

Jornadas de Automática

Guante Vestible para la Detección de Gestos usando Sensores Inerciales

Benito, A.^a, Bernat, L.^a, García, G.J.^a, Ramón, J.L.^a, Jara, C.A.^a, Ubeda, A.^a

^a Human Robotics Group, Universidad de Alicante, Edificio Escuela Politécnica Superior III, Crta. de San Vicente del Raspeig, S/N, 03690 San Vicente del Raspeig, España.

To cite this article: Benito, A., Bernat, L., García, G.J., Ramón, J.L., Jara, C.A., Ubeda, A. 2024. Wearable Glove for Gesture Detection using Inertial Sensors. *Jornadas de Automática*, 45. <https://doi.org/10.17979/ja-cea.2024.45.10945>

Resumen

En este artículo se presenta un guante vestible que permite obtener la información de las posiciones relativas de la mano mediante la incorporación de varios sensores inerciales que registran los ángulos de movimiento de los dedos junto con la orientación de la palma de la mano. El dispositivo se ha validado preliminarmente clasificando tres gestos básicos de la mano con resultados muy prometedores, lo cual indica que el uso del guante es viable para el reconocimiento de un mayor número de gestos. En el futuro, se ampliará la base de datos de gestos clasificados y se incorporarán herramientas en tiempo real de detección para poder aplicarse en la comunicación directa entre personas sordas y su entorno o la evaluación de movimientos en el ámbito de la rehabilitación de la mano.

Palabras clave: Interfaces Hombre-Máquina, Aprendizaje Automático, Tecnología asistiva e ingeniería de rehabilitación

Wearable Glove for Gesture Detection using Inertial Sensors

Abstract

This paper presents a wearable glove that allows for obtaining the relative hand position information by incorporating several inertial sensors that record the angles of movement of the fingers as well as the position of the hand. The device has been preliminarily validated by classifying three basic hand gestures with very promising results indicating that the use of the glove is feasible for the recognition of a larger number of gestures. In the future, the database of classified gestures will be expanded and real-time detection tools will be incorporated to be applied in direct communication between deaf people and their environment or the evaluation of movements in the field of hand rehabilitation.

Keywords: Man-machine interfaces, Machine Learning, Assistive technology and rehabilitation engineering

1. Introducción

La detección de gestos es un área muy explorada en el campo de la inteligencia artificial y es especialmente útil en el ámbito de la asistencia a personas sordas para interpretar automáticamente el lenguaje de signos (Galván-Ruiz et al., 2020) o en el ámbito de la rehabilitación motora para evaluar el control motor de la mano en, por ejemplo, un proceso de rehabilitación física (Tchantchane et al., 2023).

Existen varias formas de detectar gestos automáticamente de forma efectiva, siendo la visión artificial

una de las alternativas más utilizadas (Oudah et al., 2020). Aunque este método no requiere ningún tipo de instrumentación sobre el usuario es muy dependiente de condiciones externas como la iluminación, el tipo de tecnología de captura de imagen o la posición del usuario respecto a las cámaras.

Una alternativa interesante para evitar esta problemática es el uso de sistemas de sensores vestibles, por ejemplo, mediante el uso de sensores inerciales colocados en puntos específicos del miembro superior y la estimación de la cinemática resultante de la realización de diferentes posturas

o gestos (Jeon et al., 2022). Esta alternativa es más compleja por la necesidad de emplear múltiples sensores en una zona relativamente pequeña. El uso de wearables o guantes vestibles permite medir el movimiento relativo de las falanges y la muñeca para evaluar el gesto que está realizando el usuario (Lin et al., 2018). En este caso, la opción más precisa, si se quiere determinar por completo la cinemática articular de la mano, es emplear un sensor por cada falange. Esta alternativa es compleja a nivel de instrumentación y adquisición de datos y puede no ser necesaria si lo que se pretende es evaluar gestos sencillos de comunicación con el entorno. Además, el hecho de incluir un número alto de sensores inerciales incrementa los costes del dispositivo y la complejidad de interconexión.

