Chord Label Personalization through Deep Learning of Integrated Harmonic Interval-based Representations

Hendrik Vincent Koops[1], W. Bas de Haas[2], Jeroen Bransen[2], and Anja Volk[1]

[1] Department of Information and Computing Sciences,
Utrecht University, Utrecht, the Netherlands
h.v.koops@uu.nl, a.volk@uu.nl
[2] Chordify, Utrecht, the Netherlands
bas@chordify.net, jeroen@chordify.net

Abstract

Harmony annotations are at the core of a wide range of studies in music information retrieval and Automatic Chord Estimation (ACE) in particular. Nevertheless, annotator subjectivity makes it hard to derive one-size-fits-all chord labels. Annotators transcribing chords from a recording by ear can disagree because of personal preference, bias towards a particular instrument, and because harmony can be ambiguous perceptually as well as theoretically by definition. These reasons contributed to annotators creating large amounts of heterogeneous chord label reference annotations. For example, on-line repositories for popular songs (e.g. Ultimate Guitar, Chordify) often contain multiple, conflicting versions. Approaches that aim to integrate conflicting versions can be used to create a unified view that can outperform individual sources. Nevertheless, this approach is built on the intuition that one single correct annotation exists that is best for everybody, on which ACE systems are almost exclusively trained and evaluated.

In this paper, we propose a first solution to the problem of finding appropriate chord labels in multiple, subjective heterogeneous reference annotations for the same song. We propose an automatic audio chord label estimation and personalization technique using the harmonic content shared between annotators. We create an harmonic bird's-eye view from different reference annotations, by integrating their chord labels at the level of harmonic intervals. More specifically, we introduce a new feature that captures the shared harmonic interval profile of multiple chord labels, which we deep learn from audio. First, we extract Constant Q (CQT) features from audio, then we calculate Shared Harmonic Interval Profile (SHIP) features from multiple chord label reference annotations corresponding to the CQT frames. Finally, we train a deep neural network to associate a context window of CQT to SHIP features. From the deep learned shared harmonic interval profiles, we can create chord labels that match a particular *annotator vocabulary*, thereby providing an annotator with familiar, and personal chord labels.

We test our approach on a 20-song dataset with multiple reference annotations, created by annotators who use different chord label vocabularies. In an

experiment we compare training of our chord label personalization system on multiple reference annotations with training on a commonly used single reference annotation. In the first case we train a DNN on SHIPs derived from a dataset containing 20 popular songs annotated by five annotators with varying degrees of musical proficiency. In the second case, we train a DNN on the *Isophonics* single reference annotation. *Isophonics* is a peer-reviewed, and *de facto* standard training reference annotation used in numerous ACE systems.

We show that by taking into account annotator subjectivity, our system is able to personalize chord labels from multiple reference annotations. Comparable high accuracy scores for each annotator show that the model is able to learn a SHIP representation that is meaningful for all annotators, and from which chord labels can be accurately personalized for each annotator. Furthermore, our results show that personalization using a commonly used single reference annotation yields significantly worse results. From the results presented in this paper, we believe chord label personalization is the next step in the evolution of ACE systems.