



Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias de la Computación
Doctorado en Ingeniería del Lenguaje y del Conocimiento

Metodología para el Reconocimiento de Comandos
Gestuales Definidos por el Usuario siguiendo el Enfoque de
Aprendizaje One-Shot

Una disertación presentada en cumplimiento de los requisitos para obtener el grado de
Doctor en Ingeniería del Lenguaje y del Conocimiento

Presenta:

David Augusto Céspedes Hernández

Junio, 2022

Directores:

Dr. Juan Manuel González Calleros

Dr. Jean Vanderdonckt

Comité evaluador:

Dra. Josefina Guerrero García

Dra. María Auxilio Medina Nieto

Dr. José Pascual Molina Massó

Dedicatoria

Mi estancia en el programa de doctorado estuvo marcada por los cambios. Pasamos por un sismo que hizo que tuviéramos que trabajar a distancia por un tiempo, por marchas y paros de actividades en la universidad, y finalmente por una pandemia que nos hizo cambiar completamente la forma de trabajar y que marcó un parteaguas en la forma en la que los procesos educativos y laborales se llevan a cabo.

Vivir todos esos cambios hizo que la adaptación fuera determinante, y adaptarme no hubiera sido posible sin las personas que se mantuvieron a mi alrededor durante esos momentos. Por eso, quiero agradecer a Lili por estar siempre cerca durante más de 12 años ya. Por ayudarme a crecer y ser siempre una motivación para seguir adelante. No pude haber tenido mejor compañera de pandemia y me enorgullece también haber sido testigo de tu crecimiento. T.A.C.N.Y.P.S.

Quiero agradecer también a mis papás y a mi hermana, porque han estado presentes siempre. Preocupándose por mí, siendo un apoyo en cada momento, y estando al pendiente.

A mis amigos y familia, tanto a quienes se mantuvieron cerca como a quienes se agregaron durante los últimos años y meses. Mamá de Lili, Papá de Lili, Marco, Patty, Memo, Clau, Dra. Josefina, familia necaxista, Necaxistas CF y Octonautas, por sus valiosos aportes y su tiempo y atención dedicados a mí.

Finalmente, a mi amigo y director de tesis, Dr. Juan Manuel. Por su paciencia a lo largo de estos casi 10 años de trabajo conjunto, por las oportunidades que me ha dado para desarrollarme profesionalmente, y por su gran esfuerzo para hacerme el trabajo más fácil.

Agradecimientos

Agradezco al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el otorgamiento de la beca con número de apoyo 306387 para la realización de este doctorado, que tiene como resultado el presente trabajo de tesis. Asimismo, agradezco por el otorgamiento de la beca mixta con número de convocatoria 291250, que me permitió realizar una estancia en el extranjero.

Agradezco a la Universidad de Castilla-La Mancha y específicamente a los profesores Dr. Pascual González López y Dr. José Pascual Molina Massó por las facilidades brindadas para la realización de la estancia de investigación.

Gracias a mi director de tesis, el Dr. Juan Manuel González Calleros por sus comentarios y asesorías, todos sus comentarios fueron invaluable para este trabajo de investigación.

Je tiens à remercier le professeur Jean Vanderdonckt, en sa qualité de codirecteur, pour avoir pris le temps d'examiner mon travail et pour sa volonté de collaborer.

Finalmente, agradezco el tiempo invertido en la lectura y envío de comentarios al comité evaluador: Dra. Josefina Guerrero García, Dr. José Pascual Molina Massó, y Dra. María Auxilio Medina Nieto, gracias por su valioso aporte al trabajo de investigación.

Tabla de Contenido

Resumen	10
Abstract	11
1. INTRODUCCIÓN	12
1.1. Planteamiento de la investigación.....	13
1.2. Definición del problema.....	13
1.3. Objetivos de la investigación.....	14
1.4. Justificación de la investigación.....	14
1.5. Preguntas de investigación.....	15
1.6. Hipótesis.....	15
1.7. Organización del documento.....	15
2. MARCO TEÓRICO	17
2.1. Ergonomía y movimiento corporal.....	17
2.1.1. Antropometría.....	18
2.1.2. Postura humana.....	19
2.1.3. Actividad humana.....	19
2.2. Aprendizaje automático.....	20
2.2.1. Técnicas de aprendizaje automático.....	21
2.2.2. Enfoque de aprendizaje <i>one-shot</i>	22
2.3. Interacción Humano-Computadora.....	23
2.3.1. Interfaces Naturales de Usuario.....	24
2.3.2. Usabilidad.....	25
2.3.3. Experiencia de Usuario.....	27
2.3.4. Modelado de contexto.....	28
2.3.5. Tareas abstractas.....	28
2.4. Conclusiones del capítulo.....	29
3. ESTADO DEL ARTE	32

3.1.	Modelos de usuario	33
3.2.	Trabajos relacionados	34
3.2.1.	Interacción gestual usando dispositivos basados en visión.....	38
3.2.2.	Interacción gestual utilizando dispositivos vestibles.....	44
3.2.3.	Interacción gestual por medio de señales inalámbricas.....	45
3.3.	Vocabularios de gestos definidos por el usuario	46
3.4.	Conclusiones del capítulo	48
4.	METODOLOGÍA.....	50
4.1.	Identificación de Regiones de Interés	50
4.2.	Segmentación de la Información	51
4.3.	Obtención de Datos	51
4.4.	Aprendizaje A-Priori	56
4.5.	Aprendizaje A-Posteriori	60
4.6.	Reconocimiento Gestual	62
4.7.	Asociación de Gestos Reconocidos y Significados Semánticos	62
4.8.	Conclusiones del capítulo	63
5.	RESULTADOS	66
5.1.	Navegación en Ambientes Virtuales utilizando Comandos Definidos por el Usuario 66	
5.2.	Identificación de un vocabulario gestual para interacción con múltiples aplicaciones 71	
5.3.	Un vocabulario gestual asociado a tareas abstractas	74
5.4.	Entrenamiento preliminar de un clasificador gestual siguiendo un enfoque de aprendizaje <i>one-shot</i>	78
5.5.	Conclusiones del capítulo	80
6.	DISCUSIÓN DE LOS RESULTADOS	83

6.1. Navegación en Ambientes Virtuales utilizando Comandos Definidos por el Usuario	83
6.2. Identificación de un vocabulario gestual para interacción con múltiples aplicaciones	85
6.3. Un vocabulario gestual asociado a tareas abstractas	87
6.4. Entrenamiento preliminar de un clasificador gestual siguiendo un enfoque de aprendizaje <i>one-shot</i>	90
6.5. Conclusiones del capítulo	93
7. CONCLUSIONES Y TRABAJO FUTURO.....	96
REFERENCIAS	102
<i>Apéndice A. Estructura de datos para almacenamiento de información antropométrica descriptiva de un gesto.....</i>	
	115
<i>Apéndice B. Reporte de caracterización de gestos del estudio sobre identificación de un vocabulario gestual para interacción con múltiples aplicaciones.....</i>	
	122
<i>Apéndice C. Reporte de caracterización de gestos del estudio sobre identificación de un vocabulario gestual para interacción con tareas abstractas.....</i>	
	125
<i>Apéndice D. Fragmentos de los archivos JSON y ARF generados para el experimento reportado en la sección 5.4 sobre entrenamiento de un clasificador gestual siguiendo un enfoque de aprendizaje one-shot.....</i>	
	131
Fragmento del archivo JSON	131
Fragmento del archivo ARFF	133
<i>Apéndice E. Formato de consentimiento para participación en los estudios preliminares realizados</i>	
	136

Tabla de tablas

Tabla 1. Relación entre tareas abstractas y sus concretizaciones.	29
Tabla 2. Relación entre elementos del marco epistémico y conceptos descritos en el marco teórico.	30
Tabla 3. Relación entre elementos del marco epistémico y conceptos revisados dentro del estado del arte.	48
Tabla 4. Instrumento propuesto para registro de tareas e incorporación de resultados preliminares.	58
Tabla 5. Partes del cuerpo asociadas a sus posibles movimientos de acuerdo con el modelo GVUM.	59
Tabla 6. Relación entre elementos del marco epistémico y conceptos abordados dentro de la metodología propuesta.	63
Tabla 7. Tareas para observar durante el experimento para identificación de un vocabulario gestual para interacción con múltiples aplicaciones.	72
Tabla 8. Tareas seleccionadas para el experimento de identificación de un vocabulario gestual asociado a tareas abstractas.	75
Tabla 9. Relación entre elementos del marco epistémico y conceptos abordados dentro de la presentación de resultados.	80
Tabla 10. Relación entre elementos del marco epistémico y resultados reportados tras los experimentos realizados.	94

Tabla de figuras

Figura 1. Representación del Modelo Virtual de Usuario Genérico de Kaklanis et. al (2013).....	33
Figura 2. Distribución de los dispositivos para adquisición de información utilizados en la literatura. Elaboración propia.....	35
Figura 3. Distribución de las partes del cuerpo para las que se lleva a cabo reconocimiento gestual en la literatura. Elaboración propia.....	36
Figura 4. Frecuencia de uso de técnicas de clasificación para reconocimiento gestual empleadas en la literatura. Elaboración propia.	37
Figura 5. Metodología de siete etapas propuesta para interacción basada en gestos incluyendo el flujo de la información.....	50
Figura 6. Resultado de configuración del sensor Microsoft Kinect para obtención de información.....	53
Figura 7. proyecto Kinect Unity3D Demo de NeurorehabLab	54
Figura 8. Capturas de la utilización de las muestras incluidas en Vitrubius.	55
Figura 9. Captura de pantalla de la aplicación para monitoreo de posiciones y ángulos de las articulaciones de interés superponiendo un esqueleto en el que tales puntos se destacan.....	56
Figura 10. Software para captura de gestos junto con los archivos que se generan en formato JSON y ARFF.	60
Figura 11. Configuración del experimento siguiendo el enfoque de Mago de Oz y la disposición en un escenario real.....	68
Figura 12. Vocabulario de gestos propuesto para la interacción con ambientes de RV.	68
Figura 13. Evaluación general del sistema de navegación en ambientes virtuales en términos de tiempo promedio en segundos para cada tarea realizada, cantidad de errores cometidos en total, tareas no finalizadas en total, y la media de las dimensiones de usabilidad de acuerdo con el instrumento CSUQ.	70
Figura 14. Representación gráfica del vocabulario de gestos definidos por usuarios incluyendo el identificador de las tareas asociadas a ellos y sus etiquetas genéricas. .	74

Figura 15. Extracto de la representación gráfica del vocabulario de gestos obtenido durante el estudio asociado a tareas abstractas.....	77
Figura 16. Representación formal del gesto A006 junto con su representación gráfica.....	78
Figura 17. Posturas adoptadas para la realización del experimento sobre entrenamiento de un clasificador gestual para tareas de navegación en ambientes virtuales siguiendo un enfoque de aprendizaje one-shot.....	79
Figura 18. Matriz de confusión arrojada durante el proceso de entrenamiento siguiendo el enfoque de aprendizaje one-shot.....	79
Figura 19. Resultados de la realización de pruebas sobre el conjunto de gestos generado de forma aleatoria.....	80
Figura 20. Relación entre tiempos de respuesta y edades de los usuarios, y entre tiempos de respuesta y calificaciones de los usuarios para gestos realizados con la cabeza.....	87
Figura 21. Reporte de precisión alcanzada en experimento realizando modificaciones sobre el número de muestras consideradas y la cantidad de atributos utilizados resaltando como mapa de calor en rojo los resultados menos deseables y en verde los más deseables.....	91
Figura 22. Reporte de precisión alcanzada en experimento realizando modificaciones sobre el número de muestras consideradas y la cantidad de atributos utilizados resaltando como mapa de calor en rojo los resultados menos deseables y en verde los más deseables.....	92

Resumen

La comunicación humana se ha estudiado desde diferentes enfoques, lo que ha dado lugar a contribuciones en varias disciplinas. Desde el punto de vista de las ciencias de la computación, los descubrimientos realizados en el área han inspirado el desarrollo de las Interfaces Naturales de Usuario (NUI), mecanismos de interacción destinados a replicar la forma en que las personas se comunican, de modo que el intercambio de información con los sistemas computacionales se produzca de manera similar. Las interfaces gestuales son un tipo específico de NUI centrado en el análisis de la relación entre el movimiento del cuerpo y los significados semánticos. Aunque, desde un punto de vista técnico, las propuestas encontradas en la literatura han demostrado una alta eficiencia y precisión en el reconocimiento gestual, diversos autores han manifestado la falta de naturalidad en la interacción con aplicaciones basadas en gestos, lo que lleva a la conclusión de que las NUI no suelen ser tan naturales como se cree. Además, los gestos son específicos de una cultura y de un idioma, lo que los hace ambiguos, incompletamente especificados y difíciles de relacionar con el significado semántico cuando se desconoce el contexto. En esta disertación, se propone una metodología para permitir el desarrollo de aplicaciones basadas en gestos, considerando que la precisión y la eficiencia en las tareas de reconocimiento no deben verse afectadas, pero priorizando la flexibilidad para permitir el uso de gestos que sean adecuados para diferentes contextos de usuario a través de la exploración de conjuntos de gestos definidos por el usuario y técnicas de aprendizaje automático, y utilizando un enfoque de aprendizaje de una sola toma.

Abstract

Human communication has been studied from different approaches resulting in contributions to several disciplines. From the computer sciences point of view, the findings made in the area have inspired the development of Natural User Interfaces (NUI), interaction mechanisms aimed at replicating the way in which people communicate, so the information exchange with computational systems happens in similar fashion. Gestural interfaces are a specific type of NUI focused on analyzing the relationship between body motion and semantic meanings. Although, from a technical perspective, proposals found in the literature had proven high efficiency and accuracy on gestural recognition, several authors had reported lack of naturalness in the interaction with gesture-based applications, leading to the conclusion that NUIs are not usually as natural as they claim to be. Moreover, gestures are culture and language specific, which makes them ambiguous, incompletely specified, and difficult to match with semantic meaning when the context is unknown. In this dissertation, a methodology for enabling the development of gesture-based applications is proposed, considering that accuracy and efficiency in recognition tasks must not be affected, but prioritizing the flexibility for allowing the use of gestures that are suitable for different user contexts through the exploration of user-defined gesture sets and Machine Learning techniques, and using a one-shot learning approach.

1. INTRODUCCIÓN

El estudio de la comunicación humana, abordado desde diferentes enfoques, ha resultado en contribuciones a múltiples disciplinas. Desde la perspectiva de las ciencias computacionales, los hallazgos realizados en el área han inspirado el desarrollo de mecanismos de interacción que imitan la manera en la que las personas se comunican entre ellas, de tal forma que el intercambio de información con sistemas computacionales ocurra similarmente. Así, en el campo de la Interacción Humano-Computadora (HCI por sus siglas en inglés), se han desarrollado trabajos relacionados con el uso de pantallas táctiles, superficies interactivas, interfaces tangibles, interacción vocal e interacción gestual, resultando en la acuñación del término Interfaces Naturales de Usuario (NUI por sus siglas en inglés) (Jagodziński y Wolski, 2015).

A pesar de que múltiples investigaciones mencionan que aproximadamente el 65 por ciento de la comunicación humana se lleva a cabo a través de mecanismos no verbales (Anders, 2015), es posible observar en la literatura que se da mayor atención al análisis de comunicación por vía vocal y escrita. El énfasis en el reconocimiento gestual (considerando expresiones faciales y movimientos corporales), se ha dado mayormente durante las últimas dos décadas como consecuencia del desarrollo de algoritmos de aprendizaje automático (ML por sus siglas en inglés), así como de técnicas de procesamiento digital de imágenes y de señales, junto con dispositivos tales como cámaras de profundidad, cámaras estereoscópicas y dispositivos vestibles.

El análisis y la interpretación de gestos representan, en sí, un problema complejo que tradicionalmente es abordado siguiendo un enfoque de tres pasos: 1) obtención de datos de una fuente; 2) identificación de regiones de interés, es decir, dependiendo del gesto a reconocer se identifican gestos candidatos en términos de tiempo y espacio; y 3) el reconocimiento del gesto, normalmente logrado por medio del uso de técnicas de ML, y que resulta en la asociación del gesto con significados específicos. Es importante señalar que para que este proceso funcione apropiadamente, se requiere llevar a cabo el entrenamiento de clasificadores o la configuración de plantillas y, por lo tanto, los gestos que pueden ser reconocidos son limitados y predefinidos por diseñadores de experiencia de usuario (UX) (Liu y Wang, 2018).

En este trabajo de disertación se propone una metodología para permitir el desarrollo de aplicaciones cuya interacción se basa en comandos gestuales. Se debe tomar en cuenta que la precisión y la eficiencia en las tareas de reconocimiento no deben ser afectadas, priorizando la flexibilidad para permitir el uso de gestos adecuados para diferentes contextos de uso, a través de la exploración de conjuntos de gestos definidos por el usuario y de técnicas de ML, y empleando un enfoque de aprendizaje *one-shot* (de una sola toma).

1.1. Planteamiento de la investigación

Si bien, desde una perspectiva técnica, en la literatura es posible encontrar propuestas que han mostrado una alta eficiencia y precisión en cuanto al reconocimiento gestual tales como las de Kim et. al (2017), Zou et. al (2017) y Liu et. al (2017), autores como Norman (2010), Norman y Nielsen (2010), y Gonzalez-Calleros et. al (2018), refieren falta de naturalidad en la interacción con aplicaciones basadas en gestos, llevando a la conclusión de que las NUI no siempre son tan naturales como su nombre lo sugiere.

1.2. Definición del problema

Las interfaces gestuales son un tipo específico de NUI enfocado en analizar la relación entre movimientos corporales y significados semánticos. A pesar de que, desde una perspectiva técnica, se pueden encontrar trabajos en la literatura que muestran altos niveles de precisión y eficiencia en el reconocimiento gestual, múltiples autores han reportado falta de naturalidad en aplicaciones que hacen uso de interacción basada en gestos, llegando a la conclusión de que la interacción haciendo uso de estas interfaces no satisface las expectativas que tienen los usuarios en cuanto a su usabilidad y a una correcta UX. Además, los gestos son específicos a culturas y lenguajes, haciéndolos ambiguos, incompletamente especificados, y causando que su emparejamiento con significados semánticos sea complicado cuando se desconoce el contexto de uso (Chandler y Schwarz, 2009).

1.3. Objetivos de la investigación

El presente trabajo de disertación tiene como objetivo general, proponer una metodología para la definición de conjuntos de gestos definidos por el usuario para su reconocimiento en contextos determinados, utilizando técnicas de procesamiento digital de imágenes y de clasificación siguiendo un enfoque de aprendizaje *one-shot* (de una sola toma), para favorecer la interacción natural.

Este objetivo general se deriva en los siguientes objetivos específicos:

O. E. 1. Realizar un mapeo de usuario en 3D por medio de dispositivos basados en visión y API (Interfaz de Programación de Aplicaciones) para permitir el monitoreo de movimientos y el reconocimiento gestual.

O. E. 2. Diseñar e implementar una base de datos de gestos a partir de información obtenida durante estudios preliminares para permitir la comparación y el análisis.

O. E. 3. Proponer un método para etiquetado de gestos utilizando una notación genérica para soportar las tareas de reconocimiento y de asociación de significados para contextos específicos.

O. E. 4. Implementar reconocimiento gestual utilizando técnicas de ML y siguiendo un enfoque de aprendizaje *one-shot* para permitir interacción natural.

1.4. Justificación de la investigación

La presente investigación encuentra su motivación al observar que el reconocimiento gestual se aplica actualmente en diversos dominios tales como: interacción niño-computador; identificación y detección forense; monitoreo médico y rehabilitación; navegación y manipulación en ambientes virtuales; aprendizaje a distancia; seguridad; entrenamiento atlético; y entretenimiento. Adicionalmente, empresas como Google, Microsoft, Apple y Orbbec, por mencionar algunas, realizan esfuerzos en el desarrollo de Hardware y Software enfocándose en el estudio del reconocimiento gestual haciendo pertinente su tratamiento desde una perspectiva de investigación e ingeniería.

1.5. Preguntas de investigación

El problema presentado en líneas anteriores puede ser abordado a través de las siguientes preguntas de investigación:

P. I. 1. ¿Cómo se puede lograr la interacción gestual de manera que resulte natural para los usuarios?

P. I. 2. ¿Qué algoritmos, técnicas, métodos y estrategias se deben incluir en una metodología para facilitar la implementación de aplicaciones con interacción basada en gestos?

P. I. 3. ¿Qué implicaciones resultan de la utilización de gestos definidos por el usuario en una evaluación de usabilidad y experiencia de usuario, realizada sobre aplicaciones basadas en interacción gestual?

1.6. Hipótesis

A través de la definición de una metodología que incluya técnicas de clasificación y de procesamiento digital de imágenes, junto con algoritmos y estrategias de aprendizaje automático, es posible facilitar el desarrollo de aplicaciones con interacción basada en gestos y mejorar la apreciación de su usabilidad y la experiencia de usuario.

1.7. Organización del documento

El resto del documento de disertación se divide en capítulos. En el capítulo dos se reporta el marco teórico, incluyendo la descripción de conceptos relevantes para comprender la investigación realizada. El capítulo tres contiene el estado del arte, considerando trabajos relacionados, que se toman como base para la presente investigación y que, a la vez, permiten contar con marcos de referencia para la evaluación y comparación de los resultados obtenidos. En el capítulo 4 se describe de manera detallada la propuesta de una metodología para el reconocimiento de comandos gestuales definidos por el usuario siguiendo un enfoque de aprendizaje *one-shot*, destacándose como principal aporte de este trabajo de investigación y profundizando en los pasos que contempla. Los resultados obtenidos de estudios

preliminares, así como aquellos que se obtienen tras la aplicación y evaluación de la metodología propuesta, se presentan en el capítulo 5. La discusión de los resultados reportados se aborda en el capítulo 6, y finalmente, en el capítulo 7 se escriben las conclusiones tras la realización del trabajo y se introducen líneas de trabajo futuro que se identificaron en relación con la metodología propuesta.

2. MARCO TEÓRICO

Existen distintos enfoques metodológicos a seguir al llevar a cabo procesos investigativos. Sin embargo, de acuerdo con Daros (2002), las fases generales de tales procesos son el establecimiento de un problema, la documentación de un marco teórico, el diseño metodológico, y la realización de lo planeado.

Dentro del capítulo de introducción, se presenta una problemática referente a la complejidad existente en el reconocimiento y análisis de gestos corporales realizados por usuarios de un sistema con fines de interacción. Tomando en cuenta la relación que debe existir entre las subsecuentes etapas de la investigación, se definieron conceptos a ser explorados y definidos como referencia para el proceso investigativo y que fungen como enlace con el diseño metodológico.

En este sentido, las siguientes secciones se dedican a la presentación de conceptos relacionados con la problemática identificada. En la sección 2.1 se describen de manera breve la ergonomía y el movimiento corporal haciendo especial énfasis en su estudio desde una perspectiva de ingeniería y que permita definir correctamente los alcances del presente trabajo.

Dentro de la sección 2.2, se abordan técnicas para ML junto con su aplicabilidad para un enfoque de aprendizaje *one-shot*, favoreciendo así el enlace entre el diseño metodológico y el alcance de los objetivos presentados en el capítulo de introducción.

Posteriormente, en la sección 2.3 se introducen conceptos propios de la disciplina de HCI que aportan al trabajo instrumentos y aspectos que deben ser considerados en el diseño metodológico, pero que también serán de utilidad para llevar a cabo la evaluación de la propuesta.

2.1. Ergonomía y movimiento corporal

De acuerdo con la Asociación Internacional de Ergonomía (IEA), la ergonomía es la disciplina científica que se ocupa de la comprensión de las interacciones entre los seres humanos y otros elementos de un sistema. La intención que se sigue con el estudio de esta disciplina es poder aplicar la teoría, los principios, los datos y los métodos de

diseño para optimizar el bienestar humano y el rendimiento general del sistema en el que este se desenvuelve (Bridger, 2017).

La ergonomía contribuye a la implementación de sistemas seguros a través de tres directrices: 1) la adopción de un enfoque de sistemas usando un proceso sistemático e iterativo, 2) la orientación al diseño, y 3) la optimización del rendimiento y el bienestar de las personas (IEA, 2020). Así, es posible observar la relevancia que tiene la consideración de aspectos relacionados a la ergonomía en etapas tempranas de diseño de productos e interacciones.

Dentro del análisis que se puede realizar del movimiento corporal en seres humanos, los gestos son de particular interés, al ser acciones físicas significativas (estáticas o dinámicas) que involucran el desplazamiento físico de dedos, manos, brazos, piernas, cabeza o cara, que permiten no solo comunicarse en una forma no verbal, sino que también pueden ayudar a enriquecer la comunicación verbal (Mitra y Acharya, 2007).

2.1.1. Antropometría

La antropometría es una rama de la ergonomía dedicada a la medición de las dimensiones del cuerpo humano y a su evaluación (Doležal y Hrženjak, 2019). El objetivo de la antropometría es medir, de la manera más precisa posible, el cuerpo humano, con el fin de identificar sus características morfológicas y fisiológicas. Los dominios de aplicación de esta disciplina son diversos. En la literatura es posible encontrar trabajos de autores que se dedican a la utilización de las mediciones de grupos de personas para llevar a cabo el diagnóstico de padecimientos, para determinar posibles riesgos sanitarios asociados a la morfología, para describir guías de diseño de productos y para la creación de simuladores virtuales entre otras aplicaciones (Laal et. al, 2018), (Hanson et. al, 2009), (Littman et. al, 2007) y (Bastanfard et. al, 2004).

Con respecto a la relación entre la antropometría y las ciencias computacionales, es posible observar en la literatura que además de las aplicaciones sobre la creación de simuladores virtuales, se han aplicado modelos ingenieriles para la predicción de eventos a partir de datos obtenidos de forma antropométrica (Pytel et. al, 2015). Igualmente, la conjunción de ambas áreas de la ciencia ha permitido llevar a cabo

trabajos para la identificación y descripción de personas (Hamad et. al, 2014), para el desarrollo de sistemas de rehabilitación física (Favre et. al, 2012), y para reconocimiento gestual (Zhang et. al, 2010).

2.1.2. Postura humana

La postura es parte importante para entender el comportamiento humano. La complejidad de su tratamiento reside en la gran cantidad de posturas que son permitidas por los amplios grados de libertad de las articulaciones del cuerpo (Boulay et. al, 2005). De acuerdo con especialistas en ergonomía, el cuerpo humano tiene un total de 86 grados de libertad (ben-Aaron, 2003). El análisis de las capacidades de cada una de las articulaciones que proporcionan tales grados de libertad ha sido objeto de estudio de diversos autores debido al interés que se tiene en esta información para múltiples contextos (Boone y Azen, 1979), (Mosler et. al, 2017).

El estudio de la postura humana es relativamente nuevo dentro de las ciencias médicas. Su evaluación permite, principalmente, obtener información con respecto a posibles patologías músculo esqueléticas (do Rosário, 2014). Las técnicas para analizar la postura humana son diversas y se apoyan en la utilización de instrumental óptico, médico y tecnológico (Caputo et. al, 2018), (Chevtchenko et. al, 2018).

La postura, para efectos de este trabajo, se puede relacionar con gestos estáticos, para cuyo reconocimiento se toma en cuenta la posición de puntos de articulación, la distancia entre tales puntos y las rotaciones que realicen en los distintos ejes con referencia a una posición inicial de reposo.

2.1.3. Actividad humana

En algunas propuestas, el estudio de posturas es suficiente para identificar la información requerida para alcanzar una solución a problemáticas específicas (Li et. al, 2019). Sin embargo, para ciertos dominios se hace evidente la necesidad de contar con un análisis de comportamiento más extenso en el tiempo y que soporte el monitoreo de una sucesión de poses corporales, ya que las dimensiones agregadas de tiempo y

trayectoria de desplazamiento en el espacio permiten obtener información adicional (Cippitelli et. al, 2016).

Desde un punto de vista computacional, la complejidad agregada al análisis de actividades humanas con respecto al reconocimiento de posturas es tal, que hace necesaria la utilización de algoritmos, técnicas y equipos que permitan el rastreo de múltiples puntos de forma simultánea en contextos de tres dimensiones y que, en el mejor de los casos, exige la iteración de procedimientos de identificación y reconocimiento (Kellokumpu et. al, 2005).

En el contexto en que este trabajo de investigación tiene lugar, las actividades humanas son equivalentes a los gestos dinámicos, y son considerados como la sucesión de dos o más posturas observadas en un espacio de tiempo determinado.

2.2. Aprendizaje automático

Diversos autores señalan que el término ML fue acuñado en el año 1959 por Arthur Lee Samuel (Syam y Sharma, 2018), y utilizado para describir la intención de hacer que las computadoras pudieran aprender de reglas y experiencias con respecto al desempeño de tareas, para luego trabajar con grandes volúmenes de información de manera eficiente (Zhou, 2018).

Actualmente, la existencia de grandes conjuntos de datos y la necesidad de procesarlos para generar información en distintos dominios como el militar y en servicios de salud, ha propiciado un aumento en la cantidad de algoritmos y técnicas de ML y en la búsqueda de opciones que aporten beneficios en cuanto a tiempo de ejecución, poder computacional y simplicidad para sus usuarios (Dey, 2016).

Para el presente trabajo de investigación, es de interés conocer alternativas que podrán ser de utilidad para ser aplicadas en procesos de clasificación de datos obtenidos a través de distintas fuentes. Las siguientes subsecciones se dedican a brindar información sobre tal propósito.

2.2.1. Técnicas de aprendizaje automático

Las múltiples técnicas de ML que se encuentran documentadas han sido categorizadas por autores en revisiones a la literatura. En la propuesta de Dey (2016) se hace una clasificación de las técnicas de ML de acuerdo con los enfoques que estas siguen para sus procesos de entrenamiento.

