Journal of Assistive Technologies
HeadBanger: controlling switchable software with head gesture
Matthew Oppenheim

Article information:

Users who downloaded this article also downloaded:

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Peer-reviewed paper

HeadBanger: controlling switchable software with head gesture

Matthew Oppenheim

Matthew Oppenheim is Honorary Researcher at School of Computing and Communications, Lancaster University, Lancaster, UK.

Abstract

Purpose – The purpose of this paper is to present a novel non-contact method of using head movement to control software without the need for wearable devices.
Design/methodology/approach – A webcam and software are used to track head position. When the head is moved through a virtual target, a keystroke is simulated. The system was assessed by participants with impaired mobility using Sensory Software's Grid 2 software as a test platform.
Findings – The target user group could effectively use this system to interact with switchable software.
Practical implications – Physical head switches could be replaced with virtual devices, reducing fatigue and dissatisfaction.
Originality/value – Using a webcam to control software using head gestures where the participant does not have to wear any specialised technology or a marker. This system is shown to be of benefit to motor impaired participants for operating switchable software.

Keywords Assistive technology, Software, Head gesture, Depth camera, Virtual control, Webcam
Paper type Research paper

Introduction

The principal contribution of this paper is in demonstrating a head gesture recognition system for mobility impaired users, using readily available webcam technology, which is easy and intuitive to operate and that could be practically implemented. This provides an extra interface for the target user group to interact with their switchable software, which is used for communication and environmental control. The participants do not need to wear any specialised technology or targets and the system is non-contact.

Human computer interaction (HCI) devices such as push buttons and head tracking mice exist to help people with mobility restrictions or control problems to interact with software using head motion. The limited range of physical interfaces available and the practical restrictions of their placement can lead to user fatigue and dissatisfaction. Having virtual controls that can be placed within the comfortable range of motion of the participant has the potential to reduce user fatigue and add an extra tool for helping those who have restricted mobility to interact with software.

The need for a non-contact and non-invasive human computer interface using head gestures was identified by the Assistive Technologist and Occupational Therapist professionals at Beaumont College in Lancaster who have extensive experience of working with people with disability. Beaumont College is a member of the charity Scope and has around 94 students, aged 18-23, many of whom have restricted mobility due to reasons such as cerebral palsy. The range of mobility varies with the individual, so there is no generic solution to enable the students at the College to effectively interact with software. Each student requires a custom arrangement of aids.

Switchable software such as Sensory Software's Grid 2 is used to enable communication and to control the environment using home automation devices such as network enabled lights and curtain opening devices. The aim of the head motion recognition system presented here is to create an extra tool to enable the students to easily interact with their software.

An evaluation of the system by volunteers from the target user group at Beaumont College demonstrated the usability and effectiveness of this system to control the Sensory Systems Grid 2 software.

Related work and devices

Current approaches for tracking head movement for HCI by disabled users include optical methods and inertial sensor methods.

A variety of commercial optical products aimed at disabled users exist, such as HeadMouse by Origin Instruments. This emulates a computer mouse by using an optical camera to track a small disposable target that is stuck onto the user's face. Prior research into virtual mice includes using a PC camera to track the nose position of participants to implement a head controlled mouse, using the participants' blinks to emulate clicking the mouse (Kim and Ryu, 2006).

Head tracking to control video games led to the development of several products, such as the TrackIR produced by NaturalPoint. This can track the head with six degrees of freedom, by using a small infra-red LED array that is attached to the user's head. A software development kit is available by license for the commercial development of games. Open source versions of this hardware exist, such as Free Track. This has the advantage of freely available open source software.

Gyroscope, accelerometer and magnetometer sensors mounted onto a helmet enable the 3D positioning of the head to facilitate the use of software by participants with cerebral palsy (Raya *et al.*, 2012). Both head mounted gyroscope sensors and head mounted accelerometer sensors have been implemented to enable disabled participants to control wheelchairs by using head movement (Rechy-Ramirez *et al.*, 2012; Pajkanovic and Dokic, 2013). These sensor-based systems necessitate the participant wearing bulky and intrusive instrumentation on their heads. It is unlikely that the target user group would be able to don this equipment unaided. Head mounted sensor systems can lead to false activations unless the operator has a method to disable them when the functionality they provide is not required.

