A novel head gesture based interface for hands-free control of a robot

Anja Jackowski and Marion Gebhard
Sensors and Actuators in Medical Engineering (SAM)
University of Applied Sciences, Westphalia
45877 Gelsenkirchen, Germany
Email: {anja.jackowski, marion.gebhard}@w-hs.de

Axel Gräser
Institute of Automation (IAT)
University of Bremen
28359 Bremen, Germany
Email: ag@iat.uni-bremen.de

***Abstract*—Within this work a novel head gesture based interface for hands-free control of a collaborative robot is developed and evaluated. Based on previous work, robot control is divided into several control groups for intuitive head motion based control. The switching commands to select and switch between these robot control groups are given by four gestures performed with head movements. The head movements were measured using a nine-axis inertial measurement unit.**

The control of a robot with the novel interface was evaluated objectively as well as subjectively. The objective evaluation contains the measurement of time needed for the given control task and the number of trials for switching between the different groups of robot control. The subjective evaluation was carried out with a questionnaire. All subjects were able to perform the given task of controlling a robot arm with the head gesture based interface.

I. INTRODUCTION

Usually, interfaces to operate with robots are controlled with the hands. For disabled people who cannot use their hands or for people whose hands are already occupied with other activities, such a manual control mode is impossible to use.

This work presents a novel hands-free interface to control a robot arm with head motions. Disabled people, e.g. with tetraplegia[1], may restore their autonomy with a semi-autonomous care giving robot and use the hands-free interface for robot control. In general, humans have three easy to be used degrees of freedom (DOFs) for head motions, namely roll φ, pitch ϑ and yaw ψ. To control a robot with head motion, these DOF have to be translated to the movement of a robot arm.

A. State of the art in hands-free interfaces

Lobo-Prat et al. [1] give a review of hands-free interfaces for the control of assistive devices. Hands-free interfaces can be controlled for example with speech, brain activity, movements of the eyes or the tongue or, as in this work, with head movements. Speech controlled interfaces are sensitive to ambient noises. In non-invasive brain computer interfaces (BCIs), the brain activity is measured with EEG (electroencephalography) electrodes placed on the head. The low information transfer rate restricts the use of EEG-based BCI for hands-free control of a robot.

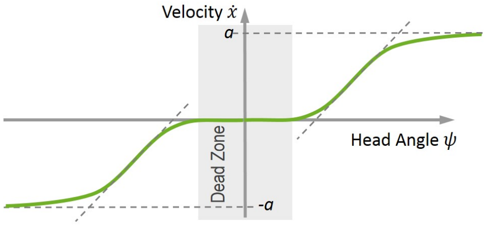

Fig. 1. The head angle yaw ψ is for example mapped to the velocity $\dot{x}$. Due to the shape of the Gompertz function, there is a dead zone. Small head movements do not cause a movement of the robot.

Most of recent hands-free interfaces which are based on head movements are limited to the control of devices with two degrees of freedom, e.g. a mouse cursor or a wheelchair ([2], [3] and [4]). But in [5] a head motion based hands-free device for 3D control of a robot is presented. The head motions are measured with a nine-axis inertial measurement unit (IMU). The head motions are translated to an intermediary cursor control of the human machine interface (HMI). Using this interface developed by Rudigkeit et al., the user has to switch between two control modes: the control of a mouse cursor on a screen and the control of a robot arm or any other device. The user has to select the different robot control groups on a graphical user interface (GUI). This selection enables robot control mode and the user has to return to cursor control to switch to another control group. In this work, a novel interface was developed and evaluated in order to allow a control by switching directly between the different control groups.

B. Control Mode

In [6] two modes for head motion-based control have been compared. The mapping of the head angle to the velocity of the robot arm was evaluated as the preferred control mode. Therefore, this control mode is also used within this work. The angle of the head is mapped to the velocity of the robot. The neutral position of the head is defined at the beginning of the control. Increasing the head angle relative to this neutral position accelerates the robot. Decreasing the head angle decelerates the robot. The head angle is mapped to the velocity

---

[1] Tetraplegia is the partial or total paralysis of all limbs and torso.