El primer tipo de ML que se destaca es el que engloba a los algoritmos supervisados. Dentro de esta categoría se pueden encontrar, entre otros, los árboles de decisión, los clasificadores Bayesianos, las redes neuronales supervisadas y las máquinas de vector de soporte. Este tipo de clasificadores requieren de asistencia externa. Normalmente, para su utilización, las entradas son divididas en conjuntos de entrenamiento y para pruebas. El conjunto de entrenamiento contiene información etiquetada que permite que los algoritmos identifiquen patrones para ser aplicados al conjunto de prueba para evaluación (Kotsiantis et. al, 2007).

El segundo tipo dentro del ML es el aprendizaje no supervisado. Este tipo de algoritmos busca aprender pocas características sobre los datos de entrada. Posteriormente, cuando se introducen nuevos datos al modelo, este utiliza tales características aprendidas para reconocer la clase a la que estos nuevos datos pertenecen. Su uso principal se observa para reducción de dimensionalidad y agrupamiento (*clustering*). Los principales algoritmos dentro de este tipo son Análisis de Componentes Principales, las redes neuronales no supervisadas y K-medias.

En la literatura se observa también un tercer tipo de algoritmo de ML que reúne características de los dos tipos mencionados previamente, los algoritmos de aprendizaje semi supervisado. Este tipo de algoritmos es especialmente útil en áreas como la minería de datos donde se tiene información no etiquetada y llevar a cabo este preprocesamiento es una labor tediosa (Zhu y Goldberg, 2009). Existen diversas categorías dentro del aprendizaje semi supervisado, las principales son los modelos generativos, el auto entrenamiento, y las máquinas de vector de soporte para aprendizaje transductivo (Zhu, 2005).

El cuarto tipo de ML es el aprendizaje por refuerzo, el cual toma decisiones con base en qué acciones tomar de forma que la salida sea mejor. El modelo no tiene

conocimiento sobre tales acciones hasta que una situación específica se presenta. Así, la decisión tomada afecta situaciones y, por tanto, las decisiones a tomar en el futuro. Este tipo de aprendizaje se apoya de fuerte manera en la búsqueda por prueba y error (Sutton, 1992).

El aprendizaje multitarea es considerado el quinto tipo de ML. Este tipo de algoritmos tienen el principal objetivo de apoyar a que otros algoritmos se desempeñen de mejor manera. Cuando se aplica un algoritmo de este tipo a una tarea, este recuerda el procedimiento que se efectuó para resolver el problema o para alcanzar una conclusión. Luego, usa este conocimiento para intentar alcanzar la solución de otro problema o tarea aplicando un enfoque de transferencia inductiva con respecto a otros algoritmos adicionales. La premisa en la que se apoyan esta clase de algoritmos es que, al compartir la experiencia o aprendizaje entre dos mecanismos de ML, el proceso puede ser más rápido (Caruana, 1997).

El sexto tipo de algoritmo es llamado de ensamble o combinado. Cuando varios mecanismos de ML se combinan para formar uno solo se obtiene un mecanismo de aprendizaje de ensamble. De este tipo de algoritmos, se ha observado que generalmente tienen un mejor desempeño que sus contrapartes individuales (Opitz y Maclin, 1999). Dos de las técnicas de aprendizaje de esta clase más utilizadas son el método de impulso (*boosting*) y el de empaquetado (*bagging*).

El último tipo de algoritmo de ML por describir es el de aprendizaje basado en instancias. En este tipo de aprendizaje, el algoritmo aprende un tipo específico de patrón y trata de aplicarlo a cualquier dato que ingresa. La complejidad de estos algoritmos aumenta junto con el tamaño de los datos. Un ejemplo de este tipo de técnicas es el algoritmo de los K vecinos más cercanos (*K-nearest neighbor*) (Hu et. al, 2016).

2.2.2. Enfoque de aprendizaje *one-shot*

Desde la década de 1960, investigadores de las áreas de estudio de la interacción y del conocimiento han usado el término de aprendizaje *one-shot* como una referencia de la adquisición de conocimiento tras la exposición a la información una sola vez (Mamoshina et. al, 2018). Durante 1980 se propusieron métodos para aplicar este

enfoque como una mejora y característica deseable de los algoritmos de ML (Pollack y Barnden, 1988). Además, Vo (1994) propuso el uso de técnicas siguiendo este enfoque para permitir la adaptación de sistemas de reconocimiento para nuevos usuarios.

Actualmente, es común encontrar propuestas aplicando técnicas de aprendizaje *one-shot* para llevar a cabo Procesamiento de Lenguaje Natural (NLP) (Zhang y Zhao, 2018), sin embargo, el entrenamiento exhaustivo reporta aún mejores resultados en cuanto a precisión y tiempo de respuesta.

Tomando en cuenta que el lenguaje corporal es parte de la comunicación humana y considerando que las tareas comprendidas en el NLP son equivalentes a las que se llevan a cabo para el reconocimiento gestual, entendiendo además la necesidad de que múltiples usuarios interactúen con sistemas usando gestos, las diferencias culturales que entre ellos pueden existir, y que no es viable pedir a cada usuario que repita un gesto las veces suficientes para entrenar un clasificador de manera tradicional, es posible observar la necesidad de contar con métodos apoyados por el enfoque de *one-shot* para llevar a cabo interacción gestual (Zhang et. al, 2017), (Escalante et. al, 2017) y (Konečný y Hagara, 2014).

El reto actual sobre la aplicación del enfoque de aprendizaje *one-shot* específicamente para aplicaciones basadas en gestos se encuentra en alcanzar un nivel de precisión alto equivalente al de otras alternativas y en disminuir los tiempos de respuesta para mejorar la percepción de parte de los usuarios en cuanto a usabilidad y UX.

2.3. Interacción Humano-Computadora

Como su nombre lo indica, la HCI es la disciplina dentro de las ciencias computacionales que se encarga del estudio de la relación entre las personas y los ordenadores, de su entendimiento mutuo, y del desarrollo de sistemas que faciliten a sus usuarios la realización de tareas (Bansal y Khan, 2018). Este propósito puede ser aproximado siguiendo distintas perspectivas, siendo una de ellas el estudio de las modalidades de interacción.

Entre las modalidades de interacción que se encuentran en la literatura (White, 2018), destacan para este trabajo de investigación aquellas referidas como Interfaces Naturales de Usuario (NUI).

2.3.1. Interfaces Naturales de Usuario

La comunidad de HCI acuñó el término de NUI para referirse a aquellas interfaces que hacen uso de mecanismos de interacción que se asemejan a la forma en la que las personas se comunican entre ellas. El origen del término, de acuerdo con algunos autores, se da al explorar nuevas modalidades de interacción no relacionadas con el uso de dispositivos como el *mouse* (Mann, 2002). Dentro de las NUI, se pueden encontrar interfaces vocales, interacción táctil, dispositivos vestibles e interfaces gestuales entre otras opciones (Fernandez et. al, 2016).

En el área de las Ciencias de la Computación, el reconocimiento gestual se aplica a diversos dominios tales como la interacción niño-computadora; identificación y detección forense; rehabilitación y monitoreo médico; navegación y manipulación en ambientes virtuales; aprendizaje a distancia; seguridad; entrenamiento atlético; y aplicaciones de entretenimiento (Schlömer et. al, 2008), (Liang y Ouhyoung, 1998), (Daniloff et. al, 1982) y (Jia et. al, 2007). Además, existe un constante desarrollo de hardware y software soportando esta modalidad de interacción.

El reconocimiento gestual es un problema para el que típicamente se proponen soluciones que incluyen dos etapas: identificación gestual, y asignación de significados a gestos. La identificación gestual se logra comúnmente a través de dispositivos vestibles, usando técnicas basadas en visión, o por medio del procesamiento de señales inalámbricas. La interacción a través de dispositivos vestibles exige que los usuarios carguen equipos incómodos (Norman, 2010), mientras que las técnicas basadas en visión suponen lidiar con otros problemas principalmente relacionados con la oclusión (Malima et. al, 2006), y el procesamiento de señales inalámbricas no ha sido explorado profundamente e involucra el uso de dispositivos como radares y enrutadores (Wan et. al, 2014). Cada clase de dispositivo varía en cuanto a distintas dimensiones, incluyendo precisión, resolución, latencia, rango de movimientos detectables, comodidad para el usuario, y costo (Mitra y Acharya, 2007).

2.3.2. Usabilidad

De acuerdo con la norma ISO/IEC 25010 de la Organización Internacional de Normalización (ISO), la usabilidad se define como la “capacidad del producto de software para ser entendido, aprendido, usado y resultar atractivo para el usuario” (ISO, 2011). En otras definiciones como en la aportada por Nielsen (1994) se indica que además de ser “un atributo relacionado con la facilidad de uso”, la usabilidad es la “aceptación práctica y provechosa del producto”, extendiendo el concepto de satisfacción de usuario con el objetivo de evaluación. Complementando esto, la norma ISO 9241-11 menciona que la usabilidad es “el grado en que un producto puede ser utilizado por usuarios específicos para alcanzar objetivos específicos con efectividad, eficiencia y satisfacción en un contexto específico de uso” (ISO, 2018).

El diseño de sistemas usables se basa en 5 componentes reportados en investigaciones al respecto (Cappel y Huang, 2007), y que pueden ser asociados a preguntas:

- ¿Qué tan fácil es para los usuarios cumplir con tareas básicas la primera vez que se enfrentan a un diseño? (facilidad de aprendizaje).
- Una vez que los usuarios conocen el diseño, ¿Qué tan rápido pueden desempeñar tareas? (eficiencia).
- Cuando los usuarios vuelven a utilizar el diseño tras un período de ausencia, ¿Qué tan fácilmente recuperan la habilidad en su uso? (memorabilidad).
- ¿Cuántos errores cometen los usuarios?, ¿Qué tan graves son los errores?, y ¿Qué tan fácil es para los usuarios recuperarse de los errores cometidos? (manejo de errores).
- ¿Qué tan agradable es el uso del diseño para los usuarios? (satisfacción).

La descripción de estos componentes es consistente con las dimensiones de la usabilidad que son reportadas por otros autores Interaction Design Foundation, (2020).

La evaluación de la usabilidad puede ser aproximada por medio de dos enfoques, el primero considerando la interacción de los sistemas y los usuarios, mientras que, en el

segundo se pone a prueba el diseño de un producto con respecto a criterios establecidos y previamente validados.

Si bien, en la literatura se pueden encontrar diversos protocolos para la evaluación de usabilidad en productos de software con usuarios, es de especial interés para este trabajo la técnica del Mago de Oz (Höysniemi et. al, 2004). Este método consiste en diseñar una experiencia similar a la que un usuario tendría al interactuar con un sistema sin necesidad de llegar a la implementación de esta. A través de la utilización de recursos multimedia y prototipos de interfaz, es posible así llevar a cabo experimentos que permitan observar cómo se comportaría el usuario ante ciertas tareas que se espera que sea capaz de realizar con el sistema a desarrollar y así, identificar errores potenciales en etapas tempranas de desarrollo.

La técnica del Mago de Oz ha sido utilizada a lo largo de la historia del desarrollo de sistemas interactivos, y en particular, en el campo del desarrollo de las NUI como una forma para recolectar datos en ambientes de realidad mixta (Dow et. al, 2005) y para encontrar comandos gestuales para interacción entre otros dominios de aplicación. La selección de esta estrategia para la presente propuesta se basa en la intención de evitar el desarrollo de sistemas complejos y enfatizar el esfuerzo en el análisis gestual sobre la implementación de mecanismos de reconocimiento.

A partir del uso de una técnica como la reportada en líneas anteriores, es posible obtener datos de la interacción entre usuarios y un producto. Con el fin de generar información relevante a partir de tales datos, es necesario hacer uso de instrumentos como el Cuestionario de Usabilidad de Sistemas Computacionales (CSUQ por sus siglas en inglés), que se compone de 19 preguntas, con una escala Likert de 7 puntos acerca del uso del sistema (preguntas 1 a la 8), calidad de la información (preguntas 9 a la 15), calidad de la interfaz (preguntas 16 a la 18) y satisfacción general con el sistema (pregunta 19). Al ser aplicado a usuarios de un producto, el análisis de los resultados realizado en forma de promedios y desviaciones estándar brinda información que permite identificar áreas de mejora y diversidad de las opiniones recibidas (Lewis, 2018).

Otra propuesta sobre instrumentos para la evaluación de usabilidad de sistemas fue realizada por John Brooke en 1986 y llamada Escala de Usabilidad de Sistemas (en

inglés *System Usability Scale, SUS*). Este cuestionario hace uso de una escala Likert de 5 puntos y cuenta con 10 preguntas sobre la apreciación de la usabilidad de un sistema. La forma en la que se procesan los resultados de este cuestionario difiere del CSUQ ya que, para este caso, se pide que se siga el siguiente algoritmo:

- Para cada una de las preguntas impares, se debe restar uno a la puntuación dada ($y_i = x_i - 1$, x es el valor respondido entre 1 y 5, i es el número impar de la pregunta entre 1 y 10).
- Para cada una de las preguntas pares, se debe restar la puntuación dada a 5 ($y_i = 5 - x_i$, x es el valor respondido entre 1 y 5, i es el número par de la pregunta entre 1 y 10).
- Tomar los valores obtenidos y sumarlos todos para obtener una puntuación total. Luego, multiplicar 2.5 por el valor obtenido ($u = 2.5 * \sum_{i=1}^{10} y_i$).

El resultado de estas operaciones (u) da un número en una escala de 100 puntos que, si bien no representa un porcentaje, puede dar ideas claras de la apreciación de los usuarios con respecto al sistema con el que interactuaron. De acuerdo con la literatura, calificaciones sobre 80 puntos son deseables, calificaciones alrededor de los 70 puntos indican pocos aspectos por mejorar, y calificaciones menores a los 50 puntos indican la presencia de múltiples errores graves de usabilidad (Lewis y Sauro, 2018).

Por otra parte, la evaluación sin usuarios se conoce en la literatura como evaluación heurística y se realiza enfrentando un diseño ante una serie de criterios predefinidos que deberían ser satisfechos. Ejemplos de criterios para evaluación heurística son los reportados en (Bastien y Scapin, 1992) y (Nielsen, 1995).

2.3.3. Experiencia de Usuario

El concepto de usabilidad, en combinación con otros como la utilidad, accesibilidad y la credibilidad, se engloban en un término de carácter más general llamado UX Interaction Design Foundation, (2020).

El concepto de UX es definido como la percepción y la respuesta de las personas con respecto al uso o a la anticipación de la utilización de un producto, sistema o servicio (Schrepp et. al, 2014).

En la revisión a la literatura de Maia y Furtado (2016), se destacan aspectos importantes de la UX, tales como las técnicas más utilizadas para llevar a cabo su evaluación, siendo estas la aplicación de cuestionarios, entrevistas, encuestas en línea, grabaciones en video y observación. Así mismo, destacaron como componentes de la UX (aspectos a evaluar), contar con un diseño usable, atractivo, útil, que el usuario pueda controlar, que le genere un sentido de pertenencia y que satisfaga sus requerimientos.

2.3.4. Modelado de contexto

Considerando la condición descrita en el capítulo de introducción, sobre que los gestos son específicos a culturas, lenguajes y contextos, es importante contar con un mecanismo que permita llevar a cabo la representación de estas características, favoreciendo el proceso analítico y las tareas de identificación gestual y asociación semántica.

En (Calvary et. al, 2003), se propone la representación del contexto como un conjunto de tres aspectos fundamentales: el usuario final del sistema, la plataforma de cómputo (incluyendo tanto software como hardware) con la que el usuario debe llevar a cabo una tarea interactiva, y el ambiente físico en el que las acciones tienen lugar. Así mismo, se define una interfaz sensible al contexto como aquella interfaz que muestra la capacidad de ser consciente del contexto (*context awareness*) y de reaccionar a los cambios de este. Esta comprensión del contexto ha sido adoptada y explotada ampliamente por la comunidad de investigación de HCI debido su expresividad y a que reduce la complejidad en el análisis (Sassi et. al, 2017), (Molina et. al, 2012).

2.3.5. Tareas abstractas

Definir las instrucciones a realizar en términos de tareas estándar o abstractas, proporciona la oportunidad de generalizar los resultados tras la culminación de un experimento y predecir comandos que pueden ser utilizados para otras concretizaciones de ellas. Es decir, que, a partir de un estudio de elicitación para un

contexto específico, sea posible generar una serie de comandos que pueden ser empleados en otro contexto.

Con el objetivo de modelar sistemas informáticos, Constantine (2003) propuso las tareas estándar que se llevan a cabo en los mismos, junto con los componentes y prototipos abstractos que les dan soporte. Estas tareas estándar, determinadas por verbos, son acciones que se efectúan típicamente y que, al agregarles sustantivos, definen tareas específicas de las soluciones informáticas creadas. Resulta de interés que las tareas corresponden a funciones interactivas y no a componentes de interfaz. Las tareas abstractas identificadas son las siguientes: operar, ir a, iniciar, detener, completar, seleccionar, crear, borrar, eliminar, modificar, mover, duplicar, ejecutar, alternar y ver. La relación entre estas tareas y concretizaciones de estas se muestra en la Tabla 1.

Tabla 1. Relación entre tareas abstractas y sus concretizaciones.

Acción estándar	Ejemplo de uso
Acción / operación	Imprimir tabla de símbolos, colorear forma seleccionada
Iniciar / Ir / Hacia	Realizar prueba de consistencia, confirmar compra
Detener / Terminar / Completar	Terminar sesión de inspección, interrumpir prueba
Seleccionar	Selector de miembro de grupo, selector de objetos
Crear	Nuevo cliente, diapositiva en blanco
Eliminar, Borrar	Romper conexión, borrar formulario
Modificar	Cambiar dirección de envío, editar detalles de un cliente
Mover	Agregar a lista de direcciones, mover arriba/abajo
Duplicar	Copiar dirección, duplicar diapositiva
Ejecutar (y regresar)	Formatear objetos, asignar distribución de impresión
Alternar	Activar/desactivar negritas, modo encriptado
Ver	Mostrar detalles de archivo, cambiar a resumen

Es importante mencionar la existencia de otras propuestas en este mismo sentido como la de Wüthrich (1999), en donde se resume la interacción con ambientes virtuales en tres acciones atómicas: seleccionar, posicionar y deformar.

2.4. Conclusiones del capítulo

Dentro del presente capítulo se presentaron conceptos y teorías que resultan relevantes para la realización de este trabajo de investigación. En la Tabla 2 se puede

observar de manera resumida la relación que existe entre los objetivos específicos y las preguntas de investigación con respecto a los conceptos aquí descritos.

Tabla 2. Relación entre elementos del marco epistémico y conceptos descritos en el marco teórico.

Elemento del marco epistémico	Conceptos del marco teórico	Justificación
O. E. 1. Realizar un mapeo de usuario en 3D por medio de dispositivos basados en visión y API para permitir el monitoreo de movimientos y el reconocimiento gestual.	Ergonomía y comportamiento humano.	Los componentes incluidos en el marco teórico que hacen referencia al estudio del comportamiento humano hacen posible definir los puntos a analizar y a considerar para la modelación de usuarios.
O. E. 2. Diseñar e implementar una base de datos de gestos a partir de información obtenida durante estudios preliminares, para permitir la comparación y el análisis.	Ergonomía, comportamiento humano y HCI.	El conocimiento de las métricas y puntos de interés permitirá el diseño de la estructura de la base de datos. Los conceptos relacionados a HCI permitirán obtener la información correspondiente a gestos y las tareas abstractas serán la base para determinar los gestos de interés.
O. E. 3. Proponer un método para etiquetado de gestos utilizando una notación genérica, para soporte a las tareas de reconocimiento y de asociación de significados.	HCI y Tareas abstractas.	La asociación de gestos con tareas de carácter genérico permite su etiquetado y posterior asociación con significados específicos.
O. E. 4. Implementar reconocimiento gestual utilizando técnicas de ML y siguiendo un enfoque de aprendizaje <i>one-shot</i> para permitir interacción natural.	ML y HCI.	La implementación del reconocimiento gestual requiere soporte de técnicas y enfoques. El modelo de contexto proporciona información adicional para identificar el significado de un gesto al ser realizado por un usuario, en un ambiente y por medio de una plataforma específica.
P. I. 1. ¿Cómo se puede lograr la interacción gestual de manera que resulte natural para los usuarios?	HCI.	Se requieren métricas e instrumentos para determinar si las NUI están siendo correctamente implementadas y la apreciación de los usuarios con respecto a estas.
P. I. 2. ¿Qué algoritmos, técnicas, métodos y estrategias se deben incluir en una metodología para facilitar la implementación de aplicaciones con interacción basada en gestos?	ML y HCI.	Se debe proponer la utilización de técnicas de ML que soporten los modelos propuestos con el fin de alcanzar la concretización de la metodología propuesta para desarrollo de NUI.
P. I. 3. ¿Qué implicaciones resultan de la utilización de gestos definidos por el usuario en una evaluación de usabilidad y	Ergonomía y HCI.	Por medio de técnicas e instrumentos de evaluación, será necesario validar los productos basados en interacción natural que

experiencia de usuario, realizada sobre
aplicaciones basadas en interacción gestual?

sean implementados por medio del uso de la
metodología propuesta.

El marco teórico desarrollado permite, además, sentar las bases de la investigación del estado del arte cuyo reporte se incluye en el siguiente capítulo.

3. ESTADO DEL ARTE

En los proyectos de investigación, una revisión temprana a la literatura permite definir los fundamentos de la investigación, facilita enfocarse en temas relevantes para su realización, y apoya en la definición y acotación de preguntas de investigación. Con el fin de llevar a cabo una revisión sistemática de la literatura, en el presente trabajo se siguió una metodología de 8 pasos propuesta por Okoli y Schabram (2010). Como parte del protocolo definido, se determinó considerar únicamente trabajos con menos de 10 años de antigüedad, escritos en inglés y en español, y publicados en revistas de investigación, conferencias organizadas por IEEE y ACM, o en capítulos de libro.

La cadena de búsqueda utilizada incluyó los siguientes términos y combinaciones de ellos: “interacción gestual”, “aplicación basada en gestos”, “interfaces de usuario naturales”, “aprendizaje *one-shot*”, “reconocimiento gestual”, “gestos definidos por el usuario”, “identificación gestual”, “reconocimiento de acciones humanas”, “reconocimiento de actividades humanas”, y “reconocimiento de movimientos” tanto en inglés como en español. El proceso seguido para revisión de la literatura, de acuerdo con las características descritas, resultó en la documentación de trabajos que permiten conocer el estado del arte propio de la investigación como se puede observar en la subsección 3.2.

Así, a través del seguimiento de la metodología fue posible encontrar en la literatura los trabajos de diversos autores abordando el problema de reconocimiento gestual desde distintos enfoques, haciendo uso de dispositivos tales como cámaras de profundidad, cámaras de visión estéreo y 3D, acelerómetros y otros artículos vestibles, así como aplicando técnicas para ML como los modelos ocultos de Markov, redes neuronales, máquinas de vector de soporte y máquinas de estado finito entre otros (McCormick et. al, 2014), (Wu, et. al, 2016), (Pisharady y Saerbeck, 2015), (LaViola Jr, 2015). Las siguientes subsecciones se dedican a la descripción de los trabajos hallados y de interés para la comprensión del presente trabajo de disertación.

3.1. Modelos de usuario

Para los propósitos de este trabajo de investigación, y alineado con la definición del problema, esta subsección se enfoca en la descripción de métodos para modelar los movimientos que los usuarios potencialmente pueden realizar.

Una propuesta al respecto es la de Kaklanis et. al (2013), en la que se detalla un Modelo de Usuario Virtual Genérico (GVUM) y se asocia con un modelo de capacidades.

De acuerdo con lo que se puede observar en el modelo GVUM presentado en la Figura 1, para llevar a cabo el análisis de capacidades, el cuerpo humano se forma de seis componentes principales: cuello, columna vertebral, extremidad superior izquierda, extremidad superior derecha, extremidad inferior izquierda, y extremidad inferior derecha.

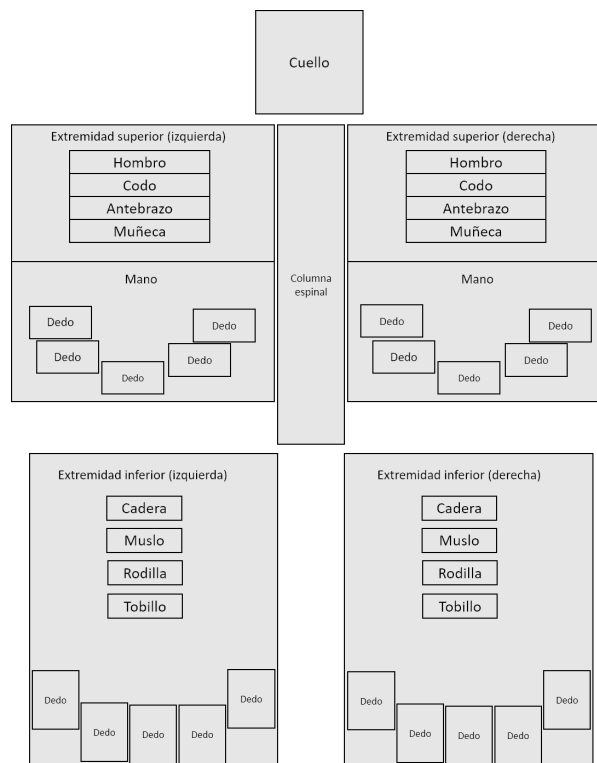


Figura 1. Representación del Modelo Virtual de Usuario Genérico de Kaklanis et. al (2013).

En un nivel más detallado, las extremidades superiores incluyen hombros, codos, antebrazos, muñecas, y manos con dedos, mientras que las extremidades inferiores incluyen a su vez la cadera, muslos, rodillas, tobillos y dedos del pie.

Al utilizar el GVUM, es posible modelar a los usuarios de manera anatómica, pero no lo es representar su interacción con las interfaces de un sistema. Para satisfacer este objetivo, el modelo se acompaña de un modelo de capacidades que consta de 108 clases correspondientes a sistemas, partes del cuerpo, características y movimientos realizables tales como rotaciones, flexiones, extensiones, aducciones y abducciones por mencionar algunos. La capacidad para identificar los movimientos que los usuarios pueden realizar al interactuar con sistemas permite llevar a cabo experimentos en los que se identifiquen los gestos que estos usan con mayor frecuencia y su empareje con tareas específicas.

Una segunda opción para la representación de usuarios de un sistema son los estereotipos de usuario. Una metodología para desarrollar estereotipos de usuario se conoce como Personas. Las Personas son representaciones de los objetivos y del comportamiento de un grupo real de usuarios. Normalmente, el proceso metodológico para su definición inicia con entrevistas u observación y resulta en la generación de entidades ficticias (Singh, 2019). La utilidad de las personas radica en su fácil comprensión y en su capacidad para presentar cualidades relevantes de usuarios potenciales de un producto. Entre las ventajas que se reportan en la literatura con respecto al uso de este enfoque del desarrollo de Personas, se destacan que estas suelen ser más fáciles de recordar y que la traducción de sus necesidades a requerimientos es sencilla (An et. al, 2018).

El desarrollo de Personas es comúnmente integrado a procesos de diseño centrado en el usuario (Vallet et. al, 2020), ya que promueve fuertemente el entendimiento de las necesidades del grupo de usuarios de un producto, al tiempo que genera un sentido de pertenencia en el equipo de diseño quienes tienen así la sensación de trabajar para alguien específico y no solo para satisfacer un documento de requerimientos.

3.2. Trabajos relacionados

Siguiendo el protocolo mencionado para la revisión sistemática de la literatura y considerando la pertinencia para el trabajo de investigación, se seleccionaron 60 trabajos relacionados notando que la mayoría incluían como parte de sus objetivos el

desarrollo de aplicaciones basadas en interacción gestual. Así, se llevó a cabo el análisis en términos de las tres etapas que comúnmente se aplican para llevar a cabo el reconocimiento gestual, obteniendo información acerca de los dispositivos utilizados como fuente para la adquisición de datos, las partes del cuerpo cuyos gestos se desea reconocer, y las técnicas de clasificación empleadas en las propuestas.

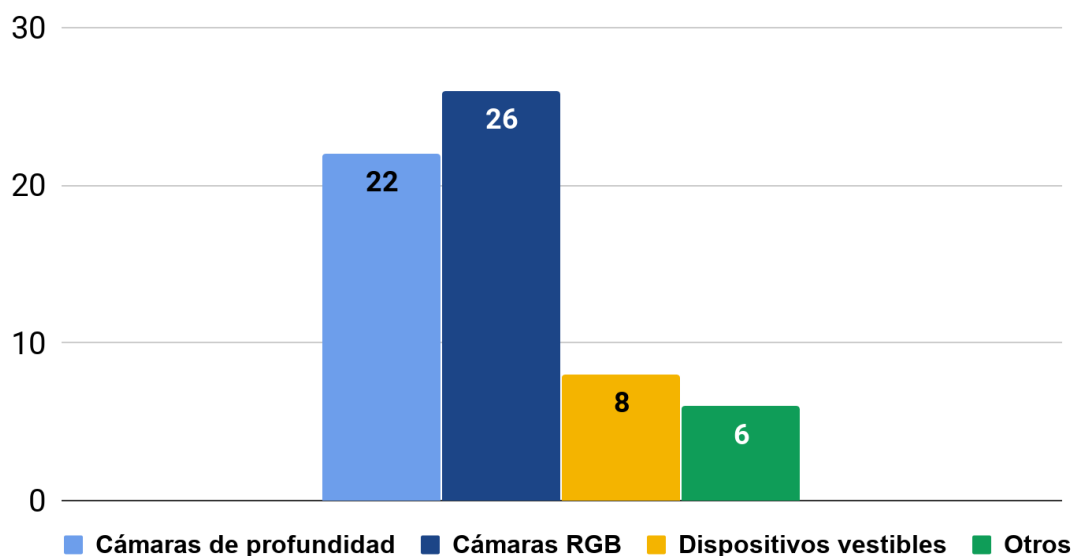


Figura 2. Distribución de los dispositivos para adquisición de información utilizados en la literatura. Elaboración propia.

