A HANDS-FREE HUMAN-MACHINE INTERFACE FOR ELECTRIC WHEELCHAIR CONTROL BASED ON FACE AND TORSO MOVEMENTS

ISSN 1590-8844
International Journal of Mechanics and Control, Vol. 26, No. 01, 2025
https://doi.org/10.69076/jomac.2025.0005

A HANDS-FREE HUMAN-MACHINE INTERFACE FOR ELECTRIC WHEELCHAIR CONTROL BASED ON FACE AND TORSO MOVEMENTS

Lorenzo Baglieri* Daisuke Matsuura** Tsune Kobayashi*** Giuseppe Quaglia*

* Politecnico di Torino, DIMEAS – Department of Mechanical and Aerospace Engineering[1]

** Institute of Science Tokyo - Department of Mechanical Engineering[2]

*** Institute of Science Tokyo[2] and JTEKT Corporation[3]

ABSTRACT

As the global population ages, mobility impairments are becoming a significant challenge, increasing the need for innovative assistive technologies. Traditional electric wheelchairs require manual input, which may not be feasible for users with severe disabilities. This paper presents a novel hands-free Human-Machine Interface (HMI) designed to enhance electric wheelchair control using face and torso movements. The proposed system integrates a vision-based motion capture method with a shared control strategy that considers both the rider's. A stereovision setup detects and tracks head and torso movements, which are then processed to generate control commands for the wheelchair. The experimental results demonstrate the feasibility of this system, highlighting its potential to improve wheelchair maneuverability in confined indoor environments while providing an intuitive and non-invasive control method. Challenges related to lighting conditions and system robustness are also discussed, with future work aimed at refining control accuracy and adaptability for different users.

Keywords: Human-Machine Interface (HMI), hands-free control, vision-based tracking, assistive technology, mobility assistance.

1 INTRODUCTION

The establishment of a highly aged society causes various and complex problems concerning welfare such as the rising number of people with walking diseases and the necessity of working-aged people for caregiving. In the European zone, the number of people with mobility limitations is over 120 million people by 2020 [1]. According to studies, up to 10% of people have difficulties or are unable e to use a traditional power wheelchair, another 10% have problems steering without help, and 40% struggle with steering [2].

Contact author: Lorenzo Baglieri

[1]Corso Duca degli Abruzzi 24, Torino, 10129, Italy.
[2]2-12-1 Ookayama, Meguro, Tokyo, 152-8550, Japan.
[3]1-1 Asahi, Kariya, Aichi, 448-8652, Japan.
E-mail: lorenzo.baglieri@polito.it, matsuura.d.aa@m.titech.ac.jp, kobayashi.t.cs@m.titech.ac.jp, giuseppe.quaglia@polito.it

To tackle this problem, some researchers have worked on a wheelchair with an alternative propulsion system [3] and mobility mechanism and haptic interface for electric wheelchairs to provide fine operability for riders and caregivers [4]. Although these researchers have contributed to improving the usability of manual- and electric wheelchairs, there was a common problem that the operation of those wheelchairs required manual input by the riders or caregivers. This research is thus going to develop an innovative human-machine interface (HMI) that allows the riders or caregivers to operate the electric wheelchair without occupying their hands, which results in the improvement of the users' daily activities. The developed HMI system will be implemented in an electric and omnidirectional wheelchair to perform better in narrow indoor environments [5] to demonstrate the feasibility of the proposed concept. Regarding the development of innovative Human-Machine Interfaces (HMI) to drive wheelchairs in a novelty way closer to the driver's need, especially the achievement of the free-hand interface, there are various ways such as EEG [6], eye gaze tracking [7], tactile sensation [8], IMU and visual information [9], and use of tongue [10]. The goal of the