Por este motivo, en este artículo se presenta un dispositivo de bajo coste para la detección de gestos básicos de la mano que consiste en un guante vestible con únicamente 6 sensores inerciales situados en los extremos de los dedos y la palma de la mano que permite obtener las posiciones estáticas de la mano respecto a una referencia fija. El dispositivo se ha evaluado preliminarmente para discriminar algunos gestos principales de interacción con resultados muy satisfactorios. La clasificación inicial se muestra muy prometedora mediante el uso algoritmos no lineales como *k-NN* (*k-Nearest Neighbors*), *Random Forest* o *Naive Bayes*. En el futuro, se pretende incrementar el número de gestos en la base de datos para poder implementar un reconocimiento en tiempo real y validar la utilidad del dispositivo en un entorno realista como es el caso de aplicaciones de comunicación para personas sordas o sistemas de evaluación de la rehabilitación física de la mano.

2. Diseño del guante vestible

2.1. Estructura del guante y distribución de sensores

El diseño propuesto para el sistema de detección de gestos se basa en la integración vestible de varios sensores inerciales que permiten obtener la posición estática de los dedos y la palma de la mano. Para ello se emplea un guante de tela, al que se añaden los siguientes elementos:

- 6 sensores inerciales altIMU 10 v-6 de Pololu donde cinco se han cosido en el anverso de cada falange distal de manera vertical y el restante en el centro del dorso de la mano (Figura 1). Esta colocación posibilita obtener un mayor rango de movimiento para poder medir la inclinación de los dedos al gesticular y poder interpretar su orientación con el centro de la palma de la mano.
- Un multiplexor Adafruit TCA9548 que es empleado para poder adquirir y coordinar la entrada de cada una de las líneas de datos que envían los sensores inerciales.
- Una placa ESP32 WROOM en la que se introduce un software con ARDUINO IDE y que captura y envía los datos a la línea de procesamiento.

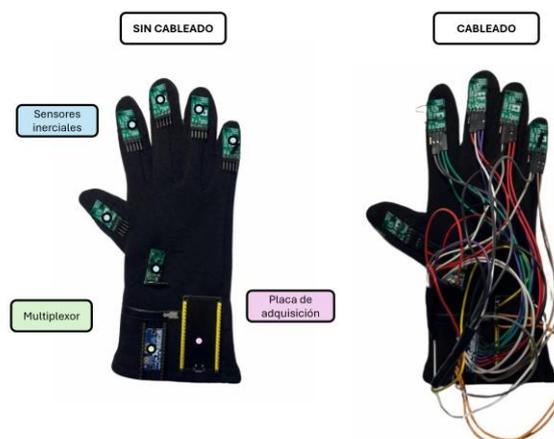


Figura 1: Aspecto general del guante diseñado y colocación de los sensores y elementos de adquisición

2.2. Comunicación entre sensores y al PC

Los sensores inerciales transmiten datos de posición a través del protocolo I2C, utilizando los puertos SDA y SCL, hacia el multiplexor. Este dispositivo gestiona los distintos canales de comunicación solicitados por la placa ESP32, que está conectada a los puertos SDA y SCL del multiplexor. La necesidad de un multiplexor surge para alternar el acceso a las direcciones fijas generadas por los sensores.

Cada sensor se enlaza al multiplexor a través de conexiones específicas: el primer sensor se conecta de SDA a SDA1 y de SCL a SCL1, y así sucesivamente con los siguientes cinco sensores. Además, tanto los sensores como el multiplexor comparten las conexiones de alimentación y tierra.

Finalmente, el multiplexor conecta su salida SDA al pin GPIO 21 y su salida SCL al GPIO 22 de la placa ESP32, responsables de la integración y el manejo de señales entre los dispositivos (Figura 2). La placa ESP32 se conecta al PC mediante un puerto USB, empleado como puerto serial (COM) y envía los datos de todos los canales recibidos para que el PC pueda procesar la información.

