PersonalSoundtrack:

Context-aware playlists that adapt to user pace

Greg T. Elliott
University of California Irvine
Arts Computation Engineering
Irvine, CA 92697-2775 USA
gelliott@uci.com

Bill Tomlinson
University of California Irvine
430A Computer Science Building
Irvine, CA 92697-3440 USA
wmt@uci.edu

CHI 2006, April 22–27, 2006, Montréal, Québec, Canada.
ACM 1-59593-298-4/06/0004.

Abstract

This paper describes a mobile music player, PersonalSoundtrack, that makes real-time choices of music based on user pace. Standard playlists are non-interactive streams of previously chosen music, insensitive to user context and requiring explicit user input to find suitable songs. The context-aware mobile music player described here works with its owner's library to select music in real-time based on a taxonomy of attributes and contextual information derived from an accelerometer connected wirelessly to a laptop carried under the arm. We are in the process of evaluating this prototype with 25 users who will compare the system's context-sensitive playlist to random shuffle. On the basis of user feedback and analysis, a hand-held device will be implemented for testing in less constrained mobile scenarios. PersonalSoundtrack allows users to experience their music with both mind and body, providing a unique embodied experience of their personal music library. In mobile environments where attention is a limited resource, users can spend less time deciding what music to enjoy and more time enjoying it.

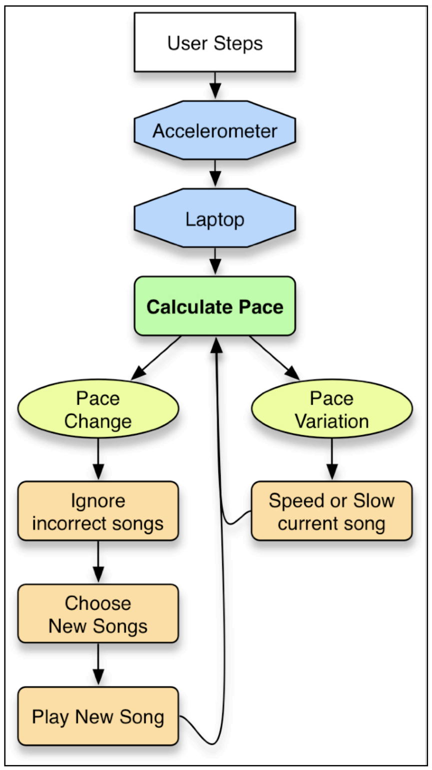

Figure 1. PersonalSoundtrack system diagram showing inputs, actions, and flow of control.

Keywords
Interaction design, interactive music, mobility, wearable computing, emotional state, inherent feedback

ACM Classification Keywords
H.5.2 Information interfaces and presentation (e.g., HCI): User Interfaces

Introduction
Portable music players (e.g. the iPod) allow users to listen to music in multiple mobile scenarios. They listen on the way to class, on the subway, during bike rides, jogs, or workouts, etc. Many users attempt to plan for mobile activity by pre-defining playlists that correspond to specific activities or moods. From Suchman [1], plans alone do not dictate actions, but instead provide scaffolding that individuals can use to organize action. Thus, users attempt to follow previous plans while continuously adapting their actions to the environment [2]. Pre-defined playlists cannot adapt to such ever-changing situations without explicit user input. Manually selecting music requires both user attention and memory when mobile navigation inherently demands the majority of user resources [2]. Context-aware playlists that automatically choose music in real-time and in response to user movement, can better match the unpredictability of mobile activity. By monitoring user pace, the mobile music player becomes a personal DJ that automatically slows, speeds, and changes songs to match the user's movement.

The system described here detects user pace and chooses songs by comparing SPM (steps-per-minute) to BPM (beats per minute) (Figure 1). To start the music, a user holds the system, places earphones on, and begins walking. Music is seamlessly matched to the user's speed, putting the user "in tune" with the music. By continuously adapting to user pace, the device remains "in tune" with its user without explicit control; however, PersonalSoundtrack uses a simple mechanism to learn inappropriate song choices should users decide to explicitly skip songs.

