. Thus, we build our systems on chroma features which have been shown to capture tonal information and to be invariant against timbral variations to a large extent. In the following, $\mathbf{c} = (c_0, c_1, \dots, c_{11})^T \in \mathbb{R}^N$ denotes a chroma vector of dimension $N := 12$. The entries c_0, c_1, \dots, c_{11} indicate the energy of the twelve chroma classes C, C[♯], ..., B. Because of the octave invariance, a chroma vector shows a cyclic behaviour: a circular shift of the chroma entries corresponds to a transposition of the music. For the chroma extraction, we use a pitch filter bank approach from the Chroma Toolbox [6]. We use a feature resolution of 10 Hz and normalize the features columnwise with respect to the ℓ^1 norm so that $\|\mathbf{c}\|_1 = 1$.

In this paper, we introduce some basic concepts of quantifying tonal complexity. Thus, we do not optimize the chroma extraction by considering higher partials. For further improvements, it may be necessary to consider more advanced chroma computation methods such as the Harmonic Pitch Class Profiles [7], or the Nonnegative Least Squares chroma introduced in [8]. In order to account for the logarithmic behavior of loudness perception, we apply a logarithmic compression before the normalization step. Each energy entry e is replaced by $\log(\eta \cdot e + 1)$ with a positive constant η . This procedure has been shown to improve chord detection in [9]. Inspired by that work, we use $\eta = 100$ for our experiments.

3.2. Complexity Features

Based on normalized chroma features, we now derive more abstract features modeling tonal complexity. The basic idea is to calculate a measure for the flatness of a chroma vector. Our feature values should be scaled to unit range and increase for growing tonal complexity:

$$0 \leq \Gamma \leq 1. \quad (6)$$

On a local scale, the least complex tonal item may be an isolated musical note. This is represented by a sparse pitch class distribution

$$\mathbf{c}^{\text{sparse}} = (1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)^T. \quad (7)$$

This item should obtain the lowest feature value $\Gamma(\mathbf{c}^{\text{sparse}}) = 0$. In contrast, an equal distribution \mathbf{c}^{flat} corresponding to a simultaneously played chromatic scale should yield the highest value, e.g., $\Gamma(\mathbf{c}^{\text{flat}}) = 1$.

We have tested several statistical measures for their suitability to capture such characteristics. In the following, we present three selected features which are used for the experiments.

- As the first feature $\Gamma_{\text{Entr}}(\mathbf{c})$, we use the Shannon entropy of the chroma vector, which describes the amount of information in a data sequence.

$$\Gamma_{\text{Entr}}(\mathbf{c}) = -\left(\sum_{n=0}^{N-1} c_n \cdot \log_2(c_n)\right) / \log_2(N). \quad (8)$$

- Next, we use a flatness measure $\Gamma_{\text{Flat}}(\mathbf{c})$ calculated as the quotient between the geometric and the arithmetic mean of the chroma vector

$$\Gamma_{\text{Flat}}(\mathbf{c}) = \left(\prod_{n=0}^{N-1} c_n\right)^{1/N} / \left(\frac{1}{N} \sum_{n=0}^{N-1} c_n\right). \quad (9)$$

- For the third complexity feature, we take into account the harmonic similarity of the pitch classes. To this end, we reorder the chroma values to a perfect fifth ordering $\mathbf{c}^{\text{fifth}}$:

$$c_n^{\text{fifth}} = c_{(n \cdot 7 \bmod 12)}. \quad (10)$$

We obtain a circular distribution of the pitch class energies similar to the circle of fifths. Therefore, we interpret the chroma values as angular data and calculate the angular deviation describing the spread of the pitch classes:

$$\Gamma_{\text{Fifth}}(\mathbf{c}) = \left(1 - \left|\frac{1}{N} \sum_{n=0}^{N-1} c_n^{\text{fifth}} \exp\left(\frac{2\pi i n \pi}{12}\right)\right|\right)^{1/2}. \quad (11)$$

For all features, the boundary conditions $\Gamma(\mathbf{c}^{\text{flat}}) = 1$ and $\Gamma(\mathbf{c}^{\text{sparse}}) = 0$ are fulfilled. With this set of features, we consider several flatness-related aspects of a chroma vector.