above and other researchers is the development of a wheelchair that fits better with human life standards. For example, increasing the quality of life of users by allowing them to manage their mobility autonomously. From this point of view, the implementation of an HMI adapted to the user's disability while making assistance for the user in navigation possible is very important. The HMI system proposed by the authors estimates the intentions of the wheelchair users, the rider and the caregiver. The main objective is to obtain shared control of the wheelchair's movements between the two users' intentions. A set of different rules and various driving solutions divides the action priority of the two users, while one person can drive a wheelchair the other can make only some adjustments and vice-versa. To make the natural and non-invasive HMI system, a vision system to measure the movements of the face and torso of the rider and caregiver is desirable. Measured movements are used to generate commands for the wheelchair. Thanks to this it is possible to drive a wheelchair. Either of the rider's-, caregiver's intentions, or a combination of them is taken depending on the situation by a user's intention estimation and selection algorithm. This is the basis of our HMI that the authors are willing to achieve in the future. This aim also is addressed in the UN Sustainable Development Goals n. 3, Good Health and Well-Being.

1.1 COMPOSITION AND WORKING PRINCIPLE OF THE WHEELCHAIR

The proposed system is an electric wheelchair having omnidirectional capability and hands-free interface software aiming to enable a "side-by-side" style movement in which a wheelchair user and caregiver move around by side naturally, again, without using their hands for maneuvering. Figure 1 shows the subsystems that compose the wheelchair. The inputs come from different cameras (Rider camR, Rider camL and Caregiver Camera) that acquire the face and torso movements of the two persons, this movement are represented by the rider's head (rh) and rider torso (rt) reference frame. The acquisition phase and the elaboration of this information happens in a Host computer, that is a PC with an adequate GPU and an Ubuntu with ROS operating system. The elaboration phase is split into three different steps: Face and torso detection, Intention estimator and Motion Planner. The first step is identifying, from the acquired images, the movements of the face and torso of the rider and/or the caregiver. Once the movements are detected, the "Intention Estimator" chooses the overall direction of the wheelchair by overlapping the two inputs from the rider and caregiver. Finally, target motion is generated in the Motion Planner. In the Microcontroller the target motion is transformed into velocity commands for the traction motors and angular commands for the steering motors, to achieve the target motion. The following sections will focus on the Face and Torso Acquisition and the Intention Estimator, more specifically the case treated is the only one related to the rider.

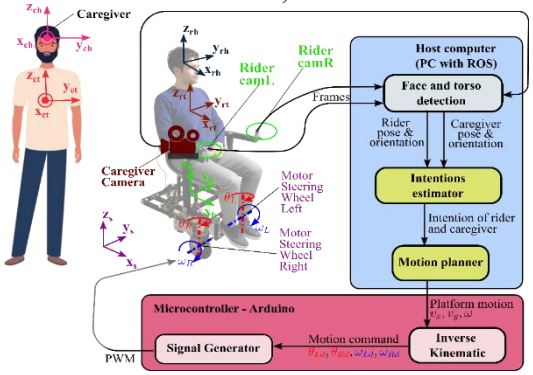

Figure 1 Description of all the hands-free side-by-side wheelchair subsystems.

2 FACE AND TORSO ACQUISITION

2.1 PROCESS VARIABLE

In this paragraph, it will be explained that all the variables used to explain the process of the acquisition of the face and torso movements (for brevity this process will be explained in detail only for the face variable, but the same concept can be applied for the shoulder or anything similar). Table I explains all the References Systems (R.F.s), the positions, the Rotation Matrix (R.M) and the Transformation Matrix (T.M.) that are taken into consideration for the calibration and for measuring the face pose. With the Reference System, it is intended that three axes are orthogonal respecting the right-hand rules for the crossing product among each other's. Regarding the position it is intended that a point in 3D-space in a certain R.F. is described with a 3x1 vector.

