XPOD – A HUMAN ACTIVITY AND EMOTION AWARE MOBILE MUSIC PLAYER

S. Dornbush K. Fisher K. McKay A. Prikhodko Z. Segall
University of Maryland Baltimore County, United States

INTRODUCTION

With close to 1.5 billion cellular handsets in use and worldwide sales last year of over 500 million units, mobile phones are by far the most prevalent mobile device. Second only to the mobile phone, the MP3 player is the most visible and influential mobile device. Candidates to convergence with the mobile phone, such as the Sony-Ericsson W800c Walkman, indicate the significant impact the technological evolution of the music player could have in the development of the next generation of mobile devices and integrated mobile user interfaces. In this paper, we consider the notion of collecting human emotion and activity information from the user, and explore how this information could be used to improve the user experience with mobile music players. Although the device chosen is a mobile MP3 player, our results could also have significant implications in the mobile phone arena.

In particular, this paper proposes a mobile MP3 player, XPod, which is able to automate the process of selecting the song best suited to the emotion and the current activity of the user. Other attempts to relate user activity to mobile devices [7, 9], have targeted the mobile phone user experience.

The proliferation of mobile music players has been rapid, and is predicted to continue, albeit at different rates than mobile phones. JupiterResearch predicts:

"MP3 Player Sales To Drive Online Music Market U.S. sales of MP3 players are expected to increase 35 percent this year to 18.2 million units, according to JupiterResearch. That growth rate is expected to be enough to drive demand for online music services and stores. Sales are expected to grow more than 10 percent annually through 2010, and the MP3 market is expected to have a base of 56.1 million units by the end of the decade, up from 16.2 million in 2004. Separately, market researcher In-Stat said the worldwide online music market is expected to increase 134 percent this year, reaching $1 billion for the first time." [13]

In spite of those predictions, there are a number of issues limiting the user base for such devices. For an individual who only has basic computer skills, the world of MP3 compression, file transfers and endless streams of various file types can be intimidating and difficult to understand. The XPod project is an exercise in developing a mobile music player that is able to eliminate, or simplify significantly, some of the many challenges that users are facing today.

The XPod concept is based on the idea of automating much of the interaction between the music player and its user. The XPod project introduces a "smart" music player that learns its user's preferences, emotions and activity, and tailors its music selections accordingly. The device is able to monitor a number of external variables to determine its user's levels of activity, motion and physical states to make an accurate model of the task its user is undertaking at the moment and predict the genre of music would be appropriate. The XPod is relying on its user to train the player as to what music is preferred and under what conditions. After an initial training period, the XPod is able to use its internal algorithms to make an educated selection of the song that would best fit its user's emotion and situation. This project is a continuation of the Xpog system [8]; both projects were developed as part of a course in wearable computing at UMBC.

PHYSIOLOGICAL DATA, EMOTION AND HUMAN ACTIVITY DETECTION

The increasing sophistication of devices able to provide reliable data about human activity and emotions has made possible the development of real time human-aware mobile systems. Physiological data of the kind gathered by the BodyMedia SenseWear [5] sensor package has been used previously to predict emotions and activity of the user [1, 4]. Data reported in [4] shows that using the SenseWear device, one could determine with reasonable accuracy a number of human emotions. The study [4] demonstrated high recognition accuracy when detecting sadness (92% accuracy), anger (88%), surprise (70%), fear (87%), frustration (82%) and amusement (83%). We are using the data gathered from a new, real time streaming version of the BodyMedia SenseWear to detect different levels of user activity and emotion.

XPOD IMPLEMENTATION

XPod utilizes a client/server configuration where all central processing tasks are performed on a stand alone computer and the final results (such as song selection and song information) are sent to the XPod client device. The development system employs a laptop running Windows XP as a server and a Dell Axim X3i PDA running the PocketPC 2003 operating system as a client. The laptop was used to execute all back end processes necessary for song selection and to store the song repository and song metadata. The PDA was used as a mobile terminal that provided a user interface with the server and the hardware necessary to output music to the user. The laptop and the PDA utilize WiFi networking to establish a two-way connection that allows the user to rate music, skip to the next song, and control the playback of music. Once the user makes a request to listen to music, the server wirelessly streams the music to the PDA, which in turn plays the musical selection.

