



Universidad Politécnica
de Madrid



**Escuela Técnica Superior de
Ingenieros Informáticos**

Máster en Ciencia de Datos

Trabajo Fin de Máster

**Minería de datos aplicada a sistemas de
captura del movimiento.**

Autor: Raúl Moldes Castillo.

Madrid, julio 2024

**Este Trabajo Fin de Máster se ha depositado en la ETSI
Informáticos de la Universidad Politécnica de Madrid.**

Trabajo Fin de Máster
*Máster en **Ciencia de Datos***

*Título: **Minería de datos aplicada a sistemas de captura del
movimiento.***

Julio, 2024.

*Autor: **Raúl Moldes Castillo.***

Tutor

Pablo Toharia Rabasco
ETSI Informáticos
Departamento de Arquitectura y
Tecnología de Sistemas
Informáticos
Universidad Politécnica de Madrid

Co-Tutor

Luca Piovano
CEDINT
Laboratorio de Realidad Virtual y
Visualización de Datos
Universidad Politécnica de
Madrid

Resumen

En los últimos años, la Realidad Virtual (RV) ha experimentado un crecimiento significativo como herramienta para ofrecer experiencias inmersivas. Su capacidad para simular entornos y situaciones que serían difíciles o peligrosos de recrear en el mundo real la convierte en una tecnología valiosa en campos como la educación, la medicina, el entretenimiento o la seguridad vial.

Para lograr un alto grado de realismo en la RV, es fundamental modelar con precisión el movimiento de los avatares que forman parte de los escenarios virtuales. La captura del movimiento (*Motion Capture* o *MoCap*) es una técnica esencial en este proceso. Consiste en monitorizar y registrar el movimiento de objetos físicos en el espacio tridimensional. Los datos obtenidos a través de sistemas de cámaras, como el sistema de sensores exteroceptivos *Optitrack*; o los sensores inerciales integrados en cascos de RV como son el *HTC Vive-Pro* o el *Oculus Rift*; se pueden utilizar para dicho propósito.

El Centro de Domótica Integral de la UPM (CEDINT), colabora con el Instituto Nacional del Automóvil (INSIA-UPM), en un proyecto centrado en predecir colisiones entre peatones y vehículos en escenarios de RV. Para ello, los investigadores del CEDINT han desarrollado una aplicación que introduce al usuario en situaciones de riesgo de colisión con un vehículo a motor. Esta aplicación permite registrar el movimiento de los usuarios que se introducen en las simulaciones utilizando un sistema de captura de movimiento.

En este Trabajo de Fin de Máster, a través de técnicas de minería de datos y aprendizaje no supervisado, se busca clasificar el movimiento de los usuarios en una serie de patrones. De esta forma se espera facilitar la mejora de la experiencia inmersiva gracias a la integración de estos patrones de movimiento en la programación de avatares virtuales. Adicionalmente, se pueden obtener conclusiones valiosas que permitan comprender como es el comportamiento de los peatones en situaciones de riesgo.

Abstract

In recent years, Virtual Reality (VR) has seen significant growth as a tool for providing immersive experiences. Its ability to simulate environments and situations that would be difficult or dangerous to recreate in the real world makes it valuable in fields such as education, medicine, entertainment, and road safety.

To achieve a high degree of realism in VR, accurately modeling the movement of avatars within virtual scenarios is essential. Motion capture (*MoCap*) is a crucial technique in this process. It involves monitoring and recording the movement of physical objects in three-dimensional space. Data obtained from camera systems, such as the Optitrack system with exteroceptive sensors, or inertial sensors integrated into VR headsets like the *HTC Vive Pro* or *Oculus Rift*, can be used for this purpose.

The UPM Center for Integral Home Automation (CEDINT) collaborates with the National Institute of Automotive Research (INSIA-UPM) on a project focused on predicting collisions between pedestrians and motor vehicles in VR scenarios. The researchers at CEDINT have developed an application that immerses users in collision risk situations with motor vehicles. This application records user movement during simulations using a motion capture system.

In this master's final project, data mining and unsupervised learning techniques are used to reduce user movement to a series of patterns. By integrating these movement patterns into virtual avatar programming, the goal is to enhance the immersive experience. Additionally, valuable insights can be gained into pedestrian behavior in risky situations.

Tabla de contenidos

1	Introducción.....	1
1.1	Antecedentes.....	1
1.2	Sistemas de captura de movimiento.	1
1.3	Objetivos del Trabajo.....	3
1.4	Estructura del Trabajo de Fin de Máster	4
2	Estado del Arte.	5
2.1	Validación de sistemas de captura del movimiento.....	5
2.2	Caracterización del movimiento humano.....	5
2.3	Identificación de patrones de movimiento.....	6
3	Metodología.....	7
3.1	Proceso seguido.....	7
3.2	Tecnologías y métodos empleados.	9
3.2.1	Obtención de los datos.....	9
3.2.1.1	Obtención de los datos recogidos por Optitrack.	11
3.2.1.2	Obtención de los datos recogidos por la aplicación de realidad virtual.	12
3.2.2	Exploración de los datos.	13
3.2.2.1	Diferencias en el sistema de referencia espacial.	13
3.2.2.2	Identificación de zonas donde la detección es peor.....	14
3.2.3	Preparación de los datos.	16
3.2.4	Minería de datos.	17
3.2.4.1	Análisis del movimiento.	17
3.2.4.2	Evaluación de la corrección biomecánica de los datos.....	22
3.2.4.3	Extracción de características.	24
3.2.4.4	Identificación de patrones del movimiento	30
4	Resultados y discusión.	32
4.1	Evaluación de la fiabilidad y la corrección de las medidas.....	32
4.1.1	Análisis de la visibilidad que tienen las cámaras de cada Rigid Body.	32
4.1.2	Relación de la coherencia biomecánica con el nivel de visibilidad de cada cruce.....	35
4.1.2.1	Resultados obtenidos empleando LCSS.....	36
4.1.2.2	Resultados obtenidos para las medidas interpoladas empleando DTW.	39
4.1.3	Conclusiones de la evaluación de las medidas y del algoritmo de interpolación.....	40
4.2	Generación de una base de datos de características.....	41
4.3	Identificación de patrones de movimiento.....	42
4.3.1	Reducción de la dimensionalidad.....	42

4.3.2	Selección del número de clústeres.	43
4.3.3	Resultados del agrupamiento.....	46
4.3.3.1	Parametrización de los grupos.	49
4.3.3.2	Análisis de los resultados obtenidos en función del escenario de Realidad Virtual.	50
4.3.3.3	Análisis de los resultados obtenidos en función de la ocurrencia o no de un accidente.	51
5	Conclusiones, trabajos futuros y limitaciones.	52
5.1	Evaluación de la calidad de los datos.	52
5.2	Generación de una base de datos de características y parametrización del movimiento.	53
5.3	Líneas futuras.....	53
5.4	Limitaciones.....	54
5.4.1	Limitaciones del sistema de recogida de datos.	54
5.4.2	Limitaciones del desarrollo.	54
6	Anexo 1. Descripción de los conjuntos de datos de partida.	55
6.1	Archivos de log de Optitrack.....	55
6.2	Archivos de log de la Aplicación.....	55
7	Bibliografía	57

1 Introducción

1.1 Antecedentes.

La alta incidencia de fallecimientos de peatones en accidentes de tráfico [1], [2] hace que el comportamiento de este grupo poblacional sea un factor de gran interés a la hora de mejorar la seguridad vial en las ciudades, además de ser un elemento clave para el desarrollo de Sistemas Avanzados de Asistencia al Conductor (ADAS, de su acrónimo en inglés *Advanced Driver Assistance Systems*).

En los últimos años, la Realidad Virtual (RV) se ha convertido en una herramienta de gran utilidad para ofrecer experiencias inmersivas que serían imposibles de reproducir en el mundo real. El uso de técnicas de Realidad Virtual para la simulación de impactos entre peatones y vehículos ha demostrado ofrecer una ventaja significativa a la hora de reproducir las condiciones de la colisión en un entorno controlado y seguro [2]. De esta forma, se puede recoger información acerca del comportamiento de los peatones en los instantes previos a un accidente. Esto permite entender las condiciones del impacto, minimizando tanto su incidencia como sus consecuencias.

El Centro de Domótica Integral de la Universidad Politécnica de Madrid (CEDINT-UPM), colabora con el Instituto Nacional del Automóvil (INSIA-UPM) para el desarrollo de un sistema capaz de predecir potenciales colisiones entre peatones y vehículos en escenarios de Realidad Virtual [3].

Esta línea de investigación arrancó con el proyecto OPREVU (*Uso de técnicas de Realidad Virtual para Optimizar sistemas de identificación de Usuarios Vulnerables mediante la Predicción de sus reacciones en la fase previa del atropello* RTI2018-096617-B-I00). Este proyecto ha sido financiado por el Ministerio de Ciencia, Innovación y Universidades y por el fondo europeo FEDER, en la Convocatoria 2018 de Proyectos I+D+i “*Retos Investigación*” del Programa Estatal de I+D+i orientada a los Retos de la Sociedad. Su objetivo último es la identificación de Usuarios Vulnerables en Carretera (UVC) a partir de la caracterización de las reacciones de los peatones en la fase previa a un accidente. Para ello, el CEDINT desarrolló una aplicación que introduce al usuario en una serie de escenarios virtuales donde se simula una colisión con un vehículo a motor [3].

A partir de la información recogida en las simulaciones, se obtienen variables que modelan las reacciones del usuario, y se entrenan modelos predictivos para identificar a los grupos más vulnerables [3].

Partiendo de un conjunto de datos recogido por un sistema de captura de movimientos, el presente Trabajo de Fin de Máster tiene el objetivo de identificar patrones en el movimiento de los peatones en situaciones de riesgo. Para ello, se aplican técnicas de minería de datos y aprendizaje no supervisado para identificar patrones habituales y tendencias.

1.2 Sistemas de captura de movimiento.

La Captura del Movimiento (*Motion Capture* o *MoCap*) es una técnica que tiene como objetivo monitorizar el movimiento de objetos físicos en el espacio tridimensional. Su aplicabilidad abarca una amplia gama de campos, como la animación, la Realidad Virtual, la biomecánica y el análisis motor. Dentro de la amplia variedad de sistemas de Captura del Movimiento que existen, se distinguen dos grupos principales: los sistemas basados en sensores interoceptivos y los sistemas basados en sensores exteroceptivos.

Los **sensores interoceptivos** monitorizan el estado interno del objeto a partir de su aceleración angular y posicional. Posteriormente realizan una integración sobre los valores estimados para obtener la posición y la orientación. La principal desventaja de los sensores que se incluyen en esta categoría es que introducen un error de derivación, el cual se propaga con el tiempo. Por este motivo, las medidas recogidas por sensores interoceptivos son menos fiables a medida que avanza el proceso de recogida de datos [4].

Alternativamente, los **sensores exteroceptivos** estiman la posición y orientación del objeto con respecto a un marco de referencia externo. Generalmente ofrecen una estimación más exacta que los sensores interoceptivos. Un ejemplo de este tipo de sensores es el sistema de cámaras *Optitrack* [5], que emplea un montaje de cámaras 2D/3D. Estas cámaras son capaces de detectar balizas compuestas por conjuntos de marcadores LED. Su principal problema es su limitada portabilidad, además de estar sujetos a posibles oclusiones e interferencias del propio entorno que pueden impedir la localización de un sensor por parte de las cámaras, introduciendo valores nulos [6], [7].

El Centro de Domótica Integral de la Universidad Politécnica de Madrid cuenta con un montaje *Optitrack* capaz de detectar la posición y orientación de un conjunto de cinco balizas (*pucks*) activas que contienen ocho marcadores LED cada una. Se trata de un sistema *outside-looking-in*, que consta de 12 cámaras *Slim13E* activas colocadas en el techo, cubriendo un área de 7 x 5 m (Ver Ilustración 1). Las cámaras se conectan a través de un *Switch NETGEAR* a una estación de trabajo que recibe los datos. Este sistema en su conjunto permite recoger la posición y orientación de las manos, los pies y la pelvis de los usuarios de aplicaciones de Realidad Virtual.



Ilustración 1. Disposición de las cámaras Optitrack alrededor de la sala de ensayos.

De manera adicional, el CEDINT utiliza los sensores inerciales integrados en diferentes cascos de realidad virtual (*Head Mounted Displays* o *HMD*). Algunos ejemplos son el *Oculus Rift* [8] o el *HTC Vive Pro* [9]. Este último emplea un conector inalámbrico (*HTC Wireless Adapter*) para recoger los datos posicionales

sin interferir en el movimiento de los individuos. El *HTC Vive Pro* cuenta con una pantalla DUAL AMOLED que ofrece un campo de visión de 110 grados y una resolución de 2880 x 1600 píxeles. Posee sensores *Steam VR Tracking*, giróscopos, un sensor de proximidad y un sensor IPD (Detector de movimiento por infrarrojos o *Infrared Motion Detector*) [9].

En este Trabajo de Fin de Máster he empleado tanto *Optitrack* como el *HTC Vive Pro* como fuentes de datos.

1.3 Objetivos del Trabajo.

El objetivo principal del presente Trabajo de Fin de Máster es identificar y parametrizar patrones en el movimiento de los usuarios de una aplicación de Realidad Virtual enfocada en la seguridad vial. En el proyecto mencionado en el apartado anterior, se emplea únicamente la posición y orientación de la cabeza de los sujetos para caracterizar las reacciones. Este Trabajo pretende ir un paso más allá, identificando patrones y tendencias en el movimiento humano a partir de las extremidades y la pelvis.

Los datos se obtienen a partir del sistema de cámaras *Optitrack* que posee el CEDINT para registrar el movimiento de los usuarios de sus aplicaciones de Realidad Virtual. Por las características inherentes a los sistemas exteroceptivos, las cámaras pueden sufrir oclusiones que resulten en datos faltantes. El software de gestión de *Optitrack* (que recibe el nombre de *Motive*) tiene la capacidad de triangular la posición de un sensor siempre que las cámaras detecten dos o más de sus marcadores. Para obtener la orientación, al menos tres marcadores son necesarios. De acuerdo con las recomendaciones de *Optitrack*, a partir de cuatro marcadores se obtienen valores de posición y orientación óptimos.

Con el propósito de mejorar la estimación que ofrece *Motive*, así como para obtener valores de posición y orientación incluso cuando las cámaras no son capaces de detectar los elementos, los investigadores del CEDINT desarrollaron un algoritmo de interpolación, el cual integra en sus aplicaciones de Realidad Virtual. Durante la recogida de datos, las cámaras emiten los valores estimados de posición y orientación de cada uno de los marcadores que componen un sensor. Todo ello se envía a la estación base y a las aplicaciones de Realidad Virtual que el CEDINT desarrolla, donde el algoritmo de interpolación actúa para obtener un valor definitivo.

La fiabilidad de este algoritmo, así como del sistema de recogida de datos en su conjunto, no ha sido verificada hasta la fecha. Para garantizar la consecución del objetivo establecido, es necesario previamente analizar el grado de corrección y coherencia de las medidas de posición y orientación. De esta forma se garantiza que estas medidas sean válidas para lograr el objetivo principal, evitando que una fuente de datos poco fiable afecte a las conclusiones.

Por tanto, para lograr el objetivo principal, se definen los siguientes objetivos técnicos:

- **Objetivo Técnico 1 (OT1):** Evaluar la fiabilidad del sistema de captura de datos, tanto a partir de los parámetros que el propio sistema ofrece (como el número de marcadores detectados para cada sensor), como empleando otras métricas basadas en restricciones biomecánicas de la marcha humana.
- **Objetivo Técnico 2 (OT2):** Extracción de características a partir de los datos del movimiento de los usuarios. Generación de una base de datos de características.

- **Objetivo Técnico 3 (OT3):** Identificación de patrones de movimiento similares a partir de las características obtenidas. En primer lugar, se aplican técnicas de análisis multivariado y reducción de la dimensionalidad para obtener las variables más importantes en la identificación de patrones. Posteriormente se selecciona el número óptimo de grupos en función de las características de los datos. Finalmente se utiliza un algoritmo de aprendizaje no supervisado para agrupar las trayectorias en clústeres similares.

1.4 Estructura del Trabajo de Fin de Máster

La estructura del documento en los siguientes capítulos sigue la estructura definida a continuación:

- En el **Capítulo 2** se expone el estado del arte actual. Se incluye una revisión de otros trabajos en los que se han utilizado sistemas de captura del movimiento para parametrizar y caracterizar la marcha humana.
- En el **Capítulo 3** se desarrolla la metodología llevada a cabo. En primer lugar, se describe el proceso seguido en el desarrollo del trabajo, para después desarrollar brevemente las tecnologías y técnicas empleadas en cada fase.
- En el **Capítulo 4** se describen los resultados obtenidos, además de elaborar una breve discusión. La estructura de este capítulo pretende desarrollar como se han ido consiguiendo los objetivos planteados.
- En el **Capítulo 5** se detallan las conclusiones del trabajo, y se proponen líneas futuras para continuar con la investigación.

2 Estado del Arte.

2.1 Validación de sistemas de captura del movimiento.

En la literatura, existen diferentes trabajos centrados en la validación de sistemas de captura de movimiento. En [6] los autores ofrecen una revisión sistemática de las diferentes tecnologías y sistemas de captura del movimiento humano que existen en la actualidad, además de analizar su aplicabilidad en campos como la biomecánica, el deporte, la Realidad Virtual y la animación 3D. De manera adicional proporcionan un breve resumen de las propiedades y limitaciones de cada uno de estos sistemas. En cuanto a los sistemas ópticos (como es *Optitrack*), las principales limitaciones que los autores destacan son su elevado costo, y su difícil portabilidad. De manera adicional, estos sistemas son sensibles a posibles oclusiones que pueden producirse durante la captura de datos. La adecuada colocación de las balizas, así como la disposición del montaje requiere de tiempo y de conocimiento experto para realizarse correctamente.

Otros trabajos se centran en la valoración del desempeño de un sistema de captura del movimiento concreto. En [4] los autores sugieren una alternativa basada en redes neuronales para corregir el error de derivación característico de los *Head Mounted Displays (HMDs)*. Este problema se aborda realizando una predicción de la posición y orientación de la cabeza del usuario, lo que compensa el error de derivación y mejora la experiencia del sujeto en el entorno virtual. Por otra parte, en [10] se evalúa la efectividad del *Oculus Rift* frente a un sistema de sensores exteroceptivos basado en cámaras *Optitrack*. Los autores consideran los datos posicionales recogidos por *Optitrack* como referencia, y evalúan la precisión del *Oculus Rift* frente a los recogidos por las cámaras. Se obtiene un error de traslación de 1.83cm y un error de orientación de 0.77°.

En lo que respecta exclusivamente a los sistemas *Optitrack*, en [7] se realiza una revisión sistemática sobre sus aplicaciones en el análisis biomecánico. Entre los artículos que se incluyen en dicha revisión, cabe destacar [11], donde se propone un sistema de captura del movimiento basado en 12 cámaras *Optitrack* similar al que posee el CEDINT, combinado con un sensor magnético. Los autores justifican su uso en aplicaciones de Realidad Virtual utilizadas en procesos de rehabilitación de la marcha humana.

En el presente Trabajo de Fin de Máster pretendo evaluar los datos obtenidos por el sistema *Optitrack* utilizando características biomecánicas de la marcha para verificar la validez de los datos y detectar posibles anomalías.

2.2 Caracterización del movimiento humano.

En cuanto a la caracterización del comportamiento de los peatones, en [12] se exploran los principales patrones que existen a la hora de cruzar intersecciones señalizadas. Los autores identifican los principales factores que afectan al comportamiento de los individuos, así como la posibilidad de interaccionar con vehículos bajo ciertas condiciones de tráfico.

En el proyecto que precede a mi Trabajo [3], se utilizó la posición de la cabeza como la referencia de la posición del peatón. A partir de la posición se calcula la velocidad del peatón; y a partir de la orientación de la cabeza con respecto al eje vertical se estima el nivel atencional. Los autores definen un parámetro llamado el *Mínimo Ángulo Atencional (MAA)*, calculado a partir del arco-tangente entre la distancia del peatón y del coche al punto donde se produciría el impacto en el momento que se inicia la frenada.

Posteriormente, se emplearon algoritmos de aprendizaje automático (vecinos cercanos, árbol de decisión, máquinas de soporte vectorial y *random forest*); para tratar de predecir la ocurrencia o no de un accidente según las diferentes variables. Algunos de los resultados de [3] indican que aquellos peatones que reducen su velocidad tienen una menor probabilidad de sufrir un accidente que aquellos que aceleran o no reaccionan.

En este Trabajo de Fin de Máster pretendo ir un paso más allá en la caracterización del movimiento humano. Mientras que en [3] se utiliza únicamente la posición y orientación de la cabeza, en este Trabajo empleo también la posición y orientación de la cintura, las manos y los pies de los sujetos con el objetivo de elaborar un análisis más profundo.

2.3 Identificación de patrones de movimiento.

Otros proyectos en los que se han utilizado técnicas no supervisadas para reducir la complejidad del movimiento humano a una serie de patrones incluyen [13], donde se emplea un algoritmo de agrupamiento para detectar acciones anómalas en un vídeo. Fuera del campo de la visión por computador, en [14] los autores proponen el Análisis de Clústeres Alineados (ACA, de sus siglas en inglés *Aligned Cluster Analysis*) y el Análisis Jerárquico de Clústeres Alineados (HACA, o *Hierarchical Aligned Cluster Analysis*), algoritmos que desarrollan como una extensión del método *Kmeans* y el clustering espectral aplicados a series temporales. De esta forma elaboran una descomposición jerárquica de la marcha humana.

En mi Trabajo de Fin de Máster utilizo las características extraídas para posteriormente aplicar algoritmos de agrupamiento con el objetivo de identificar patrones y tendencias en el movimiento humano.

3 Metodología

3.1 Proceso seguido.

En este apartado se describe el proceso seguido a lo largo del desarrollo del Trabajo, desde la recogida de los datos hasta la obtención de los resultados finales. Siguiendo un orden cronológico, se distinguen las siguientes fases o etapas:

- **Fase 0:** Entendimiento y familiarización con el problema inicial.
- **Fase 1:** Obtención de los datos.
- **Fase 2:** Exploración de los datos para valorar la calidad general, e identificar posibles anomalías.
- **Fase 3:** Preparación de los datos para la extracción de conocimiento posterior.
- **Fase 4:** Minería de datos. Una vez los datos han sido explorados y tratados, en esta fase se extrae el valor contenido en ellos para generar nuevos conocimientos.
- **Fase 5:** Presentación de los resultados.

Para lograr los objetivos técnicos definidos en el apartado 1.3, se requiere de cumplir con los paquetes de tareas definidos en la Tabla 1.

Objetivos Técnicos	Paquete de Tareas	Descripción
OT1, OT2, OT3	PT1	Exportación de los datos recogidos por las cámaras Optitrack a formato .csv.
	PT2	Exportación de los archivos de log de la aplicación (que contienen los datos interpolados) a formato .csv
OT1, OT2, OT3	PT3	Exploración de los datos para identificar los requisitos de cara al tratamiento posterior.
OT1, OT2, OT3	PT4	Preparación de los datos en función de las necesidades que se identifican en PT3.
OT1	PT5	Calcular el grado de corrección de las medidas a partir de métricas basadas en restricciones biomecánicas.
	PT6	Evaluar el grado de visibilidad de las cámaras para cada elemento.
	PT7	Evaluar la coherencia biomecánica de los datos en función del grado de visibilidad de las cámaras utilizando los datos originales.
	PT8	Evaluar la coherencia biomecánica de los datos en función del grado de visibilidad de las cámaras utilizando los datos interpolados.
	PT9	Comparar el grado de coherencia antes y después de aplicar la interpolación.
OT2	PT10	Uso de técnicas de ingeniería de características para obtener las variables que caracterizan el movimiento humano.
	PT11	Generación de una base de datos de características del movimiento.
OT3	PT12	Aplicar técnicas de reducción de dimensionalidad para obtener las variables que más explican la varianza del conjunto de datos.
	PT13	Selección del número óptimo de perfiles.
	PT14	Aplicar un algoritmo de aprendizaje no supervisado para identificar patrones en el movimiento de los usuarios.
	PT15	Identificar los parámetros que definen cada perfil mediante tests estadísticos.

Tabla 1. Descripción de los paquetes de tareas.

La asignación de los paquetes de tareas a cada fase del desarrollo del Trabajo de Fin de Máster se expone en la Tabla 2.

Fase	Paquetes de Tareas
Fase 1	PT1 y PT2
Fase 2	PT3
Fase 3	PT4
Fase 4	PT5, PT10, P12, PT13, PT14
Fase 5	PT6, PT7, PT8, PT9, PT11, PT15

Tabla 2. Asignación de los paquetes de tareas a cada una de las fases del Trabajo.

El diagrama de GANTT de la Ilustración 2 muestra la duración de las diferentes fases del proceso de forma simplificada, desde su inicio en noviembre de 2023 hasta su finalización en julio de 2024.