Depth cameras have been used for recognising pre-defined hand gestures. The Microsoft Kinect depth camera has been used for recognising sign language, e.g. Chai *et al*. (2013) and Microsoft Research Asia (2013).

Eye trackers, such as the X2 range manufactured by Tobii which use the position of the eye pupil to control software are in use at Beaumont College. Some people find eye trackers difficult to use due to the requirement for consciously controlling the eye position. Eye tracking devices require calibration before use. At Beaumont, this technology tends to be used by students who have no other mode of interaction.

The system presented in this paper allows the participant to activate virtual controls through head movement, where the controls can be placed within the participant's comfortable range of motion and relative to their head position. The participant does not have to be placed in a certain position, rather the tracking system adapts to their natural sitting position. The participant can move their head in whichever way is comfortable and they are not restrained to interacting in a single plane, which is the case when using a head mouse. For instance, a nodding motion may be more desirable than moving the head sideways. No sensors or other equipment needs to be attached to the head. The enabling technology is unobtrusive compared with other technologies used for head movement tracking.

Hardware and software

Figure 1 shows the head tracking system being tested at Beaumont College, Lancaster and identifies the main components. The system has been called HeadBanger as this appeals to the students at Beaumont who are the target user group.

| Figure 1 | The HeadBanger system during testing at Beaumont College |
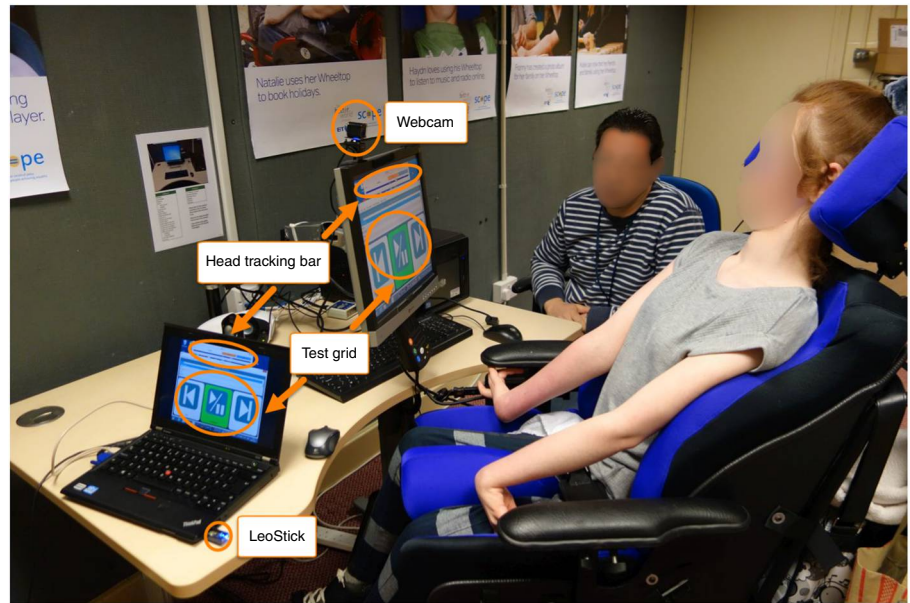
Notes: The HeadBanger software is running on the laptop, with an external monitor attached that mirrors the laptop's screen. The webcam is used to image the student's head. The LeoStick dongle is used to emulate keyboard strokes to interact with the test grid
Source: Author's own

Development

Several technologies were tested to arrive at the current implementation of HeadBanger. Initial development and testing of the idea of using virtual switches operated through head motion was implemented using depth camera technology; the Microsoft Kinect for Windows and the Asus Xtion Pro Live (Oppenheim, 2014). This work demonstrated the potential of the system to be useful for the students at Beaumont, but had several obstacles preventing the system from being implemented in a real world setting:

1. The minimum range that the depth cameras can reliably image over is too high for these devices to be mounted onto a wheelchair for individual use.

2. The depth cameras are physically intrusive for the target user group to have mounted on to their wheelchairs. Informational interviewing with the students at Beaumont identified that the target user group prefers discreet assistive technology.

