

Jornadas de Automática

Guiado de una silla de ruedas robótica mediante un dispositivo portable de seguimiento ocular

Frigola, M^a, Secchieri, L^b, Aranda, J^a, Casals, A.^a

^a Centre de Recerca en Enginyeria Biomèdica, Universitat Politècnica de Catalunya, Jordi Girona 1-3, Barcelona, España.

^b Università degli Studi di Padova, Dipartimento di Ingegneria dell'Informazione

To cite this article: Frigola, M., Secchieri, L., Aranda, J., Casals, A., 2024. Guidance of a robotic wheelchair using a wearable eye-tracking device. *Jornadas de Automática*, 45. <https://doi.org/10.17979/ja-cea.2024.45.10850>

Resumen

La necesidad de aportar una solución a la interfaz de usuario de un dispositivo robótico cuando la capacidad de actuación es muy limitada a causa de una discapacidad severa, supone el diseño de dispositivos que puedan adaptarse a estas capacidades de forma efectiva. En el proyecto Posmofya, se pretende desarrollar una plataforma robotizada de múltiples grados de libertad consistente en una silla de ruedas que incorpora control postural, además de una estructura tipo exoesqueleto de extremidad superior para asistir también en la actuación de brazos y mano. En este artículo se describe la problemática en el desarrollo de una interfaz de guiado basada en el seguimiento ocular, *eye-tracking*, con la particularidad de hacerlo desde un dispositivo portable y ergonómico (unas gafas equipadas con micro-cámaras). El sistema es usado como interfaz de selección de modo de operación de la plataforma y en particular para el guiado efectivo en tiempo real de la silla de ruedas, pudiendo controlar de forma continua la dirección y velocidad de avance/retroceso. Finalmente, se presentan los resultados de un test preliminar comparando el guiado con el sistema propuesto respecto al guiado usual con una palanca de mando (joystick).

Palabras clave: Interfaces inteligentes, Interacción multimodal, Tecnología asistencial, Percepción y sensorización, Soporte al operador humano, sistemas embebidos.

Abstract

The need to provide a solution to the user interface of a robotic device when the action capability of the user is very limited due to a severe disability, requires the design of devices that can adapt to these capabilities effectively. In the Posmofya project, the aim is to develop a multiple degrees of freedom robotic platform consisting of a wheelchair that incorporates postural control and also an upper limb exoskeleton-type structure to also assist the user in manipulation procedures. This article describes the problems in the development of a guided interface based on eye tracking, with the particularity of doing it from a portable and ergonomic device (glasses equipped with micro-cameras). The system is used as an interface for selecting the platform's operating mode and in particular for the effective real-time guidance of the wheelchair, being able to control continuously the direction and forward/reverse speed. Finally, the results of a preliminary test are presented comparing the guidance with the proposed system with respect to the usual guidance with a joystick.

Keywords: Intelligent interfaces, Multi-modal interaction, Assistive technology and rehabilitation engineering, Perception and sensing, Human operator support, Embedded systems.

1. Introducción

Cuando la capacidad de actuación de las personas es muy limitada a causa de una discapacidad severa, las ayudas tecnológicas que la robótica puede aportar a estos usuarios deben superar la barrera que en ocasiones presentan las

interfaces comúnmente usadas y disponibles en el mercado. Las limitaciones operativas de un usuario con discapacidad llevan consigo la necesidad de adaptar la interfaz de control de dispositivos robóticos a sus posibilidades. Este requisito comporta mayores exigencias cuando se trata de controlar

múltiples grados de libertad de forma segura, fiable y con respuesta en tiempo real.

