Supporting Augmented and Alternative Communication using a Low-Cost Gestural Device

Matt Wheeler, Flynn Wolf and Ravi Kuber

UMBC

{ mattw1, flynn.wolf, rkuber } @umbc.edu

ABSTRACT

In this paper, we describe an exploratory study to determine the feasibility of using a low-cost gestural headset to support communication. Findings have shown tasks involving facial gestures, such as blinks and smiles, can be performed and detected by an Augmented and Alternative Communication (AAC) system within a shorter period of time compared to brow movements. As tasks increase in complexity, rates of accuracy and time taken remain relatively constant for blinking gestures, highlighting their potential in AAC interfaces. We aim to refine such a system to better address the needs of individuals with disabilities, by limiting input errors from involuntary movements and examining ways to reduce interface navigation time. Insights gained from the study offer promise to interface designers seeking to widen access to their interfaces using gestural input.

Categories and Subject Descriptors

H.5.2 User Interfaces – *Input devices and strategies*

General Terms

Human Factors.

Keywords

Augmented and alternative communication;

1. INTRODUCTION

The ability to communicate with others is often taken for granted. However, the case is not so straightforward for individuals with conditions such as Motor Neuron Disease, where speech may be limited and movement may be difficult. The inability to communicate effectively may affect self-perception, independence, and access to health care [1]. In response to this situation, Augmentative and Alternative Communication (AAC) devices have been developed, with the aim of presenting characters, words, or symbols to communicate with others, which can be selected through eye blinks or through performing other gestures. These devices are known to be expensive, and are often constrained to a small range of vocabulary.

In this paper, we report the results of an exploratory study undertaken to identify the feasibility of using a low-cost gestural headset to support personal communication. More specifically, we examined the efficacy of using a variety of facial gestures to select targets on a GUI, in order to construct sentences. The study represents a first step towards using low-cost gestural technologies to augment personal communication among individuals with physical disabilities.

2. RELATED WORK

AAC devices have been developed which can be accessed either directly (e.g. with a head pointer, joystick etc.) or indirectly (e.g. scanning techniques are used to move through a range of letters or icons, and a switch can be used to select the item). More recently, gestural technologies have been used to support individuals who experience challenges communicating. Lievesley et al. [3] evaluated the efficacy of using a gestural device to control a computer using facial expressions and conscious thought. Results showed that the device was slower and less accurate compared to head switches. However, the solution described in their paper was thought to offer promise to disabled users to control assistive technologies.

As gestural technologies improve in fidelity and reduce in price, they offer a feasible alternative to individuals with physical disabilities, and provide a means for interaction with an interface using available voluntary movements. Gestural technologies also alleviate the need for speech-based input, which can pose difficulties for some individuals with physical disabilities. In the research described in this paper, we aim to examine the efficacy of detecting facial gestures to support personal communication processes.

3. INTERFACE DESIGN

An AAC system was developed using Java. The GUI resembles a communication board (Figure 1), often used by individuals with disabilities, with vocabulary and corresponding icons arranged in three 4x4 grids. Similar to other AAC technologies, vocabulary is presented on multiple screens, due to the constraints of the interface. These can be accessed through tabs towards the top of the screen. The Emotiv Epoc neuroheadset (Figure 2) was selected to interface with the system, as it could be used to detect expressions using EEG sensors, and head motion using gyroscopic features embedded within the device.

To use the system, the user can move his/her head to control the on-screen cursor. He/she is able to construct sentences by selecting a term by performing a range of facial gestures. Blinking, eye-brow raises and smiling were selected for our study, as pilot tests had revealed these could be recognized with greater levels of accuracy compared with other facial

expressions. Selected words appear in the lower panel on the screen. The user is then able to search for successive words or symbols across the 3 tabbed screens, to compose the remainder of the sentence. Tabs can be selected using the same gestures.

Figure 1: AAC Interface. Each of the 3 screens displays 16 words/icons. Each screen can be accessed via tabs at the top of the interface.