Figure 1. Xpod user interface.

The music playback is initiated by pressing the "Play" button on the XPod PDA interface. Once the user has indicated his desire to begin to listen to music the server would then examine the data stream flowing from the BodyMedia device to determine the user's current state. The BodyMedia device, using a series of sensors, measures the rate of body movement, acceleration, and body heat from the user and wirelessly transmits the data obtained to the server.

To make a decision based on the state of the user's body, the server relies on a series of algorithms which determine the average and the standard deviation of values obtained from the BodyMedia sensors and compare the results to a pre-determined range that would correspond to "active", "passive" and "resting" states.

Once the state of the user is determined, that information is passed to the neural network engine, which compares the user's current state, time, and activity levels to past user song preferences matching the existing set of conditions and makes a musical selection. The neural network engine obtains user preference information by querying a database that contains past user song selection information. The neural network makes its decision based on the user's past song ratings and the BodyMedia information that was associated with those selections. For instance, suppose the XPod user enjoys techno music during jogging sessions that are regularly scheduled every evening at 6pm. During the next jogging session, the neural network compares the status of the user's body motion - Active, the time of the day - 6:00 PM and past musical preferences and playback a Techno song.

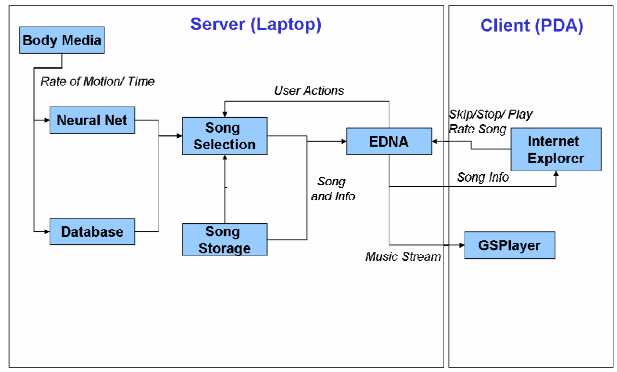

Figure 2. Xpod system architecture.

During the neural network training stage the user listens to music played by the XPod and rates his song satisfaction using the GUI on the PDA. The user preference applies not only to a single song, but to the artist and genre of the song. To make sure that the neural network is able to correctly reference song information, each song's ID3 tags are modified to indicate the genre, artist, song title, album title and beats-per-minute (BPM). The song BPM was obtained by using the MixMeister BPM Analyzer application [10] to analyze the entire song library and write the BPM information to each song's ID3 tags. By comparing past user body states, BPM, genres and authors of the past song selections to the existing set of variables, the neural network makes an informed decision and selects an appropriate song. The song selection is then forwarded to the edna streaming server application [11], which makes a call to the music library and wirelessly transmits the song to the GSPlayer MP3 player application [12] on the PDA (see Figure 2).

EXPERIMENTATION WITH SENSEWEAR

The first step towards supplying user's physical state information to the XPod was to create an algorithm that identifies the state of the user. The user's state was mapped to one of three distinct activity levels: "resting", "passive", or "active". In order to accurately determine the user's activity level, the BodyMedia device was used to gather physiological data. The device gathers data on transversal and longitudinal acceleration, galvanic skin response (GSR), skin temperature, heat flow, and near-body temperature.

To learn an exact method to determine user's body state, several experiments were designed and tested with various individuals in an effort to understand the data and value ranges produced by the BodyMedia. To ensure that the testing covers as many possible ranges of activities and that results are not limited to single individual, various people of different genders, ethnicities and athletic abilities were monitored while performing four different activities: lying down, sitting at a desk, walking, and running.

Experiment 1: Lying Down
The first test measured five subjects lying down in order to study a resting state. The following two graphs show the acceleration readings for one subject laying down over a ten minute period.