El proyecto Posmofya desarrolla una plataforma híbrida consistente en una silla de ruedas con control postural y provista de grados de libertad adicionales tipo exoesqueleto de extremidad superior para asistir a personas con movilidad reducida también en las extremidades superiores (Casals, 2022). La plataforma está concebida para usuarios con limitaciones de movilidad, no solo para desplazamientos, sino también de su capacidad de actuación con brazos y manos, y por tanto el diseño de la interfaz para un control eficiente y seguro de los múltiples grados de libertad es de especial relevancia. En este contexto, la interfaz propuesta debe aportar soluciones adaptables a la tipología de las necesidades del usuario (Minguella, 2024), (Sharmila, 2020) e incluir una integración multimodal de los sensores para complementar la información limitada de algunos de ellos, así como proporcionar la redundancia necesaria que garantice una interpretación segura de la voluntad del usuario (Muhammad Sidik, 2018).

El seguimiento de la mirada, por su interés como sistema intuitivo y “manos libres” de guiado, ha dado lugar a diversos dispositivos de interfaz comerciales para facilitar su integración en múltiples aplicaciones. En este trabajo se presenta un sistema de guiado basado en el seguimiento ocular, *eye-tracking*, integrado en unas gafas portables equipadas con cuatro cámaras y evitando la necesidad de disponer de una pantalla de interfaz, como la mayoría de sistemas de guiado de este tipo. El problema que aparece respecto a estos sistemas es que la información de dirección ocular es relativa a las gafas y estas no están adheridas a la silla, sino a la cabeza del usuario, y por tanto debe añadirse a la consigna de control la orientación relativa entre cabeza del usuario y silla. Al ser importante no solamente la dirección, sino la distancia del punto objetivo que permita determinar también la velocidad del desplazamiento en la movilidad o la posición del objeto de interés en manipulación, es necesario la detección y localización de objetos o patrones en un escenario complejo (Wang, 2024), lo que representará también una de las contribuciones del trabajo.

La interfaz debe además dar la posibilidad de operar en distintos modos. Por una parte, poder decidir entre los modos de control: control manual mediante joystick o control de guiado mediante la mirada (manos libres). Por otra parte, debe también poder seleccionar los modos de operación de la plataforma híbrida, la silla robotizada. Para este trabajo se han definido los siguientes modos de operación: *Conducción*, *Acomodación* y *Manipulación*.

En la sección 2 se presenta la metodología seguida, con la descripción de los componentes del sistema de guiado, la interfaz, sus funciones y la operativa para seleccionar los modos de operación. La sección 3 explica las pruebas realizadas y resultados. Finalmente, la sección 4 expone las conclusiones y trabajos futuros.

2. Metodología

Uno de los objetivos del proyecto es que la conducción de la silla de ruedas robotizada sea colaborativa; entendiendo por colaborativa que el usuario tenga un control a alto nivel y el

computador embarcado tome la decisión de los detalles de control de más bajo nivel. La experimentación se hace sobre una silla Permobil C300 a la que se dota de grados de libertad adicionales con un exoesqueleto de miembro superior. El esquema de la figura 1 muestra los sensores que se han añadido a la silla o al usuario, indicados de 1 a 6, un dispositivo EEG, unas gafas de seguimiento de la mirada, *eye tracking*, una cámara frontal y un micrófono incorporados ambos en las gafas, un sensor inercial (IMU) en la silla y otro en las gafas, y además un módulo de medidas de distancia por láser (LIDAR) para detectar obstáculos en las maniobras marcha atrás. Como elementos de cálculo un computador embarcado, con el nombre DRIVE AI (12), que tiene por objetivo interpretar la voluntad del usuario y enviar las órdenes de control de más bajo nivel a un módulo llamado uC Bridge (10). El módulo uC Bridge que la silla pueda ser guiada manualmente mediante el joystick (JSM) o bien por el módulo DRIVE AI.

La silla permite varios modos de actuación, el primer modo, llamado de *Conducción*, permite el guiado para maniobrar en el entorno utilizando la mirada. El segundo modo, *Acomodación*, permite ajustar el ángulo del respaldo, el ángulo del asiento y el ángulo del reposapiés. El tercero está previsto que sea el modo *Manipulación* que servirá para activar el brazo robot. Existe un modo *Neutro*, que sería el equivalente a una silla de ruedas normal accionable únicamente con el joystick de maniobra.

