POLITECNICO DI TORINO
Repository ISTITUZIONALE

Vision Systems and IMU Signals to Design a Hand-Free Driving HMI

Original
Vision Systems and IMU Signals to Design a Hand-Free Driving HMI / Baglieri, L., Matsuura, D., Kobayashi, T., Quaglia, G.. - STAMPA. - 163:(2024), pp. 283-291. (IFToMM Italy 2024 Turin (ITA) September 11–13, 2024) [10.1007/978-3-031-64553-2_33].

Availability:
This version is available at: 11583/2992669 since: 2026-01-15T09:23:57Z

Publisher:
Springer

Published
DOI:10.1007/978-3-031-64553-2_33

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
Springer postprint/Author's Accepted Manuscript

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/978-3-031-64553-2_33

(Article begins on next page)

28 November 2026

Vision systems and IMU signals to design a hand-free driving HMI

Lorenzo Baglieri[1], Daisuke Matsuura[2], Tsune Kobayashi[3] and Giuseppe Quaglia[1]

[1] Politecnico di Torino, Italy
lorenzo.baglieri@polito.it, giuseppe.quaglia@polito.it
[2] Tokyo Institute of Technology, Japan
matsuura.d.aa@m.titech.ac.jp
[3] JTEKT Corporation and Tokyo Institute of Technology, Japan
kobayashi.t.cs@m.titech.ac.jp

Abstract. This paper presents an experimental analysis of a prototype measurement system to construct a hand-free Human-Machine Interface (HMI) for controlling wheelchairs, integrating face orientation algorithms and Inertial Measurement Units (IMUs). The aim is to enable individuals with mobility limitations to achieve independent mobility and social interaction. Moreover, the to-be-designed HMI can be an alternative approach to the classical joystick, for people who have health conditions that do not allow the effective use of manual devices. The HMI system utilizes cameras and IMUs to capture the user's head and torso movements. Experimental results are discussed in comparison of IMU and camera data with the Motion Capture system on accuracy and ease of instrumentation, and it has been demonstrated not only the essential reliability of IMU but also the need for fusion algorithms to enhance camera data accuracy. Thanks to the fusion processing of left and right camera images beside a rider's face, face orientation detection becomes robust while requesting no inconvenience for the rider such as wearing specially-made mounting devices, etc. The findings suggest a promising approach for developing inclusive and efficient wheelchair control systems.

Keywords: Human-Machine Interface (HMI), Face orientation algorithm, Motion Capture, Vision system, driving logic, service robots.

1 Introduction

The population of medium age is constantly increasing and linked also people with walking diseases. In the European zone, the number of people with mobility limitations is over 120 million people by 2020 [1]. According to studies, people with severe disabilities have a very hard time learning how to operate a traditional power wheelchair. Up to 10% of people find difficult or impossible to use a power wheelchair, another 10% find it impossible to steer without help, and 40% struggle with steering [2]. Moreover, omnidirectional movements performs better in narrow indoor environment [3].

Nowadays the use of new Human-Machine Interfaces (HMI) enables the development of intelligent wheelchairs. Various approaches developed innovative ways of wheelchair control with a free-hand interface: using EEG [4], voice commands [5],

hand gestures [6], eye gaze [7], tactile information [8], Sip-and-Puff (the use of air pressure variation by inhaling or exhaling via a tube-like object, to generate command signals) [9], tongue [10]. The actual challenge in the research environment is developing a wheelchair that fits with human life standards, developing HMI adapted to user disability and capability to assist the user in navigation (i.e. obstacle avoidance allows more safety operation of the wheelchair).
This article describes research to contribute to the development of a new wheelchair that gives back independent mobility, inclusivity and social interaction thanks to a free-hand and side-by-side wheelchair. This aim also is addressed in the UN Sustainable Development Goals n. 3, Good Health and Well-Being. The under-developing device utilizes a rider's natural movement for driving an omnidirectional wheelchair, capable of realizing the three planar speeds v_x, v_y and ω. A better and faster comprehension of the rider's intention is the advantage of using this type of interface. This paper presents how the sensor of the HMI interface of the wheelchair is used to acquire input from the rider. The acquired inputs are body movements from the head and torso. The accuracy of these signals will be evaluated to achieve the best estimation of the rider's movement which will be transformed into a velocity reference signal for the wheelchair.

