Leila Vaziri
ISE575b-report #10

Visualizing Expressive Performance in Tempo-Loudness Space
Authors: Jorg Langner, Werner Goebl

This study introduced a novel approach for an integrated visualization of two performance parameters as tempo and loudness. Its properties and behavior demonstrated with examples from several expressive performances. This approach includes data acquisition from audio recordings and via MIDI, data reduction by continuous smoothing over a certain time window, and a two-dimensional display which shows the performance on the screen, or like snapshot for a particular excerpt of the performance.

One of the problems of performance analysis is the enormously large amounts of information the researcher has to work with. This method shows efficient display and analysis of large amount of performance data.

An integrated analysis technique was developed in which tempo and loudness are processed and displayed at the same time. Both curves are smoothed with a window size corresponding ideally to the length of a bar. These two performance parameters are then displayed in a two-dimensional performance space on a computer screen. The trajectory of its tail shows geometric shapes which are different for different performances.

The behavior of the algorithm and insights illustrated with performances of two musical excerpts by Chopin and Schubert. Expert performances of Chopin's E major Etude and an algorithmic performance of Schubert's G flat major Impromptu are compared with performances by famous pianists. These two excerpts were chosen because articulation is constant throughout the whole excerpt and analysis can concentrate on tempo and dynamics.

The visualization needs two main steps in processing; (1) data acquisition and (2) the gathered data need to be reduced over the certain time window corresponding to a certain granularity of display. The timing information of expressive performances in MIDI format has the advantage of having each onset defined. The track level is a unit of score time defining the resolution at which tempo changes are measured. The track level is usually faster than the beat as indicated through the time signature. The software analysis of audio date, proposes a possible beat track by displaying the beats as vertical lines over the amplitude envelope of the audio signal. The system provides audio feedback in the form of percussion track mixed with the original sound. For MIDI and audio data, the loudness information was taken from recording of the played back MIDI file or audio file itself. The audio signal is converted into the frequency domain and bundled into critical bands. After determining spectral and temporal masking effects, the loudness sensation is computed from the equal loudness levels which in turn are calculated from the sound pressure level in decibel. From the loudness envelope, one loudness value is taken for further data processing.

The tempo and loudness data were smoothed using overlapping Gaussian windows. The window size will be chosen corresponding to the average performed duration of a bar, resulting typically in a window size of around 3 seconds. The smoothed data is displayed in a two-dimensional space of tempo against loudness. To elaborate the impression of time, the trajectory of the initial red dot decreases in size over time. This enlarges the impression of a virtual space. To indicate some types of structural properties of the score, the dot is enlarged and changed in color at phrase boundaries.

One of the experiments was based on Chopin's E Major Etude. The Chopin Etude has a very homogenous texture. The track level was set to the semiquavers level. The second of the two main steps in data processing are: data reduction via smoothing. Smoothing leads to a worthwhile reduction of amount of complexity of data. The smoothed curves are the basis for the animations. Many observations emerge from the combined display that reveals both the similarities and differences between the performances. The similarity between the performances was the

expressive trajectory tends to go to the lower left side of the space at phrase boundaries. Tempo apex occurs before the loudness apex in all three performances. A certain tendency towards a counterclockwise movement of the expressive trajectories was found in all three performances. The differences between the performances are outstanding on the first glimpse. One of the differences between the famous pianist and the two others is that he does not slow down too much at the end of the section. He planned to play the whole etude and not only the first 21 measures, as other pianists did. The shape of a path is very much dependent on the window size of the smoothing window applied to the data. Smoothing introduces some artifacts. Turning points or peaks in data change their location and extent through smoothing. Analysis with visualization technique should be accompanied by conventional data display for detailed analysis.

In second experiment on Schubert's G flat Major Impromptu, the main strength of the display is to explain relations between tempo and dynamics. This relationship was modeled like" the faster the louder, the slower the softer" by Neil Todd. Windsor and Clarke used the model to test how much of expressive variation of a real performance could be explained. They came up with the preferred algorithmic performance using different parameters for timing and dynamics. The "Hybrid performance" is not a diagonal as would have been expected for these reasons, one because of different parameters were chosen for timing and intensity; second because of the linear relationship between timing and dynamics in terms of MIDI velocity units is not linear when dynamics is expressed. Todd's model responds to the symmetrical grouping structure. The comparison of a performance model to a professional performance was conducted not to prove Todd's model simple one, which it was planned to be, but to demonstrate the behavior of the two-dimensional display on real and artificial data. It also show the strength of highlighting relations between the two performance parameters displayed.

The display in this study shows both the tempo and the dynamic shaping of a performance and elucidates the interaction between these two. The smoothing procedure could be supported by a perceptual hypothesis: "The perceived tempo of an expressive music performance is more stable than the tempo rustling form the played note onsets measured". The larger the smoothing window, the more the smoothed display will deviate from the data. The trajectories of the visualization technique evoke a visual impression of gestural motion. The idea of the two dimensional space could be reversed, using it as an interactive control of music performance. The control space is called perceptual expressive space and spans a space between adjectives like hard, soft, light, and heavy. It should be better if the experiment involve more performance analysis, also based on different composers. Understanding the meaning of certain repeating shapes of trajectories, comparison of the performance trajectories with emotional responses of musically trained listener can be interesting. Another approach can be study of the relations between the two-dimensional time series with Artificial intelligence.