Para pasar de un modo a otro se han implementado varios métodos: mediante pulsadores, mediante comandos de voz y un tercer método mediante la mirada. El método de cambio de modo mediante la mirada, se ha realizado utilizando una técnica llamada de carrusel temporal, de uso muy habitual en interfaces para personas con fuerte discapacidad: si se cierran los ojos se activa el carrusel que irá pasando de un modo a otro cada cuatro segundos. Cuando se llega al modo Neutro es necesario volver a abrir los ojos y volver a cerrarlos para seguir en el carrusel. El cambio de modo se va indicando al usuario de forma sonora mediante la repetición de notas auditivas.

El procedimiento para realizar la conducción guiada por la mirada comporta el uso de diferentes sensores y su procesado en relación a las funciones implicadas: detección de las pupilas, estimación de la profundidad de la mirada, corrección de los ángulos de colocación de las gafas, estimación de los ángulos de la cabeza y conversión de la mirada a órdenes de movimiento.

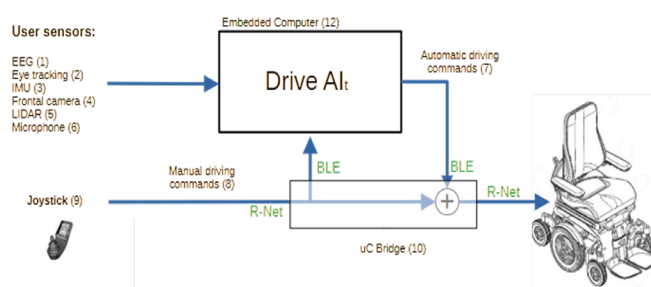


Figura 1: Esquema general del módulo de conducción

2.1 Detección de pupilas

Las gafas utilizadas son las *Tobii© Pro Glasses 3* que disponen de una cámara frontal para captación del entorno y dos cámaras en cada cristal para captar los movimientos de la pupila. El procesado para la detección de las pupilas se basa en una binarización mediante un umbral adaptativo, a partir de un preprocesado previo para evitar falsos positivos: un filtro de mediana que reduce el efecto de los reflejos, seguido de un procesado morfológico por erosión para realzar las regiones negras. A partir de los contornos detectados en la imagen con las pupilas realzadas, se extraen los círculos mínimos envolventes y se consideran pupilas aquellos círculos cercanos al centro de las imágenes con un diámetro dentro de unos rangos establecidos. La figura 2 muestra la posición de la pupila en el plano imagen para cada una de las cámaras, indicando en rojo el contorno y en verde el centro. En el caso de no detectar un contorno válido, se descarta el cuadro correspondiente. Este método proporciona una detección precisa de la pupila que será utilizada tanto para determinar la dirección de la mirada como para estimar la profundidad de la mirada.

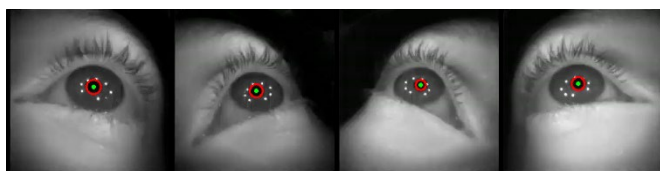


Figura 2: Detección de las pupilas a partir de las cámaras en las gafas *Tobii Pro Glasses3*.

Nota: Los puntos blancos alrededor de la pupila son el reflejo de la iluminación por infrarrojo.

2.2 Estimación de la profundidad

Para estimar la posición 3D del punto de fijación de la mirada se parte de la posición detectada de las pupilas en el apartado anterior. Esta parte supone cierta dificultad puesto que la pupila rota sobre una esfera (aproximación del globo ocular) cuyo centro no queda claramente determinado y depende fuertemente de la posición en que se colocan las gafas y la particular anatomía del usuario. Cabe decir que incluso determinando con precisión el centro del globo ocular, el centro de rotación del ojo y el eje de visión generalmente no coinciden (Iskander, 2018). Para evitar establecer un modelo excesivamente multi-paramétrico se ha escogido un método en dos etapas, una primera etapa de calibración-regresión, y una segunda etapa de corrección de la posición de las gafas.