2 Wheelchair free-hand working principle and requirements

Fig. 1a shows the working principle of the measurement system corresponds to constructing the hand-free HMI of a wheelchair composed of two control levels, two motor steering wheels, a passive castor wheel, two cameras (Cam_R, Cam_L) and two IMU (IMU1 and IMU2). The control is split between high-level and low-level, the first one is executed in a PC with ROS and the second one is an Arduino microcontroller that performs as an I/O interface for the PC. Each motor steering wheel (L, left and R, right) has a powerful motor, controlled in PWM, to manage the wheel rotation speed $\omega_{R/L}$ and a second motor connected to the steering subsystem to manage the steering angle $\theta_{R/L}$. The movement reference signal for the wheelchair is obtained from the information acquired from cameras and IMU. The cameras' acquired frames are processed by a face orientation detection algorithm which calculates the overall head orientation angle $\theta_{xh}, \theta_{yh}, \theta_{zh}$[11]. The IMU information can come from the head and the torso. Here, from the ease of use and comfortability points of view, it is better not to require any mounting device on the rider's head, so the use of cameras for head motion and IMU for catching torso movement can be said suitable.

2.1 System reference frames and calibration of the IMU and the face orientation recognition

Fig. 1a shows the system reference frame (r.f.) dispositions and Fig. 1b describes the system r.f. tree, with the relationship between parent and child r.f.. The IMU1 r.f. is a child r.f. of the rider's head r.f. and IMU2 is a child r.f. of torso r.f.. Red arrows highlight the desired input and the green arrow is the odometry of the wheelchair.

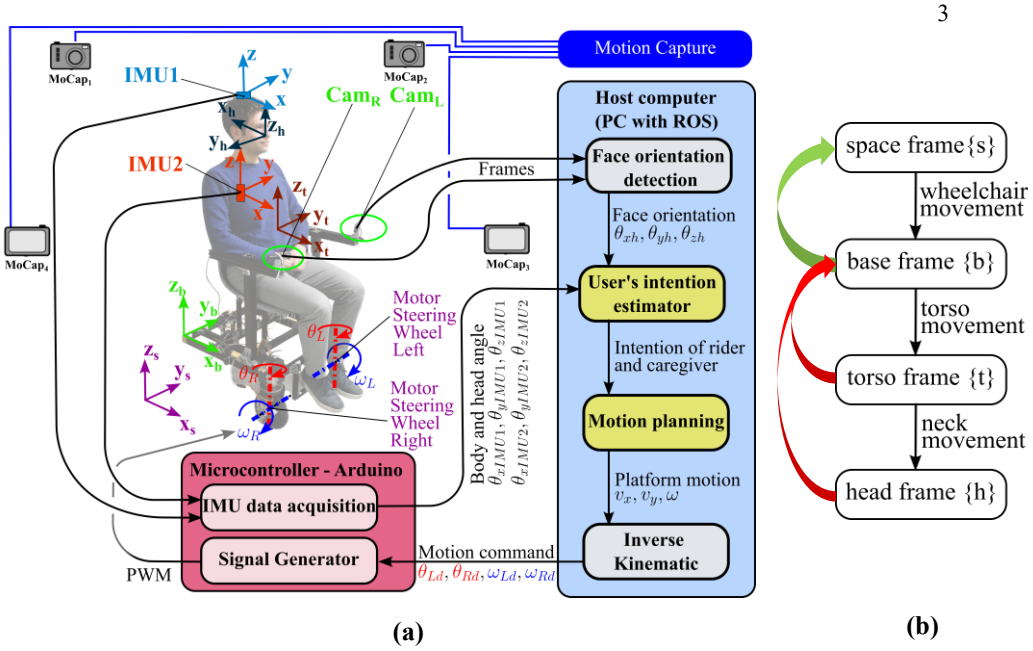


Fig. 1. (a) Wheelchair free hands drive working principle, software and hardware structure with reference frames. (b) Reference frames tree.

The IMU sensors used are 9-axis, which contains an accelerometer, gyroscope and magnetometer. It fuses all the sensor information to obtain the rotations around the x, y and z-axis. The calibration process has two important steps. The first is to detect the rotation matrix between the IMU and the part of the body under measuring. Fig. 1a shows how the head or torso r.f. and the IMU r.f. can be misaligned. The easier way to obtain the calibration matrix is by using the angles that IMU measures when the body or head is in the rest position. The second calibration step is to refer the torso or head measurement to the wheelchair base r.f. (red arrows) because IMU measurement is always referred to the space r.f.. The odometry of the wheelchair (green arrow) can be, measured by placing an additional IMU in the wheelchair r.f.. The subtraction between head or torso IMU and odometry signal gives the measurements of head or torso w.r.t. wheelchair base.

