METU Computer Engineering
SE542 – Human Computer Interaction

Speech Controlled Mobile Games

PROJECT REPORT

Fall 2014-2015

1708668 - Cankat Aykurt

1502210 - Murat Ezgi Bingöl

1679588 - Zeliha Şentürk

Description

Speech recognition is getting more and more important in human computer interaction. Its popular usage areas are in-car systems, automatic translation, interactive voice response, speech-to-text reporters, robotics, etc. There are many improvements in the field of speech recognition which let people do their tasks in a more effective, efficient and easier way most of the time.

Swipe-based third person view mobile games are also the trendiest ones nowadays. As of December 2014, there are five games in top 25 free apps in Google Play Store as you can see in Figure 1. We can give Subway Surfers, Temple Run, and Despicable Me as the popular examples.

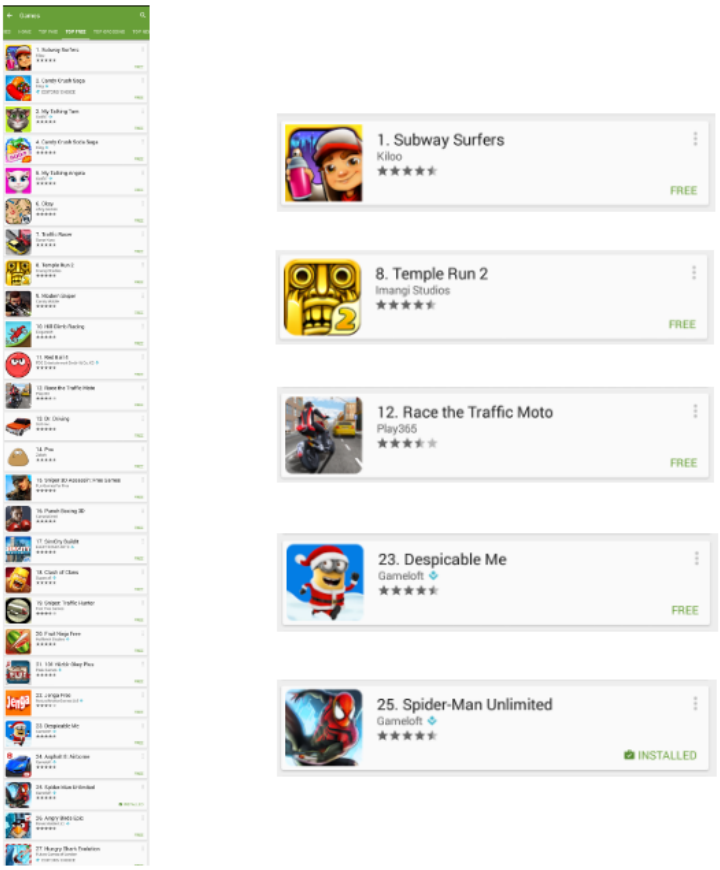

FIGURE 1 - FIVE GAMES IN TOP 25 FREE APPS IN GOOGLE PLAY STORE

Because of both the popularity of the game genre and expectations about enhancing the gameplay with additional speech control, we thought it would be good for people to have the option of playing games by giving commands via speech. Our application will allow to define speech commands, recognize the defined speech commands during the gameplay (such as

"up", "down", "right", "left", etc.) and will transform these commands into swipe actions on the screen of the system.

Walkthrough of the system;

- Define speech commands corresponding to swipe actions

FIGURE 2 - DEFINING SPEECH COMMANDS FOR SWIPE ACTIONS

- Browse previously installed applications on the mobile phone
- Select the target application (swipe based mobile game)
- Reflect defined speech commands to control the target application

Design Guidelines

The following design guidelines are being taken into consideration in design process:

- State Transparency:

  Necessary feedback is provided to the user in the form of a graphic state display. These feedbacks include both visual ones about the system parameters and auditory ones. During the whole recording process, the user knows what can be spoken and mostly freedom is provided for the user to choose what to say. Status displays are rapidly interpretable, without detracting from the task-related activity. They are simple and not obtrusive. It is planned to be careful about and design to not overwhelm the user with feedback.