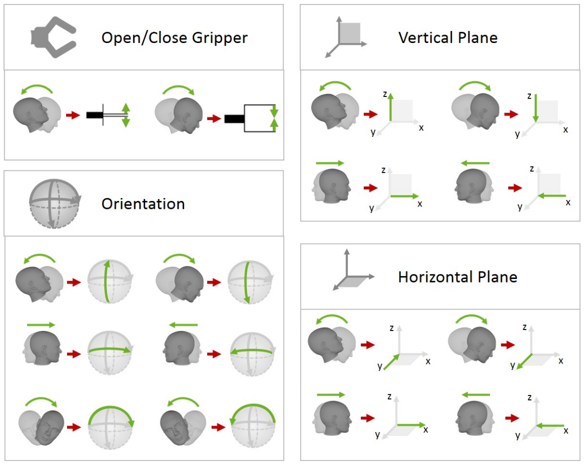

Fig. 2. The seven degrees of freedom of the robot are divided into four control groups for intuitive control with head movements.

of the robot with a Gompertz function (e.g. for the DOF yaw ψ and the velocity $\dot{x}$):

$$\dot{x}\left(\psi\right) = \begin{cases} a \cdot e^{b \cdot e^{c \cdot \psi}} & \text{if } \psi \geq 0 \\ -a \cdot e^{b \cdot e^{-c \cdot \psi}} & \text{else.} \end{cases} \tag{1}$$

The parameter a corresponds to the maximum velocity of the robot, b sets the displacement along the head angle (e.g. yaw ψ) and c sets the growth rate. The sigmoidal mapping contains a dead zone (Fig. 1), which ensures that small head movements around the neutral position are not translated into real robot motions.

C. Control Structure

The six DOFs of the robot arm and the gripper (7^{th} DOF) have to be controlled with three DOFs of the head motion (roll φ, pitch ϑ and yaw ψ). The seven DOFs of the robot are therefore divided into four control groups, which can be controlled with head movements intuitively. Four head gestures are introduced in order to switch easily between the four control groups. This is a small but important modification compared with the control structure described in [5]. Fig. 2 shows the mapping of the groups:

Open/Close Gripper	Pitch ϑ movements open or close the gripper.
Orientation	The orientation of the head (φ, ϑ, ψ) is mapped to the rotation of the gripper.
Vertical Plane	Yaw ψ and pitch ϑ movements change the gripper's position in the vertical plane in x and z directions.
Horizontal Plane	Yaw ψ and pitch ϑ movements change the gripper's position in the horizontal plane in x and y directions.

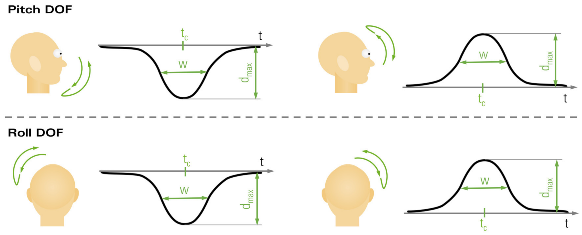

Fig. 3. The four head gestures "Nodding Down", "Nodding Up", "Bending Left" and "Bending Right" are recognized with an analytical algorithm. The head angle can be approximated by a Gaussian function while performing a gesture. [7]

To switch between the different control groups, the user has to perform head gestures.

