Strike-A-Tune: Fuzzy Music Navigation Using a Drum Interface

Adam R. Tindale, David Sprague, and George Tzanetakis
Department of Computer Science
University of Victoria
art,dsprague,gtzan@csc.uvic.ca

Abstract

A traditional music library system controlled by a mouse and keyboard is precise, allowing users to select their desired song. Alternatively, randomized playlist or shuffles are used when users have no particular music in mind. We present a new interface and visualization system called Strike-A-Tune for *fuzzy* music navigation. Fuzzy navigation is an imprecise navigation approach allowing users to choose preference related items. We believe this will help users to play music they want to hear and re-discover infrequently played songs in their music library, thus combining the best aspects of precision navigation and shuffles. We have designed an interface using an electronic drum to communicate with a visualization and playback system.

1. Motivation

There are many pieces of software available for organizing and playing music. With rare exception, these programs utilize the keyboard and mouse, devices that are very precise, for navigating the music collection. For large collections of music, problems arise. Users will more often pick their favourite artists over others. The other typical method of navigation is random mode or the iTunes "Party Shuffle."

We propose a new type of navigation: *fuzzy navigation*. People are often in the mood for music they like, without being concerned about the particular song playing. Two websites, Pandora.com and last.fm, have become very popular by playing music based on user-guided feedback. Strike-A-Tune puts the user in direct control of the navigation utilizing the user's own collection. The interface's imprecision introduces adjustable navigation variability. Due to the tactile feedback of a drum, the user is aware of the instrument in his or her hands [1].

2. Previous Work

Two existing systems [3, 6] use a personal MP3 player to adjust an exercise routine by selecting music based on the user's performance. These interfaces provide fuzzy music navigation but none are designed for music re-discovery. The Music Rainbow system [7], despite also using a circular interface based on song similarity, is designed for music discovery and permits precision music selection.

Murdoch and Tzanetakis' [5] radio drum research looked at using a drum to navigate large music collections using rhythm based navigation and SOM dimensionality reduction. For dimensionality reduction, 3D song positions would be accessed by the 3D position of the drum sticks.

The CircleView system by Keim et al. [4] uses a visually similar approach to ours with segments of a circle representing an attribute, distance from the exterior representing time, and colour representing the average data value for that time. Appan et al.[2] also use the analogy of water ripples to display communication patterns over time.

3. Drum Interface

Electronic drum pads come equipped with a piezo sensor that can transduce the movement of the drumhead into an audio signal. The drum software interface, implemented in Marsyas [10] takes an audio stream and segments it into discrete hits. Once a hit is detected, a fixed window of 4096 samples are captured to be processed. This window is sent to the feature extraction section of the software where the resulting feature vector is delivered to multiple Gaussian classifiers. The output of the classifiers is collected and sent to the visualization via Open Sound Control. The system is based on a real-time implementation of the system described in [9].

The categories of gesture used in the system are: Radial Position(0-1) and Magnitude (0-1). The radial position refers to the point on the drumhead where the drum was struck. When classifying the difference between Normal and Rim, the system achieved 85.0% accuracy. Radial accuracy was tested by dividing the radius of the drum into 2, 3, and 4 section; the system achieved an accuracy of 89%, 81.2%, and 73.2% respectively.

4. The Disk Visualization

The disk visualization presents rings of song data with the currently playing/chosen song in the center of the disk (see Figure 1(i)). Rings closer to the center have fewer songs positioned on the ring but more song details. Shneiderman's information visualization mantra of "Overview, zoom and filter, details-on-demand" is thus supported [8]. An overview of the music is provided by outside rings, song data is filtered based on the currently playing song, and song details are provided on demand by highlighting the song. Songs on a given ring are currently positioned evenly around the ring in song segments. The current implementation presents 1 song in the center, 10 songs on the first ring, 100 songs on the second ring, and 1000 songs on the outside ring. Ring number thus conveys 1 dimension of information while position on the ring convey a second dimension. Ring position and segment order can also convey a single variable, like song similarity (used in the current implementation). Currently, songs on a ring are positioned according to similarity to the center song. Songs on rings closer to the center are also more similar to the center song than songs on outer rings. Each song segment is coloured according to a third nominal dimension (eg. genre). Strike-A-Tune supports up to three dimensions of song data for each library entry, in addition to traditional text based information.

