Interactive Visualization of Expressive Piano Performance

Motivation

By following the ideas of Ben Shneiderman’s Information Visualization Mantra, multiple visualization views have been smoothly integrated into a music analysis workflow. Shneiderman’s mantra suggests that the data should be accessible at various levels of detail, which enables the user to get a coarse overview, as well as a detailed visualization of selectable data subsets. Overview visualizations have been realized as 3D views, which can be interactively manipulated, while 2D charts are used to show the detail information.

Features

Overview, then zoom and filter, then details on demand

By following the ideas of Ben Shneiderman’s Information Visualization Mantra, multiple visualization views have been smoothly integrated into a music analysis workflow. Shneiderman’s mantra suggests that the data should be accessible at various levels of detail, which enables the user to get a coarse overview, as well as a detailed visualization of selectable data subsets. Overview visualizations have been realized as 3D views, which can be interactively manipulated, while 2D charts are used to show the detail information.

Direct interaction widgets

In order to keep the user interface very simple and intuitive, direct interaction widgets - which allow the user to manipulate parts of the visualizations (e.g., time locator planes in the 3D visualizations) by directly dragging them with the mouse.

Linking

One performance of a given score can be analyzed simultaneously in several views (the views show the same position in performance or “real” time). Moreover, several performances of the same piece of music can be analyzed in multiple instances of the application. The applications can be linked over TCP/IP, so that they all show the same positions in score time. This gives the user the ability to compare different performances with respect to certain expressive features (e.g., tempo or dynamics structuring).

Moreover, network linking allows visualizations to be distributed among multiple workstations. This is a useful feature, since real-time visualizations of highly dynamic content (as music is) is likely to exhaust the processing resources of a single workstation (since laptops are a target environment for the application, this is even more likely). Another problem that can be circumvented with distributed visualizations is the limited screenspace of a single workstation.

Visualization metaphors

3D Piano Roll

A well known visualization metaphor which is used heavily in musicology and music production, is the Piano Roll representation (a leftover from the early days of music performance research, where musicologists tried to analyze music from paper rolls used to steer reproducing pianos). It is used heavily in sequencing applications and it is well understood by musicians and researchers.

The classical piano roll representation has been extended in several ways. First, it has been embedded into 3-space, by using the third dimension for encoding of MIDI velocity (louder notes are higher). Second, colors are used to show articulation (staccato or legato playing). Legato notes are rather greenish, staccato notes are shown in red. Thirdly, tempo has been visualized on the ground plane of the piano roll. The first method was to draw the bar lines at their corresponding positions in performance time (the wider the space between bar lines, the slower the tempo), the second one was to use colors. Faster passages are marked by yellowish colors at the ground plane, bluesh colors correspond to slower passages.

3D Performance Worm

The Performance Worm is a visualization metaphor that shows the development of tempo and loudness values over time (since audio data has not been used as input, it was not possible to calculate perceived loudness values; instead, MIDI velocity has been integrated over all voices for a coarse dynamics measure), by plotting circle-like shapes at discrete points in time. By fading out points further away in time, a pseudo-3D representation is achieved. This metaphor has been extended to a full 3D visualization, that allows the user to navigate the worm and zoom into regions of interest. Moreover, articulation is also sketched, by modulating the shape of the worm (to achieve this effect, articulation values are used as control parameters for superellipsoids, which in turn are used to define the cross section of the worm; the worm itself is modeled as a Cardinal Spline, along which the superquadric cross sections are swept).

Multi-resolution tempo/velocity/pedal visualization

Tempo is computed from the time intervals between two played notes, normalized by the time interval between the corresponding locations.

The tempo visualizations allow for interactive selection of the metrical level (Bar/Beat/Sub-beat) that is used for the calculation of the tempo.

Time range/ROI (Region of Interest) specification

Since audible information is inherently time-dependent, interesting features in music are distributed primarily along the time dimension. This fact is reflected in the user interface by a simple, yet effective, principle: Let the user navigate to an interesting point in time, let her or him hear the music at that time, and let her or him constrain the area surrounding the current point of interest. In the 3D view, the current position and the lower/upper bound of the region of interest can be adjusted with so-called time planes, that can be dragged directly in the 3D visualization domain. In the 2D view, the selection of the user is visualized with an adapted fisheye view method, that magnifies the region of interest (the focus region) and minimizes the context information.

Technical info

The implementation has been realized in J2SE 1.5. The Tritonus OpenSource JavaSound implementation has been used for stable MIDI playback. 3D rendering is done with JOGL and the Xith3D scenegraph library.