Eric Cheng
2.22.06
ISE 599

Visualizing Expressive Performance in Tempo-Loudness Space
Jörg Langner and Werner Goebl

In this paper, Langner and Goebl present a method for the visualization of tempo and loudness characteristics in a performance. Their method consists of two basic steps: the acquisition and reduction (smoothing) of performance data, and the mapping of that data to an animated representation in tempo-loudness space.

Timing and loudness data were retrieved from both MIDI grand piano recordings and conventional audio recordings. The authors note that the MIDI recordings have the advantage of clearly defined onsets, but that the precision of some computer-monitored pianos is not much higher than that of data obtained from audio recordings. Once the onsets are determined, they are matched to a symbolic score of the piece so that the onsets of the track level can be automatically determined. Track level is defined as the unit of score time at which tempo variations are measured. In the case of the Chopin Etude they examined, the track level was the sixteenth note. In the case of the audio recordings, timing information was obtained using interactive software for automatic beat detection.

Loudness data were obtained either from a recording of the MIDI file or the audio recording itself. The files, which were in WAV format, were converted to loudness envelopes in sones using a MATLAB implementation of Zwicker's loudness model (Zwicker and Fastl 2001). For each tracked point in time, the highest loudness value within half an inter-track interval was the value used. This windowing procedure accounted for the fact that important loud notes sometimes did not coincide with track time.

Both tempo and loudness data were smoothed using overlapping Gaussian windows, with window size being defined as the time from the left to the right point of inflection in seconds (corresponding to two standard deviations). Choice of window length essentially determined the scale on which the data were examined. For example, a window length of a quarter note would bring out more local phenomena while a window length of 4 bars would show more global and long-term developments.

The visualization is implemented as an animation over time. A red dot moves in synchrony with the music, leaving behind a trace of its trajectory in tempo-loudness space, with the x-axis corresponding to tempo, and y-axis corresponding to loudness. As time progresses, the trajectory of the red dot fades out and decreases in size, evoking an impression of a third dimension: time.

This implementation was used to examine performances of Chopin's Etude in E-Major and Schubert's Impromptu in G-flat Major. Several interesting observations emerged from their analysis. In the three performances of the Etude, the expressive trajectory tends to move towards the lower left side of the space at phrase endings, corresponding to a decrescendo and ritardando towards the end of phrases. Interestingly, a counterclockwise trajectory, corresponding to a strong crescendo and ritenuto, was found at the musical apex of the excerpt in all three performances. In examining the Schubert Impromptu, the authors implemented an algorithmic performance following the "faster-louder" model proposed by Todd (1992) and compared this to a recording by Alfred Brendl. The visualization highlighted the phrasing cues present in Brendel's performance, which were absent in Todd's model.

In conclusion, Langner and Goebl present a fascinating and illuminating method for music visualization. The transformation of aural stimuli to the visual realm produced interesting insights into performance practice, and could serve as a valuable tool for further analysis of expressive performance strategies. Judging from the animations posted online, subtle expressive cues were made clear and noticeable through the visualization method. It would be interesting to extend this method to include other musical parameters and to explore other possible visual representations.