A partir de la posición de las pupilas se estima la profundidad de la mirada utilizando un modelo de aprendizaje automático, una regresión gaussiana, *Gaussian Process Regression (GPR) model*, (Rasmussen, 2005). Se han comparado dos enfoques distintos: un modelo de regresión simple de los datos sin linealizar y una estándar, un modelo de regresión con linealizado. El error cuadrático medio obtenido del modelo linealizado mediante una función logarítmica fue $MSE = 14,67$ cm, mientras que con el modelo no linealizado se consiguió un $MSE = 12,62$. A pesar de que el modo linealizado da un MSE mayor, la figura 3 ilustra que su

comportamiento es mejor a distancias cercanas. Dada la relación exponencial entre la vergencia del ojo y la distancia a un objeto, la entrada de datos linealizada a través de una transformación logarítmica mejora la precisión del modelo para distancias cortas, al tiempo que mantiene una respuesta razonable para objetos más lejanos.

2.3 Estimación de la colocación de las gafas

Cada vez que el usuario se coloca de nuevo las gafas, estas pueden quedar ligeramente desplazadas en relación a cómo fueron colocadas durante la calibración. Unas gafas descolocadas afectan fuertemente el resultado del cálculo de la posición 3D de la mirada. Para mitigar el efecto de la colocación de las gafas es necesario realizar una rectificación de la posición de las pupilas detectadas en el apartado anterior. De los posibles seis valores a determinar en la colocación de las gafas (3 ángulos y 3 desplazamientos) se ha corregido únicamente el ángulo roll y un ligero avance en la colocación.

Analizando la altura de las pupilas y suponiendo que el usuario tiene los ojos equilibrados (no estrábico) se puede determinar la inclinación (roll) y mediante la distancia entre pupilas se puede determinar si las gafas están avanzadas, Figura 4.

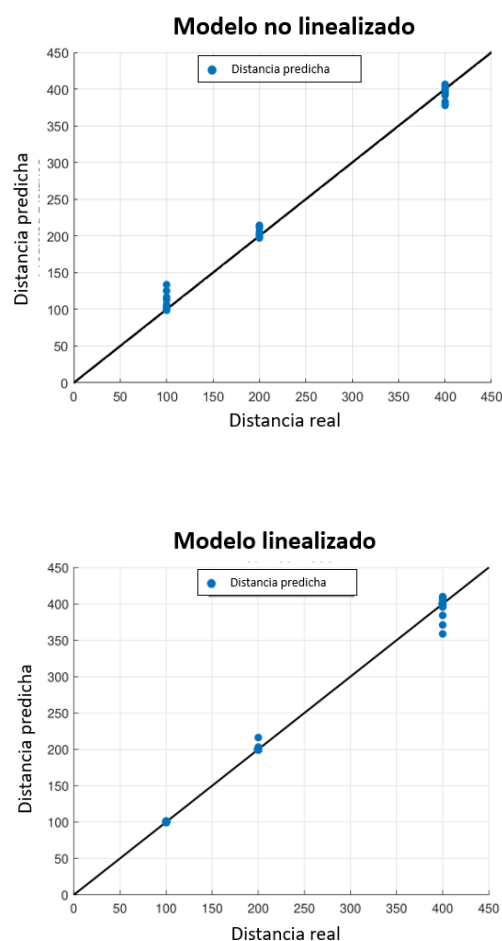


Figura 3. Comparación de la distancia real versus distancia predicha utilizando la linealización de los datos



Figura 4. Diferencia de alturas entre pupilas debido a una incorrecta colocación de las gafas (líneas horizontales roja y azul) y corrección teórica del centro de las pupilas (azul claro)

