Gaze communication systems for people with ALS

John Paulin Hansen (IT University of Copenhagen, Denmark)
Håkon Lund (Royal School of Library and Information Science, Denmark)
Hirotaka Aoki & Kenji Itoh (Tokyo Institute of Technology, Japan)

1. Introduction

Gaze interaction systems are powerful communication tools for people who are not able to control their hands. The technology has been available for more than 20 years (Majaranta and Raiha, 2000) and is still improving. It supports severely disabled people such as Amyotrophic Lateral Sclerosis (ALS) patients who have lost their ability to control muscle movement. They can communicate with their family just gazing at a particular part of a computer screen with characters or commands.

Gaze communication systems work by a video camera that records the fast movements of one or both eyes. In some systems the camera is mounted on the computer screen. Other systems have a camera that can be positioned independently between the user and the computer. Most system use infrared light that is shine into the users eye. This gives a clear image and a reflection (a so called “glint”) in the eye that is easy for the computer vision system to identify. The user is required to look at 5 - 12 points on the screen before start, to calibrate the location of the pointer with their actual fixation points. This process normally takes 1 - 2 minutes. The systems cost between 10.000 $/_ and 20.000 $/_ (in year 2006, but prices are expected to drop)..

On-screen keyboards (e.g., “Point for Windows”, “Wivik” and others) are well suited in combination with eye trackers, and they have been used in type-to-talk systems for decades. Character sets can be arranged in a traditional QWERTY order, in alphabetic order or according to frequency-of-use. Some systems offer predictions of the most likely next characters and words or give access to a dictionary of words related to a specific context such as “dinnertime”. While it's unlikely that a gaze controlled communication system will ever achieve communication rates comparable to unimpeded speech, which may be more than 150 words per minute for a fast-talking person, (300 Japanese characters per minute), the long term goal is to achieve an input rate that is comparable to the QWERTY keyboard for typing, as this is usually sufficient to partake in on-line conversations (“chatting”). Present day eye gaze communication is efficient at a level that compares to text input methods on mobile devices (PDA and mobile phones), typically around 10 words per minute/ 20 Japanese characters per minute. One particular system, Dasher (Ward & Mackay, 2000), shows promise for expert users reaching the level of Morse code and cursive handwriting, typically around 25 words per minute/50 Japanese characters per minute.

Some of the more expensive systems have sufficient precision to allow control of a mouse-pointer in a Windows-environment. The screen resolution may have to be at a rather low level (e.g. 600 x 800 pixels), and larger-than-normal icon size may have to be applied. A zooming principle that works with normal size icons and at a higher resolution is also provided by some of the manufactures. Advanced gaze tracking

systems has dwell-time selections of icons, menus and links. Dwell selection works by looking steady at an object or letter for a preset time. Most users start with a dwell time setting at more than one second, but some experienced users can control a dwell time setting that are less than half a second. Most users like to have longer dwell times when they are tired in order to avoid unintended selections. Some systems allow the user to select by eye blinking, but this can be very tiresome to use for typing.

There are only a few reports on gaze control of wheelchairs and this is still an area of research and development, because gaze cannot be reliably controlled in a dynamic environment where, for instance, other people walk around. While it can be annoying to do an unintended selection on a personal computer, it may cause a serious accident if it happens while driving.

2. Communication abilities at different stages of ALS

The main recommendation that we would like to give people with ALS is to start preparing for gaze interaction before it is actually needed. It will take some time to learn to master it well, typically 5 hours of continuous typing (Itoh et. al., 2006). Most of the gaze communication systems also support other kind of input, so it should be possible to introduce the gaze communication interface when the user can still move a joystick, mouse, or head tracker. If the users start preparing early, most of the systems for western languages may adapt to the users specific vocabulary by remembering the words that is typed.

The progression through the stages of ALS is gradual, and the fatigue factor often makes it necessary for the user to switch to a less demanding input method during the day. Table 1 shows a possible development of disabilities associated with ALS. There are many individual differences from this general pattern as some people remain at the same stage for a very long time, some people may be able to speak even when they have lost control of most other muscles and some people never progress to the later stages. So **Table 1** is only meant to illustrate, that in general, it is wise to consider what might come next, and to have a range of input means ready for this.