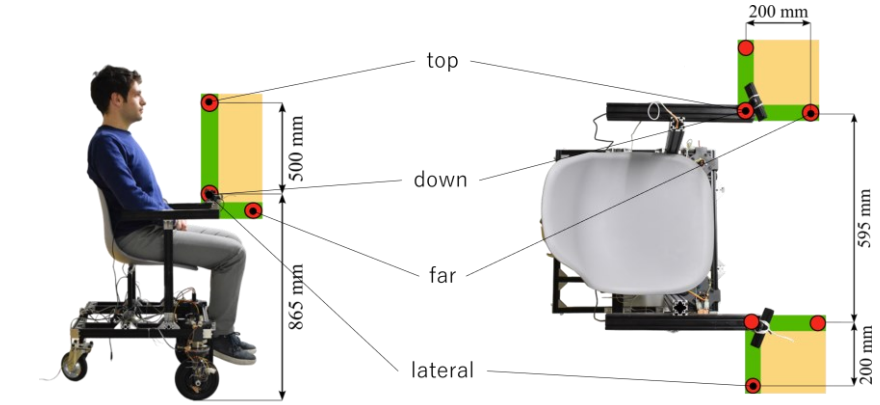
Cameras r.f. measures head orientation w.r.t. base r.f., so it requires only calibration between cameras r.f. and head. Face orientation recognition detects the human face in the camera frames and the physiognomic characteristics (eyes, nose and mouth), which permit the identification of face orientation. Also in this case the easier way to obtain the calibration matrixes is by using the angles that cameras measure when the head is in rest position.

3 Head motion measurement experiment

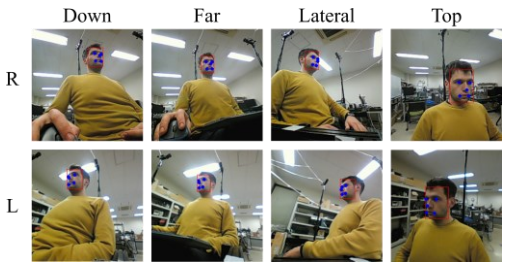
The comparison has been done between cameras, IMU and a Motion Capture system, used as a reference sensor, with the experimental setup shown in Fig. 1a. The aim is to

estimate the quality of the acquired data to understand the feasibility of the hand-free HMI and find the optimal camera position.

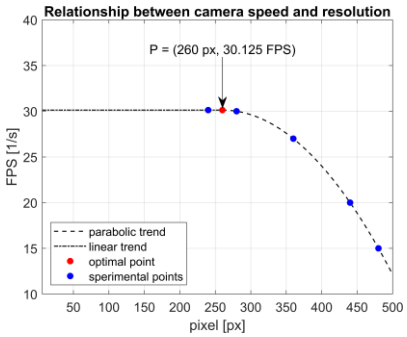
Fig. 2a shows some red circles corresponding to the different camera positions that were investigated during the experiment and Fig. 2b details the camera point of view for each position. The top position is the one with the closest view, but it also obstructs the rider's vision. Lateral and Far positions have a wider field of views but occupy more space. Down has a good position but does not present any pros in the field of view. Moreover, the acquired frames are square because the external parts are affected by lens distortion. Fig. 2c shows the experimental correlation between frame size and frame per second (fps). According to the results, 260 px for each side of the frames is the best compromise that still ensures 30 fps. This result resolution will be used during all the following experiments. The experiment consists of recording yaw, pitch and roll head motion both with camera and IMU to compare with the Motion Capture data. The acquisition of cameras was reiterated for each of the four camera positions.



(a)



(b)



(c)

Fig. 2. (a) Different camera position investigated during the experiment. (b) Camera point of view in different camera position (c) Optimal pixel resolution to obtain best compromise between acquisition velocity (Frames Per Second, FPS) and resolution.

4 Experiment results

4.1 Camera results

Yaw movements detected by the two cameras are reported in Fig. 3. Three different diagrams of the same raw data but with different sizes of elements for the median filters are reported. The raw data are not enough accurate and by using at least a median filter with 7 elements it is possible to remove all the outliers that affect the measuring process. It can be also noticed that comparing the yaw rotation of the camera and of motion capture the positive rotations are better detected by the left camera and vice-versa. This behavior depends on the progressive loss of information about face details when the head rotates in the opposite direction w.r.t. camera position.