Table I - parameter for camera calibration process and of the acquisition of the rider landmarks for the calibration process

General variables	
$camL$	Reference System of the left camera
$camR$	Reference System of the right camera
T_{camL}^{camR}	Transformation matrix from $camR$ to $camL$
Calibration process variables	
M	Reference system of the calibration markers
p_{camL}^{M}	Position of the calibration marker from camL
p_{camR}^{M}	Position of the calibration marker from camR
Stereovision process variables	
$p_{camL}^{l,n}$	Position of a nth landmark from camL
$p_{camR}^{l,n}$	Position of a nth landmark from camR
A_{camL}	Projection matrix of rh in camL RF
A_{camR}	Projection matrix of rh in camR RF
$x_{camL}^{l,n}$	Pixel coordinates of a nth landmark in camL
$x_{camR}^{l,n}$	Pixel coordinates of a nth landmark in camR
T_{camL}^{rh}	Transformation matrix from rh to $camL$

With the Rotation Matrix, it is intended that the rotation around three coordinates to orient the origin R.F. like the measured R.F. and it is described with a 3x3 matrix. The Transformation Matrix is intended as a 4x4 matrix which includes a rotation matrix and a position matrix,

$$T = \begin{bmatrix} R & \boldsymbol{p} \\ \mathbf{0} & 1 \end{bmatrix} \quad (1)$$

Where R is the 3x3 R.M., $\boldsymbol{p}$ is the 3x1 position vector, $\mathbf{0}$ is a 1x3 zero vector and 1 is the scale factor

2.2 CALIBRATION PROCESS

Figure 2 shows how the calibration setup is made. The process includes two cameras (camL and camR) and a calibration target marker (M). The target is composed of a chessboard structure with 17 AruCo markers to make a unique pattern to be recognized and also an easy pattern to compare. In fact, the chessboard structure, with the real size of a square is uploaded in a software program to execute the calibration.

During several calibration trials, at least 20, coordinates of the target markers, ($\boldsymbol{p}^{\boldsymbol{M}}_{\boldsymbol{camL}}$ and $\boldsymbol{p}^{\boldsymbol{M}}_{\boldsymbol{camR}}$), are taken from the left- and right cameras. With the acquisition of the markers from the point of view of the cameras, it is possible to compare the single image to the expected, already uploaded in the software as mentioned above, to evaluate the intrinsic parameters of the camera, like focal length, f_x, f_y; principal point, c_x , c_y and skew factor, s . This parameter will constitute the intrinsic matrix of each camera

$$K = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

In addition, fusing the two points of view from the two cameras on a common object and knowing the size of that object make it possible to perform triangulation for estimating the calibration matrix, namely, the transformation matrix from camR to camL, T^{camR}_{camL}. This matrix is calculated as the result of the minimum error optimization across the triangulation in all the measurements that are kept by the two cameras. The transformation matrix is useful to calculate the extrinsic matrix of the camera.

The extrinsic matrix is a 3x4 matrix composed of the rotation matrix, R, between the world reference frame and the camera reference frame. The last column has the position of the camera reference frame from the origin reference frame, $\boldsymbol{p}$:

$$E = [R \quad \boldsymbol{p}] \quad (3)$$

In this process, for simplicity, the world origin is put on the camL reference frame, so for the left camera the R matrix is the identity matrix and the $\boldsymbol{p}$ vector is a zero vector. For the right camera, the R and the $\boldsymbol{p}$ vector come out from the transformation matrix, from camR to camL, calculated during the calibration. Finally, it becomes possible to calculate the two projection matrices, A_{camR} and A_{camL}, multiplying the intrinsic matrices with the extrinsic one

$$\begin{aligned} A_{camR} &= K_{camR}E_{camR}, \\ A_{camL} &= K_{camL}E_{camL} \end{aligned} \quad (4)$$

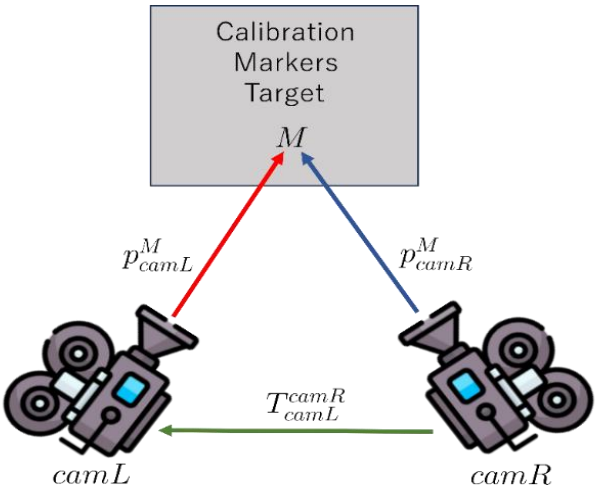

Figure 2 Calibration process.