Scenarios
PersonalSoundtrack can adjust to many scenarios, providing unique personal experiences using familiar music. Feedback received from beta testers inspired the following hypothetical scenarios:

- *Morning Jogger* – A woman grabs her headphones and heads out for an early morning jog. She begins a warm-up walk and with each step music fades in to match her pace. Each beat of the song occurs when she takes a step, synchronizing her with her music. As she speeds to a jog, the music matches her pace by cross-fading into a faster song. Suddenly she feels a surge of energy and decides to sprint to end her jog. The device chooses an upbeat song that emotionally fuels her sprint. She slows to a casual walk to cool down, and the music slows and adapts. Her music has reflected her morning workout, providing a pleasing emotional and physical experience beyond a simple jog.
- *Indecisive Walker* – Bored on his 10-minute walk to the subway, a man has neither time nor desire to stop walking and create a playlist; however, randomly selected music is less than satisfying. He is not sure what he wants to hear, so he chooses the context-aware playlist and a song begins to play. While the song is not his favorite, he doesn't mind hearing it. As he walks, he notices he is walking to the beat of the song, or rather, the beat is reflecting his steps. He

begins to enjoy the nuances as each step coincides with a bass pluck or a kick drum hit. His attention drifts back to the world around him, allowing the music to adapt to him as he walks.

- *Late for class* – Leaving in a hurry, a boy walks quickly toward the bus stop. Music fades in at a heightened tempo, increasing his anxiety and speed. As he nears the bus stop, he notices the bus has already arrived. Fearing he might miss it, he runs to catch it. His quick steps cause the music to change to a maddeningly fast song as he races to stop the bus. Reaching the doors just in time, he grabs a seat and relaxes as his music cross-fades into a slower beat.

Related Work

Many researchers have created innovative, context-aware music devices. Mazé and Jacobs [3] developed SonicCity: a wearable jacket that sensed light, noise, movement, and proximity to algorithmically generate music in response to the environment. While the user's movement was used to introduce randomness or set initial tempo, music was primarily generated in response to environmental input [3,4], rather than selected from a user's music library based on user context as in this project. Strachan et al. [5] designed gpsTunes, an mp3 player that uses mobile GPS to guide users to locations by panning and changing music volume. Bassoli et al. [6] created the tunA project which allows users to listen to what other nearby users are listening to using hand-held devices.

Drawing inspiration from SonicCity, we applied a modified notion of embodied music interaction to pop culture. By playing music from the user's library of songs, PersonalSoundtrack integrates with the existing and successful iPod platform. This greatly extends the reach and accessibility of the project, as it can easily be incorporated into everyday use. SonicCity algorithmically generates music, acting as a digital instrument that relies on the creativity of its user, making it less suited for general use.

Furthermore, SonicCity and tunA rely on external factors for context, so that music primarily reflects the environment. In contrast, PersonalSoundtrack prioritizes personal movement in determining context. This is useful in two ways: public and private use is appropriate, and the musical experience is intimately tied to the user. The system provides a novel personal experience of one's music, where the user is always "in tune" with their music without conscious effort. Neither attention nor visual interface is required, making it ideal for mobile contexts. Control is implicit and highly personal: music is chosen to match whatever pace the user finds comfortable.

Tempo and Context-Aware Music

Mobile devices often suffer from the limitations of user attention and poor integration into the environment [7,8]. At home, one can take time to carefully select songs to play; in mobile contexts, attention is a limited resource that results in short bursts of attention [8,9]. Given this mobile constraint, automatic music selection is possible by appealing to the affective system, that can quickly assess valence without conscious thought [10]. The affective system responds well to tempo and rhythm as they strongly influence human emotion [11]. Based on the dimensional model of emotion [12], PersonalSoundtrack follows a three-dimensional version (Arousal, Valence, Stance) [13] were tempo (arousal) directly affects enjoyment and receptivity (valence and stance). The device attempts to synchronize user arousal with music tempo to promote positive valence and open stance. When users explicitly skip songs, the device learns and modifies future selections. This simple learning mechanism is described below.