Como se puede observar en la Figura 2, el 77% de los trabajos analizados hacen uso de dispositivos basados en visión para obtener información sobre movimientos realizados por el usuario y para permitir la interacción gestual. Según lo reportado, el motivo principal por el que se decidió usar este tipo de dispositivos es que los equipos como cámaras y sensores *Leap Motion* resultan ser menos invasivos que los dispositivos vestibles. Los enfoques basados en visión se enfrentan a problemáticas específicas como la oclusión, sin embargo, este problema ha sido resuelto por medio de la utilización de una mayor cantidad de dispositivos para obtener la información sin pérdidas significativas (Tran y Trivedi, 2011) y (Kuo et. al, 2016).

Por otra parte, en 13% de los trabajos analizados se propone el uso de dispositivos vestibles tales como guantes interactivos, lentes para realidad aumentada, brazaletes para detección de movimientos de los brazos, y trajes de cuerpo completo equipados

con sensores. Si bien este tipo de dispositivos se adhieren al cuerpo de los usuarios y permiten un mejor monitoreo de los movimientos realizados en términos de precisión y eficiencia, el reto cuando se utilizan para permitir interacción gestual, está en proveer comodidad y en evitar que la necesidad de vestir un dispositivo para interactuar con las aplicaciones interfiera en el desarrollo de las tareas (Norman, 2010).

El resto de las propuestas exploran el uso de otros medios como señales de *Wi-Fi* y radar logrando resultados prometedores cercanos a los de las alternativas basadas en visión, pero con la desventaja de tener un costo mayor y la falta de disponibilidad para su uso masivo (Wan et. al, 2014), (Al-qaness y Li, 2016).

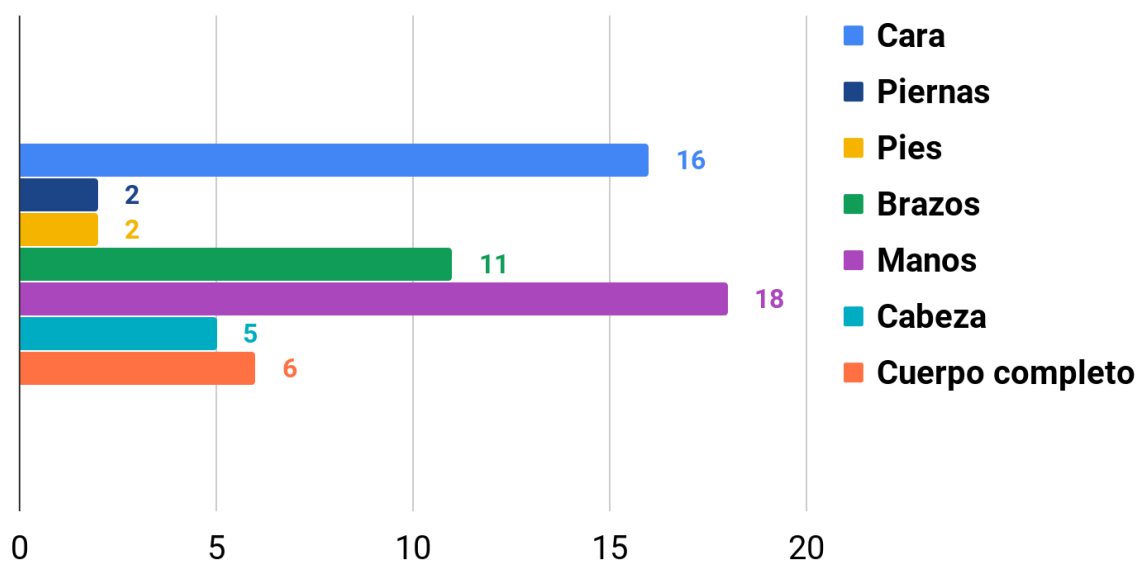


Figura 3. Distribución de las partes del cuerpo para las que se lleva a cabo reconocimiento gestual en la literatura. Elaboración propia.

Con respecto a las partes del cuerpo para las que se lleva a cabo reconocimiento gestual en los trabajos analizados, como se puede ver en la Figura 3, el 27% de ellos se enfocan en el estudio de gestos de las manos usando sensores Leap Motion, cámaras de profundidad, dispositivos vestibles y cámaras RGB. Este hallazgo es consistente con lo que se reporta en otras revisiones de la literatura del área (Mitra y Acharya, 2007) y se relaciona con dos principales motivos: que los usuarios comúnmente interactúan con dispositivos electrónicos usando solamente sus manos, y que múltiples estudios se

dedican al desarrollo de aplicaciones para la traducción del lenguaje de señas (Ghotkar et. al, 2014), (Halim y Abbas, 2015).

En contraste, solo el 7% de los trabajos revisados se enfocan en la descripción de interacción utilizando las extremidades inferiores por medio de dispositivos vestibles y cámaras de profundidad. Las extremidades inferiores conformadas por la cadera, piernas y pies, son menos expresivas que su contraparte superior, sin embargo, si se consideran escenarios de navegación en ambientes virtuales, usuarios con movilidad restringida en las extremidades superiores, o contextos en los que la parte superior del cuerpo ya está siendo utilizada para un propósito distinto al de interacción con aplicaciones, vocabularios de gestos que incluyen la parte inferior del cuerpo cobran sentido y relevancia.

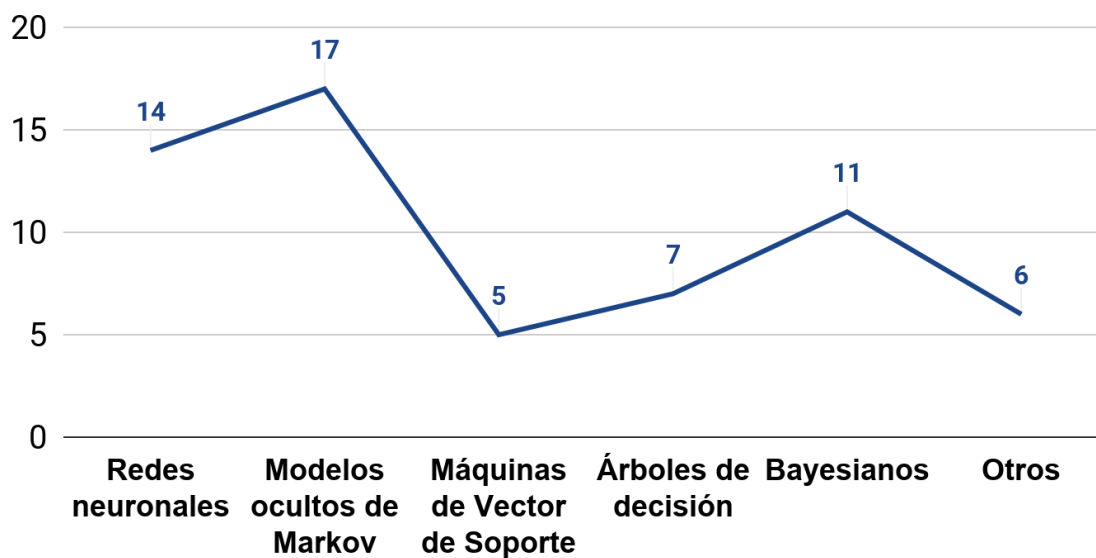


Figura 4. Frecuencia de uso de técnicas de clasificación para reconocimiento gestual empleadas en la literatura. Elaboración propia.

Para este trabajo de investigación, fue relevante encontrar que solamente el 10% de los trabajos analizados consideran la interacción con todo el cuerpo, y que la mayoría de ellos, aproximadamente el 83%, utilizan cámaras de profundidad como dispositivos de adquisición de información. Además, en la revisión a la literatura fue posible observar que existe un incremento significativo en patentes y publicaciones acerca de detección de movimiento corporal desde 2007, coincidiendo con la introducción de

cámaras de profundidad y sensores infrarrojos para propósitos domésticos y con la inversión en NUI por parte de compañías como Microsoft, Sony y Google.

Finalmente, para tener una mejor idea acerca de las herramientas para clasificación que actualmente se aplican para llevar a cabo el reconocimiento gestual y para promover su utilización como parte de la metodología propuesta, fue posible notar, como se ilustra en la Figura 4, que las técnicas utilizadas más comúnmente son los modelos ocultos de Markov y las redes neuronales, mientras que el uso de máquinas de vector de soporte y árboles de decisión no ha sido suficientemente explorado. En las siguientes subsecciones se proporciona información relevante acerca de las propuestas revisadas, organizada de acuerdo con el tipo de dispositivos usados para su implementación.

3.2.1. Interacción gestual usando dispositivos basados en visión

La propuesta de Chen et. al (2003), aborda el desarrollo de un método para reconocer gestos realizados con las manos sobre fondos estacionarios. En esta propuesta, se utiliza un algoritmo de monitoreo y extracción en tiempo real para delimitar y rastrear la mano que se está moviendo. Luego, mediante el Descriptor de Fourier y un análisis de movimientos para caracterizar propiedades de espacio y tiempo se lleva a cabo el entrenamiento de un modelo oculto de Markov para clasificación de gestos. Como parte de este trabajo, se condujo un experimento en el que se solicitó a veinte usuarios realizar veinte gestos en tres ocasiones, resultando en la captura de 1,200 (20 x 20 x 3) registros. Con esta información, la propuesta alcanzó una precisión entre el 85 y el 93% al reconocer gestos previamente definidos. El reconocimiento gestual efectuado siguió un enfoque de tiempo real y se apoyó en el uso de técnicas de procesamiento digital de imágenes. Desafortunadamente, los investigadores no especificaron el dispositivo utilizado para la captura de datos, y los resultados de su experimento no fueron contrastados con trabajos relacionados.

Otra propuesta que considera el enfoque de tiempo real es la reportada en (Li y Jarvis, 2009). En esta investigación, se utilizó una cámara de profundidad, se aplicaron múltiples técnicas para preprocesamiento de los datos, se llevó a cabo el reconocimiento usando el método de emparejamiento de *Chamfer*, y se identificaron

trayectorias por medio de una máquina de estado finito. El experimento realizado en esta propuesta consistió en evaluar reconocimiento de formas e identificación de trayectorias. Para el reconocimiento de formas, se realizaron pruebas con ocho gestos previamente definidos, logrando una precisión del 97%, mientras que, para la detección de trayectorias, la precisión reportada fue del 88%.

En el trabajo reportado en (Marin et. al, 2014), se presentaron dos algoritmos para reconocimiento de gestos producidos con las manos utilizando sensores *Leap Motion* y Kinect. En este trabajo en particular, el reconocimiento gestual fue realizado siguiendo un enfoque similar al que se aplica en otros campos del NLP, es decir, obteniendo características por medio de dispositivos o sensores, entrenando un clasificador de Máquina de Vector de Soporte, generando clases, y finalmente llevando a cabo el proceso de reconocimiento. Los resultados fueron presentados en términos de la descripción de un experimento en el que catorce usuarios realizaron diez gestos previamente definidos en diez oportunidades, resultando en un total de 1,400 (14 x 10 x 10) registros. Tales gestos fueron capturados al mismo tiempo usando ambos sensores, *Leap Motion* y Kinect. Al utilizar el sensor *Leap Motion*, la mayor precisión que alcanzaron fue del 81%, mientras que al utilizar el sensor Kinect, se obtuvo una precisión del 90%. Finalmente, al combinar la información obtenida por medio de ambos dispositivos, se logró una precisión máxima del 91%.

Igualmente, en la literatura es posible encontrar trabajos que muestran aplicaciones del reconocimiento gestual en contextos reales. Tal es el caso del trabajo en (Molchanov et. al, 2015), en el que se destacó que los sistemas de reconocimiento gestual aportan de manera significativa a las interfaces automotrices ya que brindan seguridad y comodidad. En tal trabajo, se enfocaron en proponer un algoritmo basado en el uso de redes neuronales convolucionales para identificar los gestos de conductores, alcanzando una efectividad del 77.5% al probar un conjunto específico de gestos. La propuesta no utilizó ningún dispositivo de entrada, sino que se utilizó el conjunto de datos VIVA (Martin et. al, 2015), conteniendo 885 registros correspondientes a 19 gestos efectuados por 8 participantes dentro de un vehículo y capturados usando un dispositivo Kinect.

Siguiendo el enfoque de primero identificar y luego reconocer, la propuesta de Chen et. al (2007), se basa en dos niveles para el reconocimiento de gestos con las manos en tiempo real. En el nivel inferior, las características *Haar* y el algoritmo de aprendizaje *AdaBoost* (Freund y Schapire, 1997) se aplican para el reconocimiento de posturas, proporcionando un rendimiento en tiempo real y una alta precisión de reconocimiento, mientras que en el nivel superior se implementa el reconocimiento lingüístico de gestos con la mano, utilizando un análisis sintáctico basado en gramática libre de contexto a través de una estructura de cascada paralela de clasificadores. El experimento reportado consistió en recoger alrededor de 450 muestras realizando 4 gestos diferentes con diferentes escalas y rotaciones utilizando una cámara web. Para cada gesto, se utilizó una cantidad determinada de clasificadores, con precisiones entre el 97 y el 98%.

Los sistemas de reconocimiento robustos para permitir la interacción gestual fueron objeto de estudio de Ren et. al (2011) y Ren et. al (2013). Los autores de ambas propuestas centraron su trabajo en abordar los dos grandes retos del reconocimiento gestual: la identificación de objetos y la detección de gestos. Para ello, y apoyado por el uso de un sensor Kinect y un brazalete negro en la muñeca del usuario como referencia, la segmentación de la mano se realizó utilizando el mapa de profundidad y el color del dispositivo de entrada, mientras que el reconocimiento de gestos se consigue a través de la medida de la disimilitud y la coincidencia de plantillas. Como dominios de aplicación, la computación aritmética y los juegos se identifican y utilizan como casos de estudio. No se informó de experimentos con usuarios sobre esas propuestas, lo que dificulta la estimación de la exactitud del algoritmo y su comparación con otras alternativas.

Así mismo, se patentó un sistema y un método para reconocer los gestos de las manos (Grzeszczuk et. al, 2004). En esta patente, es interesante notar que el proceso típico para el reconocimiento de gestos se descompone en cuatro etapas: 1) procesamiento de imágenes, 2) generación de plantillas, 3) identificación y 4) clasificación. Considerando el interés de este trabajo en un enfoque de aprendizaje *one-shot*, se puede observar que en tal propuesta se aborda la posibilidad de crear dinámicamente plantillas que ayuden a un mayor reconocimiento. La identificación de

las manos y el reconocimiento de los gestos se lograron en este trabajo utilizando una cámara estéreo como dispositivo de entrada y aplicando la segmentación de color y la concordancia de plantillas.

Al igual que en otros campos de interés para la HCI, el reconocimiento de gestos corporales es aplicable a contextos en los que las modalidades tradicionales no pueden ser utilizadas, así como cuando las características propias de los usuarios hacen necesario proporcionar medios alternativos de entrada. Este es el caso de la propuesta de Starner et. al (1998), en la que se aplica el reconocimiento de gestos realizados con las manos para el desarrollo de dos sistemas de reconocimiento e interpretación del Lenguaje de Señas Americano, y que resultó en la aplicación de dos experimentos. En el primero, la cámara estaba montada sobre un escritorio delante de los participantes. A los participantes se les pidió que realizaran signos durante 1 a 3 segundos, obteniendo 500 frases de las cuales 22 fueron descartadas por errores, 384 fueron tomadas para entrenamiento y 94 fueron seleccionadas sólo para fines de prueba. Este experimento resultó en una precisión de 91.9% al clasificar el conjunto de pruebas, y 94.1% al realizar el entrenamiento de un modelo oculto de Markov.

Para el segundo experimento, la cámara estaba montada sobre una gorra que un sujeto llevaba puesta mientras otra persona realizaba señales delante suyo. Una vez más se obtuvieron 500 frases, de las cuales 400 se utilizaron para entrenamiento y 100 para la realización de pruebas. En este experimento se observó que antes de realizar las señas, los participantes permanecían con las manos en posición de reposo. Esta posición de reposo se añadió al conjunto de entrenamiento como "silencio", pero sus resultados de reconocimiento no forman parte del reporte de precisión. Para este experimento, la precisión del conjunto de pruebas fue del 97.8%, mientras que durante el entrenamiento del clasificador se obtuvo un valor del 99.3%.

Para el presente trabajo de disertación, es interesante observar en esa propuesta tanto el dominio de aplicación, como el hecho de que dentro de ella se considera el uso de mecanismos basados en visión. Además, es importante resaltar que la alta precisión alcanzada se debe a la implementación de un clasificador utilizando conjuntos de entrenamiento y de prueba, y que la utilización de nuevos gestos definidos por el usuario no se aborda.

El trabajo de Patsadu et. al (2012), tuvo como objetivo determinar las posiciones de los usuarios (de pie, sentados y acostados). Para este propósito, se instaló una cámara Kinect en un ambiente interior, aproximadamente a un metro sobre el suelo, y se realizaron experimentos colocando a los usuarios a 2 y 3 metros de la cámara. Seis usuarios participaron en el experimento, produciendo 10,800 registros, que se dividen en conjuntos de entrenamiento y pruebas de 7,200 y 3,600 elementos respectivamente. Luego, los registros obtenidos fueron preprocesados aplicando una normalización. Posteriormente, los clasificadores se entrenaron usando el conjunto de datos para entrenamiento, y se obtuvieron resultados de precisión para cada uno de ellos: Redes neuronales de retro propagación 100%, Máquinas de vector de soporte 99.75%, Árboles de decisión 93.19% y Clasificadores Bayesianos 81.94%.

Respecto a los resultados obtenidos, los autores de esa propuesta señalaron que la mayoría de los errores de clasificación se deben a las similitudes entre el hecho de tener un sujeto alto sentado y un sujeto de menor altura de pie. Además de las limitaciones dadas por el hardware que sólo permite rastrear 20 puntos conjuntos, los autores sólo consideraron 3 posiciones para ser rastreadas, pero esto lleva a la suposición de que más posiciones pueden ser rastreadas también.

Junto con el reconocimiento de gestos, otra actividad de interés para la comunidad de investigación es el proporcionar herramientas que permitan a los desarrolladores no expertos la implementación de reconocimiento de gestos en nuevas aplicaciones. Tal es el caso del trabajo de Ibanez et. al (2014), en el que se realizaron dos experimentos.

En una primera sesión del trabajo de investigación presentado en líneas anteriores, se creó un conjunto para entrenamiento con 7 gestos, teniendo en cuenta que los gestos diseñados implicaban movimiento en las articulaciones de la parte superior e inferior del cuerpo. Se pidió a cuatro usuarios que realizaran cada uno de los 7 gestos 20 veces, lo que resultó en 560 muestras. La precisión se calculó utilizando un enfoque de validación cruzada de 10 iteraciones, obteniendo 0.989 y 0.991 para modelos ocultos de Markov y Deformación de Tiempo Dinámico (DTW) respectivamente. Este experimento tuvo como objetivo determinar la precisión de un reconocedor de gestos para que funcionara como base para el segundo experimento.

En la segunda sesión que se llevó a cabo, reclutaron 10 desarrolladores. El primer paso consistió en que ellos debían diseñar aplicaciones controladas por los 7 gestos previamente definidos. El segundo paso consistió en implementar realmente las aplicaciones utilizando una herramienta proporcionada. Finalmente, las aplicaciones generadas fueron probadas en términos de calidad de código y tiempo de desarrollo, obteniendo mejores resultados cuando se utilizó la herramienta propuesta, que cuando se usó un enfoque basado en reglas. En este trabajo, un sensor Kinect fue utilizado como dispositivo de entrada.

Teniendo en cuenta que este trabajo de disertación tiene como objetivo permitir a los usuarios definir su propio vocabulario de gestos corporales para interactuar con aplicaciones soportadas por interacción gestual, es de interés estudiar trabajos realizados siguiendo el enfoque de aprendizaje *one-shot*. En este sentido, la propuesta de Konečný y Hagara (2014), tuvo como objetivo describir dos métodos siguiendo el enfoque de este tipo de aprendizaje para el reconocimiento de gestos. Para la validación de ambos métodos se utilizó el conjunto de datos *ChaLearn* (Guyon et. al, 2012), que comprende más de 50,000 secuencias de gestos. Este conjunto de datos fue creado usando un sensor Kinect y proporciona información de profundidad junto con imágenes RGB como referencia. Los resultados se expresan en términos de la precisión y de la complejidad de los algoritmos propuestos.

De acuerdo con el enfoque de aprendizaje *one-shot*, también se propuso un método de reconocimiento de gestos de varios niveles (Jiang et. al, 2015). La propuesta en este trabajo sigue un proceso de tres capas: clasificador de componentes de movimiento, clasificador de componentes de localización y clasificador de componentes de forma. Por lo tanto, se realizaron diferentes experimentos, utilizando sólo una capa de clasificación a la vez. La precisión se reporta en términos de distancia *Levenshtein* (Yujian y Bo, 2007), y tiempo de reconocimiento de gestos, obteniendo un valor de 37.53% en 0.54 segundos cuando se utiliza sólo la primera capa, 29.32% y 6.03 segundos cuando se utiliza sólo la segunda capa, 39.12% y 6.64 segundos cuando se utiliza la tercera capa, 24.36% y 2.79 segundos cuando se utilizan las dos primeras capas, y finalmente 19.45% y 3.75 segundos cuando se aplican las tres capas.

3.2.2. Interacción gestual utilizando dispositivos vestibles

Si bien, el objetivo del presente trabajo consiste en el estudio de la interacción basada en gestos haciendo uso de dispositivos basados en visión, los resultados, al ser de carácter metodológico, pueden ser extensibles y aplicables a una variedad de dispositivos de entrada tales como los dispositivos vestibles.

Con respecto al uso de acelerómetros, Schlömer et. al (2008), propusieron una biblioteca para el reconocimiento de gestos. Para ello, se reclutaron 6 usuarios, una mujer y cinco hombres, de edades comprendidas entre los 19 y los 32 años, todos ellos con poca experiencia en el uso del mando de Wii y ninguno de ellos con alguna discapacidad. Se pidió a los usuarios que realizaran 5 gestos 15 veces. Con esa información, el sistema fue entrenado y probado, obteniendo una precisión promedio de 90% en ambos procesos utilizando el algoritmo K-mean, un modelo oculto de Markov, y un clasificador Bayesiano.

En (Liu et. al, 2009), introdujeron un sistema para permitir la interacción usando gestos de las manos que son definidos por el usuario por medio de acelerómetros, junto con un prototipo con fines de prueba en el llevar a cabo autenticación de usuarios mediante gestos. Los autores de tal propuesta implementaron su sistema de reconocimiento en una computadora y usaron el control remoto de una consola de videojuegos Wii para permitir la interacción del usuario. Ocho participantes (7 hombres y una mujer) participaron en el experimento. Todos ellos eran estudiantes de pregrado o posgrado, entre 20 y 30 años, diestros. Para cada participante, los gestos se recolectaron durante 7 días en un período de tres semanas. Cada día se pedía a los participantes que realizaran 8 gestos específicos sosteniendo el mando de Wii diez veces, lo que dio como resultado una base de datos de 4,480 gestos (560 gestos por usuario). Se observó que los gestos varían cuando se pide a los usuarios que los realicen en días diferentes.

El experimento realizado en el trabajo descrito en líneas anteriores arrojó los siguientes resultados: sin adaptar las plantillas y tratando a cada usuario por separado, el sistema fue entrenado y probado utilizando un proceso de 10 iteraciones con una precisión del 93.5%, mientras que al adaptar las plantillas y seguir el mismo

procedimiento, se logró una precisión del 98.6%. Después de esto, los autores realizaron otro experimento en el que desarrollaron una aplicación usando el reconocimiento gestual como modalidad de interacción. El objetivo de la aplicación era permitir a los usuarios definir identificadores de gestos, y luego ser autenticados por el sistema. A partir de este experimento, llegaron a la conclusión de que memorizar un identificador en forma de un gesto es menos difícil que memorizar un ID de texto pre compuesto, y que no pudieron encontrar pruebas suficientes que demuestren que realizar un gesto es más difícil que escribir una cadena de texto. La concordancia de plantillas se logró mediante DTW.

3.2.3. Interacción gestual por medio de señales inalámbricas

De la misma manera que se mencionó la posibilidad de llevar a cabo una extensión de la propuesta metodológica del presente trabajo para hacer uso de dispositivos vestibles, el uso de equipos que trabajan con señales inalámbricas es viable, haciendo que el análisis de estudios en que hacen uso de ellos sea relevante.

En el trabajo de Pu et. al (2013) introdujeron un sistema de reconocimiento de gestos que aprovecha las señales inalámbricas llamado *WiSee*. La propuesta se basa en el uso de dispositivos *Wi-Fi* dentro de una casa. Para la detección de gestos, se miden los desplazamientos Doppler (Lee et. al, 1965), mientras que para el reconocimiento se propone un algoritmo de concordancia de patrones. Los autores de este trabajo realizaron experimentos destinados a detectar 9 gestos en diferentes condiciones, como en escenarios de línea de visión, no línea de visión y a través de la pared, en los que la persona se encuentra en una habitación diferente a la del transmisor inalámbrico y el receptor, con una precisión media del 94%.

Este artículo es interesante porque introduce un sistema de reconocimiento de gestos que no se basa en técnicas comunes basadas en la visión, sino en el análisis de las señales *Wi-Fi* y sus desplazamientos cuando los usuarios realizan gestos y alteran las ondas. También es interesante que, aunque los dispositivos y algoritmos que se utilizan son diferentes y siguen un enfoque diferente al del presente trabajo de disertación, la base de la tarea es la misma: primero identificar y luego reconocer. Este

trabajo es limitado, ya que requiere la definición de patrones antes de su ejecución y no permite a los usuarios definir gestos personalizados.

3.3. Vocabularios de gestos definidos por el usuario

Como se mencionó en la introducción, uno de los retos actuales relacionados al reconocimiento gestual consiste en la especificidad de los gestos a culturas y lenguajes. Debido a tal condición, para permitir la interacción siguiendo una modalidad gestual se requiere entrenar a los usuarios o diseñar soluciones ajustadas a un determinado grupo de estos con contextos similares. De acuerdo con la literatura, entrenar a los usuarios puede no ser la mejor alternativa ya que éstos deben aprender grandes cantidades de gestos para interactuar con aplicaciones de uso común (Black et. al, 2017). Además, los principios de usabilidad y UX indican que las interfaces de usuario deben ser auto explicativas, fáciles de utilizar, fáciles de recordar, e intuitivas para los usuarios (Rusu et. al, 2015), considerando que los entrenamientos para el uso de sistemas son innecesarios y no deseables.

Tradicionalmente, los estudios de elicitación de gestos, basados en el procedimiento de estudios de predicción (Wobbrock et. al, 2005), se conducen con el propósito de identificar aquellos gestos correspondientes a tareas específicas de acuerdo con la percepción de un grupo de usuarios. Este enfoque que se ha convertido en un estándar en HCI, busca que después de obtener los vocabularios de gestos tomando en cuenta la información proporcionada por los usuarios, estos vocabularios sean utilizados como comandos dentro de una aplicación para impactar de manera positiva las evaluaciones de usabilidad y de UX.

Sin embargo, a pesar de que las propuestas basadas en estudios de elicitación reportan resultados alentadores en términos de satisfacción general y motivación con respecto al uso de interfaces gestuales (Smith y Gilbert, 2018), es importante observar que la mayoría de los estudios del área no son estadísticamente representativos, haciendo que sigan estando limitados a grupos específicos de usuarios con contextos similares, y llevando a la conclusión de que si una aplicación debe estar disponible para

diferentes poblaciones, múltiples estudios de elicitación deberán ser realizados, encareciendo proyectos y reduciendo su viabilidad.

Adicionalmente, se ha observado durante la aplicación de experimentos de elicitación que aislar a los usuarios de la influencia de productos digitales y tecnológicos es un reto complejo. Es común observar en los resultados de estos estudios la influencia y presencia de metáforas digitales para la ejecución de comandos específicos (Hoff et. al, 2016). En este sentido, investigadores del área han propuesto el uso de distintas estrategias a ser aplicadas en los experimentos de elicitación. Tal es el caso de Morris et. al (2014), quienes proponen la realización de estos experimentos añadiendo actividades físicas y solicitando la producción de múltiples gestos por tarea para reducir la probabilidad de que éstos sean afectados por experiencias previas y por influencias externas.

Otra derivación interesante relacionada al campo de reconocimiento gestual es la creación de gestos configurados para interactuar con aplicaciones específicas. Con respecto a la Realidad Virtual (RV), se definió un vocabulario para la interacción usando gestos de las manos (Piumsomboon et. al, 2013). Para ello, se reclutaron veinte participantes para un experimento, doce hombres y ocho mujeres, de edades comprendidas entre los 18 y los 38 años, con una experiencia mínima en RV y diferentes niveles de experiencia en el uso de dispositivos de pantalla táctil, y otras interfaces de gestos en el aire como los dispositivos Wii y Kinect. Se pidió a los participantes que realizaran gestos para 40 tareas que les fueron presentadas a través de la RV. Después de definir un conjunto completo de gestos, se pidió a los usuarios que los realizaran tres veces más para que pudieran autoevaluar su expresividad y facilidad de uso. La actividad fue monitoreada usando cámaras web y un sensor de profundidad. La prueba resultó en la obtención de 800 gestos que fueron analizados y clasificados para conformar un conjunto o vocabulario. Se informó y discutió el nivel de acuerdo, la facilidad de uso y la similitud de los gestos.

3.4. Conclusiones del capítulo

Dentro de este capítulo se siguió una metodología para llevar a cabo una revisión sistemática de la literatura. La relación de los resultados de esta revisión y los elementos presentados como objetivos y preguntas de investigación se presenta en la Tabla 3. Relación entre elementos del marco epistémico y conceptos revisados dentro del estado del arte.

Tabla 3. Relación entre elementos del marco epistémico y conceptos revisados dentro del estado del arte.