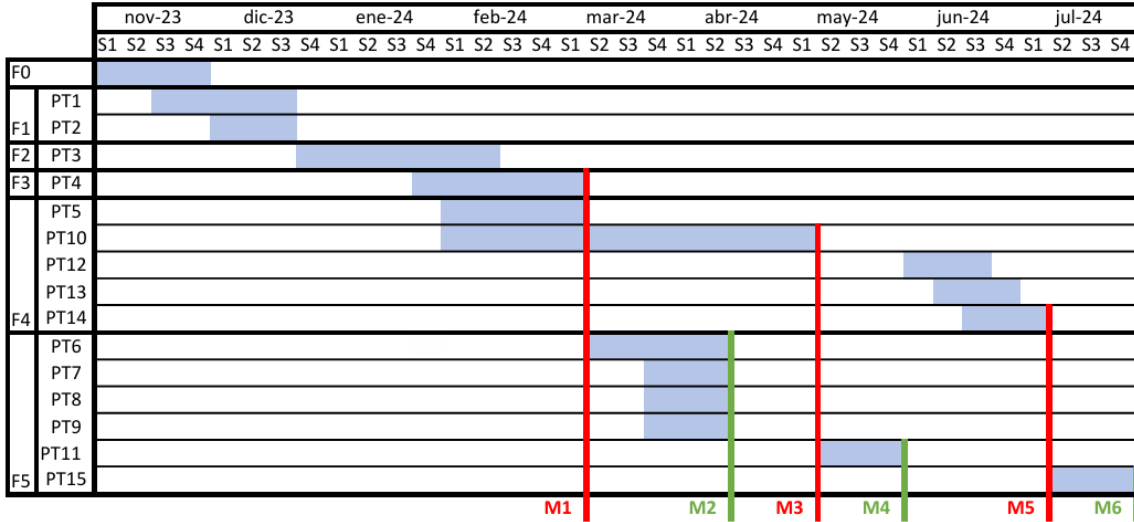


Ilustración 2. Diagrama de GANTT del Trabajo de Fin de Máster.

Las metas marcadas en rojo (M1, M3 y M5) marcan el inicio de la fase de presentación de resultados de cada objetivo técnico. Por otra parte, las metas marcadas en verde (M2, M4 y M6) marcan la consecución de cada objetivo técnico. La descripción de cada una de las metas se muestra en la Tabla 3.

Meta	Descripción
M1	Inicio de la fase de presentación de resultados de OT1.
M2	Consecución de OT1.
M3	Inicio de la fase de presentación de resultados de OT2.
M4	Consecución de OT2.
M5	Inicio de la fase de presentación de resultados de OT3.
M6	Consecución de OT3.

Tabla 3. Descripción de cada una de las metas establecidas.

Dado que los objetivos técnicos están interrelacionados, no se puede iniciar la fase de minería de datos del OT3 antes de que se haya completado la consecución del OT2. De la misma manera, la fase de presentación de resultados del OT2 no puede iniciarse hasta que se haya verificado la fiabilidad del sistema de recogida de datos. En el diagrama de PERT de la Ilustración 3 se muestran las dependencias entre los Paquetes de Tareas que constituyen el Trabajo de Fin de Máster.

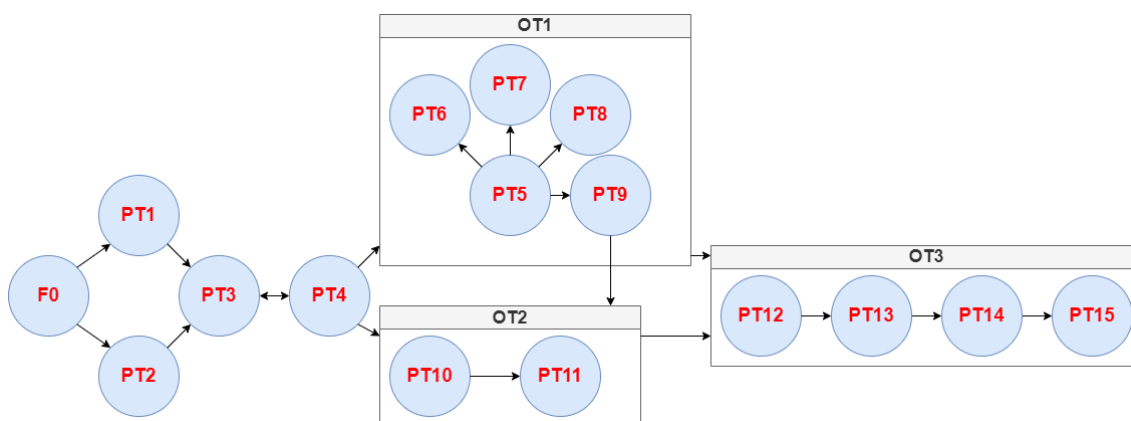


Ilustración 3. Diagrama de PERT del Trabajo de Fin de Máster.

Los paquetes de tareas 1, 2, 3, y 4 son necesarios para conseguir cualquiera de los Objetivos Técnicos planteados. La obtención de características del movimiento (PT10) puede realizarse de forma paralela a la consecución del Objetivo Técnico 1, aunque la base de datos de características no es definitiva (PT11) hasta que el OT1 se ha completado. Para realizar cualquier tarea relativa al OT3 es necesario haber cumplimentado previamente los Objetivos Técnicos 1 y 2.

3.2 Tecnologías y métodos empleados.

En este apartado se describen brevemente tanto las diferentes tecnologías utilizadas lo largo de cada fase del Trabajo como los métodos y técnicas empleadas para llevar a cabo cada una de estas fases. La Fase 0 no se incluye, ya que únicamente supone el entendimiento del problema a resolver y el establecimiento de los objetivos, los cuales se han desarrollado en el Capítulo 1.

3.2.1 Obtención de los datos.

En esta fase se extraen los datos de partida para su posterior análisis y procesado. Adicionalmente, se realiza una exploración inicial para comprender su estructura y características.

Los datos se recogieron durante una serie de ensayos realizados con 19 usuarios voluntarios siguiendo el protocolo y metodología descritos en [3]. Al disponer el CEDINT de un sistema de captura de movimiento *Optitrack*, se utilizaron las cámaras y sensores correspondientes para registrar la posición de la pelvis y las extremidades de los individuos que participaron en las simulaciones. También se utilizó el casco *HTC Vive Pro* para monitorizar la posición de la cabeza de los sujetos.

Cada uno de los participantes experimentó 4 escenarios virtuales diferentes en los que cruzaba un paso de cebra regulado por semáforo y donde se forzaba una situación de riesgo personal (colisión con un vehículo).

Los mencionados escenarios reproducen escenas de atropello reales que tuvieron lugar en la ciudad de Madrid, en concreto en la Avenida de los Toreros, la Avenida de Emilio Muñoz, la Avenida de los Hermanos García Noblejas, y la Avenida de Machupichu [3]. Antes de iniciar las simulaciones, los usuarios pasaron por un escenario de familiarización diseñado con un doble objetivo: acostumbrarse a moverse en un entorno de Realidad Virtual y medir la velocidad de marcha de los usuarios. El escenario de la Avenida de los Toreros se realiza a continuación de cumplimentar la simulación en el escenario de familiarización. El objetivo de este escenario es preparar al usuario para que mantenga un elevado nivel atencional de cara a la posible aparición de un vehículo en los ensayos que se realizarán a continuación. Es por eso por lo que el vehículo en este escenario aparece por el primer carril (Ver Ilustración 4).



Ilustración 4. Vista en perspectiva del escenario ambientado en la Avenida de los Toreros.

Las características de cada escenario aparecen de forma resumida en la Tabla 4. La densidad del tráfico hace referencia a la cantidad de interferencias presentes a lo largo del paso de peatones (ya sean otros peatones o vehículos virtuales). La visibilidad del cruce representa la capacidad que tiene el usuario para visibilizar la zona donde se produce la colisión en el momento que el vehículo empieza a aproximarse. En el escenario de la Avenida de Machupichu, esta visibilidad está limitada por la presencia de un autobús en el primer carril. En la Ilustración 5 se muestra una vista en perspectiva de este escenario.

Escenario	Lado de aparición del vehículo	Carril aparición vehículo	de del	Visibilidad del cruce	Densidad del tráfico
Av. de los Toreros	Izquierda del peatón.	Primer carril.		Buena	Normal
Av. Machupichu	Izquierda del peatón	Segundo carril.		Mala	Normal
Av. Hnos. G-Noblejas	Izquierda del peatón	Tercer carril		Buena	Baja
Av. E. Muñoz	Derecha del peatón	Tercer carril		Normal	Alta

Tabla 4. Descripción de cada uno de los escenarios donde se simula un atropello [3].



Ilustración 5. Vista en perspectiva del escenario ambientado en la Avenida de Machupichu.

Previo al arranque de cada sesión de ensayos, se calibraba el área de juego necesaria para la simulación, garantizando que no hubiese ningún obstáculo para los usuarios y a la vez asegurando el adecuado funcionamiento de cada cámara y sensor empleado.

Durante estos ensayos, las cámaras Optitrack y el HTC Vive Pro recogían la posición y orientación de cada una de las partes del cuerpo a monitorizar. Posteriormente, un sistema integrado en la aplicación recibía estos datos y realizaba una interpolación sobre los mismos en caso de que la medida obtenida por las cámaras indicara una detección poco fiable. Por tanto, las fuentes de datos con las que he contado en este Trabajo de Fin de Máster son las siguientes:

- El **log de eventos de la aplicación** de realidad virtual desarrollada en el CedInt y empleada en OPREVU [3].
- El **log de eventos de las cámaras Optitrack** del CEDINT.

En las siguientes secciones (3.2.1.1 y 3.2.1.2) desarrollo las características específicas de los datos de partida según cada fuente.

3.2.1.1 Obtención de los datos recogidos por Optitrack.

Los datos recogidos por el sistema de cámaras *Optitrack* se obtienen a partir de *Motive* [5], un software que se utiliza para calibrar y configurar los parámetros en la recogida de datos.

Las cámaras recogen, con una frecuencia de muestreo de 60Hz, la posición y orientación de cada uno de los *Rigid Bodies* que se le indiquen en la fase de calibración. Un *Rigid Body* es un sensor compuesto por ocho marcadores LED. Para cada fotograma, las cámaras *Optitrack* son capaces de registrar la posición de cada uno de los marcadores que componen un determinado *Rigid Body* mediante la reconstrucción tridimensional de la imagen 2D que reciben. Los *Rigid Bodies* a monitorizar en este caso se colocan en la pelvis, las manos y los pies del usuario.

Si en un fotograma el sistema de cámaras no puede posicionar un marcador determinado debido a oclusiones u otros factores del entorno, reconstruye su

localización asumiendo que todos los marcadores del *Rigid Body* están a la misma distancia de su centroide. El mínimo de marcadores necesarios para poder reconstruir la posición del *Rigid Body* se estableció en dos durante la fase de calibración, de manera que cuando *Optitrack* detecta menos de dos marcadores para un determinado sensor no devuelve ningún valor.

Utilizando el software de *Motive*, se exportan cada una de las tomas a formato CSV para su posterior análisis y procesamiento. Cada uno de estos archivos contiene las medidas de posición y orientación recogidas para cada *Rigid Body* tal y como las obtiene el sistema de cámaras. Estos archivos incluyen también una columna *Time* que hace referencia al instante en que se emitió la medida.

3.2.1.2 Obtención de los datos recogidos por la aplicación de realidad virtual.

La aplicación de Realidad Virtual que se utiliza para recoger la información de las reacciones de los usuarios frente a posibles atropellos fue desarrollada en el CEDINT empleando el motor de desarrollo de videojuegos Unity3D [15]. La frecuencia de muestreo en este caso es de 10Hz.

Dicha aplicación se comunica tanto con las cámaras *Optitrack* como con el casco *HTC Vive Pro* a través de sus respectivas APIs. La aplicación utiliza estos datos para animar en tiempo real al avatar que representa al usuario, para que su posicionamiento espaciotemporal quede reflejado con la máxima aproximación posible.

Antes de utilizar estos datos, se realiza una comprobación para verificar que son coherentes. En concreto, se comprueba el número total de marcadores detectados para cada sensor, la distancia y la diferencia de orientación con respecto a la última posición registrada.

En función de esta información, se aplica una interpolación sobre los datos en crudo para obtener un valor final. El algoritmo de interpolación define cuatro niveles de interpolación: *NONE* (0), *POSITION* (1), *ORIENTATION_LAST* (2), *ORIENTATION_INIT* (3) y *BOTH_INIT* (4); los cuales se describen en la Tabla 5. En esta tabla, un estado se define mediante una posición y una orientación de un sensor. Se considera un estado como válido si la distancia al estado previo es inferior a 0.2m. La lógica empleada por el algoritmo se describe en el diagrama de flujo de la Ilustración 6.

NIVEL	BOTH INIT	POSITION	ORIENTATION LAST	ORIENTATION INIT	NONE
Interpolar posición	Calcula posición a partir de la posición establecida en la calibración.	Calcula posición a partir de la posición establecida en la calibración.	Un marcador detectado. Se calcula la posición a partir de la posición del marcador. Dos o más marcadores detectados. Calcula la posición local a partir del centroide de los marcadores.		No aplica.
Interpolar orientación	Estado previo válido. Calcula la orientación a partir del estado previo. Estado previo no válido. Calcula orientación a partir de la orientación establecida en la calibración.		Calcula la orientación a partir del estado previo.	Calcula orientación a partir de la orientación establecida en la calibración.	No aplica.

Tabla 5. Descripción de cada nivel de interpolación.

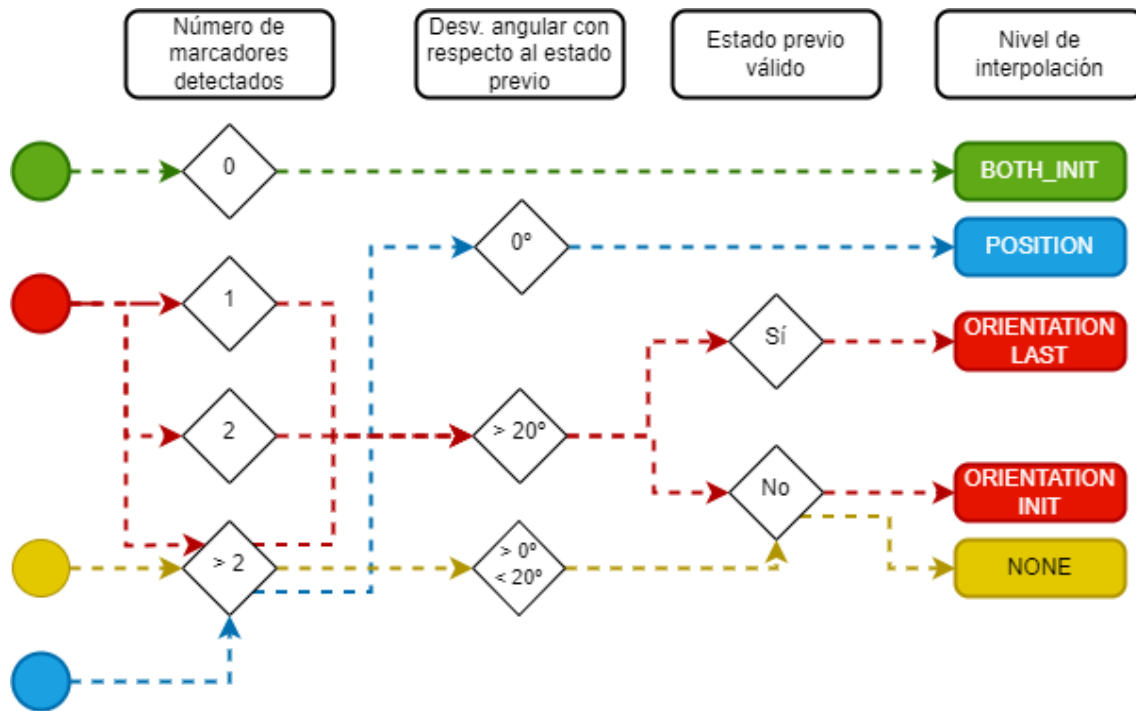


Ilustración 6. Diagrama de flujo con la lógica del algoritmo de interpolación.

Las medidas, modificadas o no, de cada sensor se escriben dinámicamente en los archivos de *Log* de la aplicación. Estos archivos contienen las posición y orientación de cada uno de los *Rigid Bodies* (sensores) después de aplicar el algoritmo de interpolación. Además, para cada medida, se registra el número total de marcadores detectados y el nivel de interpolación aplicado. Los archivos incluyen una columna con el *timestamp*, o instante temporal en el que se registró la medida, en formato UNIX. En los *Logs* se escriben también los eventos que tienen lugar durante la simulación. En el Anexo 1 se incluye una tabla con la descripción de estos eventos

3.2.2 Exploración de los datos.

En esta fase se realiza un análisis de las características de los datos para identificar patrones, tendencias y valores atípicos, con el objetivo de identificar necesidades de cara al posterior procesamiento o preparación de los datos. La descripción tanto de la semántica como del tipo de dato de cada variable de los conjuntos de datos de partida se incluye en el Anexo 1.

Tanto en esta etapa como en la siguiente la principal herramienta que he utilizado para explorar, analizar y realizar transformaciones sobre los datos ha sido *Pandas* (2.2.2) [16], una librería de *Python* centrada en la manipulación de datos. Para generar las figuras que se muestran a lo largo de todo el proyecto, he utilizado las librerías *Matplotlib* [17] y *Seaborn* [18]. A lo largo de todo el Trabajo de Fin de Máster, la versión de *Python* empleada ha sido la 3.11.8 [19].

A continuación, se exponen cada una de las necesidades que se identificaron en esta fase y que se trataron en la fase de preparación:

3.2.2.1 Diferencias en el sistema de referencia espacial.

Inicialmente, las medidas recogidas por *Optitrack* y la aplicación no comparten el mismo sistema de referencia. Esto se debe a que la situación del centro de

coordenadas en cada escenario de la aplicación de realidad virtual es diferente, mientras que *Optitrack* lo sitúa en el centro de la sala de ensayo.

De manera adicional, *Optitrack* emplea un sistema de coordenadas diestro, mientras que Unity emplea un sistema zurdo, de manera que los ejes x y z tienen el sentido opuesto. La situación y orientación de los ejes de coordenadas en cada escenario virtual con respecto a la sala de ensayo se muestra en la Ilustración 12.

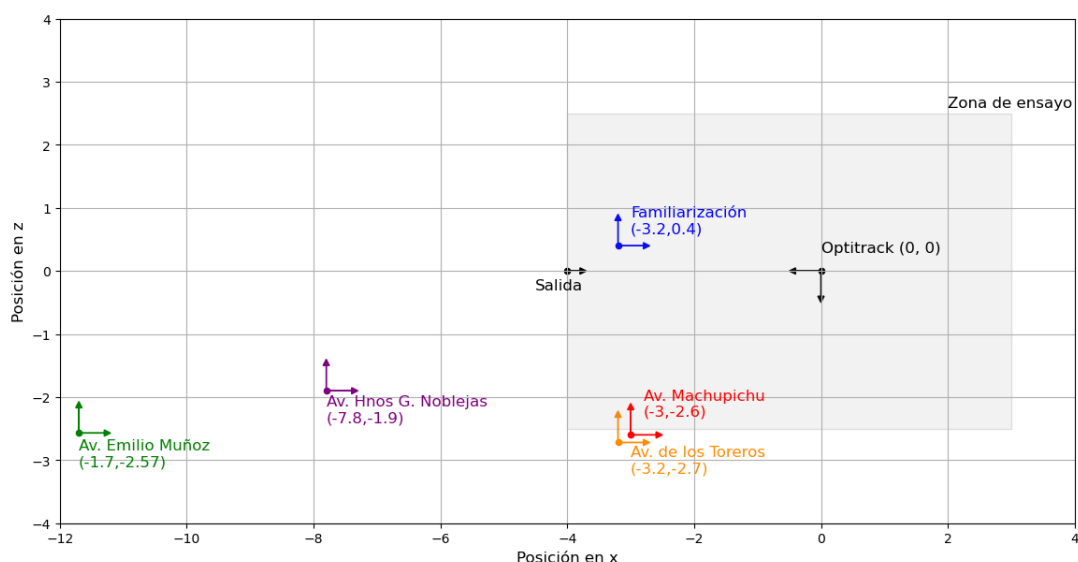


Ilustración 7. Situación del centro de coordenadas en cada escenario virtual con respecto a la sala de ensayos. Realizado con Matplotlib [17]

Como se observa en la Ilustración 7, en los escenarios ambientados en la *Avenida de Emilio Muñoz* y la *Avenida de los Hermanos García Noblejas*, el eje de coordenadas está situado por detrás de la sala de ensayo. La longitud del paso de cebra en estos escenarios virtuales supera el espacio disponible en la sala de ensayo del laboratorio de Realidad Virtual del CEDINT. Por tanto, al inicio de la simulación, el avatar se desplaza de forma automática hasta situarlo en el inicio de la zona de juego (marcado con un círculo rojo en la Ilustración 3).

Este movimiento no se corresponde a un avance real del usuario hacia la sala de ensayo y aun así la aplicación registra estos datos para los escenarios mencionados.

3.2.2.2 Identificación de zonas donde la detección es peor.

De acuerdo con las recomendaciones de *Optitrack* [5], la adecuada disposición del montaje de cámaras es crucial para que la detección sea óptima. En determinadas zonas de la sala de ensayos, la detección puede ser peor debido a que un menor número de cámaras tienen a los sensores en su campo de visión. La disposición del montaje de cámaras en la sala de ensayo se muestra de forma esquemática en la Ilustración 7. Las cámaras se sitúan a una altura aproximada de 3m sobre el suelo, y cubren un área de 7 x 5 m, y se orientan hacia el centro

de la sala, de manera que la región de la sala de ensayo en la que tienen la máxima capacidad de visión se encuentra en esta zona central.

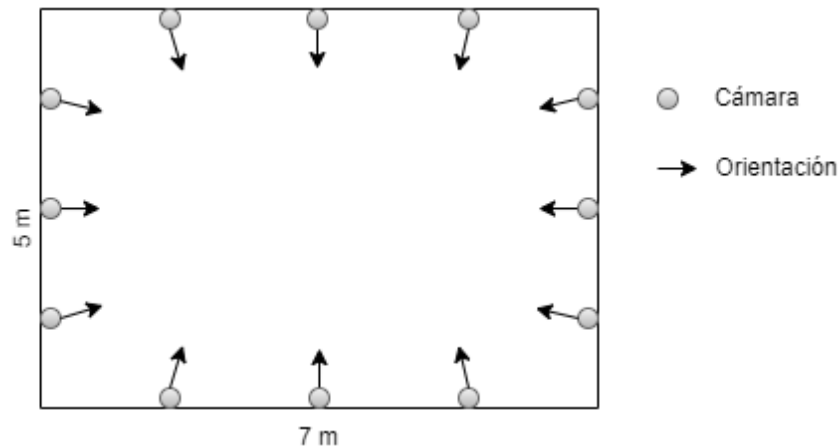


Ilustración 7. Disposición del montaje de cámaras alrededor de la sala de ensayo.

En el mapa de calor de la Ilustración 8 se muestra la distribución del nivel de visibilidad de la pelvis a lo largo de la sala de ensayo (representada en forma de rectángulo). Para cada cuadrado de 20 cm se muestra el valor mediano del número de marcadores detectados en esa zona. Según este valor, se le asigna un color en la escala de color secuencial que va desde rojo claro a rojo oscuro, distinguiendo tres niveles:

- **Mala visibilidad** (cero a dos marcadores detectados).
- **Visibilidad normal** (dos a cuatro marcadores detectados).
- **Buena visibilidad** (más de cuatro marcadores detectados).

La selección de estos rangos se ha realizado debido a las características del algoritmo de interpolación descrito en el apartado 3.2.1.2.

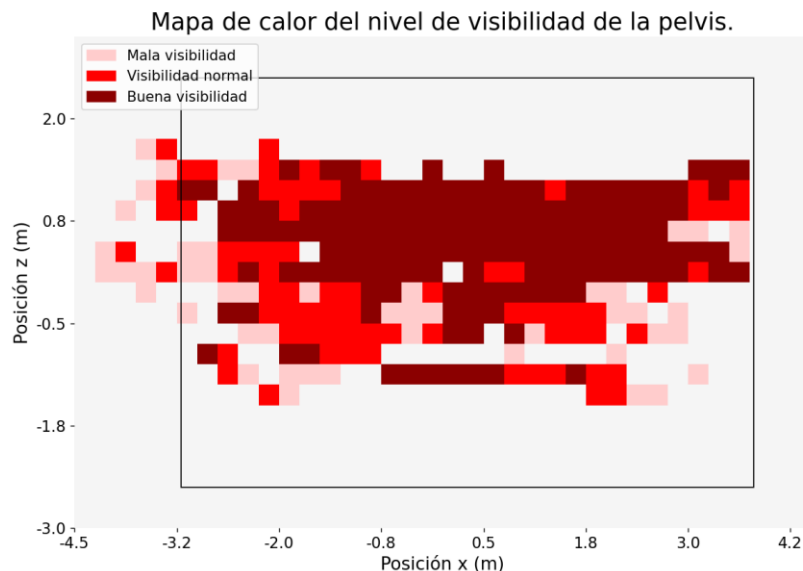


Ilustración 8. Mapa de calor del nivel de visibilidad que tienen las cámaras de la pelvis en la sala de ensayo.

Se observa como en los extremos de la sala la visibilidad de la pelvis es peor, mientras que en la zona central las cámaras tienen una mejor visibilidad de este elemento.

3.2.3 Preparación de los datos.

A partir de las necesidades identificadas en la fase de exploración, se realizan las siguientes transformaciones sobre los datos antes de pasar a las siguientes etapas.

