Simone Eidam, Jens Garstka, and Gabriele Peters,Towards Regaining Mobility Through Virtual Presence for Patients with Locked-in Syndrome, accepted for: 8th International Conference on Advanced Cognitive Technologies and Applications (COGNITIVE 2016), Rome, Italy, March 20-24, 2016.

Towards Regaining Mobility Through Virtual Presence for Patients with Locked-in Syndrome

Simone Eidam*, Jens Garstka†, and Gabriele Peters†

Human-Computer Interaction, Faculty of Mathematics and Computer Science
University of Hagen
D-58084 Hagen, Germany
Email: eidam.simone@gmail.com*, {jens.garstka, gabriele.peters}@fernuni-hagen.de†

***Abstract*—The emergent technology of virtual presence systems opens up new possibilities for locked-in syndrome patients to regain mobility and interaction with their familiar environment. The classic locked-in syndrome is a state of paralysis of all four limbs while retaining full consciousness. Likewise, there is a paralysis of the vocal tract and respiration. Thus, the central problem consists in controlling a system only by eyes. In this paper, we present a prototype of a communication interface for patients with locked-in syndrome. The system allows the localization and identification of objects in a view of the local environment with the help of an eye tracking device. The selected objects can be used to express needs or to interact with the environment in a direct way (e. g., to switch the lights of the room on or off). The long term goal of the system is to give locked-in syndrome patients a larger flexibility and a new degree of freedom.**

***Keywords*–Biomedical communication; Human computer interaction; Eye tracking.**

I. INTRODUCTION

It is undoubtedly a major challenge for locked-in syndrome (LIS) patients to communicate with their environment and to express their needs. Patients with LIS have, for example, to face severe limitations in their daily life. LIS is mostly the result of a stroke of the ventral pons in the brainstem [1]. The incurred impairments of the pons cause paralysis, but the person keeps his or her clear consciousness. The grade of paralysis determines the type of LIS and has been classified in classic, total and incomplete LIS. Incomplete LIS means that some parts of the body are motile. Total LIS patients are like classic LIS patients completely paralyzed. However, the latter ones still can perform eyelid movements and vertical eye movements that can be used for communication. Therefore, several communication systems for classic LIS patients have been designed in the past.

This paper provides, in contrast to already existing solutions, an alternative way to enable LIS patients to use eye gaze and eye gestures to communicate with their environment. In the presented prototypic environment, the patients will see exemplary scenes of the local environment instead of the typically used on-screen keyboard. These scenes contain everyday objects, e. g., a book the impaired person wants to read, which can be selected using a special eye gesture. After selection, the patient can choose one of various actions, e.g., “I want to read a book” or “please, turn the page over”. A selection can either lead to a direct action (light on/off) or to a notification of a caregiver via text-to-speech.

In a long-term perspective, the aim is to build a system where the screen shows a live view of the environment captured by a virtual presence systems (VPS). The LIS patient should also have the ability to control the VPS. This enables the patient to directly interact with its environment. To be able to perform direct actions with some everyday objects (e. g., a light switch), the system has to be extended with an object-recognition approach.

The text is structured as follows: Section II describes related work and other communication systems using eye tracking. Section III describes the prototype design with the implementation of eye gesture recognition and simple object recognition. In Section IV and V, the evaluation results will be presented and discussed. This work will be concluded alongside a description of future work in Section VI.

II. RELATED WORK

This section gives a brief overview on eye tracking and already existing systems that support LIS patients with their communication.

A. *Eye Tracking*

Many existing eye tracking systems use the one or other kind of light reflection on eyes to determine the direction of view. The human eye reflects incident light at several layers. The eye tracking device used for controlling the prototype employs the so-called method of dark-pupil tracking. Dark-pupil-tracking belongs to the video-based eye tracking methods. Further examples are bright-pupil- and dual-Purkinje-tracking [2].