3. The power requirements for running these devices could be an issue as all of the systems on the wheelchair need to be run from battery. The Microsoft Kinect requires 12V, for example.

This paper improves on this early work in two important ways:

1. Implementing the head tracking using webcam technology opens the method to being realistically integrated with a wheelchair mounted communication aid. This would enable it being used in a real world setting.

2. The virtual switches are placed relative to the participant's body position. The depth camera-based system requires the participant to be in a fixed position relative to the depth camera. This is not always a practical solution for somebody who already has mobility issues. The webcam system recognises the participant's head and neck and places the targets relative to the participant. This allows the student to move freely relative to the position of the webcam. In addition, this lowers the cognitive load of using the system compared with having to move to a fixed location to interact with a system.

Hardware

The system has been tested using a Microsoft LifeCam and a laptop (Lenovo X230) webcam. A multi-core processor is required to run the head tracking software. The system has been successfully tested on an Intel core i5 laptop (Lenovo X230) and on PCs with an AMD Phenom II X6 processor or an Intel i5 core processor. Problems have been encountered with computers using Intel core i3 processors as these are not powerful enough for the head tracking component of the system.

Software

The HeadBanger system uses the open source software FaceTrackNoIR. FaceTrackNoIR was developed to track head position and orientation using a webcam. This is to enable game players to interact with their software through head motion, e.g. to be able to look out of the rear of a simulated cockpit by slightly turning their head. A screen shot of this software is shown in Figure 2. The author of FaceTrackNoIR is aware of and supports the use of FaceTrackNoIR for the HeadBanger project.

The head position and orientation information can be extracted from FaceTrackNoIR through a virtual joystick control. This was implemented using a script written in Python 3.4. This script generates the graphical user interface (GUI) which displays the participant's head position relative to moveable virtual targets. Examples of this GUI are shown in Figures 3 and 4.

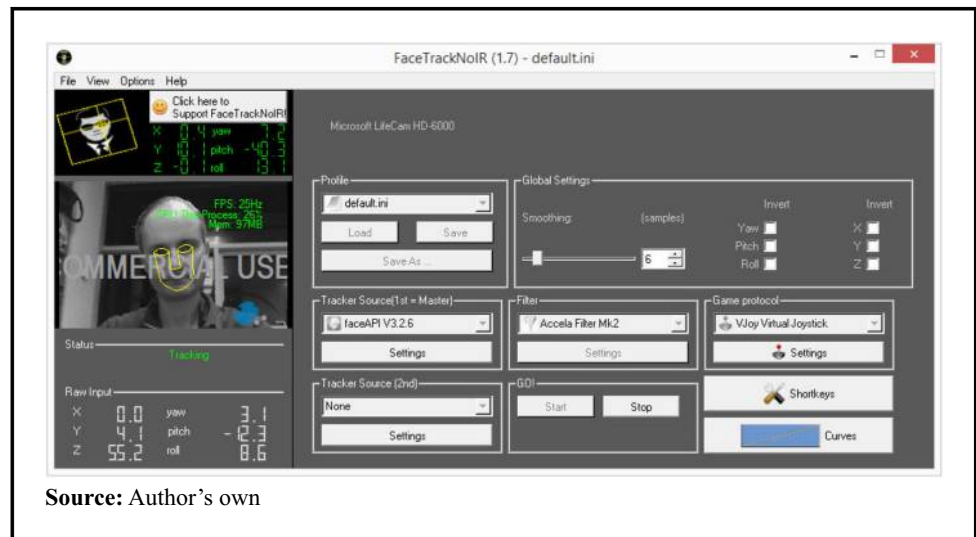**Figure 2** FaceTrackNoIR interface showing a head being tracked

Source: Author's own

Figure 3 HeadBanger interface showing repositioned targets

Source: Author's own
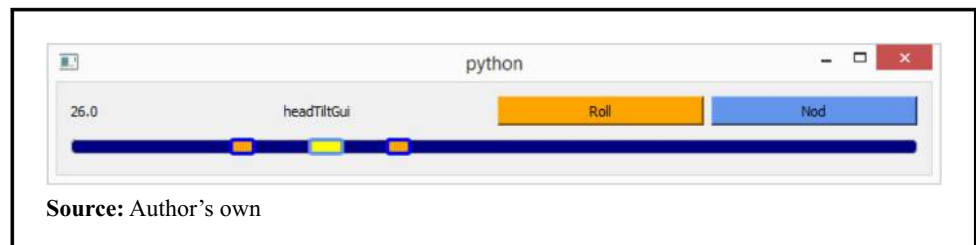
Figure 4 HeadBanger interface showing an activated target

