HIGH-LEVEL CHORD FEATURES EXTRACTED FROM AUDIO CAN PREDICT PERCEIVED MUSICAL EXPRESSION

Jochen Steffens [1] **Steffen Lepa**[1] **Martin Herzog**[1]
Andreas Schönrock[1] **Geoffroy Peeters**[2] **Hauke Egermann**[3]
[1] Audio Communication Group, TU Berlin, Germany
[2] UMR STMS (IRCAM-CNRS-UPMC), Paris, France
[3] York Music Psychology Group, University of York, UK
jochen.steffens@tu-berlin.de

ABSTRACT

We investigated the relationship between high-level chord features and the perceived semantic and emotional expression of musical pieces in the context of music branding. Therefore, we first developed high-level chord features based on musicological considerations and novel MIR technologies. Inter alia, these features represent the number of chords, the proportion of major/minor chords, and the frequency of certain cadences and turnarounds. The validity of these features for predicting listeners' perceived musical expression beyond genre information was subsequently tested by means of data from two online listening experiments, where musical expression of 549 music titles had been rated on four factors, *Easy-going*, *Joyful*, *Authentic*, and *Progressive*. Results show that in all four models chord features significantly improved prediction results. Most important features turned out to be those representing the number of (unique) chords and the proportion of minor chords. Implications of results are discussed, and future work is outlined.

1. INTRODUCTION

Harmony constitutes one of the essential elements in Western tonal music. For instance, composers typically use specific harmonic progressions to express and induce particular emotional responses and to convey intended meanings. A canonical example for a relationship between harmony and expressed emotion is a minor chord being associated with sadness. While research in musicology has studied the influence of harmonic progression on emotion and meaning [1], MIR has predominantly been using chord information for genre classification [2] or cover song identification [3]. This paper thus bridges the gap between both disciplines and presents a specific use case, namely the prediction of perceived musical expression through high-level chord features in the context of music branding (ABC_ DJ project; http://abcdj.eu). In particular, we aim on demonstrating that chord features can not only help to predict genre membership, but are also able to explain perceived musical expression beyond genre information. Therefore, we first developed high-level chord features based on musicological considerations and novel MIR technologies and then tested their predictive capability by means of data from two online listening experiments.

2. DEVELOPMENT OF NOVEL CHORD FEATURES

High-level chord features were developed based on the chord progression extracted by the IRCAMchord algorithm [4] and the key/mode estimation using the IRCAMkeymode algorithm [5]. A first group of features deals with the number of chords in a certain segment. The feature *chords_total* is defined as the total number of chords divided by the track duration (in seconds) while *chords_unique* is defined as the number of unique chords divided by the track duration. In addition, the feature *chords_func* is defined as the number of functional (with regard to the locally estimated key/mode) chords divided by the total number of chords. These three features were established as measures for harmonic complexity. A measure to potentially estimate the harmonic tension of a musical piece is *chords_until_tonic* which is defined as the average number of chord changes until the next tonic occurs. Lastly, *chords_minor* and *chords_major* are defined as the number of minor and major chords respectively, divided by the total number of chords. A second group of high-level chord features focused on detecting specific cadences and turnarounds in a music track. For instance, *Authenticad* describes the number of *Authentic* cadences (i.e. V-I chord progressions), in a segment, and *turnaround_Blues* denominates the typical basic Blues chord progression (I-IV-I-V-IV-I).

3. METHOD

To validate the chord features developed with regard to their predictive capability in terms of listeners' perceived musical expression, two online experiments were conducted, with a total of 10.047 participants (49.9% female) from three different countries (UK, Spain, Germany), three different age cohorts, three different educational backgrounds, and both genders (country-wise crossed-quotas).

Participants were requested to rate four (study part 1) or six (study part 2) randomly assigned 30-seconds music excerpts stemming from a pool of 549 music titles, representing 10 different genres and 61 sub-genres. They rated the excerpts by means of the self-developed *General Music Branding Inventory* (GMBI; [6]) measuring the perceived musical expression in the branding context and comprising the four orthogonal factors *Easy-going*, *Joyful*, *Authentic*, and *Progressive*. To measure the influence of harmonic properties of the music titles on these dimensions, four linear regression models were estimated with the high-level chord features as independent variables and the GMBI factor scores averaged across participants as dependent variables. To investigate the additional explanatory gain of chord features beyond genre information, genre memberships tagged by experts were used as control variables.

A two-step procedure was chosen to identify important chord features and to detect interaction effects between features and genre. In a first step, stepwise regressions were computed where all nine dummy-coded genre variables (redundant category: world music) were entered at once in a first block, before gradually adding the chord features. After that, final general linear models were estimated with the previously identified significant chord features and the genre tags. Additionally, all potential interaction terms between the remaining chord features and genre tags were computed and tested for significance ($\alpha = .05$).