For video-based systems, a light source (typically infrared light) is set up in a given angle to the eye. The pupils are tracked with a camera and the recorded positions of pupil and reflections are analyzed. Based on the pupil and reflection information, the point of regard (POR) can be calculated [2]. In Figure 1, the white spot just below the pupil shows a reflection of an infrared light on the cornea. This reflection is called the glint. In case of dark-pupil tracking, it is important to detect both, the pupil center and the glint. The position of the pupil center provides the main information about the eye gaze direction while the glint position is used as reference. Since every person has individually shaped pupils, a onetime

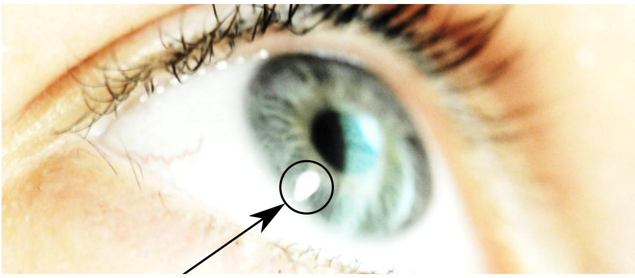

Figure 1. The reflection of the infrared light on the cornea.

Figure 2. An example scene used with this prototype.

calibration is needed. In case of a stationary eye tracker, also the distance between the eyes is determined to calculate the position of the head relative to the eye tracker.

B. Other Systems

There are many prototypes that have been developed in order to support LIS patients with their communication. Many of them are video-based eye tracking systems. One of the first systems was the communication project ERICA developed in 1989 [3]. With the help of the system users were enabled to control menus with eyes. They were able to play computer games, to hear digitized music, to use educational programs and to use a small library of books and other texts. Additionally, ERICA offered the possibility to synthesize speech and control nearby devices. Currently available and commercial communication systems for LIS patients are basically based on ERICA. These systems include the Eyegaze Edge Talker from LC Technologies and the Tobii Dynavox PCEye Go Series. The Tobii solution provides another interaction possibility called "Gaze Selection" in addition to an eye controlled mouse emulation. It allows a two stage selection, whereas starring at the task bar on the right side of the screen enables a selection of mouse options like right/left button click or the icon to display a keyboard. Subsequently, starring on a regular GUI-element triggers the final event (such as "open document"). Two-stage means that the gaze on the target task triggers a zoom-in event. It is said, that this interaction solution is more accurate, faster and reduces unwanted clicks in comparison to a single stage interaction.

Furthermore, current studies present alternative eye based communication systems for LIS patients. For example, the prototype developed by Arai and Mardiyanto [4], which controls the application surface using an eye gaze controlled mouse cursor with the eyelids to trigger the respective events. This prototype offers the possibility to phone, to visit websites, to read e-books, or to watch TV. An infrared sensor/emitter-based eye tracking prototype was developed from Liu et al. [5], which represents a low-cost alternative to the usual expensive video-based systems. With this eye tracking principle, only up/down/right/left eye gaze moves can be detected as well as staying in the center using the eyelids to trigger an event. By using the eye movement, the user can move a cursor in a 3×3 grid from field to field. And by using the eyelids, the user can finally select the target field. Barea, Boquete, Mazo, and Lpez [6] developed another prototype which is based on electrooculography. This prototype allows by means of eye movements to control a wheelchair allowing an LIS patient to freely move through the room.

All prototypes that have been discussed so far are based on an interaction with static contents on screen, for example of a displayed 2-D keyboard. However, the prototype presented in this contribution shows a path to select objects in a 2-D picture by a simulated object recognition. This allows an evaluation of the system without the need of a full recognition engine. The latter will lead to a selection of real objects in the patient's proximity.

III. OUR METHOD

Now, we will present the concept and implementation details of our method.

A. Concept

The following section provides an overview of the basic concept of this work. As already mentioned, the impaired person will see an image of a scene with typical everyday objects. This image is representative for a real scene, which is to be captured by a camera and analyzed by an object recognition framework in future work. Figure 2 shows an image of one possible scene. The plant can be used by a LIS patient to let a caregiver know, that one would like to be in the garden or park, the TV can be used to express the desire to watch TV, while the remote control directly relates to the function of the room light. The red circle shown at the center of the TV illustrates the point of regard (POR) calculated by the eye tracker. The visual feedback by the circle can be activated or deactivated, depending on individual preferences.

