

RunDMC: A Gesture-Controlled MP3 Player with Vibro-Tactile Feedback and Trigger Mechanism

Anya Kogan, Jake Ritter, Dan Kestranek, Amy Chang

College of Computing and GVU Center

Georgia Institute of Technology

Atlanta, GA 30332-0280 USA

{akogan, jritter, dan.kestranek, amy.chang}@gatech.edu

ABSTRACT

RunDMC is a gesture-controlled mp3 player that also provides vibro-haptic feedback to aid gesture formation in stationary and non-stationary situations. This iteration offers a trigger mechanism that not only signals the device to recognize gestures in order to eliminate errors such as falsely recognized gestures during casual movement, but also allows the user to correct his or her gestural command. This paper focuses on the design of a user study to evaluate the effectiveness of the vibro-tactile feedback by subjecting participants to blind and non-blind conditions.

INTRODUCTION

Although gesture interfaces allow users to interact with small devices such as mp3 players and cellphones, current interaction models of gesture interfaces face problems such as falsely triggered commands caused by unintentional gestures during casual movement and the inability for users to correct themselves before the device executes their command. RunDMC addresses these particular problems by investigating the trigger mechanism employed by a previous iteration, the Gesture Watch.

Instead of a push-to-gesture mechanism that signals the device to recognize gestures, RunDMC employs a triggering mechanism that also allows the user to correct him or herself and re-input the desired gesture before the device executes the command. Furthermore, the incorporation of haptic feedback eliminates the need for the user's visual attention, which could be divided in non-stationary situations.

In order to test the effectiveness of the vibro-tactile feedback, a user study was designed based on the availability of the user's visual attention in a non-stationary situation. This user study compares the user's performance in forming correct gestures in a 'blind' condition and 'non-blind' or normal condition. The following sections of this paper focus on the design process of the user study in order to create realistic 'blind' and normal conditions to accurately evaluate the effectiveness of the device's feedback.

PREVIOUS AND RELATED WORK

In their paper on the Gesture Watch (2007), Kim *et al* [2] describe their protocol for an evaluation of their device.

Overall, their goal was to determine the effectiveness of the device within four conditions--standing vs. walking and indoor vs. outdoor. They were concerned with determining if the sensors could be employed in the sunlight, which in a previous iteration of a similar device (the gesture pendant) could not.

As in our evaluation, the design consisted of a 2x2 within subject design, with each participant performing 5 unique gestures four times each (20 total). Each participant took part in a brief training period until they could demonstrate each gesture with ease. For feedback, they were shown the testing interface with visualizations of the gesture recognition. All tasks were given through a set of Bluetooth headphones as the researchers followed the user with a laptop running the recognition system. The participants were given a general walking path, but no set boundaries to stay within. Our evaluation, on the other hand, had a carefully laid track that was meant to require part of the overall cognitive effort in completing the tasks. By splitting the participant's attention between two tasks (walking and making gestures), we would be able to simulate a more realistic scenario.

The prior study provided very little feedback from the system. Their participants would only hear a short beep whenever a gesture was recognized (correct or not). This is a key difference from our study, in which we strategically gave noticeable feedback to users through haptic vibrations.

In order to analyze the results, several parameters were logged in the original study. In particular, the gestures were recorded as those requested versus those actually performed. This was key in determining the accuracy with which the gestures could be used and we strove to record comparable information for our analysis. Every gesture trigger event, whether intentional or not, was recorded. The time taken to start a gesture following a command and the time taken to complete the actual gestures were both recorded.

The results from their study were highly positive, with an overall recognition rate of 95.5%. However, their participants were free to visually observe the device throughout the study. This necessity limits the utility of such a device

to situations that have easy device visibility. Our evaluation went further by attempting to examine the gesture recognition rate in situations where device visibility would be prohibited. We believe that the combination of our new interaction model and the haptic feedback provided will extend the usefulness of the device by allowing for vision-free navigation.

OUR WORK

Evaluation

The aim for our study was to evaluate whether haptic feedback is helpful for using gestures in the space above a device, particularly when users cannot look directly at it. We hypothesize that the haptic feedback would greatly increase the gesture accuracy in the blind condition when the participant could not look at the watch. To a lesser extent, it may also improve gesture recognition in the normal condition.