D. Gesture Recognition

In [7] one head gesture is used to return from robot control mode to cursor control mode. Here, we use four different gestures to switch immediately between the four control groups. This approach simplifies the robot control for the user considerably. Fig. 3 shows the four different gestures ("Nodding Down", "Nodding Up", "Bending Left" and "Bending Right") that have to be recognized. These four head gestures can be classified with the analytical algorithm presented in [7]. The recognition algorithm which is based on pitch ϑ and roll φ motions of the head is triggered when a threshold of the activity, characterized by the linear acceleration, is exceeded. As long as the activity goes on, the head angle is captured. First of all, it is checked whether one DOF is dominating. Afterwards, a Gaussian function (2) is fitted to the data using the gradient method.

$$d = d_{max} \cdot e^{-\left(\frac{t-t_c}{w}\right)^2} \tag{2}$$

The regression coefficient R^2 and the calculated amplitude d_{max} and peak width w are compared to defined thresholds to classify the gestures.

II. METHODS

Within this work, a novel control interface is generated as described in section I.

The head angle is mapped to the movement of the robot arm using a Gompertz function (Sec. I-B) with the parameters $b = -30$ and $c = -10\ \text{deg}^{-1}$. The maximum velocity of the robot, described by parameter a, was adjusted with different scaling factors for each control group. The range of head motion used for robot control was empirically defined by the limits $\varphi_{max} = 20$ deg, $\vartheta_{max} = 15$ deg and $\psi_{max} = 25$ deg. The neutral position is calibrated at the beginning of every use of the interface. For this calibration, the user's gaze is directed towards the center of the robot and the user holds the head in a stable position for five seconds. The head angle is calculated relative to this neutral position.

To calibrate the head gestures, the user has to perform each gesture for three times. The parameters of the fitted