- Se aplica una traslación y rotación a las medidas recogidas por la aplicación para unificar el sistema de referencia. El sistema de referencia definitivo es el que establece *Optitrack* (Ver Ilustración 7).
- En cada toma, se selecciona únicamente el intervalo de tiempo desde que el individuo empieza a andar hasta que vuelve a la posición inicial y el ensayo termina. Se seleccionan únicamente estas medidas debido a los siguientes motivos:
 - Al principio y al final de cada toma los usuarios no realizan ningún movimiento, por lo que la información que se puede extraer de estas medidas es escasa.
 - En los extremos de la sala de ensayo, la detección por parte de las cámaras es peor, como se describe en el apartado 3.2.2.2.
 - Existe una discrepancia entre la información recogida por *Optitrack* y por la aplicación en el intervalo de tiempo anterior a que el sujeto empiece a andar en el caso de los escenarios de la *Avenida de Emilio Muñoz* y de la *Avenida de los Hnos. García Noblejas* (ver apartado 3.2.1.2).

En la Ilustración 9 se muestra una comparativa de las medidas de posición en los ejes XZ, y de orientación con respecto al eje Y del sensor de la pelvis de un usuario en el escenario ambientado en la Avenida de Emilio Muñoz. El intervalo entre la primera medida válida y la última medida válida corresponde al tramo en el que el usuario está en movimiento, desde que parte de la posición inicial en el eje longitudinal (x) hasta que finalmente vuelve a esta posición.

Comparativa de las medidas de posición y orientación en el escenario de la Avda. de Emilio Muñoz para la pelvis.

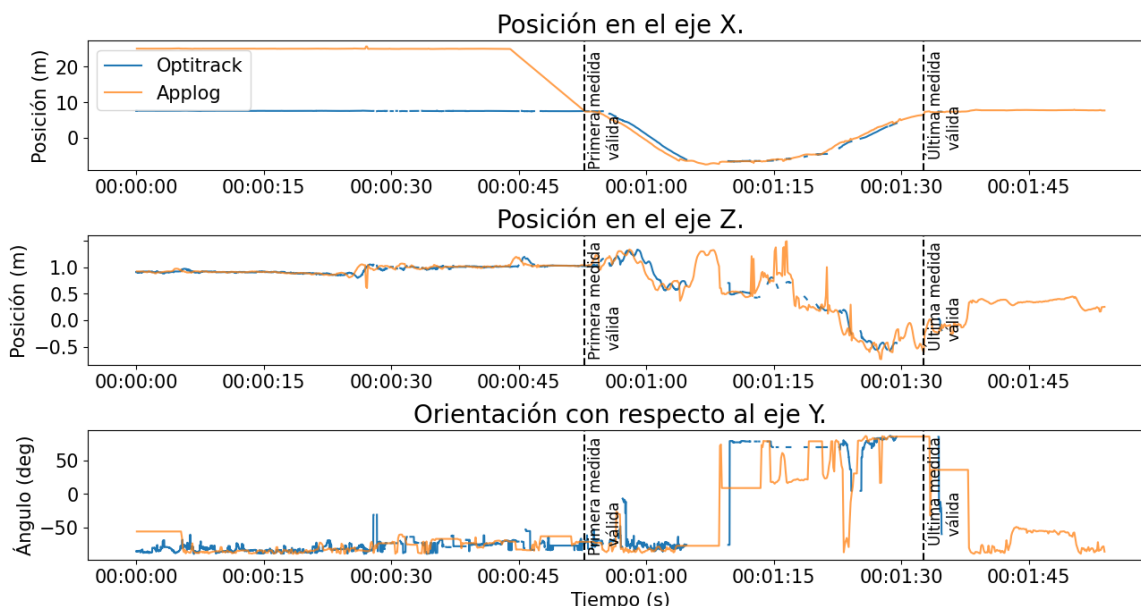


Ilustración 9. Comparativa entre las medidas de orientación y posición para un usuario en el escenario de Emilio Muñoz.

Finalmente, se descartan aquellas tomas en las que la información recogida por las cámaras no es representativa del ensayo ejecutado. En concreto, se descartaron todos los ensayos en los que el porcentaje de valores nulos en los datos registrados por *Optitrack* durante el movimiento era superior al 90% del total de medidas. De los 88 ensayos de los que se disponía en un inicio, se descartó un 10.9 % (10 ensayos).

En los siguientes apartados no se utilizaron los datos recogidos durante estos ensayos ya que no se podía garantizar que las medidas interpoladas tuviesen un mínimo de fiabilidad, al no disponer de datos de los sensores.

3.2.4 Minería de datos.

Esta es una de las fases más importantes del Trabajo de Fin de Máster. En esta etapa se transforma la información de partida en conocimiento valioso. La estructura de este apartado consta de tres partes:

- Una **primera parte** en la que presento un análisis del movimiento de cada parte del cuerpo, con el objetivo de justificar la selección de los parámetros a calcular.
- Una **segunda parte** en la que se detallan las técnicas que he utilizado, así como el proceso seguido a la hora de calcular las características que modelan el movimiento de los peatones.
- Una **tercera parte** en la que se desarrollan los algoritmos empleados y el proceso llevado a cabo a la hora de obtener perfiles del movimiento a partir de estas características.

Respecto a las tecnologías empleadas, en esta fase además de las ya mencionadas librerías de *Python*, he empleado *Scikit-Learn* [20], *NumPy* [21], y *SciPy* [22].

3.2.4.1 Análisis del movimiento.

Comprender la biomecánica natural de la marcha humana es fundamental en este Trabajo de Fin de Máster para seleccionar adecuadamente los parámetros a calcular.

El modelado de la marcha humana requiere no tan sólo de conocimientos matemáticos e informáticos, sino también de un profundo conocimiento de la mecánica del movimiento humano.

En [23] se describen las principales características de la marcha, la cual consta de dos fases: fase de apoyo y fase de balanceo. La marcha humana se define a partir de una serie de parámetros, como son la duración y longitud del paso, la separación entre pies, la velocidad, similitud y simetría. Estos parámetros se pueden extraer a partir de la información contenida en las trayectorias de las extremidades, la pelvis, la cintura y la cabeza de los peatones.

En la aplicación de Realidad Virtual que se utilizó en los ensayos donde se recogieron los datos que se emplean en este Trabajo, se incluye un escenario de familiarización con el entorno. El propósito de este escenario es familiarizar al usuario con el entorno virtual, facilitando que adquiera confianza de manera que su comportamiento se asemeje lo más posible al que tendría en el mundo real. En este contexto, se le solicita al usuario cruzar repetidamente un paso de peatones, eliminando todas las interferencias (tráfico u otros peatones) para descartar cualquier factor externo que pueda constituir un obstáculo. Los datos recabados en este entorno se utilizarán como referencia comparativa con aquellos obtenidos en escenarios que presentan riesgos de atropello.

A partir de los datos recogidos en este escenario se puede analizar cómo es el movimiento natural de las extremidades, la cabeza y la pelvis. Para ello en esta sección empleo los datos interpolados, ya que no contienen valores nulos.

3.2.4.1.1 Análisis del movimiento de la cintura.

De todas las partes del cuerpo de las que se recoge movimiento, la pelvis es la que presenta un menor número de grados de libertad, por lo que su movimiento es el más sencillo de modelar y analizar.

En la gráfica de líneas de la Ilustración 10 se muestra la evolución de la posición de la pelvis de un usuario en el ensayo de familiarización. El eje horizontal codifica el tiempo, y el eje vertical codifica la posición en un eje determinado. Cada punto representa una medida, y la calidad de la detección en ese punto se ha codificado a través del color según la misma escala secuencial definida en la sección 3.2.2.2.

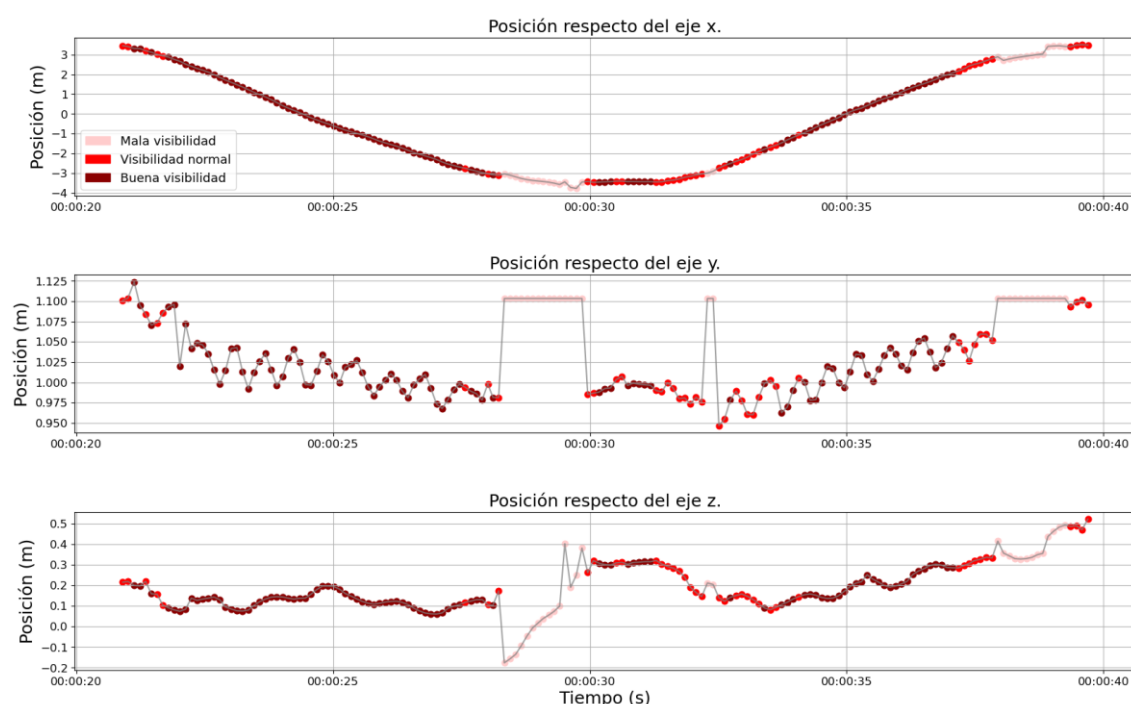


Ilustración 10. Movimiento de la pelvis con respecto a los tres ejes.

El análisis visual permite detectar un segmento con mala detección en la zona central, que se corresponde al extremo final de la sala donde el usuario se sale del campo de visión de las cámaras. De manera adicional, se observa que en la segunda mitad del recorrido hay varios puntos con mala detección en los que la interpolación introduce anomalías en el movimiento.

Dado que las medidas recogidas en los extremos de la sala tienen una baja fiabilidad (como se detectó en el apartado 3.2.2.2), se eliminan estas medidas y se separan los ensayos en idas o vueltas. En este apartado únicamente se utilizarán los datos de la ida (el segmento que corresponde a la primera vez que el usuario cruza la sala de ensayo antes de llegar al extremo final). Tras filtrar las medidas para obtener este segmento de la trayectoria, se obtiene el resultado que se muestra en la Ilustración 11.

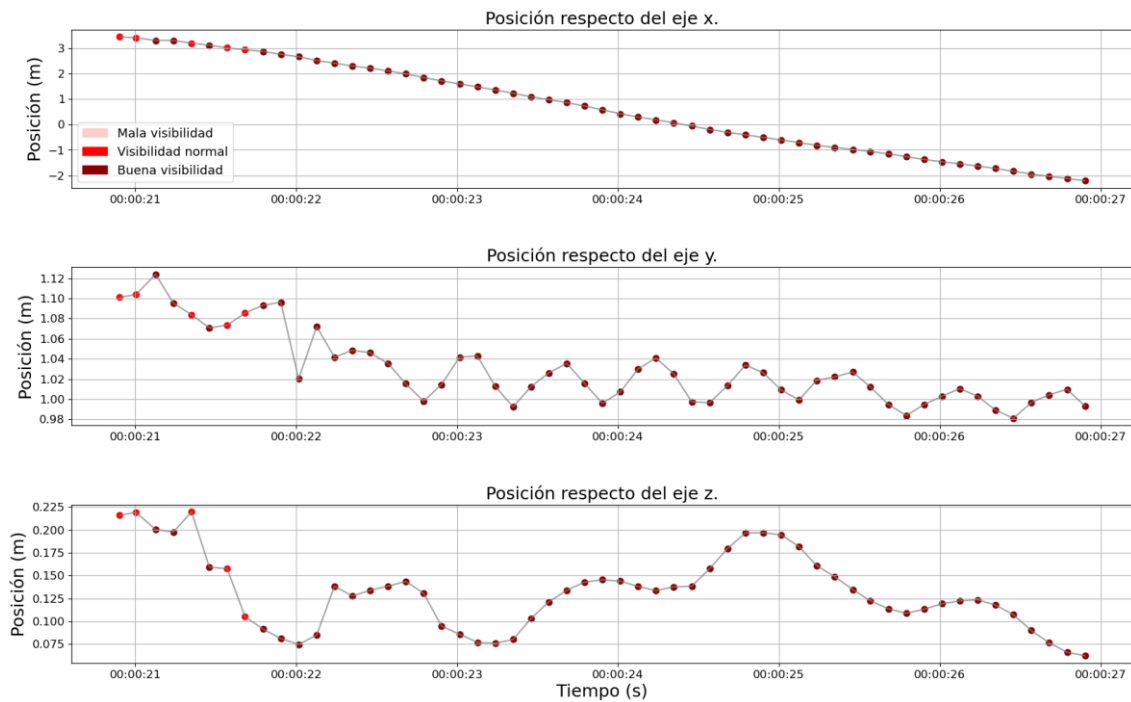


Ilustración 11. Movimiento de la pelvis durante la ida.

Se observa como la pelvis posee en el eje x un movimiento lineal con una clara tendencia descendente, a medida que el usuario se mueve hacia el final de la sala de ensayo. La pendiente de la trayectoria en el eje x corresponde a la velocidad del usuario en esta dirección.

En el eje y, se pueden detectar ligeras oscilaciones de baja amplitud (5 cm) en la posición de la pelvis, que corresponden a las oscilaciones del sensor y por tanto no se consideran para modelar el movimiento del peatón.

En el eje z, se observan las desviaciones del peatón con respecto a la línea base de la trayectoria (movimientos a la derecha y a la izquierda).

En la Ilustración 12 se muestra esta misma trayectoria en el plano XZ. El rectángulo gris marca los límites de la sala de ensayo, mientras que el rectángulo verde denota la zona que se emplea en el análisis.

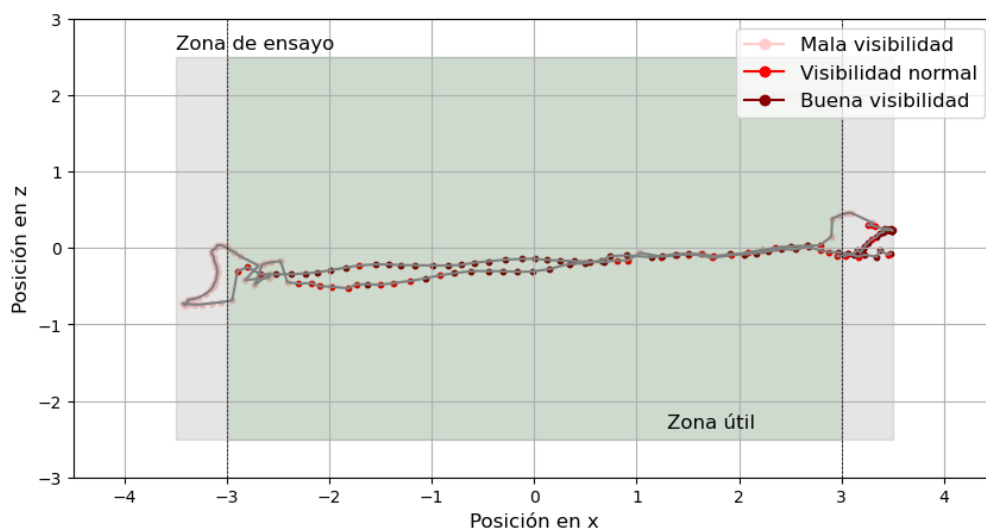


Ilustración 12. Trayectoria del usuario en el plano XZ.

3.2.4.1.2 Análisis del movimiento de los pies.

La trayectoria del pie izquierdo a lo largo del primer cruce de la sala de ensayo se muestra en la Ilustración 13. Además de la información que ya aportaba el movimiento de la cintura, se observa que el movimiento de los pies tiene una componente periódica tanto en el eje x como en el eje y. En el eje z no se aprecia esta periodicidad ya que este eje es perpendicular al plano transversal, que es donde se produce el movimiento de las extremidades inferiores al caminar.

La mínima posición en el eje vertical del pie se encuentra aproximadamente a 10 cm del suelo. Esto se debe a la colocación de los sensores durante los ensayos. En el caso del pie, el sensor se colocaba en el tobillo para minimizar los artefactos producidos por su balanceo.

En el eje x el pie sigue una trayectoria escalonada, con tramos de pendiente negativa (que corresponden a la fase de balanceo de la marcha), seguidos por tramos de pendiente nula (que corresponden a la fase de apoyo). Cabe recalcar que en la Ilustración 13 se están mostrando únicamente los datos de la ida. Durante la vuelta, la fase de balanceo correspondería a tramos con pendiente positiva.

La componente periódica de estas señales correspondería a la duración total de un paso. De la misma forma, puede obtenerse la longitud de un paso a partir del avance del pie en el eje x en un periodo.

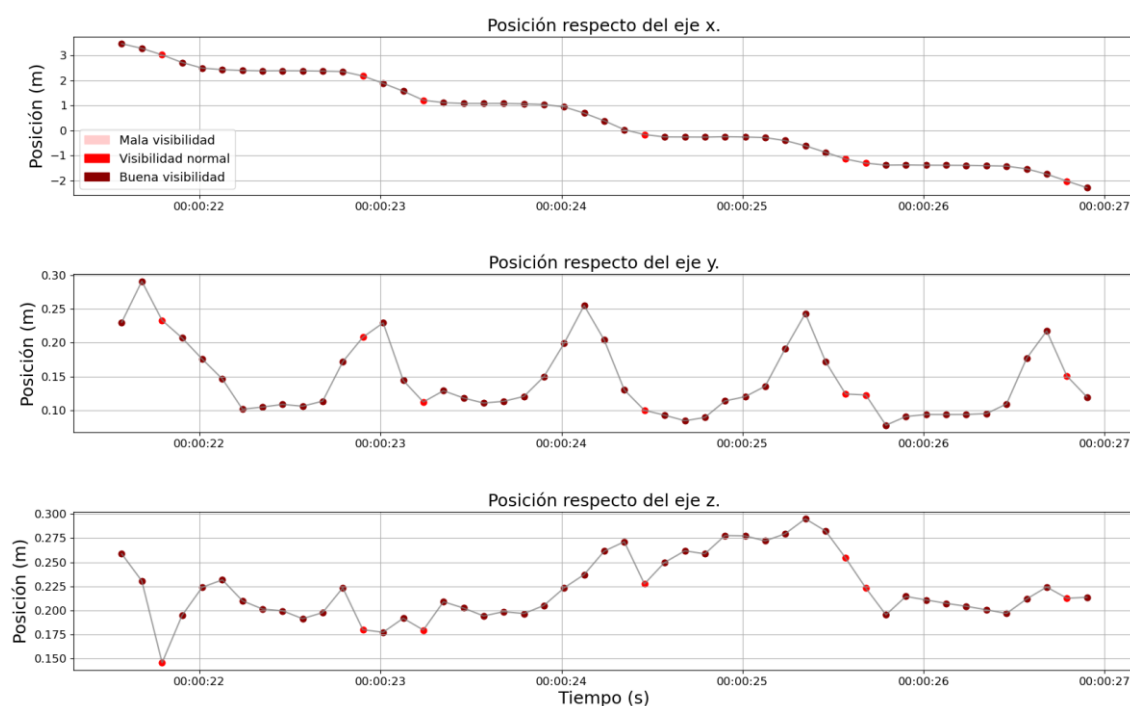


Ilustración 13. Trayectoria del pie izquierdo a lo largo de los tres ejes.

Si visualizamos el movimiento de los pies en su conjunto, como se muestra en la Ilustración 14, vemos que ambos pies tienen un movimiento muy similar, aunque en los ejes x e y las señales presentan un desfase que corresponde a la mitad de un paso.

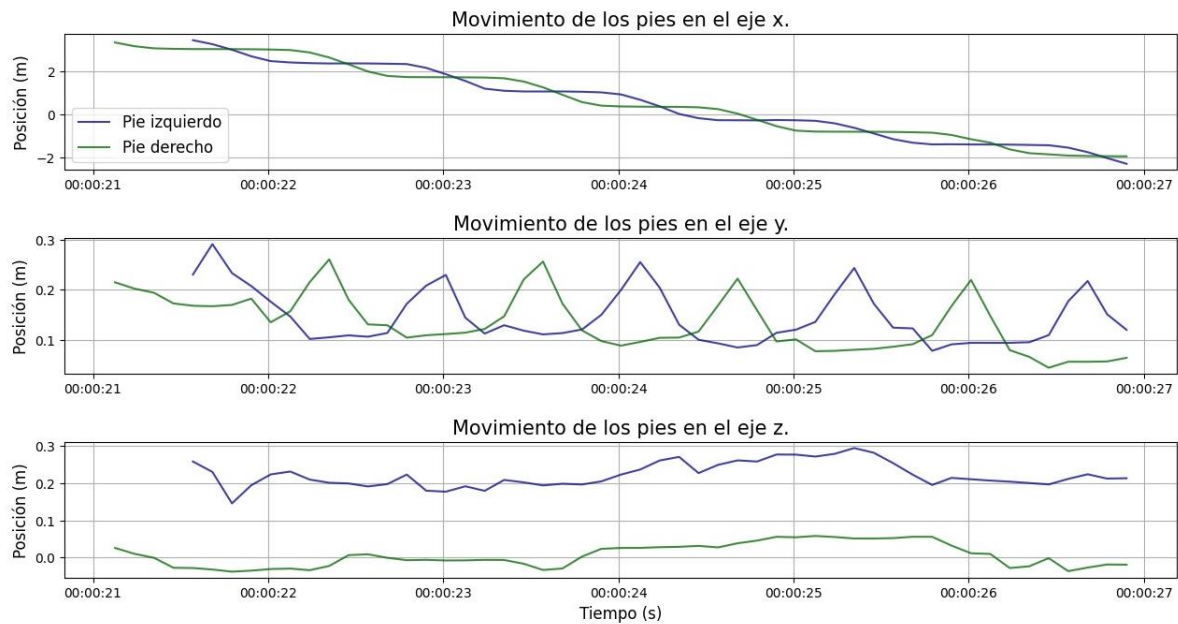


Ilustración 14. Comparativa en las trayectorias de los pies respecto a los tres ejes.

3.2.4.1.3 Análisis del movimiento de las manos.

El movimiento de las manos es probablemente el más difícil de analizar por el mayor número de grados de libertad que presentan frente a la cintura o incluso los pies. La sincronía entre el movimiento de las extremidades superiores e inferiores durante la marcha presentan simetría impar. Esto significa que la mano del lado derecho se mueve de forma similar al pie izquierdo, y viceversa. En la Ilustración 15 se muestra una comparativa de la evolución de la posición de la mano y el pie del lado izquierdo con respecto a los ejes x e y. Se observa como presentan un desfase que corresponde a la mitad de un periodo.

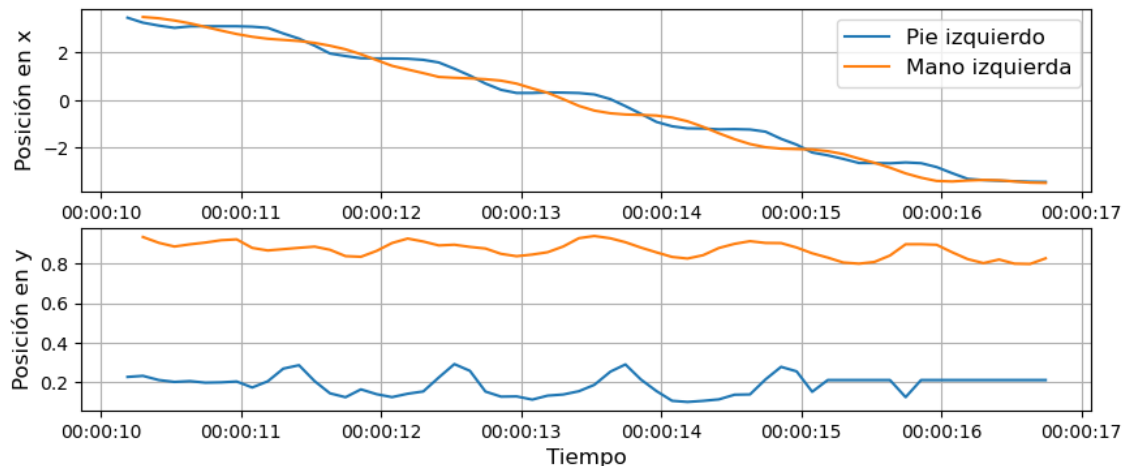


Ilustración 15. Evolución de la posición de la mano y el pie izquierdos en los ejes x e y.

Aunque el movimiento habitual de la mano sea similar al del pie del lado contrario, se pueden producir eventos asociados a la mayor libertad de movimientos que tienen las extremidades superiores con respecto a las inferiores. En la Ilustración 16 se muestra una comparativa en el movimiento en el eje vertical de la mano izquierda y la pelvis (en este caso, se representa un usuario diferente al que se visualiza en la Ilustración 15 para destacar estos

eventos). En rojo se observa un punto en el que el usuario levanta la mano izquierda, colocándola por encima de la pelvis.