Figure 2: Emotiv Epoc Neuroheadset (www.emotiv.com)

4. STUDY DESIGN

Twelve participants (aged 18-29) were recruited for the study. Individuals without disabilities were selected due to the exploratory nature of the study, and the difficulties associated with recruiting individuals with physical disabilities. Participants were asked to wear the headset, and were presented with a series of calibration tasks. Participants were asked to reproduce a set of sentences using head movements and gestures. Examples included statements which may be used by individuals with disabilities in an educational environment (e.g. "Want more volume classmates", meaning the speaker wants classmates to speak more loudly). Nine sentences were presented for each of the three gestural conditions (smiling, blinking and brow movements). The tasks required identifying terms from up to three tabs, as this was deemed to be a realistic task faced by individuals with disabilities when interacting with constrained AAC interfaces. Tasks were presented in a randomized order.

5. RESULTS

Findings showed that participants were able to perform tasks with an 88.2% rate of accuracy (SD: 7.8%), spending on average 21.9s (SD: 4.9s) to construct sentences using the headset. A repeated measures ANOVA performed on the task time taken showed the presence of a significant effect ($F_{(2, 22)}=6.053$, $p<0.01$). Posthoc analysis (Bonferroni corrected) confirmed that tasks involving blinking and smiling gestures were performed and detected within a shorter period of time using the system compared with tasks involving raised brow movements (Blinking - M: 20.0s, SD: 4.6s, $p<0.05$; Smiling - M: 20.3s, SD: 3.6s, $p<0.05$).

Participants were asked to construct sentences using terms and icons spanning up to three tabs. Findings showed that as more tabs were accessed, task time increased (1 tab - 20.2s, 3 tabs – 24.4s). In contrast, levels of accuracy appeared to decrease when constructing sentences using more tabs (1 tab – 90.3%, 3 tabs – 86.2%). The rate of decline was more pronounced for tasks performed involving smiling and brow raising gestures.

6. DISCUSSION

The results from our study highlighted the efficacy of specific gestures to support the personal communication processes. Tasks involving blinking (20.0s) and smiling gestures (20.3s) were

detected within a shorter period of time, compared with tasks involving brow raises (24.9s, $p<0.05$). Comments revealed that when raising the brow, the cursor would inadvertently move as participants would move their heads slightly, resulting in the selection of an unintended target. As a result, time would be taken to recover from the error. Participants reported the raised brow input method seemed less natural to perform compared with blinking or smiling, adding an additional level of complexity to the interaction.

Regardless of the number of tabs spanned, results for accuracy and time taken appeared to be most consistent when performing tasks involving blinking gestures. Although our earlier studies had highlighted challenges associated with blink detection [2], the results described in this paper may have in part been attributed to additional gestural interface training provided to participants. Time was also invested in training the software to recognize the facial gestures performed by each participant.

Participants suggested that integrating predictive capabilities within the system could provide one method of reducing task time and fatigue incurred, which would specifically reduce the burden on individuals with physical disabilities. Suggestions also including presenting the user with a 'Did You Mean?' option, in the event that an error was made when composing a sentence. This would enable the user to recover from the error, rather than spending time and mental effort deleting and re-selecting terms. Participants thought the interface could easily be customized, enabling disabled users and/or their caregivers to upload images and terms relevant to the user, addressing the generic nature of AAC devices. This would be particularly useful, as over time, the user's interests may change. Vocabulary would need to be updated accordingly.

7. CONCLUSIONS AND FUTURE WORK

The study has revealed that tasks involving blinking and smiling gestures were performed in a shorter period of time compared with tasks involving brow raises. Blinks were found to be detected most consistently, irrespective of sentence length.

Further work will examine ways in which algorithms for detecting gestures can be refined to cater to the abilities of individuals with limited movement capabilities, such as individuals able to make only small brow movements, or those who make involuntary movements which may affect interaction with the interface. A study will then be performed with individuals with disabilities to compare our solution with other commercially-available AAC devices, to determine the merits of gesture-based interaction to support personal communication.

8. REFERENCES

[1] Augmented Communication Inc., 2012 http://www.augcominc.com/whatsnew/impacts.html/

[2] Krishnaswamy, K. and Kuber, R., 2012. Toward the Development of a BCI and Gestural Interface to Support Individuals with Physical Disabilities. In *Proc. ASSETS'12*, 229-230.

[3] Lievesley, R. et al., 2011. The Emotiv Epoc neuroheadset. An inexpensive method of controlling assistive technologies using facial expressions and thoughts? *Journal of Assistive Technologies*, 5 (2), 67–82.