Elemento del marco epistémico	Conceptos del estado del arte	Justificación
O. E. 1. Realizar un mapeo de usuario en 3D por medio de dispositivos basados en visión y API para permitir el monitoreo de movimientos y el reconocimiento gestual.	Modelos de usuario, trabajos relacionados, vocabularios de gestos definidos por el usuario.	Conocer alternativas para el modelado de usuarios permite definir la estructura del modelo y las características a tomar en cuenta. Tener un vocabulario de gestos definidos por usuarios permite además saber cuáles son los gestos comunes y en los que se debe poner especial atención.
O. E. 2. Diseñar e implementar una base de datos de gestos a partir de información obtenida durante estudios preliminares, para permitir la comparación y el análisis.	Modelos de usuario, vocabularios de gestos definidos por el usuario	A partir de la definición de una técnica para modelado de usuario es posible definir la estructura que seguirá la base de datos de gestos. Con el fin de definir los registros que deberán aparecer en la base, es importante conocer vocabularios de gestos definidos por usuarios.
O. E. 3. Proponer un método para etiquetado de gestos utilizando una notación genérica, para soporte a las tareas de reconocimiento y de asociación de significados.	Vocabularios de gestos definidos por el usuario.	Dentro de las primeras etapas de desarrollo de esta propuesta, se puede realizar el etiquetado de un conjunto de gestos definidos por un grupo de usuarios. Posteriormente, esta información podrá ser complementada
O. E. 4. Implementar reconocimiento gestual utilizando técnicas de ML y siguiendo un enfoque de aprendizaje <i>one-shot</i> para permitir interacción natural.	Trabajos relacionados.	En los trabajos relacionados se puede observar mecanismos, estrategias y técnicas aplicadas para llevar a cabo el reconocimiento gestual. Esto permite aprender de las experiencias de otros autores y hacer uso de las opciones que resulten de mayor conveniencia para este trabajo.
P. I. 1. ¿Cómo se puede lograr la interacción gestual de manera que resulte natural para los usuarios?	Trabajos relacionados, vocabularios de gestos definidos por el usuario.	Conocer la experiencia de propuestas anteriores y enriquecerla con la generación de vocabularios de gestos definidos por el usuario permitirá hacer evaluaciones de usabilidad y UX, así como modificar parámetros enfocándose en la búsqueda de la respuesta a esta pregunta de investigación.

<p>P. I. 2. ¿Qué algoritmos, técnicas, métodos y estrategias se deben incluir en una metodología para facilitar la implementación de aplicaciones con interacción basada en gestos?</p>	<p>Modelos de usuario, trabajos relacionados, vocabularios de gestos definidos por el usuario.</p>	<p>Alineando la propuesta con los mecanismos clásicos para llevar a cabo interacción gestual, es necesario contar con elementos que ayuden en tareas de diseño, de implementación y de validación.</p>
<p>P. I. 3. ¿Qué implicaciones resultan de la utilización de gestos definidos por el usuario en una evaluación de usabilidad y experiencia de usuario, realizada sobre aplicaciones basadas en interacción gestual?</p>	<p>Trabajos relacionados, vocabularios de gestos definidos por el usuario.</p>	<p>La respuesta a esta pregunta requiere la realización de procesos de validación. Para llevar a cabo estas evaluaciones, es necesario conocer propuestas realizadas, los resultados que estas reportan y tener un marco de referencia de gestos por detectar.</p>

La identificación de técnicas, modelos, y herramientas permite definir las bases para la proposición de una metodología para la definición de conjuntos gestuales para interacción. En el capítulo siguiente se presenta el trabajo realizado al respecto.

4. METODOLOGÍA

Como se mencionó en capítulos anteriores, el reconocimiento gestual es una actividad que se lleva a cabo típicamente siguiendo métodos de tres pasos: 1) obtención de datos, 2) identificación, y 3) reconocimiento.

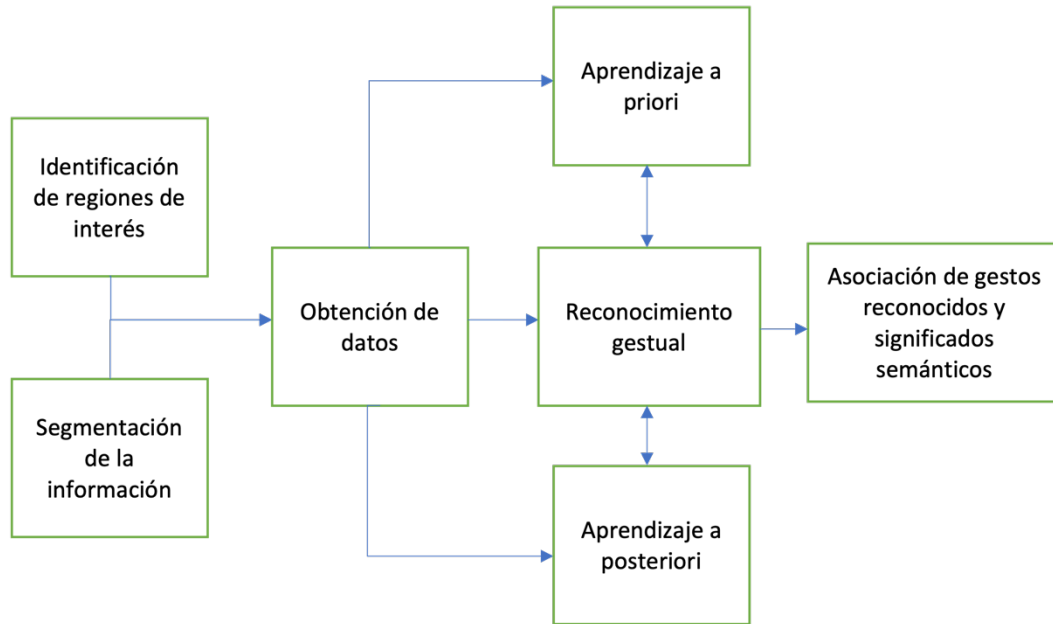


Figura 5. Metodología de siete etapas propuesta para interacción basada en gestos incluyendo el flujo de la información.

El enfoque tradicional funciona correctamente bajo circunstancias controladas en las que solo gestos específicos de determinadas partes del cuerpo se esperan, pero si se desea permitir la definición de nuevos gestos para interacción, trabajo adicional debe ser llevado a cabo. Así, en el presente capítulo se propone una metodología de siete etapas para llevar a cabo el reconocimiento de gestos definidos por el usuario. La metodología propuesta, presentada en la Figura 5, se detalla en las secciones subsecuentes.

4.1. Identificación de Regiones de Interés

El preprocesamiento de información es una buena práctica en los proyectos de clasificación ya que ayuda a la reducción del tiempo computacional y previene esfuerzo innecesario. Específicamente para el reconocimiento gestual, es deseable pre procesar

los datos para identificar dónde y cuándo se presentan los gestos. En términos de espacio, como la metodología propuesta contempla el uso de cámaras de profundidad, la información sobre la localización del usuario se obtiene de manera simple y el fondo y otros objetos son fácilmente removibles gracias al uso de API. Sin embargo, si se utilizan otros dispositivos como cámaras RGB, será necesario aislar los cuerpos de los usuarios de otras fuentes de ruido haciendo uso de algoritmos específicos para tal fin.

Mientras tanto, en términos de tiempo, propuestas como la de Wittenburg et. al (2006) proveen métricas y permiten la detección de posibles gestos, resultando en trozos de información anotados como gestos candidatos.

4.2. Segmentación de la Información

La metodología propuesta incluye un segundo paso de preprocesamiento además del de identificación de regiones de interés previamente mencionado. Esta etapa consiste en tomar los gestos candidatos, etiquetarlos, y preparar la información para ser reconocida o procesada para reentrenamiento en tiempo de ejecución.

En la literatura es posible encontrar trabajos dedicados a la investigación de mecanismos automáticos para etiquetar gestos reportando resultados prometedores (Wang et. al, 2018). En este caso, la idea es utilizar etiquetas genéricas independientes de significados, ya que estos serán asignados en etapas posteriores.

Lo que se obtiene tras la ejecución de esta etapa es la anotación de la información obtenida y su preparación para ser reconocida o incluida en el modelo de clasificación dependiendo de si es que ya existe o no dentro del modelo de clasificación.

4.3. Obtención de Datos

En la literatura, es posible observar diversas alternativas para obtener datos para proyectos de reconocimiento gestual como se mencionó en los capítulos de marco teórico y estado del arte. En la presente propuesta, se considera específicamente el uso de técnicas basadas en visión. Sin embargo, su aplicabilidad a proyectos que hacen uso de dispositivos vestibles y de procesamiento de señales inalámbricas es posible, dado que la metodología propuesta es independiente de plataformas.

La obtención de datos dentro de esta metodología se debe llevar a cabo en dos distintos momentos. Primero, al llevar a cabo el entrenamiento del clasificador como se describe en la sección anterior, y segundo, cuando los usuarios interactúan con el sistema permitiendo reconocer un gesto o actualizando el catálogo de gestos reconocibles.

Especialmente, para efectos explicativos, este trabajo se concentra en el uso de un arreglo de cámaras de profundidad para este propósito, y la meta es generar modelos 3D de usuarios a partir de los datos que estas proporcionan y evitando así la oclusión. Tales modelos facilitan la extracción de características antropométricas para la clasificación y para el proceso de aprendizaje *one-shot*. Para completar esta fase de manera exitosa, es necesario identificar correctamente regiones de interés y segmentar la información obtenida en términos de espacio y de tiempo.

Para llevar a cabo la implementación de esta primera etapa de la metodología, se utilizó una computadora con procesador AMD Ryzen 3[®] 3200G con Gráficos Radeon Vega 8[®] de 3.60 GHz, 16 GB de memoria RAM y con sistema operativo Windows 10 Pro[®], conectada mediante un puerto USB 3.0 a un sensor Microsoft Kinect v2[®]. Así mismo, se realizó de manera previa la configuración del dispositivo, se resolvieron conflictos asociados a compatibilidad y se llevó a cabo la instalación del SDK¹ de Kinect para Windows v2.0 incluyendo los proyectos de demostración y las librerías para trabajar con el sensor. Como resultado de esta configuración inicial, se logró garantizar que el sensor funcionaba correctamente y que era posible obtener de él los datos necesarios como se puede observar en la Figura 6 en donde se muestra una captura de la interfaz del software *Microsoft Kinect Studio*.

Considerando los objetivos del presente trabajo y alineándolos con las características del sensor Microsoft Kinect, se observó la necesidad de contar con herramientas de software que facilitaran la extracción de la información que se requería. En particular, se observó la necesidad de adquirir información relacionada con la posición de las distintas articulaciones en observación además de poder determinar su orientación y el ángulo que existe entre ellas.

¹ <https://www.microsoft.com/en-us/download/details.aspx?id=44561>

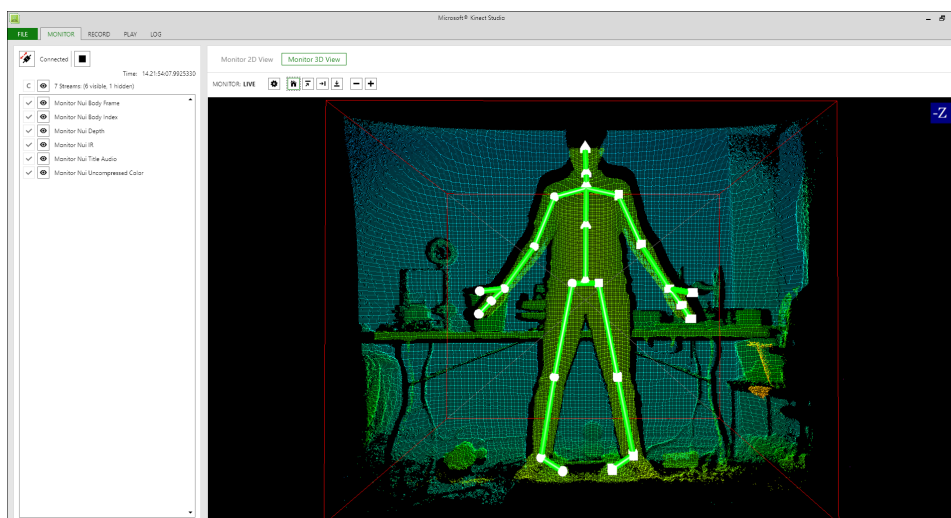


Figura 6. Resultado de configuración del sensor Microsoft Kinect para obtención de información.

Así, tras la consideración y descarte de diversos recursos con base en el análisis de su documentación, disponibilidad actual y ajuste a los equipos con los que se contó, se decidió trabajar con el proyecto Kinect Unity3D Demo de NeurorehabLab² que es un complemento de la plataforma de desarrollo Unity, tomando en cuenta que de acuerdo con su documentación, esta herramienta permite obtener datos acerca de la posición de articulaciones, la orientación de estas y la replicación de movimientos corporales en un avatar humanoide en un ambiente 3D. El proyecto Kinect Unity3d Demo depende para su funcionamiento de un software adicional llamado Reh@Panel que se encarga de obtener la información del sensor y transmitirla por medio de un puerto de red a Unity, donde es recibida e interpretada para su procesamiento y presentación. La utilización del proyecto se ilustra en la Figura 7.

Posteriormente a la exploración de las funcionalidades del recurso Kinect Unity3D Demo, se llevó a cabo la depuración de los scripts incluidos a fin de comprender el flujo de datos. Durante este proceso, fue posible observar que a pesar de que en la documentación se señala que el panel *Reh@Panel* envía información correspondiente a la posición de las articulaciones y sobre las rotaciones de estas, y que dentro de los scripts incluidos existen las estructuras para recibir tales datos, esa información no está

² <https://neurorehabilitation.m-iti.org/tools/en/kinect-demo>

siendo realmente enviada, haciendo que la utilización de esta herramienta no satisfaga los requerimientos de este proyecto de investigación.

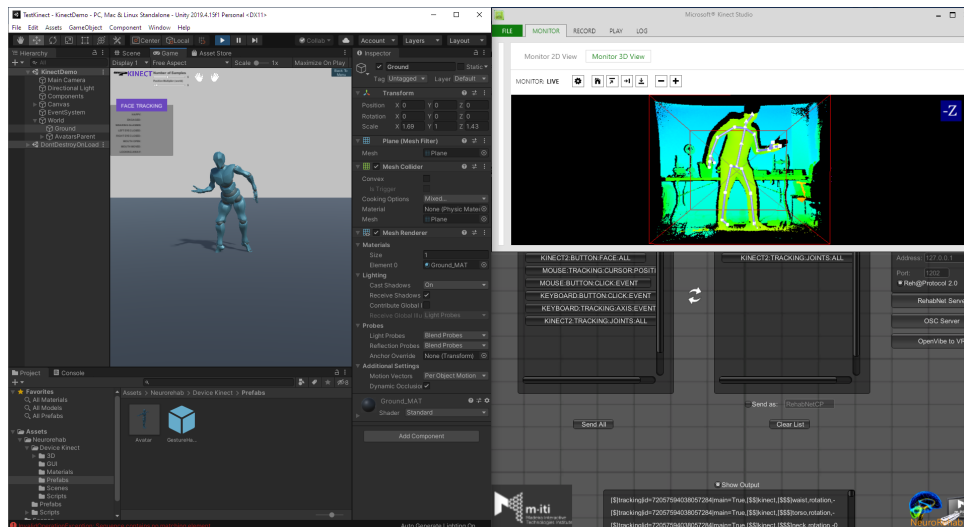


Figura 7. proyecto Kinect Unity3D Demo de NeurorehabLab

Tras la experiencia adquirida, se continuó con la investigación sobre herramientas que permitieran el monitoreo de las articulaciones de interés, haciendo especial énfasis en la necesidad de identificar la posición, la orientación y los ángulos asociados a ellas. Siguiendo estos criterios, se identificó Vitruvius³ como una alternativa que de acuerdo con la documentación proporcionada cumpliría con lo requerido, agregando además la capacidad para crear avatares en ambientes 3D, el soporte para dispositivos diversos y la posibilidad de trabajar también con rasgos faciales.

Con el objetivo de verificar la pertinencia de la utilización de Vitruvius para el presente proyecto, se probaron inicialmente las muestras incluidas en el proyecto, cuya captura se muestra en la Figura 8, y posteriormente se llevó a cabo un proceso de depuración de código sobre las mismas, arrojando resultados favorables.

³ <https://vitruviuskinect.com/>



Figura 8. Capturas de la utilización de las muestras incluidas en Vitrubius.

El análisis del código realizado, demostró que para obtener los datos necesarios, era necesario hacer uso de referencias del SDK de Microsoft Kinect y llamar a funciones específicas de Vitrubius que complementan la información obtenida. Con esto como base, se desarrolló una aplicación en Visual 2019 utilizando C# como lenguaje de programación. Dentro del desarrollo, se mantiene constante monitoreo de los datos que llegan al sensor Microsoft Kinect en cuanto a la posición de cada una de las articulaciones de interés en los ejes X, Y y Z y se calcula también la información de los ángulos en los mismos puntos, tomando como referencia las articulaciones contiguas. Es decir, para la articulación correspondiente al cuello, se toman como referencia las pertenecientes a la cabeza y a los dos hombros, para la del hombro izquierdo, se toman la del cuello y la del codo izquierdo, para el codo izquierdo se toman la del hombro izquierdo y la de la muñeca izquierda y se continúa de la misma forma para cada uno de los puntos de interés restantes. El objetivo que se persigue con el desarrollo de este programa, es crear las bases para satisfacer el objetivo específico OE1 de este trabajo de investigación, al tiempo que se obtiene la información necesaria para soportar las siguientes etapas de la metodología propuesta. La interfaz del programa desarrollado superponiendo una imagen de un esqueleto en el que resaltan las articulaciones de interés se muestra en la Figura 9.

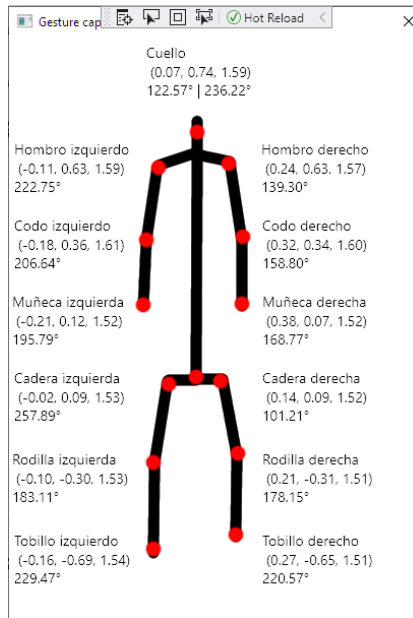


Figura 9. Captura de pantalla de la aplicación para monitoreo de posiciones y ángulos de las articulaciones de interés superponiendo un esqueleto en el que tales puntos se destacan.

Cabe destacar que, hasta este punto, la aplicación desarrollada no cuenta con ninguna guía visual que permita validar la información de manera sencilla y que los datos que se obtienen son relativos a la posición del sensor Microsoft Kinect en un escenario 3D.

4.4. Aprendizaje A-Priori

Los movimientos humanos se encuentran limitados por las articulaciones del cuerpo y por los grados de libertad de estas. El número de posiciones que los usuarios pueden adoptar, si bien es finito, es lo suficientemente grande como para intentar entrenar un clasificador capaz de reconocerlos todos. Además, al considerar que los gestos pueden ser estáticos o dinámicos, y agregando características de los movimientos tales como la dirección y la velocidad, la cantidad de registros con los que el clasificador debiera ser entrenado incrementa drásticamente, haciendo que su implementación no sea viable.

Sin embargo, durante la realización de la revisión a la literatura, se observó como un hecho relevante de los estudios de elicitación que los usuarios declararon que algunos de los gestos efectuados les parecían más naturales y preferibles con respecto

a otros (Modanwal y Sarawadekar, 2018), llevando a pensar que del universo de los gestos que las personas pueden hacer, los usuarios solamente usarían un número limitado para interactuar con sistemas.

Comúnmente, los estudios de elicitación se enfocan en hallar gestos con los que a los usuarios les gusta interactuar con aplicaciones específicas. Esto es correcto para determinados desarrollos personalizados, pero si se piensa en algo de carácter genérico, los estudios de elicitación de gestos pueden brindar además información relevante sobre cómo se comportan los usuarios durante sus interacciones.

Como parte de este trabajo de disertación se propone el uso de este tipo de estudios con el propósito de obtener información sobre gestos sin asociarles significados específicos. Es decir, durante los experimentos, se da especial énfasis a los movimientos que los usuarios realizan cuando se les pide llevar a cabo una tarea determinada.

Así, se desarrolló un protocolo de estudio de elicitación basado en la técnica del Mago de Oz que consta de tres etapas. En la primera etapa, se debe obtener el consentimiento de parte de los usuarios para hacer uso de la información que proporcionan de acuerdo con lineamientos éticos y legales. Igualmente, en esta etapa preparatoria, se debe llevar a cabo la caracterización de los usuarios conforme a la información que para el experimento en curso resulte útil y se deben definir las tareas que se observarán durante el experimento. Para este fin se propone hacer uso de un formato que, además de detallar las instrucciones a llevar a cabo, es de utilidad para registrar los resultados del experimento. La estructura del instrumento propuesto para enlistar las tareas a llevar a cabo y para posteriormente añadir los resultados del estudio durante la segunda etapa, se presenta en la Tabla 4. Dentro de este instrumento, el identificador de la tarea es un dato unívoco para describir la tarea realizada, el nombre de la tarea es una denominación descriptiva en lenguaje natural de la acción a realizar en el contexto específico, la tarea estándar relacionada es el nombre de la tarea de carácter abstracto con la que se guarda una relación, el identificador del gesto es una clave distintiva que permita describir cada uno de los gestos realizados, el tiempo es una medición en segundos realizada desde que se solicita al participante realizar una acción hasta que termina de efectuarla, y la calificación es la valoración que el participante da al gesto realizado por sí mismo.

Tabla 4. Instrumento propuesto para registro de tareas e incorporación de resultados preliminares.

Identificador de tarea	Nombre de la tarea	Tarea estándar relacionada	Identificador del gesto	Tiempo	Calificación
------------------------	--------------------	----------------------------	-------------------------	--------	--------------

Una vez que el experimento ha sido concluido y que se han obtenido los datos en cuanto a gestos realizados, tiempos y calificaciones, como en videos y notas de parte del observador y del controlador del experimento, se procede a solicitar a los participantes que respondan el cuestionario CSUQ con el fin de tener datos que revelen información adicional acerca de su percepción de la usabilidad del sistema simulado o implementado. Finalmente, se debe llevar a cabo el procesamiento de la información reunida.

Esta tercera etapa de procesamiento tras la realización de experimentos de elicitación se deriva a la presentación de resultados. Además de los formatos con tareas y detalles de la interacción, y de los resultados de promedios y desviaciones estándar reportados por el instrumento CSUQ sobre las dimensiones de usabilidad, es necesario presentar el vocabulario de los gestos recolectados.

La presentación del conjunto de gestos, se puede realizar por medio de representaciones textuales informales (e. g. levantamiento del brazo derecho), representaciones textuales de los movimientos de acuerdo con notación como la incluida en el modelo GVUM presentado en el capítulo del estado del arte (e. g. flexión del hombro derecho, extensión del codo derecho), por medio de representaciones gráficas o animaciones, o formalmente mediante el uso de estructuras de datos que permitan la representación de los gestos desde un punto de vista computacional. A partir del conocimiento de esta información, se genera la Tabla 5 en la que se presenta un resumen de las partes del cuerpo del modelo GVUM junto con los movimientos posibles para ellas.

Tras este análisis, como parte de la presente propuesta se creó una estructura de datos en la que se puede insertar información antropométrica que describa los movimientos realizados por el usuario. En el apéndice A se muestra la estructura sin datos, lista para almacenar la información recibida.

Tabla 5. Partes del cuerpo asociadas a sus posibles movimientos de acuerdo con el modelo GVUM.

Parte del cuerpo	Movimientos posibles de acuerdo con el modelo GVUM
Cuello	Flexión, extensión, flexión lateral derecha, rotación lateral derecha, flexión lateral izquierda y rotación lateral izquierda.
Columna	Flexión, extensión, flexión lateral derecho, rotación lateral derecho, flexión lateral izquierda y rotación lateral izquierda.
Hombros	Rotación externa, extensión, flexión, rotación interna, abducción y aducción.
Codos	Extensión, hiperextensión y flexión.
Antebrazos	Supinación y pronación.
Muñecas	Extensión, flexión, desviación radial y desviación ulnar.
Manos	Supinación y pronación.
Dedos	Aducción, extensión, abducción, flexión e hiperextensión.
Cadera	Aducción, abducción, rotación interna, flexión, rotación externa y extensión.
Muslos	Flexión y extensión.
Rodillas	Flexión y extensión.
Tobillos	Dorsiflexión, flexión plantar, eversión e inversión.
Dedos del pie	Extensión y flexión.

Posteriormente, tomando como base el desarrollo de la aplicación para detección de gestos utilizando el sensor Microsoft Kinect reportada en la sección 4.3, se agregó la posibilidad de exportar una captura de tales datos a una estructura que permita su análisis.

Así, considerando los datos que esta versión del desarrollo es capaz de arrojar y la intención de utilizarlos para el entrenamiento de un clasificador, se decidió agregar al software la opción de exportar los datos a una estructura en formato JSON en la que las llaves son etiquetas que describen cada uno de los datos capturados para las articulaciones de interés (posición en eje x, posición en eje y, posición en eje z, ángulo con respecto a articulaciones contiguas), mientras que los valores son los datos que el sensor recolecta. De igual manera, se incorporó a la aplicación la opción de exportar los datos recolectados a un formato ARFF (Formato de Archivo Atributo-Relación) definiendo un total de 53 atributos y tomando en cuenta que las clases serán los distintos gestos que se agreguen al archivo. La aplicación modificada junto con un extracto de los archivos que esta genera se presenta en la Figura 10.

De esta manera, usando la aplicación creada, se abre la posibilidad de capturar posturas corporales y traducirlas a una estructura que permite llevar a cabo el proceso

de entrenamiento de un clasificador de gestos – estáticos – y sentar las bases para llevar a cabo el reconocimiento gestual. Los gestos que resulten comunes durante los experimentos realizados, elegidos por medio de tasas de acuerdo deben ser seleccionados, etiquetados, y utilizados para entrenar un clasificador de manera tradicional.

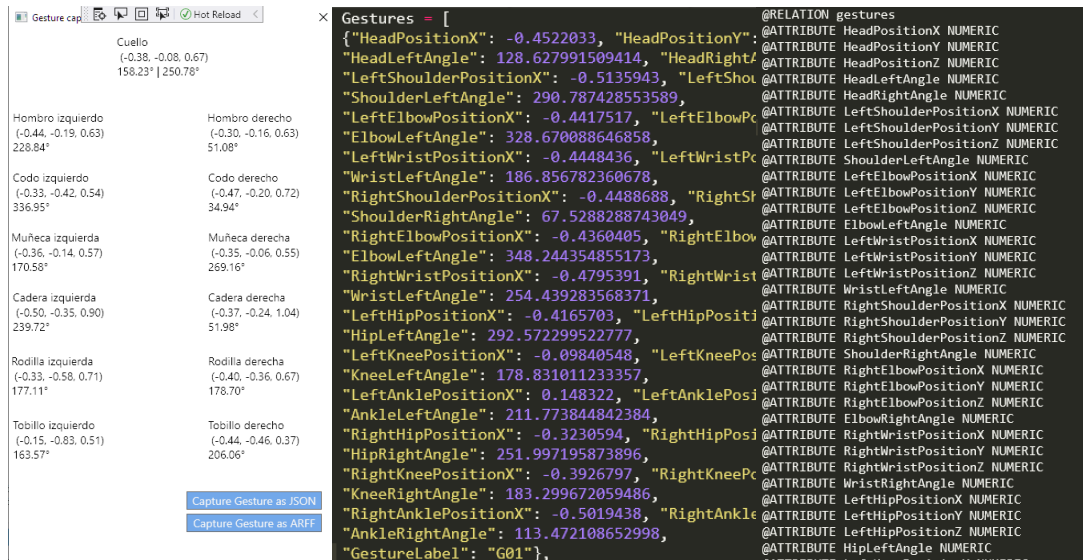


Figura 10. Software para captura de gestos junto con los archivos que se generan en formato JSON y ARFF.

De esta etapa, se debe obtener como producto un clasificador entrenado para el que un gesto determinado debe devolver una etiqueta si el gesto es reconocido, o un valor nulo en caso contrario. Adicionalmente, el análisis permite proponer significados candidatos en términos de tareas abstractas. Para este fin, es importante asegurar el entendimiento de la correspondencia entre las tareas que se solicitaron a los usuarios durante los estudios y tareas canónicas abstractas (Constantine, 2003).

4.5. Aprendizaje A-Posteriori

Comúnmente, las propuestas sobre reconocimiento gestual consideran, o bien tener una etapa de entrenamiento previo a la interacción, o son abiertas a permitir la utilización de técnicas de ML durante la ejecución de tareas. La presente propuesta consta de un enfoque mixto en el que un clasificador entrenado es enriquecido con información que los usuarios proporcionan en tiempo de ejecución.

Así, en adición a la etapa previamente descrita en la que un clasificador se entrena utilizando datos obtenidos en estudios de elicitación, se requiere la definición de una fase de aprendizaje posterior. Típicamente, cuando se agrega información a un clasificador con el fin de permitir la identificación de nuevas clases, se requiere obtener y etiquetar grandes cantidades de datos, entonces, el clasificador se entrena nuevamente y se requiere la realización de pruebas. Sin embargo, para el dominio de la interacción basada en gestos, la adquisición de datos implica que grupos de usuarios repitan el mismo gesto cientos de veces antes de poder utilizar tal movimiento para interactuar con un sistema. Esto es un gran inconveniente que impacta considerablemente la UX y que requiere la inversión de tiempo y esfuerzo extraordinarios.