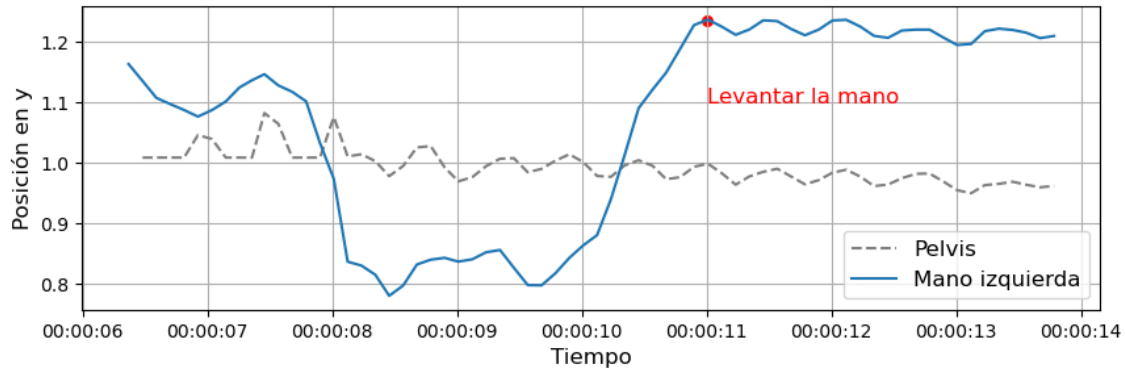


Ilustración 16. Evento levantar la mano.

Por tanto, y analizando las trayectorias de diferentes usuarios, se identifican como eventos separar la mano del cuerpo, ya sea en vertical o en horizontal.

3.2.4.1.4 Análisis del movimiento de la cabeza.

En [3], se emplea la posición y orientación de la cabeza para localizar al peatón. En este Trabajo utilizo la posición de la pelvis ya que esta posee menos grados de libertad. De manera adicional, la pelvis presenta un movimiento similar a la cabeza, una vez se considera la separación en altura entre ambas. Sin embargo, la orientación de la cabeza aporta información adicional acerca del estado atencional del usuario.

Dado que los valores de orientación recogidos por la aplicación se definen en forma de ángulos de Euler, la orientación alrededor del eje vertical permite detectar si el sujeto mira a la derecha o a la izquierda, mientras que la orientación con respecto al eje transversal permite detectar si el sujeto mira hacia arriba o hacia abajo.

3.2.4.2 Evaluación de la corrección biomecánica de los datos.

Para evaluar la coherencia de los datos se comprueba en qué grado se cumplen las restricciones biomecánicas del movimiento humano. El uso de métricas basadas en restricciones de la biomecánica humana para evaluar el sistema de captura de movimiento permite verificar el grado de efectividad del algoritmo de interpolación al comparar los resultados obtenidos empleando los datos originales.

La trayectoria de la cabeza y la cintura en el plano horizontal debería ser similar. Calculando la similitud entre ambas se puede obtener una métrica del grado de corrección del sistema de recogida de datos. De la misma forma, y teniendo en cuenta que en los ensayos participaron únicamente sujetos sanos, las trayectorias de los pies deberían tener un elevado grado de similitud si la recogida de datos se ha realizado correctamente.

Como estimaciones de la coherencia o corrección de los datos, se han propuesto dos medidas que se emplean para parametrizar la disparidad entre dos series temporales:

- El algoritmo de **Deformación Temporal Dinámica o Dynamic Time Warping (DTW)**[24] permite calcular la distancia entre dos trayectorias de diferente longitud. Dadas dos trayectorias, el algoritmo busca el alineamiento óptimo entre ambas, de manera que permite comparar las trayectorias de objetos que se muestrean a frecuencias diferentes. En la

Ilustración 17 se muestra el funcionamiento de este algoritmo, frente a la distancia euclídea calculada punto a punto.

El algoritmo DTW acepta un parámetro: el tamaño de la ventana de alineamiento. He establecido dicho parámetro en tres muestras, lo que corresponde a tres décimas de segundo en el caso de las medidas recogidas por la aplicación.

El valor de distancia obtenido se divide entre la longitud de la secuencia más corta para obtener un valor que correspondería a la *disparidad promedio por medición* entre las señales durante toda la trayectoria.

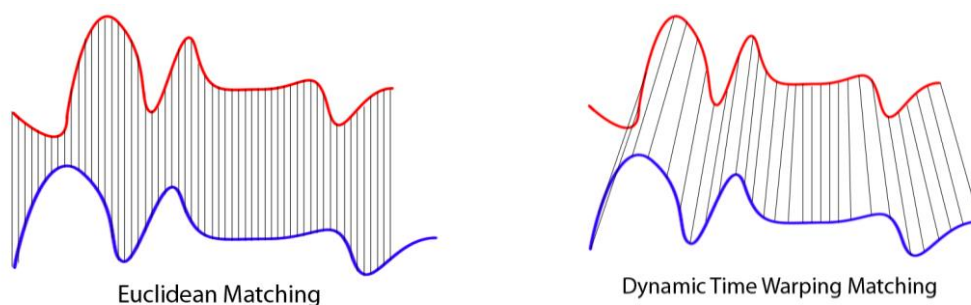


Ilustración 17. Distancia euclídea punto a punto frente a la distancia DTW.

- El algoritmo de la **Subsecuencia más Larga o Longest Common Subsequences (LCSS)**[25] permite obtener la proporción de tiempo en la que la distancia entre dos elementos es inferior a un umbral establecido (0.1 m). Es un algoritmo más robusto frente a los valores atípicos que DTW, pero presenta el problema de que no es capaz de diferenciar entre dos trayectorias con subsecuencias comunes similares pero diferentes tamaños de brechas intermedias [26]. Para que el resultado de este algoritmo no se vea sesgado por la longitud total de la secuencia, se empleará la proporción de la secuencia total que cubre la LCSS, en lugar del valor original. De esta forma se obtiene un valor normalizado entre 0 y 1.

Con el objetivo de aprovechar las ventajas de ambos algoritmos, se emplearon ambas métricas para calcular la similitud o distancia entre la cabeza y la cintura y entre pies, permitiendo ofrecer una evaluación de la corrección de las medidas obtenidas mediante el sistema de captura de movimiento utilizado.

La **LCSS** ofrece una medida del grado de mejora conseguida por el algoritmo de interpolación. Por otra parte, la distancia calculada mediante **DTW** aporta un valor concreto en el grado de coherencia mecánica de las medidas.

Antes de realizar el cálculo de la similitud, las trayectorias se estandarizan restando la media y dividiendo entre la desviación estándar. De esta forma se anula la separación en altura entre la cabeza y la pelvis; y la separación en el eje longitudinal entre los pies. De manera adicional, previamente a calcular la similitud entre los pies, el desplazamiento temporal que se observa en la Ilustración 14 se corrige. En la Ilustración 18 se muestra un ejemplo de este proceso.

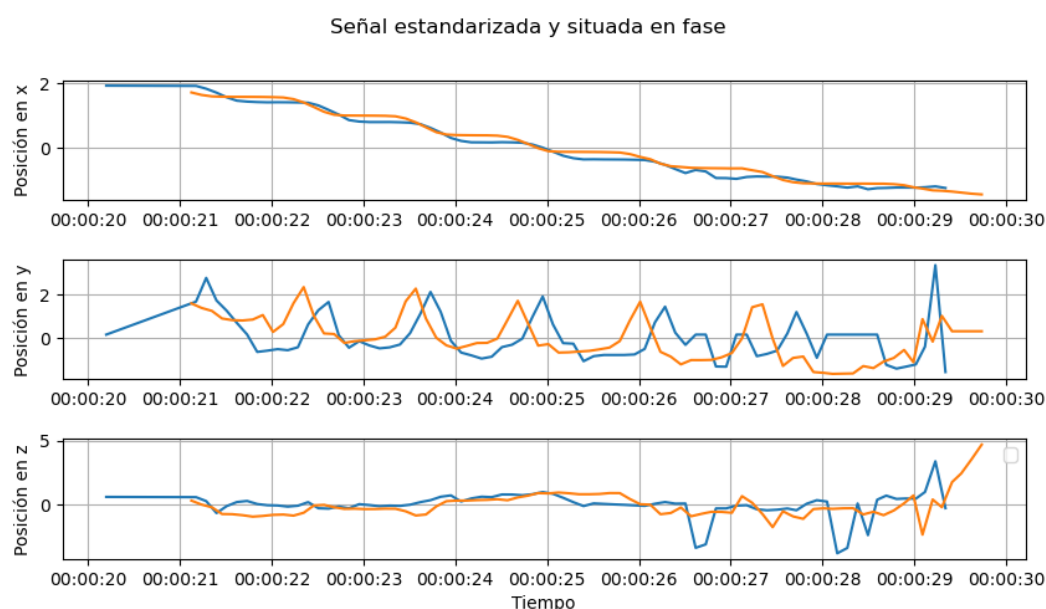
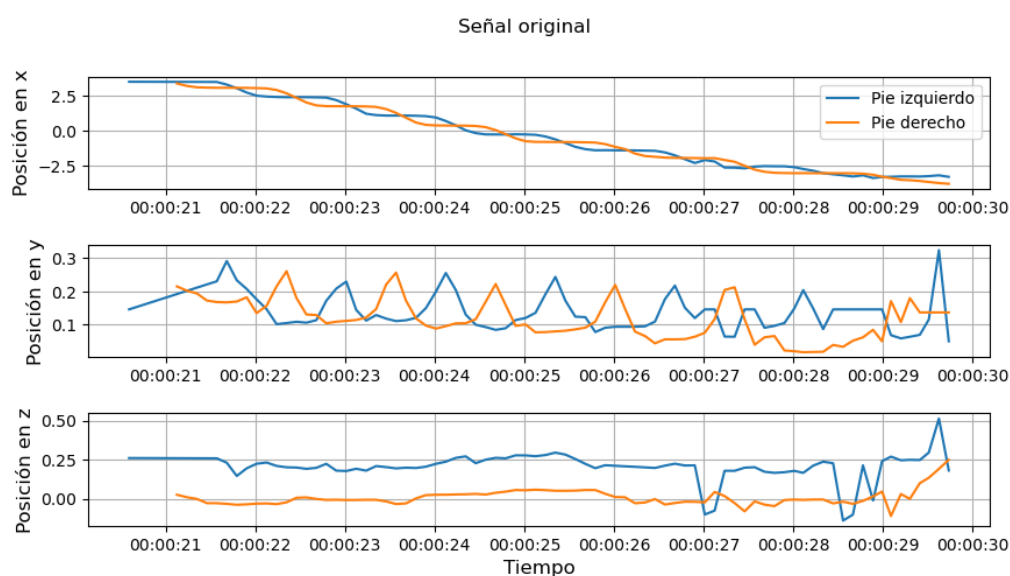


Ilustración 18. Ajuste de la señal de los pies antes de realizar los cálculos de distancia.

3.2.4.3 Extracción de características.

A partir de las conclusiones obtenidas en el análisis realizado en el apartado 3.2.4.1, se obtienen las características que modelan el movimiento.

El cálculo de características se realiza separando la ida y la vuelta de cada ensayo, de manera que para cada usuario y escenario se obtienen dos trayectorias parametrizadas: una con los datos de la ida, y otra con los datos de la vuelta.

3.2.4.3.1 Extracción de características a partir del movimiento de la pelvis.

La pelvis es la mejor opción para representar la posición y orientación del peatón, puesto que tiene un menor número de grados de libertad que el resto de los elementos que se monitorizan.

3.2.4.3.1.1 Velocidad de avance del usuario.

La **velocidad del usuario** puede calcularse diferenciando los valores de la posición en el eje longitudinal (x) y dividiendo entre el periodo de muestreo. Se calcula para cada cruce de la sala de ensayos la velocidad mediana en valor absoluto.

3.2.4.3.1.2 Desviación del usuario con respecto al eje de la trayectoria.

La **máxima desviación del eje** de la trayectoria en el eje transversal se obtiene restando a todos los valores de la posición con respecto al eje z su valor mediano. Posteriormente se obtiene el máximo de esta diferencia en valor absoluto. En la Ilustración 19 se muestra un ejemplo de cómo se calcula este parámetro.

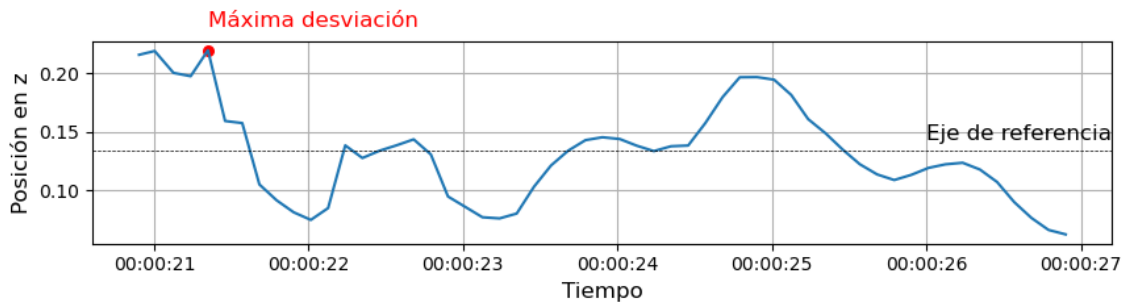


Ilustración 19. Cálculo de la máxima desviación del eje de la trayectoria.

3.2.4.3.1.3 Similitud cabeza-cintura.

La **similitud** entre las trayectorias de la cabeza y la cintura puede obtenerse tal y como se describe en el apartado 3.2.4.2.

3.2.4.3.1.4 Detección del evento: pararse y retroceder

Finalmente se calcula el **número de veces** que el usuario se detiene o da marcha atrás durante un cruce.

Este último parámetro se obtiene simbolizando la trayectoria del peatón a lo largo del eje longitudinal. La simbolización de series temporales es una técnica basada en la asignación de etiquetas a diferentes tramos de una secuencia de valores numéricos, convirtiéndola en una secuencia de símbolos. Esto facilita la detección de eventos, entre otras ventajas.

El enfoque habitual al simbolizar una secuencia consta de dos pasos. En primer lugar, se segmenta la secuencia calculando el valor promedio para ventanas de un determinado tamaño. Posteriormente se utiliza un algoritmo de asignación de símbolos que asigna un símbolo a cada ventana [27], [28].

En la Ilustración 20 se muestra la lógica del algoritmo de asignación de símbolos implementado.

```
def simbolizar (serie, symbols: dict) # SYMBOLS ES UN DICCIONARIO CONDICIÓN: VALOR
    serie_segmentada = segmentar_serie (serie, TAMAÑO DE VENTANA)
    serie_simbolizada = serie ()
    for indice in range (0, len (serie_segmentada), TAMAÑO_DE_VENTANA):
        if indice < TAMAÑO_DE_VENTANA:
            # A LA PRIMERA VENTANA SE LE ASIGNA EL SÍMBOLO POR DEFECTO
            serie_simbolizada = asignar_simbolo (serie_simbolizada, indice, list
(symbols. values) [0])
        else:
            # ASIGNAR EL PRIMER SÍMBOLO PARA EL QUE SE CUMPLE SU CONDICIÓN
            for condición, valor in symbols. items():
                if eval(condición):
                    serie_simbolizada = asignar_simbolo (serie_simbolizada, i,
valor)
                    break
            return serie_simbolizada

def asignar_simbolo (serie_simbolizada, window_index, symbol):
    # ASIGNAR EL SÍMBOLO A TODOS LOS ELEMENTOS DE LA VENTANA ACTUAL
    for j in range (window_index, window_index+1-TAMAÑO _DE_VENTANA)
        serie_simbolizada[j] = symbol
    return serie_simbolizada
```

Ilustración 20. Algoritmo de asignación de símbolos.

El tamaño de la ventana de segmentación es de 10 muestras, lo que se corresponde con un tamaño de ventana de 1 segundo dada la frecuencia de muestreo de Unity (10Hz). La serie que se simboliza es la posición en el eje x. Para cada ventana, se calcula el avance con respecto a la ventana anterior y la velocidad media intra-ventana. A partir de estos parámetros se asigna el símbolo correspondiente según el diccionario de símbolos de la Tabla 6. Si el usuario avanza más de 10 cm en intervalo de un segundo se considera que se hay movimiento (hacia delante o hacia detrás). Si además la velocidad instantánea es superior a 1.1m/s se considera que está avanzando rápido.

Símbolo	Descripción	Condición.
S	Usuario parado	$abs(avance) < 0.1m$
A1	Usuario avanza lento.	$avance > 0.1m \ \&\& \ abs(velocidad) < 1.1 \ m/s$
A2	Usuario avanza rápido.	$avance > 0.1m \ \&\& \ abs(velocidad) > 1.1m/s$
R1	Usuario retrocede despacio.	$avance < -0.1m \ \&\& \ abs(velocidad) < 1.1m/s$
R2	Usuario retrocede rápido.	$avance < -0.1m \ \&\& \ abs(velocidad) > 1.1m/s$

Tabla 6. Diccionario de símbolos empleado.

El resultado de la simbolización se muestra en la Ilustración 21. Ilustración 21 Las cruces (x) indican los puntos donde se asignan símbolos. Se utiliza el color para codificar el tipo de evento detectado.

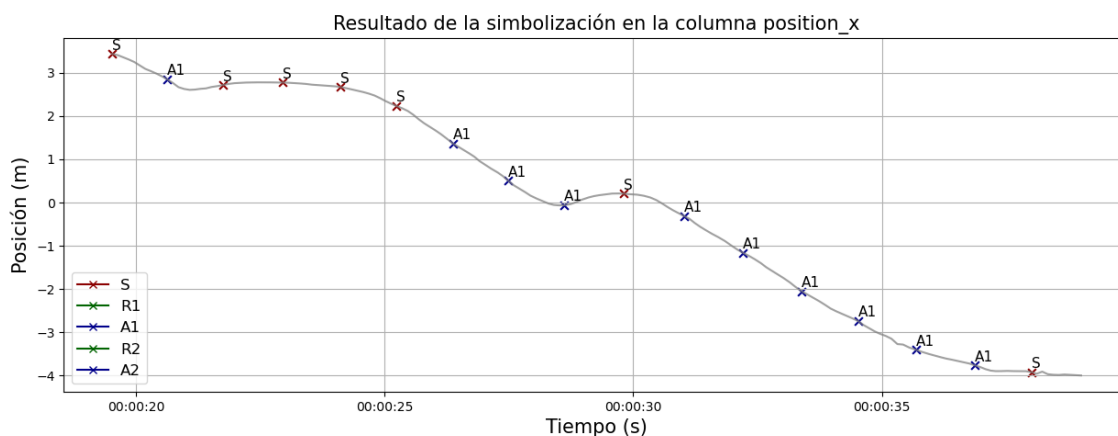


Ilustración 21. Resultado de la simbolización de la posición de la cintura.

3.2.4.3.2 Extracción de características a partir del movimiento de los pies.

La mayoría de las características que definen la marcha humana pueden extraerse a partir de la información contenida en el movimiento de los pies.

3.2.4.3.2.1 Periodo de la marcha.

La posición vertical del pie permite obtener el **periodo de la marcha** (o duración de un paso) realizando la Transformada de Fourier de la secuencia temporal.

La Transformada de Fourier es una transformación que se aplica sobre una secuencia para simplificar la información contenida en ella en forma de sus componentes de frecuencia [29]. Posteriormente se pueden extraer las frecuencias dominantes a partir de los picos más marcados en la transformada. En este caso se obtiene la frecuencia del paso a partir del máximo pico en la Transformada de Fourier en la banda de frecuencias 0.5 – 2.5Hz (entre 0.5 y 2.5 pasos por segundo).

3.2.4.3.2.2 Longitud media de un paso.

Una vez se calcula el periodo, se puede obtener fácilmente la **longitud de un paso** calculando el avance en el eje longitudinal del pie en cada periodo. Realizando la mediana se obtiene una estimación de la longitud del paso en cada cruce.

3.2.4.3.2.3 Amplitud y ángulo de la marcha.

La **amplitud de la marcha** se define como la separación entre pies en el eje transversal (z). Para cada cruce, se calcula la amplitud mediana.

El **ángulo** de la marcha corresponde al ángulo que forman los pies con respecto al eje vertical (y). Se obtiene también la mediana de este parámetro.

3.2.4.3.2.4 Similitud de la marcha.

La **similitud** de la marcha es la similitud entre las trayectorias de ambos pies, y se calcula tal y como se describe en el apartado 3.2.4.2.

3.2.4.3.3 Extracción de características a partir del movimiento de las manos.

Las manos constituyen los elementos cuyo movimiento es más difícil de caracterizar debido a que poseen un mayor número de grados de libertad que el resto.

Algunas de las características del movimiento de las manos, como el periodo de oscilación, pueden obtenerse de manera similar a como se realiza el cálculo para los pies. Del mismo modo, se podría calcular la similitud en el movimiento de las manos. Sin embargo, el valor obtenido podría verse afectado por la aparición de eventos como levantar o separar la mano del cuerpo.

3.2.4.3.3.1 Detección de los eventos: levantar la mano y separar la mano.

Dado que los parámetros de la marcha ya se calculan a partir de las trayectorias de los pies, cuyo movimiento es más estacionario que el de las manos, las trayectorias de las manos únicamente se han empleado para detectar los mencionados eventos a partir de la simbolización de su trayectoria.

Empleando el algoritmo de simbolización de la Ilustración 20 se simbolizan las trayectorias de las manos. En este caso, para cada segmento, se calcula la distancia de la mano a la pelvis en el eje vertical y horizontal. De esta forma se pueden detectar los eventos incluidos en el diccionario de símbolos de la Tabla 7, donde dp es la distancia de la mano a la pelvis. Si la distancia entre la mano y la pelvis supera su valor mínimo durante la trayectoria por más de diez centímetros, se considera que la mano está lejos del cuerpo. Si esta separación se da en el eje vertical, se considera que la mano está separada en el plano vertical. Si se da en el eje transversal, se considera que la mano está separada del cuerpo en el plano transversal.

Símbolo	Descripción	Condición
M00	Mano relajada.	$dp_y < [\min(dp_y) + 0.1] \ \&\& \ dp_z < [\min(dp_z) + 0.1]$
MUU	Mano separada en el plano vertical.	$dp_y > [\min(dp_y) + 0.1] \ \&\& \ dp_z < [\min(dp_z) + 0.1]$
MLL	Mano separada en el plano horizontal.	$dp_y < [\min(dp_y) + 0.1] \ \&\& \ dp_z > [\min(dp_z) + 0.1]$
MUL	Mano separada en los planos vertical y horizontal.	$dp_y > [\min(dp_y) + 0.1] \ \&\& \ dp_z > [\min(dp_z) + 0.1]$

Tabla 7. Diccionario de símbolos diseñado.

Detectados los eventos, se pueden calcular el número de veces que el usuario realiza cada acción a lo largo de cada trayectoria, obteniendo información acerca de las acciones que el usuario realiza durante el movimiento.

3.2.4.3.4 Extracción de características a partir del movimiento de la cabeza.

En [3], se utiliza la cabeza como estimación de la posición del peatón. En este Trabajo de Fin de Máster utilizo la pelvis para dicho propósito. El objetivo es doble: por un lado, la pelvis posee un menor número de grados de libertad que la cabeza. Por otra parte, pretendo ofrecer un enfoque diferente al que se utiliza en el mencionado proyecto para caracterizar los movimientos de los sujetos. Sin embargo, la orientación de la cabeza permite obtener información adicional que no se puede conseguir únicamente a partir de los datos de la cintura.

3.2.4.3.4.1 Discretización de la orientación.

La orientación de la cabeza alrededor de los ejes vertical y transversal permite conocer la dirección de la mirada del usuario. En la Ilustración 22 se muestra la evolución de la orientación de la cabeza con respecto a ambos ejes para un usuario en el escenario de la Avenida de Machupichu. En un umbral de 30

grados alrededor de una orientación basal de 0 grados todos los valores se consideran como puntos en los que el usuario mira hacia el frente. Fuera de ese umbral se considera que la mirada se dirige a la dirección correspondiente según el sentido de la orientación. Si observamos la orientación alrededor del eje z, vemos que en la parte inicial se producen picos que se atribuyen a una mala colocación del casco. Para evitar que estos valores afecten al resultado final, se consideran únicamente las medidas de orientación que en valor absoluto no superen 60°. Este valor se corresponde de manera aproximada al máximo rango de movimiento del cuello [30].