3.3. Scale Dependence

The measurement of complexity crucially depends on the time scale of the observation. On a chromagram with fine resolution, the measures give an estimate of the complexity of chords and local scales. Regarding coarser levels, the complexity of several bars or a whole section is calculated. Using a chroma histogram as input, the complexity value refers to the full movement. To examine the dependence of our proposed features, we visualize them hierarchically on different time scales, using the scape plot technique by Sapp [10]. Examples can be found in Section 4.2.

4. EVALUATION

4.1. Chord Study

To better understand the behavior of the proposed features, we analyze their behaviour for different local items of tonality such as pitches, intervals, chords, and scales. First, we do this for synthetic versions of these items and calculate the feature values Γ for idealized binary templates. For example, the major chord template is $\mathbf{c}^{\text{major}} = \tilde{\mathbf{c}}^{\text{major}} / \|\tilde{\mathbf{c}}^{\text{major}}\|_1$ with

$$\tilde{\mathbf{c}}^{\text{major}} = (1, \varepsilon, \varepsilon, \varepsilon, 1, \varepsilon, \varepsilon, 1, \varepsilon, \varepsilon, \varepsilon, \varepsilon). \quad (12)$$

To avoid degenerations in formulas due to zero entries, we use a small value $\varepsilon = 0.05$ for the silent pitch classes. Second, we analyze real audio recordings of the same chords, played on a piano for approximately 3 sec. For both representations, we calculate the values of the three features. As items, we used a single pitch, a fifth interval, the four basic triads, seven types of seventh chords, four types of ninth chords, and three scales (pentatonic, diatonic, and chromatic). The results of this study are shown in Figure 1.

First, let us start with the results for ideal chord templates. Even if we expect $\Gamma = 0$ for the single pitch (No. 1), this is not the case due to the non-zero ε entries. With growing number of notes, Γ_{Entr} increases monotonically. A seventh chord obtains higher complexity values than a triad for this feature. For three or more notes per chord, Γ_{Flat} behaves similar. Both Γ_{Entr} and Γ_{Flat} only react on the number and intensity of sounding pitches but do not depend on their interrelations. They correspond to some degree of polyphony of the local chords. In contrast, Γ_{Fifth} accounts for the ordering of the pitches. For example, the value for a diminished triad (No. 5) is higher than for a major triad (No. 3), since the diminished triad has a larger spread on a perfect fifth axis. Symmetric divisions of the octave such as the augmented triad (No. 6) and the full-diminished seventh chord (No. 12) obtain high complexity values. In contrast, the pentatonic scale with 5 pitches has a relatively small Γ_{Fifth} value, since all pitches are fifth-related.