De esta manera, como parte de la propuesta se considera el uso del enfoque de aprendizaje *one-shot*, de manera que los usuarios únicamente deban proporcionar los gestos en una ocasión, y que el sistema, como se mencionó en la descripción de la etapa de obtención de datos, genere un modelo 3D del usuario, y aplique una serie de variaciones menores en la información para generar tantas muestras como se requiera para reentrenar el clasificador de gestos. Se considera la variación en cuanto a escala, posición y rotación de puntos de articulación, pero características adicionales se pueden agregar para lograr una mayor precisión.

Tomando como base el desarrollo reportado en la sección 4.3 y actualizado para satisfacer los requerimientos descritos en la sección 4.4, se identifica la posibilidad de aplicar las transformaciones a los datos obtenidos desde ese código, generando así archivos en los que, en vez de contar con un único registro para cada gesto detectado, cuenten con un número determinado de muestras para cada gesto, considerando que tales muestras son generadas de forma artificial con parámetros establecidos.

Para una primera prueba realizada, se consideró realizar transformaciones desplazando las posiciones reportadas en el eje X dos metros a la izquierda y dos metros a la derecha con separación de un centímetro resultando en 400 muestras nuevas; en el eje Y, un metro hacia arriba y un metro hacia abajo con separación de un centímetro resultando en 200 muestras adicionales; en el eje Z, dos metros hacia adelante y dos metros hacia atrás con separación de un centímetro, obteniendo así 400 muestras más;

con variaciones aleatorias de -5 y 5° en todos los ángulos asociados a las articulaciones medidas, resultando en 200 muestras agregadas; y finalmente, generando 400 muestras que son resultado de transformaciones aleatorias de hasta un metro en cada dirección en los ejes X, Y y Z, y de hasta 5 grados en cada dirección en los ángulos asociados a cada articulación, haciendo un total de 1601 muestras por cada gesto leído.

Con el conjunto de datos generado a partir de las transformaciones propuestas, se busca resolver no únicamente la ausencia de información suficiente para llevar a cabo el entrenamiento de un clasificador, sino que, además, sea posible resolver problemáticas asociadas a la estatura y complexión de la persona que realiza el gesto a analizar, y sobre las capacidades en cuanto al rango de movimiento de los individuos.

4.6. Reconocimiento Gestual

Una vez que la interacción de los usuarios con los sistemas comienza y tras la obtención de la información segmentada, se necesita probar si los datos correspondientes al gesto realizado son reconocibles por el clasificador entrenado. Si los gestos no son reconocidos, es necesario llevar a cabo el entrenamiento a-posteriori. En caso contrario, se asigna significado semántico a la acción realizada para decidir qué acción llevar a cabo.

El producto que se obtiene de esta etapa es la etiqueta correspondiente al gesto que fue reconocido o agregado al modelo de clasificación.

4.7. Asociación de Gestos Reconocidos y Significados Semánticos

Dependiendo del contexto de uso, los gestos pueden tener distintos significados. El uso de etiquetas identificando cada gesto reconocible permite mapear los gestos de manera que sean significativos para el contexto de uso. Entonces, el contexto de uso debe ser caracterizado igualmente, dado que contiene información que es relevante para la interpretación de un gesto específico.

Esta última etapa de la metodología se enfoca en la asociación de gestos con el significado más probable para ellos y de ejecutar acciones en el sistema que

correspondan con tal información. En esta fase, es necesario desarrollar además un mecanismo tal que los usuarios puedan utilizarlo para declarar si lo que el sistema interpretó no es correcto. De esta manera, las aplicaciones basadas en gestos pueden realmente adaptarse a la forma en que los usuarios se comunican y, por lo tanto, enriquecer la interacción.

4.8. Conclusiones del capítulo

Dentro del presente capítulo se presentó la propuesta de una metodología para llevar a cabo distintas etapas relacionadas con la identificación y el reconocimiento gestual. La metodología consta de 7 etapas que consideran desde la obtención de una muestra hasta la asociación de un gesto con un significado semántico, pasando por procesos de abstracción y segmentación, reconocimiento y haciendo uso del entrenamiento de un clasificador basado en aprendizaje automático. La relación de los conceptos presentados con respecto al marco epistémico se muestra en la Tabla 6.

Tabla 6. Relación entre elementos del marco epistémico y conceptos abordados dentro de la metodología propuesta.

Elemento del marco epistémico	Concepto en la metodología propuesta	Justificación
O. E. 1. Realizar un mapeo de usuario en 3D por medio de dispositivos basados en visión y API para permitir el monitoreo de movimientos y el reconocimiento gestual.	Modelos de usuario, estructuras JSON y ARFF, identificación de movimientos y característica relevantes de articulaciones de interés.	El modelado de usuarios por medio de sus características físicas estableciendo métricas y estructuras para su representación hacen que los gestos realizados puedan ser monitoreados, estudiados e interpretados.
O. E. 2. Diseñar e implementar una base de datos de gestos a partir de información obtenida durante estudios preliminares, para permitir la comparación y el análisis.	Modelos de usuario, estructuras JSON y ARFF, aprendizaje a priori, aprendizaje a posteriori	A partir de la definición de una técnica y estructura para modelado de usuarios es posible crear una base de datos de gestos. Los procesos de entrenamiento a priori y a posteriori tienen el objetivo de proveer la información necesaria para alimentar una base de datos con información predeterminada y enriquecerla con nuevos datos que se recolecten tras la interacción del sistema con el usuario.
O. E. 3. Proponer un método para etiquetado de gestos utilizando una	Etapas de obtención de datos, estructuras JSON y ARFF,	Tras la realización de la etapa de obtención de datos, se asigna a la información recolectada

<p>notación genérica, para soporte a las tareas de reconocimiento y de asociación de significados.</p>	<p>aprendizaje a priori, aprendizaje a posteriori.</p>	<p>un identificador. Esto con el fin de facilitar su adición a las estructuras JSON y ARFF y posteriormente, de permitir la identificación de gestos predeterminados y la agregación de gestos definidos por el usuario. La notación utilizada en esta propuesta permite una rápida identificación manual de gestos al incluir dentro de ella letras que indican el origen del gesto y números consecutivos.</p>
<p>O. E. 4. Implementar reconocimiento gestual utilizando técnicas de ML y siguiendo un enfoque de aprendizaje <i>one-shot</i> para permitir interacción natural.</p>	<p>Obtención de datos, estructuras JSON y ARFF, entrenamiento a priori, entrenamiento a posteriori.</p>	<p>Las estructuras propuestas para almacenar la información recolectada de los usuarios, hace posible la aplicación de transformaciones que generen múltiples muestras sintéticas que representen ligeras variaciones sobre gestos realizados y hacen posible el entrenamiento a posteriori. Esto aunado a la suposición de la preexistencia de una base de datos, resultado de un proceso de entrenamiento a priori, hacen posible el reconocimiento gestual tras la obtención de una sola muestra por parte de los usuarios.</p>
<p>P. I. 1. ¿Cómo se puede lograr la interacción gestual de manera que resulte natural para los usuarios?</p>	<p>Estructuras JSON y ARFF, entrenamiento a posteriori</p>	<p>Permitir a los usuarios la definición de gestos con los que desean interactuar con un sistema, les permite utilizar comandos que les parecen naturales para llevar a cabo sus propósitos. Las estructuras JSON y ARFF junto con el proceso descrito como entrenamiento a posteriori, permite el reconocimiento de gestos propuestos por el usuario.</p>
<p>P. I. 2. ¿Qué algoritmos, técnicas, métodos y estrategias se deben incluir en una metodología para facilitar la implementación de aplicaciones con interacción basada en gestos?</p>	<p>Identificación de regiones de interés, segmentación de la información, obtención de datos, entrenamiento a priori.</p>	<p>La revisión a la literatura mostró que los métodos clásicos de procesamiento digital de imágenes alcanzan porcentajes altos en cuanto a la segmentación de recursos en términos de espacio y tiempo, siendo esto de vital importancia para la detección y reconocimiento gestual. Asimismo, los procesos y algoritmos actuales de aprendizaje automático tales como los clasificadores Bayesianos, árboles de decisión y máquinas de vector de soporte, atienden de manera correcta y eficiente la tarea de reconocimiento.</p>
<p>P. I. 3. ¿Qué implicaciones resultan de la utilización de gestos definidos por el usuario en una evaluación de usabilidad y experiencia de usuario, realizada</p>	<p>Entrenamiento a priori, entrenamiento a posteriori, reconocimiento gestual, asociación de gestos</p>	<p>Contar con la definición de mecanismos que permitan llevar a cabo en entrenamiento de un clasificador de gestos y su enriquecimiento con otros gestos que el usuario propone, así como con procedimientos para asociar tales gestos</p>

sobre aplicaciones basadas en interacción gestual?	reconocidos y significados semánticos.	con tareas específicas para un contexto, permite la definición de procedimientos para estudiar la apreciación de los usuarios ante la interacción gestual con gestos definidos por ellos mismos y tener un marco de comparación con respecto a mecanismos de interacción convencionales.
--	--	--

El desarrollo y la ejecución de las distintas etapas que forman parte de la propuesta metodológica resultaron en la realización de estudios preliminares que reportaron información relevante. Estos resultados se muestran en el siguiente capítulo.

5. RESULTADOS

Dentro de la presente sección se presentan los resultados alcanzados durante la realización del trabajo de investigación. En las siguientes secciones es posible observar los hallazgos realizados en los trabajos preliminares y así mismo, se describen las salidas obtenidas tras la ejecución de los programas implementados.

5.1. Navegación en Ambientes Virtuales utilizando Comandos Definidos por el Usuario

La multimodalidad es un paradigma poderoso utilizado para elevar el nivel de realismo y para facilitar la interacción en ambientes de Realidad Virtual (RV). El desarrollo de sistemas que permiten entradas multimodales y en específico la creación de mundos virtuales y su navegación, involucra múltiples retos incluyendo la búsqueda de técnicas que soporten la interacción en tres dimensiones, y para satisfacer requerimientos funcionales alcanzando niveles deseables de eficiencia y accesibilidad entre otras dimensiones de usabilidad y de UX (Boudoin et. al, 2008).

Al seguir un enfoque de desarrollo basado en modelos (MDA) para el mejor entendimiento de los requerimientos de interfaces de usuario, es importante comenzar por la identificación de tareas cotidianas y significativas a ser realizadas dentro de una aplicación. Tales tareas se definen como secuencias de acciones coordinadas o lógicas, y pueden ser válidas para diferentes aplicaciones; luego, las tareas identificadas pueden ser descompuestas en acciones elementales similares a las tareas propuestas por Constantine (2003).

La utilidad de las aplicaciones 3D es, de hecho, proporcional a la habilidad de sus usuarios para interactuar con la información que contienen, y que puede ser colocada en múltiples ubicaciones dentro del ambiente virtual y alcanzable desde distintos ángulos, haciendo que sea necesario que los usuarios sean capaces de navegar eficientemente en el entorno visualizando el escenario y objetos de estudio desde diferentes perspectivas. Así, este estudio se enfocó en la navegación dentro de ambientes de RV, siendo esta la tarea que más comúnmente se efectúa en tales

aplicaciones (Bowman et. al, 1997), mientras que otras tareas como la selección y manipulación se dejaron fuera del alcance de este.

A pesar de que la navegación es una actividad primaria de los ambientes virtuales, y de que tareas equivalentes en contextos reales se logran a través del movimiento y posicionamiento del cuerpo, se pudo observar durante la revisión al estado del arte sobre trabajos relacionados a la interacción gestual, que pocos autores se enfocaron en la investigación acerca de la navegación en mundos virtuales. El diseño efectivo de la navegación en ambientes de RV, permite a los usuarios tener la sensación de movimiento fácil e intuitivo. Un buen diseño típicamente resulta en dos productos: investigación para comprender principios cognitivos dentro de la navegación, y diseño personalizado para crear técnicas de navegación específicas para la aplicación de interés.

Alineado con las preguntas de investigación presentadas en el capítulo de introducción, se diseñó un experimento en el se pide a usuarios llevar a cabo tareas específicas en ambientes de RV utilizando en un momento interacción gestual y en otro, interfaces gráficas convencionales, siguiendo la técnica de Mago de Oz, descrito en el capítulo del marco teórico. El desempeño de los usuarios interactuando con ambas alternativas se contrastó para obtener conclusiones con respecto a la facilidad de uso y preferencia de usuarios utilizando instrumentos validados.

Para el dominio de estudio en específico, las tareas mínimas necesarias para llevar a cabo la interacción con el mundo virtual, identificadas a través de un proceso de abstracción considerando interfaces de usuario de un sistema existente son: acercar (*zoom in*), alejar (*zoom out*), ir a un lugar (*fly to place*), inclinar vista (*tilt view*), y guardar ubicación (*save location*).

La configuración del experimento, como se muestra en la Figura 11, incluyó dos computadoras (PC1 y PC2), la primera para controlar un sensor Kinect y una cámara web con el fin de obtener la información de los movimientos de los usuarios, y permitiendo la documentación del experimento; mientras que la segunda se encontraba conectada a un proyector para brindar la retroalimentación y para simular la interacción de los usuarios con el sistema.

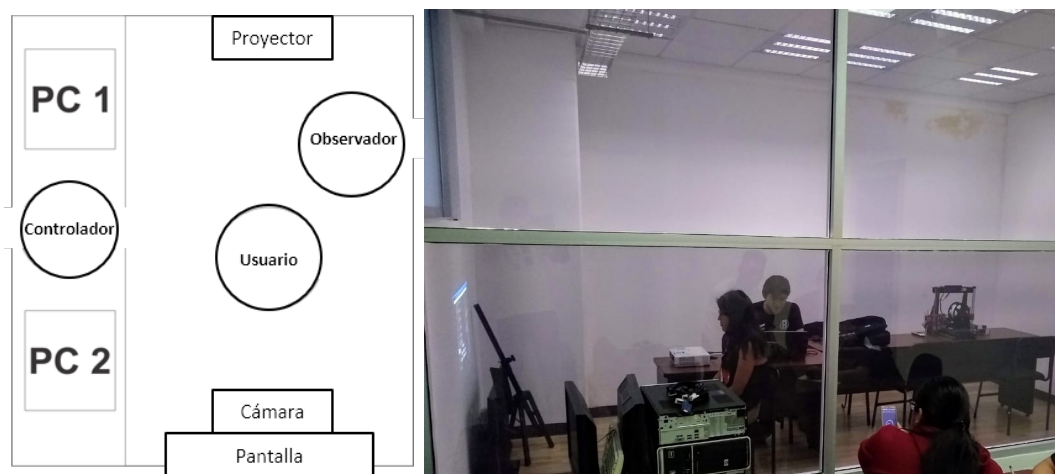


Figura 11. Configuración del experimento siguiendo el enfoque de Mago de Oz y la disposición en un escenario real.

Los participantes que formaron parte del experimento eran responsables de proveer gestos ante una aplicación simulada, y de interactuar con tal aplicación siguiendo un conjunto de instrucciones. Se designó un observador del experimento, quien se ubicó junto al usuario con el fin de brindarle soporte en caso de dudas. El controlador del experimento administró los dispositivos para documentar el experimento y para explicar cómo realizó las actividades cada usuario.

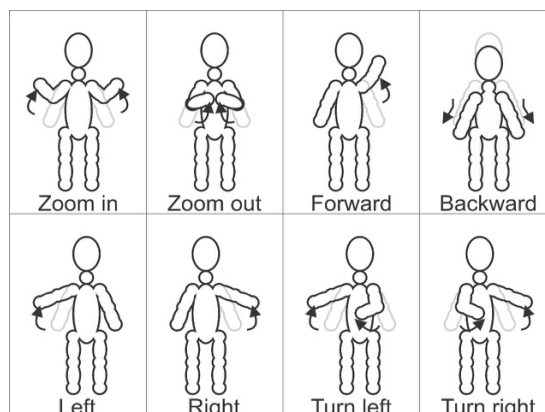


Figura 12. Vocabulario de gestos propuesto para la interacción con ambientes de RV.

En el experimento, se involucraron estudiantes de la carrera de computación de una universidad pública. Ellos fueron instruidos en el uso de la herramienta y recibieron una breve descripción y ejemplos de cómo utilizar la plataforma, así como sobre la forma de aplicar la técnica de Mago de Oz. Posteriormente, se les solicitó encontrar voluntarios en la universidad.

Así, se consiguió la participación de ocho estudiantes de pregrado, siete hombres y una mujer de entre 18 y 23 años. A cada uno de ellos se le solicitó indicar los gestos que llevarían a cabo para realizar una serie de tareas, repitiendo este proceso hasta que los usuarios estaban satisfechos con los gestos proporcionados. De esta forma, un conjunto de 64 gestos fue obtenido. Tales gestos fueron caracterizados, comparados y agrupados resultando en un conjunto de gestos definidos por el usuario para navegación en ambientes virtuales como se presenta en la Figura 12.

Antes del experimento se solicitó a los participantes brindar consentimiento para el tratamiento de la información de su participación usando el formato que se muestra en el Apéndice E y se les aplicaron entrevistas con el fin de obtener información relevante sobre ellos. Los ocho participantes manifestaron tener preferencia por el uso de la aplicación basada en gestos sobre la aplicación basada en la interfaz gráfica debido, principalmente, a que les parecía más divertida. Incluso, uno de los usuarios mencionó que le gustaría ejercitarse usando sistemas basados en gestos.

Mientras el experimento se llevaba a cabo, se comparó el desempeño de los usuarios y sus preferencias al interactuar con la aplicación para navegación. En el experimento con la interfaz gráfica, todos los usuarios fueron capaces de completar las actividades y los tiempos de respuesta nunca fueron de más de un minuto. Por el contrario, al usar gestos las actividades llegaron a tomar más de cuatro minutos. Claramente el desempeño fue peor en el segundo escenario. Un resumen de la retroalimentación y de los problemas a los que cada uno de los usuarios se enfrentaron se presenta en la siguiente lista:

1. Dificultades para encontrar el lugar objetivo. El usuario constantemente se perdió en el ambiente virtual.
2. No se detectaron problemas.
3. No se detectaron problemas.
4. Falló al completar la tarea. Muy confundido, la mayoría de los problemas se relacionaron con encontrar caminos, múltiples errores al tratar de recordar el vocabulario definido. El usuario no sabía a donde ir. Finalmente, la frustración causó que el usuario abandonara el experimento.

5. Múltiples dudas, frustración y desesperación mientras se intentaba navegar el ambiente virtual. Constantemente se perdía, en un punto determinado trató de hablarle al Sistema para recibir asistencia.
6. Satisfecho con la tecnología. El usuario quería utilizar la aplicación por más tiempo.
7. Frustración menor. El usuario solo realizaba gestos con una mano para manipular la aplicación.
8. Algunos problemas al encontrar caminos, memoria, dudas sobre el lenguaje de gestos definido.

La tarea de acercar (*zoom in*) fue realizada más comúnmente moviendo ambos brazos arriba y lejos del tórax. De manera recíproca, alejar (*zoom out*) se realizó acercando ambos brazos al tórax. Moverse hacia adelante (*forward*) se representó como el levantamiento del brazo izquierdo mientras que moverse hacia atrás (*backward*) fue caracterizado como encorvarse enviando ambos brazos adelante. Para desplazarse a la izquierda (*left*) y a la derecha (*right*), los usuarios eligieron respectivamente mover el brazo izquierdo hacia la izquierda y el brazo derecho a la derecha. Finalmente, para girar a la izquierda (*turn left*) o a la derecha (*turn right*), los usuarios movieron ambos brazos hacia un lado o al otro.

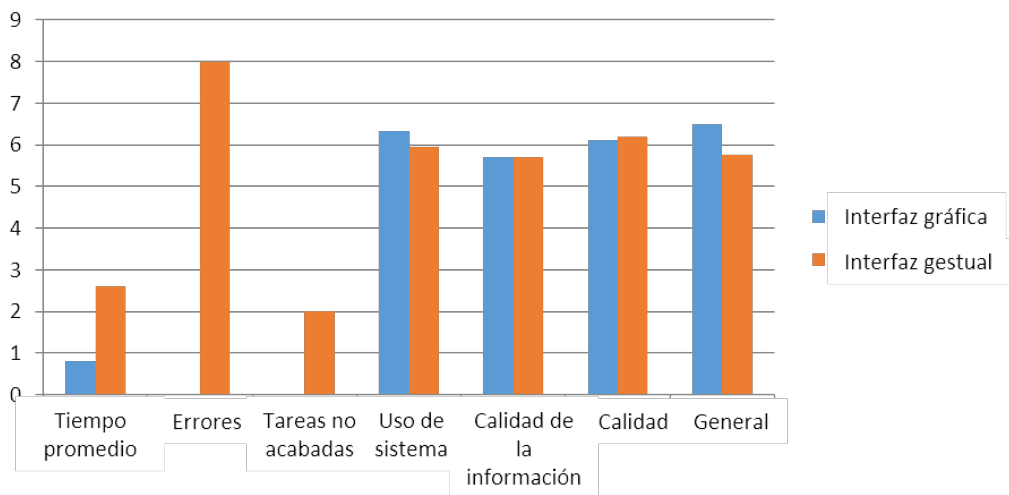


Figura 13. Evaluación general del sistema de navegación en ambientes virtuales en términos de tiempo promedio en segundos para cada tarea realizada, cantidad de errores cometidos en

total, tareas no finalizadas en total, y la media de las dimensiones de usabilidad de acuerdo con el instrumento CSUQ.

Con el fin de contrastar los resultados obtenidos se efectuaron dos experimentos, uno usando teclado y ratón sobre una interfaz gráfica, y otro usando la simulación de interacción gestual como se describió en párrafos anteriores. Durante ambos experimentos, se recolectaron datos del desempeño a través de métricas y se complementaron con información cualitativa obtenida por medio de instrumentos de evaluación y de la misma observación.

Después de cada una de las pruebas, se solicitó a los participantes responder el cuestionario CSUQ para obtener datos acerca de su apreciación con respecto al sistema usado en términos de usabilidad. Esos resultados junto con la evaluación de los tiempos promedio de ejecución, de los errores cometidos y de tareas no terminadas se reportan en la Figura 13.

5.2. Identificación de un vocabulario gestual para interacción con múltiples aplicaciones

De acuerdo con lo expresado en los objetivos y considerando la información reportada en la Tabla 5, esta sección se dedica a la presentación de un estudio realizado con la finalidad de proponer un vocabulario de gestos definidos por usuarios por medio de la observación de los movimientos de sus articulaciones.

Para llevar a cabo este estudio, se contó con 70 participantes, con poca experiencia en el uso de aplicaciones basadas en interacción gestual. Los participantes fueron separados en 5 grupos con 14 personas cada uno. El primer grupo, dedicado al estudio de gestos del cuello, consistió en 9 mujeres y 5 hombres entre 19 y 79 años. El segundo grupo que fue creado para observación de gestos realizados con los brazos fue conformado por 6 mujeres y 8 hombres con entre 17 y 67 años.

El tercer grupo, que trabajó con gestos de las manos, contó con un número igual de hombres y mujeres cuyas edades se encontraban entre los 20 y los 63 años. El cuarto grupo, para gestos de las piernas, contó con 7 hombres y 7 mujeres con edades entre 13 y 53 años. Finalmente, el quinto grupo que fue creado para la observación de gestos de los pies tuvo 7 mujeres y 7 hombres entre 12 y 52 años. La distribución de los

usuarios en estos grupos se llevó a cabo de acuerdo con la puntuación que estos obtuvieron por su desempeño en pruebas de creatividad, esto debido a la influencia documentada de este factor dentro de estudios similares (Motaghi, 2015).

Tras la selección de los participantes y su agrupamiento, se llevaron a cabo experimentos para identificar vocabularios de gestos relacionados a las partes del cuerpo mencionadas previamente. Si bien este tipo de experimentos se realiza comúnmente implementando aplicaciones, se decidió utilizar prototipos y aplicar un enfoque de Mago de Oz descrito en el marco teórico. Así, se llevó a cabo la configuración del experimento siguiendo la disposición mostrada en la Figura 11.

Durante el experimento se les pidió a los usuarios proporcionar gestos (usando solo partes del cuerpo limitadas) para llevar a cabo 14 tareas dentro de distintas aplicaciones y contextos. Las 70 tareas que fueron definidas se presentan en la Tabla 7.

Tabla 7. Tareas para observar durante el experimento para identificación de un vocabulario gestual para interacción con múltiples aplicaciones.

Parte del cuerpo (aplicación)	Tareas
Cabeza (Instagram)	Encontrar perfiles (N1), Agregar imagen (N2), Seleccionar imagen (N3), Tomar fotografía (N4), Grabar video (N5), Ver notificaciones (N6), Ver perfiles (N7), Dar 'me gusta' (N8), Crear historia (N9), Agregar Emoji (N10), Agregar hashtag (N11), Aceptar solicitud de seguidor (N12), Etiquetar amigo (N13), Editar perfil (N14).
Brazos (Domótica)	Abrir puerta (A1), Cerrar puerta (A2), Encender luces (A3), Apagar luces (A4), Encender aire acondicionado (A5), Apagar aire acondicionado (A6), Encender TV (A7) Apagar TV (A8), Subir volumen de TV (A9), Bajar volumen de TV (A10), Cambiar canal de TV (siguiente) (A11), Cambiar canal de TV (anterior) (A12), Encender bomba de agua (A13), Apagar bomba de agua (A14).
Manos (Facebook)	Ver publicaciones (H1), Ver notificaciones (H2), Ver mensajes (H3), Crear grupo (H4), Dar 'me gusta' (H5), Dar 'me encanta' (H6), Escribir mensaje (H7), Agregar amigo (H8), Cargar fotografía (H9), Etiquetar amigo (H10), Bloquear usuario (H11), Crear evento (H12), Iniciar video llamada (H13), Iniciar llamada (H14).
Piernas (Juego Crash Bandicoot)	Mover hacia abajo en el menú principal (L1), Mover hacia arriba en el menú principal (L2), Seleccionar opción en el

	menú principal (L3), Mover adelante en 3D (L4), Mover a la izquierda en 3D (L5), Golpear una caja en 3D (L6), Saltar obstáculo en 3D (L7), Destruir caja en 3D (L8), Mover adelante en 2 ½D (L9), Saltar obstáculo en 2 ½D (L10), Golpear caja en 2 ½D (L11), Destruir caja en 2 ½D (L12), Agacharse en 3D (L13), Salir de minijuego en 2 ½D (L14).
Pies (Teléfono Android, reproductor de medios, navegador Web)	Responder llamada telefónica (F1), Terminar llamada telefónica (F2), Ir a pantalla de inicio (F3), Iniciar búsqueda por voz (F4), Play/Pausa (F5), Salir de reproductor (F6), Ir a canción anterior (F7), Modo aleatorio (F8), Subir volumen (F9), Bajar volumen (F10), Mover cursor a la izquierda (F11), Mover cursor a la derecha (F12), Acercarse (zoom in) (F13), Alejarse (zoom out) (F14).

Cuando los usuarios llevaron a cabo gestos para cada una de las tareas, se recolectaron datos sobre el tiempo de respuesta y una calificación proporcionada por ellos mismos con respecto a que tan natural piensan que es el gesto brindado para la tarea especificada. Asimismo, se grabó el gesto correspondiente en un clip de video. En total, se obtuvieron 980 registros que fueron procesados a través de un etiquetado y comparación manual, para posteriormente calcular una tasa de acuerdo.

Los gestos con las mayores tasas de coincidencia para cada tarea fueron caracterizados de acuerdo con la información de la Tabla 5 que corresponde a movimientos posibles para las distintas articulaciones del cuerpo de manera consistente con el modelo GVUM. Además, esa descripción se complementó con una descripción de los gestos en lenguaje natural y con una representación gráfica etiquetada de manera genérica para su rápida identificación. Esta información en conjunto, mostrada en apéndice B, e ilustrada en la Figura 14, formó un vocabulario de gestos corporales para interacción con múltiples aplicaciones.

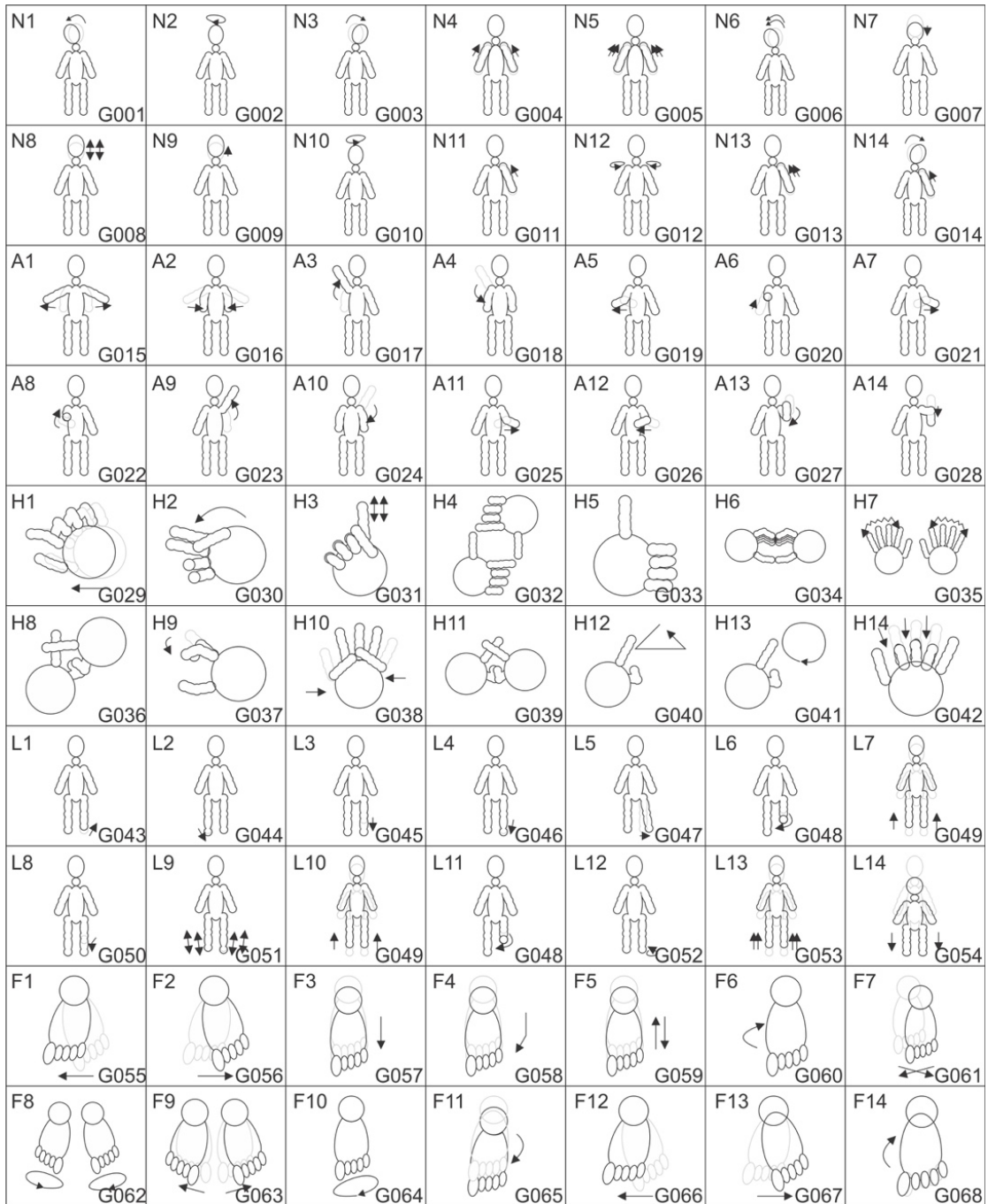


Figura 14. Representación gráfica del vocabulario de gestos definidos por usuarios incluyendo el identificador de las tareas asociadas a ellos y sus etiquetas genéricas.