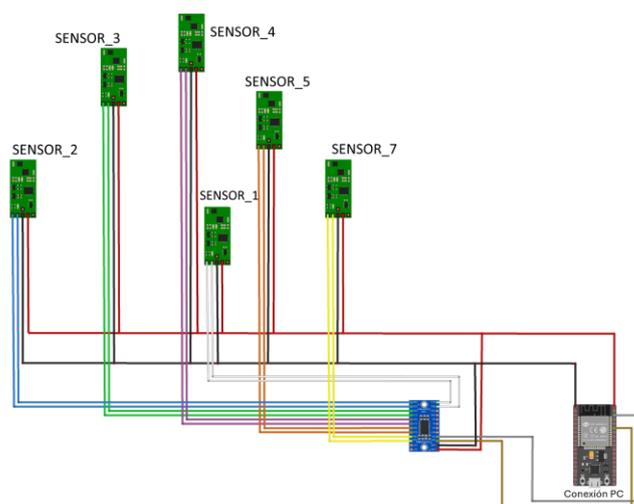


Figura 2: Esquema de conexión de los distintos elementos de comunicación entre sensores.

2.3. Registro y estructura de los datos

Los datos obtenidos de los sensores inerciales se envían en crudo a través de la placa ESP32 y se procesan mediante un programa desarrollado en Arduino. Este programa se basa en el uso de matrices MDC (Matrix of Directive Cosines) para representar la orientación de los sensores respecto a un sistema de referencia fijo. Emplea una matriz ortogonal de 3x3, donde cada elemento es el coseno del ángulo entre dos ejes específicos de los sistemas de referencia fijos y móviles. Así, se determina, directamente de las matrices MDC, la orientación de cada sensor en términos de *Yaw*, *Pitch* y *Roll*, que corresponden a las coordenadas de un sistema de tres ejes (X, Y, Z) en el espacio. Finalmente, el programa muestra el tipo de salida, que en este caso es analógica, y el número específico del sensor del que provienen los datos. Previamente a la toma de datos, los sensores inerciales se calibran para sincronizar los sistemas de referencia. Para ello, se realiza una calibración sobre una posición inicial fija, colocando la mano en posición horizontal sobre la mesa.

La salida obtenida (ángulos) se transmite desde la ESP32 a un programa de Python donde los datos se formatean para poder representar en pantalla la posición de cada IMU y para obtener una salida con la que se entrenará el modelo. Esta salida recoge, en formato JSON, los datos que incluyen el número del sensor y los valores de *Yaw*, *Pitch* y *Roll*.

3. Clasificación de gestos de la mano

3.1. Registro experimental

Se ha realizado una validación preliminar con un único usuario para evaluar la utilidad del guante en el reconocimiento de gestos. Para la primera prueba del modelo se han seleccionado tres gestos distintos realizados con la mano derecha: pulgar levantado - movimiento 1, dedos índice y corazón levantados (símbolo de victoria) - movimiento 2, y palma de la mano extendida hacia arriba - movimiento 3 (Figura 3, arriba).

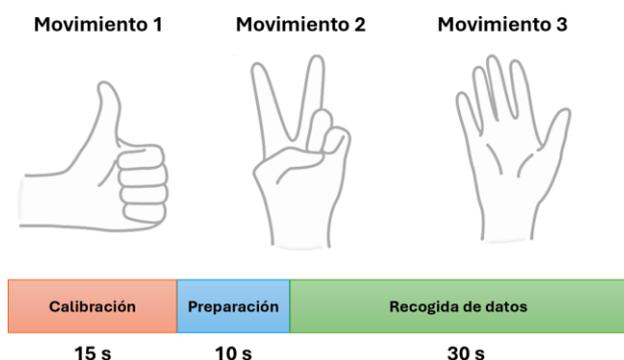


Figura 3: Gestos clasificados (arriba). Protocolo de registro (abajo).