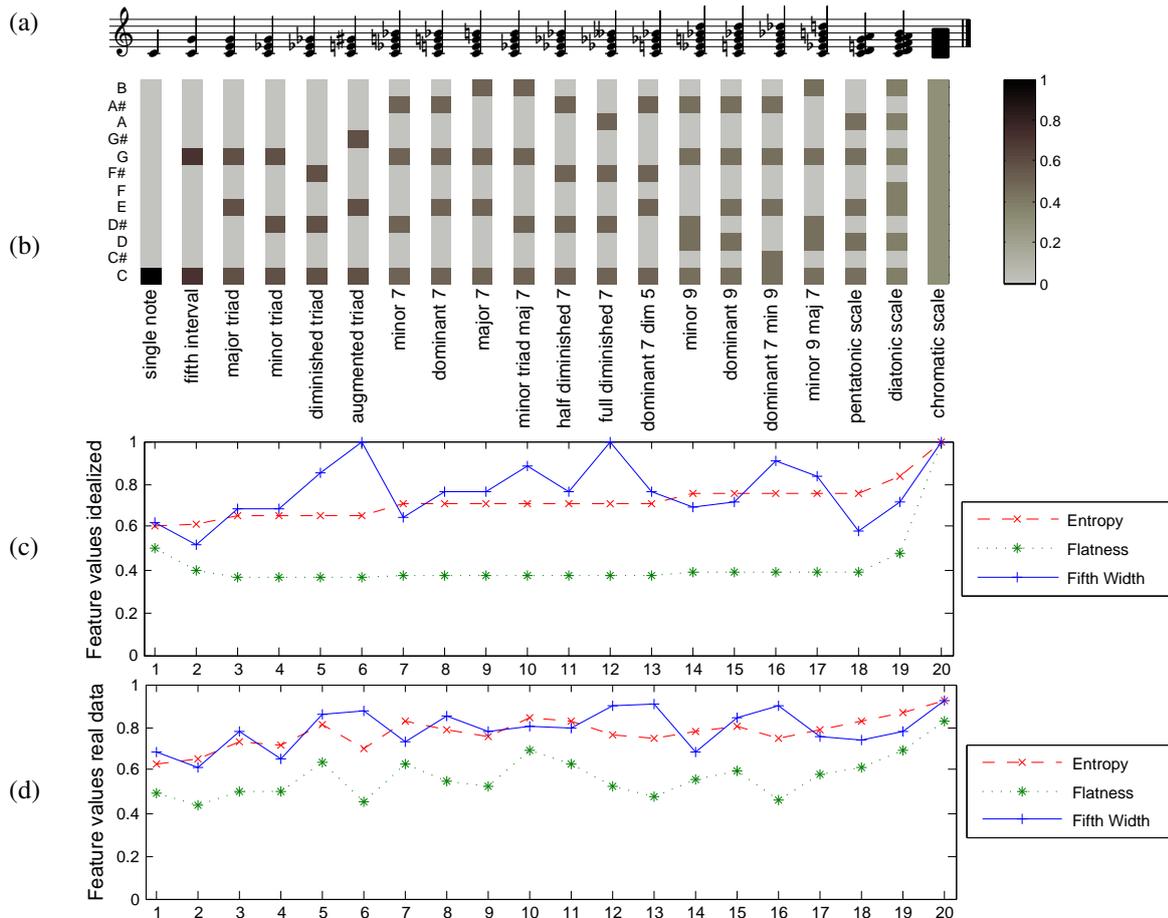


Figure 1: Complexity feature values Γ for 20 tonal items. The musical notations of the items is given in (a), the ideal chroma templates of the items in (b). (c) shows the values for the ideal templates and (d) for the recorded piano chords.

For the recorded chords, differences in intensity appear in the chroma vector, although the chords were played with equal loudness, approximately. The features react on these variations so that the above-mentioned observations are less clear for the real piano chords. Γ_{Flat} has turned out particularly sensitive to this effect. To improve the robustness of the features, more elaborate chroma features could be useful.

4.2. Study on the Beethoven Sonatas

Next, we study Beethoven’s sonatas in the recording of Barenboim. Even though they are not standard sonata examples of their time but full of surprising ideas and changes, general trends can be observed. In the upper part of Figure 2, we show three scape plots as introduced in Section 3.3. To compute the plots, the original 10 Hz chroma vectors are averaged at different window sizes. The horizontal axis gives the position of the segment in seconds, the vertical axis corresponds to the length of the segment. The lowest row describes a local level, the highest one gives a single value for the full recording. The feature value for the respective segment is encoded by the color. For all movements, we see a dark region indicating high complexity for the development phases. The similarity between exposition and repetition can be recognized. Regarding the fine structure, we see bright phases corresponding to the themes and dark phases describing the higher complexity in the transition phases. In the development, the global complexity is always high, but not the local one. This may arise from development parts without complex chords, but with complex modulations, covering distant keys within a short segment. Looking at the development of the “Appassionata” Op. 57 in Figure 2 (b), a modulating phase is followed by a long segment in $A\flat$ major, indicated by a white section starting at 240 sec.