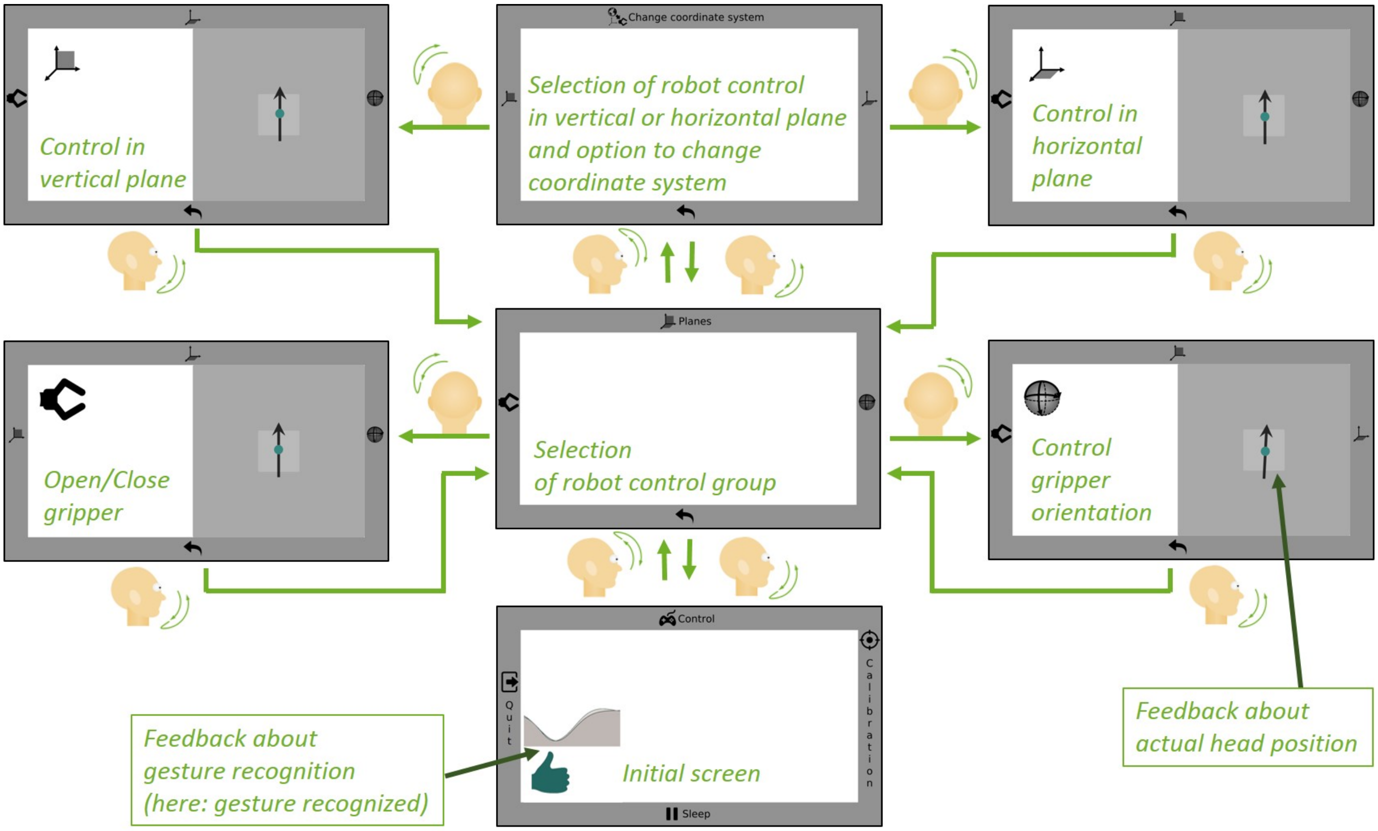

Fig. 4. Structure of the control interface presented with screenshots and additional information (green). The different control groups and features are selected with head gestures. For simplicity, the immediate switching between the four robot control groups is not shown explicitly.

Gaussian function are calculated for every gesture. If the parameters d_{max}, w and the R^2-value are in defined intervals, the parameters amplitude d_{max} and peak width w are taken into account for user-dependent gesture recognition. If the gesture does not meet the requirements, the subject has to repeat the gesture again. The arithmetic mean of the three repetitions is calculated for the amplitude d_{max} and the peak width w. The user-dependent thresholds are set to the computed mean ∓ 20 %. If the threshold of the amplitude would be computed higher than 20 degrees, it is set to $d_{max} = 20$ deg, if the threshold for the peak width would be below 0.28 seconds it is set to $w = 0.28$ s in order to guarantee a straight forward gesture performance.

A. Control Interface

At the beginning of every session, the neutral position is calibrated. To include the user-dependent thresholds, the gesture calibration is carried out.

After the calibration routines, the user has the choice to select one of the four options which are shown at the borders of the screen. The selection is induced by one out of the four head gestures: The option at the top is chosen with the gesture ”Nodding Up”. To select the option at the bottom, the subject performs the gesture ”Nodding Down”. For the left option, the gesture ”Bending Left” has to be performed, for the right one the gesture ”Bending Right”.

The subjects can switch between the different control groups and features as shown in Fig. 4. The selection of a control group enables robot control within this group (Fig. 2). With the head gestures, the subject can switch between the different control groups without droping out of the robot control mode. In robot control mode, one out of the four control groups is active only. The icon of this group is shown in the upper left corner. The other three control groups are available with head gestures and are shown at the borders of the screen as previously described. ”Nodding Down” stops robot control mode. In robot control mode, the subject receives feedback about the actual head position on the right. The head position is demonstrated by an arrow. The range of head motion that results in a change of robot velocity is shown as large square. The small square in the middle represents the dead zone in which head movements do not result in robot movements. If a gesture is performed, feedback about the performance is given in the lower left corner, as shown on the screen at the bottom of Fig. 4. Another option for the robot control mode is to change the coordinate system from world coordinates to gripper coordinates, which was not needed for the control task in this work.

B. Experimental Setup

The control interface and the feedback are shown on a 27-inch-screen behind and above the range of motion of the robot.

