ABSTRACT

For most of the time, we enjoy and appreciate music performances as they are. Once we try to understand the performance not in subjective terms but in an objective way and share it with other people, visualizing the performance parameters is indispensable. In this paper, a figure for visualizing performance expressions is described. This figure helps people understand the cause and position of the performance expression as it has expressive cues, which coincide with the cognitive meaning of musical performance, and not by using only MIDI parameter values. The differences we hear between performances are clarified by visualized figures.

CR Categories: H.5.5 [Sound and Music Computing]: Systems; J.5 [Arts and Humanities]: Performing arts

Keywords: Music Performance, Expressive Cue, Performance Visualization, Understanding Performance

1 INTRODUCTION

With performed music the resulting sounds are interpretation of a printed notation. Musical scores have been the basis of performances. Besides the usual staff notation, Karkoschka has introduced several new score types—some resemble a contour map, some like planisphere—which are used by modern composers [8]. In his book, he says that music can even be performed from Kandinsky’s paintings.

If we want to understand the difference between more than two performances or why a note seems to have more impact, and share a clear understanding of the performance with other people on an objective basis, we need scientific rather than artistic performance visualization, since a picture is worth a thousand played notes and musical reviews.

Visualizing performance realizes the conversion from auditory to visual media. In this conversion, some musical information can be described by two different types of representations, while other information only by one. The purpose of performance-visualization research in general is to find the maximum overlap of performance and visualization data. The sharable information between auditory and visual media is different depending on which abstract level the musical information is described. The shared information in Figure 1 is described using MIDI1.

In this paper, we propose a figure which visualizes a melody of a piano performance with its expressive cues (tempo, articulation, and sound level) that represents the musical meaning instead using MIDI parameter values. With this type of figure, especially people who are not well-trained in music can understand the cause and effect of a performance expression clearly. We think this figure is usable in visual data mining of musical sound database. It also helps music novices to create a performance of their preference by giving them a visual music environment.

In Section 2, we classify the music visualization and describe works related to this research. We refer to issues particular to music visualizations. In Section 3, a case of visualizing performance expressions is described. Finally, in Section 4, we summarize the current study and describe our future plans for using the visualization of performance expressions in actual applications.

2 MUSIC VISUALIZATION

While “music visualization” is restricted to visualizing music with computer graphics for non-artistic purposes, visualizing sound is already indispensable in research and applications. Besides drawing sound as waves, Pickover has developed an interesting and impressively symmetrical figure [12]. Sobieczky visualized the consonance of a chord on the diagram based on a roughness curve [15].

User interface of Desk Top Music (DTM) software provides users with several types of visualized music figures. Usually, sequence software includes a staff notation and a piano-roll figure, as well as a table of MIDI events. Some introductory DTMs show icons of musical instruments and their players.

2.1 Related works

Although visualization has not been studied for music by many researchers so far, two types have been developed.

1. Augmented score

Since conventional staff notation does not precisely represent the composers’ intentions, Oppenheim proposed a tool for including it [11]. Kunze proposed several three-dimensional figures for composers to use as a graphical tool when they write their music [9]. Watanabe proposed a virtual score with focus+context to support music learning [6].

2. Performance visualization

---

Hiraga proposed simple figures to help analyzing musical performances [5]. Smith proposed mapping 3D graphics based on MIDI parameters [14]. Foote’s checkerboard-type figure [3] shows the resemblance between played notes by using musical-performance data. Miyazaki’s comp-i enables users to generate music of their own by using a rich set of functions with a 3D performance-visualization interface [10].

Hereafter in this paper, we use the term music visualization in a narrower sense meaning using computer graphics to depict music excluding music visualization for artistic purposes. Namely we restrict music visualization to visualize music data on a computer display for a scientific purpose, a method to aid musicians to create their music, or a method to help people understanding music with the visual aids. Above all, our research visualizes performance expression.

2.2 Problems of visualizing music

Three issues need to be solved in our visualization of performance expressions.

1. Which abstract level of music should be used to visualize an expression.

Sequence software handles musical performance on a computer using MIDI base. Basic events of MIDI are Note On and Note Off which represent the start and end of a played note, respectively, and they have parameters for the pitch (which note to play) and the velocity (the volume of the note) in 128 levels along with the time interval from the previous MIDI event. MIDI based visualization is summarized as follows:

   - Only MIDI parameter values are shown.
   - Parameter values are independent and have no relationships with parameters of other MIDI events.
   - Musical structure is not considered.