2.3 FACE POSE AND ROTATION ESTIMATION

Figure 3 shows the configuration of cameras and riders when the system is used to collect the rider's movements. The five purple crosses on the face of the rider represent the landmarks that should be acquired: two eyes, a nose and two mouth corners. These five landmarks are useful to calculate the pose of the face with respect to a camera reference frame. This paragraph will only explain how to solve the triangulation problem and calculate the position of the landmarks from one of the two cameras.

The starting point is the pixel coordinates of the landmarks in each camera's viewpoint, $x^{l,n}_{camL}$ and $x^{l,n}_{camR}$. As shown in the Figure 3, the pixel coordinates of the landmarks are the projection of the 3D position of those landmarks on the camera's screen coordinate. This projection is calculated by multiplying the projection matrix by to position vector as

$$\begin{aligned} x^{l,n}_{canL} &= A_{camL}p^{l,n}_{camL}, \\ x^{l,n}_{canR} &= A_{camR}p^{l,n}_{camR} \end{aligned} \quad (5)$$

In eq. (6) there are two unknown vectors: $p^{l,n}_{camR}$ and $p^{l,n}_{camL}$. Using the transformation matrix between the two cameras it

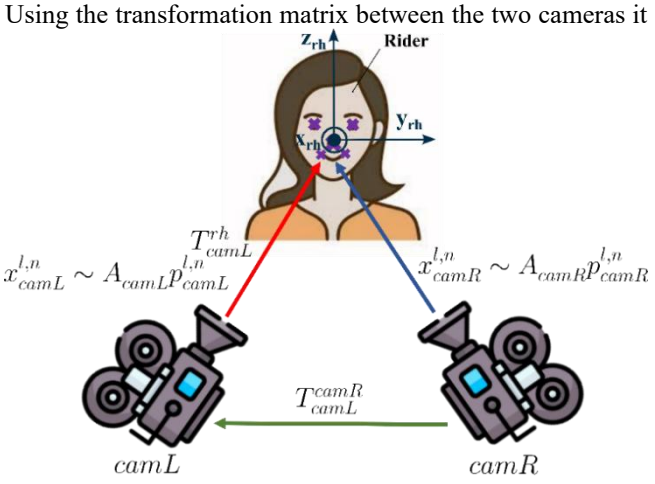

Figure 3 Stereovision working principle.

is possible to reduce the unknown vectors to only one that makes the system solvable

$$\begin{bmatrix} A_{camL} \\ A_{camR}T_{camR}^{camL} \end{bmatrix} p_{camL}^{l,n} = \begin{bmatrix} x_{canL}^{l,n} \\ x_{canR}^{l,n} \end{bmatrix} \quad (6)$$

This system contains a 4x3 matrix, which implies that the position vector is a 4-element vector and the pixel coordinate is a 3-element vector. The additional number in the pixel coordinate is the perspective division, which can be considered a scaling factor according to how close or far the object is from the camera. In the position coordinate vector, the 4th element is also a sort of scale factor but is always 1. From eq. (7) the position of the landmarks from the camL reference frame can be finally calculated. The position of the landmarks is used in the next section to calculate the position and orientation of the rider's head to estimate the one's intention.

3 INTENTION ESTIMATOR

3.1 FACE REFERENCE FRAME

This paragraph explains the use of the landmark position to calculate the reference frame of the rider's head. Figure 3 shows the landmark and how the reference frame should be placed accordingly. The first axis identified is the z-axis, it is placed from the nose marker and passing in the middle of the eyes' markers. The second marker identified is the y-axis, it starts from the nose landmark and the direction is between the left eye and the left corner of the mouth. The y-axis is depurated from any components on the z-axis to be sure that is perpendicular to the z-axis

$$\overrightarrow{y_{rh}} = \overrightarrow{y_{rh}} - (\overrightarrow{y_{rh}} \cdot \overrightarrow{z_{rh}})\overrightarrow{z_{rh}} \quad (7)$$

Finally, the x-axis is the axis that comes out from the nose landmark and it is calculated as the cross product of the y-axis and z-axis.