In our study design, the participants wear the gesture watch on the left hand, so as to apply the gestures with the right hand. The gestures were designed so that a right-handed person could map the mp3 controls to a set of traditional music controls (i.e. move right to go to next song). The gestures chosen were play, pause, next, previous, volume up, volume down and fast forward. After completing a preliminary learning session, each participant would complete 4 random-order conditions which consist of 6 unique gestures, performed 5 times within each condition. With 30 gestures per condition, the test includes a total of 120 gestures per participant. The conditions are in a 2x2 design - with vs. without tactile feedback and normal vs. blind track setups.

The normal condition includes a clear track set on the ground with strict boundaries participants must stay within.

The blind condition includes a set of arrows on the ceiling grid designed to discourage participants from looking directly at the device while walking.

Figure 3 (see Appendix) shows an image of these ceiling arrows. Both tracks use the same route as seen in the schematic in Figure 1 (see Appendix), but only one of them is set up at a time as necessitated by a given condition. During the normal condition the user will follow only the ground track. During the blind condition the user will follow only the arrows on the ceiling.

In the preliminary session, participants learn the gestures under the guidance of the experimenter. They are free to repeat the gestures for as many times as they need until the gestures are easily recalled. Tactile feedback is provided throughout this session. The experimenter also takes this time to collect the demographic data and sign the consent form. Once the gestures are learned, a short practice session is run. The participant demonstrates each gesture with tactile feedback and without (this order is mixed between subjects), while hearing the mp3 music via the Bluetooth headphones, and walking around the track.

In the main session, the participants perform the gestures according to the block that they are in. The blocks for both the Preliminary and Main sessions are summarized in Table 1 and 2 below. The order of the blocks will be mixed between-subjects and the gestures presented within them will also be randomized for each participant.

To analyze our results, we plan to log the participant performance accuracy of gesture commands, the time to start and complete each gesture, the track distance completed, and the overall workload using the NASA-TLX. We would compare the results across all 4 conditions using an ANOVA analysis.

Preliminary Session				
Learning Session	Normal Practice Session		Blind Practice Session	
Block A (no TF)	Block B (no TF)	Block C (TF)	Block B (no TF)	Block C (TF)
6 gestures	6 gestures	6 gestures	6 gestures	6 gestures

Table 1: Summary of Preliminary Session. 'TF' denotes 'tactile feedback.'

Main Session			
Normal condition		Blind condition	
Block 1 (no TF)	Block 2 (TF)	Block 3 (no TF)	Block 4 (TF)
6 gestures x 5	6 gestures x 5	6 gestures x 5	6 gestures x 5

Table 2: Summary of Main Session. 'TF' denotes 'tactile feedback.'

Interaction Model

We designed a new interaction model in order to provide the start of browsing capability for our device. The model allows the user to observe their input before executing the command. They may also cancel an input if it is incorrect and reenter their desired command. Figure 4 (see Appendix) shows the workflow of the new interaction model.

DISCUSSION

Evaluation Discussion

For our device we used vibration motors to produce vibrotactile feedback. As suggested by previous research, in order to clearly discern the patterns produced by the vibrotactile feedback, we would need to separate them in excess of two motor widths. Otherwise, the unique tactile feed-

back patterns mapped to the individual mp3 gestural commands were indiscernable outside of a merely binary sense. Furthermore we discovered it is more effective to convey the haptic feedback after the user has completed a gesture rather than convey haptic feedback as the user performs a gesture. This is due to several factors such as the speed at which the user performs a gesture timed in comparison to the duration and frequency of the feedback. An alternative and more precise way of conveying haptic feedback is electro-tactile feedback which utilizes electrical currents.

In designing our user study we needed to test for user performance in using our device while his or her attention is available or unavailable in a non-stationary condition. One way to measure user performance is the NASA TLX workload. We would also observe the user as he or she completes the user study. The factors we would note are related to how well the user stays within the designated track while using the device. Although investigating whether or not the user steps off the track can be easily observed, seeing whether or not the user looks down at his or her wrist during the blind condition is less noticeable. Therefore for the blind condition we designed a track with arrows on the ceiling grid so that it would be obvious when the user looked down at his or her wrist. Yet, this method raises some questions about the evaluation protocol that remain to be investigated, such as whether or not the user's neck becomes sore from constantly looking up to follow the arrows on the ceiling and whether or not the participant is prone to tripping.