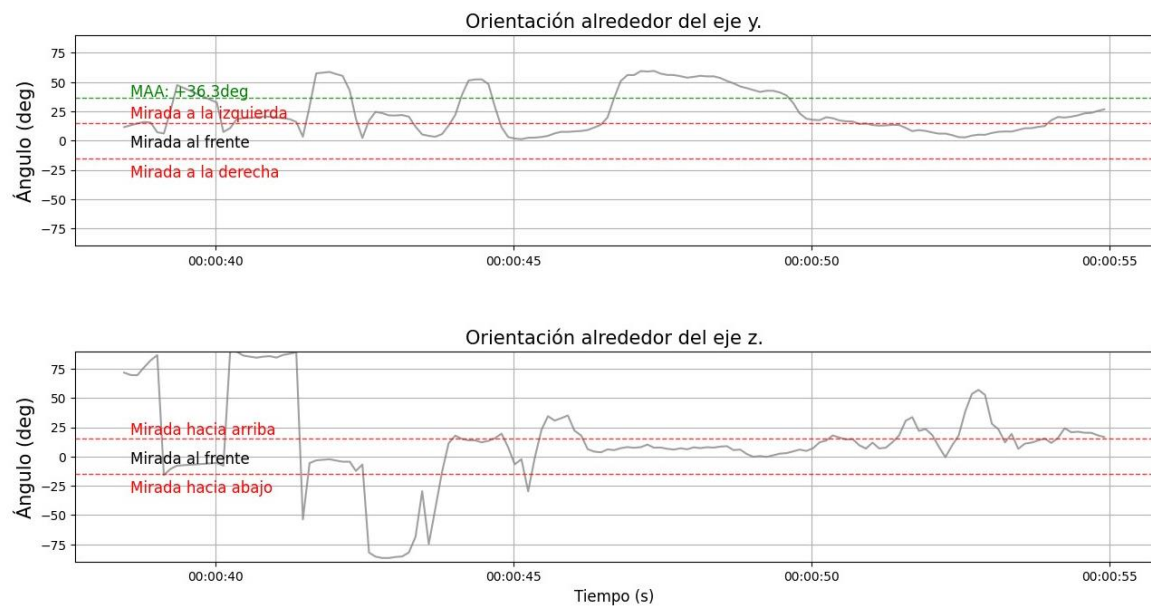


Ilustración 22. Evolución de la orientación de la cabeza en el escenario de la Avenida de Machupichu.

El *Mínimo Ángulo Atencional* es la mínima orientación de la cabeza alrededor del eje vertical necesaria para que el peatón sea capaz de ver el coche que presenta un riesgo potencial de colisión. Se calcula tal y como describen los autores en [3], a partir del arco-tangente del ángulo que forma la línea que une al vehículo y el peatón con el eje longitudinal, en el momento que el coche empieza a frenar. El valor concreto en cada caso se incluye en la Tabla 8. Los valores establecidos para los escenarios de la Avenida de Machupichu y la Avenida de los Hnos García-Noblejas son los que aportan los autores de [3]. Los valores de los escenarios de la Avenida de los Toreros y la Avenida de Emilio Muñoz se han calculado siguiendo las instrucciones que brindan estos mismos autores.

Escenario	MAA (°)
Av. de Emilio Muñoz	-34.6
Av. de los Toreros	81.5
Av. de los Hnos G-Noblejas	43.4 [3]
Av. de Machupichu	36.9 [3]

Tabla 8. Valores de MAA para cada escenario.

En el escenario de la *Avenida de Emilio Muñoz*, el coche aparece por la derecha, por lo que el valor de MAA es negativo. En el resto de los escenarios el coche aparece por la izquierda. En el escenario de la *Avenida de los Toreros* se estima

que el sujeto tiene que girar la cabeza casi 90° con respecto al eje vertical para visualizar el vehículo, ya que el coche aparece al inicio del paso de peatones (en el primer carril).

De esta forma se puede obtener el **porcentaje de tiempo atencional** (PTA, o proporción del tiempo en el que el usuario mira hacia el coche) y la proporción de tiempo que el usuario mira en cada una de las direcciones establecidas (arriba, abajo, a la izquierda y a la derecha) durante la trayectoria.

3.2.4.3.4.2 Estimación de la altura del usuario.

La **altura del usuario** puede estimarse a partir de la mediana de la posición vertical de la cabeza. Aunque no es un parámetro relacionado con el movimiento, he creído conveniente calcularlo porque considero que puede guardar relación con el resto de las variables.

3.2.4.4 Identificación de patrones del movimiento

En este apartado desarrollo el proceso seguido de cara a obtener los resultados finales del Trabajo de Fin de Máster. A partir de las características del movimiento obtenidas, se entrenan algoritmos de aprendizaje no supervisado que agrupan las trayectorias en función de su similitud. Dado que las características obtenidas tienen diferentes escalas, antes de realizar el agrupamiento todas las variables se estandarizan de manera que la escala no influya en los resultados. Los métodos empleados [31], [32] se describen en los siguientes apartados:

3.2.4.4.1 *Clustering jerárquico aglomerativo.*

Este algoritmo empieza asignando cada instancia a un grupo diferente. Sucesivamente en cada iteración, agrupa las instancias en base a un criterio de unión hasta que todas las instancias pertenecen al mismo clúster.

He seleccionado este algoritmo por su fácil interpretabilidad al permitir visualizar el resultado del clustering en un dendrograma. El criterio de unión empleado es el criterio de unión *Ward*, que trata de minimizar la suma del cuadrado de las distancias intra-clúster [33].

La principal desventaja que posee el algoritmo es que no permite la reasignación de instancias a nuevos clústeres, de manera que otros métodos, como el clustering particional, pueden arrojar mejores resultados si el número de perfiles a obtener se conoce con anterioridad.

3.2.4.4.2 *Clustering particional (k-means)*

Los algoritmos de clustering particional enfocan la asignación de grupos como un problema de optimización. El objetivo es dividir las instancias en un número determinado de particiones (*k*) de manera que se minimice la suma de los cuadrados de las distancias de cada instancia al centroide del clúster al que pertenece.

Este tipo de algoritmos permiten la reasignación de instancias a nuevos clústeres, a diferencia del clustering aglomerativo. Sin embargo, los resultados están sesgados por la selección del parámetro *k* (número de clústers), y por la inicialización de los centroides de los grupos. Además, su interpretabilidad es inferior que la que poseen algoritmos de clustering jerárquico.

La variante de *k-means* que he utilizado reasigna los centroides de cada clúster en cada iteración [34]. La inicialización de los centroides no es arbitraria, si no que la posición de los centroides se inicializa de manera que se maximice la distancia entre los grupos en la primera iteración.

3.2.4.4.3 Métricas de evaluación de los resultados del agrupamiento.

A la hora de seleccionar el número óptimo de clústeres, así como para evaluar la separabilidad y homogeneidad del resultado final, se han utilizado las siguientes métricas.

El **coeficiente de silueta** [35] se calcula para cada muestra. Evalúa el resultado del agrupamiento en clústeres a partir de la distancia media de cada instancia al resto de instancias del mismo grupo; y de la distancia media al siguiente grupo más cercano. Su valor varía en el rango $[-1,1]$, donde un valor de -1 indicaría que los clústeres están mal definidos; un valor cercano a 0 indicaría que la distancia entre clústeres no es relevante; y un valor de 1 indicaría que la definición de los clústeres es perfecta.

El **índice de Calinski-Harabasz** [36] se define como la ratio de la dispersión entre clústeres e intra-clúster (siendo la dispersión la suma de las distancias al cuadrado).

3.2.4.4.4 Proceso seguido en la obtención de patrones de movimiento.

El enfoque que he llevado a cabo a la hora de obtener perfiles del movimiento se basa en la explotación de los beneficios de los dos métodos de clustering descritos:

- En primer lugar, he utilizado el análisis de componentes principales para identificar aquellas variables que mejor explican la varianza del conjunto de datos y así reducir la dimensionalidad.
- Se selecciona el número adecuado de grupos a partir de los resultados obtenidos mediante el algoritmo de clustering aglomerativo y de su interpretación. El número óptimo de clústers (valor de k) es aquel para el cual, dadas las características del conjunto de datos, se maximiza la homogeneidad intragrupo a la vez que se maximiza la heterogeneidad entre grupos. Esto se determina analizando el dendrograma obtenido por el algoritmo jerárquico; y los resultados obtenidos en los coeficientes de silueta y de Calinski-Harabasz.
- Posteriormente se analizan los resultados obtenidos mediante *k-means* para el número de grupos seleccionado.
- El uso de pruebas estadísticas como la prueba de Kruskal-Wallis [37] permiten identificar las variables más influyentes en la separabilidad de los grupos.
- Finalmente se calculan los parámetros que distinguen cada grupo.

Los resultados obtenidos siguiendo esta metodología se describen y comentan en el capítulo siguiente.

4 Resultados y discusión.

En este apartado se presentan los resultados obtenidos a la hora de lograr la consecución de cada uno de los objetivos planteados al inicio del Trabajo de Fin de Máster. Posteriormente elaboro una breve discusión de las conclusiones que se pueden extraer de los mismos.

4.1 Evaluación de la fiabilidad y la corrección de las medidas.

En esta sección se presenta el resultado del informe de evaluación de la corrección de las medidas, que constituye la consecución del OT1. Aunque una parte de este ya se introdujo como parte de la fase de exploración de los datos, en este apartado se pretende ofrecer una evaluación más completa.

En primer lugar, se evalúa el grado de visibilidad de las cámaras durante el ensayo, empleando el porcentaje de cada toma en la que la detección no es óptima. De esta forma se puede determinar hasta qué punto se ha necesitado de una interpolación en cada trayectoria.

Posteriormente se evalúa el grado de consistencia de los datos a partir de nociones de biomecánica. Durante la marcha, las trayectorias de la cabeza y la cintura deberían ser similares. De manera adicional, el movimiento de los pies debe guardar cierta sincronía. Partiendo de estas bases se evalúa la mejora conseguida al aplicar el algoritmo descrito en el apartado 3.2.1.2. De manera adicional, el uso de mediciones basadas en restricciones biomecánicas habilita la identificación de trayectorias que contienen información ruidosa, ya sea por factores como problemas en la visibilidad de los sensores (debido a una colocación inadecuada o a oclusiones causadas por la dinámica del movimiento), o por eventos inesperados en el contexto de la ejecución de las tareas, como quitarse el casco al final de una toma.

Todo ello permite establecer una base sólida de cara a lograr los siguientes objetivos que se plantearon en este Trabajo de Fin de Máster.

4.1.1 Análisis de la visibilidad que tienen las cámaras de cada Rigid Body.

El nivel de visibilidad que tienen las cámaras *Optitrack* a lo largo de un cruce puede medirse a través del porcentaje de tiempo en el que la detección de un sensor (*Rigid Body*) determinado es sub-óptima. Una medición se considera con detección sub-óptima si el número de marcadores detectados para un sensor determinado es inferior a tres. Este valor se selecciona de acuerdo con las características del algoritmo descrito en el apartado 3.2.1.2 y las recomendaciones en la documentación técnica de *Optitrack* [5].

En la Ilustración 23 se muestra una comparativa en el porcentaje de tiempo con mala detección separando la ida hasta el extremo de la sala de ensayo (verde) de la vuelta al punto de partida (naranja).

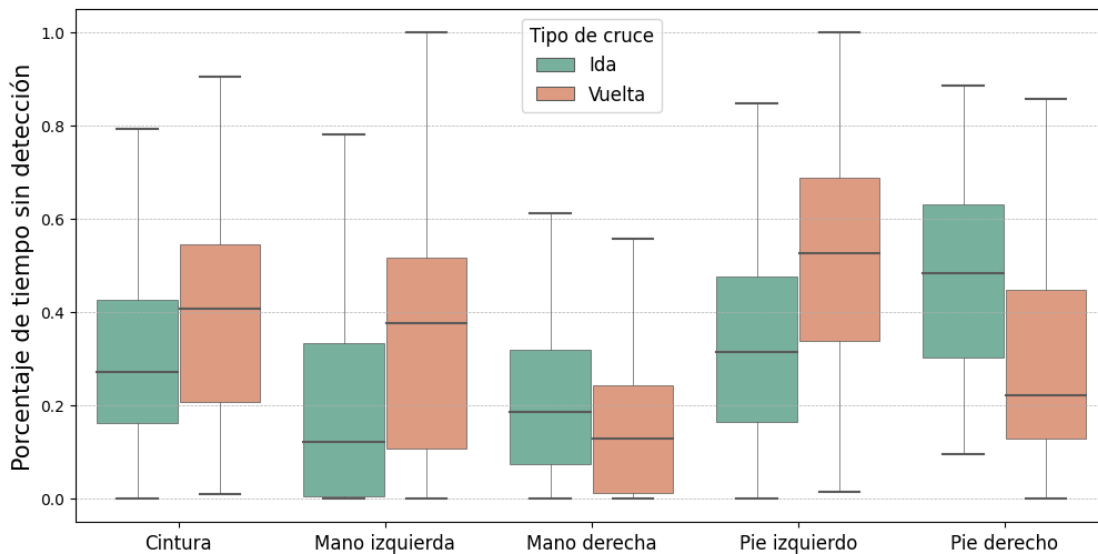


Ilustración 23. Comparativa en el porcentaje de tiempo con mala detección para cada *Rigid Body*.

A nivel visual, se observa como la distribución del porcentaje de tiempo con mala detección es superior en el caso de los pies frente al resto de *Rigid Bodies*. La configuración de cámaras *Optitrack* utilizada presenta dificultades para detectar marcadores a nivel del suelo. Una posible explicación es la aparición de reflejos que introducen ruido en las señales capturadas por las cámaras. Además, la distancia en altura entre las cámaras y los marcadores es otro factor para considerar, ya que la señal puede degradarse y no alcanzar adecuadamente los marcadores.

Mediante la prueba de *Mann Whitney* [37] se verifica la hipótesis alternativa (H1) de que el porcentaje de tiempo con mala detección es superior en el caso de los pies que para el resto de *Rigid Bodies*. El p-valor obtenido (**6.771e-21***) indica que hay evidencia estadísticamente significativa para determinar que el porcentaje de tiempo con mala detección es mayor en el caso de los pies.

De manera adicional, la visibilidad que tienen las cámaras de todos los *Rigid Bodies* empeora de manera significativa durante la vuelta al punto de partida, a excepción de la mano y el pie derecho. En las extremidades del lado izquierdo, la detección es óptima durante la ida hasta el final de la sala, mientras que para las extremidades del lado derecho es mejor durante la vuelta al punto de partida.

Se utiliza la prueba de *Wilcoxon* [38] para comparar los resultados obtenidos a lo largo de la ida y la vuelta. La hipótesis alternativa (H1) es el porcentaje de tiempo sin detección aumenta durante la vuelta al punto de partida. Los resultados se muestran en la Tabla 9.

Rigid Body	Mediana porcentaje tiempo de detección del de sin		Prueba de Wilcoxon	
	Ida	Vuelta	Estadístico	p-valor
Cintura	27.14 %	40.88 %	2393	3.5459e-5*
Mano derecha	18.55 %	12.84 %	1094	0.9598
Mano izquierda	12.31 %	37.72 %	2473	1.5648e-8*
Pie derecho	48.33 %	22.31 %	440	0.9999
Pie izquierdo	31.40 %	52.70 %	2811	8.9399e-10*

Tabla 9. Resultados de la prueba de Wilcoxon.

Los *p-valores* obtenidos confirman que las extremidades del lado izquierdo del cuerpo se detectan peor durante la vuelta al punto de partida en la sala de ensayo. En las extremidades del lado derecho ocurre lo contrario. Esto podría indicar que, por lo general, las cámaras del lado izquierdo de la sala tienen una mejor visibilidad que las del lado derecho. En el montaje implementado en el CEDINT, las cámaras del lado izquierdo se colocan a una altura ligeramente mayor que las del lado opuesto. Al tratarse de un estudio observacional, no se puede inferir una causalidad directa entre la altura de las cámaras y la capacidad de detección de los sensores, pero este podría ser un punto de partida para futuros experimentos. Otros factores no considerados en este Trabajo de Fin de Máster (como la orientación de las cámaras, las características de las lentes, o el propio nivel de batería de los sensores durante los ensayos), podrían estar afectando a los resultados.

El hecho de que la detección de la pelvis empeore durante la vuelta al punto de partida podría ser consecuencia de los posibles movimientos bruscos que el usuario realiza cuando aparece el vehículo (al final de la ida), y que provoca que las cámaras dejen de localizar los sensores. Esta localización tardaría unos segundos en recuperarse, provocando que los datos de la vuelta al punto de partida tengan generalmente una menor fiabilidad. De nuevo, al tratarse de un estudio observacional no pueden extraerse conclusiones definitivas respecto a la causalidad de la mala detección. En la Ilustración 24 se muestra una comparativa del porcentaje de tiempo con mala detección de la pelvis a lo largo de la ida y la vuelta en cada escenario. Se observa que en todos los casos (también en el escenario de Familiarización, donde no se realiza una simulación de atropello), este porcentaje de tiempo con mala detección aumenta durante la vuelta al punto de partida, siendo este incremento más marcado en el escenario de Emilio Muñoz. Esto podría atribuirse a que habitualmente las vueltas no forman parte del ensayo, sino que simplemente el usuario vuelve al punto de partida para quitarse los sensores.

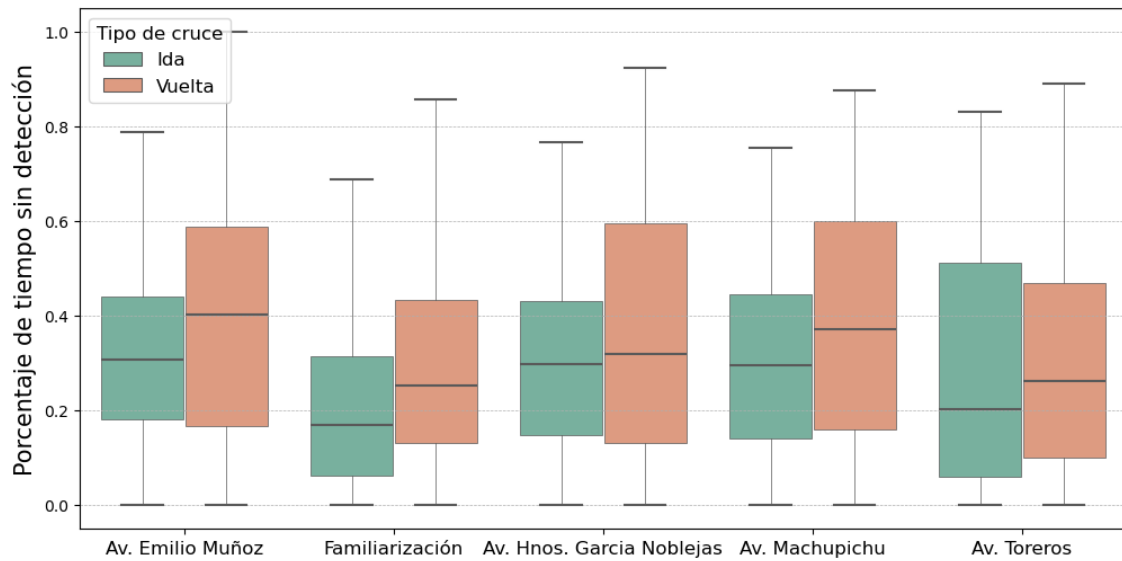


Ilustración 24. Comparativa en el incremento en el porcentaje de tiempo con mala detección de la pelvis para cada escenario.

Utilizando de nuevo la prueba de Wilcoxon se rechaza la hipótesis nula (H_0) de que las distribuciones son iguales en cada escenario; y se verifica la hipótesis alternativa (H_1) de que el porcentaje de tiempo con mala detección aumenta durante la vuelta al punto de partida. Los resultados (Ver Tabla 10) favorecen H_1 en todos los casos a excepción del escenario de la *Avenida de los Toreros*, aunque en este caso el p-valor obtenido es muy cercano al nivel de significancia establecido (0.05).

Los resultados obtenidos parecen indicar que existen otros factores no relacionados con la simulación del atropello que provocan una peor detección de la pelvis durante la vuelta al punto de partida. Algunos de estos factores podrían estar relacionados con el movimiento de los sensores durante la prueba, las características del montaje de cámaras u otras variables no consideradas en este Trabajo de Fin de Máster.

Escenario	Mediana del porcentaje de tiempo sin detección		Prueba de Wilcoxon	
	Ida	Vuelta	Estadístico	p-valor
Av. Emilio Muñoz	0.3083	0.4035	1757	0.0230*
Familiarización	0.1684	0.2532	2390	0.0007*
Av. Hnos García Noblejas	0.2980	0.3198	1758	0.0229*
Av. Machupichu	0.2957	0.3721	1934	0.0006*
Av. Toreros	0.2024	0.2621	2104	0.077

Tabla 10. Resultados de la prueba de Wilcoxon.

4.1.2 Relación de la coherencia biomecánica con el nivel de visibilidad de cada cruce.

En este apartado presento los resultados obtenidos empleando la metodología descrita en el apartado 3.2.4.2. Para ello se clasifican las trayectorias en niveles

de visibilidad en función del porcentaje de tiempo con mala detección. Los mencionados niveles de visibilidad cuales se describen en la Tabla 11:

Nivel de visibilidad	Nombre	Descripción
Nivel 2	Buena visibilidad	Porcentaje de tiempo sin detección inferior al 25%.
Nivel 1	Visibilidad normal	Porcentaje de tiempo sin detección en el intervalo 25 - 50%
Nivel 0	Mala visibilidad	Porcentaje de tiempo sin detección superior al 50%

Tabla 11. Descripción de los niveles de visibilidad establecidos.

En la Ilustración 25 se muestra el número de trayectorias asignadas a cada nivel de visibilidad para los pies y la pelvis. En el caso de los pies la mayor parte de tomas tiene un porcentaje de tiempo sin detección superior al 50%, mientras que en el caso de la pelvis el número de tomas con mala detección es más bajo.

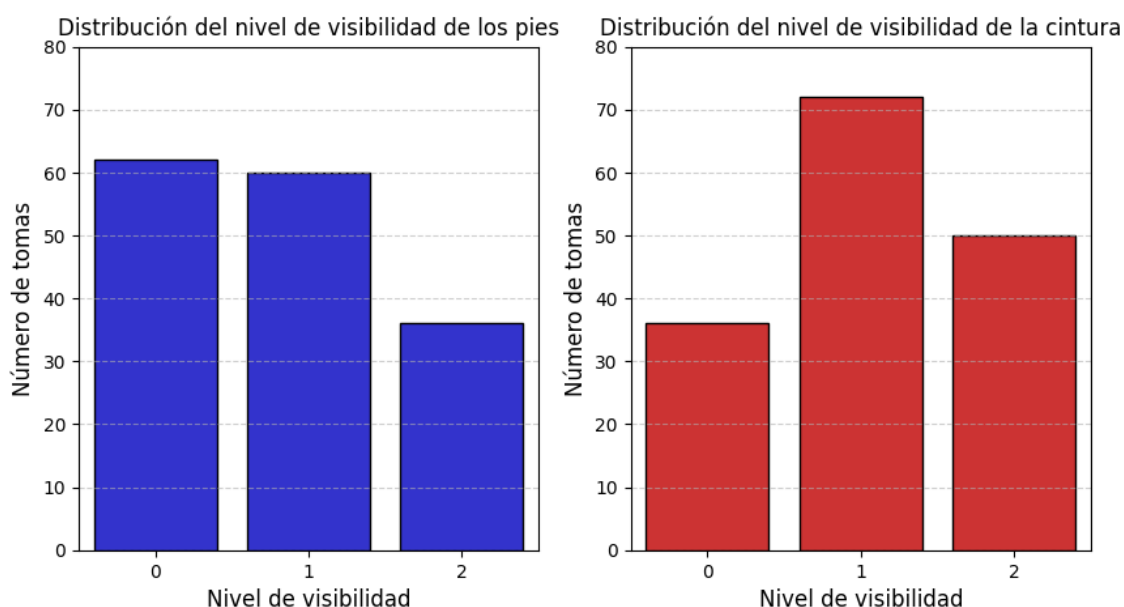


Ilustración 25. Distribución de las trayectorias según el nivel de visibilidad.

4.1.2.1 Resultados obtenidos empleando LCSS.

Esta sección incluye un resumen de los resultados obtenidos al evaluar la coherencia biomecánica de cada cruce según el nivel de visibilidad. La coherencia biomecánica se estima a través de la longitud de la subsecuencia común más larga (LCSS) entre la cabeza y la cintura; o entre los pies.

4.1.2.1.1 Resultados obtenidos a través de las medidas originales

En este apartado, analizo la relación que tienen los valores de la longitud de la subsecuencia común más larga calculada empleando las medidas originales en función los niveles de visibilidad que se han definido en el apartado 4.1.2.

En la Ilustración 26 se muestra una comparativa de los resultados obtenidos en la longitud de la subsecuencia común más larga entre la cabeza y la cintura y entre los pies, según el nivel de visibilidad.

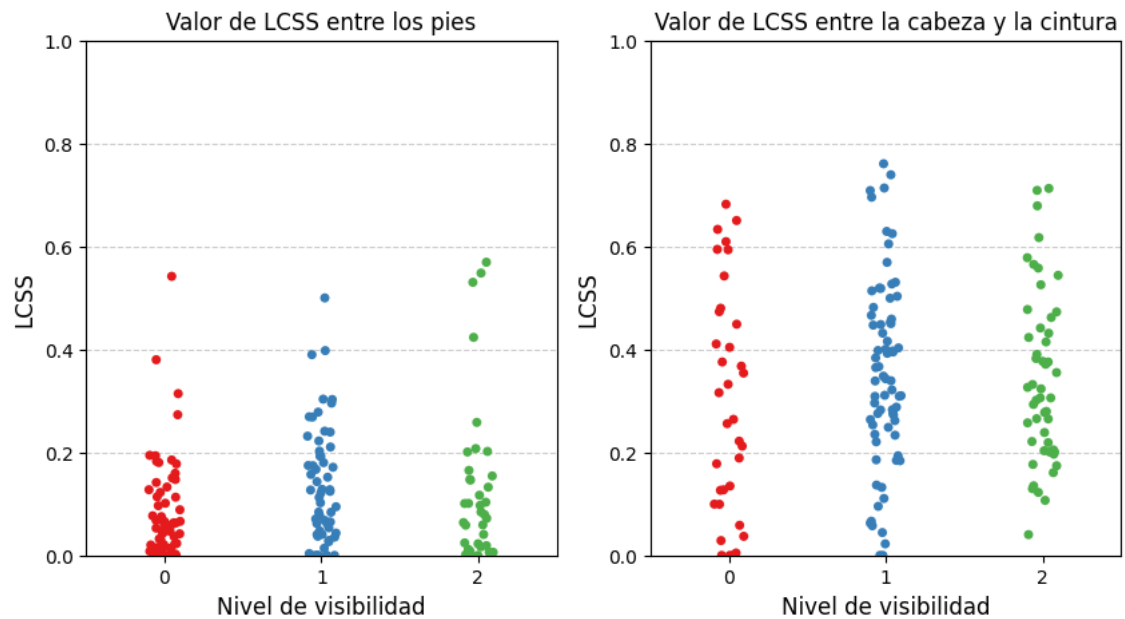


Ilustración 26. Comparativa en los valores de LCSS obtenidos entre los pies y entre la cabeza y la cintura.