Source: Author's own

Communicating with Sensory Software's Grid 2 software was problematic as the Grid 2 software disables the usual methods of external software control, e.g. by broadcasting keystrokes by using a software command. Sample code was sent from Sensory Software to show how to trigger their Grid 2 software purely through software. However, this requires the software to be run using an administrator's account, which made it difficult to distribute the code for testing. The solution found involved using a Freetronics LeoStick hardware board. This was used to emulate a physical keyboard, so that when a target was activated, a software signal from the Python script caused the LeoStick to emulate a physical keystroke. The LeoStick has the form factor of a small USB stick and can be seen in Figure 1.

HeadBanger has been tested on Windows 7 and Windows 8.1. The Windows operating system (OS) was chosen as the target user group use this OS to run their communication software (Sensory Software's Grid 2).

HeadBanger interface

Figures 3 and 4 show examples of the HeadBanger GUI. Figure 3 shows an example of the display shown to the participant where the targets have not been triggered. The head position is indicated by the position of the yellow marker on the blue bar and the targets by the orange markers. The targets can be repositioned using a mouse to accommodate the comfortable range of motion of the participant. A sideways (roll) or front to back head motion (nod) can be tracked by selecting the relevant button control. Figure 4 shows the head tracking bar with an activated target. In addition to the visual queue given by the change of colour, audio confirmation is generated by playing a drum beat when a target is activated. Each of the targets has a different drum beat to help the participant discern which has been activated. As many of the target user group have visual impairment, this audio confirmation was identified as an asset by the Occupational Therapists at Beaumont.

Evaluation

Ethical approval was obtained from Lancaster University to conduct user studies at Beaumont College. In addition, Beaumont College have their own internal protocols and risk assessments to cover user studies. Permission was obtained from all of the participants to appear in a project video.

Six students were recruited from the student population at Beaumont College, aged 18-21, four male, two female to evaluate the system. The students were selected by the Technologists working at Beaumont as being suitable candidates to test, assess and give feedback on the technology. The participants were rewarded with a £20 gift token after the user studies were complete, but were unaware of the reward at the time of the user studies. All of the participants were wheelchair users. Each of the participants had a different level of mobility. Five of the participants rely on using the Grid 2 software for creating speech using a communication device while one had not used the Grid 2 software before.

An interaction scenario based around a media player was set up using Sensory Software's Grid 2 package. This evaluation was used as a part of the Services for Enabling Technology Testing (SETT) project being led by National Star College for the Joint Information Systems Committee to

set up a framework for assessing novel assistive technology. The media player grid was set up by one of Beaumont College's Technologists who has several years' experience with designing grids for the students at Beaumont College. Five of the six participants were familiar with using Grid 2, but the specific grid and interaction scenarios had not been seen by these participants before. The interaction scenarios were:

1. toggle player on/pause;

2. advance one track;

3. rewind one track; and

4. stop the player.

Figure 5 shows the Grid 2 interface for the music player. The play/pause button can be toggled while the other controls each have a single function. The activated control changes colour to give a visual indicator to the participant that the control has been successfully operated. This grid system works using a "scan and select" system, where each grid square in turn lights up for a period of time and the participant selects the active square using either of the head targets. The speed of scanning can be set in the Grid 2 software. Five of the six participants were familiar with this mode of interaction from using switchable grid systems with their communication devices. Normally a students would have the scan speed and pattern of their grids tuned to their individual needs by a Technologist. For testing, a single scan rate was used throughout. The Grid 2 software accepts up to eight switches, each of which can be configured to have a separate function. The target user group has a range of cognitive abilities due to the individual nature of their disability, so for the purpose of assessing the technology, a simple single switch interaction was chosen. Either of the virtual head switches operated the same control. If the technology is adopted, each of the controls could be configured to operate different functions, e.g. one control could reverse the scan direction, while the second acts as the selector.

