6.835 Intelligent Multi-Modal Interfaces
Protoype Studio Feedback

Demircan Tas - 922147419
tasd@mit.edu

April 2022

2 Prototype Studio: April 7th

2.1 Early Bird - Isabel Rosa

A virtual pet that obeys gesture and speech commands.

Things that I like:

1. Glitching if the speech commands give the bird character.

2. Visual design of the bird is cute.

3. I like the pun with the bird's name.

Things that can improve:

1. Earl can respond to voice commands especially when they do not register correctly.

2. Users can be provided with hints regarding gesture and speech commands.

3. Expected tasks can be communicated to the user with ques or hints.

2.2 Beacon - Zack Schmitz, Leroy Sibanda, Dave Ludgin

An augmented reality based *beacon* that helps users locate each other.

Things I like:

1. The utility aspect is convincing.

2. Pointing in AR is intuitive.

3. Works with both right and left hands.

Things than can improve:

1. The system should be tested in crowded, urban settings.

2. Personalized beacons can be sold as downloadable content.

3. Speech recognition could be improved.

2.3 Talk to the Hand - Kelly He, Donald Liu

An educational software for teaching sign language.

What I like:

1. Flashcard idea is intuitive and familiar.

2. The objective is interesting and beneficial.

3. Sign language is a good use case for using hand gestures.

What could improve:

1. Users could receive information in smaller chunks.

2. Quiz mode should provide hints.

3. This project would benefit from facial expression detection.

2.4 Dance To Your Own Beat - Alex Crease, Natha Singhasaneh

Software that generates music based on users' dancing.

What I like:

1. The idea is fun! makes me want to dance!

2. User testing is diverse and rigorous.

3. Detects *tips* and *taps*.

What could be better:

1. Program could start from a rhythm that adapts to the user.

2. Visual feed back could include predicted beats and bars.

3. Recorded performances can be used as challenge games for other users. Skeletons can be connected to avatars.

2.5 Point Music - Raul Alcantara

Gesture based sound sample switch. Feels like rapping with your hands.

What I like about it:

1. Entertaining subject, feels original.

2. Test case is well defined.

3. Experiments are well quantified.

What could improve:

1. Sound samples could have karaoke-like bars that fill as each sample progresses.

2. Sensitivity is too high.

3. Vertical hand position can be connected to playback speed; both enabling an easier learning experience, and adding another level of expression and control.

2.6 Multimodal Google Street View - Antonio Berrones, Tiffany Chen

Gesture based navigation for Google Street View

What is good about it:

1. It feels like a taxi ride. Adds a story layer.
2. Movement is relatively fluent.
3. Gestures add a sense of immersion.

What could be better:

1. A "keep going" command would be nice
2. Bookmarks and *jump to* commands would be useful.
3. Commands that use both hands could be intuitive and useful.

2.7 EDiTH: Emotion Detection in Telehealth - Jin Kim, Sandy Jean-Charles

An emotion detector for facilitating remote patient-doctor connection.

What is good about it:

1. Improving non-verbal communication is valuable in the presented context.
2. Presents a good use case for emotion detection.
3. Supporting statistics are provided.

What could be improved:

1. Detected emotions can be displayed as graphics instead of text.
2. Interface should not obscure the image of the patient.
3. Interactive cropping could keep focus on the patient.

2.8 Who said that - Amanda Tong

A system utilizing lip reading to specify the speaker in a crowded conversation.

What works:

1. Unique and creative use case.
2. Interesting choice of example scene.
3. Crops into lips succesfully for reading.

What could be better:

1. Isolation can be triggered and defined by eye tracking.
2. Lip reading performs poorly.
3. Is accurate lip reading even necessary? Can we not simplify the system?