La LCSS es una medida sensible a los valores nulos, que se producen cuando las cámaras no son capaces de detectar un elemento determinado.

En primer lugar, se observa que los valores son notablemente superiores al comparar las trayectorias de la cabeza y la cintura. Los mencionados elementos tienen menos grados de libertad de movimiento y por tanto la probabilidad de obtener trayectorias similares en el caso de la pelvis y la cabeza es muy superior. De manera adicional, la detección de la pelvis es superior a la de los pies (Ver apartado 4.1.1), lo que podría estar contribuyendo también a estos resultados.

Utilizando la prueba de Mann Whitney se pretende verificar la hipótesis alternativa (H1) de que la corrección biomecánica de las medidas es superior para un nivel de visibilidad 1 o 2 frente a un nivel 0. Los resultados se presentan en la Tabla 10.

Métrica	Valores medianos			Test de Mann-Whitney-U	
	Nivel 0	Nivel 1	Nivel 2	Estadístico	p-valor
LCSS (entre pies)	0.0603	0.1126	0.0090	3658.0	0.0075*
LCSS (cabeza-cintura)	0.4365	0.4835	0.4328	9913.5	0.0487*

Tabla 10. Resultados de la prueba de Mann Whitney al comparar los niveles 1 y 2 con el nivel 0.

Tomando un nivel de significancia de 0.05, se han obtenido resultados significativos en ambos casos. Los p-valores obtenidos indican que, en las trayectorias con un porcentaje de tiempo con mala detección superior al 50%, las medidas de posición estimadas por el sistema de cámaras presentan más incoherencias. Esto se podría atribuir a la presencia tanto de valores nulos (el sistema no es capaz de localizar un sensor) como de valores erróneos.

4.1.2.1.2 Comparativa con los resultados obtenidos utilizando las medidas interpoladas.

Siguiendo el mismo procedimiento que se llevó a cabo en el apartado anterior, se obtiene la distribución de la longitud de la subsecuencia común más larga en función del nivel de visibilidad de cada cruce, esta vez empleando las

medidas interpoladas por la aplicación. El resultado obtenido al comparar con las medidas originales se muestra en la Ilustración 27.

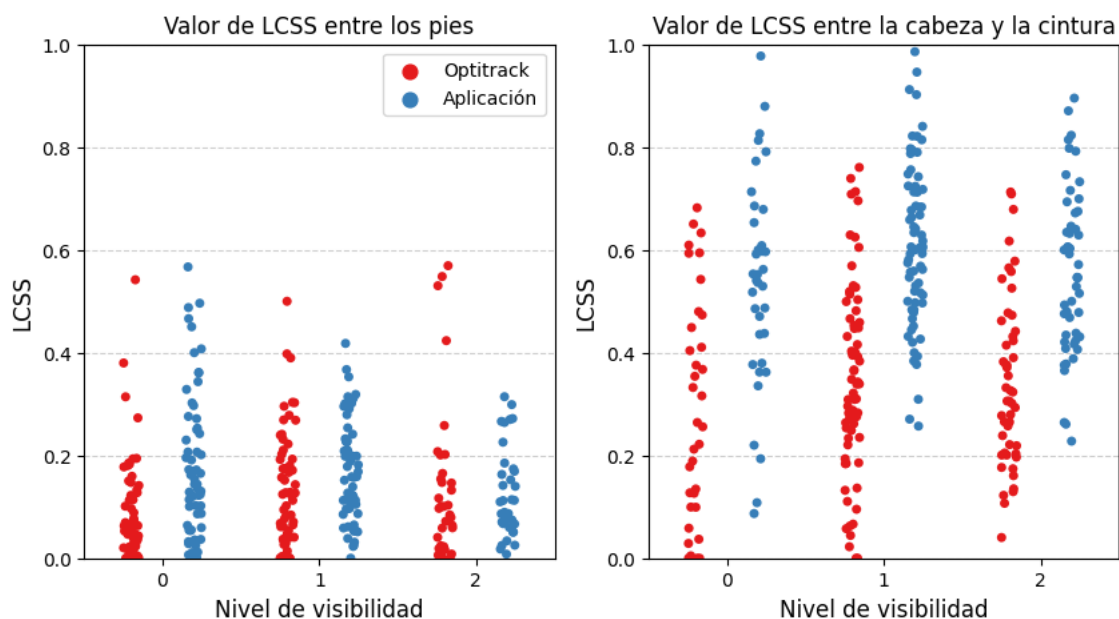


Ilustración 27. Comparativa en los valores de LCSS entre las medidas originales y las medidas interpoladas.

La interpolación realizada sobre las medidas de posición mejora significativamente la longitud de la subsecuencia común más larga, especialmente entre la cabeza y la cintura. Esto hace que, para un nivel de visibilidad bajo, el valor mediano de LCSS sea superior al obtenido para un nivel de visibilidad alto. Para verificar que existe esta mejora, he utilizado la prueba de *Wilcoxon* para muestras pareadas. La hipótesis alternativa (H_1) es que con las medidas interpoladas se obtiene un valor de LCSS más alto que empleando las medidas originales. Los resultados se muestran en la Tabla 11.

Nivel de visibilidad	Métrica	Valores medianos		Test de Mann-Whitney-U	
		Medidas interpoladas	Medidas originales	Estadístico	p-valor
Nivel 0	LCSS (entre pies)	0.0603	0.1440	1558.0	2.2816e-5*
	LCSS (cabeza-cintura)	0.2599	0.5463	603.0	1.9790e-6*
Nivel 1	LCSS (entre pies)	0.1168	0.1586	1089.0	0.1001
	LCSS (cabeza-cintura)	0.3410	0.6034	2537.0	3.3691e-12*
Nivel 2	LCSS (entre pies)	0.0906	0.0995	375.0	0.2594
	LCSS (cabeza-cintura)	0.3146	0.5464	1243.0	2.4558e-12*

Tabla 11. Resultados de la prueba de *Wilcoxon* al comparar las medidas interpoladas con las originales para cada nivel de visibilidad.

Los p-valores obtenidos indican que la mejora conseguida por el algoritmo de interpolación es estadísticamente significativa, especialmente en el incremento de la LCSS calculada entre las trayectorias de la cabeza y la cintura. En el caso de las trayectorias de los pies, se han obtenido mejoras significativas en las

trayectorias con un nivel de visibilidad bajo, pero no en el resto. De manera adicional, en el caso de la similitud entre la cabeza y la cintura, los resultados son más significativos con los niveles de visibilidad más altos.

Todo ello indica que el algoritmo de interpolación mejora la coherencia biomecánica de las medidas posicionales especialmente cuando se aplica sobre la cintura. Al aplicar el algoritmo para reconstruir la posición de los pies, se obtienen resultados más coherentes que los obtenidos empleando las medidas originales cuando el nivel de visibilidad es bajo. Con un mayor nivel de visibilidad esto no ocurre, lo que podría atribuirse a la mayor libertad de movimientos de los pies. Esta mayor libertad de movimientos podría estar dificultando la estimación de su posicionamiento de forma óptima, frente al movimiento de la pelvis y la cabeza, el cual es más lineal y predecible.

4.1.2.2 Resultados obtenidos para las medidas interpoladas empleando DTW.

En esta sección se muestran los valores de distancia calculada mediante el algoritmo de deformación temporal dinámica. En este caso se emplean únicamente las medidas interpoladas para obtener un valor de distancia definitivo para cada trayectoria. No se realiza la comparativa con las medidas recogidas por *Optitrack* ya que estas incluyen valores nulos que impiden realizar el cálculo utilizando esta métrica. Los valores obtenidos según el nivel de visibilidad se muestran en la Ilustración 28.

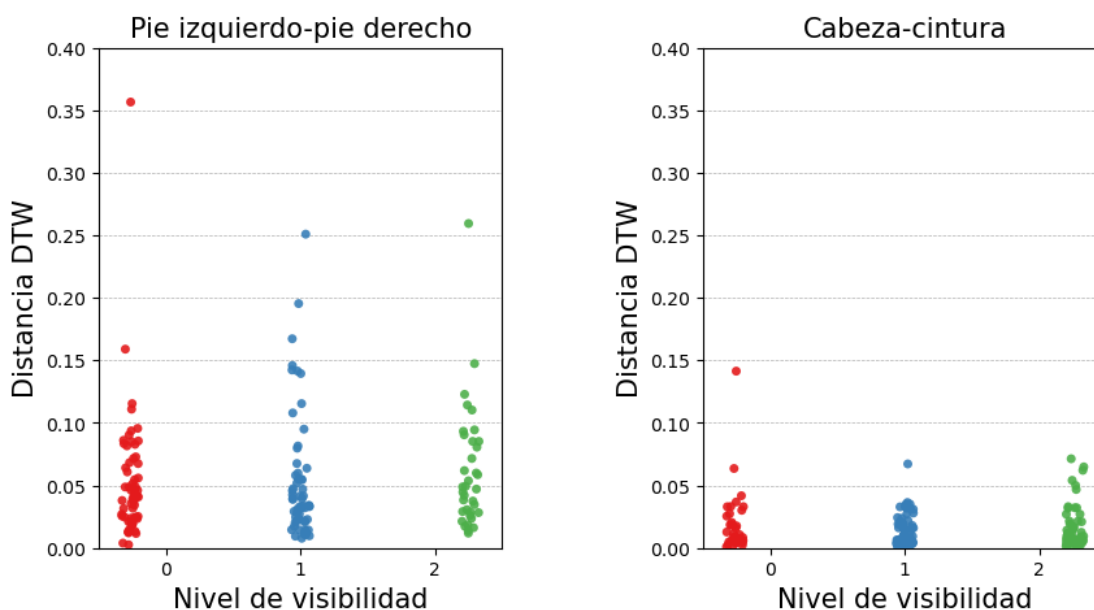


Ilustración 28. Comparativa en la distancia DTW entre la cabeza y la cintura y entre pies, según el nivel de visibilidad.

En todos los casos se obtienen resultados muy cercanos a 0, lo que indicaría que las trayectorias no tienen una gran disparidad. La semántica de esta métrica puede interpretarse como la *disparidad media entre dos señales*. En todos los casos, los valores medianos son inferiores a 0.05 m (Ver Tabla 12).

Cabe destacar que la variabilidad de la distribución es notablemente superior en el caso de la disparidad entre las trayectorias de los pies frente a las de la cabeza y la cintura.

Se realiza la prueba de Kruskal Wallis [37] para verificar si existen diferencias significativas en la disparidad en función del nivel de visibilidad. Los resultados se muestran en la Tabla 12. No se obtienen resultados significativos en ninguno

de los casos, siendo el *p-valor* aún más bajo en el caso de los pies frente a la cintura. Esto podría indicar que el nivel de detección afecta más a la similitud entre pies que a la similitud entre las trayectorias de la cabeza y la cintura. Sin embargo, los resultados de la prueba no ofrecen evidencia estadística significativa en ninguno de los casos.

Métrica	Valores medianos			Test de Kruskal Wallis	
	Nivel 0	Nivel 1	Nivel 2	Estadístico	p-valor
DTW (entre pies)	0.0416	0.0339	0.0478	2.2224	0.3234
DTW (cabeza-cintura)	0.0109	0.0066	0.0098	3.9343	0.1399

Tabla 12.Resultados de la prueba de Kruskal-Wallis.

Llama la atención una trayectoria anómala que se aleja mucho del centro de la distribución, obteniendo un valor de disparidad entre las trayectorias de los pies de casi 40 cm. Además, esta trayectoria es la misma que obtiene el valor más alto de disparidad entre la cabeza y la cintura (14 cm), incluyéndose en el grupo de trayectorias con un mal nivel de visibilidad. Se corresponde a la vuelta al punto de partida de un ensayo ambientado en el escenario de Emilio Muñoz.

A lo largo de esta trayectoria el usuario se queda inmóvil al final de la sala (Ver Ilustración 29). Dado que en los extremos de la sala la detección es peor, esto podría haber provocado que los datos de esta toma no sean fiables.

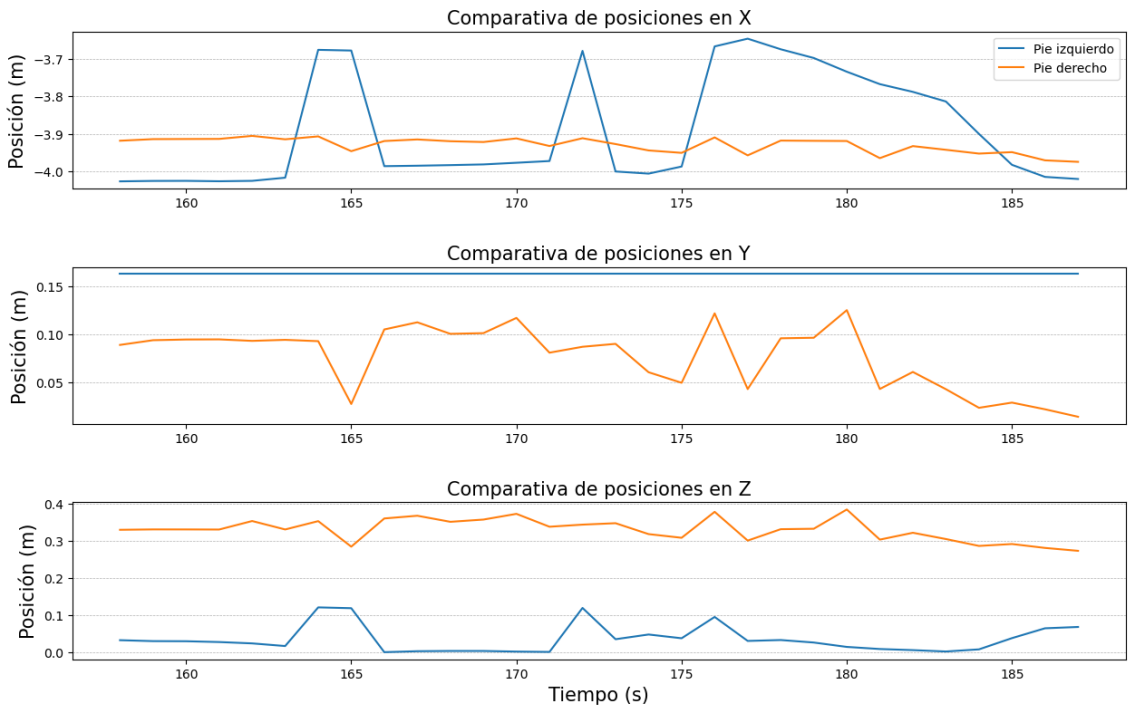


Ilustración 29. Trayectoria de los pies del ensayo atípico

4.1.3 Conclusiones de la evaluación de las medidas y del algoritmo de interpolación.

En los apartados 4.1.2.1 y 4.1.2.2 he utilizado dos métricas para evaluar la validez de los datos de partida a partir de la disparidad entre las trayectorias de los pies y de la cabeza y la cintura de los sujetos. La Longitud de la Subsecuencia Común más Larga permite ofrecer una comparativa con las medidas antes y después de aplicar la interpolación. Por otra parte, mediante el algoritmo de Deformación Temporal Dinámica se obtiene un valor de la disparidad obtenida con las medidas interpoladas.

Los resultados de esta evaluación podrían indicar que el mayor número de grados de libertad de movimiento que los miembros inferiores presentan afecta tanto al valor de disparidad obtenido como a la capacidad del algoritmo para estimar su posición. En el caso de las trayectorias con un mayor porcentaje de tiempo con mala detección se obtiene una mejora significativa en todos los casos gracias al mencionado algoritmo.

Por otra parte, los valores de disparidad obtenidos se distribuyen en torno a valores cercanos a 0, lo que parece indicar que de forma general las medidas son coherentes. En los siguientes apartados se utilizan únicamente los datos correspondientes a la ida, ya que durante la vuelta la detección de las cámaras empeora, como se demuestra en el apartado 4.1.1. De manera adicional, es durante la ida cuando se realiza la simulación del atropello, por lo que es donde se encuentra la información más interesante de cara a la identificación de patrones de movimiento en situaciones de riesgo de tráfico urbano.

4.2 Generación de una base de datos de características.

El proceso seguido para la consecución de este objetivo se desarrolla en su mayor parte en el apartado 3.2.4.3.

Parámetro de la marcha	Descripción
<i>times_separating_left_hand (vertical)</i>	Número de veces que el usuario separa la mano izquierda del cuerpo en el plano vertical.
<i>times_separating_right_hand (vertical)</i>	Número de veces que el usuario separa la mano derecha del cuerpo.
<i>times_separating_left_hand (horizontal)</i>	Número de veces que el usuario separa la mano izquierda del cuerpo en el plano horizontal.
<i>times_separating_right_hand (horizontal)</i>	Número de veces que el usuario separa la mano derecha del cuerpo en el plano horizontal.
<i>proportion_look_right</i>	Proporción de tiempo en que el usuario mira a la derecha.
<i>proportion_look_left</i>	Proporción de tiempo en que el usuario mira a la izquierda.
<i>proportion_look_up</i>	Proporción de tiempo en que el usuario mira hacia arriba.
<i>proportion_look_down</i>	Proporción de tiempo en que el usuario mira hacia abajo.
<i>proportion_attending</i>	Proporción de tiempo atencional.
<i>stop_retreat</i>	Número de veces que el usuario se para o vuelve hacia atrás a lo largo de la trayectoria.
<i>step_duration</i>	Duración de un paso. Para su estimación se emplea la señal obtenida a partir del pie izquierdo ya que este se detecta mejor durante la ida que el pie derecho.
<i>step_length</i>	Longitud median de un paso.
<i>walking_width</i>	Separación mediana de los pies.
<i>walking_angle</i>	Valor mediano del ángulo formado por ambos pies.
<i>max_deviation</i>	Máxima desviación de la línea base de la trayectoria en el eje z.
<i>speed</i>	Velocidad mediana del usuario.
<i>height</i>	Altura del usuario

Tabla 13. Características que modelan el movimiento.

Para cada cruce de la sala de ensayo (ida) se calculan los parámetros definidos en la Tabla 13. Se emplean únicamente los datos de la ida por que contienen la información más relevante, tal como se describe en el apartado anterior. Además, la visibilidad que las cámaras tienen de los sensores empeora de manera general durante la vuelta al punto de partida (Ver apartado 4.1.1)

4.3 Identificación de patrones de movimiento.

En esta sección explico el proceso seguido para preparar los datos de cara a obtener perfiles de movimiento. A continuación, justifico la selección del número de perfiles a partir del conjunto de datos de características. Posteriormente describo los resultados obtenidos en el agrupamiento, y por último presento los parámetros que definen cada patrón de movimiento.

4.3.1 Reducción de la dimensionalidad.

El cálculo de la distancia entre instancias en el espacio de características puede verse afectado si la dimensionalidad del conjunto de datos es muy grande. En este sentido, el análisis de componentes principales puede ser útil para identificar aquellas variables que más influyen en la separabilidad de los datos.

El porcentaje de contribución de cada una de las variables sobre el espacio de componentes principales se muestra en la Ilustración 30. En rojo, se visualiza la contribución acumulada. Las variables *step_duration*, *walking_width*, *proportion_look_up* y *proportion_look_down* apenas aportan varianza al conjunto de datos. El porcentaje de contribución (en conjunto) de las mencionadas características a la variabilidad de los datos no supera el 5%. La variable que más contribuye es *times_separating_right_hand (vertical)*.

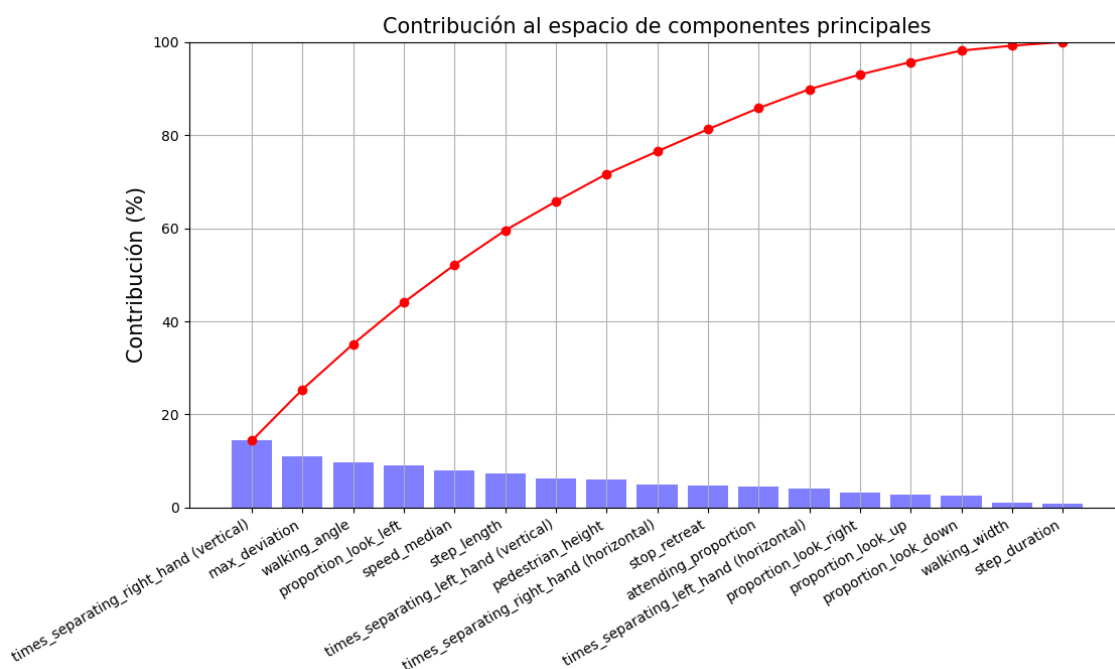


Ilustración 30. Porcentaje de contribución de cada variable sobre el espacio de componentes principales.

Por otra parte, la matriz de correlación de todo el conjunto de datos se presenta en la Ilustración 31. Se observa que hay una dependencia lineal fuerte entre las variables *step_length* (longitud del paso) y *speed_median* (valor mediano de la velocidad). Existe también una correlación negativa importante entre las variables *proportion_look_up* con *proportion_look_down*; y entre *proportion_look_left* con *proportion_look_right*, como cabría esperar. Finalmente, la anchura de la marcha (*walking_width*) está muy positivamente correlada con la proporción del tiempo que el usuario mira a la derecha (*proportion_look_right*).

Considerando lo anterior, las variables *proportion_look_up*, *proportion_look_down*, *walking_width* y *step_duration* no se considerarán en el resto del análisis. Eliminar estas variables permite reducir la dimensionalidad de los datos, además de retirar algunas de las dependencias lineales más fuertes. Todo esto se consigue a la vez que se mantiene el 95% de la varianza del conjunto de datos.

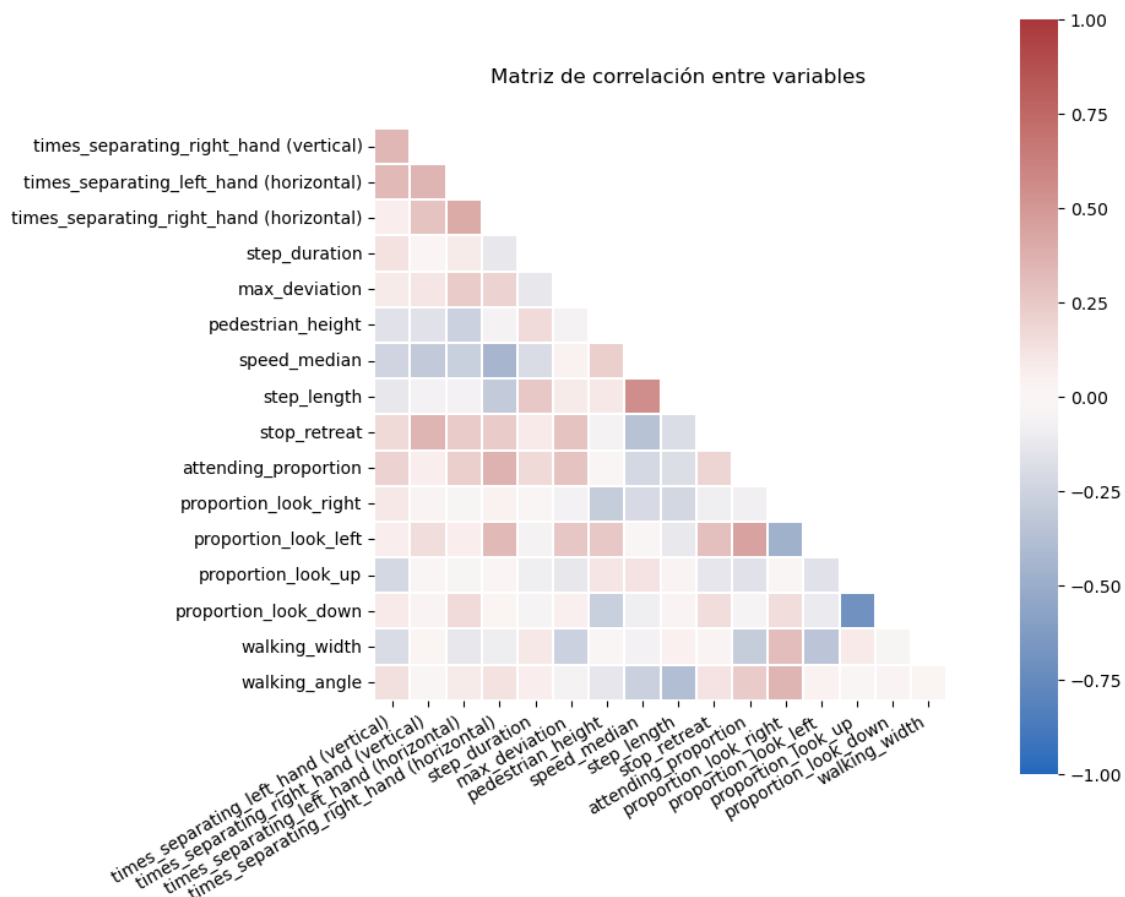


Ilustración 31. Matriz de correlación.