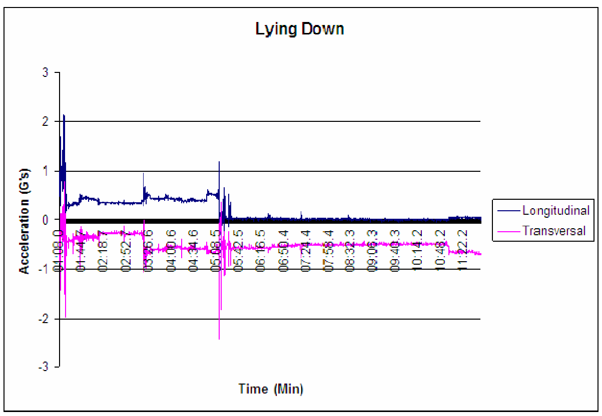

Figure 3. Lying down.

The graphs demonstrate that the acceleration values for both the transversal and the longitudinal accelerations are very steady and low. The other measured values such as GSR, temperature, and heat flow measurements also show consistently little to no fluctuations and because there is very little activity, the values do not change significantly. The GSR value, which is an important factor in determining body state, is consistently low and remains below 0.2 microSiemens. The results of this test were user-independent, as all subjects yielded similar responses.

Experiment 2: Sitting Down
The second test took measurements of five subjects in a relaxed state while sitting down and using a computer. Once again the acceleration values proved to be the only significant measurement as temperature values and GSR showed very little change and maintained a value close to 0.2 microSiemens. The acceleration values were steady over short intervals but showed some fluctuations over time.

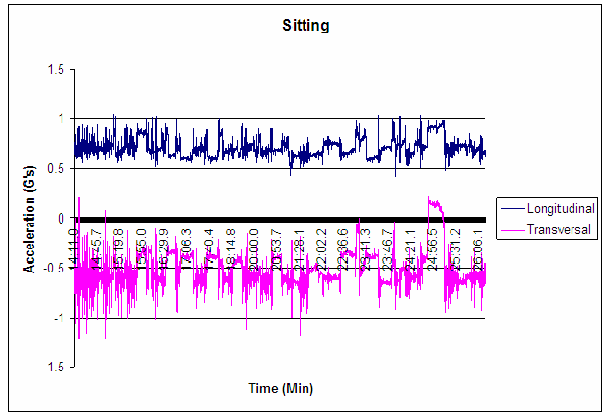

Figure 4. Sitting down.

Experiment 3: Walking
The next test measured subjects walking on a treadmill at a consistent pace for approximately ten minutes. The results of the walking test demonstrate significant differences over the previous tests. The acceleration readings display a much higher deviation over short intervals than previous the tests. Furthermore, the average acceleration value is very consistent over the course of the test and has a value of about 1.0 G.

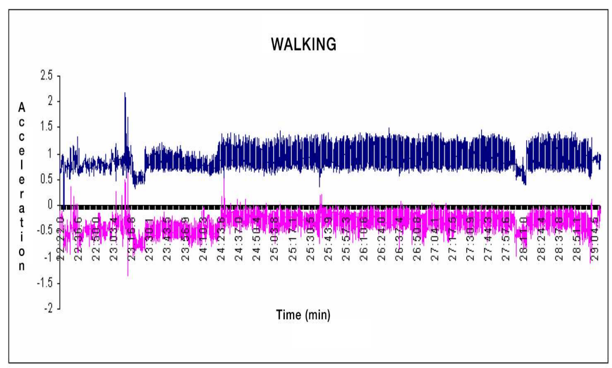

Figure 5. Walking.

Experiment 4: Running
The final test monitored several subjects while running on a treadmill. The average GSR values demonstrated a significant increase to values up to 3.5 microSiemens during the experiment. The acceleration readings also showed a sizable increase with respect to both the average value and the deviation values.

