Interactive Poster: 3D Visualization of MIDI Dataset

Reiko Miyazaki       Issei Fujishiro
Graduate School of Humanities and Sciences, Ochanomizu University, Japan
reiko@inv.is.ocha.ac.jp; fuji@is.ocha.ac.jp

1. Background and Purpose
When editing and playing a MIDI dataset, users usually rely on so-called sequence software systems, which offer a variety of editwindows, such as staff and piano-roll, for different tasks. Since a limited number of related parameters appear in each of the editwindows, the users are forced to open multiple windows of small sizes side by side, and thus making it difficult for them to grasp the global musical structures embedded in a multi-channel sequence of note events. In this poster, we present a system, called comp-i (Comprehensible MIDI Player – Interactive), which provides a designated 3D virtual space, where the users are allowed to perform visual exploration of a given MIDI dataset in an immersive and intuitive manner. Although a few studies on 3D music visualization can be found in the literature (e.g.,[1],[2]), the existing systems offer restricted functionalities of overview and detailed analysis. We have carefully designed a spatial substrate and retinal properties for the comp-i virtual space, along with a rich set of operations according to Shneiderman’s Visual Information Seeking Mantra [3].

2. System Overview
The comp-i system accepts a standard MIDI file (SMF), and selects from the file, three primary MIDI elements, namely, note-on, note-off and set-tempo, to construct a corresponding 3D virtual space. Note-on and note-off are fundamental MIDI messages to start and stop a single note sound, respectively, and include channel, pitch, and sound volume parameters. Set-tempo is a MIDI event that sets the tempo parameter.

Fig.1 shows a complete view of the comp-i virtual space, where multiple channel layers are stacked along the depth axis. Each layer contains point nodes for bar lines, and a series of cylinders each of which corresponds to a single note sound. For visibility and distinction, cylinders on different layers are assigned different semitranslucent colors. The three parameters pitch, volume, and tempo of a note sound are encoded as the height, diameter, and color saturation of the corresponding cylinder, respectively. As a given MIDI dataset is played through a 3D virtual sound device, a scan-plane orthogonal to the layers is moving from left to right to indicate the current position of the note sounds.

The comp-i system offers one or more operations for each of Shneiderman’s four major information seeking task categories [3]. Indeed, the users can take advantage of perspective view to grasp the entire dataset within the virtual space. 3D illumination gives object shadows on the floor, and conveys the right information of MIDI object geometry (overview), whereas the users can make their viewpoints as close to selected objects as they like (zoom). The users are also allowed to permute the channel layers of interest, control the visibility of retinal/auditory properties of the objects, and choose arbitrary positions to start/stop playing (filter). Furthermore, the scales and quantitative properties are possible to be displayed along with the channel layers and the scan-plane (Details-on-Demand) (Fig.2).

After the users comprehend the outline of a given MIDI dataset, they are allowed to alter the projection to orthographical for accurate editing work. The system provides two ways to edit the MIDI dataset: direct manipulation of objects for novice users and textbox-based specification for experts. Choice of particular projections of the virtual space supersedes typical GUIs of sequence software systems. For example, the top orthographical view can give a clear view of multi-channel information involving both volume and tempo (Fig.3 (a)-(a’)). The side view of the virtual space makes it possible for the users to look over both pitch and volume of multiple channels simultaneously (Fig.3 (b)-(b’)). These visualizations cannot be obtained easily through the traditional sequence software systems.

3. Towards Visualizing Musical Structure
We are still extending the comp-i system so that it can provide the users with a focus+context display mechanism to perform visual MIDI data mining. A circular form is a good candidate for a new spatial substrate to visualize the global structure of a long MIDI sequence effectively. In addition, a ConeTrees[4]-based multi-resolution technique is adopted to support interactive similarity search reflecting the hierarchical structure of music pieces (e.g., sentence, phrase, and motif) (Fig.4).

References
Figure 1: Overview of comp-i virtual space.

Figure 2: Information seeking task categories supported by comp-i.

Figure 3: Choice of particular projection of the comp-i virtual space supersedes typical GUIs of a sequence software ((a)(b): sequence software, (a')(b'): comp-i).

Figure 4: Exploiting musical structure within ConeTrees-based spatial substrate.