Fig. 4 shows the overall comparison between cameras and Motion Capture in all types of rotation and all types of camera positions. Camera data are processed, as mentioned before, with a 7 elements filter to remove noise and outlier values. Pitch rotation angle values over 30° are difficult to measure; only cameras in the top position can detect this rotation since it is high enough to catch the entire pitch range of motions. For roll rotations, there are no evident errors attributable to head orientation, the only errors are related to camera position and noise. For yaw rotations, as mentioned above, there are some errors when the absolute value of the head rotation is more than 30° because only the camera placed in the direction of rotation can give the correct measurement.

4.2 IMU results

Fig. 5 shows the raw results acquired by IMU and compared with Motion Capture data. The two data fit very well with each other and the maximum misalignment is 2.5° at the maximum of head rotation. Moreover, these movements were done at the highest

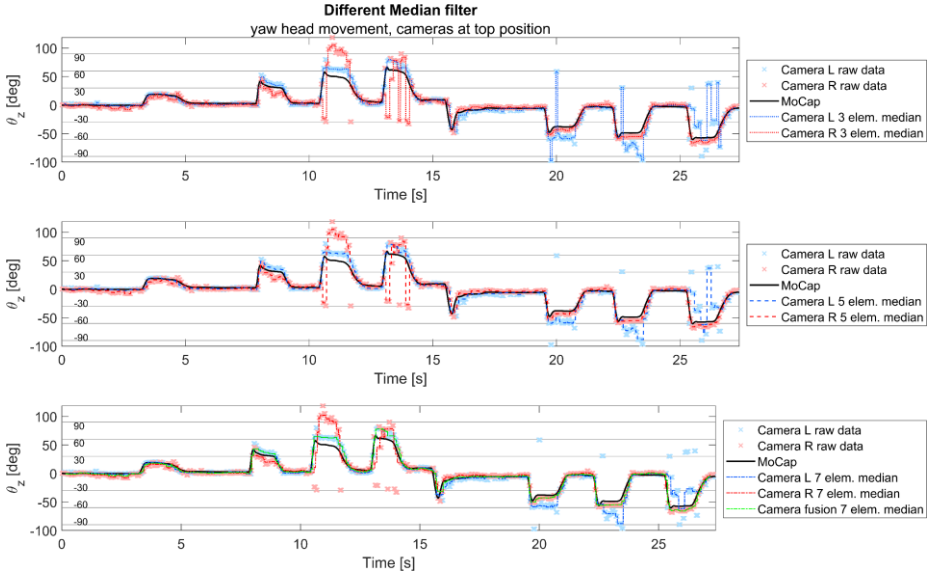


Fig. 3. Raw acquisition of cameras and data processing with different type of median filter.

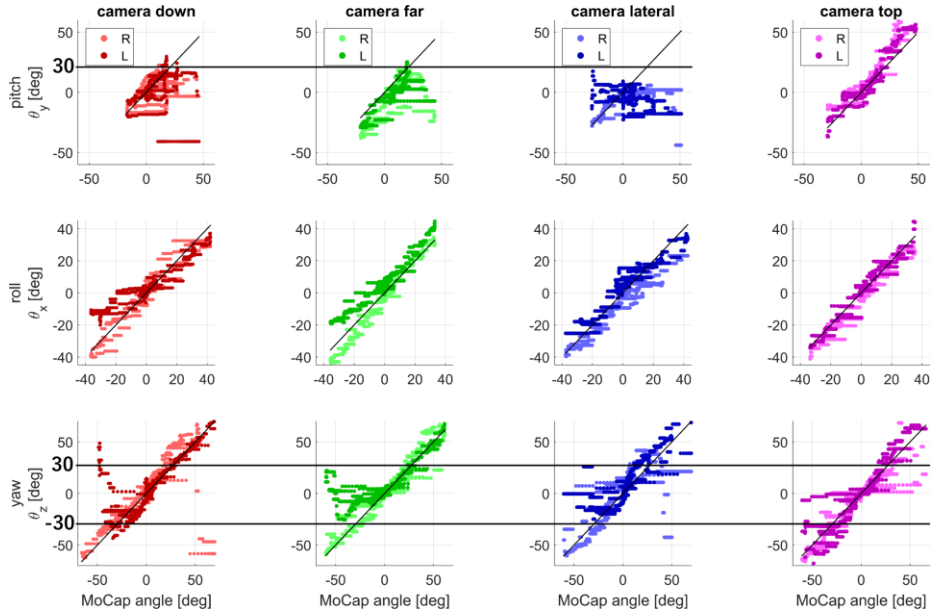


Fig. 4. Comparison between cameras filtered and motion capture data.

possible speed, which of course will affect the measurement precision and accuracy due to some drift effects of the sensors. Altogether the IMU results are reliable and not affected by noise or static errors.