Upon observation in the pilot test, we discovered that the participant's neck does indeed become sore after about one and a half minutes into the test. We therefore suggest an alternative test track involving eye-level markers as depicted in Figures 5 and 6 (see Appendix). We also observed that the ceiling track is clear enough that the users did not get lost, although they would stop occasionally or step off the track at sharp corners. However, these problems should be eliminated in the new track design as shown in Figures 5 and 6, in which the markers are placed about two inches above the participant's head. The placement of these markers allows the participant to comfortably follow the track while keeping their visual attention adequately divided to test the effectiveness of the tactile feedback in aiding gesture formation. One of the benefits of the former track design with arrows on the ceiling grid was that they would not need to be adjustable to the participant's height. To address this criterion the markers in Figures 4 and 5 are also easily adjustable to the participant's height.

Not only did we want to investigate the effectiveness of the haptic feedback, but we also wanted to investigate improving the designs of the gestures themselves. We designed the gestures to map naturally with common mp3 controls. However we discovered that these gestures often depend on their directionality, the angle of the user's arm (such as whether or not the arm is straight out in front or to the side), and their relation to the device. To temporarily control for these factors for the sake of the pilot user study we

designed the device for right-handed users. Based on how quickly the participants learned the gestures in the training session of the pilot user study this suggested that the gestures map well to their functionality at least for the right-handed user. Whether or not the gestures are as natural for the left-handed user remains to be investigated.

Technical Discussion

Changes to Hardware. Since Project 1, the device has seen the addition of four vibration motors attached to an additional prototyping board. The vibration motors are model number 310-101 from www.sparkfun.com and are rated at 2.5 to 3.8 volts and 75 mA. The vibration motor prototype board connects directly to four pins on the PIC microcontroller of the Bluetooth accelerometer. The PIC microcontroller can only supply ~25mA, which is not enough to run the vibration motors at a substantial strength. To increase the current throughput to the motors, we utilize a transistor for each motor with a shared direct connection to the 9v battery power supply. The transistors allow for about ~70mA to flow directly to the vibration motors when the PIC microcontroller's pins are raised high. To raise a PIC microcontroller pin to high, a capital letter A-D is sent via Bluetooth. To lower a pin, a lower case letter is sent over the Bluetooth link.

Additional hardware changes include upgrades to the infrastructure of Project 1. All of the wires except for the 9v connector wires and the rainbow ribbon cable from the PIC microcontroller to the proximity sensor prototype board were replaced with thinner soft wires. Also, we shortened the new soft wires from the length of the original wires to shave off the extra bulk when trying to compactly package the device. The overall size of the device did increase due to the addition of the vibration motors and the extra prototype board.

Changes to Software. The Expect script from Project 1 received the expansion of the capability to handle vibration outputs. After much deliberation amongst the group members, a late decision called for the rewriting of the entire software in a different language - C++. We decided to implement the same logic from the Expect script inside of Peter Presti's 'accelcap' C++ program that communicates with the device via serial over Bluetooth. C++ provided a more robust and better documented solution to our needs than the much outdated Expect/TCL scripting language. Certain features such as the ability to sleep works much better in C++ than Expect since Expect's 'after' command doesn't block inputs - leading to difficulty reading in data only when desired. C++ also ran much faster due to several reasons. Firstly, we didn't have to create the new expect script process - we just directly hijacked the accelcap software with our functions and programming logic. Secondly, we didn't have to create new processes every time we wanted to trigger a vibration motor. To trigger a vibration motor in the Expect script, a separate C program is created to call a system() command wrapped around an echo shell command that sends a single capital or lower case letter to the device. Finally, to invoke other processes such as the

Totem Movie Player, we could use a simple inline system() call instead of starting a new C program with a system() call or using Expect's 'spawn' command which creates a new process as well as rigging the process up to have the 'expect' command called on it - thus adding superfluous overhead. Everytime we sent a command to Totem Movie Player (essentially a Linux commandline command), we called spawn on it (since Expect's other processing creating command 'exec' creates the new process and blocks until the new process terminates).