4.3.2 Selección del número de clústeres.

En la Ilustración 32 se incluye el dendrograma correspondiente al algoritmo de clustering aglomerativo, que permite visualizar en el eje vertical cómo se incrementa la distancia entre clústeres a medida que el algoritmo agrupa las instancias en cada iteración. En el dendrograma, se observa como en la primera iteración, el algoritmo asigna cada instancia a un clúster diferente. Sucesivamente, las instancias se van agrupando hasta formar un solo grupo.

Observando como evoluciona la distancia entre grupos (eje vertical) a medida que estos se forman es posible tomar una idea de las características del conjunto de datos.

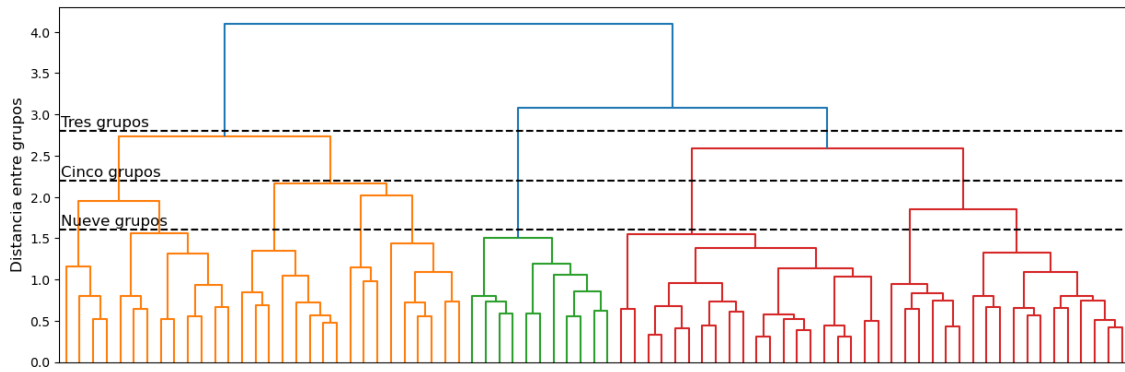


Ilustración 32. Dendrograma resultante del algoritmo de clustering aglomerativo.

Al inicio, todas las instancias (ensayos) se asignan a un clúster diferente. A medida que se incrementa la distancia entre clústeres, estas instancias se van agrupando. Para una distancia entre clústeres de 1.6 ya hay nueve grupos diferenciados. Con una distancia entre clústeres de 2.2 se obtienen cinco grupos. Si se incrementa este umbral de distancia progresivamente, dos de ellos se unen formando el grupo rojo. Con una distancia entre grupos de 3.2 se obtienen tres grupos: dos de ellos (*rojo y amarillo*) formados a través de la unión de los clústeres preexistentes; y una última agrupación (*verde*) que contiene únicamente once trayectorias.

Por otra parte, empleando las métricas seleccionadas, se evalúa el desempeño del algoritmo de clustering aglomerativo para cada número de clústeres. Los resultados se incluyen en la Ilustración 33.

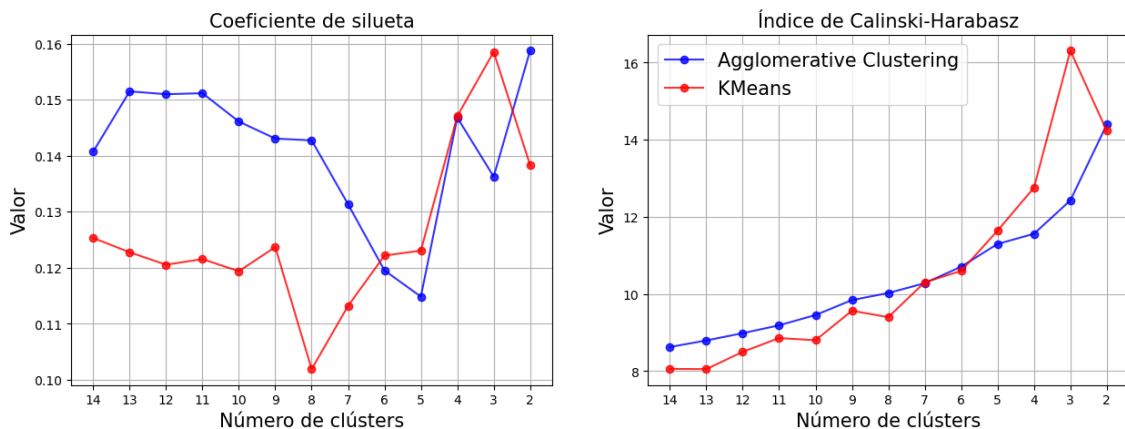


Ilustración 33. Resultados de la evaluación del algoritmo de clustering aglomerativo junto al método K-means.

En el caso del algoritmo jerárquico(azul), el máximo valor del coeficiente de silueta y del índice de *Calinski-Harabasz* se obtiene para dos grupos. El coeficiente de silueta se reduce drásticamente al pasar de dos a tres grupos, pero vuelve a aumentar al pasar a cuatro. Si nos fijamos en el dendrograma se observa que, al pasar de tres a dos grupos, el valor de distancia entre grupos apenas aumenta, lo que indicaría que con dos grupos se obtienen clústeres mejor definidos empleando el algoritmo jerárquico. Utilizando el método K-means, los valores más altos se obtienen para tres grupos. Mientras que si nos fijamos únicamente en la evaluación del algoritmo de clustering jerárquico se seleccionaría un número de clústeres de dos; al evaluar el algoritmo k-means los resultados indican que el valor óptimo es de tres.

Esto demuestra la ventaja de los algoritmos particionales sobre los algoritmos jerárquicos, al no permitir estos últimos la reasignación de instancias durante el agrupamiento.

En la Tabla 14, se muestran los valores del coeficiente de silueta y el índice de Calinski-Harabasz obtenidos mediante el algoritmo jerárquico junto a los obtenidos mediante *k-means* utilizando el correspondiente valor del parámetro *k*. Los valores marcados en negrita son los valores óptimos.

Métrica	Número de clústeres	Resultado (C. Aglomerativo)	Resultado (K-means)
Índice de Calinski-Harabasz	5	11.2991	11.6433
	4	11.5625	12.7582
	3	12.4285	16.3129
	2	14.3606	14.2311
Coefficiente de Silueta	5	0.1148	0.1231
	4	0.1468	0.1476
	3	0.1363	0.1585
	2	0.1582	0.1383

Tabla 14. Resultados numéricos de la evaluación de los algoritmos de clustering.

Otra forma de seleccionar el número óptimo de clústeres empleando el algoritmo *k-means* sería a partir del gráfico de sedimentación. Este gráfico permite visualizar cómo se reduce la inercia (suma de los cuadrados de las distancias intra-clúster) a medida que se incrementa el número de grupos. En la Ilustración 34 se muestra esta reducción. Empleando el método del *codo*, se obtiene que el número óptimo de grupos es de tres.



Ilustración 34. Gráfico de sedimentación del algoritmo de *k-means*.

En el apartado siguiente se analizan las características de los grupos obtenidos empleando el algoritmo *k-means* para tres clústeres. Los valores del coeficiente de silueta y el índice de Calinski-Harabasz obtenidos mediante K-Means han sido, respectivamente 0.1585 y 16.3129, superiores en ambos casos a los obtenidos mediante el algoritmo de clustering jerárquico. Esto indica que los clústeres más homogéneos y mejor diferenciados se obtienen con este algoritmo y estos parámetros.

4.3.3 Resultados del agrupamiento.

En biplot de la Ilustración 35 se muestra la proyección de las instancias en el plano formado por las dos componentes principales del conjunto de datos. El porcentaje de varianza explicado por estas dos componentes principales es del 40% del total.

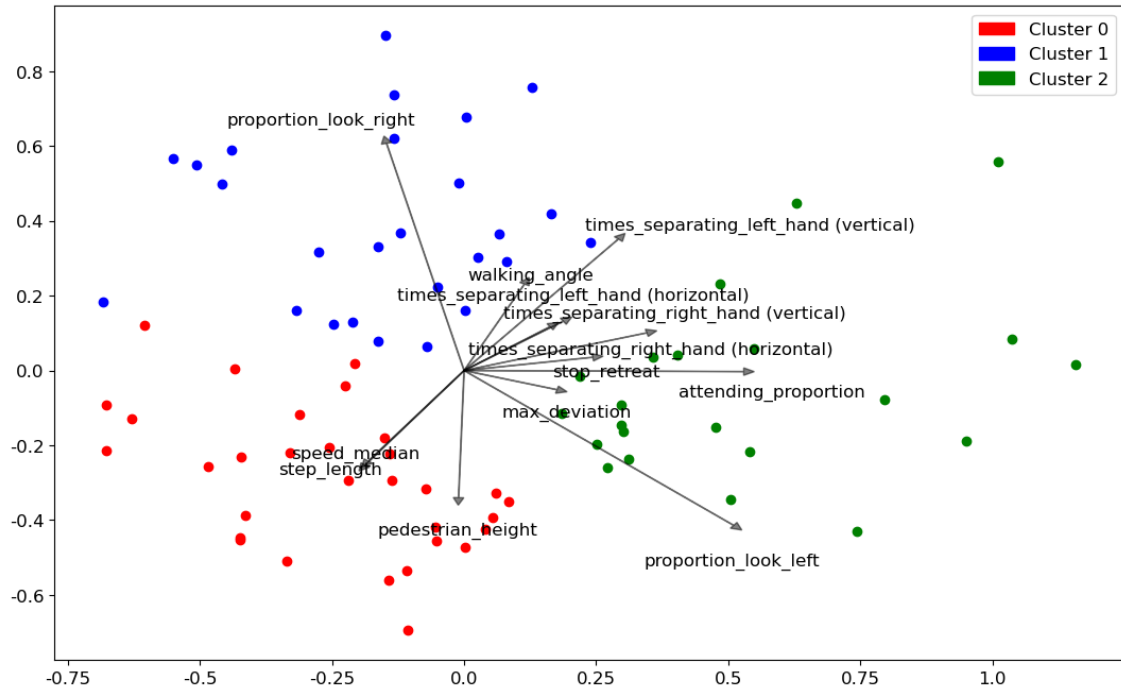


Ilustración 35. Proyección de los tres clústeres sobre el plano de componentes principales.

A partir del análisis visual, se puede determinar que el **grupo 0** (rojo), incluye las trayectorias de los individuos con mayor altura y que caminan con mayor velocidad y longitud del paso.

El grupo 1 (azul), contiene aquellas trayectorias en las que el sujeto pasa más tiempo mirando a la derecha. **El grupo 2** (verde), engloba las trayectorias en las que el usuario tiene mayor porcentaje de tiempo atencional y que pasan más tiempo mirando a la izquierda. Además, en este último grupo se incluyen los usuarios que tienden a separar más las manos del cuerpo.

La homogeneidad de cada clúster, calculada como la suma de los cuadrados de las instancias al centroeide para cada grupo se muestra en la Ilustración 36. Se observa que el grupo 2 es menos homogéneo que los grupos 0 y 1.

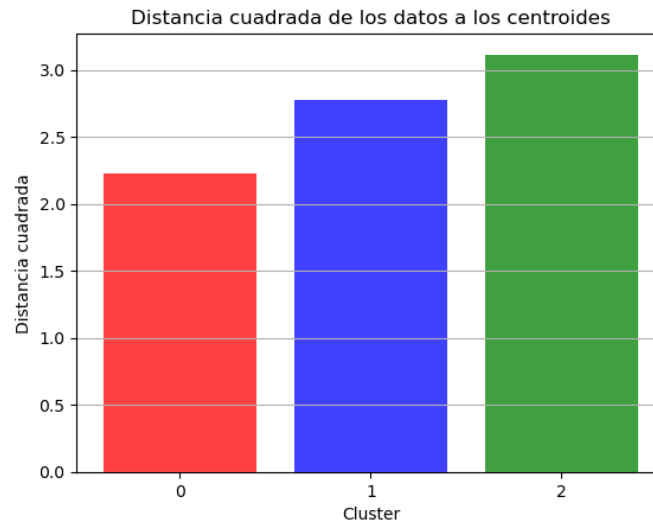


Ilustración 36. Homogeneidad calculada como la suma de los cuadrados de las distancias al centroide.

He utilizado la prueba de Kruskal-Wallis [37], [38] para identificar aquellas variables en las que las diferencias entre los grupos son estadísticamente significativas. En la Tabla 16 se incluyen los resultados.

Estableciendo un nivel de significancia de 0.05, las variables marcadas con un asterisco son aquellas en las que se encuentran las mayores diferencias estadísticas entre grupos. En la Ilustración 37 se muestra una comparativa de la distribución de cada una de estas características para cada grupo.

Variable	Valores medianos			Resultado de la prueba de Kruskal-Wallis.	
	Clúster 0	Clúster 1	Clúster 2	Estadístico	p-valor
<i>times_separating_left_hand (horizontal)</i>	1.0	1.0	2.0	12.243	0.0021*
<i>times_separating_right_hand (horizontal)</i>	1.0	1.0	2.0	16.1948	0.0003*
<i>times_separating_left_hand (vertical)</i>	0.0	1.0	1.0	15.055	0.0003*
<i>times_separating_right_hand (vertical)</i>	0.0	1.0	1.0	5.6297	0.0569
<i>max_deviation</i>	0.2261	0.1837	0.3011	10.7439	0.0463*
<i>pedestrian_height</i>	1.61	1.52	1.54	8.8810	0.0112*
<i>speed_median</i>	0.8934	0.5134	0.4143	27.797	9.202e-07*
<i>step_length</i>	0.7063	0.3823	0.2988	26.3099	1.9368e-06*
<i>stop_retreat</i>	2.0	2.0	3.0	11.6845	0.0029*
<i>attending_proportion</i>	0.0	0.0764	0.43348	25.7298	2.5874e-6*
<i>proportion_look_right</i>	0.0000	0.03178	0.0000	48.2629	3.3099e-11*
<i>proportion_look_left</i>	0.5429	0.3047	0.8356	47.4474	3.6624e-10*
<i>walking_angle</i>	9.2085	27.7279	27.3915	18.2586	0.0001*

Tabla 15. Resultados de la prueba de Kruskal Wallis.

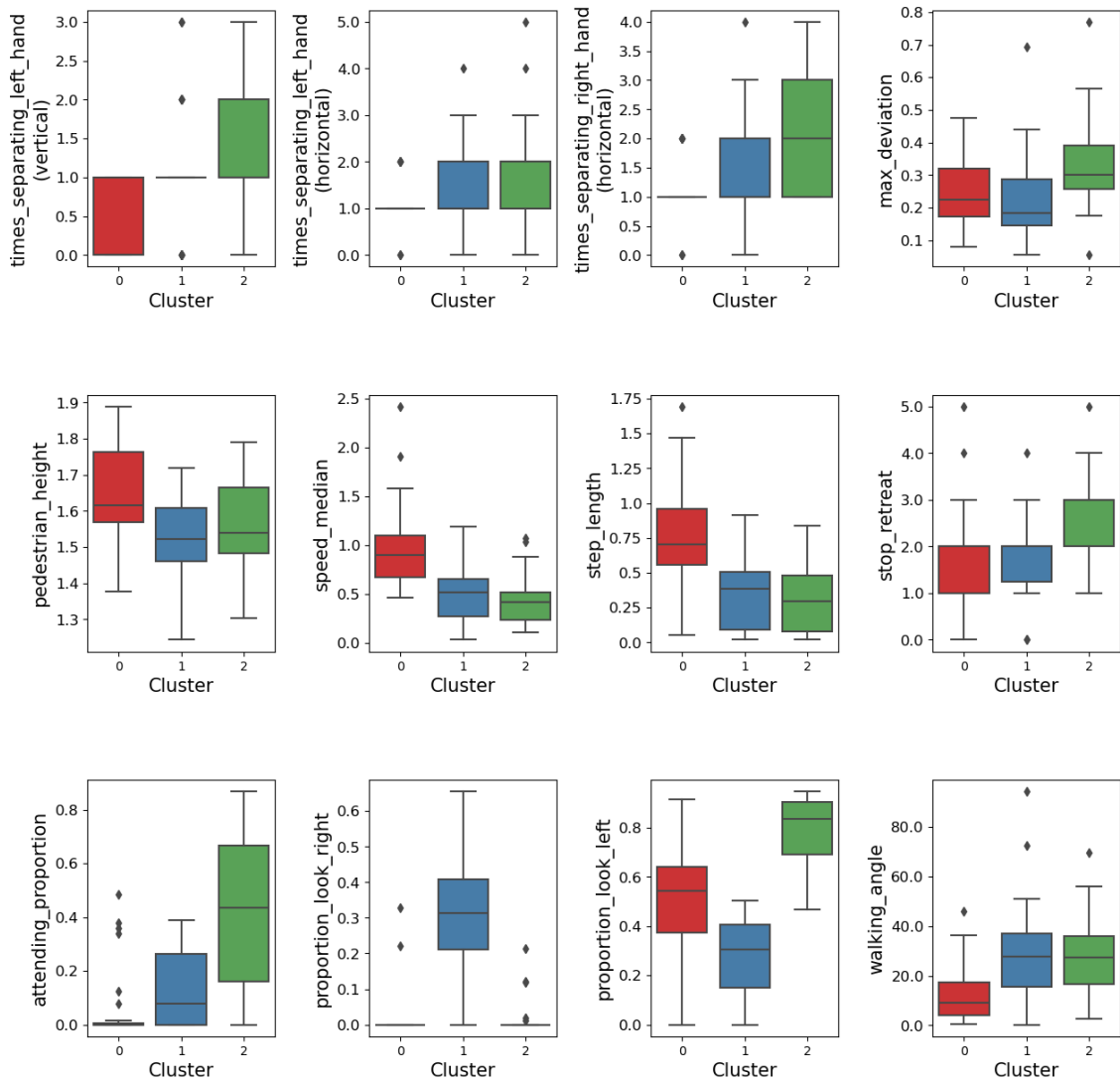


Ilustración 37. Comparativa en la distribución de cada una de las características para cada uno de los grupos.

El **grupo 2** es el menos homogéneo de los tres, con un valor de inercia intra-clúster de 11.53. En él se incluyen las trayectorias en las que el sujeto presenta un porcentaje de tiempo atencional mayor, y además son aquellas en las que el usuario pasa más tiempo mirando a la izquierda. Esto tiene sentido ya que en la mayoría de los escenarios donde tiene lugar la simulación del atropello, el vehículo aparece por el lado izquierdo del peatón. La longitud del paso, así como la velocidad, en este grupo es menor a los otros dos grupos. Además, los usuarios de este grupo son aquellos que tienden a separar más las manos del cuerpo y tienden más a pararse y retroceder.

El algoritmo engloba en el **grupo 0** a aquellos usuarios que caminan más rápido, y tienen menor porcentaje de tiempo atencional. Este grupo presenta una menor homogeneidad intra-clúster que el grupo 1.

El **grupo 1** contiene las trayectorias en las que el usuario pasa más tiempo mirando hacia la derecha. Este grupo es el más heterogéneo y por tanto el que más variabilidad presenta entre las instancias que lo componen. El porcentaje de tiempo atencional en este grupo es superior al grupo 0, pero inferior al grupo 2.

4.3.3.1 Parametrización de los grupos.

El resumen de las características de los patrones de movimiento obtenidos se incluye en la Tabla 16. Para cada variable continua he calculado la mediana y la desviación estándar, mientras que para las variables discretas se calcula la moda, el valor máximo y el mínimo.

Variable	Parámetro	Grupo 0	Grupo 1	Grupo 2
<i>times_separating_left_hand</i> (horizontal)	Moda	1	1	2
	Máximo	2	4	5
	Mínimo	0	0	0
<i>times_separating_right_hand</i> (horizontal)	Moda	1	1	1
	Máximo	2	4	4
	Mínimo	0	0	1
<i>times_separating_right_hand</i> (vertical)	Moda	0	1	1
	Máximo	1	3	3
	Mínimo	0	0	0
<i>max_deviation</i>	Mediana	0.2261	0.1836	0.3011
	Desv. estándar	0.1064	0.1420	0.1464
<i>pedestrian_height</i>	Mediana	1.6122	1.5214	1.5391
	Desv. estándar	0.1497	0.1203	0.1308
<i>speed_median</i>	Mediana	0.8934	0.5134	0.4142
	Desv. estándar	0.4354	0.3070	0.2855
<i>step_length</i>	Mediana	0.7062	0.3823	0.2988
	Desv. estándar	0.3421	0.2711	0.2482
<i>stop_retreat</i>	Moda	2	2	3
	Máximo	5	4	5
	Mínimo	0	0	1
<i>attending_proportion</i>	Mediana	0.0	0.0792	0.4335
	Máximo	0.1338	0.1364	0.3087
<i>proportion_look_right</i>	Mediana	0.0	0.3134	0.0
	Desv. estándar	0.0701	0.1832	0.055
<i>proportion_look_left</i>	Mediana	0.5432	0.3047	0.8352
	Desv. estándar	0.2490	0.1488	0.1339
<i>walking_angle</i>	Mediana	9.2086	27.7279	27.3915
	Desv. estándar	10.6571	20.6778	17.1583

Tabla 16. Parámetros calculados para cada grupo.

4.3.3.2 Análisis de los resultados obtenidos en función del escenario de Realidad Virtual.

En esta sección, se incluye en la Ilustración 38 una comparativa de la distribución de los grupos obtenidos para cada escenario modelado en la aplicación de Realidad Virtual.

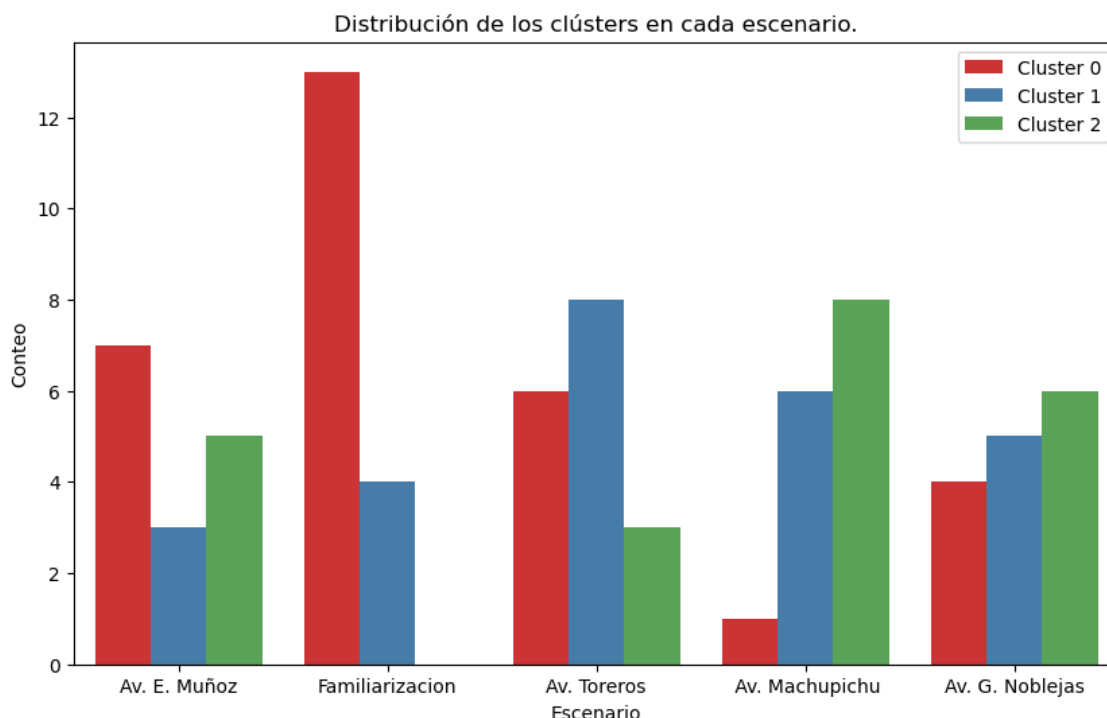


Ilustración 38. Distribución de los grupos para cada escenario.