To test the coarse structure hypothesis, we plot the averaged Γ_{Fifth} values of the main parts for the 28 head movements in sonata form (Figure 2, lower part). The complexity in the development phase is always highest, with four exceptions. One case is the sonata Op. 109, where the development shows almost no modulations. Rather, the movement consists of alternating parts of similar harmonic structure. In the G minor sonata Op. 49, No. 1, the development contains a long stable $E\flat$ major part and thus is not rated with a high complexity. In contrast, the recapitulation of this movement yields a high Γ_{Fifth} value, clearly higher than for the exposition. A reason for this observation may be the modulation structure of the sonata form in minor. In the exposition, the second theme usually stands in the relative major key and thus contains mainly one diatonic scale. In the recapitulation, this part is transposed to the tonic (minor key) which includes pitches from the harmonic and melodic minor scales, leading to a higher complexity. A similar effect can be observed for other movements in minor key (Op. 2, No.1 or Op. 10, No.1). A contrasting scenario is found in Op. 79, where a stable exposition section is followed by a strongly modulating development, covering the local keys E major, C major, C minor, $E\flat$ major, and the change back to G major. In future work, the discussion of further details could be combined with analyses of modulations such as the ones presented in [11].

Regarding global complexity, the hypothesis 4 of increasing values over time does not hold. The scores for the late works change substantially—a hint to high individuality of the compositions—in contrast to the early sonatas, which show a similar complexity structure among each other. Within the late sonatas, we find the most extreme values: the light and tonally constant Op. 101 in E major, in contrast to the last sonata Op. 111 in C minor with complex harmony full of dissonances and a polyphonic development.