2.4 Estimación de la orientación de la cabeza

Para que la silla de ruedas se mueva en la dirección de la mirada se hace necesario sumar o añadir a los ángulos de rotación de los ojos, cuáles son los ángulos de giro de la cabeza respecto a la silla, tomando como base los ejes de referencia de la silla de ruedas. Al no haber cámaras frontales ni pantallas instaladas solidariamente en la silla de ruedas, como sí ocurre en general en los sistemas clásicos de guiado mediante la mirada, se hace necesario estimar los ángulos de rotación de la cabeza de forma indirecta. Los ángulos de balanceo (*Roll*) y de abatimiento (*Pitch*) de la cabeza son fáciles de estimar mediante el sensor inercial (IMU) incorporado en las gafas, figura 5.

Gracias a que el sistema usuario-silla se mueve normalmente con una baja dinámica, el vector gravedad predomina sobre los demás valores de aceleración. A partir de la medida del acelerómetro $a = (ax, ay, az)$, cuando la silla no sufre aceleraciones, solo queda la influencia de la gravedad, con lo cual $a \approx g$. En estos casos, los ángulos de inclinación de la cabeza se pueden calcular mediante las ecuaciones 1 y 2:

$$Pitch = \arcsin(ax/g) \quad (1)$$

$$Roll = \arctan2(ay, az) \quad (2)$$

$$Yaw = \arccos(a' \cdot w / (|a'| \cdot |w|)) \quad (3)$$

Un tema más complejo es la estimación del giro en horizontal de la cabeza o ángulo de guiñada (*Yaw*) en relación a los ejes de la silla.

Una posibilidad para estimar el ángulo de giro de la cabeza en relación a la silla es el uso de estimaciones diferenciales utilizando dos IMUs. Una IMU situada en el asiento, y la IMU que lleva incorporada las gafas, figura 6. El ángulo que mantienen ambas medidas de la aceleración: a' y el vector w en el plano horizontal, ecuación 3), nos da una estimación del ángulo *Yaw* que mantiene la cabeza del usuario en relación a la silla.

Otra posibilidad es utilizar una estimación del norte magnético mediante las medidas proporcionadas por el magnetómetro, que también lleva incorporado las gafas sensorizadas. La estimación del ángulo girado mediante el norte magnético no proporciona la referencia en relación al sistema de coordenadas de la silla de ruedas, inconveniente que obliga a incorporar un segundo magnetómetro en la base de la silla. Mediante el segundo magnetómetro se pueden establecer el ángulo de giro de la silla y el ángulo de giro de la cabeza en relación al norte magnético y la diferencia de ambos valores proporciona el ángulo de la cabeza en relación al sistema de coordenadas de la silla.

Otra técnica utilizada es fijar un imán colocado en el respaldo de la silla cerca de la cabeza, solidario con los

movimientos de la silla de ruedas. Si no hay ninguna perturbación del campo magnético, el magnetómetro apunta aproximadamente en la dirección donde se encuentra el imán $m = (mx, my, mz)$. Mediante un método de calibración del campo magnético circundante al imán, se puede estimar el ángulo real de las gafas y en consecuencia el de la cabeza en relación a la silla. La calibración del campo magnético circundante al imán se hace necesaria debido a que este varía fuertemente según la posición. El problema asociado con estos métodos de estimación del norte o de un campo magnético artificial (imán) es que queda fuertemente afectado por los cambios posturales del respaldo de la silla y por la presencia de dispositivos (ferromagnéticos) que estén cerca de la cabeza del usuario: tales como teléfonos móviles, reposacabezas, etc.

Las pruebas realizadas no nos han proporcionado resultados concluyentes sobre que técnica proporciona mejores estimaciones. En espera de una evaluación más exhaustiva se ha utilizado la técnica del norte magnético.