5 Camera information fusion

To overcome the problem related to the lack of visibility of the rider's face and reduce the noise associated with that measurement was developed a camera fusion algorithm to improve the results obtained from cameras. Fig. 6 depicts the overall working of the camera information fusion that takes data from face detection. These two algorithms together compose the face orientation block. For each block that performs pitch, roll, or yaw fusion there is a different equation to use in the best way the data acquired in each configuration. The equations for roll (θ_{xh}), pitch (θ_{yh}) and yaw (θ_{zh}) are the following:

$$\theta_{xh} = mean(\theta_{xL},\ \theta_{xR})$$

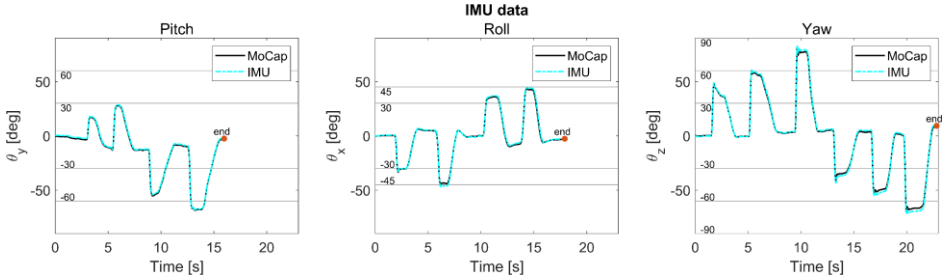


Fig. 5. IMU raw data comparison with Motion Capture.

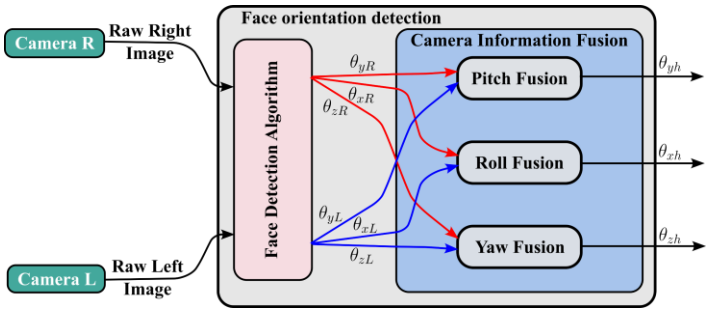


Fig. 6. Working principle of the camera fusion information algorithm.

$$\theta_{yh} = \begin{cases} if\ \theta_{zR} < -30 \rightarrow \theta_{zh} = \theta_{zR} \\ if\ \theta_{zL} > 30 \rightarrow \theta_{zh} = \theta_{zL} \\ else \rightarrow \theta_{zh} = mean(\theta_{zL},\ \theta_{zR}) \end{cases}$$

$$\theta_{zh} = \begin{cases} if\ \theta_{zR} < -30 \rightarrow \theta_{zh} = \theta_{zR} \\ if\ \theta_{zL} > 30 \rightarrow \theta_{zh} = \theta_{zL} \\ else \rightarrow \theta_{zh} = mean(\theta_{zL},\ \theta_{zR}) \end{cases}$$

Fig. 7 shows the results of the camera data filtered and fused. Now the data are more concentrated becoming more similar to the IMU signal. The positive pitch angle can be well detected only with the camera in the top position while, thanks to the fusion, the yaw movement is better detected. Omitting unreliable angle information, especially of pitch angle, is thus an important considerable point in the development of actual HMI.