Hardware

The prototype has been implemented using a Mac laptop and a Bluetooth accelerometer for rapid testing and development. Mac laptops are an ideal environment for testing as they support wireless Ethernet, Bluetooth, and multiple programming languages. Since many people are accustomed to walking while carrying a laptop, it does not impose an undue burden on beta testers that might skew reactions. We have plans to develop a hand-held version based on the Gumstix [14] platform. We use a 3-axis Freescale MMA7260Q accelerometer, reporting at a rate of 150hz with a sensitivity of 1.5+-g. A simple pedometer was tested, but was significantly less accurate than the accelerometer.

A wireless accelerometer allowed the designers to test multiple placements for accurate user pace detection. Popular iPod holding areas were examined, such as hip, arm, and pocket, as well as novel holding areas such as wrist, shoulder, and ankle. Hip placement of the accelerometer was the most reliable and accurate. Wrist and ankle movements heavily depend on gait, while arm and pocket locations decrease sensitivity. User steps are detected by measuring subtle impact along the vertical axis of the body, working well with high-impact (running) and low-impact (bicycling) movement. Software controls are used to normalize noisy data.

Software

Initial software was written in Java, using the Quicktime for Java API for sound, and the PlaceLab API [15] for wireless access point detection. While not discussed explicitly in this paper, location detection is implemented and its potential use is being explored. An HSQL database imports and stores the user's iTunes library, WiFi access points, song probabilities, etc. In processing output from the accelerometer, several sensitivity controls were needed. First, two thresholds were required for changing a song's rate and for changing songs. Second, a moving average of SPM was needed to find the perceived user pace, versus the actual user pace. Precisely timed and typical user steps lead to SPM variations near ±5 BPM and ±15BPM respectively, due to physical gait and environmental obstacles. As a result, the perceived SPM is determined by a system that is difficult to influence at first, but becomes easier to change if the user's pace change is consistent. Random fluctuations in SPM have little influence on perceived SPM, while consistent changes (e.g. changing from walking to running) strongly influence perceived SPM. Third, after beta testing, it was determined that many users felt they were walking in-time to the music, but technically were moving too fast or too slow. The system should not punish users who, within reason, believe they are walking to the beat but are not, nor should it assume the user is wrong. The device determines if the user is unable to walk to the song's BPM, or if the user has purposefully changed pace. This was solved by averaging the user's previous fifteen steps. If the average pace is stable and within the range required to change songs, the music will not change because the system assumes the user believes they are walking to the beat. Surprisingly, it is desirable that the system is highly insensitive to most pace changes, while remaining sensitive to deliberate pace changes.

This software is appropriate for walking, jogging, running, biking, jump rope, etc. Certain activities such as biking may require a BPM that is slower than actual pedal revolution (SPM), as pedal revolutions might be disproportionately faster than walking or running. This issue will be addressed in future work.

The devices used a simple machine learning approach to affect the probability of songs being played based on the following equation

$$p = 100 - 25s + t$$

where p is the likelihood of a song being played, s is the number of times the song has been skipped within the first fifteen seconds of play, and t is the total number of song choices made by the device that session. We plan to explore a more complex learning mechanisms [16] that use variables such as location, terrain, time of day, weather, etc.

A simple graphical interface provides two controls: one to adjust the device's sensitivity, and one to skip songs. In the final prototype, the "Next Song" button will be implemented in hardware, and a hardware version of the sensitivity control may be included depending on the user study.

User Evaluation in Progress

To evaluate this project, the experimental hypothesis is that the selection mechanism will produce results that are more appealing to people than random shuffle. In order to test this hypothesis, we added an option to PersonalSoundtrack that allows it to play randomly shuffled music instead of pace-matched songs. Each user carries a laptop under his/her arm, while listening to headphones plugged in to it. The accelerometer is worn on the waist. This same interface can play both random shuffle and pace-matched songs, so subjects cannot differentiate song selection based on interface.

We plan to observe the behavior of 25 college students in their everyday environments. Subjects bring in their music library for use in the experiment, and BPM is hand-calculated for a subset of songs. Subjects are asked to walk to three specific points on the UCI campus while listening to music, each walk lasting about 10 minutes. The experimenter rides ahead and waits for the subject at each point. While each session is different, all are located in the same general area as the others. The device randomly chooses to use random or pace-matched playlist for the first session, and the non-chosen playlist for the second session. For the final session, subjects are asked to choose to hear music in the style of session one or session two. In each walk, users can skip songs as often as they like.