- Input Control:

  A robust continuous listening system is active during the gameplay. A push-and-hold protocol is used for recording speech commands at the beginning of the game.

- Error Recovery:

Offered an "undo" capability to change existing speech command at the beginning of the game but does not provide an "undo" capability during the game because the game's nature is not suitable for this kind of behavior. Confirmation control is also used to move between recording steps.

- Error Detection:

  None. By defining own commands, error risk is minimized. In case, of spelling a wrong command, the game behaves like wrong swipe action or does nothing.

- Error Correction:

  The natural haptic swipe input is still active to provide an alternative input mode that overrides speech input.

- Logging Performance:

  Users are allowed to report unusual behavior. In case of crashes, a log is created automatically and an enquiry is shown to send log report to the developer.

- Integration with an application:

  This is a very simple application which provides speech interface that does not take into account the state of the application.

Sharing of Workload

In alphabetical order:

- Cankat AYKURT: Conceptual design, development of alternative designs, low-fidelity prototyping, evaluation activities, reporting.
- Murat Ezgi BİNGÖL: Development of alternative designs, high-fidelity prototyping, evaluation activities, presentation, reporting.
- Zeliha ŞENTÜRK: Development of alternative designs, low-fidelity prototyping, evaluation activities, reporting.

User Profiles

Potential users of this project are basically the same as the target users of swipe based mobile games. Since today almost everyone has smart phones and is familiar with the mobile applications, it would not be wrong to make an assumption that a large group of people can be classified as our potential users.

Prototype Design

Low Fidelity Prototypes:

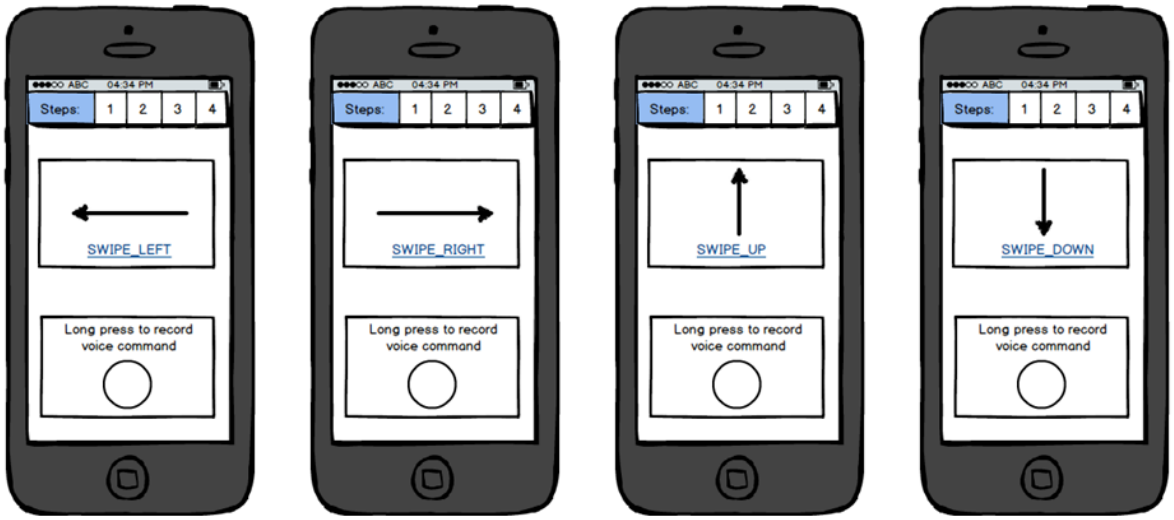

FIGURE 3 – LOW FIDELITY PROTOTYPES

High Fidelity Prototypes:

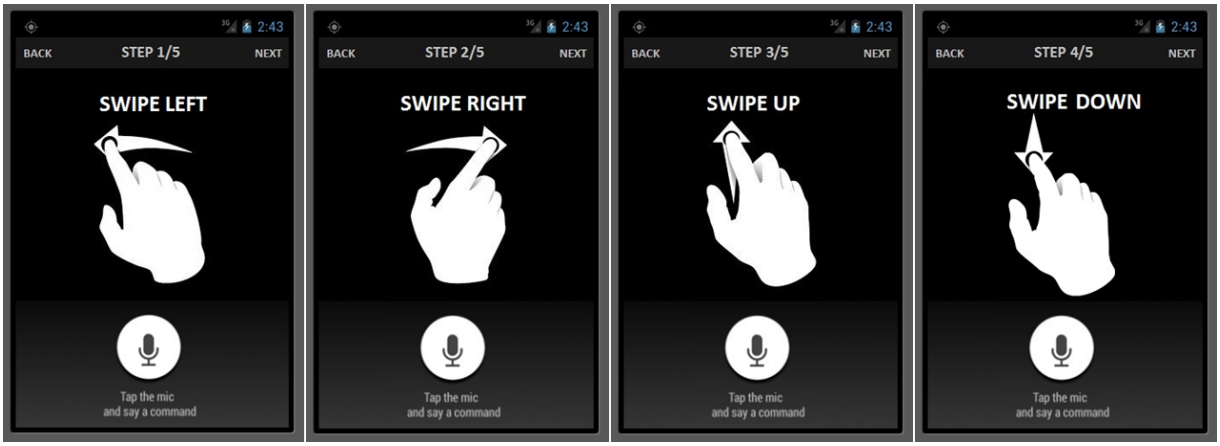

FIGURE 4 – HIGH FIDELITY PROTOTYPES

Gameplay Screenshots & Video URL

Following figures show how our application reflects the user commands to provide a visual representation.

FIGURE 5 – REPRESENTATION OF "UP" COMMAND

FIGURE 6 – REPRESENTATION OF "LEFT" COMMAND

Subway Surfers Demonstration

http://tinypic.com/r/qox509/8

Spider-Man Unlimited Demonstration

http://tinypic.com/r/28we9ux/8

User Evaluation

10 people (4 females, 6 males) aged between 20 and 29 participated in our test cases. 3 of the applicants (2 females, 1 male) had very little or none experience with third person view mobile games, while the others were used to the dynamics of such games. They were firstly asked to

play the game Subway Surfers in its normal mode, without using any speech commands. Then they were asked to play the game through our application which allows usage of speech commands as well as touch based commands.

After conducting the tests; we observed that using our application, users tend to make less erroneous moves and were able to play the game for a longer duration. Table 1 summarizes our test results according to the duration of game played by the users.

Attempts	#1 (No Practice)	#2
Subway Surfers *(normal mode)*	2' 35''	2' 54''
Subway Surfers *(through our application)*	3' 20''	3' 49''

TABLE 1 - RESULTS OF EXPERIMENTS

Evaluation Critique and Redesign

The users who tested our system stated that the system does not provide them a feedback on whether or not it understood the commands or listens to their voice. Considering these assessments provided by the users, we decided to add a small auditory icon on the screen indicating that the command is being heard and processed by our application.

Conclusion

Speech based user interfaces are becoming more and more popular with the developments in speech recognition. Also, the popularity of third person view swipe based mobile games is increasing rapidly. In this project, we aim to develop an application that converts speech commands into swipe commands with providing additional speech interface to replace swipe based commands for third person view mobile games to build effective, efficient, and easier interaction. We prepared both low fidelity prototypes and high fidelity prototypes to enhance product evaluation, and also we created videos to show how this application is going to work like. Swipe based mobile game users have positive opinion to use the applications by talking instead of swiping interaction.

References

1. Android SDK

   http://developer.android.com/sdk/index.html

2. Google Play Store

   https://play.google.com/store

3. Speech Interface Guidelines

   http://www.speech.cs.cmu.edu/air/papers/SpInGuidelines/SpInGuidelines.html

4. Balsamiq Mockups

   http://balsamiq.com/products/mockups/