Project 2 also received the addition of a Java software framework for logging data during the user study. Standard Out and Standard Error data from the accelcap software is piped directly into a separate java program that can then write the data to a text file depending on whatever criteria we needed. Java was the software of choice due to it's ease of use regarding strings, it's ability to quickly and easily create graphical user interfaces with buttons, and it's built in KeyListener ability. The Java KeyListener proved to be most useful since the command line is in use by the accelcap program so reading from Standard.in to grab user key presses was not an option. The framework is an empty shell that can be filled with any number of java components like buttons or text fields that the proctor of the user study can interact with to direct the logger.

FUTURE WORK

Future Improvements

Optimizing the Device. Because this project is a prototype and a proof-of-concept build, there are a number of improvements that could be made given more time and material access. Optimizing the device would be an adequate project for future groups in this same class.

The first point that could be vastly improved is the size of the device. While our infrared sensors are very effective not only at determining if there is something in front of them, but also at determining how far away the object is, they are quite bulky in size. While we were doing research on these components, we found a set of infrared detectors that only reported binary results (all we need for this device) that were about a third of the size of the ones we're currently using. Changing out these sensors would dramatically cut down on the size of the unit and make the device more comfortable to wear. Another point where the hardware could be improved is the vibration pads. The vibration pads that we used are moderately effective, but as mentioned before, the haptic response could be improved using electro-tactile feedback instead of vibro-tactile feedback. While this would require an almost complete redesign of the hardware, the user study we have designed would work just as well to test the other form of haptic feedback. Another part of the device that warrants more research is the power supply. While we did not have the time to delve into this subject, it is obvious that using a 9-volt battery is not an effective way to power the device as battery life lasts about 15 minutes. The current draw is just

a bit much to try to run both the Bluetooth board and the vibration motors off of the single battery. However, solutions to this problem would require further study.

Another point we would have liked to improve is removing the laptop from the system. One solution to this problem would be replacing the laptop with a beagle board running a very limited version of Linux. A final, and slightly obvious, hardware improvement would be combining the circuits designed into one custom fabricated board (perhaps adding it into the Bluetooth accelerometer board).

As far as the software is concerned, there are also a number of improvements we would have liked to make. While we finally moved away from the Expect script and into the programming language of C++, there are still more improvements to make. We would have liked to object orient the programming a little bit more than it is, allowing it to be ported to different operating systems (as of now it only works on Linux). Also, there is a distinct area of improvement should someone choose to try to implement GART in order to teach the device gestures as opposed to our form of more specific programming. All of these would go a long way to improving the device itself, and would provide for good projects in later classes for groups less concerned with designing or applying a user study.

CONCLUSION

The motivation behind the RunDMC prototype is to investigate viable gestural interfaces for non-stationary contexts and the effectiveness of haptic feedback. The inclusion of the trigger mechanism introduces a new interaction model that reduces the errors in falsely recognized gestures as well as allows the user to correct his or her command before the device executes it. In order to investigate the effectiveness of the haptic feedback we focused on testing the device in blind and non-blind conditions. Because our initial blind condition in which participants followed ceiling arrows proved to be infeasible, in the end we decided that marking the track as illustrated in Figures 5 and 6 would be the most practical for future studies. In the future the device can potentially be optimized for better haptic feedback, mobility, and wearability. With electro-tactile feedback the haptic feedback patterns could be potentially more recognizable. Finally smaller sensors could reduce the device to the size of a wristwatch and would make the device more wearable.

REFERENCES

1. Kim, J., He, J., Lyons, K. and Starner, T. The Gesture Watch: A Wireless Contact-free Gesture based Wrist Interface. 11th IEEE International Symposium on Wearable Computers, Boston, MA, Oct 11-13, 2007.
2. Lee, S. and Starner, T. "Haptic Gesture – Effect of tactile feedback on gesture-based mobile interaction" (unpublished project description)

APPENDIX
Pilot User Study Images

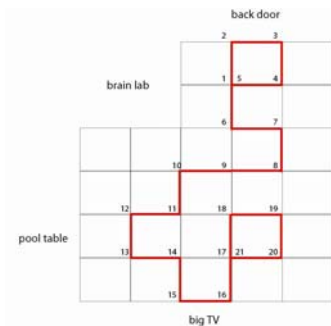


Figure 1: A schematic of the track followed by both blind conditions as shown in Figures 2 and 3.