Each student typically spent 20 minutes assessing the system, then answered a series of questions on how they found using the system. The questionnaire was devised for the SETT project. The IBM usability satisfaction questionnaire (Lewis, 1993) is often used for HCI user study assessment, but previous experience with working with the students at Beaumont has shown that this is impractical due to the difficulty for the students with working through the document.

Each of the participants who had prior experience of the Grid 2 software had a different level of motor control and used a different interface for their normal interaction with the Grid 2 software.

Figure 5 Grid 2 music player interface. The play button is active

Source: Author's own

The interfaces normally used by the participants included physical head switches, hand operated push buttons and joysticks. Using a virtual control system instead of physical controls was novel for all of them.

Results

Of the six participants, five demonstrated successful task completion using the system once they became familiar with the system. For the successful participants, it typically took under one minute to adjust the position of the virtual targets to where they could be comfortably operated. One participant was unable to successfully interact with the system within the time limit of the user study. With further training she may well have become proficient with the technology as she demonstrated the correct intent to operate the controls, but not the necessary motor control. This was a different participant than the one who had not used the Grid 2 software before. We decided to end her user study so as not to have the student "sin-bin" the technology. She may be invited for further testing with a future revision of the technology.

Five of the six participants showed a preference for the sideways head motion to interact with HeadBanger. This was not the same five who had previous experience with the Grid software. The participant who had a clear preference for the nodding head motion chose this gesture due to the nature of his physical mobility characteristics, despite this making it more difficult to watch the monitor at the same time as interacting with the controls. On a scale of 1-4, with 1 being poor, 2 being ok, 3 being good and 4 being very good, the five participants who successfully interacted with the system all scored the general usability of HeadBanger as either a 3 or 4 (two graded the system as overall good, three as overall very good).

Once a participant became competent at interacting with the system, the time for task completion was controlled by the scan rate of the grid controller. This scan rate was the same for all of the participants.

Discussion and conclusions

The user studies and feedback from the students indicate that the webcam-based head tracking technology presented here is of practical use to enable the operation of switchable software through head motion. Five of the six participants became effective users of the technology. The sixth student may well have become proficient given a further training period.

The ability to move the targets was critical to the system being successful as each of the students who volunteered to participate had a unique range of comfortable head motion. For instance, one student could only move her head in a narrow range tilted to one side, whereas another found having targets arranged around the vertical to be preferable.

There is no "one size fits all" solution for any assistive technology. Each of the students at Beaumont have a different physical and cognitive ability. One of the challenges in designing any assistive technology is to make it useful for as wide a range of potential users as possible. HeadBanger was designed to be adaptable to cater for an individual's comfortable range of head motion by having movable targets and two axis of interaction. Suitable candidates for the technology need to be identified by an Occupational Therapist and the positions of the targets on the system tuned to the individual.

A variety of algorithms for gesture recognition have been implemented in prior work, e.g. Kalman filters by Raya *et al.*(2012). Kratz and Rohs (2010) investigated 3D gesture recognition using a Nintendo Wii controller. However none of these methods are 100 per cent accurate and false triggers occur during routine usage. Using virtual targets is a simpler approach that removes the inaccuracies associated with pattern matching algorithms. The position of the targets and the axis of head motion defines the gesture that the participant makes to trigger a virtual control.

The communications software used by the target user group runs on tablet computers using the Windows OS. Installing a webcam-based head tracking system like HeadBanger onto the wheelchair mounted communications system used by the students at Beaumont College is feasible as the webcam runs from the power provided by a USB socket and all of the software

required by the system operates under the same OS which is already installed on their communication systems (Microsoft Windows). Tablets with a forwards facing webcam may be able to employ the built in webcam, further reducing the cost of implementation.

Development

HeadBanger was written as a proof of concept for using webcam-based head tracking as assistive technology for communication as a part of a research project, not as a commercial product. The system uses open source software as a component for head tracking. The legality of using this in a commercial system would need to be explored. The interface relies on the PyQt graphics library which would need to be licenced for commercial use, or recoded using a licence-free library.