6 Conclusions

The overall behavior of cameras is improved thanks to the sensor fusion. Although the best position is the top one, it obstructs too much the rider's field of view. However, thanks to this analysis, it is possible to define the best position around the down one, but shifted in direction to the top and a little bit to the far position. Examining qualitatively the quality of the camera filtered and fuse data, it is possible to say that the

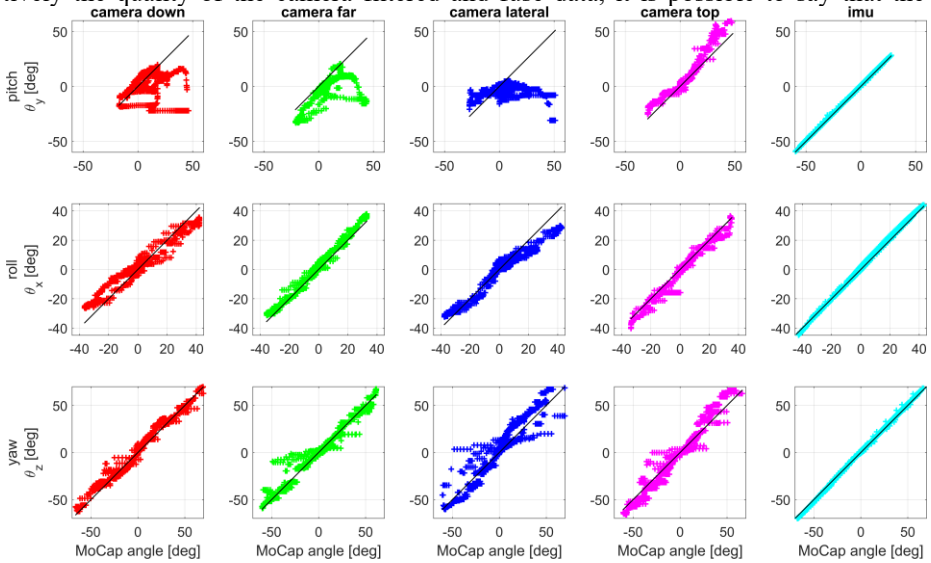


Fig. 7. Comparison between cameras filtered fused, imu and motion capture data.

convergence behavior is similar to IMU and can be fused to obtain a better control signal of the platform. In particular, the IMU signal is indispensable in the case of positive pitch, otherwise, this information will never be acquired properly.

Acknowledgments 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 Tokyo Institute of Technology.

7 References

1. Persons with disabilities - Employment, Social Affairs & Inclusion - European Commission, https://ec.europa.eu/social/main.jsp?catId=1137
2. Fehr, L., Langbein, W.E., Skaar, S.B.: Adequacy of power wheelchair control i terfaces for persons with severe disabilities : A clinical survey. 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. Tagliavini, L., Quaglia, G.: On the Design of MoviWE.Q: An Omnidirectional Electric-Powered Wheelchair for Indoor Mobility. In: Okada, M. (ed.) Advances in Mechanism and Machine Science. pp. 311–321. Springer Nature Switzerland, Cham (2023)
4. Iturrate, I., Antelis, J., Minguez, J.: Synchronous EEG brain-actuated wheelchair with automated navigation. In: 2009 IEEE International Conference on Robotics and Automation. pp. 2318–2325 (2009)
5. Pacnik, G., Benkic, K., Brecko, B.: Voice operated intelligent wheelchair - VOIC. In: Proceedings of the IEEE International Symposium on Industrial Electronics, 2005. ISIE 2005. pp. 1221–1226 vol. 3 (2005)
6. Manawadu, U.E., Kamezaki, M., Ishikawa, M., Kawano, T., Sugano, S.: A hand gesture based driver-vehicle interface to control lateral and longitudinal motions of an autonomous vehicle. In: 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC). pp. 001785–001790 (2016)
7. Matsumotot, Y., Ino, T., Ogsawara, T.: 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). pp. 262–267 (2001)
8. Matsuura, D., Kiyomoto, K., Kobayashi, T.: Development of a simple human-cooperative transportation system based on tactile information. ICPE2022. Proceedings of the 19th International Conference on Precision Engineering,
9. Mougharbel, I., El-Hajj, R., Ghamlouch, H., Monacelli, E.: Comparative study on different adaptation approaches concerning a sip and puff controller for a powered wheelchair. In: 2013 Science and Information Conference. pp. 597–603 (2013)
10. Lund, M.E., Christiensen, H.V., Caltenco, H.A., Lontis, E.R., Bentsen, B., Andreasen Struijk, L.N.S.: Inductive tongue control of powered wheelchairs. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology. pp. 3361–3364 (2010)
11. fisakhan: fisakhan/Face_Pose, https://github.com/fisakhan/Face_Pose, (2024)