

## **Summary of the Research Proposal**

### **Music Structure Extraction for Music Summarization**

With the recent explosion in the quantity of digital audio libraries and databases, content descriptions of standard metadata, such as MPEG-7, play an important role in efficiently managing and retrieving audio files. As repetition and transformations of music structure have created a unique identity of music itself, it is believed that music structural descriptions (which subsume temporal, harmonic, rhythmic, melodic, polyphonic, motivic and textual information) can provide better quality access and a powerful way of interacting with audio content. Moreover, it can also serve as a linkage between low-level descriptions and higher-level descriptions of audio (e.g. audio summarization, audio fingerprinting, etc.), which would be very useful and applicable for music commercials and the movie industry.

My research aims at studying methods in detecting and extracting musical structure descriptors from music audio pieces, with the ultimate goal of building a short summary from a long music audio piece. In order to build a short summary out of a long music piece, objective evaluation of the significance of musical structure will be studied.

Current research in this area has been mostly focusing on a few specific low-level features attributes, such as MFCC, pitch-chroma, etc. in describing the musical content of sound signal (M. Goto, 2003; B. Logan and S. Chu, 2000; M. Cooper, and J. Foote, 2002; Xu, Y. Zhu, and Q. Tian, 2002, G. Petter, A.L. Burthe, X. Roder, 2002). Hence in our proposed research, we will explore and integrate some other low-level features attributes, which have been previously disregarded, for better description of musical content of audio signal.

Signal segmentation facilitates partitioning audio streams into short regions for further analysis. Finding appropriate boundary truncations is an indispensable process for certain content-based applications. Instead of depending on only low-level features descriptors, our research approach will make use of higher level analysis technique, such as beat detection, phrase detection, etc. in order to achieve better segment truncation.

From the theoretical point of view, my research will help in increasing our knowledge of music structure. From the practical point of view, it will not only provide an powerful way of interacting with audio content but also facilitate in time-saving browsing and assessing of music audio pieces.

References:

M. Goto, "A Chorus-Section Detecting Method for Musical Audio Signals," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing* vol. V, pp. 437-440 (2003).

Logan and S. Chu, "Music summarization using key phrases," in *International Conference on Acoustics, Speech and Signal Processing* pp. 7] B. Logan and S. Chu, "Music summarization using key phrases," in *International Conference on Acoustics, Speech and Signal Processing* pp. II-749-752 (2000).

M. Cooper, and J. Foote, "Automatic Music Summarization via Similarity Analysis," in *Proceedings of ISMIR* pp. 81-85 (2002).

Xu, Y. Zhu, and Q. Tian, "Automatic Music Summarization Based on Temporal, Spectral and Cepstral Features," in *Proceedings of IEEE International Conference on Multimedia and Expo* pp. 117-120 (2002).

G. Peeters, A.L. Burthe, X. Roder, "Toward Automatic Music Audio Summary Generation from Signal Analysis," in *Proceedings of ISMIR: Third International Conference on Music Information Retrieval* pp. 94-100 (2002).