4. RESULTS

The stepwise regression revealed three chord features to be significantly associated (all ps $< .01$) with the perceived *Easy-goingness* of the music tracks, namely *chords_unique* (standardized regression coeff. $\beta = 0.102$), *chords_func* ($\beta = 0.122$), and *chords_minor* ($\beta = 0.147$). The subsequent general linear model revealed two significant interaction effects, *chords_unique X Jazz* ($\beta = 0.373$, $p < .05$) and *chords_func X Folk* ($\beta =$ -0.552, $p < .05$). The model explains 32.8% of the variance (${R^2}_{\text{adj}} = .261$). Here, chord features explain 4.6%, and the inclusion of interaction effects increased the explained variance by 0.9%. For the second factor, *Joyful*, the stepwise regression showed that also three chord features significantly predicted (all ps $< .01$) the ratings of the participants, namely *chords_total* ($\beta = 0.208$), *chords_unique* ($\beta =$ -0.166), and *chords_minor* ($\beta =$ -0.156). The subsequent linear model revealed no significant interaction effects. The model explains 23.7% of the variance (${R^2}_{\text{adj}} = .220$), while the contribution of chord features is 6.1%. Regarding the perceived *Authenticity* of a music track, the stepwise procedure only revealed *chords_unique* ($\beta = 0.193$, $p < .01$) to be a significant predictor of the perceived music expression. In this case, no significant interaction effects were found either. This linear model explains 38.1% of the variance (${R^2}_{\text{adj}} = .369$), and the contribution of *chords_unique* is 3.5%. Lastly, the two predictors *chords_unique* ($\beta =$ -0.132), and *chords_minor* ($\beta = 0.086$) were shown to significantly predict the perceived *Progressiveness* of a track (both ps $< .01$). Again, no interaction effects were found. The model explains 50.4% of the variance (${R^2}_{\text{adj}} = .494$). Here, chord features only explain 1.7% of the variance.

5. DISCUSSION

Results show that chord features significantly contributed to the prediction of perceived musical expression on all of its four dimensions. While the absolute values of explained variance (4.1% on average) may appear rather low, it has to be considered that this quantifies the additional explanatory gain above a very high-level and holistic music feature such as genre membership that can itself be predicted by chord features [2]. Therefore, we consider the developed chord features to provide a valuable additional set of predictors for diverse MIR scenarios such as music branding. The most important chord features, occurring in at least 50% of the models, were *chords_unique*, *chords_total*, and *chords_minor*. *chords_unique* was included in all four models and can be regarded as an adequate feature for measuring harmonic complexity. For example, the *Joyfulness* of a song was positively predicted by *chords_total*, but negatively predicted by *chords_unique*. This implies that a music title was regarded joyful when a lot of harmony changes occurred, but only if these changes led back to only a small set of different harmonies. Except for the model predicting the factor *Easy-going*, no interaction effects between chord features and genre were observed. Thus, one might conclude that the found effects of chord progressions on perceived musical expression are stable across different genres. The promising results inspire the development of further novel features detecting bass notes as a measure of harmonic stability and additional notes (e.g. sixths, ninths) as a measure of harmonic colorfulness.

Acknowledgements: This research has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 688122.

6. REFERENCES

[1] L. Meyer, Emotion and Meaning in Music. Chicago: University of Chicago Press, 1956.

[2] A. Anglade, R. Ramirez, and S. Dixon, Genre classification using harmony rules induced from automatic chord transcriptions, in Proc. of ISMIR (International Society for Music Information Retrieval), Kobe, Japan, 2009, pp. 669-674.

[3] J. Serra, E. Gmez, and P. Herrera, Audio Cover Song Identification and Similarity: Background, Approaches, Evaluation, and Beyond, Advances in Music Information Retrieval, vol. 274, pp. 307-332, 2010.

[4] H. Papadopoulos and G. Peeters, Joint Estimation of Chords and Downbeats From an Audio Signal, IEEE Trans. Audio Speech Lang. Process., vol. 19, no. 1, pp. 138-152, 2011.

[5] G. Peeters, Chroma-based estimation of musical key from audio-signal analysis. In Proc. of ISMIR, Victoria, BC, Canada, 2006, pp. 115-120.

[6] H. Egermann, S. Lepa, A. Schönrock, M. Herzog, and J. Steffens, Development and Evaluation of a General Attribute Inventory for Music in Branding, in Proc. of the 25th Anniversary Conference of the European Society for the Cognitive Sciences of Music (ESCOM), Ghent, Belgium, 2017.