An object is selected by starring a predetermined time on the object, what we call a "fixation". With a successful fixation a set of options will be displayed on the screen. A closing of the eyelids is used to choose one of these options. Depending on the selected object, a direct action (e. g., light on/off) or an audio synthesis of a corresponding text is triggered (e. g., "I would like to read a book.").

Furthermore, other eye gestures have been implemented to control the prototypes. By means of a horizontal eye movement, the object image is changed. And by means of a vertical eye movement, the visual indication of the POR can be switched on and off.

B. Implementation

The eye tracking hardware used is a stationary unit with the name RED manufactured by SensoMotoric Instruments (SMI). RED comes with a Notebook running a controller

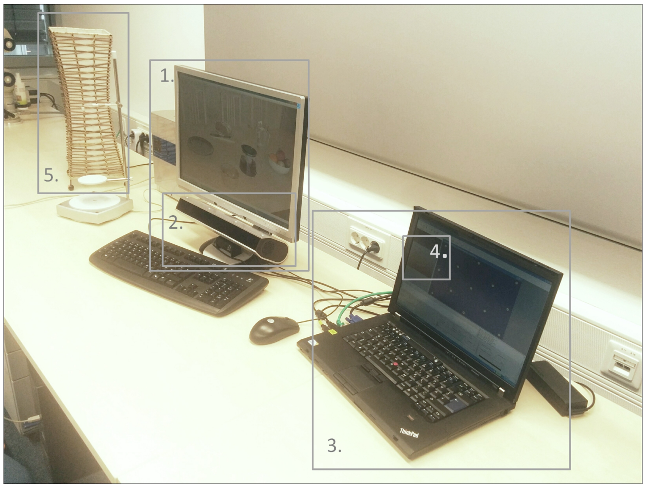

Figure 3. This figure shows the components of the prototype.

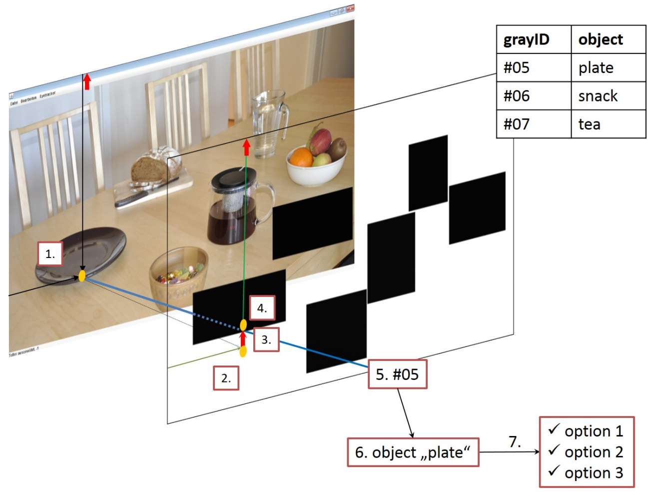

Figure 4. Simulated object recognition.

software. The latter provides a network component to allow an easy communication between the hardware and any software through a well-defined network protocol.

Figure 3 gives a brief overview of all components of our prototype. Area 1 shows the patient's components to display test scenes with different objects. The stationary eye tracking unit is shown in area 2. Area 3 shows the eye tracking workstation with the eye tracking control software in area 4. Finally, area 5 contains a desk lamp, which can be turned on and off directly with a fixation of the remote control shown in Figure 2.

C. Eye Gesture Recognition

Eye gesture recognition is based on the following principle: the received POR-coordinates from the eye tracker are stored in circular buffer. At each coordinate insertion the buffer is analyzed for eye gestures. These eye gestures are a fixation, a closing of the eyelids, and a horizontal/vertical eye movement. The following values can be used to detect these eye gestures:

- the maximum x- and y-value: $x_{\max}$, $y_{\max}$
- the minimum x- and y-value: $x_{\min}$, $y_{\min}$
- the number of subsequent zero values: c

The detection of the fixation is performed as follows:

$$|x_{\max} - x_{\min}| + |y_{\max} - y_{\min}| \leq d_{\max}, \quad (1)$$

where $d_{\max}$ is the maximum dispersion while the eye movements are still recognized as fixation. The value of $d_{\max}$ is individually adjustable.