Thus even if we hear accelerando in a phrase, we cannot find it easily by looking at onset time of each note. The impossibility of handling musical structure, such as a triplet or a measure, to say nothing of a motif or phrase, that we can understand by looking at scores, prevents users creating and judging a musical performance.

Despite MIDI data being not suitable for visualizing performance expressions, we do not know visualized figures based on the different abstract level of music.

2. A small amount of data should represent an expression

A common problem in music-visualization systems, especially if composed music or MIDI data is manipulated, is, as Kunze indicated, the small amount of data as the object of visualization. This is completely the opposite problem in scientific visualization as well as most information visualization.

3. The perpetual problem

Developing a figure style suitable for visualizing all music and that satisfies auditory and visual understanding at the same time is an infinite effort.

3 VISUALIZING PERFORMANCE EXPRESSION

3.1 Expressive cues and Chernoff’s faces

After considering the music abstraction level in performance visualization, we decided on using expressive cues. They describe in qualitative music terminology (tempo, sound level, articulation, tone onsets and decays, timbre, I(I) deviations, vibrato, and final ritardando), which is quantifiable with the performance data. Gabrielson reported that listeners find basic emotions (anger, sadness, happiness, and fear) in human performances by intentionally controlling expressive cues [4]. Then, Bresin showed a makeup experiment with his performance rendering system [1].

We tried several figures to find a more suitable way to visualize expressive cues and came up with Chernoff’s face [2] as a candidate. Chernoff, a statistician, proposed using computer-drawn faces to visualize data in many dimensions. The advantages of using Chernoff’s faces for performance visualization are as follows:

   - Since human beings are familiar with faces, we can easily detect changes.
   - A line of faces can show a time sequence (the change of expression over time).
   - A face can express many dimensions (expressive cues) at once by changing parts of the face.

The third point, the spontaneous extension of facial dimensions is significant because performance visualization will include more and more expression cues in the future. Currently, 3D computer graphics cannot be extended intuitively to show many dimensions at a time.

The faces in Figure 2 are examples of representing expressive cues. We modify the shape of parts of the original Chernoff face and omit some of them (eyebrows, for example). Each face corresponds to a note whose tempo, articulation, and sound level are visualized. The performed note sequence is from left to right.

Local tempo, calculated as IOI normalized by a note value, is mapped in the position of the eyeballs. The eyes look back (Figure 2, left-hand (2)) if the tempo is slower and look forward (Figure 2, right-hand (4)) if it is faster. The more extreme the eye movements, the more the tempo deviates from the one specified on the score.

The contour of the face and the shape of the mouth both represent articulation. Articulation is calculated as the overlap/detachment of

\footnote{Inter Onset Interval. The interval between the two Note On MIDI events.}

\footnote{The length of a note specified on a score.}
two consecutive notes divided by a note value. When a note is played rather in *staccato* then the mouth shape is narrow (Figure 2 (2)), while in *legato* it is round (Figure 2 (4)).

In contrast to local tempo and articulation, the value of the sound level can be calculated in several ways. Though the sound level is described as the velocity value in MIDI parameters, the MIDI value does not correspond to a specific decibel. Currently, we express the sound level as the difference between two consecutive notes. If a note is played more strongly than the previous one, the nose is drawn as '⟨' (Figure 2 (2)), at the same velocity '⟩' (Figure 2 (3)), and softer '⟩' (Figure 2 (4)). Since the first note has no sound level to compare to, the first note is always '⟩'. Faces (1) and (3) represent notes that are played with fidelity along a score with local tempo and articulation.

The visualized performance in Figure 2 is artificially made. By merely looking at the figure we can see that the performance starts softer and lighter, and ends lengthy in this phrase. In this way, local tempo and articulation show the deviation of a performance from the score. Deviation is important not only in understanding and analyzing performance, but also for performance rendering by a computer.