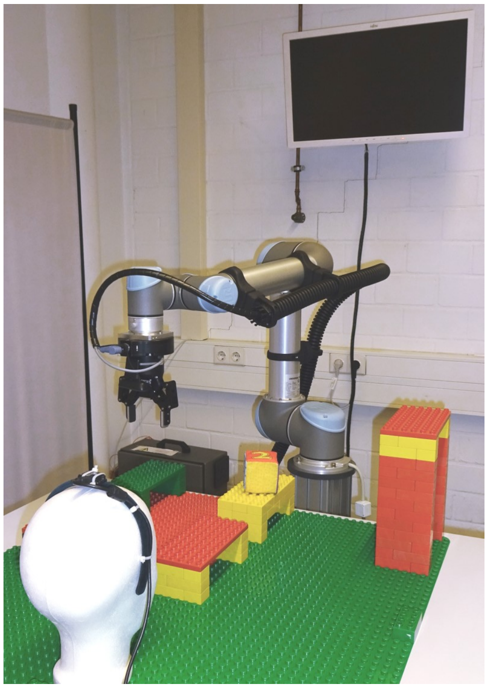

Fig. 5. Within the experiment the FSM-9 is attached to the user's head with a hairband. The control task is performed on a table in front of the user.

The controlled robot arm is an UR5 by Universal Robots A/S, Odense, Denmark [8], equipped with the adaptive gripper 2-Finger Robotiq 85 by Robotiq, Lévis, Canada [9].

The head movements are captured with the nine-axis inertial measurement unit FSM-9 by Hillcrest Laboratories, Rockville, USA [10]. Therefore, the sensor module is mounted on a hairband to attach it to the user's head. The sampling rate of the FSM-9 is set to 125 Hz. The head movements are mapped to robot movements. The update rate to refresh the robot position is 25 Hz.

Ten able-bodied subjects at the age from 22 to 53 participated in the experiment. The five males and five females with no known neck movement limitations had different previous experiences with head motion based control. Some of the subjects already controlled the robot arm with the interface described in [5], but none of them controlled the robot with the interface presented in this work.

The subjects moved and rotated a cube with the head controlled robot arm. The entire control task is performed on a table in front of them, on which different platforms are placed (Fig. 5).

C. Procedure

To evaluate the novel control interface, the subjects have to perform a defined control task.

The control structure and the control interface are presented to every subject by an introduction video. After this introduction and an information sheet, the subjects tested the head gesture based control for five minutes in a free explorative phase. Subsequent to this adaptation phase, the subjects performed the control task. They had to grip a cube from a platform, pass it to another platform, turn it into a defined orientation and put it down. The control task was designed in a way, that all control groups have to be used. The neutral position and the head gestures were calibrated in the test phase and before the control task. At the end of the control task, the subjects had to quit the application.

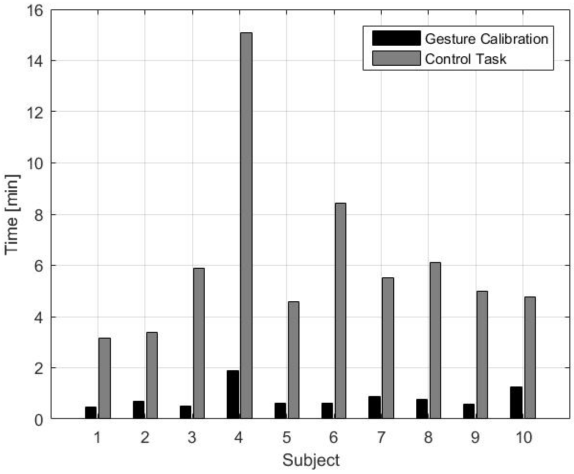

Fig. 6. Time needed by each subject to complete the gesture calibration and time required to solve the control task after completing the calibration routines.

The time needed to calibrate the head gestures and the time needed to solve the control task were measured. Moreover the number of trials for every gesture was counted. After solving the control task, all subjects completed a questionnaire with the following six statements:

1) The mental effort required for control was low.
2) The physical effort required for control was low.
3) The control was intuitive.
4) Neck fatigue was low.
5) Overall, the control was very easy.
6) The control task was very easy.