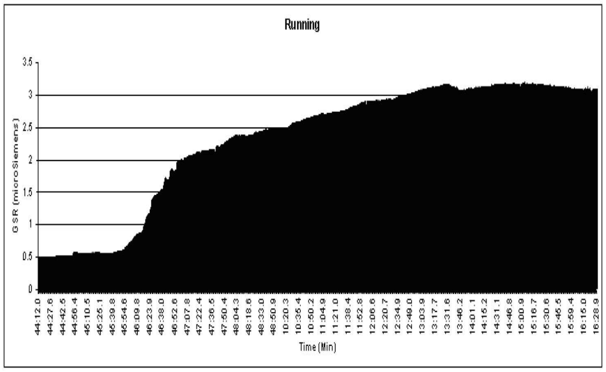

Figure 6. Galvanic skin response (GSR) measured during running test of Asian male

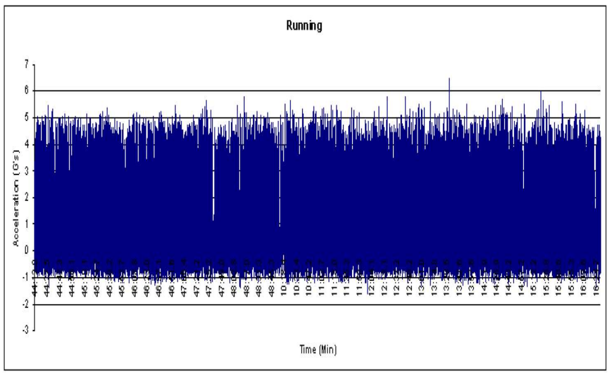

Figure 7. Longitudinal acceleration measured during running test of Asian male

RESULTS

Based on the data that was collected over multiple trials from multiple subjects it was possible to create a relation and determine a value range that would allow the XPod software to accurately determine user state. Since the value ranges between different body states showed noticeable differences, the process of distinguishing the state of the user is much easier and accurate. The relation uses several weighted calculations to compute an integer that maps to one of three states. The first part of the calculation considers the standard deviation values of the accelerations. The standard deviation value is independent of the average value and instead represents the consistency of the acceleration values. When the acceleration values fluctuate rapidly over short time intervals, it signifies increased activity by the user. The second important part of the equation considers the average GSR value over a period of time. The program is set to continuously update the mapped value and re-evaluate the state of the user.

Connecting the user state with music

The user's state is connected with the music in two ways. The first is through the preference database. When a user records their preference for a given song, their preference is stored for their current state. To determine a users preference for a given song a series of queries are executed. If an exact match for the song cannot be found in the preference database an average of similar songs is used. If an exact match for the song name cannot be found the average of other songs on the album will be used. If no other songs on the album have been rated, then the average of other songs by the artist will be used. If no songs by this artist have been rated, then the average of songs in the same genre is used. This information is used to estimate the user's preference of this song in their current state.

This series of queries is executed twice. To determine a user's preference for a song two queries are executed:

1. How much does the user like this song in this state? Only look at records that match the user's current state.

2. How much does the user like this song over all? Look at any records regardless of state.

If an exact match for the song in the given state is found, that is used. If not, an average of the two preference values is used.

The second method of connecting the user's state with the music is through the neural network. The neural network is passively trained by the users' actions. XPod records when the user skips a song and when the user listens to the entire song. If a user skips a song it is a message to XPod that this song was not an appropriate song for this time. The neural network learns under what situations a user will skip a song. If the neural net believes that a user would skip a recommended song; XPod will skip that song for the user. The neural networks are implemented using the Joone neural network package [3]. The network is a fully-connected feed-forward network.

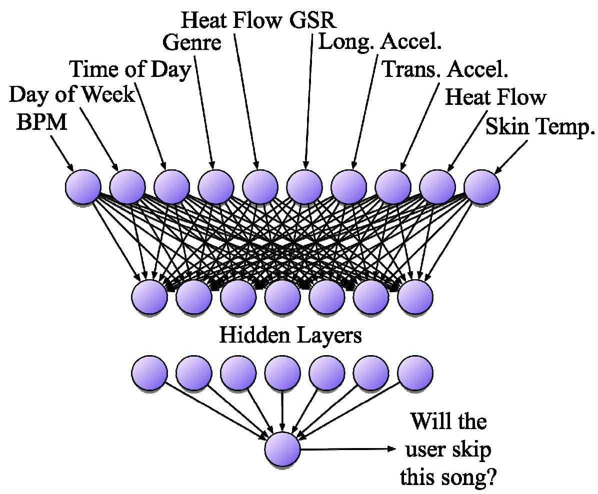

Figure 8. Neural network format.