A male ALS patient described his concern about a potential gradual loss of his eye control (i.e. going from 4^{th} to 5^{th} stage) in an e-mail he send to us "This is an important question, since many final stage people with ALS (PALS) experience that their eye muscles become weaker and weaker. This is contrary to what is written in most textbooks on ALS, but it is actually happening. I can mention that some

Table 1: Typical progression of ALS (modified from Johansen and Hansen [2006])

	Symptoms	Input devices
1st stage	Fatigue is noticeable. Reduced mobility and strength in arms and hands. Often slurred speech.	Keyboard with hand/arm rests and modified operation (sticky shift, no repeat).
2nd stage	Fatigue is a factor. Unable to move arms due to lack of strength, but mobility is usually retained in one or both hands. Severely slurred speech, largely unintelligible to outsiders.	Mouse, joystick, or reduced keyboard (5-10 keys).
3rd stage	Almost full lock-in. No speech function. Severely reduced mobility of all extremities.	One or two switches, eye or head tracking.
4th stage	Full lock-in with control of eye movements.	Eye tracking, possibly supplemented with e.g. EMG switch selection or eye blink selection.
5 th stage	Full lock-in with no eye movements	Step scan selection by bio-potentials (for instance EMG, EEG or EOG)

Japanese PALS, who have survived 20 years using ventilator, can only stare straight out. Secondly, many final stage PALS take drugs (Scopoderm, Atropin etc) against saliva. Unfortunately, these drugs interfere with the sight leading to loss of precision and accurateness."

In the 5th stage, an EMG or EEG click may be sufficient to operate the communication system that the user is familiar with from gaze interaction. But again, we recommend that EMG selections get introduced earlier on. There is another reason to start using EMG-switches: EMG clicks are potentially more robust and faster than dwell clicking, especially if the user needs a dwell time of more than half a second, so they work well in combination with gaze interaction.

3. The GazeTalk system

GazeTalk has been designed in close collaboration with a group of ALS people to allow them to communicate efficiently with others and to control computer programs for entertainment and information. Is has a very limited number of on-screen keyboards and may thus be operated even with highly disturbed gaze control. It features type-to-talk (by synthetic voice), writing, e-mail, web browsing, video playing, music (MP3) playing and smooth reading of PDF-files. GazeTalk provides direct access to the very fast Dasher typing system, and the user can switch from different language versions without leaving the system. GazeTalk is now available as freeware in a Danish, English, Italian, Chinese and Japanese version and several other language versions are expected in the years to come. It may be downloaded free-of-charge from www.cogain.org/downloads. The system can be operated by a keyboard, a mouse, a tracker-ball, a joystick, a head-tracker and an eye-tracker. It also supports step-scan selection with a mouse button, an EMG-switch or any other type of single-switch that user can control. Figure 1 shows screenshots from the English.

The original version of GazeTalk was developed for western languages. It includes a character prediction function, which predicts the most likely six letters subsequent to the last typed character and predicts the most like next word. This version employed a dynamic menu structure in which a key of a character could change its position. The Japanese version of GazeTalk has a hierarchical menu structure: At the top menu, , an entry for each of the above-mentioned Hiragana groups is allocated to each key in the lower and middle rows of the menu. When one presses, for example, the bottom left key in Figure 1 by gazing at it for a specific dwell time, then a next-level menu appears in which five Hiragana symbols composing this group are included: "na", "ni", "nu", "ne", and "no". Then, in this "character-level" menu, a Hiragana symbol can be typed by fixating one's eyes at the key which one wants to input. Thus, a Hiragana symbol can be typed by two selections for most characters. The Japanese version of GazeTalk also includes a "Kana-Kanji conversion" program. For example, a key for initiating the conversion is allocated to the right of the text area in **Figure 1**.

The next Danish version of GazeTalk will provide direct access from the users home to a large amount of digital library resources available at the Danish Royal Library. The main purpose of the project is to develop a generic model for a web-service architecture that can be addressed from within GazeTalk, providing the end user with tools for e.g. navigating and searching a book collection and downloading from it. GazeTalk will support browsing on title lists and author names. Furthermore, a simple search facility will also be implemented allowing the user to search single title words and parts of author names. The books (or texts) found by the user will be made available in the PDF - format for download and the existing PDF reader function in GazeTalk can then be used to view the downloaded files.

