Augmented Music-Understanding Interfaces

Masataka Goto

National Institute of Advanced Industrial Science and Technology (AIST)
1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan

1 Introduction

One of my research goals is to enrich music listening experiences by deepening each person's understanding of music. Music listening experiences depend on music-understanding abilities. Although the music-composing/performing abilities of musicians are often discussed, the music-understanding abilities of casual listeners have not been well discussed or studied. It is difficult, for example, to define music-understanding abilities and express the results of understanding. In addition, it is difficult to know how others understand music. Music listening is usually an individual experience, and it is impossible to directly observe how others understand elements in music. Similarly, it is common to not notice what one does not understand when listening to music. Because of this, even if listeners want to better understand music or want to improve their ability to understand music, methods to realize those wishes have not yet been established and will have to be discovered.

I therefore propose a research approach *"Augmented Music-Understanding Interfaces* that facilitates deeper understanding of music by using automatic music-understanding technologies based on signal processing. First, *visualization of music* plays an important role in augmenting people's understanding of music because understanding is deepened through seeing. By recognizing and visualizing elements in music, we let end users (listeners) without expertise understand the existence of elements and the relationships between elements. Second, *"music touch-up"* (small modifications) also helps each person's understanding of musical audio signals because understanding is deepened through editing. Music touch-up is a term I coined [1] that means making small changes to elements in existing music. It lets users naturally observe why music is composed and arranged in a certain way while casually enjoying the content modification (i.e., personalization or customization).

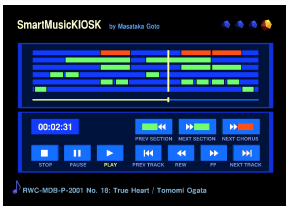
2 Three Examples of Augmented Music-Understanding Interfaces

This section introduces three examples of Augmented Music-Understanding Interfaces based on automatic music-understanding technologies. Originally, these were developed as *Active Music Listening Interfaces* [1], but I later found that they provide good examples of Augmented Music-Understanding Interfaces as follows.

SmartMusicKIOSK (visualization of music structure) [2]: This is a content-based playback-control interface for within-song browsing for popular music. A user can skim rapidly through a musical piece by easily skipping sections of no interest while viewing a "music map" (Figure 1). The music map is a visual representation of the entire song structure consisting of chorus sections (the top row) and repeated sections (the five lower rows). On each row, colored sections indicate similar (repeated) sections. This facilitates deeper understanding of music structure through music listening that focuses on it and better understanding of musical changes during repetition.

LyricSynchronizer (visualization of lyrics) [3]: This is an interface that displays scrolling lyrics with the phrase currently being sung highlighted during playback of a song. Because the lyrics are automatically synchronized with the song, a user can easily follow the current playback position and click on a word in the lyrics to listen to it. This facilitates deeper understanding of lyrics through music listening that focuses on lyrics and better understanding of messages in lyrics.

Drumix (music touch-up for drums) [4]: This is an interface for playing back a musical piece with drums as if another drummer were performing. With a real-time drum-part editing function, a user can change the volume and timbre of bass and snare drum sounds and rearrange the drum patterns of these sounds during playback. The user can casually switch drum sounds and drum patterns as the urge arises during music playback in real time. This facilitates deeper understanding of drum sounds through music listening that distinguishes between them and better understanding of how drum sounds and drum patterns can change the feeling of music.



A user can actively listen to various parts of a song while moving back and forth as desired on the visualized song structure (the "music map" in the upper window).

Figure 1: SmartMusicKIOSK [2].

3 Lessons Learned

Our experiences in designing the above interfaces revealed that the following functions are important:

- Visualization of the content of a musical piece,
- Synchronization of the visualization with music playback, and
- Provision of interactive interfaces for customizing and controlling music playback.

We also found that interfaces able to make users aware of what they tend to not notice are useful. User interaction enhances immersive music listening experiences and promotes a deep, full appreciation of music.

4 Future Direction

Since Augmented Music-Understanding Interfaces temporarily support people's understanding of music, the next step will be to develop interfaces that permanently improve one's ability to understand music. Although there are many ways to improve one's ability to understand a foreign language, such as through language schools and training materials, there are virtually no systematic means to improve music-understanding abilities. Most existing music schools and training materials including music-dictation training are intended for musicians and creators, not for casual listeners. I therefore propose a research approach *"Music-Understanding Ability Training Interfaces"* as an important future direction towards enabling a greater number of people to enjoy music in more depth from more diverse views.

ACKNOWLEDGMENTS: I thank Hiromasa Fujihara, Kazuyoshi Yoshii, and Hiroshi G. Okuno who built LyricSynchronizer and Drumix with me. This work was supported in part by CrestMuse, CREST, JST.

REFERENCES

[1] M. Goto, "Active music listening interfaces based on signal processing," in *Proc. of ICASSP 2007*, 2007.

[2] M. Goto, "A chorus-section detection method for musical audio signals and its application to a music listening station," *IEEE Trans. on ASLP*, vol. 14, no. 5, pp. 1783–1794, 2006.

[3] H. Fujihara, M. Goto, J. Ogata, K. Komatani, T. Ogata, and H. G. Okuno, "Automatic synchronization between lyrics and music CD recordings based on Viterbi alignment of segregated vocal signals," in *Proc. of ISM 2006*, pp. 257–264, 2006.

[4] K. Yoshii, M. Goto, K. Komatani, T. Ogata, and H. G. Okuno, "Drumix: An audio player with functions of realtime drum-part rearrangement for active music listening," *Trans. of Information Processing Society of Japan*, vol. 48, no. 3, 2007.