Analizar la distribución de los patrones de movimiento para cada escenario permite caracterizar el movimiento de los individuos en función de las condiciones del entorno virtual.

- En el escenario de **Familiarización**, dado que no se produce una simulación de atropello, los usuarios no tienen estímulos externos. Debido a ello, no engloba ninguna trayectoria perteneciente al grupo 2 (el que tiene mayor porcentaje de tiempo atencional).
- En el escenario de **Ventas (Avenida de los Toreros)**, el vehículo aparece a los pocos segundos de que se inicie el movimiento. El usuario se introduce en este escenario justo después de realizar el escenario de familiarización, y no se le previene de la aparición del vehículo. La distribución de instancias del grupo 2 es la menor de todos los escenarios donde se simulan atropellos. En este escenario se incluyen en su mayoría instancias del grupo 1, que es aquel en el que los usuarios tienden más a mirar a la derecha (el lado contrario al lado por el que aparece el vehículo).
- En el escenario de la **Avenida de Machupichu**, el número de instancias pertenecientes al grupo 0 es muy bajo. Este grupo engloba a los usuarios que caminan más rápido. Una posible explicación podría ser que la menor visibilidad que tienen los individuos del paso de peatones en este escenario podría forzar a que fuesen más prudentes al caminar, dando pasos más cortos, y teniendo mayor tendencia a pararse y retroceder. Esta hipótesis se vería reforzada debido a la mayor proporción de trayectorias del grupo 2 que se incluyen en este escenario.

- En el escenario de la **Avenida de los Hermanos G. Noblejas**, la distribución es más equilibrada, aunque también hay una mayor proporción de usuarios pertenecientes al grupo 2 sobre los grupos 0 y 1.
- Finalmente, en el escenario de la **Avenida de Emilio Muñoz**, la distribución de las trayectorias es favorable al grupo 0, el cual incluye a los usuarios que caminan más rápido. Hay una baja proporción de usuarios del grupo 1, que incluye a los usuarios que pasan más tiempo mirando a la derecha con un bajo tiempo atencional. Cabe recalcar que en este escenario el coche aparece por la derecha del peatón (Ver Tabla 4).

4.3.3.3 Análisis de los resultados obtenidos en función de la ocurrencia o no de un accidente.

En esta sección se muestra un análisis de los patrones de movimiento detectados en función de la ocurrencia de la colisión con el vehículo.

En la Ilustración 39 se muestra la distribución de las trayectorias accidentadas o no accidentadas en función de los patrones de movimiento identificados.

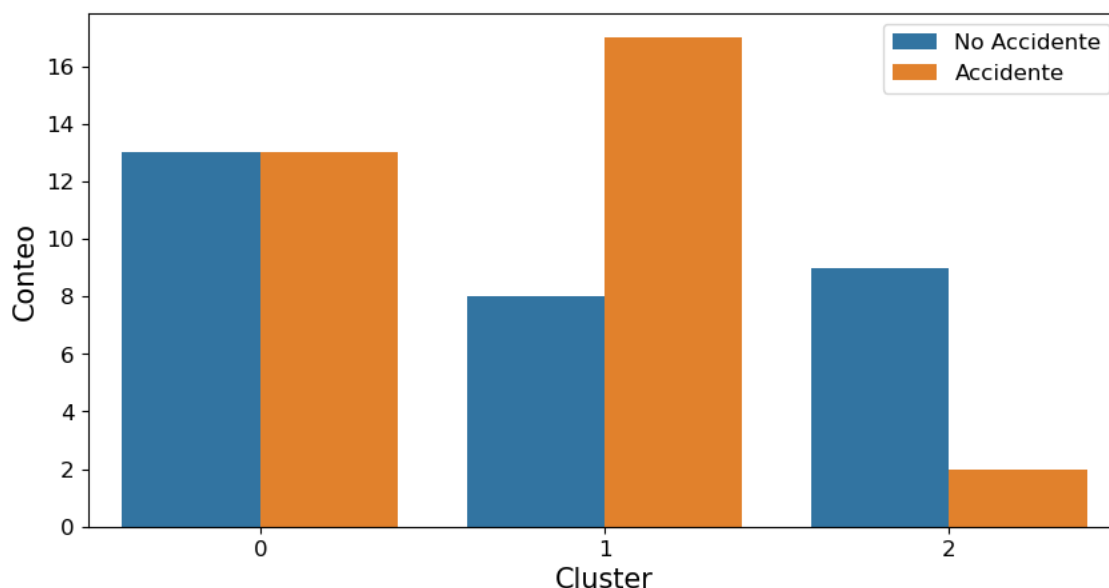


Ilustración 39. Distribución de las trayectorias accidentadas para los diferentes grupos.

Los ensayos en los que se obtuvieron los datos de los que se nutre este Trabajo de Fin de Máster forman parte de un estudio observacional. Por tanto, no puede inferirse una causalidad directa entre la ocurrencia del accidente y los patrones de movimiento identificados. Sin embargo, los resultados parecen indicar que en las trayectorias que se engloban en el grupo 2 (sujetos más prudentes, que se mueven más despacio, tienen mayor factor atencional y tienden más a pararse y retroceder cuando ven el vehículo); el riesgo de sufrir un accidente es notablemente menor a los otros dos grupos. De la misma forma, el riesgo de sufrir un accidente en las trayectorias que se engloban en el grupo 1 parece mucho mayor, mientras que en el grupo 0, la distribución de los usuarios accidentados y no accidentados es equilibrada. Este resultado abre un camino para nuevas investigaciones, que se discutirán en el siguiente capítulo.

5 Conclusiones, trabajos futuros y limitaciones.

En este Capítulo elaboro una evaluación del grado de consecución de cada uno de los objetivos técnicos que se plantearon al inicio del Trabajo de Fin de Máster, y analizo las conclusiones que se pueden extraer de los resultados obtenidos. De manera adicional, se plantean nuevas líneas para continuar la investigación en futuros trabajos, y se enuncian

5.1 Evaluación de la calidad de los datos.

El primer objetivo técnico de este Trabajo de Fin de Máster ha sido elaborar una evaluación de la completitud, coherencia y corrección de los datos posicionales proporcionados por el Centro de Domótica Integral.

Para ello, se han llevado a cabo dos análisis. En primer lugar, se elabora una evaluación del grado de visibilidad que tienen las cámaras de cada parte del cuerpo durante una trayectoria. Los resultados indican que las cámaras detectan mejor las manos y la cintura que los pies, probablemente por estar los últimos más cercanos al suelo, que puede producir reflejos que interfieran con la señal emitida por los sensores. De manera adicional, los resultados obtenidos parecen indicar que las cámaras del lado izquierdo de la sala tienen un mejor funcionamiento que las del lado derecho, y que la detección de las cámaras durante la vuelta al punto de partida empeora ligeramente.

En segundo lugar, se han propuesto dos medidas que definen la coherencia o corrección de las medidas de posición a partir de la similitud de las trayectorias de la cabeza y la cintura; y entre los pies: la **Longitud de la Subsecuencia Común más Larga** (LCSS), y la similitud calculada mediante el algoritmo de **Deformación Temporal Dinámica** (DTW).

La LCSS tiene la ventaja de que permite cuantificar el grado de mejora en la corrección biomecánica de los datos al interpolar las medidas de posición. Por otra parte, la similitud calculada mediante 1-DTW permite ofrecer una medida de la similitud global entre las trayectorias.

Los resultados obtenidos indican que el algoritmo de interpolación beneficia especialmente la similitud en el movimiento de la cabeza y la cintura. Entre los pies, el algoritmo de interpolación mejora la LCSS obtenida de manera significativa cuando las cámaras no tienen buena detección, pero esto no sucede cuando la detección es óptima. Al tratarse de un estudio observacional, no se es apropiado inferir causalidad, pero una posible línea para continuar explorando en este sentido sería analizar si la mayor libertad de movimiento de los pies hace más difícil realizar la interpolación.

En cuanto a la distancia calculada mediante el algoritmo de deformación temporal dinámica, no se obtuvieron diferencias significativas en el valor de similitud para cada nivel de visibilidad. Sin embargo, se observa que el valor máximo de distancia obtenido se reduce a medida que aumenta el nivel de visibilidad. Esto parece indicar que el peor caso esperable de disparidad en las trayectorias de los pies y de la cabeza y la cintura disminuye a medida que aumenta el nivel de visibilidad.

En cuanto a este primer objetivo, se han logrado extraer algunas conclusiones de la calidad global de las medidas de posición. Estas conclusiones han sido suficientes para identificar algunos puntos débiles en la detección del sistema de cámaras, así como para verificar que el algoritmo de interpolación actualmente implementado es válido y mejora los resultados.

5.2 Generación de una base de datos de características y parametrización del movimiento.

El segundo objetivo técnico del Trabajo se ha completado al generar una base de datos con las características del movimiento que aparecen en la Tabla 13. Posteriormente se han utilizado estas características para identificar patrones en el movimiento de los usuarios que participan en los ensayos, obteniendo tres grupos cuyas características parecen guardar cierta relación con el diseño de los escenarios integrados en la aplicación que recoge los datos. Para ello se han utilizado algoritmos de aprendizaje no supervisado, tanto jerárquicos como particionales. El algoritmo jerárquico, en combinación con el uso de métricas de evaluación de algoritmos no supervisados, ha permitido determinar el número de clústeres óptimo según el conjunto de datos, que en este caso ha sido de tres. Posteriormente se ha utilizado un algoritmo particional (K-means), para separar las instancias en grupos en función de los parámetros calculados. De esta forma se han cumplimentado los Objetivos Técnicos 2 y 3, definidos al inicio del Trabajo de Fin de Máster.

El entendimiento de las características de las trayectorias que se incluyen en cada grupo ha permitido analizar su relación con los escenarios de Realidad Virtual y con la probabilidad de sufrir o no un accidente. En concreto, algunos resultados preliminares indican que las trayectorias pertenecientes al grupo 2 (aquellas en las que los individuos caminan dando pasos más cortos, poseen mayor porcentaje de tiempo atencional y tienen mayor tendencia a pararse y retroceder) son menos propensas a que se produzca un accidente durante la simulación. Esto confirmaría lo que ya se descubrió en (3), donde los autores concluyen que un mayor porcentaje de tiempo atencional facilita que los sujetos eviten al vehículo, así como la reacción de pararse y retroceder, que es más habitual en este mismo grupo.

5.3 Líneas futuras.

El presente Trabajo de Fin de Máster deja numerosas líneas de investigación y desarrollo abiertas, tanto en el ámbito de la mejora de las experiencias inmersivas en aplicaciones de Realidad Virtual; como en la identificación de reacciones en los peatones durante la fase previa a un accidente.

Con respecto a la mejora de las experiencias de Realidad Virtual, a partir de los patrones de movimiento identificados y parametrizados se ha obtenido conocimiento acerca de cómo se mueven los usuarios que utilizan la aplicación de Realidad Virtual. Comprender el movimiento real del ser humano en estas situaciones facilita el modelado de los avatares de Realidad Virtual, permitiendo conseguir una experiencia virtual más realista y atractiva para el usuario. Esto también permitirá optimizar el algoritmo de interpolación diseñado y desarrollado por los investigadores del CEDINT. Teniendo en cuenta las características de la biomecánica humana, se puede conseguir un resultado más adecuado utilizando una interpolación diferente para cada sensor, en función de la región corporal donde se sitúa.

Otras líneas futuras podrían avanzar con el estudio del comportamiento humano en la fase previa a una colisión con un vehículo. En este sentido, sería interesante explorar como se relacionan los grupos de trayectorias con el riesgo de sufrir un accidente en cada uno de los escenarios de Realidad Virtual por separado, permitiendo descubrir si hay situaciones de tráfico o visibilidad del paso de peatones que favorezcan la ocurrencia de una colisión en función del comportamiento de los peatones.

Otras posibles líneas de investigación futura podrían ahondar aún más en la identificación de patrones de movimiento. El uso de otros algoritmos no supervisados, como el clustering probabilístico, no ha sido explorado en este Trabajo de Fin de Máster. En futuros Trabajos se podrían explorar algoritmos de esta índole para tratar de obtener conclusiones diferentes.

5.4 Limitaciones.

En este apartado se describen algunas de las limitaciones del Trabajo, tanto por parte del sistema de recogida de datos como del propio desarrollo.

5.4.1 Limitaciones del sistema de recogida de datos.

Como se explica en este trabajo desde un primer momento, el sistema de captura de movimiento *Optitrack*, aunque es preciso en la localización de los objetos en el espacio tridimensional, está sujeto a la aparición de valores nulos.

Por este motivo, los investigadores del CEDINT se vieron obligados a desarrollar un algoritmo de interpolación para lograr que la recogida de datos fuese completa. Sin embargo, la fiabilidad de este algoritmo no había sido testeada hasta ahora. Siguiendo la metodología descrita en el apartado 3.2.4.2, se ha tratado de identificar cuáles trayectorias contienen información más útil, completa y veraz para detectar patrones de movimiento de forma fiable. A pesar de ello, los datos recogidos representan una estimación de lo que sería el movimiento real del usuario y puede no corresponderse con la realidad en todos los casos.

Como se explica en el apartado 5.3, una posible línea para futuros desarrollos sería la optimización de este algoritmo, aplicando una interpolación diferente en función del elemento corporal que se monitoriza.

5.4.2 Limitaciones del desarrollo.

En cuanto a las limitaciones del desarrollo del Trabajo de Fin de Máster, cabe destacar la no existencia de una referencia objetiva a la hora de evaluar el sistema de captura de movimiento. La evaluación del sistema se ha realizado a partir de restricciones de la biomecánica humana, aunque no se conoce el grado de cumplimiento de dichas restricciones que sería necesario para considerar que los datos son correctos.

De manera adicional, respecto a la identificación de patrones de movimiento en la fase previa a un accidente, cabe destacar que no está claro que el movimiento de las personas en entornos virtuales se corresponda de manera inequívoca con el movimiento en situaciones de riesgo real. En [39] se realiza una comparativa del valor obtenido en los parámetros de la marcha en el mundo real y en un entorno virtual. Los autores concluyen que la longitud del paso y la velocidad de la marcha van en aumento a medida que los usuarios ganan experiencia en el entorno virtual, asemejándose más a lo que ocurriría en una situación real.

Teniendo en cuenta estas limitaciones, cabe recalcar que los resultados obtenidos en este Trabajo son parte de una línea de investigación, y únicamente pretenden establecer un camino para poder continuar avanzando en futuros proyectos.

6 Anexo 1. Descripción de los conjuntos de datos de partida.

6.1 Archivos de log de Optitrack.

Variable	Tipo de dato	Descripción semántica.	Unidades
<i>Time</i>	Variable temporal.	Instante temporal en que se registra la medida.	Segundos (s).
<i>RIGID_BODY_position_AXIS</i>	Variable numérica continua.	Posición del Rigid Body en el eje axis (x y o z).	Metros (m)
<i>RIGID_BODY_orientation_AXIS</i>	Variable numérica continua.	Orientación del Rigid Body alrededor del eje axis (ángulos de Euler).	Grados (°)

Tabla 17. Descripción de los archivos en crudo recogidos por Optitrack.

6.2 Archivos de log de la Aplicación.

Variable	Tipo de dato	Descripción semántica.	Unidades
<i>id</i>	Variable numérica discreta.	Identificador de fila (clave suplente).	No aplica.
<i>timeStamp</i>	Variable temporal	Instante temporal en que se tomó la medida.	Milisegundos en formato UNIX.
<i>type</i>	Variable categórica.	Tipo de medida. Las medidas de tipo <i>TRACKER</i> hacen referencia a los Rigid Bodies. Las medidas <i>EVENT</i> hacen referencia a los eventos que tienen lugar durante la simulación (Ver Tabla 19).	No aplica.
<i>name</i>	Variable categórica.	Nombre del elemento al que hace referencia la medida. Los elementos de tipo <i>TRACKER</i> son <i>CABEZA</i> , <i>CINTURA</i> , <i>MANO_I</i> (mano izquierda), <i>PIE_I</i> (pie izquierdo), <i>MANO_D</i> (mano derecha) y <i>MANO_I</i> (mano izquierda).	No aplica.
<i>transform</i>	Tupla de seis elementos. Cada elemento es una variable numérica continua.	Los tres primeros elementos hacen referencia a la posición con respecto a los tres ejes; y los tres últimos a la orientación.	Metros para los tres primeros elementos, grados para los tres últimos.
<i>markers</i>	Variable numérica discreta.	El número de marcadores detectados para cada <i>TRACKER</i> en esa medición.	Número de marcadores [0,8].
<i>interpolation</i>	Variable categórica	El nivel de interpolación que aplicó el algoritmo en esa medida.	No aplica.

Tabla 18. Descripción de los archivos en crudo recogidos por la Aplicación.

Evento	Tipo de evento	Descripción
<i>Approach_start</i>	EYE_TRACKER_EVENT	El avatar se empieza a mover de forma automática a la sala de ensayo.
<i>Approach_end</i>	EYE_TRACKER_EVENT	Termina el movimiento automático del avatar.
<i>Simulation_start</i>	CAR_EVENT	Inicio de la simulación
<i>Simulation_end</i>	CAR_EVENT	Fin de la simulación
<i>Break_start</i>	CAR_EVENT	El vehículo empieza a frenar.
<i>Break_end</i>	CAR_EVENT	El vehículo termina de frenar.
<i>Horn_start</i>	CAR_EVENT	El vehículo empieza a pitar.
<i>Horn_end</i>	CAR_EVENT	El vehículo termina de pitar.
<i>Accident</i>	CAR_EVENT	Se produce la colisión con el vehículo.

Tabla 19. Descripción de los eventos que se encuentran en los archivos recogidos por la aplicación.

7 Bibliografía

- [1] Observatorio Nacional de Seguridad Vial, “Avance de las principales cifras de la Siniestralidad Vial,” 2020.
- [2] J. De Winter, P. M. van Leeuwen, R. Happee, and others, “Advantages and disadvantages of driving simulators: A discussion,” in *Proceedings of measuring behavior*, 2012, pp. 28–31.
- [3] Á. Losada, F. J. Páez, F. Luque, and L. Piovano, “Application of Machine Learning Techniques for Predicting Potential Vehicle-to-Pedestrian Collisions in Virtual Reality Scenarios,” *Applied Sciences (Switzerland)*, vol. 12, no. 22, Nov. 2022, doi: 10.3390/app122211364.
- [4] M. Jimenez, I. Becerra, and U. Ruiz, “Toward inertial position tracking for head-mounted displays: a dataset and a deep learning approach evaluation,” *Virtual Real*, vol. 27, no. 3, pp. 2605–2621, Sep. 2023, doi: 10.1007/s10055-023-00831-x.
- [5] Optitrack, “Optitrack documentation,” <https://docs.optitrack.com/>. Accessed: Feb. 24, 2024. [Online]. Available: <https://www.optitrack.com/>
- [6] L. L. Gómez Echeverry, A. M. Jaramillo Henao, M. A. Ruiz Molina, S. M. Velásquez Restrepo, C. A. Páramo Velásquez, and G. J. Silva Bolívar, “Human motion capture and analysis systems: a systematic review,” *Prospectiva*, vol. 16, no. 2, pp. 24–34, 2018.
- [7] G. Nagymáté and R. M. Kiss, “Application of OptiTrack motion capture systems in human movement analysis,” *Recent Innovations in Mechatronics*, vol. 5, no. 1., Jan. 1970, doi: 10.17667/riim.2018.1/13.
- [8] Meta Quest, “Oculus Documentation,” <https://developer.oculus.com/documentation/>.
- [9] HTC Vive, “HTC Vive Pro User Guide,” <https://developer.vive.com/resources/hardware-guides/vive-pro-specs-user-guide/>.
- [10] R. Monica and J. Aleotti, “Evaluation of the Oculus Rift S tracking system in room scale virtual reality,” *Virtual Real*, vol. 26, no. 4, pp. 1335–1345, Dec. 2022, doi: 10.1007/s10055-022-00637-3.
- [11] S. Zhang, B. Y. Tan, K. H. Leo, and R. Q. F. Tham, “A hybrid human motion tracking system for virtual rehabilitation,” in *2011 6th IEEE Conference on Industrial Electronics and Applications*, 2011, pp. 1993–1998.
- [12] T. L. Stańczyk, L. Jurecki, M. Jaśkiewicz, S. Walczak, and R. Janczur, “Researches on the reaction of a pedestrian stepping into the road from the right side from behind and an obstacle realized on the track,” *Journal of KONES*, vol. 18, no. 1, pp. 615–622, 2011.
- [13] D. M. Gavrila, “The visual analysis of human movement: A survey,” *Computer vision and image understanding*, vol. 73, no. 1, pp. 82–98, 1999.
- [14] F. Zhou, F. la Torre, and J. K. Hodgins, “Hierarchical Aligned Cluster Analysis for Temporal Clustering of Human Motion,” *IEEE Trans Pattern Anal Mach Intell*, vol. 35, no. 3, pp. 582–596, 2013, doi: 10.1109/TPAMI.2012.137.
- [15] Unity Technologies, “Unity Documentation.” <https://docs.unity.com/>, 2023.

- [16] Pydata, “Pandas documentation,” <https://pandas.pydata.org/>. Accessed: Mar. 23, 2024. [Online]. Available: <https://pandas.pydata.org/>
- [17] Matplotlib, “Matplotlib documentation,” <https://matplotlib.org/>. Accessed: Mar. 23, 2024. [Online]. Available: <https://matplotlib.org/>
- [18] Pydata, “Seaborn documentation,” <https://seaborn.pydata.org/>. Accessed: Mar. 23, 2024. [Online]. Available: <https://seaborn.pydata.org/>
- [19] Python, “Python 3.11.8,” <https://docs.python.org/3.11/>. Accessed: Mar. 23, 2024. [Online]. Available: <https://docs.python.org/3.11/>
- [20] Scikit-Learn, “Scikit Learn User Guide,” <https://scikit-learn.org/stable>. Accessed: Jun. 19, 2024. [Online]. Available: <https://scikit-learn.org>
- [21] NumPy, “NumPy documentation,” <https://numpy.org/>. Accessed: Mar. 23, 2024. [Online]. Available: <https://numpy.org/>
- [22] Scipy, “Scipy documentation,” <https://scipy.org/>. Accessed: Mar. 23, 2024. [Online]. Available: <https://scipy.org/>
- [23] A. Martín Nogueras, J. L. Calvo Arenillas, J. Orejuela Rodríguez, F. J. Barbero Iglesias, and C. Sánchez Sánchez, “Fases de la marcha humana,” *Revista Iberoamericana de Fisioterapia y Kinesiología*.
- [24] B.-K. Yi, H. V Jagadish, and C. Faloutsos, “Efficient Retrieval of Similar Time Sequences under Time Warping.”
- [25] M. Vlachos, G. Kollios, and D. Gunopulos, “Discovering Similar Multidimensional Trajectories,” 2002.
- [26] L. Chen and R. Ng, “On The Marriage of Lp-norms and Edit Distance,” 2004.
- [27] Y. Zhang and J. Glass, “A Piecewise Aggregate Approximation Lower-Bound Estimate for Posteriorgram-based Dynamic Time Warping,” 2011.
- [28] J. Lin, E. Keogh, L. Wei, and S. Lonardi, “Experiencing SAX: A novel symbolic representation of time series,” *Data Min Knowl Discov*, vol. 15, no. 2, pp. 107–144, Oct. 2007, doi: 10.1007/s10618-007-0064-z.
- [29] Universidad de Málaga, “Fundamentos del Procesado Digital de la Señal.”
- [30] I. D. Pinzón Ríos, “Cabeza hacia adelante: una mirada desde la biomecánica y sus implicaciones sobre el movimiento corporal humano,” *Revista de la Universidad Industrial de Santander. Salud*, vol. 47, no. 1, pp. 75–83, 2015.
- [31] C. Bielza and P. Larrañaga, *Data-Driven Computational Neuroscience*. Cambridge University Press, 2021.
- [32] Ethem Alpaydin, *Introduction to Machine Learning*, 4th ed. MIT Press, 2020.
- [33] Scikit Learn, “Scikit Learn User Guide (Hierarchical Clustering),” <https://scikit-learn.org/stable/modules/clustering.html#hierarchical-clustering>. Accessed: Jun. 19, 2024. [Online]. Available: <https://scikit-learn.org/stable/modules/clustering.html#hierarchical-clustering>
- [34] S. Lloyd, “Least squares quantization in PCM,” *IEEE Trans Inf Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982, doi: 10.1109/TIT.1982.1056489.

- [35] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J Comput Appl Math*, vol. 20, pp. 53–65, 1987, doi: [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
- [36] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics*, vol. 3, no. 1, pp. 1–27, 1974, doi: [10.1080/03610927408827101](https://doi.org/10.1080/03610927408827101).
- [37] W. H. Kruskal and W. A. Wallis, "Use of Ranks in One-Criterion Variance Analysis," *J Am Stat Assoc*, vol. 47, no. 260, p. 583, Dec. 1952, doi: [10.2307/2280779](https://doi.org/10.2307/2280779).
- [38] D. S. Moore, W. I. Notz, and M. Flinger, "Non Parametric Tests," in *Best Practice of Statistics*.
- [39] M. Nasiri, R. G. Anaraky, S. V. Babu, and A. Robb, "Gait Differences in the Real World and Virtual Reality: The Effect of Prior Virtual Reality Experience," in *Proceedings - 2022 IEEE International Symposium on Mixed and Augmented Reality, ISMAR 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 631–636. doi: [10.1109/ISMAR55827.2022.00080](https://doi.org/10.1109/ISMAR55827.2022.00080).