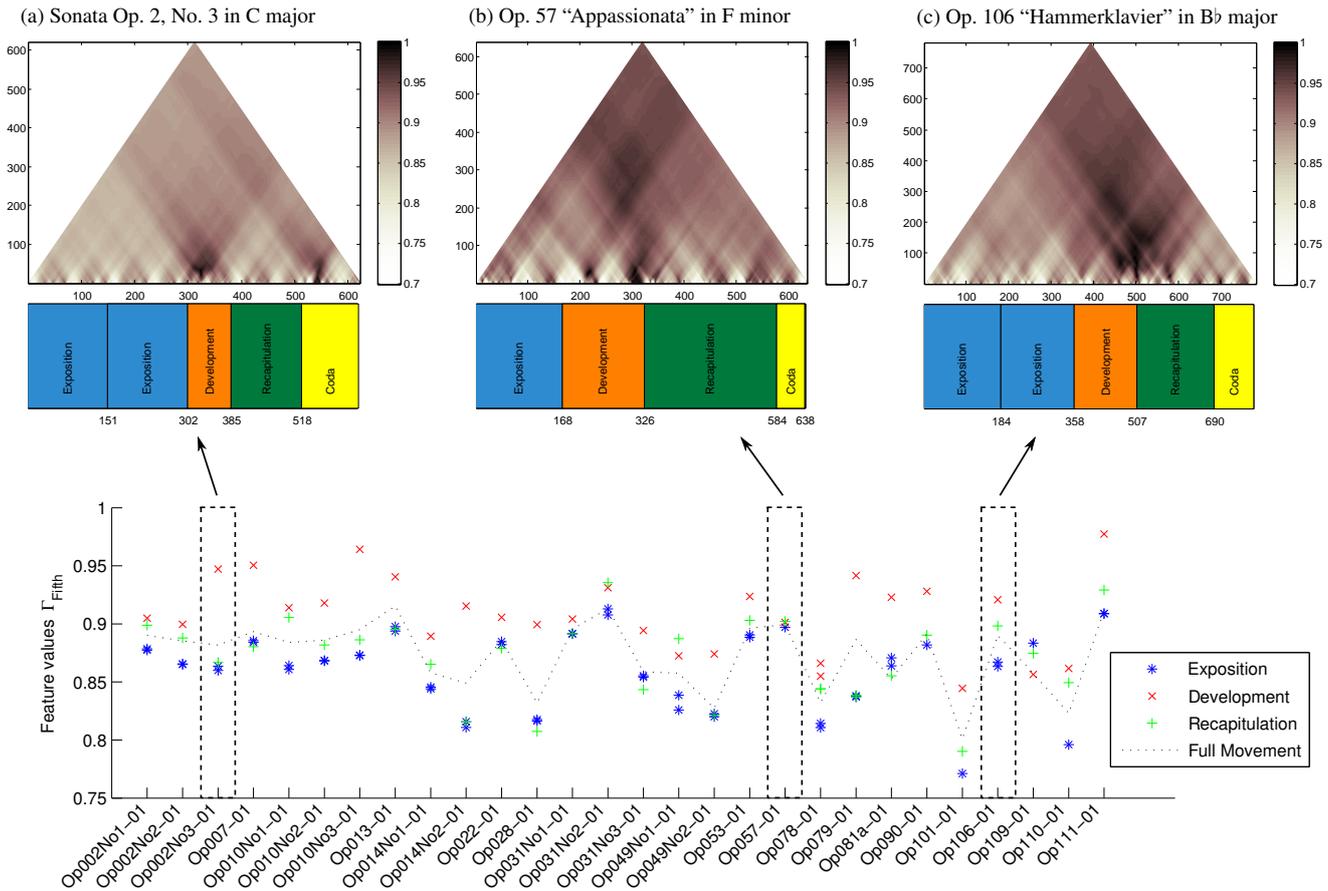


Figure 2: Scape plots using Γ_{Fifth} for the first movements of three Beethoven Sonatas (upper row). For an overview, Γ_{Fifth} is displayed for the first movements of all Beethoven sonatas which are composed in sonata form in the lower figure. The features for the individual parts are calculated on a 100 s level (0.01 Hz) and averaged.

Yet, a general trend towards higher complexity is not confirmed, in accordance with the results of [1].

5. CONCLUSION

In this paper, we have presented novel features to quantify the complexity of music regarding tonality. A set of assumptions has been compiled to define requirements to the features' characteristics. In a study with ideal chord templates as well as recorded piano chords, these assumption have been tested on the fine level. Hierarchical visualizations of complexity values for movements of Beethoven's sonatas show the features' capability to capture the structure of the sonata form. Development parts and transition phases between themes show a higher complexity, in general. This could be confirmed for most of the Beethoven sonatas' first movements.

6. ACKNOWLEDGEMENTS

This work has been supported by the German Research Foundation (DFG MU 2686/5-1). The International Audio Laboratories Erlangen are a joint institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Fraunhofer IIS. C. W. has been supported by the Foundation of German Business.

REFERENCES

- [1] D. Perttu: *A Quantitative Study of Chromaticism: Changes Observed in Historical Eras and Individual Composers*. In *Empirical Musicology Review*, volume 2(2):47–54, 2007.
- [2] A. Honingh and R. Bod: *Pitch class set categories as analysis tools for degrees of tonality*. In *Proc. of the 11th Int. Society for Music Information Retrieval Conf. (ISMIR)*, pages 459–464, 2010.
- [3] P. v. Kranenburg and E. Baker: *Musical style recognition - a quantitative approach*. In *Proc. of the Conf. on Interdisciplinary Musicology (CIM)*, pages 106–107, 2004.
- [4] M. Mauch and M. Levy: *Structural Change on Multiple Time Scales as a Correlate of Musical Complexity*. In *Proc. of the 12th Int. Society for Music Information Retrieval Conf. (ISMIR)*, 2011.
- [5] S. Streich: *Music Complexity: A Multi-Faceted Description of Audio Content*. Ph.D. thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2006.
- [6] M. Müller and S. Ewert: *Chroma Toolbox: MATLAB implementations for extracting variants of chroma-based audio features*. In *Proc. of the 12th Int. Society for Music Information Retrieval Conf. (ISMIR)*, pages 215–220, 2011.
- [7] E. Gómez: *Tonal Description of Music Audio Signals*. Ph.D. thesis, Universitat Pompeu Fabra, Barcelona, Spain, 2006.
- [8] M. Mauch and S. Dixon: *Approximate note transcription for the improved identification of difficult chords*. In *Proc. of the 11th Int. Society for Music Information Retrieval Conf. (ISMIR)*, pages 135–140, 2010.
- [9] N. Jiang, P. Grosche, V. Konz, and M. Müller: *Analyzing Chroma Feature Types for Automated Chord Recognition*. In *Proc. of the 42nd AES Int. Conf. on Semantic Audio*, pages 285–294, 2011.
- [10] C. S. Sapp: *Visual Hierarchical Key Analysis*. In *ACM Computers in Entertainment*, volume 3(4):1–19, 2005.
- [11] N. Jiang and M. Müller: *Automated methods for analyzing music recordings in sonata form*. In *Proc. of the 14th Int. Society for Music Information Retrieval Conf. (ISMIR)*, pages 595–600, 2013.