3.2 TRANSFORMING THE HEAD MOVEMENT IN RIDER MOTION INTENTION

The rider reference frame is the starting point to calculate the rider's motion intention. It is used to calculate the transformation matrix from rider's head (rh) to camL, T_{camL}^{rh}. This matrix represents the position and the orientation from the camL to the head reference but the most useful information is the movement of the rider's head from a customizable rest position. To acquire the movement of the rider head around a customizable rest position, it is important to choose a transformation matrix of a prompt rest position: $T_{camL}^{r,0}$. This transformation matrix will be settled as the origin of the rider's head translations and rotations. The rest position of the head, the one with the head aligned with the spinal column, is chosen as a rest position. Figure 4 shows how the rider movements are taken as input and then transformed into velocity commands. The rider's head's reference frame (*rh*) starts from the position *rh,0* which coincides with the reference frame with the same name. Here the rider moves at the *rh,a* position at which the polar coordinates can be calculated: distance, d_r, from the centre of the reference system and the angle, ϕ_r, of the direction w.r.t the $x_{rh,0}$ axes. The equations used to perform this calculation are below,

$$d_r = \sqrt{x_{rh,a}{}^2 + y_{rh,a}{}^2},$$
$$\Phi_r = \arcsin\left(y_{rh,a}/d_r\right) \text{ if } d_r \neq 0, \quad (8)$$
$$\Phi_r = 0 \text{ if } d_r = 0$$

With this formulation, a maximum and a minimum threshold can be set. All the velocities will have a constant behaviour regardless of the direction. Using a 5th-order polynomial function [11], explained in detail in the article written by *Taglaivini et al.*, it is possible to transform the d_r and $\theta_{z,rh}$ into linear velocity v, and angular velocity ω_z. The peculiarity of introducing the 5th-order polynomial function is to properly connect the deadband to the linear region of the input-output relationship to guarantee the achievement of comfortable behaviour of the wheelchair.

Table II summarizes all the variables that were used to transform the input signal into a command velocity signal ω_z. This can simply be done by using the direction Φ_{rt} as

$$v_x = v\cos(\Phi_r), v_y = v\sin(\Phi_r) \quad (9)$$

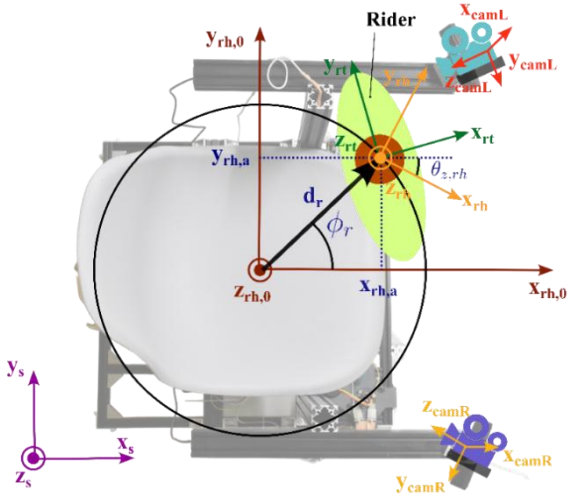

Figure 4 A rider moving from the rest position to generate a reference signal to move the wheelchair.