Se ha creado una base de datos con muestras que incluyen un conjunto de 18 valores: el yaw, pitch y roll de los seis sensores medidos. Se han realizado dos tandas de registro para cada gesto siguiendo el protocolo mostrado en la Figura 3 (abajo). Antes de comenzar cada prueba, se coloca

la mano en posición horizontal sobre una superficie plana durante 15 segundos, permitiendo que los sensores se calibren. Tras esta etapa de calibración, se realiza el gesto. Los primeros 10 segundos (preparación) se descartan para asegurar que el usuario está ya realizando el gesto de forma estática. Pasado el período de preparación, se empiezan a recoger los datos necesarios para la creación de la base de datos durante 30 segundos para cada tanda de cada gesto.

3.2. Modelado y clasificación

Para la clasificación de los gestos se han empleado clasificadores convencionales de aprendizaje automático. En concreto, se han utilizado los siguientes métodos (Sheth et al., 2022):

- **K-NN (*k-Nearest Neighbor*):** es un método no paramétrico que estima la probabilidad de que un elemento x pertenezca a una clase C_j a partir de la información previamente etiquetada basándose en la cercanía del elemento clasificado a elementos del subconjunto próximos en el espacio de representación.
- **Random Forest:** es un clasificador no lineal que se basa en una combinación de árboles de predicción para determinar con gran precisión la pertenencia a determinada clase siempre que haya un número suficiente de datos para el modelado.
- **Naive Bayes:** es un clasificador probabilístico que se basa en el teorema de Bayes y asume que cada una de las características evaluadas contribuye de forma independiente a la probabilidad de que la muestra pertenezca a determinada clase.

Para cada uno de estos modelos, se ha entrenado un clasificador que tienen tres clases posibles, las cuales son cada uno de los movimientos de la Figura 3. De manera que como entrada del modelo se tienen los datos de la cinemática de la mano obtenidos mediante los sensores inerciales, y el modelo ofrece como salida a que clase/movimiento corresponden esos datos.

3.4. Métricas de evaluación

La clasificación se ha realizado utilizando una validación cruzada dividiendo los datos en 5 subconjuntos, y empleando 4 para el entramiento del modelo y el restante para la validación en cada una de las iteraciones. Adicionalmente se han evaluado varias métricas para estudiar el rendimiento del modelo de clasificación que incluyen:

- Exactitud (E): Mide la frecuencia promedio con la que el modelo realiza una predicción correcta.

$$E = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

- Precisión (P): Mide cuál es la probabilidad de que el modelo sea correcto cuando se hizo una predicción de que algo es cierto.

$$P = \frac{TP}{(TP + FP)}$$

- Sensibilidad (S): Mide la frecuencia con la que el modelo predijo con exactitud cuando algo era realmente cierto.

$$S = \frac{TP}{(TP + FN)}$$

siendo para todos los anteriores: TP (True Positive), TN (True Negative), FP (False Positive) y FN (False Negative).

- Puntuación F1: Mide la exactitud del modelo cuando las clases están desequilibradas, centrándose en la exactitud de las predicciones y los registros positivos reales.

$$F1 = \frac{2 * P * S}{(P + S)}$$

donde P y S, son la Precisión y la Sensibilidad respectivamente.

4. Resultados y discusión

La Tabla I muestra los resultados obtenidos en la validación inicial de clasificadores.

Tabla 1: Métricas de clasificación para cada uno de los modelos

Clasificador	E	P	S	F1
k-NN	0,96	1	1	1
RF	1	1	1	1
NB	1	1	1	1

Como se puede observar en la Tabla I los resultados de clasificación para únicamente 3 gestos son prácticamente perfectos. Esto es debido a que se han seleccionado tres gestos con morfologías muy diferentes y los datos recogidos se corresponden a posiciones estáticas de la mano. El hecho de que clasificadores no lineales como los implementados obtengan buenos resultados es también esperable dada la naturaleza heterogénea de los datos de entrada: ángulos muy dispares entre las distintas configuraciones. Una línea de trabajo de futuro crítica es conocer qué capacidad de escalabilidad tiene el sistema al aumentar la base de datos de gestos en relación con la exactitud de la clasificación. Una opción interesante es incorporar el alfabeto dactilológico y evaluar si con este dispositivo sería posible clasificar con fiabilidad una comunicación completa entre una persona que utilice el lenguaje de signos y otra persona que no lo conozca.