Figure 2: Normal condition track delineated by tape on ground.

Figure 3: 'Blind' condition track delineated by arrows on ceiling grid.

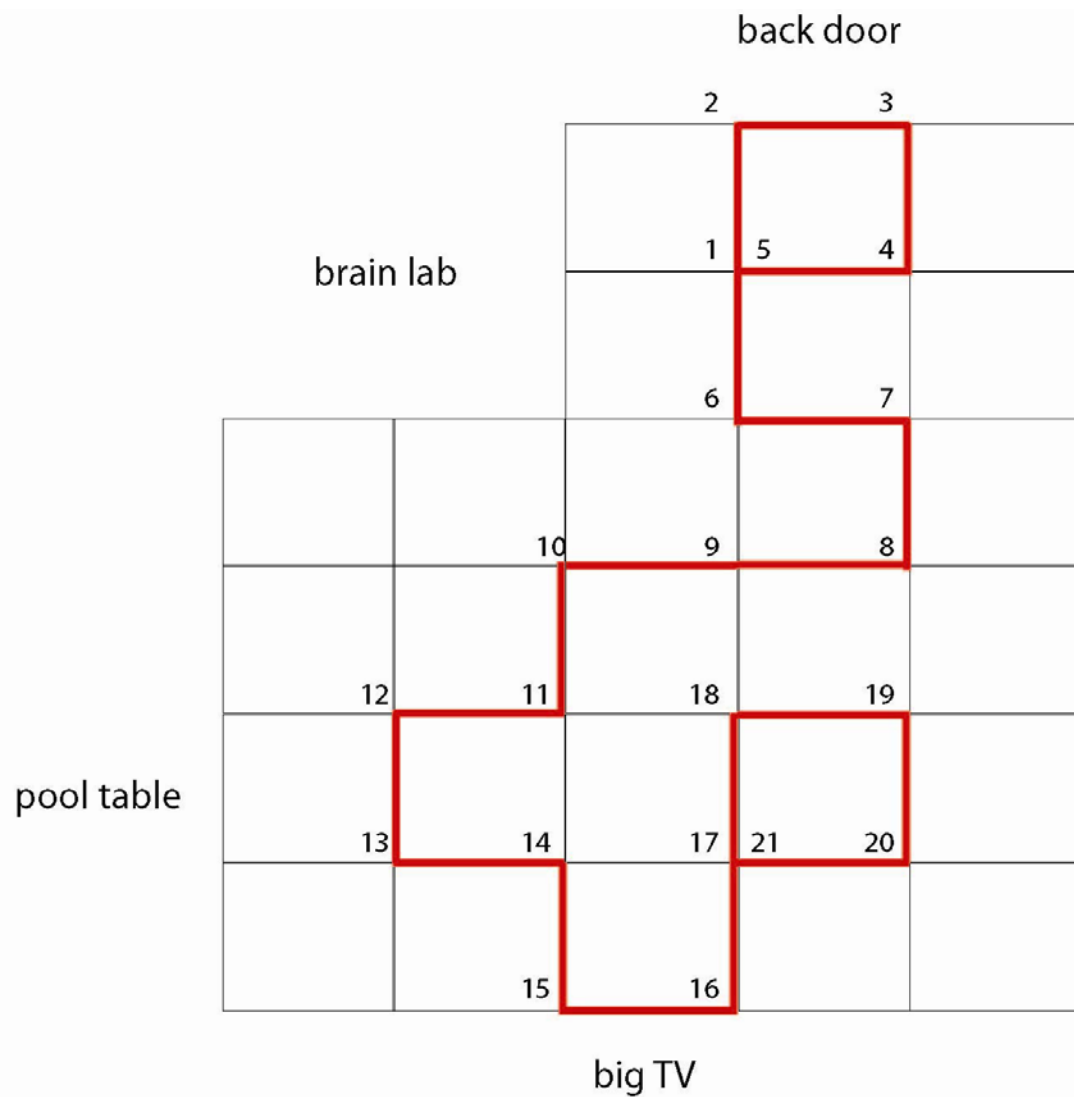


Figure 1a: A larger view of the schematic of the track for the blind and normal conditions as depicted in Fig. 1.