In the above, command velocities, v and ω_z, have been acquired in the polar coordinate. Then, they should be transformed into the reference velocities in the cartesian frame: two linear components, v_x and v_y. The deadband described in the Table are referred to the origin of the rider in the rest position (rh,0) Here, the three command velocities are obtained. Before the velocities are sent to the inverse kinematic routine they are elaborated by the Motion planner. It takes care to produce a unique velocity twist according to the rider and the caregiver's velocity intentions. In this case, the caregiver's intentions are not taken into consideration, so the only purpose of the motion planner is to add an exponential filter to smooth the velocity signal. The exponential filter used in this case considers the actual

velocity, the previous velocity and a smoothing factor. For a general variable it is possible to express it as:

$$s_t = \alpha x_t + (1-\alpha)x_{t-1}, \text{with } t > 0$$
$$s_t = x_t, \text{with } t = 0 \quad (10)$$

Where s_t is the smoothed variable, x_t is the actual variable, x_{t-1} is the previous variable and α the smoothing value (in this case 0.6)

Table II - Values and parameters that are involved in the velocity input generation

Variable	Value	Description
d_r	Obtained from measured data	Distance of the rider reference frame from the wheelchair reference frame in polar coordinates. Input of the polynomial function for linear velocity
Φ_r	Obtained from measured data	Direction of the rider reference frame from the wheelchair reference frame in polar coordinates.
$\theta_{z,rh}$	Obtained from measured data	Rotation of the head of the rider. Input of the polynomial function for angular velocity
v_{max}	0.3 m/s	Saturation for linear velocity.
$\omega_{z,max}$	0.5 rad/s	Saturation for angular velocity
d_{min}	4 cm	Dead band for the head motion to obtain a linear translation
θ_{min}	20°	Dead band for the head motion to obtain a linear translation
d_{change}	6 cm	Value for the head motion to change from polynomial to linear behavior
θ_{change}	40°	Value for the head rotation to change from polynomial to linear behavior
v_{change}	0.1 m/s	Linear velocity value to change from polynomial to linear behavior.
$\omega_{z,change}$	0.1 rad/s	Angular velocity value to change from polynomial to linear behavior.
d_{max}	10 cm	Maximum linear movements before entering the saturation zone
θ_{max}	90°	Maximum angular movements before entering the saturation zone

4 EXPERIMENTAL RESULTS

Figure 5 shows the experimental environment which is composed of an empty room with some cones that delimitates the area in which the experiment has been conducted. It is also possible to detect the monitor in which it was possible to overview the ongoing activity in the wheelchair's PC. The measurement of the wheelchair is conducted using a motion capture system. In the figure it is possible to see the tripod with two cameras and there were another two cameras symmetrically arranged on the other side of the room. The experiment aims to test the capabilities of going straight and rotating the system. During the experiment the rider was sitting on the wheelchair and was enabled the hands-free driving in order to follow a path in the perimetral lines of the cones. The experiment task was to drive around 90°-corners of the cone path. So there is an approaching phase to the cone path, until 5 s, then an approaching to the 90° corner at 13s, then the exit from the corner at 19s.

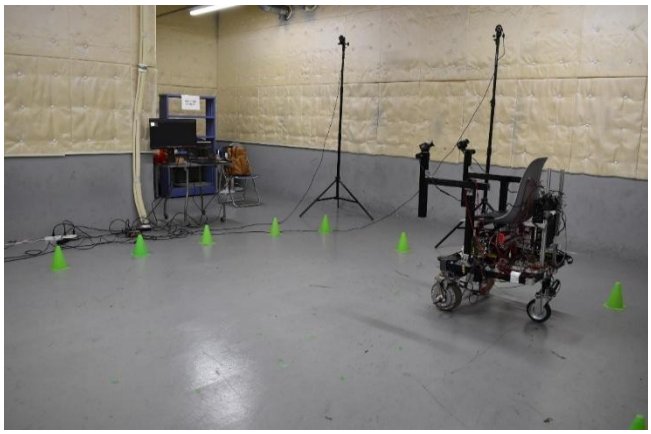

Figure 5 Experimental setup to test the hands-free driving wheelchair.