The subjects rated every statement with a value ranging from 1 (I do not agree at all) to 5 (I totally agree), according to the Likert scale. Moreover, all subjects were asked for suggestions for improvement.

III. RESULTS

All subjects were able to complete the given control task. The time, every subject required for gesture calibration and the time needed to solve the task is shown in Fig. 6. Subject 4 needed significantly more time for every part of the control, for gesture calibration as well as for the control task itself. Observation of this subject's performance has shown that this test person had a very high coupling between yaw ψ and roll φ movements of the head. Subject 4 was not able to turn the head in a single degree of freedom. Thus, robot control was very time consuming, especially the control of the orientation. Except from subject 4, the test persons

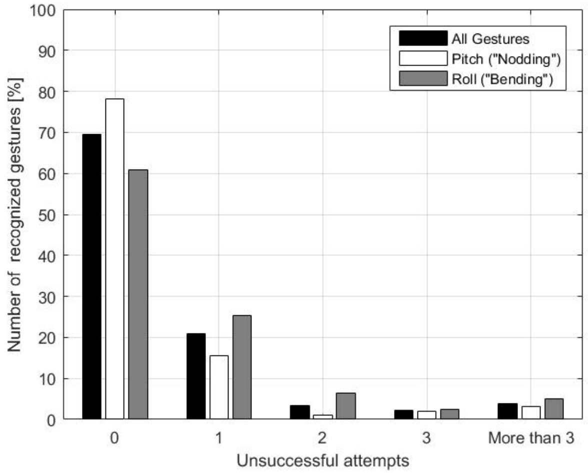

Fig. 7. Relative distribution of the recognized gestures across the number of unsuccessful attempts until correct gesture performance. The performance of all gestures is shown in the left black bars. The gestures are splitted according to the dominating DOF into gestures in pitch (”Nodding”, white bars) and roll (”Bending”, grey bars).

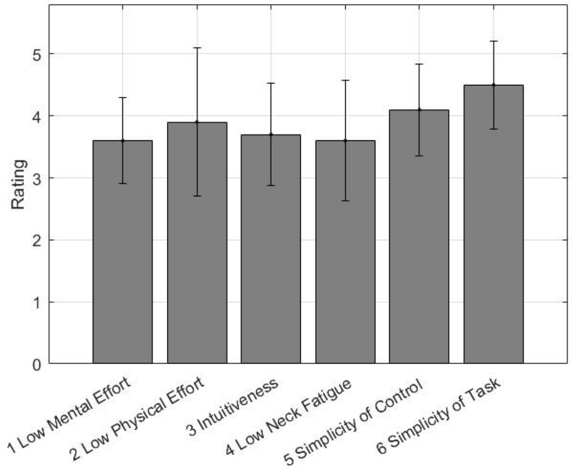

Fig. 8. Mean values and standard deviation of the answers from the questionnaire.

completed the control task in $5.20 \ \pm \ 0.70$ min and the gesture calibration in $42.22 \ \pm \ 14.87$ s.

The relative distribution of head gestures across the number of trials to perform a gesture is shown in Fig. 7. The subjects completed the head gestures in 69.49 % at the first try. Splitting the gestures according to the dominating DOF into gestures in pitch ϑ (”Nodding”) and roll φ (”Bending”) highlights, that the subjects needed less attempts to perform gestures in pitch ϑ than in roll φ. 78.13 % of the gestures in pitch ϑ, and 60.76 % of the gestures in roll φ respectively, succeeded at the first try. Except from two subjects, the test persons performed all the gestures in pitch ϑ at the first or at least second try.

In all tests, no gesture is classified as a wrong gesture. The gestures are either performed and classified correctly or the performance did not meet at all the requirements for a gesture and no gesture was recognized.