The detection of a closing of the eyelids is realized by counting the amount c of subsequent coordinate pairs with zero values for x and y. Zeros are transmitted by the eye tracker, when the eyes couldn't be recognized. This occurs on the one hand when the eyelids are closed, but on the other hand when the user turns the head or disappears from the field of view of the eye tracker. Therefore, this event should only be detected if the number of zeros corresponds to a given time interval:

$$(c > c_{\min}) \wedge (c < c_{\max}) \quad (2)$$

All variables $c_{\min}$ and $c_{\max}$ can be customized by the impaired person or the caregiver, respectively.

The combination of these two different approaches is a benefit, because object selection is realized through the fixation while option selection is done by closing the eyelids. The latter allows the LIS patient to rest the eyes while the option panel is open. Hence, the patient can calmly look over the offered options in order to get an overview.

For the horizontal eye gesture detection, a given range of x-values must be exceeded while the y-values remain in a small range, and vice versa for the vertical eye gesture. As already mentioned, the horizontal eye movement is used to switch between different images. But this functionality is not a part of a later system and is merely a simple additional operation to present a variety of objects while using this prototype. The vertical eye movement (vertical eye gesture) is used to enable or disable the visual feedback of the POR. While the visual presentation of the POR may interfere with the passing of time, the marker can be used to check the accuracy of the eye tracking.

D. Simulated Object Recognition

Figure 4 shows schematically the principle of the simulated object recognition. It is based on a gray-scale image that serves as a mask for the scene image. On this mask the available objects from the scene image are filled with a certain gray value. Thus, each object can be identified by a unique gray value ($grayID$). The rear plane illustrates the screen. The coordinates that correspond to a fixation of an object (1.) refer to the screen and not to a potentially smaller object image. Thus, these raw coordinates require a correction by an offset (2. & 3.). The corrected values correspond to a pixel (4.) whose value (5.) may belong to one of the objects shown. In case of the example illustrated in Figure 4 this pixel has a gray value of 5 which corresponds to the object "plate" (6.). Finally, either all available options will be displayed (7.) or nothing will happen in the case the coordinates do not refer to a known object.

IV. RESULTS

The prototype has been tested by five non-impaired persons to analyze its basic usability. Figure 5 briefly illustrates the results of the usability test. It shows whether a test person

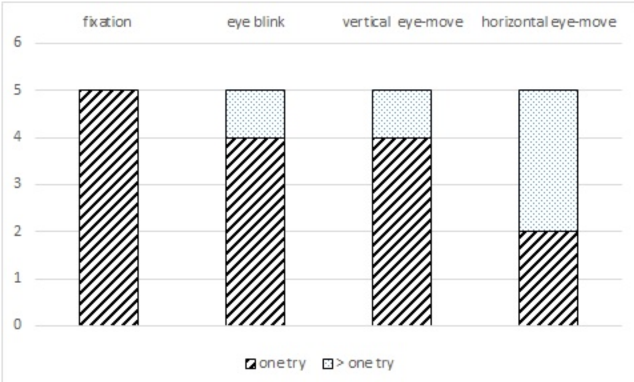

Figure 5. Bar diagram of the eye gesture recognition.

(subject) required one or more attempts to use a specific function successfully. During these tests, the subjects were able to validate the detected position of the eye tracker by means of the POR visualization. The diagram shows that none of the test persons had problems with the fixation. While the options were selected due to closing the eyelids, only one subject required several attempts. The same applies to the vertical eye movement. In a second pass, it turned out that precisely this subject requires other settings for a successful eye gesture recognition. Thus, more time for training and personal settings will help to achieve better results. However, it should be stated that this combination of object selection via fixation and option selection by closing the eyelids turned out to be a workable solution. Figure 5 further shows that three of five test persons had difficulties to deal with the horizontal eye movement. Interviews with the subjects showed that it appears to be very difficult to control the horizontal eye movement to get a straight motion. Apart from that, it must be considered that in general LIS patients are not able to do horizontal eye movements.