5.3. Un vocabulario gestual asociado a tareas abstractas

En la sección anterior, se presentó un vocabulario gestual para interacción con múltiples aplicaciones y se indicó la posibilidad de asociar varias de las tareas que

formaron parte del estudio con tareas de orden abstracto. Hacer esto, permitió la creación de un vocabulario con gestos genéricos que puede ser implementado en aplicaciones de distintos dominios a través de la concretización de tareas como lo referido en la Tabla 1.

Así, se decidió llevar a cabo un nuevo estudio en el que se simuló la interacción de los participantes con la plataforma de transmisión Netflix. En la Tabla 8 se muestran las tareas definidas como parte del protocolo del experimento junto con su tarea abstracta asociada.

Tabla 8. Tareas seleccionadas para el experimento de identificación de un vocabulario gestual asociado a tareas abstractas.

Id	Tarea	Tarea abstracta asociada
1	Iniciar sesión	Seleccionar, Iniciar / ir a
2	Buscar película o serie	Seleccionar, accionar / operar
3	Ajustar volumen	Modificar
4	Reproducir video	Iniciar / ir a
5	Cambiar capítulo	Llevar a cabo (y regresar), seleccionar
6	Activar subtítulos	Seleccionar, cambiar estado
7	Cambiar idioma	Llevar a cabo (y regresar), seleccionar
8	Ir adelante 10 segundos	Accionar / operar
9	Ir atrás 10 segundos	Accionar / operar
10	Agregar a lista de reproducción	Crear
11	Cerrar sesión	Detener / terminar / completar
12	Pausar video	Detener / terminar / completar
13	Desactivar subtítulos	Seleccionar, cambiar estado
14	Saltar introducción	Detener / terminar / completar
15	Silenciar	Cambiar estado
16	Salir del video	Detener / terminar / completar
17	Mover hacia arriba	Mover
18	Mover hacia abajo	Mover
19	Mover a la izquierda	Mover
20	Mover a la derecha	Mover
21	Elegir contenido	Seleccionar, Iniciar / ir a
22	Cambiar perfil	Seleccionar, Modificar
23	Ver en pantalla completa	Ver

Con respecto a la configuración del experimento, se empleó la misma disposición que en la subsección anterior y descrita dentro de la sección 4.1, siguiendo el enfoque de Mago de Oz. Durante el experimento, se contó con la participación de 90 usuarios divididos en 6 grupos, cuidando una distribución uniforme de las personas en cuanto a

su sexo, mano hábil y edad. Estos 6 grupos fueron creados para captar datos de expresiones faciales, gestos de cuello y hombros, gestos de brazos, gestos realizados con las manos, gestos de las piernas, y gestos efectuados con los pies.

A los participantes se les pidió llevar a cabo gestos para las 23 tareas seleccionadas y presentadas en la Tabla 8 usando únicamente partes específicas del cuerpo. Además de la grabación de los gestos realizados, se registraron los datos correspondientes a tiempos de respuesta y la calificación de parte del usuario para la naturalidad de cada gesto. En total, se obtuvieron 2070 muestras, las cuales fueron procesadas manualmente, llevando a cabo un etiquetado y comparación, para finalmente asignarles una puntuación de nivel de acuerdo.

Los gestos que obtuvieron las puntuaciones de nivel de coincidencia más altas para cada una de las tareas se consideraron como parte de la propuesta de vocabulario final, se representaron de manera gráfica y se describieron tanto de manera textual como a través de la notación del GVUM con la intención de poder almacenarlos en una estructura formal que abra la posibilidad de entrenar un clasificador con los datos organizados.

De esta manera, durante el estudio de elicitación se obtuvieron 124 gestos. Tales gestos y combinaciones de ellos fueron utilizados por los participantes para expresar comandos dentro de la plataforma de transmisión. Como se mencionó en líneas anteriores, con respecto a expresiones faciales, fue posible identificar 16 gestos distintos.

Para movimientos realizados con cuello y hombros, se observaron 21 gestos. Así mismo, se encontraron 22 gestos realizados con los brazos, 23 gestos de las manos, 27 gestos efectuados con las piernas, 15 más que fueron caracterizados para interacción con los pies.

Todos los gestos fueron etiquetados usando una notación genérica, asociados a las tareas para las que fueron realizados y descritos tanto de forma textual como en términos del GVUM. La información obtenida se presenta en el apéndice C. Posteriormente, estos gestos fueron representados de forma gráfica para su fácil identificación como se muestra en la Figura 15.

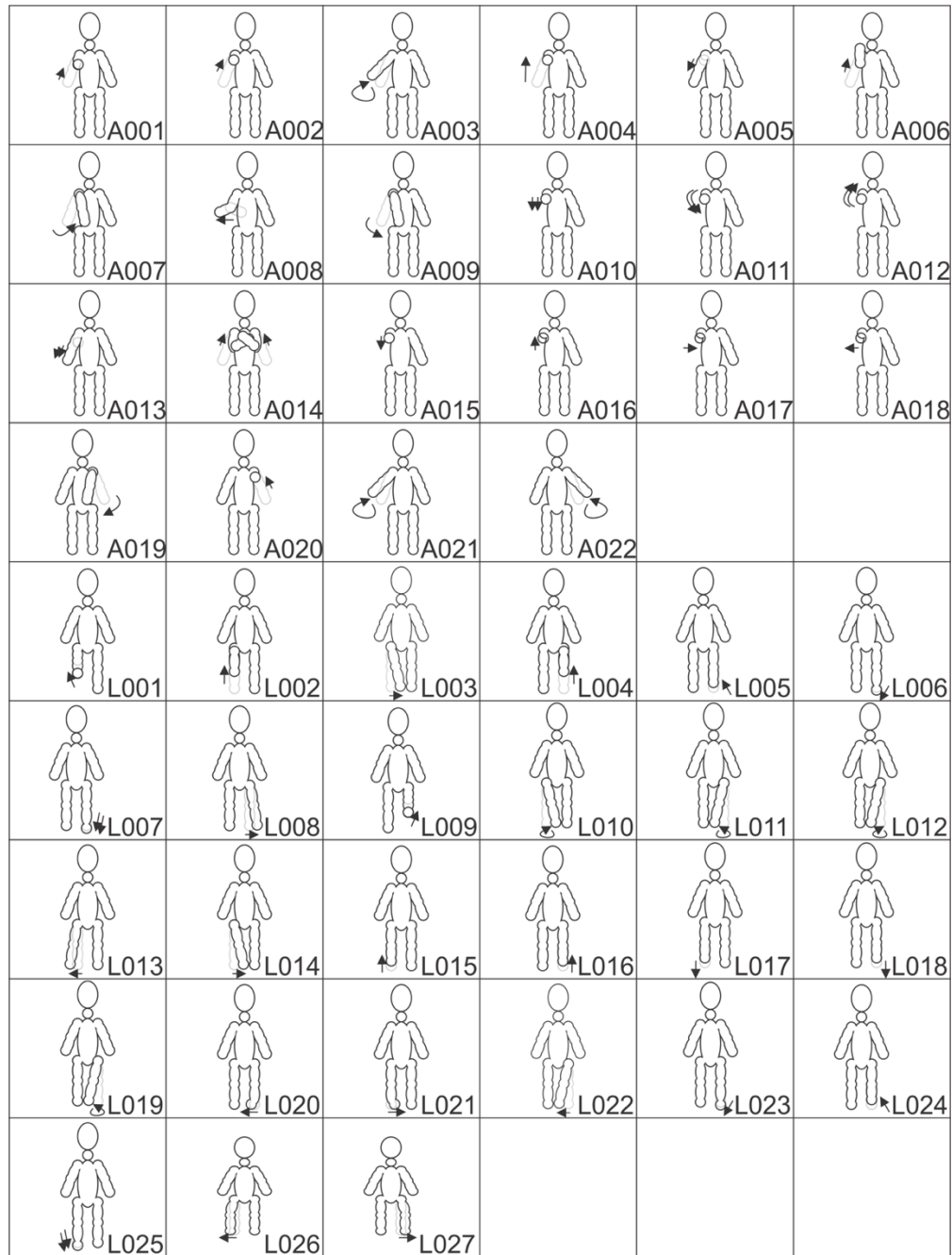


Figura 15. Extracto de la representación gráfica del vocabulario de gestos obtenido durante el estudio asociado a tareas abstractas.

Finalmente, se utilizó la estructura del apéndice A para representar de manera formal los distintos gestos obtenidos. Es relevante mencionar que esto fue realizado de manera manual basándose en la observación de puntos de articulación y asignando valores aproximados. Un ejemplo de esta traducción se puede observar en la Figura 16.

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<capability>
  <motor>
    <face> ☐
    </face>
    <neck> ☐
    </neck>
    <leftUpperLimb> ☐
    </leftUpperLimb>
    <rightUpperLimb>
      <shoulder>
        <flexion>0</flexion>
        <extension>90</extension>
        <adduction>0</adduction>
        <abduction>0</abduction>
        <externalRotation>0</externalRotation>
        <internalRotation>0</internalRotation>
      </shoulder>
      <elbow>
        <flexion>90</flexion>
        <extension>0</extension>
        <hyperExtension>0</hyperExtension>
      </elbow>
      <forearm> ☐
      </forearm>
      <wrist> ☐
      </wrist>
      <hand> ☐
      </hand>
    </rightUpperLimb>
    <leftLowerLimb> ☐
    </leftLowerLimb>
    <rightLowerLimb> ☐
    </rightLowerLimb>
  </motor>
</capability>
```




Figura 16. Representación formal del gesto A006 junto con su representación gráfica.

5.4. Entrenamiento preliminar de un clasificador gestual siguiendo un enfoque de aprendizaje *one-shot*

Dentro de las secciones 4.3 y 4.4 se describió el desarrollo de una herramienta que permite captar la información que se recolecta por medio de un sensor Microsoft Kinect para, posteriormente, organizarla en estructuras que permitan su transmisión y procesamiento habilitando la posibilidad de analizar los gestos identificados y de realizar tareas de entrenamiento de un clasificador gestual. Igualmente, en la sección 4.5 se especifican las transformaciones propuestas para generar múltiples muestras a partir de un solo registro obtenido.

Dentro de las pruebas realizadas sobre el software desarrollado y a fin de validar su funcionamiento para los propósitos de este trabajo, se llevó a cabo un experimento en el que se solicitó a un participante adoptar distintas posturas cuyos datos fueron capturados utilizando la herramienta implementada y procesados de acuerdo con las transformaciones propuestas, obteniendo así un conjunto de 8 posturas correspondientes con el vocabulario obtenido en la sección 5.1 para navegación en

ambientes virtuales que se muestran en la Figura 17 y que fueron expresadas en 12,808 registros con 53 atributos cada uno.

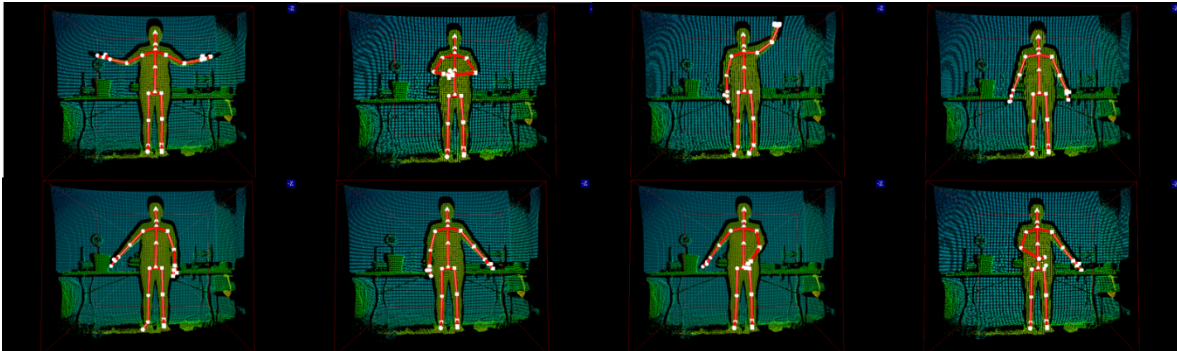


Figura 17. Posturas adoptadas para la realización del experimento sobre entrenamiento de un clasificador gestual para tareas de navegación en ambientes virtuales siguiendo un enfoque de aprendizaje one-shot.

Tras la captura de los datos, se generaron los archivos JSON y ARFF correspondientes incluidos parcialmente en el apéndice D, y que fueron utilizados para llevar a cabo el entrenamiento de un clasificador Bayesiano obteniendo como resultados generales una precisión del 1.00, exhaustividad de 1.00 y valor-F de 1.00, mostrando que el 100% de las instancias fueron correctamente clasificadas durante el proceso de entrenamiento como se puede notar en la matriz de confusión que se muestra en la Figura 18.