### 3.2 Case: Visualizing three performances of Etude 10-3 by Chopin

We used three performances of Etude 10-3 by F. Chopin (subtitled “Sadness” and widely known as “Farewell” in Japan). We asked a pianist, Chica, to play two performances according to the interpretations of two editions—Cortot and Paderewski. Because space is limited, we can only show the first four measures of the melody (Figure 3). The slight difference of the editions gives musicians different feelings when they perform them. Chica finds that in the Cortot edition the sadness is like the end of “Gone with the wind”, while Paderewski is sadness in calm. The other performance is played by another pianist who uses the Zenon edition (a Japanese score maker), which is based on the Peters edition.

The visualization of the three performances are shown in Figure 4. Each face represents a performed note in the melody. The vertical bar shows the bar in the score. From the figure, we can see that the two performances by Chica vary in tempo and articulation more than the third performance, which is rather regular. The move of sound level is almost the same in these performances. These explanations by looking at figures fit the listener’s understanding.

The upper face line in Figure 5 (1) shows the difference of the Cortot performance (PC) from the Paderewski one (PP). Different from Figure 4, the nose shows the difference in velocity between the two performances. By listening to them, we can recognize that (1) PP is played more slowly, and that (2) C-sharp in the third measure, which is the highest and the most impressive note in these measures is a hesitation in PP. The figure showing the differences (Figure 5 (1)) indicates that the tempo differs regularly (PC is played faster). It also confirms the difference in expression of C-sharp in the third measure (fourth face), and can see that the difference lies in the articulation. In PC, G-sharp (the note prior to C-sharp) is played more lightly, or conversely, G-sharp is played longer and lengthy in PP. In the figure of difference, what we hear as “hesitation” is indicated at the position as the difference of articulation.

The lower face line in Figure 5 (2) shows the difference of PC from the Zenon performance (PZ). PZ sounds louder and faster, which in the figure is indicated by the direction of the noses and eyeballs. We can tell that the third note in the second measure (F-sharp) in PC is distinctively a light performance (shorter and softer).

### 3.3 Discussions

We showed the performance visualization in Figures 4 and 5 to two users.

One is an amateur pianist. She recognized the deviation of local tempo and articulation of the performance from the score. Since deviation is important for understanding and analyzing performances, as well as for a performance created by computer, she thought she could get hints from the figures to generate a performance either by using DTM software or a performance rendering system ⁴. She felt the change in sound level shown on the figure, which is almost the same as in the performances in Figure 4, was as she had heard. In the figure that shows difference (Figure 5), what she had heard as a “hesitation” was indicated at the position as a difference in articulation.

The other is also a music amateur who has studied music theory. He recognized the role of the musical notes and the musical struc-

---

⁴The output of a performance rendering system is a expressive performance as if it is played by a human player.
Figure 5: Performance differences. (1) Cortot – Paderewski, (2) Cortot – Zenon

ture. Though we visualized all notes uniformly in this figure, each note has a different role in a music. Depending on it can be visualized accordingly. Since musical structure is the nature of music, it will be a help to understand performance from visualization.

4 CONCLUSIONS AND FUTURE PLANS

We described visualizing performance expressions by using modified Chernoff faces. The performance expressions are not described in visualized by MIDI parameter values but by expressive cues that can explain them qualitatively musical meaning. The same shape can compare two performances. With the figures, we can clearly understand the tendency and slight differences of performances. This approach is a solution for handling the abstract level indicated in 2.2.

In the future we will focus on grouping information based on performance data, such as accelerando or diminuendo, clearly on the figure to enhance understanding of the performance. Another challenge is to introduce continuous data, such as that of the half-pedal on the discrete note icons. After all, since grouped information also describes continuous change over a range, this issue is an inevitable problem.

We plan to include performance visualization in our case-based performance rendering project HHH [7] so users can choose a case performance as the source of performance expression from a performance database at a glance and not through listening. This is the extraction of auditory data by using visual data mining. Another plan is to make use of the figure at the performance rendering concours Rencon [13]. Rencon is a project to pursue evaluation methods for performance rendering systems in the style of music contest where clarifying the elements to be reviewed and their standards is important. The visualization is useful for sharing common understanding and coming to an agreement in judging performances which are subjective in nature.

ACKNOWLEDGEMENT

This research is supported by The Ministry of Education, Culture, Sports, Science and Technology as Grant-in-Aid for Exploratory Research.

REFERENCES