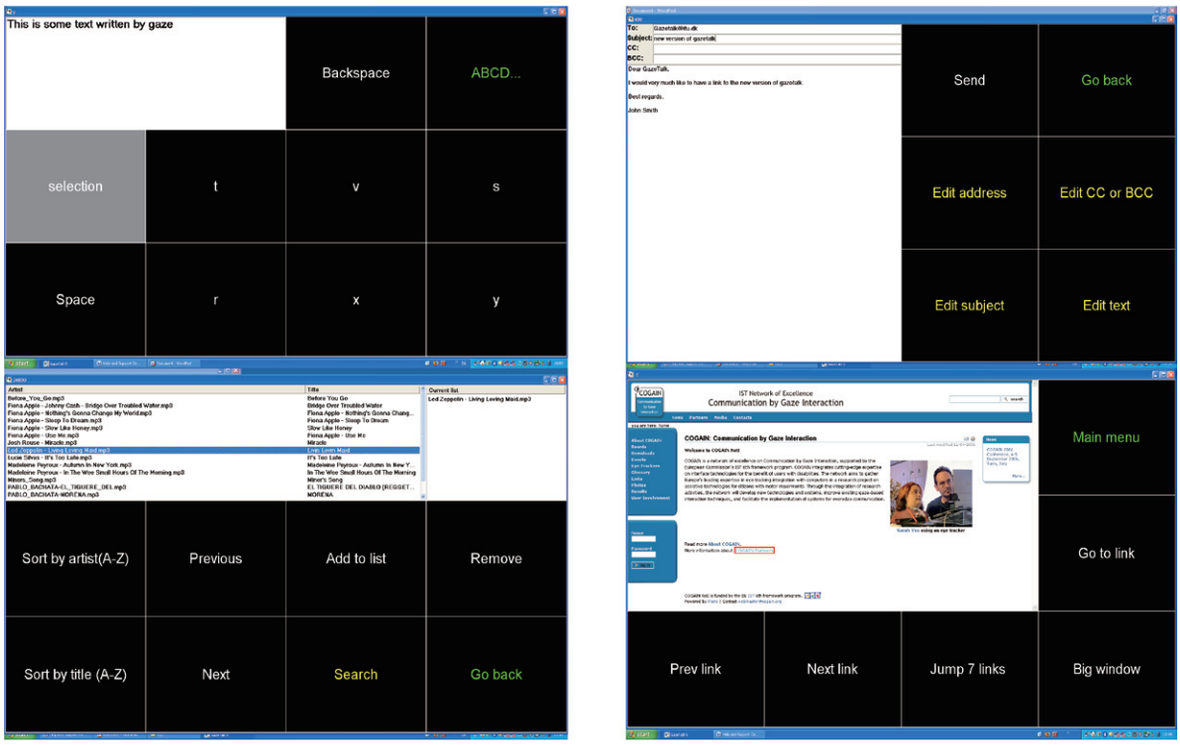

Figure 1: Typing interface (top left), e-mail interface (top right), music player (bottom left) and web browser (bottom right) in GazeTalk.

4. Conclusion

Gaze interaction appears to be one of the most desirable alternatives for individuals who cannot direct cursor movement with their hands. It should be introduced for people with ALS at an early stage in order for them to learn how to master gaze communication system gradually. There is a range of dedicated gaze communication systems available, and one particular system, GazeTalk, which supports communication on several stages of ALS has been presented. With the GazeTalk system the user can speak, write, send e-mail, browse the Internet and enjoy music and video.

Litterature:

Majaranta, P., and Räihä, K. J. 2002. Twenty Years of Eye Typing: Systems and Design Issues. In Proceedings of the Symposium on ETRA 2002: Eye Tracking Research & Applications Symposium 2002, New Orleans, LA, 15-22.

Johansen, A. S. and Hansen, J. P. 2006. Augmentative and alternative communication: the future of text on the move, Universal Access in the Information Society, Jan 2006, 1 - 25.

Itoh, K., Aoki, H., and Hansen, J. P. 2006. A comparative usability study of two Japanese gaze typing systems. In Proceedings of the 2006 Symposium on Eye Tracking Research & Applications (San Diego, California, March 27 - 29, 2006). ETRA '06. ACM Press, New York, NY, 59-66.

Ward, D. J., and MacKay, D. J. C. 2002. Fast Hands-Free Writing by Gaze Direction. Nature 418, p. 838, August 22.