The subjective evaluation is shown in Fig. 8. The answers from the questionnaire point out that the control requires a low mental and physical effort ($3.6 \ \pm \ 0.70$ and $3.9 \ \pm \ 1.20$) and neck fatigue is low ($3.6 \ \pm \ 0.97$) according to most of the subject’s assessments. The intuitiveness of the control was rated with $3.7 \ \pm \ 0.82$. The overall control and especially the given control task are rated as very easy with $4.1 \ \pm \ 0.74$ and $4.5 \ \pm \ 0.70$.

IV. DISCUSSION

Although the gestures in pitch ϑ are more successful than the gestures in roll φ, some subjects also had problems with the performance of these gestures. Thus, the gesture recognition should be more adaptive. The currently used analytical gesture recognition algorithm prescribes the Gaussian shape of the gestures. This Gaussian shape does not necessarily correspond to the intuitive head movements of the users. In future work, the system will include individual arbitrary gestures to be defined by the user as switching commands. Individual gestures could be trained and recognized with machine learning algorithms. With the use of a more general gesture recognition, a larger number of gestures can be defined. This will lead to more functionality without the loss of immediately switching between the functions. Various parameters e.g. mental effort of this approach will be analysed in future work.

Compared to the interface presented in [5], the novel interface has some advantages and disadvantages. Due to immediately switching between the control groups, the control with the head gesture based interface is probably faster than the control with the GUI-aided two-staged interface. With the use of a graphical user interface, all options can be shown at a glance and can be selected with buttons. Thus more functionality can be reached without requiring a higher mental effort. The two interfaces will be compared with each other under defined conditions to evaluate their advantages and disadvantages and choose the most suitable for future work.

Moreover, the interface will be tested with disabled people.

In conclusion, the novel interface offers the potential of being a promising alternative to existing hands-free interfaces.

ACKNOWLEDGMENT

The authors would like to thank Nina Rudigkeit and Julius Heinke for their outstanding support and all the volunteers who took part in the experiment.

This work was funded by the Federal Ministry of Education and Research of Germany.

REFERENCES

[1] J. Lobo-Prat, P. N. Kooren, A. H. Stienen, J. L. Herder, B. F. Koopman, and P. H. Veltink, “Non-invasive control interfaces for intention detection in active movement-assistive devices,” *Journal of neuroengineering and rehabilitation*, vol. 11, no. 1, p. 168, 2014.

[2] T. Elder, M. Martinez, and D. Sylvester, “Alternate computer input device for individuals with quadriplegia,” 2014.

[3] R. Raya, E. Rocon, J. A. Gallego, R. Ceres, and J. L. Pons, “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. 3049–3067, 2012.

[4] E. Perez, N. López, E. Orosco, C. Soria, V. Mut, and T. Freire-Bastos, “Robust human machine interface based on head movements applied to assistive robotics,” *The Scientific World Journal*, vol. 2013, 2013.

[5] N. Rudigkeit, M. Gebhard, and A. Gräser, “Towards a user-friendly AHRS-based human-machine interface for a semi-autonomous robot,” in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems, Workshop on Assistive Robotics for Individuals with Disabilities: HRI Issues and Beyond*. Chicago, Illinois, USA: IEEE, 2014.

[6] N. Rudigkeit, M. Gebhard, and A. Gräser, “Evaluation of control modes for head motion-based control with motion sensors,” in *Medical Measurements and Applications (MeMeA), 2015 IEEE International Symposium on*. IEEE, 2015, pp. 135–140.

[7] N. Rudigkeit, M. Gebhard, and A. Gräser, “An analytical approach for head gesture recognition with motion sensors,” in *International Conference on Sensing Technology (ICST 2015)*. Auckland, New Zealand: IEEE, 2015.

[8] Universal Robots A/S. UR5 robot. [Online]. Available: http://www.universal-robots.com/products/ur5-robot/

[9] Robotiq. Adaptive Robot Gripper 2-Finger 85. [Online]. Available: http://robotiq.com/products/industrial-robot-gripper/

[10] Hillcrest Laboratories Inc. FSM-9 Modules. [Online]. Available: http://hillcrestlabs.com/products/sensor-modules/fsm-9/