ISE 599 paper review
Title: Visualizing Expressive Performance in Tempo-Loudness Space
Author: Jorg Langner and Werner Goebl
Reviewer: Jie Liu

This paper talked about a new method for an integrated display of tempo and loudness variations as measured in expressive music performance. This visualization technique includes data acquisition from both MIDI instruments and audio recordings, data reduction by smoothing measured performance data, and real-time animated display on computer screen. In this 2D interface, the horizontal axis is tempo and the vertical axis is loudness. A dot moves through the space, leaving behind it a trajectory. The authors showed that this trajectory may be interpreted as the intrinsic performance path of a particular performance.

In the introduction, the authors mentioned that a special focus was given to piano performance, because the expressive parameters are relatively few: tempo, loudness, articulation and pedaling. On the other hand, those parameters are somehow easy to obtain. In this paper, the tempo data is obtained from both MIDI file and audio recording. Here the authors mentioned a beat tracking software, proposed by Dixon, to analyze the audio data, find the onsets of musical notes and propose a possible beat track by displaying the beats as vertical lines over the amplitude envelope of the audio signal. I have no idea about this software, but I know beat tracking from audio is still a challenging problem currently. So I would not be surprised that error will be induced by using this software.

Then the authors talked about the smoothing technique. From the graphs shown in this paper, I found that the real tempo data was not smooth, although Cambouropoulos has mentioned that human prefer tempo smoothness. That should be one of the reasons why smoothing is needed in this paper, because the authors want to show a smooth trajectory later. However, I think the non-smoothness of the tempo shown in the graph could be explained by the fact that the sampling time is too big, which is about 1 second. I know it would be harder to get the tempo data if the sampling time is too small but I would like to see the results here. Also, I would like to see the trajectory if the tempo data is not smooth and see how different it is compared with the smooth data.

Then 2 case studies are described. In the first case study, the authors claimed that similar trajectories mean that the pianists sort of follow similar performing strategies. On the other hand, the shape of a trajectory is very much dependent on the window size of the smoothing window applied to the data. I think the effect of the window size is too much, because from the demo videos, we can clearly see the delay between the real maximum loudness point and the displayed point. Since this shifting effects could be removed by decreasing the window size, I would raise a question that what the optimal value of the smoothing window size is. Is it possible to use varied window size throughout the whole piece?

The second case study is like a proof. They chose the Shubert's G flat major impromptu as the testing piece, which is also the testing piece of many previous works, to show the tempo-loudness trajectory. From the result, Todd's statement, the faster the louder, the slower the softer is proven reasonable. However, the trajectory is not a perfect diagonal line as expected. The first reason is in this hybrid performance, different parameters for timing and intensity were chosen. The second reason is the linear relationship between timing and dynamics in terms of MIDI velocity units is not actually linear.

Finally, several conclusions are drawn and several questions are discussed. This interface does reveal the features of expressive performance, but in this paper, the number of tested performances is relatively small. The results for more performances are expected. Also, there is an interesting open question for the following four choices: Pianists tend (1) to get louder but not softer at tempo maxima, (2) to get slower but not faster at loudness maxima, (3) to get softer but not louder at tempo maxima, and (4) to get faster but not slower at loudness maxima. I would say that all of the 4 choices are possible but there are more examples in (1) and (3) than in (2) and (4)

After I finish reading this paper, I got one idea, which is in the opposite way compared with the interface in this paper. Suppose we can draw the trajectory in the tempo-loudness space first, and then we can control the tempo and loudness of the performance according to the trajectory. I think it is an application of the expression synthesis. However, there are still some problems. For example, how can we control the moving speed on the trajectory and how can we get the total length of the trajectory based on a specific performnce.