
Finding two-level structure in field recordings of folk music

Ciril Bohak and Matija Marolt

University of Ljubljana, Faculty of Computer and Information Science
{ciril.bohak, matija.marolt}@fri.uni-lj.si

Abstract. Ethnomusicological field recordings are audio recordings of folk music that also contain interviews with performers. Such field recordings may contain several hours of recorded audio that consists of different content such as speech, singing, instrumental music, event recordings etc. It is easy for a human to split field recordings into meaningful parts (units) according to their content types, such as songs, interviews as well as stanzas and verses. Manual segmentation is a very time consuming task for large archives of field recordings.

In this paper we present a two-level segmentation algorithm. The algorithm segments a recording on two levels: on the first (high) level, a recording is split into units, representing parts of the same content type (e.g. singing, speech), while on the second (low) level, high level parts containing singing and instrumental music are split into individual repeating parts (stanzas). We first test and compare different algorithms for high and low level segmentation on a collection of field recordings of Slovenian folk music. Next we present a novel approach, which combines high and low level segmentation. A probabilistic model is used for high level segmentation, while for low level segmentation a combination of vocal pauses detection and self-similarity of chroma vectors is used. The algorithm makes use of the low level boundaries between stanzas to improve estimation of high level boundaries between different units. We evaluate the performance of the proposed algorithm on a collection of field recordings and show that the two-level approach outperforms the high level one.

References

- MAROLT M. (2009): Probabilistic segmentation and labeling of ethnomusicological field recordings. *In Proceedings of ISMIR 2009, Kobe, Japan, 75-80.*
- MÜLLER M., GRÖSCHE P., WIERING F. (2009): Robust segmentation and annotation of folk song recordings. *In Proceedings of ISMIR 2009, Kobe, Japan, 735-740.*

Keywords

FOLK MUSIC, SEGMENTATION, MUSIC STRUCTURE EXTRACTION