```

=== Confusion Matrix ===
      a   b   c   d   e   f   g   h  <-- classified as
1601  0   0   0   0   0   0   0 | a = G01
  0 1601  0   0   0   0   0   0 | b = G02
  0  0 1601  0   0   0   0   0 | c = G03
  0  0  0 1601  0   0   0   0 | d = G04
  0  0  0  0 1601  0   0   0 | e = G05
  0  0  0  0  0 1601  0   0 | f = G06
  0  0  0  0  0  0 1601  0 | g = G07
  0  0  0  0  0  0  0 1601 | h = G08
  
```

Figura 18. Matriz de confusión arrojada durante el proceso de entrenamiento siguiendo el enfoque de aprendizaje one-shot

En una segunda etapa, se repitió el entrenamiento del clasificador Bayesiano sustrayendo del conjunto de entrenamiento ocho registros, uno de correspondiente a cada gesto y que habían sido generados de forma aleatoria como parte del proceso de

aprendizaje a posteriori reportado en la sección 4.5, obteniendo durante este proceso valores de precisión, exhaustividad y valor-F de 1.0 nuevamente.

A continuación, el modelo generado se evaluó sobre los datos extraídos como conjunto de prueba alcanzando un 100% en la precisión de las predicciones realizadas como se puede observar en los resultados reportados en la Figura 19.

Cabe destacar que previamente a la realización de las pruebas finales sobre estos datos, se llevó a cabo un proceso de selección para asegurar que dentro del conjunto de prueba existieran muestras de cada uno de los gestos de interés y que estos se encontraran en desorden para descartar que fuera esto un factor en el reconocimiento de un gesto.

=== Predictions on user test set ===					=== Confusion Matrix ===										
inst#	actual	predicted	error	prediction	a	b	c	d	e	f	g	h	<-- classified as		
1	2:G02	2:G02	1		1	0	0	0	0	0	0	0	0	0	a = G01
2	3:G03	3:G03	1		0	1	0	0	0	0	0	0	0	0	b = G02
3	4:G04	4:G04	1		0	0	1	0	0	0	0	0	0	0	c = G03
4	6:G06	6:G06	1		0	0	0	1	0	0	0	0	0	0	d = G04
5	7:G07	7:G07	1		0	0	0	0	1	0	0	0	0	0	e = G05
6	8:G08	8:G08	1		0	0	0	0	0	1	0	0	0	0	f = G06
7	5:G05	5:G05	1		0	0	0	0	0	0	1	0	0	0	g = G07
8	1:G01	1:G01	1		0	0	0	0	0	0	0	1	0	0	h = G08

Figura 19. Resultados de la realización de pruebas sobre el conjunto de gestos generado de forma aleatoria.

5.5. Conclusiones del capítulo

Dentro de este capítulo se llevó a cabo la presentación de los resultados obtenidos durante la realización de estudios en los que se pusieron a prueba distintas fases de la metodología propuesta y cuya relación con los elementos del marco epistémico se muestra en la Tabla 9.

Tabla 9. Relación entre elementos del marco epistémico y conceptos abordados dentro de la presentación de resultados.

Elemento del marco epistémico	Concepto en los resultados presentados	Justificación
O. E. 1. Realizar un mapeo de usuario en 3D por medio de dispositivos basados en visión y API para permitir el	Entrenamiento preliminar de un clasificador gestual	Dentro de los resultados se muestra el procedimiento seguido para llevar a cabo la generación de muestras sintéticas y el

monitoreo de movimientos y el reconocimiento gestual.		entrenamiento de un clasificador gestual a partir de ellas.
O. E. 2. Diseñar e implementar una base de datos de gestos a partir de información obtenida durante estudios preliminares, para permitir la comparación y el análisis.	Definición de vocabularios de gestos asociados a tareas abstractas y específicas a un contexto.	Se realizaron experimentos para identificar vocabularios gestuales para la interacción con aplicaciones de distintos dominios, de manera que se pudiera contar con una base de datos de gestos proporcionados por usuarios permitiendo además la abstracción de sus características.
O. E. 3. Proponer un método para etiquetado de gestos utilizando una notación genérica, para soporte a las tareas de reconocimiento y de asociación de significados.	Se llevó a cabo el etiquetado de gestos asociados a tareas tanto específicas como abstractas y se propuso una notación para su etiquetado.	Dentro de los resultados del entrenamiento de un clasificador gestual se utilizaron etiquetas que enumeraban los gestos sin tener una relación directa con las características de estos. Esto permite realizar una generalización de los procesos y por lo tanto asociar gestos a tareas distintas en contextos diferentes.
O. E. 4. Implementar reconocimiento gestual utilizando técnicas de ML y siguiendo un enfoque de aprendizaje <i>one-shot</i> para permitir interacción natural.	Se realizó y reportó el proceso de entrenamiento de un clasificador gestual tomando como base muestras sintéticas creadas a partir de características determinadas.	Los vocabularios gestuales obtenidos durante estudios preliminares permitieron identificar una serie de características sobre las que se llevaron a cabo ligeras variaciones en términos de rotación y traslación para llevar a cabo el proceso de entrenamiento de un clasificador gestual.
P. I. 1. ¿Cómo se puede lograr la interacción gestual de manera que resulte natural para los usuarios?	Vocabularios de gestos definidos por el usuario, creación de muestras sintéticas a partir de muestras generadas por usuarios, identificación de la relación entre tareas abstractas y su concretización a tareas propias de un contexto.	Los estudios de usabilidad reportados muestran que la utilización de vocabularios de gestos definidos por el usuario para interacción con aplicaciones favorece la satisfacción general y abona a la fácil memorización y al aprendizaje del uso de sistemas. La definición de tareas abstractas y su concretización a contextos específicos permite la reutilización de vocabularios gestuales.
P. I. 2. ¿Qué algoritmos, técnicas, métodos y estrategias se deben incluir en una metodología para facilitar la implementación de aplicaciones con interacción basada en gestos?	Creación de vocabularios de gestos, aprendizaje automático, identificación de tareas de interacción y relación con tareas abstractas,	La eficacia de técnicas de ML radica en la calidad de los datos que se proporcionan para el proceso de entrenamiento. Al contar con vocabularios de gestos para tareas de interacción, es posible realizar una abstracción de sus características y utilizarlas para crear conjuntos de muestras sintéticas que son la base para el proceso de entrenamiento de un clasificador gestual reportando niveles óptimos de precisión y desempeño.
P. I. 3. ¿Qué implicaciones resultan de la utilización de gestos definidos por el	Los experimentos sobre la utilización de vocabularios de	Es de especial interés observar que la apreciación de los usuarios con respecto a la

usuario en una evaluación de usabilidad y experiencia de usuario, realizada sobre aplicaciones basadas en interacción gestual?	gestos definidos por el usuario para interacción en aplicaciones de diversos dominios mostraron que dimensiones de la usabilidad como la facilidad para memorizar, la facilidad para el aprendizaje y la satisfacción general son positivamente afectadas con respecto a sus contrapartes usando interacción gráfica o basada en comandos gestuales predefinidos.	utilización de interfaces gestuales es positiva, a pesar de resultar compleja en determinadas tareas y de inducir errores en la interacción.
--	---	--

Consideraciones adicionales sobre los resultados obtenidos y descritos en el presente capítulo se pueden encontrar en la Discusión de los resultados.

6. DISCUSIÓN DE LOS RESULTADOS

En el presente capítulo se profundiza en el análisis de los resultados reportados en el capítulo 5 brindando un panorama más específico de la aplicabilidad, replicación y limitantes de los mismos.

6.1. Navegación en Ambientes Virtuales utilizando Comandos Definidos por el Usuario

Durante el experimento de navegación en ambientes virtuales utilizando comandos definidos por el usuario, se observó que tareas tales como guardar lugar, mover el marcador y agregar marcador tuvieron que ser descartadas dado que existía mucha confusión en los usuarios sobre cómo podían mover su cuerpo para comunicar esas funcionalidades al sistema y que, para ellas, los participantes preferían emitir comandos vocales.

Fue posible observar de tales resultados que, a pesar de llevar a cabo interacción gestual, a los usuarios les tomó más tiempo realizar las tareas, de que cometieron más errores durante la interacción y de que no fueron capaces de completar algunas tareas, la apreciación con respecto al sistema no se vio significativamente afectada. Específicamente, al preguntar a los usuarios si usarían el sistema basado en gestos posteriormente, en la escala de 7, se obtuvo un valor promedio de 6.6 puntos, lo que sugiere un alto interés de navegar ambientes virtuales con gestos corporales. La retroalimentación fue realmente alentadora sin importar el bajo desempeño que tuvieron.

Comparando la propuesta de este estudio con otros proyectos relacionados en la literatura, el requerimiento de interactuar con aplicaciones usando gestos corporales ha sido abordado por múltiples autores, lo que, es más, existen diversos trabajos dedicados al desarrollo de técnicas de navegación a través de mapas, sin embargo, son altamente dependientes de interfaces de hardware. Estas técnicas son eficientes para tareas de navegación aisladas, pero si las acciones globales de los ambientes de RV (incluyendo tareas de navegación, selección y manipulación) son consideradas, donde

la interoperabilidad de varios dispositivos o modalidades puede ser necesaria, además de la necesidad de cambiarlos de acuerdo con las características de la tarea, pueden no funcionar apropiadamente.

Además, hay otras alternativas con las que los resultados obtenidos en este estudio preliminar deben ser contrastados, por ejemplo, existe una propuesta (Boudoin et. al, 2008) para un modelo de interacción 3D que facilita y considera observaciones previas, el cual se enfoca en mantener la misma lógica de uso a través de dispositivos, haciendo la interacción intuitiva para los usuarios. Otra técnica de interacción en la literatura, llamada plataforma de interacción visual (Aliakseyeu et. al, 2002), que, por medio de hardware, incluyendo un proyector, brinda la idea de que se tiene una computadora con un espacio de trabajo grande sobre una superficie plana que contiene una tableta digitalizadora. La mayoría del trabajo en el área incluye el análisis de NUI (Maike et. al, 2014), interfaces 3D (González-Calleros et. al, 2010), guías para la creación de ambientes de RV (Kaur, 1997), entre otros temas, pero no se encontraron trabajos acerca de la definición de un lenguaje de gestos corporales para interactuar con mapas.

Es importante resaltar el trabajo en (Tan et. al, 2001), en donde se propone una taxonomía para categorizar técnicas de navegación existentes junto con sus restricciones, para crear estructuras relacionadas a tareas de desplazamiento, permitiendo e inspirando la creación de nuevas técnicas. Finalmente, a pesar de que diversos trabajos han estudiado lenguajes definidos por el usuario para interacción con teléfonos móviles (Kray et. al, 2010) y (Ruiz et. al, 2011), tabletas (Kurdyukova et. al, 2012), realidad aumentada (Piumsomboon et. al, 2013), *tokens* activos (Valdes et. al, 2014), en una forma similar a la que se hizo en este trabajo, existe la necesidad de formalizar los lenguajes a través de la adopción de una estrategia más formal como la que se propone en (Vatavu y Wobbrock, 2016).

A pesar de que el número de participantes en tal experimento (ocho) no es estadísticamente significativo, cabe destacar que 68% de los trabajos relacionados reportaron contar con menos usuarios (Maike et. al, 2014).

6.2. Identificación de un vocabulario gestual para interacción con múltiples aplicaciones

Dentro de este estudio fue posible observar que aunque se pedía a los usuarios que utilizaran sólo los movimientos de una parte determinada del cuerpo y que sería correcto utilizar los mismos gestos para diferentes acciones, en la mayoría de los casos se sentían más cómodos utilizando gestos diferentes para cada comando, y en particular en el caso de los grupos de movimiento de cabeza y pies, comentaban que el uso de una sola parte del cuerpo no era suficiente para expresar lo que pensaban, extendiendo los movimientos a los hombros y los tobillos respectivamente. Por lo tanto, el vocabulario de gestos finales contiene también este tipo de movimientos.

Si bien el vocabulario de gestos fue creado utilizando tareas específicas en aplicaciones existentes, es posible ver que pueden ser aplicables a otros contextos similares ya que la mayoría de ellos describen actividades abstractas como encender, apagar, abrir, cerrar, moverse hacia arriba y hacia abajo.

Así mismo, un hallazgo interesante es que las tasas de coincidencia de los gestos realizados con la cabeza (media de 0,22, desviación estándar 0,08), las piernas (media de 0,32, desviación estándar 0,09) y los pies (media de 0,36, desviación estándar 0,11) no fueron significativamente inferiores en comparación con los gestos de las manos (media de 0,31 y desviación estándar 0,26) y los brazos (media de 0,31 y desviación estándar 0,12), como se esperaba, teniendo en cuenta que la interacción se lleva a cabo comúnmente con las extremidades superiores.

De acuerdo con la observación realizada durante el experimento, esto fue causado por dos razones, la primera de ellas relacionada con el modelo mental que tienen los usuarios y la segunda relacionada con que si bien no es común interactuar con las aplicaciones utilizando sólo movimientos de cabeza, piernas y pies, los grados de libertad que estas partes del cuerpo permiten reducen el número de gestos que los usuarios se sienten cómodos haciendo, reduciendo así las opciones que tienen y aumentando las probabilidades de tener movimientos similares. Además, la desviación estándar calculada sobre las puntuaciones de nivel de acuerdo mostró el valor más alto para los gestos de las manos. Con respecto a esto, fue posible observar que los gestos

que los usuarios repetían constantemente para una tarea específica, como el G033 para la tarea H5, ya se están utilizando en la interacción real entre personas con un significado similar o pueden tener una representación similar en plataformas ampliamente utilizadas y adoptadas como normas de hecho. Esto es coherente con las observaciones que figuran en (Hoff et. al, 2016) sobre la influencia de la tecnología y la cultura en la interacción gestual.

Por otra parte, otras tareas que no tienen una representación gestual estándar como H2, H3 y H4, obtuvieron puntuaciones de aceptación más bajas. Observamos que la principal causa de esto es que las manos tienen más grados de libertad y por lo tanto es posible tener más gestos de ellas. La expresividad permitida por las extremidades superiores afectó significativamente el acuerdo, considerando que sólo hubo 4 tareas para las cuales los usuarios proporcionaron gestos completamente diferentes y que tales tareas pertenecen a los grupos de brazos (A5 y A6) y manos (H3 y H12). Además, se puede notar que los gestos G048 y G049 obtuvieron la mayor puntuación de tasa de aceptación (5/14) cada uno para dos tareas, lo que dio lugar a la reducción del vocabulario gestual de 70 gestos previstos a 68.

Como parte del protocolo, también se registraron los tiempos de respuesta y se pidió a los usuarios que evaluaran los gestos que proporcionaron. De estos datos, fue interesante observar que los usuarios más jóvenes respondieron, en promedio, más rápido que los de más edad, y que cuanto más rápido los usuarios proporcionaban gestos para un comando específico, mejor evaluaban el movimiento, como se puede ver en la Figura 20 para los gestos de la cabeza. Durante el experimento, se percibió que ambas mediciones se vieron afectadas significativamente por la confianza de los usuarios con respecto a la modalidad de interacción.

Así mismo, dentro de este estudio se esperaba observar relaciones existentes entre la puntuación de acuerdo para un gesto y su tiempo de respuesta medio asociado, así como entre las puntuaciones de aceptación y las calificaciones de los usuarios para cada gesto, por medio de este análisis fue posible encontrar que existe una relación que indica que, a mejor tiempo de respuesta, mejor es la calificación de los usuarios y viceversa.

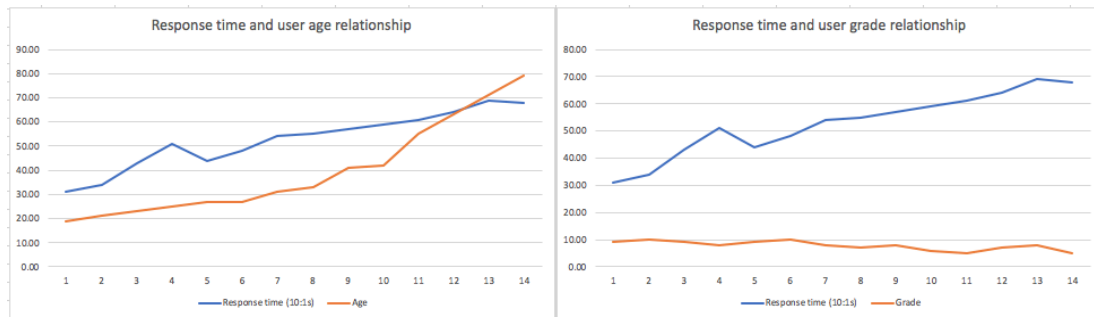


Figura 20. Relación entre tiempos de respuesta y edades de los usuarios, y entre tiempos de respuesta y calificaciones de los usuarios para gestos realizados con la cabeza.

De esta manera, se propuso un vocabulario de gestos definidos por el usuario para la interacción basada en gestos de cuerpo entero. Del experimento se obtuvieron 980 registros, que se procesaron para su etiquetado manual, comparación y cálculo de la tasa de concordancia, a fin de obtener finalmente un conjunto de 68 gestos que contiene una representación gráfica de cada gesto, etiquetas para su identificación, descripciones textuales y caracterizaciones de cada gesto que fueron basadas en los rasgos presentados en un modelo de capacidad. Se observó que las características de los usuarios, como la experiencia, la confianza, la edad y el modelo mental, así como otras facetas adicionales como la comodidad, la expresividad y el grado de libertad de las partes del cuerpo, afectan a los gestos, el tiempo de respuesta y la evaluación de ellos. Se consideraron solo superficialmente las repercusiones de factores adicionales como la creatividad y la influencia del uso de la tecnología, pero será necesario impulsar experimentos adicionales para determinar con precisión las dimensiones de esos aspectos que afectan a la interacción gestual.

Como las tareas del experimento corresponden a tareas abstractas, el vocabulario propuesto permite la implementación de aplicaciones basadas en gestos de cuerpo entero para propósitos generales, teniendo en cuenta que pueden ser necesarios gestos adicionales para la extensión y adaptación.

6.3. Un vocabulario gestual asociado a tareas abstractas

El vocabulario presentado en la sección 5.3 puede ser utilizado para la ingeniería de interfaces basadas en gestos. La asociación entre tareas abstractas y los comandos

específicos del dominio sobre el que se trabajó, así como la identificación de los gestos que los usuarios asocian a tales comandos, permite llevar a cabo una generalización del vocabulario obtenido. Sin embargo, es importante mencionar que, dada la naturaleza de las interfaces gestuales, se debe tomar en cuenta aspectos culturales, de lenguaje y cognitivos con fines de adaptación (Chandler y Schwarz, 2009).

La población que participó en el experimento estaba compuesta por 47 mujeres y 43 hombres, todos ellos mexicanos de entre 16 y 57 años (Promedio 38.19 años, desviación estándar 11.88).

Durante el experimento, se pudo confirmar que hay muchos mapeos de conceptos a gestos y viceversa, como se indica en la literatura (Mitra y Acharya, 2007). Para observar esto, nótese que, para las 23 tareas estudiadas, los usuarios proporcionaron un total de 2,070 gestos (345 gestos por cada uno de los seis grupos). De esos gestos, 124 fueron seleccionados de acuerdo con sus puntuaciones de nivel de acuerdo y tras la desambiguación y división de gestos complejos en otros atómicos. Además, como puede verse en el apéndice C, algunos de los gestos se utilizan para realizar varias tareas mientras que otros están específicamente relacionados con un comando.

Justo después de proporcionar cada gesto, se pidió a los usuarios que dieran una puntuación entre 1 y 10, de acuerdo con qué tan fácil, memorable y descriptivo pensaban que era el gesto dado para representar el comando requerido. Se esperaba ver relaciones existentes entre los tiempos de respuesta y las calificaciones de los usuarios para cada gesto y las puntuaciones de acuerdo o evaluación final, lo cual ocurrió de manera similar a lo reportado en la sección anterior.

Como parte del protocolo de experimentación seguido, se pidió a los usuarios que respondieran al CSUQ para obtener datos relativos a su apreciación del sistema en cuanto a su utilidad. Los resultados de la aplicación del cuestionario mostraron que la interacción utilizando las extremidades inferiores era menos satisfactoria para los usuarios, ya que la evaluaron con un promedio de 5.13 de 7 puntos (desviación estándar 1.17), mientras que después de simular la interacción utilizando los miembros superiores, los participantes dieron una puntuación media de 5.8 de 7 (desviación estándar 0.79). Esta diferencia puede deberse a la falta de expresividad de los miembros inferiores, pero también al hecho de que los participantes están

acostumbrados a utilizar sus manos para interactuar con las aplicaciones en lugar de sus piernas o pies.

Junto con la información cuantitativa que se reunió, es decir, el tiempo de respuesta, la puntuación de los gestos y la puntuación del acuerdo fue posible obtener conocimientos cualitativos sobre el experimento. Durante el experimento, los participantes pertenecientes a los grupos de cuello y hombros, de piernas y de pies se quejaron de lo difícil que les resultaba utilizar sólo esas partes de sus cuerpos para interactuar con una aplicación, sobre todo para realizar tareas complejas. Fue necesario introducirlos en un escenario en el que la interacción con partes limitadas del cuerpo no sólo era deseada sino crítica, por ejemplo, debido a una discapacidad, con fines de rehabilitación, o debido a las características del contexto.

En cuanto a la propuesta resultante en cuanto al vocabulario de gestos y su presentación tanto formal como no formal, es posible utilizarla para fines de capacitación de usuarios, para el diseño de aplicaciones basadas en gestos y para iniciar el entrenamiento de un clasificador automático de gestos. No obstante, es necesario hacer algunas aclaraciones. La estructura del archivo XML se creó de conformidad con el modelo de capacidad de usuario de Kaklanis et. al (2013), pero su ampliación con características suplementarias de interés es posible mediante la adición de etiquetas. De manera similar, la simplificación de la estructura también es posible mediante la eliminación de etiquetas. Esta reducción puede aplicarse no sólo a bloques completos de información (los investigadores pueden eliminar las expresiones faciales del archivo si no están interesados en seguir ese tipo de interacción), sino también a características específicas que no serán observadas.

Otra consideración interesante es que la estructura propuesta para la representación formal de los gestos incluye información opuesta, por ejemplo, flexión-extensión, inflación-contracción, abducción-aducción. Estas etiquetas opuestas pueden no estar presentes en una instanciación final, ya que tienen información repetitiva y su uso depende del enfoque que se siga.

La representación de los gestos estáticos puede lograrse mediante el uso de un solo archivo XML, mientras que para los gestos dinámicos es necesario agrupar una serie de archivos según los diferentes momentos en que se producen cambios significativos.

Deben añadirse etiquetas adicionales a la estructura básica para apoyar la inclusión del orden o la aparición y el tiempo o la rapidez.

A fin de lograr el objetivo de considerar el enfoque de aprendizaje *one-shot* como se menciona en los objetivos de este trabajo, la notación formal permite que un gesto sea comprensible para las computadoras, sin embargo, hasta ahora no sería posible completar un proceso de entrenamiento de clasificadores con una sola muestra. Por lo tanto, se identificó la necesidad de definir un mecanismo para utilizar los datos reunidos sobre un gesto para entrenar un reconocedor automático. Para ello, se propone la aplicación de ligeras variaciones sobre la información representada formalmente a fin de simular la obtención de múltiples muestras para llevar a cabo el proceso de entrenamiento.

Es importante tener en cuenta que la definición de esas transformaciones implica el conocimiento de las restricciones de movimiento, por ejemplo, los grados máximos de rotación de una articulación, la distancia máxima entre las articulaciones; así como de los niveles de tolerancia en los que se puede variar un gesto sin generar ambigüedad.

6.4. Entrenamiento preliminar de un clasificador gestual siguiendo un enfoque de aprendizaje *one-shot*

El sensor Microsoft Kinect junto con el software desarrollado y el marco de trabajo Vitrubius, permitieron llevar a cabo la detección de posturas y su caracterización para entrenar un clasificador Bayesiano cuyos detalles se describen en la sección 5.4. Sin embargo, a pesar de los resultados alentadores obtenidos para el conjunto de entrenamiento y de pruebas, es necesario precisar que no se está considerando toda la información que el sensor arroja sobre las articulaciones haciendo que los resultados de este proceso puedan no contar con la robustez que se requiere para alcanzar los objetivos de este trabajo de investigación.

Puntualmente, el sensor Kinect puede proporcionar datos sobre la orientación de las articulaciones en forma de cuaterniones que expresan rotaciones con respecto a los ejes X, Y y Z (*roll*, *pitch* y *yaw*) y que el marco de trabajo Vitrubius permite identificar.

Además, es posible encontrar dos grandes limitaciones en la implementación realizada, la primera es que los gestos son vistos siempre desde una perspectiva frontal, es decir, que para que funcione correctamente, se parte del supuesto de que el gesto siempre será realizado con la misma orientación, mientras que la segunda limitación está en que el clasificador únicamente funciona en este punto con gestos estáticos.

Adicionalmente, dentro de este estudio preliminar no se siguió un proceso formal para la determinación de la cantidad de muestras sintéticas a generar ni de las características que estas deberían tener. Por estos motivos y aunado a los resultados obtenidos, puede inferirse que el clasificador desarrollado podría haber sufrido un sobre entrenamiento.

Muestras / Atributos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
400	9.59	33.84	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
500	11.26	31.40	71.15	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
600	11.41	11.50	72.85	73.16	73.02	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
700	10.33	11.39	20.92	17.51	17.73	19.60	18.14	18.69	21.87	20.60	20.26	21.66	20.62	20.28	21.00	19.73	20.64	39.82	39.60	45.39	56.78	55.75	55.32	55.85	54.51	53.71	53.67
800	12.28	10.50	10.40	10.57	11.50	11.82	11.31	11.35	11.65	11.48	11.45	11.48	11.29	11.29	11.45	11.65	14.50	18.29	18.15	25.70	29.80	29.45	28.73	27.98	27.65	27.57	26.71
900	10.26	10.97	12.31	11.49	13.96	13.98	14.55	15.19	16.82	23.08	31.99	35.27	37.65	30.27	25.84	31.95	33.76	30.38	34.58	36.76	18.83	37.05	29.57	34.48	40.93	44.88	39.31
1000	10.45	11.08	12.18	11.58	13.72	13.70	14.20	14.70	16.15	21.64	29.36	34.29	38.25	33.60	32.78	31.46	33.06	30.42	31.25	32.68	22.66	38.33	31.89	36.02	41.42	44.72	39.99
1100	10.64	11.18	12.05	11.66	13.48	13.42	13.85	14.21	15.48	20.20	26.73	31.25	37.98	28.98	26.86	30.98	32.36	30.45	21.03	34.23	26.48	39.60	34.21	37.57	41.92	44.56	40.67
1200	10.83	11.29	11.92	11.75	13.24	13.14	13.50	13.71	14.80	18.76	24.10	28.65	31.15	24.33	27.37	30.49	31.66	30.49	24.40	20.36	30.31	40.88	36.54	39.11	42.41	44.40	41.35
1300	11.03	11.39	11.79	11.83	13.00	12.86	13.15	13.22	14.13	17.32	21.47	38.32	42.67	38.40	37.88	34.16	30.96	30.52	27.76	25.95	34.13	42.15	38.86	40.66	42.91	44.24	42.03
1400	11.24	11.43	11.85	11.72	12.21	12.36	12.58	12.81	13.03	14.70	17.30	27.92	31.98	29.62	28.31	29.45	30.87	31.43	32.17	30.14	37.46	41.56	39.31	40.78	42.25	43.72	41.97
1500	11.37	11.73	11.16	12.40	13.61	12.75	12.87	12.08	13.63	16.79	19.28	30.85	30.82	30.11	29.06	29.18	28.35	28.84	32.41	39.92	42.78	48.44	47.25	46.60	46.19	44.61	44.88
1600	11.63	11.64	11.59	11.89	11.73	11.80	11.88	11.82	11.68	11.82	12.04	13.78	15.36	16.69	29.33	28.48	29.47	31.50	38.90	41.32	45.11	44.11	43.95	43.87	43.24	43.40	43.33

Muestras / Atributos	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53
400	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
500	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
600	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
700	49.46	53.48	64.73	62.50	63.92	67.37	67.12	66.82	66.71	66.50	66.57	66.39	80.92	94.16	99.07	99.62	99.76	99.75	99.96	100.00	100.00	100.00	100.00	100.00	100.00	100.00
800	27.18	33.54	37.76	37.95	43.18	45.09	44.76	44.42	44.32	44.18	43.82	43.90	67.48	93.68	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
900	42.70	48.99	44.74	46.09	39.94	50.02	47.94	45.69	42.60	53.71	41.67	44.13	50.28	89.63	99.69	74.66	84.00	85.67	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1000	42.77	47.99	44.31	45.53	40.39	48.93	47.13	45.21	42.57	51.95	41.62	41.60	50.24	86.47	99.69	78.44	86.40	87.82	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1100	42.84	46.98	43.87	44.97	40.84	47.83	46.32	44.74	42.54	50.19	41.57	42.20	49.80	83.31	99.69	82.22	88.80	89.97	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1200	42.91	45.98	43.44	44.41	41.29	46.74	45.51	44.26	42.51	48.43	41.52	42.40	47.30	57.23	53.73	86.00	91.20	92.12	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1300	42.98	44.97	43.00	43.85	41.74	45.64	44.70	43.79	42.48	46.67	41.47	40.22	47.60	49.24	66.01	89.78	93.60	94.27	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1400	42.29	43.41	42.21	43.14	41.10	44.18	43.60	43.14	42.14	45.22	41.33	40.98	44.22	60.29	75.43	92.44	95.20	95.70	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1500	44.65	44.07	42.84	43.04	44.82	44.19	43.67	43.18	43.04	42.52	41.55	41.09	61.32	76.72	96.32	99.58	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
1600	42.43	41.40	41.34	42.02	42.00	41.99	41.98	42.19	42.08	41.70	41.23	61.05	75.39	95.52	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Figura 21. Reporte de precisión alcanzada en experimento realizando modificaciones sobre el número de muestras consideradas y la cantidad de atributos utilizados resaltando como mapa de calor en rojo los resultados menos deseables y en verde los más deseables.

En este orden de ideas, se llevó a cabo una etapa de refinamiento de la propuesta en la que se variaron de manera sistemática tanto la cantidad de muestras sintéticas generadas como el número de atributos de estas a fin de alcanzar niveles óptimos de entrenamiento, favoreciendo el desempeño y evitando problemáticas asociadas al exceso de atributos y muestras en el proceso de entrenamiento.

Los resultados de las variaciones realizadas son interesantes, ya que como se puede apreciar en la Figura 21, en los experimentos realizados con 700 muestras sintéticas o más, se nota una tendencia en los datos, de acuerdo con la cual, se alcanzan resultados óptimos al utilizar entre 42 y 46 atributos, es decir, sin incluir información relativa a los

ángulos de las articulaciones de tobillos, muñecas y rodillas, y generando entre 900 y 1100 muestras para entrenamiento, dejando fuera del proceso aquellas que fueron creadas generando rotaciones aleatorias en las distintas articulaciones.

Por otro lado, en la parte superior de la misma figura, en donde se reportan resultados obtenidos para experimentos con menos de 700 registros, es de interés notar que los patrones descritos previamente no se siguen y que a primera vista las ejecuciones del experimento parecen indicar que, al contar con pocas muestras, los atributos no son tan relevantes, y que incluso es posible alcanzar niveles de reconocimiento óptimos usando únicamente 3 características (i.e. posición de la cabeza y de ambos hombros). Esto ocurre debido a que, al contar con una menor cantidad de muestras, y considerando que las variaciones iniciales realizadas solamente contemplaban un cambio de posición de las articulaciones del cuerpo, los registros utilizados para entrenamiento y para pruebas son muy similares.

Muestras sintéticas / Atributos	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
400	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.01	0.00	0.02	0.00	0.02	0.00	0.02	0.00	0.01	0.02	0.02	0.00	0.01	0.02	0.00	0.02	0.02	0.01	0.01
500	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.01
600	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.00	0.02	0.01	0.00	0.01	0.02	0.02	0.01
700	0.00	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.02	0.02	0.00	0.02	0.00	0.02	0.01	0.01	0.00	0.00	0.01	0.02	0.01	0.01	0.02	0.02	0.02	0.02
800	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.02	0.02
900	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.02	0.02	0.02	0.00	0.01	0.00	0.01	0.00	0.01	0.00	0.00	0.02	0.01	0.01	0.02	0.03	0.02	0.02	0.03
1000	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.02	0.03	0.03	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.02	0.00	0.01	0.02	0.03	0.02	0.02	0.02	0.03
1100	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.02	0.03	0.03	0.03	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.03	0.00	0.01	0.03	0.04	0.02	0.02	0.04	0.04
1200	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.03	0.03	0.04	0.04	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.03	0.00	0.01	0.03	0.04	0.02	0.02	0.04	0.04
1300	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.03	0.04	0.04	0.04	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.04	0.00	0.01	0.04	0.05	0.02	0.02	0.05	0.05
1400	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.04	0.04	0.05	0.05	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.04	0.00	0.01	0.04	0.05	0.02	0.02	0.05	0.05
1500	0.00	0.02	0.00	0.00	0.01	0.02	0.00	0.02	0.02	0.01	0.02	0.02	0.05	0.01	0.02	0.02	0.02	0.03	0.01	0.03	0.03	0.03	0.05	0.05	0.05	0.03	0.03
1600	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.03	0.03	0.03	0.03	0.05	0.03	0.03	0.05	0.03	0.03	0.05	0.05	0.06	0.05	0.05	0.05

Muestras sintéticas / Atributos	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53
400	0.00	0.01	0.00	0.02	0.02	0.02	0.00	0.00	0.00	0.01	0.00	0.02	0.02	0.01	0.01	0.02	0.01	0.02	0.01	0.03	0.00	0.01	0.02	0.02	0.02	0.02
500	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.00	0.01	0.01	0.00	0.02	0.01	0.00	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.02	0.01	0.02	0.02	0.02
600	0.01	0.01	0.02	0.01	0.03	0.02	0.01	0.01	0.02	0.02	0.02	0.01	0.02	0.02	0.01	0.02	0.03	0.02	0.02	0.03	0.02	0.02	0.01	0.03	0.02	0.02
700	0.02	0.02	0.01	0.02	0.02	0.01	0.02	0.01	0.03	0.02	0.01	0.03	0.02	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.02	0.04	0.03
800	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.02	0.02	0.03	0.03	0.03	0.03	0.04	0.03	0.03	0.03	0.03	0.04	0.04	0.04
900	0.03	0.03	0.02	0.03	0.01	0.02	0.03	0.02	0.02	0.02	0.02	0.04	0.02	0.02	0.04	0.03	0.03	0.03	0.05	0.03	0.03	0.04	0.04	0.04	0.05	0.05
1000	0.03	0.03	0.02	0.03	0.01	0.02	0.03	0.03	0.02	0.02	0.02	0.05	0.02	0.02	0.05	0.04	0.03	0.04	0.06	0.03	0.04	0.04	0.05	0.05	0.06	0.06
1100	0.04	0.04	0.02	0.04	0.00	0.02	0.04	0.03	0.02	0.02	0.02	0.06	0.02	0.02	0.06	0.04	0.03	0.04	0.07	0.03	0.04	0.05	0.06	0.05	0.07	0.07
1200	0.04	0.04	0.02	0.04	0.00	0.02	0.04	0.04	0.02	0.02	0.02	0.07	0.02	0.02	0.07	0.05	0.03	0.05	0.08	0.03	0.05	0.05	0.07	0.06	0.08	0.08
1300	0.05	0.05	0.02	0.05	0.00	0.02	0.05	0.04	0.02	0.02	0.02	0.08	0.02	0.02	0.08	0.05	0.03	0.05	0.09	0.03	0.05	0.06	0.08	0.06	0.09	0.09
1400	0.05	0.05	0.02	0.05	0.00	0.02	0.05	0.05	0.02	0.02	0.02	0.09	0.02	0.02	0.09	0.06	0.03	0.06	0.10	0.03	0.06	0.06	0.09	0.07	0.10	0.10
1500	0.05	0.03	0.05	0.05	0.06	0.05	0.06	0.05	0.06	0.05	0.06	0.05	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.09	0.06	0.06	0.06	0.08	0.08	0.08
1600	0.05	0.06	0.06	0.05	0.05	0.06	0.06	0.06	0.05	0.06	0.08	0.08	0.06	0.08	0.08	0.08	0.08	0.08	0.08	0.09	0.08	0.08	0.08	0.09	0.09	0.09

Figura 22. Reporte de precisión alcanzada en experimento realizando modificaciones sobre el número de muestras consideradas y la cantidad de atributos utilizados resaltando como mapa de calor en rojo los resultados menos deseables y en verde los más deseables.

El hecho de que la información de los ángulos de rotación de algunas de las articulaciones sea descartable de acuerdo con la información obtenida obedece a que los gestos analizados, no hacen uso de tales características como se puede confirmar en la Figura 17, es decir, en todos los gestos usados en el presente experimento, las

muñecas, tobillos y rodillas permanecen en estado de extensión. Esto significa que, aunque para el experimento reportado, los atributos podrían no ser considerados en el proceso de entrenamiento sin afectar la calidad de los resultados, la naturaleza de los gestos a utilizar en otros contextos podría hacer necesaria su reincorporación. Aunado a esto, en los resultados se puede observar que, en todos los experimentos realizados, al considerar la totalidad de atributos se alcanzaron siempre niveles óptimos de precisión.

Al enriquecer esta interpretación con los resultados obtenidos sobre el tiempo utilizado para entrenamiento y evaluación que se reportan en la Figura 22, es posible notar que, si bien existe una tendencia que muestra un incremento en el tiempo requerido al utilizar grandes cantidades de muestras conteniendo todos los atributos seleccionados, estos tiempos aún podrían ser considerados en múltiples contextos no críticos como óptimos para respuesta. Si se desea priorizar el tiempo de respuesta junto con la precisión y la capacidad para extender o replicar la solución a múltiples dominios, se puede considerar que el experimento realizado con 800 muestras sintéticas y 53 atributos es el que obtuvo los mejores resultados en conjunto, con un 100% de precisión en la detección sobre un conjunto de pruebas, y un tiempo de entrenamiento y reconocimiento de 0.04 segundos.

6.5. Conclusiones del capítulo

Alineado con los objetivos de este trabajo de disertación, dentro del presente capítulo se muestran los resultados obtenidos tras la realización de experimentos acerca de la apreciación de las interfaces gestuales y su comparación con otras modalidades de interacción, de la definición de conjuntos gestuales por parte de usuarios para promover la UX, de etiquetado de gestos para permitir la identificación y análisis de la interacción, y de entrenamiento y reconocimiento gestual con el fin de explorar una solución basada en el enfoque de aprendizaje de una sola toma.