Apart from the latter, the usability can be assessed as stable and accurate. With a well-calibrated eye tracker, the basic handling consisting of the combination of fixation and a closing of the eyelids is perceived as comfortable. Additionally, it is possible to adjust the eye gesture settings individually at any time. This enables an impaired person to achieve optimal eye gesture-recognition results and a reliable handling.

V. Discussion

Since this work is in progress, there are different parts of this work that need to be discussed, implemented and evaluated in the near future. We list the main points – even in parts – below:

- Currently, the LIS patient cannot deactivate the eye tracking. Thus, there should be a way to disable the fixation detection. Since eye gestures based eye movements have proved to be difficult, our idea is a combination of two consecutive fixations, e. g., in the upper left and lower right corners.
- Instead of the currently used static pictures a live view of the VPS will be shown. This requires that the LIS patient can control the VPS. For this purpose, we would use the same control scheme as in the previous point, but with the lower left and upper right corners. When enabling the VPS control this way, the eyes can be used like a joystick to move the VPS. The joystick can be temporarily disabled or enabled by closing the eyelids.
- A major part of this work will be the recognition of a useful set of everyday objects. Recently, deep convolutional neural networks trained from large datasets have considerably improved the performance of object recognition. At the moment, they represent our first choice.

In addition, there are many other minor issues to deal with. However, at this point these issues are not listed individually.

VI. Conclusion and Future Work

The presented prototype demonstrates a user-friendly and alternative communication interface that allows the localization and identification of objects in a 2-D image.

In contrast to the discussed state-of-art methods, which are based on an interaction with static content on screen, the direct interaction with the environment is a benefit in two ways. On the one hand, compared to the methods that use a virtual keyboard, our method is faster and less complex. And on the other hand, compared to the methods where pictograms are used, our method eliminates the search for the matching icon. Thus, the advantage of such a system is a larger flexibility and a greater interaction area, i. e., a direct connection to controllable things like the light, a TV, or a radio.

Future work will include a live view from a VPS and the possibility to individually select objects from the local environment. This will enable the patients to select real objects for communication tasks with the help of an eye tracker. On the one hand, this ensures a more intuitive interaction where the live view provides LIS patients a new degree of freedom where they can leave behind static contents on screen for communication purposes and can interact with the real environment. On the other hand, changes within the room (displacement or exchange of objects) do not affect the interaction range of the patients.

References

[1] E. Smith and M. Delargy, “Locked-in syndrome,” *BMJ: British Medical Journal*, vol. 330, no. 7488, pp. 406–409, 2005.

[2] A. Duchowski, *Eye Tracking Methodology, Theory and Practice*. Springer-Verlag, 2007, ch. Eye Tracking Techniques, pp. 51–59.

[3] T. E. Hutchinson, K. P. White, W. N. Martin, K. C. Reichert, and L. A. Frey, “Human-computer interaction using eye-gaze input,” *IEEE Systems, Man, and Cybernetics*, vol. 19, no. 6, pp. 1527–1534, November 1989.

[4] K. Arai and R. Mardiyanto, “Eye-based human computer interaction allowing phoning, reading e-book/e-comic/e-learning, internet browsing, and tv information extraction,” *IJACSA: International Journal of Advanced Computer Science and Applications*, vol. 2, no. 12, pp. 26–32, 2011.

[5] S. S. Liu, A. Rawicz, S. Rezaei, T. Ma, C. Zhang, K. Lin, and E. Wu, “An eye-gaze tracking and human computer interface system for people with als and other locked-in diseases,” *JMBE: Journal of Medical and Biological Engineering*, vol. 32, no. 2, pp. 111–116, 2011.

[6] R. Barea, L. Boquete, M. Mazo, and E. Lpez, “System for assisted mobility using eye movements based on electrooculography,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 10, no. 4, pp. 209–218, 2002.