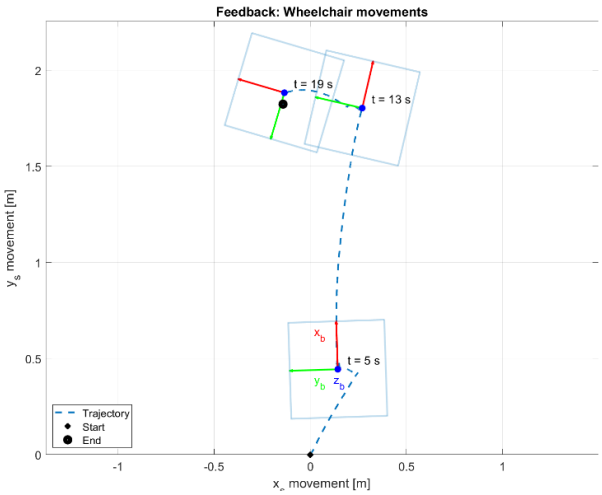

Figure 6 Trajectory and position followed by the wheelchair during the experiment.

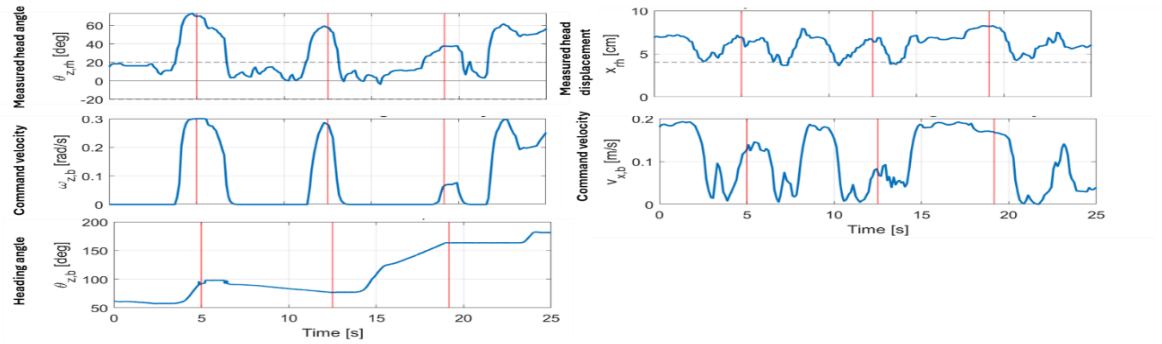

Figure 7 Head movements, velocity reference and rotation angle feedback during the experiment.

According to the user's experience, the driving method was enough satisfting although further calibration should be performed to increase the drivability of the system. Figure 6 and Figure 7 represent the results of the experiment. Figure 6 represents the head inputs, the generated velocity and the rotation angle feedback It is very interesting to notice that the control can be done in a very accurate way. The overall behaviour of the wheelchair was satisfactory; all the velocities are published according to the head movements as can be seen from the figure. The only problem is when the face is between a camera and a light, as Figure 7 represents the trajectory followed by the wheelchair during the experiment, there is a first phase, the first 5s, in which the wheelchair was approaching the perimetral line. Then the wheelchair turns to become parallel to the cone path and goes straight until it meets the corner at 13s, at this point start turning and complete another small straight trajectory. In the plot showing the wheelchair moving from the top point of view, it is possible to appreciate that the rotation and the linear velocity are not exactly as the input. The reasons for this misalignment are two. The difficulty is to put a marker in the exact center of rotation of the wheelchair so we are tracking a point that is close to the centre of rotation but not exactly the centre. Unfortunately in Figure 7 it is possible to see some accumulating delay between the input and the output, this is due to signal acquisition that happen with two different clock: motion capture clock and PC clock. In the real experience there wasn't such delay.

5 CONCLUSIONS

The papers explain how a hands-free interface is built from the head movements of the rider. The same approach can be used also for the rider's shoulder and for detecting some commands from the caregiver. Research activities are ongoing to define methods and strategies to share wheelchair control between two agents, i.e. the wheelchair rider and its caregiver and the results will provided in further papers. The results obtained in the experimental phase are interesting although some critical problems are still present. The most critical is the light position with respect to the rider and the camera because it can cause a complete loss of the rider landmarks detection and so an unexpected stop of the wheelchair motion. Future works will involve how to improve the overall drivability and, of course, how to solve the rider face illumination problems.