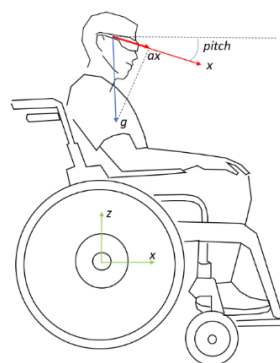


Figura 5. Ángulo de abatimiento de la cabeza (*pitch*) en relación a la gravedad (*g*)

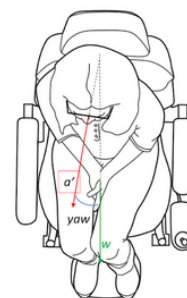


Figura 6. Ángulo *yaw* en relación a los ejes de la silla de ruedas (imagen de fondo Vector_Vault)

2.5 Conversión de la mirada a órdenes de movimiento

Una vez obtenida la dirección de la mirada en referencia a los ejes de la silla de ruedas es necesario establecer las órdenes de movimiento. Para ello es preciso adaptar la señal generada a partir del ángulo de la mirada al sistema de control de la silla.

En el caso del modelo de la silla utilizada (Permobil C300) el accionamiento se realiza mediante el envío de tramas de datos (utilizando el protocolo de comunicaciones R-Net ©) indicando la inclinación del joystick de maniobra (JSM). La inclinación del joystick se indica con un par de valores (J_x, J_y), contenidos entre los valores de -100 a 100, siendo (0,0) la posición de reposo del joystick, (0,100) avance a velocidad

máxima, (0,-100) retroceso, (100,0) giro a la derecha, (-100,0) giro a la izquierda. Es por tanto necesario mapear el ángulo de la mirada a la dirección vector (J_x, J_y) , y la velocidad se indica con el módulo de dicho vector que se calcula en función de la profundidad de la mirada.

Se ha establecido que la velocidad máxima de la silla se alcance cuando el usuario mira más allá de los 4 metros de distancia, cuando más cercana es la focalización de la mirada más lenta será la velocidad de la silla. Para poder controlar la silla marcha atrás, se ha implementado un esquema de control que se activa cuando el usuario mira muy cerca de su cuerpo, entonces la silla retrocede invirtiendo el ángulo de giro indicado por la mirada. Por motivos de seguridad, en las maniobras de marcha atrás la velocidad se mantiene fija a un 0.1 m/seg.

3. Test preliminar y resultados

Se ha realizado un test preliminar utilizando usuarios con motricidad normal debido a que el sistema está todavía en desarrollo. En estas fases de desarrollo no se ha considerado adecuado utilizar aún usuarios con baja movilidad. Se ha medido únicamente la sobrecarga en tiempo (tiempo adicional) que supone la conducción mediante la mirada en relación a la conducción con el joystick (JSM) se ha diseñado una actividad consistente en seguir un camino preestablecido que incluya cambios de dirección, cambios de sentido (180°) y márgenes estrechos.

El test supone realizar una conducción a lo largo de un pasillo suficientemente ancho de 4 metros de longitud, un giro de 90 grados a la izquierda, una conducción en un pasillo estrecho de 9 metros de largo y 1 metro de ancho, un giro de 180 grados y vuelta a la posición de origen. El total del recorrido es de aproximadamente 26 metros.

Primero, 4 usuarios con motricidad normal, sin experiencia en el manejo de sillas de ruedas motorizadas, utilizando el joystick de maniobra, han realizado el recorrido en 1 minuto 7 segundos de promedio, con una desviación máxima de ± 9 segundos.

En la segunda parte, un usuario diferente a los anteriores, con motricidad normal, ha realizado el mismo ejercicio usando el sistema de guiado con la mirada diversas veces. El tiempo invertido fue de 1 minuto 36 segundos de promedio con una desviación máxima de 13 segundos. Lo que supone una sobrecarga temporal de aproximadamente un 30%. Las variaciones más significativas se han observado en el paso por el pasillo estrecho, donde una pequeña desviación de la mirada hace que la silla derive hacia un lado y deba corregirse con cierta frecuencia, provocando pequeñas oscilaciones del recorrido.