After the final walk, subjects are interviewed and asked why they chose one style over the other, how they felt about mid-song transitions, etc. During the session, the number of songs skipped is recorded for comparison. This data will be used to help understand the device's behavior given the user's explicit actions.

Future Work

Upon completion of the user study, the researchers will finish a second prototype: a hand-held Gumstix device. Modifications will include interaction revisions, such as how the device interprets pace, how it learns from implicit and explicit interaction, and if it should change music only near the end of a song.

With the hand-held prototype, the investigators will begin a second user study that will compare the PersonalSoundtrack to random shuffle, radio, and user-defined playlists for several activities such as walking, running, and biking. This iterative design process will provide a more complete understanding of how useful and effective the device is in complex scenarios.

Conclusion

This project seeks to create a context-aware device that can meaningfully adapt to users without requiring their adaption to the device. It uses simple contextual cues to provide a more personal experience, adapting to the user on an intimate level. Because PersonalSoundtrack requires a trivial amount of learning from its users, it is immediately viable for nearly all age groups and user types.

Our digital and physical tools can significantly shape the way we work and play. It is important to explore devices that adapt to our unique work practices in order to improve HCI. PersonalSoundtrack does not require users to change how they walk, but attempts to change itself based on the user's unique walking pattern. We hope this device provides a simple example of an adaptive machine that successfully functions in the real world with minimal user learning.

Acknowledgments

We thank Bonnie Nardi, David Kirsh and the ACE (Arts Computation Engineering) Program at UCI.

Citations

[1] Suchman, L. Plans and situated actions. The problem of human-machine communication. Cambridge University Press (1987).

[2] Tamminen, S., Oulasvirta, A., Toiskallio, K., Kankainen, A. Understanding Mobile Contexts. Proc. Mobile HCI (2003), 17-31.

[3] Mazé, R., Jacobs, M. Sonic City: Prototyping a Wearable Experience. Proceedings of the 7th IEEE International Symposium on Wearable Computing (2003).

[4] Gaye, L., Holmquist, L.E. In Duet with Everyday Urban Settings: a User Study of Sonic City. Proceedings of New Interfaces for Musical Expression, (2004).

[5] Strachen, S., Eslambolchilar, P., Murray-Smith, R., Hughes, S., O'Modhrain, S. GpsTunes: controlling navigation via audio feedback. ACM International Conference Proceedings Series, Vol. 111, ACM Press (2005), 275-278.

[6] Bassoli, A., Cullinan, C., Moore, J., Agamanolis, S. TunA: A Mobile Music Experience to Foster Local Interactions. Proceedings UBICOMP (2003).

[7] Kristofferson, S., Ljunberg, F. Making place to make it work: Empirical Exploration of HCI for mobile CSCW. In Proc. GROUP'99, ACM Press (1999), New York.

[8] Oulasvirta, A., Tamminen, S., Roto, V., Kuorelahti, J. Interaction in 4-second bursts: the fragmented nature of attentional resources in mobile HCI. Proceedings of the ACM CHI'05 Conference (2005).

[9] Norman, D. Emotion & design: attractive things work better. Interactions, ACM Press (2002), 36-42.

[10] Kuo, F., Chiang, M., Shan, M., Lee, S. Emotion-based music recommendation by association discovery from film music. International Multimedia Conference (2005), 507-510.

[11] Garlan, D., Siewiorek, D., Smailagic, A., Steenkiste, P. Project Aura: toward distraction-free pervasive computing. IEEE Pervasive Computing, (2002), 22–31.

[12] Plutchik R. The Emotions, Lanham, MD, University Press of America (1991).

[13] P. Ekman and R. Davidson. The Nature of Emotion: Fundamental Questions. Oxford University Press, New York, 1994.

[14] Gumstix Circuit Boards - http://www.gumstix.com/

[15] PlaceLab - http://www.placelab.org/

[16] Mitchell, T. Machine Learning. McGraw-Hill, 1997.