It would be practical to save the target position and preferred axis of interaction for different students by using a simple database. A static HeadBanger system could use facial recognition to automatically load the relevant settings for each student. The system could be programmed to operate when a participant comes within a certain target range from the camera, giving a clear intention to use based on position. One potential group implementation discussed at Beaumont is for operating a television in a social setting. The intention to operate the television controller would be implemented by having HeadBanger focused on a designated area in front of the television, or set to operate within a prescribed range from the television.

Occupational Therapists who work with the target user group have expressed a desire for plug-and-play communication devices. The software for this project could be ported to a single board computer such as the BeagleBone Black manufactured by Texas Instruments. This could be connected with a webcam and built into a standalone head tracking device, configured for each user by a Technologist using a touch screen interface connected to the device. User profiles could be stored and recalled so that same device would be of use for different individuals.

Presenting a case for developing a commercial system would rely on demonstrating sufficient demand from the target user group to recoup the necessary investment. Using head gesture to control software could well be of use outside of assistive technology which would be useful for generating the necessary interest for developing the idea further. Software for a commercial system would need to be professionally written and tested.

The code written for this research project has been placed in a git repository and made available under an open source licence at www.assembla.com/spaces/headbanger/git/source

References

Chai, X., Li, G., Xilin, C., Ming, Z., Guobin, W. and Hanjing, L. (2013), "VisualComm: a tool to support communication between deaf and hearing persons with the Kinect", *ASSETS '13: Proceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility, Bellevue, WA, 21–23 October*.

Kim, H. and Ryu, D. (2006), "Computer control by tracking head movements for the disabled", in Klaus, M., Klaus, J., Zagler, W.L. and Karshmer, A.I. (Eds), *Computers Helping People with Special Needs*, Springer, Berlin and Heidelberg, pp. 709-15.

Kratz, R. and Rohs, M. (2010), "A $3 gesture recognizer – simple gesture recognition for devices equipped with 3D acceleration sensors", IUI, Hong Kong.

Lewis, J. (1993), "IBM computer usability questionnaires: psychometric evaluation and instructions for use", IBM Corporation, Boca Raton, FL.

Microsoft Research Asia (2013), available at: http://research.microsoft.com/en-us/collaboration/stories/kinect-sign-language-translator.aspx (accessed January 2015).

Oppenheim, M. (2014), "HeadBanger: tracking head position as a controller", *Communication Matters*, Vol. 28 No. 2, pp. 34-6.

Pajkanovic, A. and Dokic, B. (2013), "Wheelchair control by head motion", *Serbian Journal of Electrical Engineering*, Vol. 10 No. 1, pp. 135-51.

Raya, R., Rocon, E., Gallego, J.A., Ceres, R. and Pons, J.L. (2012), "A Robust Kalman algorithm to facilitate human-computer interaction for people with cerebral palsy, using a new interface based on inertial sensors", *Sensors*, Vol. 12 No. 3, pp. 3049-67.

Rechy-Ramirez, E.J., Hu, H. and McDonald-Maier, K. (2012), "Head movements based control of an intelligent wheelchair in an indoor environment", *Robotics and Biomimetics (ROBIO), 2012 IEEE International Conference IEEE*, pp. 1464-69.

Web references

www.orin.com/access/headmouse/ (accessed January 2014).

www.microsoft.com/en-us/kinectforwindows/ (accessed June 2014).

www.naturalpoint.com/trackir/products/trackir5/ (accessed March 2014).

www.free-track.net/english/freetrack/qu-est-ce-que-c-est.php (accessed March 2014).

http://therapy-box.co.uk/smarthub.aspx (accessed March 2014).

www.sensorysoftware.com/thegrid2.html (accessed March 2014).

www.freetronics.com.au/products/leostick (accessed January 2015).

http://facetracknoir.sourceforge.net/home/default.htm (accessed January 2015).

http://producttesting.org.uk/?page_id=35 (accessed July 2015).

www.ti.com/tool/beaglebk (accessed August 2015).

www.youtube.com/watch?v=73wR1FkYijs

Corresponding author

Matthew Oppenheim can be contacted at: matt.oppenheim@gmail.com

This article has been cited by:

1. Anja Jackowski, Marion Gebhard, Axel GraserA novel head gesture based interface for hands-free control of a robot 1-6. [CrossRef]