Por otra parte, este dispositivo puede mostrarse útil para la evaluación de la función motora de la mano. En este caso, se plantea evaluar la eficacia, no tanto de la detección de gestos estáticos, sino de la evaluación del movimiento de la mano durante terapia física. Esto plantea una dificultad adicional, por lo que una buena opción sería explorar algoritmos de regresión multidimensional o similares.

Por último, se plantea mejorar el dispositivo físico para que sea más ergonómico, estético y directamente vestible, eliminar la presencia de cables y optimizar la comunicación para que se pueda registrar en tiempo real la información de

los sensores inerciales, permitiendo así una clasificación en tiempo real de los gestos de la mano. Esto supondría un gran avance hacia la implementación práctica del guante en el entorno del día a día del usuario final.

5. Conclusiones

En este artículo se ha presentado un prototipo de guante vestible para la detección de gestos de la mano mediante el registro y análisis de la información obtenida de sensores inerciales situados en los dedos y la palma de la mano. Se han aplicado varios algoritmos de aprendizaje automático para clasificar tres gestos típicos de la mano con gran fiabilidad lo que demuestra la viabilidad del dispositivo como herramienta de comunicación y de evaluación del movimiento de la mano.

En el futuro se plantea escalar el dispositivo ampliando la biblioteca de gestos y plantear otras opciones como la decodificación proporcional del movimiento de la mano. Además, se pretende mejorar su ergonomía y aspecto e incorporar herramientas de clasificación en tiempo real que permitan su uso en entornos reales con el usuario final.

Agradecimientos

Este estudio ha sido desarrollado en el marco del proyecto MYOREHAB (PCI2023-143405), financiado por el Ministerio de Ciencia, Innovación y Universidades y la Agencia Estatal de Investigación MCIN/AEI/10.13039/501100011033.

Referencias

- Galván-Ruiz, J., Travieso-González, C. M., Tejera-Fetmilch, A., Pinan-Roescher, A., Esteban-Hernández, L. y Domínguez-Quintana, L., 2020. Perspective and Evolution of Gesture Recognition for Sign Language: A Review. *Sensors* 20(12), 3571. DOI: 10.3390/s20123571
- Jeon, H., Choi, H., Noh, D., Kim, T. y Lee, D., 2022. Wearable Inertial Sensor-Based Hand-Guiding Gestures Recognition Method Robust to Significant Changes in the Body-Alignment of Subject. *Mathematics* 10, 4753. DOI: 10.3390/math10244753
- Lin, B. S., Lee, I. J., Yang, S. Y., Lo, Y. C., Lee, J. y Chen, J. L., 2018. Design of an Inertial-Sensor-Based Data Glove for Hand Function Evaluation. *Sensors* 18(5), 1545. DOI: 10.3390/s18051545
- Oudah, M., Al-Naji, A. y Chahl, J., 2020. Hand Gesture Recognition Based on Computer Vision: A Review of Techniques. *Journal of Imaging* 6(8), 73. DOI: 10.3390/jimaging6080073
- Sheth, V., Tripathi, U. y Sharma, A., 2022. A Comparative Analysis of Machine Learning Algorithms for Classification Purpose. In *Proceedings of the 4th International Conference of Innovative Data Communication Technology and Application* 215, 422-431. DOI: 10.1016/j.procs.2022.12.044
- Tchantchane, R., Zhou, H., Zhang, S. y Alici, G., 2023. A Review of Hand Gesture Recognition Systems Based on Noninvasive Wearable Sensors. *Advanced Intelligent Systems* 5, 2300207. DOI: 10.1002/aisy.202300207