EXPERIMENTATION WITH XPOD

The XPod system was trained to play different music based on the user's activity level. A simple pattern was used so the state dependant customization could be verified. While a typical user would decide listening preferences based on a large number of criteria, in the experiment the user rated songs only according to their beats per minute. The experiment was to see if XPod could learn a simple pattern of user preference. In the test scenarios, while active the user rated fast songs (high beats per minute (BPM)) high and rated slow songs (low beats per minute) low. While the test user was inactive, they followed the opposite listening preferences; fast songs where rated low and slow songs where rated high. The experiment was designed so that the success could be easily validated. If the system performed as desired the average BPM would change over time.

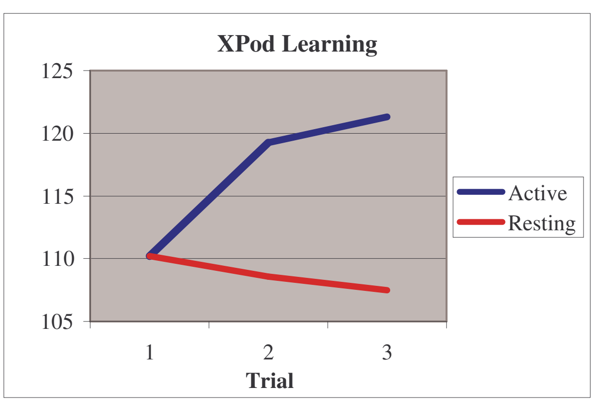

Figure 9. XPod Training Results.

The experiment consisted of six training sessions. During the active training sessions the test user ran on a treadmill. During the resting training sessions the test user was using a computer. During each session the user rated music as they listened to music being played.

XPod successfully learned the pattern of listening behavior exhibited by the test user. In the initial sessions the average BPM was the same for both active and resting. As the training proceeded the XPod learned the desired behavior and chose music to match the preferences of the test user.

In this experiment the pattern of listening was trivial, but easily testable. These algorithms should be able to learn much more complicated patterns of music preference.

CONCLUSION

XPod automates the process of manually choosing music best suited for a user's current activity. The success of the initial implementation of XPod concepts provides the basis for further exploration of human- and emotion-aware mobile systems. Our current work is investigating the notion of closed loop emotion based mobile devices. Such new user experiences should be considered possible (if not highly probable) components of future user experience and technology. While preliminary, the XPod work presented here is an attempt is to provide an activity and state aware user experience, producing a technology that may change the way people experience digital music on mobile devices.

REFERENCES

1. C. Lisetti, F. Nasoz, C. LeRouge , O. Ozyer , K. Alvarez (2003). Developing multimodal intelligent affective interfaces for tele-home health care. *Application of affective computing in human—Computer interaction* (pp 245-255), Volume 59, Issue 1-2
2. Healy. J, Picard. R, Dabek. F(1998). A new affect perceiving interface and its application to personalized music selection. *Proceeding of Workshop on perpetual user interfaces*
3. http://www.jooneworld.com/
4. Fatma Nasoz, Kaye Alvarez, Christine L. Lisetti, Neal Finkelstein. Emotion Recognition from Physiological Signals for Presence Technologies International Journal of Cognition, Technology and Work - Special Issue on Presence, Vol 6(1), 2003.
5. www.bodymedia.com
6. Segall Z., Bylund M and Frank N. The aWare Messenger – a case study in Human Aware Computing, MOBILITY 2004, Singapore, 2004.

7. Bylund M., Segall Z., Towards Seamless Mobility with Personal Servers" INFO Journal, May 2004.
8. Zary Segall, xPog - A Contextualized Human Aware Mobile Music Player, Mobility RT, Hong Kong, May 2005
9. Daniel Siewiorek, Asim Smailagic, Junichi Furukawa, Neema Moraveji, Kathryn Reiger, and Jeremy Shaffer, SenSay: A Context-Aware Mobile Phone, ISWC 2003
10. www.mixmeister.com
11. edna.sourceforge.net
12. hp.vector.co.jp/authors/VA032810/
13. www.grammy.com/news/newswatch/2005/0414.aspx