4. Conclusiones y trabajos futuros

Los primeros resultados demuestran la funcionalidad del guiado mediante la mirada sin utilizar la pantalla de referencia en la misma silla, como utilizan normalmente los sistemas de guiado de este tipo y que dificultan la operatividad al obstaculizar la interacción con el entorno. Aunque en esta fase preliminar la sobrecarga respecto al uso de joystick es mayor, el test realizado muestra la viabilidad del guiado. En fases

posteriores, una mayor experimentación y nuevos desarrollos permitirán mejorar los resultados del guiado por la mirada.

Además de las mejoras en el guiado de la navegación, en la futura investigación es importante identificar cuando la mirada es intencional o resultado de una actividad cotidiana del usuario, puesto que en este caso la interfaz debe detectar que la mirada no es una intención de control, sino que responde a la actividad o interés del usuario al margen del control de movimiento. Para evitar la activación de la interfaz cuando no hay intención voluntaria de generar nuevas órdenes o indicaciones de guiado, el llamado *Midas Touch Problem*, es preciso disponer de otra información adicional que interprete el estado o interés del usuario. Se está explorando la posibilidad de utilizar otro tipo de datos adicionales, o la propia dinámica de la mirada que permita mediante técnicas de aprendizaje interpretar conductas que discriminen el movimiento de orden voluntaria respecto a movimientos derivados de la actividad del usuario.

Agradecimientos

Esta investigación ha sido financiada por el proyecto "Orthosis-Chair Hybrid Platform to make compatible Mobility, Functionality and Acceptability for application in domestic environments" (POSMOFYA) del Ministerio de Ciencia e Innovación, Proyecto: PID2020-116091RB-I00.

La integración de la interfaz de usuario con el sistema de control de la silla utilizada para la experimentación ha contado con la información suministrada por Stephen Chavez & Specter en su proyecto *R-Net CAN library for power-wheelchairs with R-Net electronics*.

Asimismo, se ha contado con la aportación de forma desinteresada de imágenes vectoriales representativas de la diversidad, extraídas del *Vector-Vault project*. www.vector-vault.org/about

Referencias

- Casals, A. et al. A preliminary approach to a wheelchair with embedded exoskeleton capabilities. Fifth Iberian Robotics Conference. ROBOT 2022, LNNS, Vol 590, Springer
- Gemicioglu, T. et al. Gaze & Tongue: A Subtle, Hands-Free Interaction for Head -Worn Devices, CHI Conference on Human Factors in Computing Systems. 2023. DOI:10.1145/3544549.3583930
- Higa, S. Intelligent Eye-Controlled Electric Wheelchair Based on Estimating Visual Intentions Using One-Dimensional Convolutional Neural Network and Long Short-Term Memory. Sensors. 2023, DOI: 10.3390/s23084028
- Iskander, Julie & Hossny, Mo & Nahavandi, Saeid. (2018). A Review on Ocular Biomechanic Models for Assessing Visual Fatigue in Virtual Reality. IEEE Access. PP. 1-1. 10.1109/ACCESS.2018.2815663).
- Minguella, J., Frigola, M., Casals, A., 2024. Multi-material 3D printed multiple state sensor units: design and materialisation of a command joystick set. Journal of Mechanical Science and Technology. In press. DOI: pending
- Muhammad Sidik, M.H., et al. A review on electric wheelchair innovation to ease mobility and as a rehabilitation tool for spinal cord impairment patient. Int. J Eng. Technol. 2018; 10:803–815.
- Rasmussen, C.E. et Williams, C.K.I. Gaussian Processes for Machine Learning. The MIT press, 2005
- Sharmila, A., Hybrid control approaches for hands-free high level human-computer interface-a review. J Med Eng Technol. 45 (1), 2021. DOI: 10.1080/03091902.2020.1838642.
- Wang, X., et al. Dual Regression-Enhanced Gaze Target Detection in the Wild. IEEE Trans. On Cybernetics, 54 (1), 2024.
- Yeamkuan, S., Chamnongthai, K. 3D Point-of-Intention Determination Using a Multimodal Fusion of Hand Pointing and Eye Gaze for a 3D Display, Sensors, 2021. DOI: org/10.3390/s21041155