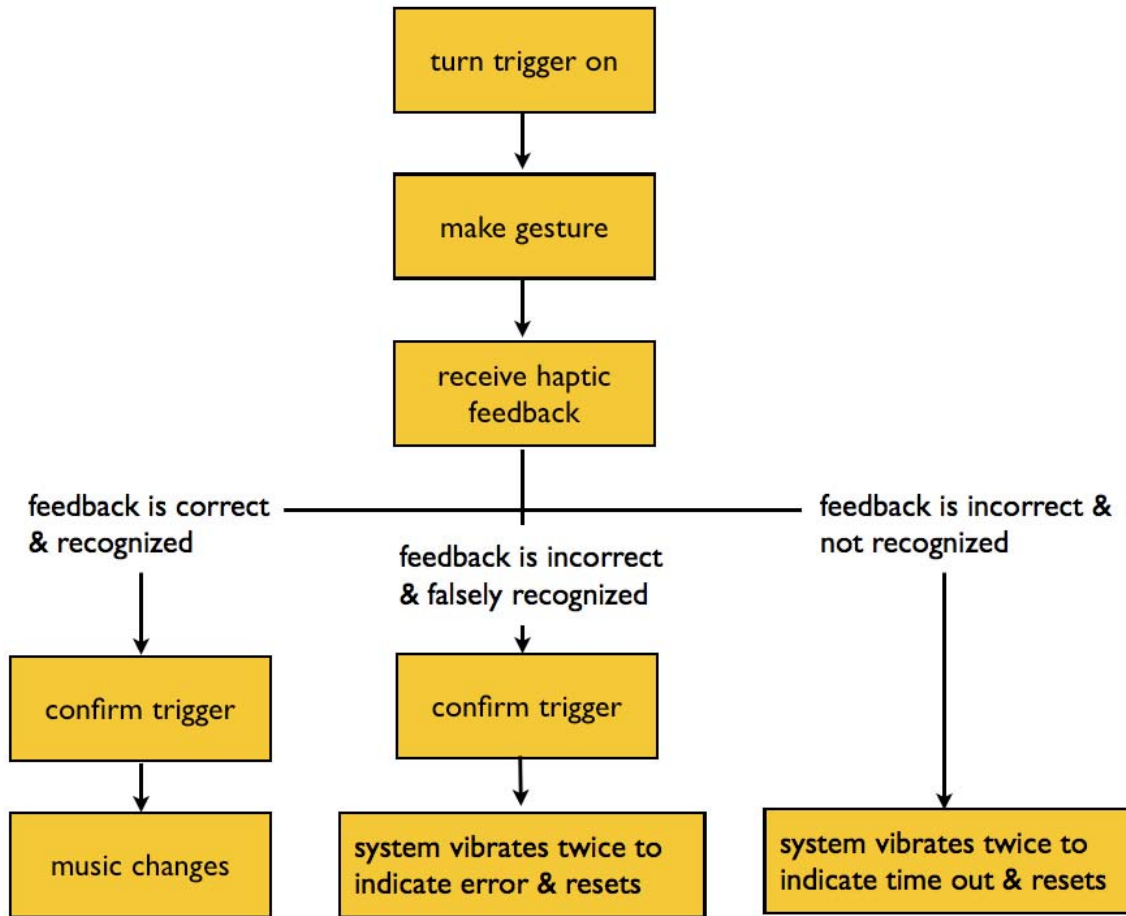


Figure 4: Workflow of interaction model.



Figure 5: A revised method of marking the track for the blind condition.



Figure 6: A closer view of the revised method of marking the track for the blind condition. The marker is about two inches above the participant's head.



Figure 7: A preliminary method of marking the track for the blind condition by hanging posters marked with arrows. Discarded due to lack a ability to adjust height for participant.



Figure 8: A preliminary method of marking the track for the blind condition by hanging tape. Discarded due to lack of directional context.



Figure 9: A preliminary method of marking the track for the blind condition by looping tape. Discarded due to lack of directional context at track intersections.