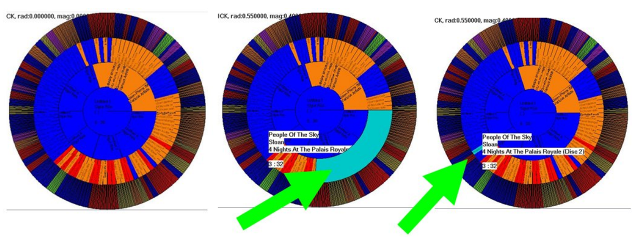

Figure 1. When a selection strike occurs, the selected song's detailed information is displayed and a blue bar animates around the selected ring until it arrives at the chosen song (ii). The chosen song is then highlighted (iii).

The disk visualization intuitively maps drum strike information to visualization interactions. The radial position from the drum center of a simple "selection" strike maps to the radius position on the disk. Radius positions (between 0 and 1) are then discretized to a ring number. Strike magnitude (0-1) represents a rotation around the chosen ring between 0 and 360 degrees. A "test of strength" analogy is used to convey how magnitude impacts a selection. A zero degree rotation selects ring segments aligned with the x axis of the center song. Stronger hits cause an animation to rotate clockwise around the ring until the chosen song is selected. This animation both attracts attention to the selected ring segment (movement is perceptually pre-attentive) and reinforces the "test of strength" analogy (similar to a carnival high striker game). See Figure 1 for details.

The metadata of one of the authors 3207 song personal music library was outputted to a tab delimited text file. Each song had 25 different pieces of metadata used to compute distance metrics between songs. We chose to implement fuzzy music navigation based on song similarity. Each metadata category was weighted based on importance and string similarity was either determined by exact string matching or the percentage of characters in common.

5. Discussion

The Strike-A-Tune system offers many advantages compared to conventional shuffle or precise mouse and keyboard based music navigation systems (ie. precise navigation). Unlike both precise and random music navigation, the benefits of fuzzy navigation improves as the library size increases. Songs the user has forgotten about have a chance at being played (unlike using precision navigation) but only songs similar the user's current preference will be played (unlike a randomized playback). The disk visualization provides an overview of music data and details for songs in focus. Finally, the intuitive nature of the interface makes interacting with the system as simple as using a traditional mouse and keyboard, but allows us to expand our interface in new directions of music exploration.

References

[1] Miguel Bruns Alonso and David V. Keyson. Musiccube: making digital music tangible. In *CHI '05: CHI '05 extended abstracts on Human factors in computing systems*, pages 1176–1179. ACM Press, 2005.

[2] Preetha Appan, Hari Sundaram, and Belle Tseng. Summarization and visualization of communication patterns in a large social network. In *The 10th Pacific-Asia Conference on Knowledge Discovery and Data Mining*, 2006.

[3] Greg T. Elliott and Bill Tomlinson. PersonalSoundtrack: context-aware playlists that adapt to user pace. In *CHI '06: CHI '06 extended abstracts on Human factors in computing systems*, pages 736–741. ACM Press, 2006.

[4] Daniel A. Keim, Jörn Schneidewind, and Mike Sips. Circleview: a new approach for visualizing time-related multidimensional data sets. In *AVI '04: Proceedings of the working conference on advanced visual interfaces*, pages 179–182. ACM Press, 2004.

[5] Jennifer Murdoch and George Tzanetakis. Interactive content-aware music browsing using the radio drum. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, 2006.

[6] Nuria Oliver and Fernando Flores-Mangas. MPTrain: a mobile, music and physiology-based personal trainer. In *MobileHCI '06: Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, pages 21–28. ACM Press, 2006.

[7] Elias Pampalk and Masataka Goto. MusicRainbow: A new user inferace to discover artists using audio-based similarity and web-based labeling, 2006.

[8] Ben Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. In *VL '96: Proceedings of the 1996 IEEE Symposium on Visual Languages*, page 336. IEEE Computer Society, 1996.

[9] Adam R. Tindale, Ajay Kapur, George Tzanetakis, and Ichiro Fujinaga. Retrieval of percussion gestures using timbre classification techniques. In *Proceedings of the 5th International Conference on Music Information Retrieval*, pages 541–544, 2004.

[10] George Tzanetakis. Marsyas: A framework for audio analysis. *Organized Sound*, 4(3):169–175, 2000.