

# INTERACTIVE CONTROLLER FOR AUDIO OBJECT LOCALIZATION AND AUTOMATIC THUMBNAIL MUSIC GENERATOR

Noriyoshi Kamado, Hiroyuki Nawata, Hiroshi Saruwatari and Kiyohiro Shikano  
Nara Institute of Science and Technology, Japan (e-mail: noriyoshi-k@is.naist.jp)

## 1. INTERACTIVE CONTROLLER FOR AUDIO OBJECT LOCALIZATION

In this demonstration, we provide a new interactive controller for audio object localization and an automatic thumbnail music generator for an arbitrary stereo mixed source. Those two products are based on the same audio coding framework enabling the extraction and localization operation of audio objects (e.g., vocal, guitar, drums) via the temporal quantization of spatial information [1].

In the last decade, the reproduction of music with a high sense of reality has been superseded with the development of sound reproduction systems with a *creative sense of reality* in which a user can operate each sound object individually in the audio signal.

Therefore, we proposed an audio object controller developed for interactive sound field reproduction using audio object localization based on spatial representative vector operations. This system allows a listener to operate an initial sound field and build a new virtual sound field by performing sound field operations on a stereo mixed source. Figure 1 shows the controller which is equipped with a capacitive touchscreen panel, and the listener can intuitively operate every audio object displayed on the touchscreen panel with his/her finger or a touch pen. The results of the subjective assessment clarified that this controller enables the listener to change the localization of audio objects without sound degradation if the gain operation is not extreme.

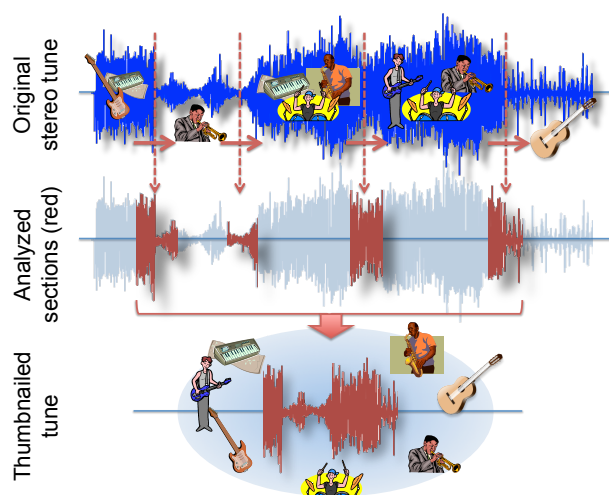
## 2. THUMBNAIL MUSIC GENERATOR

We are now able to obtain music easily over the network, regardless of time and place. On the other hand, it is complicated to search an objective music tune from a huge number of music files. Therefore, the aim of our study is to realize the system which allows us to preview thumbnail music that has only a main part of the original music tune. By previewing thumbnail music, we can comprehend the image of the tune without listening to it all over and judge easily whether the original music tune is the objective one. However, thumbnail music is generated manually now and it is difficult to generate a huge number of them manually, so the automatic music thumbnailing becomes an essential problem.

In order to automatically generate thumbnail music, we propose a new estimation method of structure changes in stereo tunes based on audio object localization and evaluated the effectiveness by the objective assessment. As a result, the proposed method can detect about 70 percent of the correct answers which are manually given. Conse-



**Figure 1.** Interactive audio object controller. This controller is demonstrated at the following URL: <http://spalab.naist.jp/aocdemo/>.



**Figure 2.** Overview of proposed music thumbnailing method.

quently, it was shown that the proposed method is effective in the music structure analysis for generating thumbnail music automatically.

## 3. ACKNOWLEDGMENT

This work was partly supported by the MIC Strategic Information and Communications R&D Promotion Programme in Japan.

## 4. REFERENCES

- [1] S. Miyabe, K. Masatoki, H. Saruwatari, K. Shikano, and T. Nomura, "Temporal quantization of spatial information using directional clustering for multichannel audio coding," Proc. of WASPAA, pp.261–264, 2009.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

© 2010 International Society for Music Information Retrieval.