AKNOWLEDGEMENT

This publication is part of the project PNRR-NGEU which has received funding from the MUR – DM 351/2022 and has been supported by a joint research contract between JTEKT Corporation and Institute of Science Tokyo.

REFERENCES

[1] 'Persons with disabilities - Employment, Social Affairs & Inclusion - European Commission'. Accessed: Feb. 26, 2024. [Online]. Available: https://ec.europa.eu/social/main.jsp?catId=1137

[2] L. Fehr, W. E. Langbein, and S. B. Skaar, 'Adequacy of power wheelchair control i terfaces for persons with severe disabilities : A clinical survey', vol. Fehr, Linda, W. Edwin Langbein, and Steven B. Skaar. 'Adequacy of power wheelchair control interfaces for persons with severe disabilities: A clinical survey.' Journal of rehabilitation research and development, 2000.

[3] G. Quaglia, E. Bonisoli, and P. Cavallone, 'A proposal of alternative propulsion system for manual wheelchair', *International Journal of Mechanics and Control*, vol. 19, pp. 33–38, Jan. 2018.

[4] L. Tagliavini and G. Quaglia, 'On the Design of MoviWE.Q: An Omnidirectional Electric-Powered Wheelchair for Indoor Mobility', in *Advances in Mechanism and Machine Science*, M. Okada, Ed., Cham: Springer Nature Switzerland, 2023, pp. 311–321. doi: 10.1007/978-3-031-45770-8_31.

[5] L. Baglieri, D. Matsuura, T. Kobayashi, and G. Quaglia, 'Development of a Mobility Platform Aiming to Achieve User-Friendly Functionality', in *Proceedings of Jc-IFToMM International Symposium Vol. 7 (2024)*, Japanese Council of IFToMM, 2024, pp. 103–110. Accessed: Sep. 23, 2024. [Online]. Available: https://www.jstage.jst.go.jp/article/jciftomm/7/0/7_103/_article/-char/ja/

[6] I. Iturrate, J. Antelis, and J. Minguez, 'Synchronous EEG brain-actuated wheelchair with automated navigation', in *2009 IEEE International Conference on Robotics and Automation*, May 2009, pp. 2318–2325. doi: 10.1109/ROBOT.2009.5152580.

[7] Y. Matsumotot, T. Ino, and T. Ogsawara, 'Development of intelligent wheelchair system with face and gaze based interface', in *Proceedings 10th IEEE International Workshop on Robot and Human Interactive Communication. ROMAN 2001 (Cat. No.01TH8591)*, Sep. 2001, pp. 262–267. doi: 10.1109/ROMAN.2001.981912.

[8] D. Matsuura, K. Kiyomoto, and T. Kobayashi, 'Development of a simple human-cooperative transportation system based on tactile information', *ICPE2022*, vol. Proceedings of the 19th International Conference on Precision Engineering.

[9] L. Baglieri, D. Matsuura, T. Kobayashi, and G. Quaglia, 'Vision Systems and IMU Signals to Design a Hand-Free Driving HMI', in *Advances in Italian Mechanism Science*, vol. 163, G. Quaglia, G. Boschetti, and G. Carbone, Eds., in Mechanisms and Machine Science, vol. 163. , Cham: Springer Nature Switzerland, 2024, pp. 283–291. doi: 10.1007/978-3-031-64553-2_33.

[10] M. E. Lund, H. V. Christiensen, H. A. Caltenco, E. R. Lontis, B. Bentsen, and L. N. S. Andreasen Struijk, 'Inductive tongue control of powered wheelchairs', in *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, Aug. 2010, pp. 3361–3364. doi: 10.1109/IEMBS.2010.5627923.

[11] L. Tagliavini, L. Baglieri, G. Colucci, A. Botta, C. Visconte, and G. Quaglia, 'D.O.T. PAQUITOP, an Autonomous Mobile Manipulator for Hospital Assistance', *Electronics*, vol. 12, no. 2, Art. no. 2, Jan. 2023, doi: 10.3390/electronics12020268.