Multimodal Attention Arousal in Head-Mounted Displays

Michelle Quin[1], Vanessa Cobus[2], Tim Stratmann[3], Susanne Boll[2,3], Orit Shaer[1], Andrew Kun[4]

[1]Wellesley College, [2]OFFIS – Institut für Informatik, [3]University of Oldenburg, [4]University of New Hampshire

Motivation

The specific motivation behind this topic was to overcome noise desensitization of staff working in noise-laden, safety- critical environments. Examples of target user groups included hospital caregivers monitoring vital patient data or cargo ship captains recognizing collision time of ships. Existing attention arousal mechanisms in such user groups consist of loud, often high-pitched auditory feedback, which cause alarm fatigue that slows user reaction rates and is detrimental to health, safety, and productivity.

NSF IRES Program

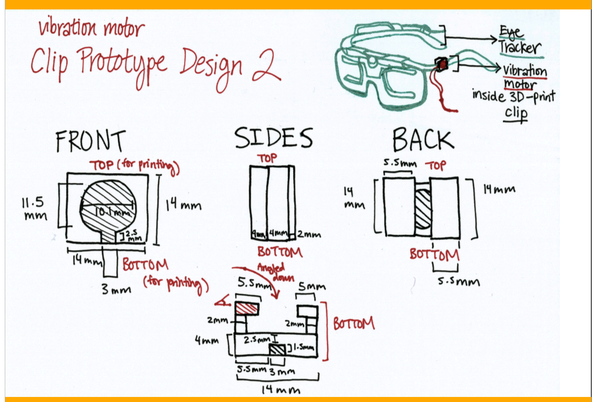
This research was done in Summer 2017 under US National Science Foundation (NSF) funding, through the International Research Experiences for Students (IRES) program “US-German Research on HCI in Ubicomp”. The program is collaboratively managed by Andrew Kun of the University of New Hampshire and Orit Shaer of Wellesley College.

The project “Multimodal Attention Arousal in Head-Mounted Displays” contributed to ongoing research in haptic feedback and interface attention guidance at OFFIS in Susanne Boll’s Media Informatics and Multimedia Systems Group at the University of Oldenburg in Germany.

Conceptual Design

- **Literature review** on auditory and vibro-tactile parameters (frequency, rate, duration, gap length, intensity, etc.)
- **Sketching, 3-D printing, and rendering** of clips and headband to attach *vibration motors* and *buzzers* to a *Tobii Eye Tracker*
- **Coding and wiring** Arduino test function files

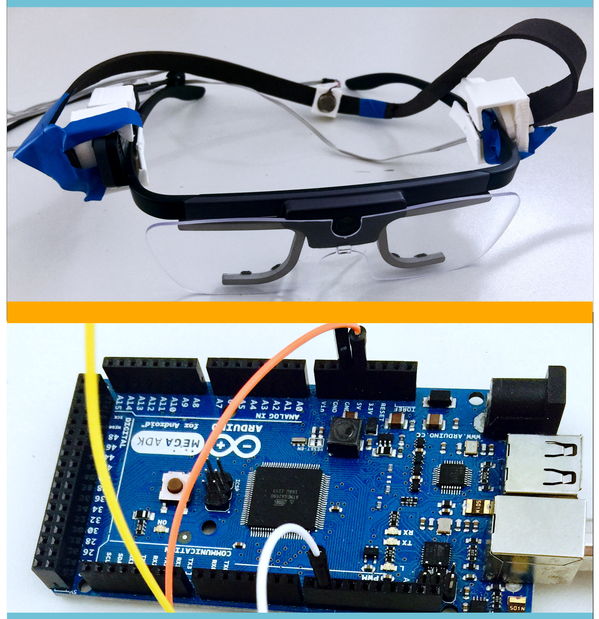
Auditory Parameters	Notes	Implications/Thoughts	For Testing
Overall	Auditory is beneficial if: message is simple/short, not referenced later, deals with events in time, calls for immediate action, visual cues unrealistic, and/or user has a lot of movement	In Head-Mounted maze experiment, vibro-tactile feedback was pitted against speech dictated intructions, and interestingly vibro-tactile had lower workload and faster completion/lower travel distance, may have potential with visually impaired people and modality (Kerdegari, "Head-Mounted" Conclusion)	It is possible to either use existing "ear-cons" or design your own, "ear-cons" aka metaphoric sounds are more successful and efficient than speech dictation; music is easy to recognize but also dependent on users' musical background if using for complex cues
Type	Defined as what type of sound - "beeping" alarm, speech dictation, etc., speech is slow, "earcon" = symbolic sound, divided into three classes: representational, abstract, and semi-abstract (Gårdenfors, "Auditory Interfaces")	If speech is used, consider the environment - voice/vocab/use polysyllabic words; better to use tones but beware false alarms (Ghirardelli, "Auditory-Visual Interactions")	Importance of correspondence between acoustic characteristics of alarm and properties of the message, for example mimicking a heartbeat to monitor heart function (Guillaume, "Judging" Underlying Perceptual Processes)
Frequency	Pitch of stimulus, tested with no silence between pulses (Baldwin, "Multimodal"), the higher the pitch the greater the perceived urgency (Guillaume, "Judging"); it is difficult to distinguish the fundamental frequency of extremely high and low pitches (Ghirardelli, "Auditory-Visual Interactions 5.2)	Pitch is not always applicable because only sounds with regular periodicity for at least a certain duration will be heard as a pitch, and mapping information with pitches only works with users who have a musical background (Gårdenfors, "Auditory Interfaces" 4.1)	In Baldwin's experiment, frequency was tested with no silence between pulses and instead a 20 ms on/offset (Baldwin, "Multimodal" General Methods)
Rate (Tempo)	Rate or tempo is expressed as duration between onset of successive tones (interonset interval, ms) and inversely related to musical concept of tempo as beats/min (Guillaume, "Judging")	The faster the rate, the greater the perceived urgency (Guillaume, "Judging"), unpredictable temporal events are more attention-getting and enhance degree of urgency (Guillaume, "Judging" Underlying	Temporal structure of alarms should have silent periods that allow time to plan action/communicate (Guillaume, "Judging")



test_file

```
130       delay(100);
131       tone(buzzer, y);
132       delay(100);
133       noTone(buzzer);
134       delay(100);
135   }
136
137   //testing incrementing frequency - harmonic up
138   //x = starting frequency, y = increment, z = cycles
139   void harmonic_up(int x, int y, int z) {
140     int i;
141     for (i=0; i<z; i++) {
142       tone(buzzer, x + y*i);
143       delay(100);
144       noTone(buzzer);
145       delay(100);
```

Prototype



User Study

For the user study, participants were asked to watch videos while wearing the prototype. They were instructed to look up at a target and back down to the video upon recognizing feedback, then fill out a survey after each parameter. 25 parameters were run twice per participant.

Findings

After looking at the 275 user responses received and running data analysis in R, we found that the three parameters that caused the most discomfort and/or pain were vibro-tactile (average ~3/5), while the three that caused the least discomfort were auditory (average ~1.25/5), with 1 being “not at all” and 5 being “very”.