La relación de los resultados reportados en este capítulo con respecto al marco epistémico se muestra en la Tabla 10.

Tabla 10. Relación entre elementos del marco epistémico y resultados reportados tras los experimentos realizados.

Elemento del marco epistémico	Concepto en los resultados	Justificación
<p>O. E. 1. Realizar un mapeo de usuario en 3D por medio de dispositivos basados en visión y API para permitir el monitoreo de movimientos y el reconocimiento gestual.</p>	<p>Se realizó el mapeo de las articulaciones de usuarios satisfactoriamente considerando su posición y el ángulo entre ellas utilizando las estructuras JSON y ARFF brindadas.</p>	<p>La utilización de las estructuras propuestas permitió la recolección y organización de la información que se obtuvo por medio de la utilización de sensores y su uso para el entrenamiento de un clasificador. El monitoreo de gestos dinámicos supone la comparación de múltiples estructuras que describan aisladamente datos recolectados en un momento determinado.</p>
<p>O. E. 2. Diseñar e implementar una base de datos de gestos a partir de información obtenida durante estudios preliminares, para permitir la comparación y el análisis.</p>	<p>Gestos obtenidos por medio de sensores basados en estudios de elicitación para tareas abstractas y organizados en las estructuras proporcionadas.</p>	<p>Los estudios preliminares presentados resultaron en la identificación de un conjunto de aproximadamente 200 gestos que los usuarios asocian con tareas abstractas o concretas y que consideran naturales de realizar con distintas partes del cuerpo. La información de estos gestos se registró de forma gráfica, textual y por medio de las estructuras JSON y ARFF propuestas.</p>
<p>O. E. 3. Proponer un método para etiquetado de gestos utilizando una notación genérica, para soporte a las tareas de reconocimiento y de asociación de significados.</p>	<p>Gestos definidos y etiquetados utilizando una notación genérica. Asociación de gestos a tareas concretas y abstractas.</p>	<p>Se destacó la necesidad de etiquetar los gestos usando notaciones simples que permitan diferenciarlos y asociarlos. Se mostró la posibilidad de utilizar una notación en la que las partes del cuerpo observadas determinarían el inicio del nombre de tales gestos y se mostró una más en la que se antepone un sufijo solamente como identificador para los procesos de entrenamiento y reconocimiento.</p>
<p>O. E. 4. Implementar reconocimiento gestual utilizando técnicas de ML y siguiendo un enfoque de aprendizaje <i>one-shot</i> para permitir interacción natural.</p>	<p>Entrenamiento de un clasificador de gestos a partir de muestras únicas y tras la generación de datos sintéticos. Reconocimiento de los gestos definidos reportando precisión del 100%. Refinación de los resultados obtenidos en un experimento inicial para identificar parámetros que permiten llevar a cabo de forma óptima el proceso de ML.</p>	<p>Tras el etiquetado de muestras únicas de gestos recolectadas de usuarios, se llevó a cabo un proceso de creación de datos sintéticos resultando en conjuntos lo suficientemente robustos para entrenar un clasificador bayesiano que logró identificar gestos de manera satisfactoria.</p>

<p>P. I. 1. ¿Cómo se puede lograr la interacción gestual de manera que resulte natural para los usuarios?</p>	<p>Estudios de elicitación con usuarios, obtención de muestras únicas a través de sensores y utilizando software especializado para generación de datos sintéticos, etiquetado de datos, utilización de estructuras para organización, entrenamiento y clasificación.</p>	<p>Dentro de los resultados se muestran resultados sobre la usabilidad de sistemas con interacción gestual y su comparación con las interfaces gráficas. Asimismo, se presentaron resultados relacionados a la capacidad para identificar gestos a partir de una única muestra tomada de usuarios.</p>
<p>P. I. 2. ¿Qué algoritmos, técnicas, métodos y estrategias se deben incluir en una metodología para facilitar la implementación de aplicaciones con interacción basada en gestos?</p>	<p>Identificación de regiones de interés, segmentación de la información, obtención de muestras de gestos de parte de usuarios y tomando en cuenta los estudios de elicitación realizados, definición y aplicación de transformaciones sobre la información recolectada para permitir el proceso de entrenamiento y clasificación.</p>	<p>Existen reportes de algoritmos y técnicas en la literatura que efectúan de forma óptima la labor de recolectar datos relevantes sobre gestos corporales. Para el caso de este trabajo en específico, se puede utilizar cualquiera que permita la obtención de información sobre la posición de articulaciones de interés. A partir de tal dato, es posible calcular el ángulo entre articulaciones y, por lo tanto, llevar a cabo los procesos de creación de muestras sintéticas, entrenamiento y clasificación.</p>
<p>P. I. 3. ¿Qué implicaciones resultan de la utilización de gestos definidos por el usuario en una evaluación de usabilidad y experiencia de usuario, realizada sobre aplicaciones basadas en interacción gestual?</p>	<p>Los usuarios participantes en los estudios realizados mostraron preferencia por los sistemas que hacen uso de interacción gestual, especialmente cuando se les permitía proponer gestos propios para ello.</p>	<p>Fue posible observar en los estudios preliminares que los usuarios evaluaron de mejor manera sistemas que hacen uso de interacción gestual que sus contrapartes de interfaces gráficas, incluso cuando cometían más errores y tomaban más tiempo para realizar tareas determinadas.</p>

La realización de los experimentos descritos y sus resultados alineados con los objetivos del trabajo de disertación, permiten la obtención de conclusiones generales y la definición de líneas de investigación para trabajos futuros. Esta información se describe en el siguiente capítulo.

7. CONCLUSIONES Y TRABAJO FUTURO

Adicionalmente a las conclusiones específicas que se brindan dentro de cada uno de los capítulos que conforman este trabajo de disertación, en el presente se incluyen las conclusiones generales del mismo organizadas de acuerdo con cada uno de los objetivos y las preguntas de investigación que se plantearon en el marco epistémico.

Para el cumplimiento del O.E.1 consistente en realizar un mapeo de usuario en 3D por medio de dispositivos basados en visión y API para permitir el monitoreo de movimientos y el reconocimiento gestual, en el marco teórico se abordaron los conceptos de ergonomía y comportamiento humano para definir los puntos a analizar y a considerar para la modelación de usuarios.

Dentro del estado del arte, se analizaron modelos de usuario, trabajos relacionados y vocabularios de gestos definidos por el usuario con el fin de conocer alternativas para definir la estructura y características del modelo a proponer. Se identificó además que, tener un vocabulario de gestos definidos por usuarios permite saber cuáles son los gestos comunes y en los que se debe poner especial atención. En la metodología, se realizaron modelos de usuario basados en estructuras JSON y ARFF buscando la identificación de movimientos y características relevantes de articulaciones de interés.

El modelado de usuarios por medio de sus características físicas estableciendo métricas y estructuras para su representación hacen que los gestos realizados puedan ser monitoreados, estudiados e interpretados. Dentro del capítulo de presentación de resultados se mostró el procedimiento seguido para llevar a cabo la generación de muestras sintéticas y el entrenamiento de un clasificador gestual a partir de ellas. Finalmente, Se realizó el mapeo de las articulaciones de usuarios satisfactoriamente considerando su posición y el ángulo entre ellas utilizando un sensor Kinect, Software para procesamiento, y estructuras de datos para la organización de datos que resulta en su posterior uso para el entrenamiento de un clasificador.

El O.E.2 consistente en diseñar e implementar una base de datos de gestos a partir de información obtenida durante estudios preliminares, para permitir la comparación y el análisis, fue abordado en el marco teórico a través de la identificación y observación de métricas y puntos de interés para realizar el diseño de la estructura de la base de

datos. Además, los conceptos relacionados a HCI permitieron obtener la información correspondiente a gestos y a las tareas abstractas que fueron la base para determinar los gestos de interés. En el estado del arte, se estudiaron modelos de usuario y vocabularios de gestos definidos por el usuario para definir la estructura y características de interés y de identificar gestos candidatos para formar parte de una base de datos.

En este mismo sentido, dentro de la propuesta metodológica, se presentaron modelos de usuario basados en estructuras JSON y ARFF conformados por características referentes a posición, traslación y rotación de articulaciones identificadas como puntos de interés. Se definieron también, los procesos de entrenamiento a priori y a posteriori para poblar una base de datos con información predeterminada y enriquecerla con nuevos datos que se recolecten tras la interacción del usuario con un sistema. Tras contar con estructuras definidas, se realizaron experimentos para identificar vocabularios gestuales para la interacción con aplicaciones de distintos dominios, de manera que se pudiera contar con una base de datos de gestos proporcionados por usuarios permitiendo además la abstracción de sus características hacia tales estructuras. Los estudios realizados resultaron en la identificación de un conjunto de aproximadamente 200 gestos que los usuarios asocian con tareas abstractas o concretas y que consideran naturales de realizar con distintas partes del cuerpo. La información de estos gestos se registró de forma gráfica, textual y por medio de las estructuras JSON y ARFF propuestas.

El O.E.3 descrito como la intención de proponer un método para etiquetado de gestos utilizando una notación genérica, para soporte a las tareas de reconocimiento y de asociación de significados fue cubierto dentro del marco teórico por medio de la exploración de la asociación de gestos con tareas abstractas para favorecer su etiquetado y posterior asociación con significados específicos en contextos determinados. Dentro del estado del arte, se analizaron trabajos relacionados con la creación de vocabularios de gestos definidos por el usuario, identificando la viabilidad de la utilización de una notación propia para un contexto, pero también de un etiquetado genérico asociado a tareas abstractas.

Dentro de la propuesta metodológica, se observó que era posible asignar un identificador a fragmentos de la información obtenida tras la utilización de sensores y de Software con el fin de facilitar su inclusión en las estructuras JSON y ARFF y posteriormente, de permitir la identificación tanto de gestos predeterminados como de otros definidos por el usuario. La notación propuesta permite una rápida identificación manual de gestos al incluir dentro de ella letras que indican el origen del gesto (i.e. la parte del cuerpo con la que fue realizado) y números consecutivos. La notación propuesta fue probada dentro del capítulo de resultados al llevar a cabo el entrenamiento de un clasificador gestual usando etiquetas que enumeraban los gestos sin tener una relación directa con sus características. Esto permite realizar una generalización de los procesos y por lo tanto asociar gestos a tareas distintas en contextos diferentes. Igualmente, se destacó la necesidad de etiquetar los gestos usando notaciones simples que permitan diferenciarlos y asociarlos con tareas. Se exploró la posibilidad de utilizar una notación en la que se antepone un sufijo solamente como identificador para los procesos de entrenamiento y reconocimiento.

Para el O.E.4. definido como implementar reconocimiento gestual utilizando técnicas de ML y siguiendo un enfoque de aprendizaje one-shot para permitir interacción natural, en el marco teórico, se trabajó por medio del estudio de la implementación del reconocimiento gestual, de técnicas y de enfoques. Se analizó el modelo de contexto con el fin de estimar la viabilidad de realizar una propuesta basada en tareas abstractas y manteniendo independencia a contextos de uso. En el estado del arte, se revisaron trabajos sobre mecanismos, estrategias y técnicas aplicadas para llevar a cabo el reconocimiento gestual, permitiendo el aprendizaje de las experiencias de otros autores y la selección de recursos que resultaron de mayor conveniencia para este trabajo.

Como parte de la propuesta metodológica y de los estudios preliminares realizados, se propusieron estructuras para almacenar la información recolectada de los usuarios por medio de dispositivos y Software, haciendo posible la aplicación de transformaciones que generen múltiples muestras sintéticas que representan ligeras variaciones sobre gestos realizados sobre determinadas características y hacen posible el entrenamiento a posteriori. Esto aunado al desarrollo de la base de datos de gestos

asociada al O.E.3, hicieron posible un óptimo reconocimiento gestual tras la obtención de una sola muestra por parte de los usuarios.

Respondiendo a la P.I.1., ¿Cómo se puede lograr la interacción gestual de manera que resulte natural para los usuarios?, en el marco teórico y el estado del arte, se buscó conocer la experiencia de propuestas anteriores y enriquecerla con la generación de vocabularios de gestos definidos por el usuario para llevar a cabo experimentos en los que hacer evaluaciones de usabilidad y UX. Posteriormente, tras la realización de tales experimentos, fue posible proponer estructuras para almacenar vocabularios de gestos definidos por expertos en interacción y por usuarios, asociándolos tanto a tareas abstractas como a acciones específicas a un contexto de uso. En el presente trabajo se incluyen resultados sobre la usabilidad de sistemas con interacción gestual y su comparación con las interfaces gráficas. Asimismo, se presentaron resultados relacionados a la capacidad para identificar gestos a partir de una única muestra tomada de usuarios.

Sobre la P.I.2., ¿Qué algoritmos, técnicas, métodos y estrategias se deben incluir en una metodología para facilitar la implementación de aplicaciones con interacción basada en gestos?, la revisión a la literatura mostró que los métodos clásicos de procesamiento digital de imágenes alcanzan porcentajes altos en cuanto a la segmentación de recursos en términos de espacio y tiempo, siendo esto de vital importancia para la detección y reconocimiento gestual. Asimismo, los procesos y algoritmos actuales de ML tales como los clasificadores Bayesianos, árboles de decisión y máquinas de vector de soporte, atienden de manera correcta y eficiente la tarea de reconocimiento. La metodología propuesta consideró la utilización de diversas técnicas y algoritmos, pero mantuvo un enfoque independiente a ellas que permite su extensión y adaptación ante la evolución de las capacidades computacionales y a la diversidad de contextos de utilización.

Finalmente, para la P.I.3., ¿Qué implicaciones resultan de la utilización de gestos definidos por el usuario en una evaluación de usabilidad y experiencia de usuario, realizada sobre aplicaciones basadas en interacción gestual?, fue posible observar en los estudios preliminares que los usuarios evaluaron de mejor manera sistemas que hacen uso de interacción gestual que sus contrapartes de interfaces gráficas, incluso

cuando cometían más errores y tomaban más tiempo para realizar tareas determinadas.

Tomando en cuenta las conclusiones reportadas en los párrafos anteriores del presente capítulo, es posible observar que, el objetivo general del trabajo de investigación de proponer una metodología para la definición de conjuntos de gestos definidos por el usuario, y para su reconocimiento en contextos determinados, utilizando técnicas de procesamiento digital de imágenes y de clasificación, y siguiendo un enfoque de aprendizaje *one-shot*, para favorecer la interacción natural fue correctamente completado.

Los resultados positivos en las evaluaciones de usabilidad y UX realizadas durante experimentos preliminares sobre interacción gestual utilizando conjuntos de gestos definidos por el usuario, la viabilidad para la utilización de una metodología basada en la identificación y el reconocimiento de gestos a partir de una sola toma, y la posibilidad de que la metodología haga uso de notaciones genéricas para describir gestos en contextos múltiples, llevan a la conclusión de que es posible confirmar la hipótesis del trabajo de investigación en la que se definió que a través de la definición de una metodología que incluya técnicas de clasificación y de procesamiento digital de imágenes, junto con algoritmos y estrategias de aprendizaje automático, es posible facilitar el desarrollo de aplicaciones con interacción basada en gestos y mejorar la apreciación de su usabilidad y la experiencia de usuario.

Durante la realización del presente trabajo se encontraron líneas de investigación que, aunque tienen relación con el dominio de acción, suponen ser abordadas desde distintas perspectivas, saliendo de los objetivos planteados en el marco epistémico. Así es como se identifica trabajo futuro a fin de complementar y reforzar lo aquí descrito. Se observa, por ejemplo, la necesidad de contar con métricas e instrumentos para determinar si las NUI están siendo correctamente implementadas y para evaluar la apreciación de los usuarios con respecto a estas. Igualmente, se puede profundizar en estudios para determinar influencia de factor novedad sobre los resultados de usabilidad y UX.

Igualmente, dentro del trabajo futuro, se encuentra la posibilidad de explorar el uso de otros tipos de clasificador a fin de contrastar resultados con los que se reportan en este trabajo e identificar potenciales áreas de mejora en cuanto al desempeño.

Finalmente, si bien se describe que desde la perspectiva de la metodología propuesta el monitoreo de gestos dinámicos supone la comparación de múltiples estructuras que describan aisladamente datos recolectados en un momento determinado, el alcance de los estudios preliminares se enfoca en el análisis de gestos estáticos. Queda entonces abierta la posibilidad de proponer notaciones y mecanismos para ampliar la solución a distintos tipos de actividad humana.

REFERENCIAS

- Aliakseyeu, D., Subramanian, S., Martens, J. B., & Rauterberg, M. (2002). Interaction techniques for navigation through and manipulation of 2D and 3D data. *ACM International Conference Proceeding Series (Vol. 23)*. (pp. 179-188). USA: ACM.
- Al-qaness, M. A., & Li, F. (2016). WiGeR: WiFi-based gesture recognition system. *ISPRS International Journal of Geo-Information*, 5(6), 92.
- An, J., Kwak, H., Jung, S., Salminen, J., Admad, M., & Jansen, B. (2018). Imaginary people representing real numbers: Generating personas from online social media data. *ACM Transactions on the Web (TWEB)*, 12(4), 1-26.
- Anders, S. (2015). What are you really saying? *Physician leadership journal*, 2(2), 82.
- Black, D., Ganze, B., Hettig, J., & Hansen, C. (2017). Auditory display for improving free-hand gesture interaction. *Mensch und Computer 2017-Tagungsband*.
- Bansal, H., & Khan, R. (2018). A Review Paper on Human Computer Interaction. *Conf. Rec. 2018 ISSN Int. Journals of Advanced Research in Computer Science and software Engineering*, 53-56.
- Bastanfard, A., Bastanfard, O., Takahashi, H., & Nakajima, M. (2004). Toward anthropometrics simulation of face rejuvenation and skin cosmetic. *Computer Animation and Virtual Worlds*, 15(3-4), 347-352.
- Bastien, J. C., & Scapin, D. L. (1992). A validation of ergonomic criteria for the evaluation of human-computer interfaces. *International Journal of Human-Computer Interaction*, 4(2), 183-196.
- ben-Aaron, M. (2003). Topological aspects of axonal regeneration. *Medical hypotheses*, 597-600.
- Boone, D. C., & Azen, S. P. (1979). Normal range of motion of joints in male subjects. *The Journal of Bone and Joint Surgery.*, 756-759.
- Boulay, B., Brémond, F., & Thonnat, M. (2005). Posture recognition with a 3d human model. *IEE International Symposium on Imaging for Crime Detection and Prevention (ICDP 2005)* (p. p. 135). London, UK: IET.

- Boudoin, P., Otmane, S., & Mallem, M. (2008). Design of a 3D navigation technique supporting VR interaction. *AIP Conference Proceedings (Vol. 1019, No. 1)* (pp. 149-153). Annaba, Algeria: American Institute of Physics.
- Bowman, D. A., Koller, D., & Hodges, L. F. (1997). Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. *Proceedings of the IEEE Virtual Reality Annual International Symposium* (pp. 45-52). Albuquerque, NM, USA: IEEE.
- Bridger, R. (2017). *Introduction to human factors and ergonomics*. Boca Raton, Florida, USA: CRC Press.
- Calvary, G., Coutaz, J., Thevenin, D., Limbourg, Q., Bouillon, L., & Vanderdonckt, J. (2003). A unifying reference framework for multi-target user interfaces. *Interacting with computers, 15(3)*, 289-308.
- Cappel, J. J., & Huang, Z. (2007). A usability analysis of company websites. *Journal of Computer Information Systems, 48(1)*, 117-123.
- Caputo, F., D'Amato, E., Greco, A., Notaro, I., & Spada, S. (2018). Human posture tracking system for industrial process design and assessment. *International Conference on Intelligent Human Systems Integration* (pp. 450-455). Dubai, United Arab Emirates: Springer, Cham.
- Caruana, R. (1997). Multitask learning. *Machine learning, 28(1)*, 41-75.
- Chandler, J., & Schwarz, N. (2009). How extending your middle finger affects your perception of others: Learned movements influence concept accessibility. *Journal of Experimental Social Psychology, 45(1)*, 123-128.
- Chen, F. S., Fu, C. M., & Huang, C. L. (2003). Hand gesture recognition using a real-time tracking method and hidden Markov models. *Image and vision computing, 21(8)*, 745-758.
- Chen, Q., Georganas, N. D., & Petriu, E. M. (2007). Real-time vision-based hand gesture recognition using haar-like features. *Instrumentation and Measurement Technology Conference Proceedings* (pp. 1-6). Warsaw, Poland: IEEE.
- Chevtchenko, S. F., Vale, R. F., Macario, V., & Cordeiro, F. R. (2018). A convolutional neural network with feature fusion for real-time hand posture recognition. *Applied Soft Computing, 73*, 748-766.

- Cippitelli, E., Gasparrini, S., Gambi, E., & Spinsante, S. (2016). A human activity recognition system using skeleton data from RGBD sensors. *Computational intelligence and neuroscience*.
- Constantine, L. L. (2003). Canonical abstract prototypes for abstract visual and interaction design. *International Workshop on Design, Specification, and Verification of Interactive Systems* (pp. 1-15). Funchal, Madeira, Portugal: Springer, Berlin, Heidelberg.
- Daniloff, J. K., Noll, J. D., Fristoe, M., & Lloyd, L. L. (1982). Gesture recognition in patients with aphasia. *Journal of Speech and Hearing Disorders*, 47(1), 43-49.
- Daros, W. R. (2002). ¿Qué es un marco teórico? *Enfoques*, 14(1 y 2), 73-112.
- Dey, A. (2016). Machine learning algorithms: a review. *International Journal of Computer Science and Information Technologies*, 7(3), 1174-1179.
- Doležal, K., & Hrženjak, R. (2019). Anthropometric measurements in helath assessment. *Sigurnost: časopis za sigurnost u radnoj i životnoj okolini*, 61(4), 357-364.
- do Rosário, J. L. (2014). Photographic analysis of human posture: a literature review. *Journal of bodywork and movement therapies*, 18(1), 56-61.
- Dow, S., Lee, J., Oezbek, C., MacIntyre, B., Bolter, J. D., & Gandy, M. (2005). Wizard of Oz Interfaces for Mixed Reality Applications. *CHI'05 extended abstracts on human factors in computing systems*, 339-1342.
- Escalante, H. J., Guyon, I., Athitsos, V., Jangyodsuk, P., & Wan, J. (2017). Principal motion components for one-shot gesture recognition. *Pattern Analysis and Applications*, 20(1), 167-182.
- Favre, J., Hayoz, M., Erhart-Hledik, J. C., & Andriacchi, T. P. (2012). A neural network model to predict knee adduction moment during walking based on ground reaction force and anthropometric measurements. *Journal of biomechanics*, 45(4), 692-698.
- Fernandez, R. A., Sanchez-Lopez, J. L., Sampedro, C., Bavle, H., Molina, M., & Campoy, P. (2016). Natural user interfaces for human-drone multi-modal interaction. *International Conference on Unmanned Aircraft Systems (ICUAS)* (pp. 1013-102). Arlington, VA, USA: IEEE.

- Freund, Y., & Schapire, R. (1997). A decision theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences* (55), 119–139.
- Ghotkar, A. S., Kharate, D., & Gajanan, K. (2014). Study of Vision Based Hand Gesture Recognition using Indian Sign Language. *International Journal on Smart Sensing & Intelligent Systems*, 7(1), 55, 56.
- González-Calleros, J. M., Vanderdonckt, J., & Muñoz-Arteaga, J. (2010). A Structured Methodology for Developing 3D Web Applications. *Integrating Usability Engineering for Designing the Web Experience: Methodologies and Principles*, 15-43.
- Gonzalez-Calleros, J. M., Guerrero-García, J., González, C., & Galicia, E. (2018). Is Natural User Interaction Really Natural? An Evaluation of Gesture-Based Navigating Techniques in Virtual Environments. *Computación y Sistemas*, 22(1).
- Grzeszczuk, R., Bradski, G. R., Chu, M. H., & Bouget, J. Y. (2004). *Washington, DC., USA Patent No. 6,788,809*.
- Guyon, I., Athitsos, V., Jangyodsuk, P., Hamner, B., & Escalante, H. J. (2012). Chalearn gesture challenge: Design and first results. *IEEE computer society conference on computer vision and pattern recognition workshops* (pp. 1-6). Providence, Rhode Island, USA: IEEE.
- Höysniemi, J., Hämäläinen, P., & Turkki, L. (2004). Wizard of Oz prototyping of computer vision-based action games for children. *Conference on Interaction design and children: building a community* (pp. 27-34). New York, NY, USA: ACM.
- Halim, Z., & Abbas, G. (2015). A Kinect-based sign language hand gesture recognition system for hearing- and speech-impaired: a pilot study of Pakistani sign language. *Assistive Technology*, 27(1), 34-43.
- Hamad, M., Thomassey, S., & Bruniaux, P. (2014). New human body shape descriptor based on anthropometrics points. *6th International Conference of Soft Computing and Pattern Recognition (SoCPaR)* (pp. 273-278). Tunis, Tunisia: IEEE.
- Hanson, L., Sperling, L., Gard, G., Ipsen, S., & Vergara, C. O. (2009). Swedish anthropometrics for product and workplace design. *Applied ergonomics*, 40(4), 797-806.

- Hoff, L., Hornecker, E., & Bertel, S. (2016). Modifying Gesture Elicitation: Do Kinaesthetic Priming and Increased Production Reduce Legacy Bias? *Proceedings of the TEI'16: Tenth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 86-91). Eindhoven, The Netherlands: ACM.
- Hu, L. Y., Huang, M. W., Ke, S. W., & Tsai, C. F. (2016). The distance function effect on k-nearest neighbor classification for medical datasets. *SpringerPlus*, *5*(1), 1-9.
- Ibanez, R., Soria, Á., Teyseyre, A., & Campo, M. (2014). Easy gesture recognition for Kinect. *Advances in Engineering Software*, *76*, 171-180.
- Interaction Design Foundation. (2020). *Usability*. Retrieved from Interaction Design Foundation: <https://www.interaction-design.org/literature/topics/usability>
- International Ergonomics Association. (2020). *Human Factors/Ergonomics (HF/E)*. Obtenido de International Ergonomics Association: <https://iea.cc/what-is-ergonomics/>
- Jagodziński, P., & Wolski, R. (2015). Assessment of application technology of natural user interfaces in the creation of a virtual chemical laboratory. *Journal of Science Education and Technology*, *24*(1), 16-28.
- Jia, P., Hu, H. H., Lu, T., & Yuan, K. (2007). Head gesture recognition for hands-free control of an intelligent wheelchair. *Industrial Robot: An International Journal*, *34*(1), 60-68.
- Jiang, F., Zhang, S., Wu, S., Gao, Y., & Zhao, D. (2015). Multi-layered gesture recognition with kinect. *The Journal of Machine Learning Research*, *16*(1), 227-254.
- Kaklanis, N., Moschonas, P., Moustakas, K., & Tzovaras, D. (2013). Virtual user models for the elderly and disabled for automatic simulated accessibility and ergonomics evaluation of designs. *Universal access in the information society*, *12*(4), 403-425.
- Kaur, K. (1997). Designing virtual environments for usability. *Human-Computer Interaction INTERACT'97* (pp. 636-639). Sydney, Australia: Springer US.
- Kellokumpu, V., Pietikäinen, M., & Heikkilä, J. (2005). Human activity recognition using sequences of postures. *IAPR Conference on Machine Vision Applications* (pp. 570-573). Tsukuba Science City, Japan: IAPR.
- Kim, S. Y., Han, H. G., Kim, J. W., Lee, S., & Kim, T. W. (2017). A hand gesture recognition sensor using reflected impulses. *IEEE Sensors Journal*, *17*(10), 2975-2976.

- Konečný, J., & Hagara, M. (2014). One-shot-learning gesture recognition using hog-hof features. *The Journal of Machine Learning Research*, *15*(1), 2513-2532.
- Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007). Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering*, *160*, 3-24.
- Kray, C., Nesbitt, D., Dawson, J., & Rohs, M. (2010). User-defined gestures for connecting mobile phones, public displays, and tabletops. *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services* (pp. 239-248). Lisbon, Portugal: ACM.
- Kuo, W. Y., Kuo, C. H., Sun, S. W., Chang, P. C., Chen, Y. T., & Cheng, W. H. (2016). Machine learning-based behavior recognition system for a basketball player using multiple kinect cameras. *IEEE International conference on Multimedia & Expo Workshops (ICMEW)* (pp. 1-1). Seattle, WA, USA: IEEE.
- Kurdyukova, E., Redlin, M., & André, E. (2012). Studying user-defined iPad gestures for interaction in multi-display environment. *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces* (pp. 93-96). Lisbon, Portugal: ACM.
- Laal, F., Madvari, R. F., Balarak, D., Mohammadi, M., Dortaj, E., Khammar, A., & Adineh, H. A. (2018). Relationship between musculoskeletal disorders and anthropometric indices among bus drivers in Zahedan city. *International Journal of Occupational Safety and Ergonomics*, *24*(3), 431-437.
- LaViola Jr, J. J. (2015). Context aware 3D gesture recognition for games and virtual reality. *SIGGRAPH '15: Special Interest Group on Computer Graphics and Interactive Techniques Conference* (pág. 10). Los Angeles, California, USA: ACM.
- Lee, R. H., Rust, D. M., & Zirin, H. (1965). The solar magnetograph of the high altitude observatory. *Applied Optics*, *4*(9), 1081-1084.
- Lewis, J. R. (2018). Measuring perceived usability: The CSUQ, SUS, and UMUX. *International Journal of Human-Computer Interaction*, *34*(12), 1148-1156.
- Lewis, J. R., & Sauro, J. (2018). Item benchmarks for the system usability scale. *Journal of Usability Studies*, *13*(3), 158-167.

- Li, G., Liu, Z., Cai, L., & Yan, J. (2019). Human standing posture recognition based on CNN and pressure floor. *Journal of Computational Methods in Sciences and Engineering, (Preprint)*, 1-10.
- Li, Z., & Jarvis, R. (2009). Real time hand gesture recognition using a range camera. *Australasian Conference on Robotics and Automation*, (pp. 21-27). Sydney, Australia.
- Liang, R. H., & Ouhyoung, M. (1998). A real-time continuous gesture recognition system for sign language. *Proceedings of the third IEEE international conference on automatic face and gesture recognition* (págs. 558-567). Nara, Japan: IEEE.
- Littman, A. J., White, E., & Kristal, A. R. (2007). Anthropometrics and prostate cancer risk. *American journal of epidemiology*, 165(11), 1271-1279.
- Liu, H., & Wang, L. (2018). Gesture recognition for human-robot collaboration: A review. *International Journal of Industrial Ergonomics*, 68, 355-367.
- Liu, J., Zhong, L., Wickramasuriya, J., & Vasudevan, V. (2009). uWave: Accelerometer-based personalized gesture recognition and its applications. *Pervasive and Mobile Computing*, 5(6), 657-675.
- Liu, X., Sacks, J., Zhang, M., Richardson, A. G., Lucas, T. H., & Van der Spiegel, J. (2017). The virtual trackpad: An electromyography-based, wireless, real-time, low-power, embedded hand-gesture-recognition system using an event-driven artificial neural network. *IEEE Trans. Circuits Syst. II Express Briefs*, 64, 1257,1261.
- McCormick, J., Vincs, K., Nahavandi, S., Creighton, D., & Hutchison, S. (2014). Teaching a digital performing agent: Artificial neural network and hidden markov model for recognising and performing dance movement. *Proceedings of the 2014 International Workshop on Movement and Computing* (pág. 70). Paris, France: ACM.
- Malima, A. K., Özgür, E., & Çetin, M. (2006). A fast algorithm for vision-based hand gesture recognition for robot control. *Proceedings of the 14th Signal Processing and Communications Applications*. Antalya, Turkey: IEEE.

- Maia, C. L., & Furtado, E. S. (2016). A systematic review about user experience evaluation. *International Conference of Design, User Experience, and Usability* (pp. 445-455). Toronto, Canada: Springer, Cham.
- Maïke, V. R., Neto, L. D., Baranauskas, M. C., & Goldenstein, S. K. (2014). Seeing through the Kinect: a survey on heuristics for building natural user interfaces environments. *International Conference on Universal Access in Human-Computer Interaction* (pp. 407-418). Heraklion, Crete, Greece: Springer International Publishing.
- Mamoshina, P., Ojomoko, L., Yanovich, Y., Ostrovski, A., Botezatu, A., Prikhodko, P., & Ogu, I. O. (2018). Converging blockchain and next-generation artificial intelligence technologies to decentralize and accelerate biomedical research and healthcare. *Oncotarget*, *9*(5), 5665.
- Mann, S. (2002). *Humanistic Intelligence as a Basis for Intelligent Image Processing*. Wiley-IEEE Press.
- Marin, G., Dominio, F., & Zanuttigh, P. (2014). Hand gesture recognition with leap motion and kinect devices. *IEEE International Conference on Image Processing (ICIP)* (pp. 1565-1569). Paris, France: IEEE.
- Martin, S., Ohn-Bar, E., Tawari, A., Møgelmoose, A., & Trivedi, M. M. (2015). Vision for Intelligent Vehicles and Application: A Challenging in-the Wild Dataset. *IEEE Conference on Computer Vision and Pattern Recognition Workshops: The Future of Datasets in Vision*. Boston, MA, USA: IEEE.
- Mitra, S., & Acharya, T. (2007). Gesture recognition: A survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *37*(3), 311-324.
- Molchanov, P., Gupta, S., Kim, K., & Kautz, J. (2015). Hand gesture recognition with 3D convolutional neural networks. *Proceedings of the IEEE conference on computer vision and pattern recognition workshops* (pp. 1-7). Boston, MA, USA: IEEE.
- Molina, A. I., Giraldo, W. J., Gallardo, J., Redondo, M. A., Ortega, M., & García, G. (2012). CIAT-GUI: A MDE-compliant environment for developing Graphical User Interfaces of information systems. *Advances in Engineering Software*, *52*, 10-29.

- Modanwal, G., & Sarawadekar, K. (2018). A New Dactylology and Interactive System Development for Blind-Computer Interaction. *IEEE Transactions on Human-Machine Systems*, *48*(2), 207-212.
- Morris, M. R., Danielescu, A., Drucker, S., Fisher, D., Lee, B., & Wobbrock, J. O. (2014). Reducing legacy bias in gesture elicitation studies. *Interactions*, *21*(3), 40-45.
- Mosler, A. B., Crossley, K. M., Thorborg, K., Whiteley, R. J., Weir, A., Serner, A., & Hölmich, P. (2017). Hip strength and range of motion: normal values from a professional football league. *Journal of science and medicine in sport*, 339-343.
- Motaghi, H. (2015). *Creativity and technology in the context of creative industries*. Montréal, Canada: Université du Québec à Montréal.
- Nielsen, J. (1994). Usability inspection methods. *Conference companion on Human factors in computing systems* (págs. 413-414). Boston, Massachusetts, USA: ACM.
- Nielsen, J. (1995). How to conduct a heuristic evaluation. *Nielsen Norman Group 1*, 1-8.
- Norman, D. (2010). The way I see it: Natural user interfaces are not natural. *Interactions*, *Volume 17, Number 3*, 6-10.
- Norman, D. A., & Nielsen, J. (2010). Gestural interfaces: a step backward in usability. *Interactions*, *17*(5), 46-49.
- Okoli, C., & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *Sprouts: Working Papers on Information Systems*, *10*(26), 10-26.
- Opitz, D., & Maclin, R. (1999). Popular ensemble methods: An empirical study. *Journal of artificial intelligence research*, *11*, 169-198.
- Organización Internacional de Normalización. (2011). *Modelo de Calidad del producto de Software*. Obtenido de ISO/IEC 25010: <http://iso25000.com/index.php/normas-iso-25000/iso-25010>
- Organización Internacional de Normalización. (2018). *Modelo de la Ergonomía de la Interacción Humano-Sistema*. Obtenido de ISO 9241-11: <https://www.iso.org/standard/63500.html>
- Patsadu, O., Nukoolkit, C., & Watanapa, B. (2012). Human gesture recognition using Kinect camera. *(JCSSE) International Joint Conference on Computer Science and Software Engineering* (pp. 28-32). Bangkok, Thailand: IEEE.

- Pisharady, P. K., & Saerbeck, M. (2015). Recent methods and databases in vision-based hand gesture recognition: A review. *Computer Vision and Image Understanding*, *141*, 152-165.
- Piumsomboon, T., Clark, A., Billinghamurst, M., & Cockburn, A. (2013). User-defined gestures for augmented reality. *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 955-960). Paris, France: ACM.
- Pollack, J. B., & Barnden, J. A. (1988). *A Work-shop on High-Level Connectionist Models*. Las Cruces, NM, USA.: New Mexico State University Las Cruces Computing Research Lab.
- Pu, Q., Gupta, S., Gollakota, S., & Patel, S. (2013). Whole-home gesture recognition using wireless signals. *Proceedings of the 19th annual international conference on Mobile computing & networking* (pp. 27-38). New York, NY, USA: ACM.
- Pytel, K., Nawarycz, T., Ostrowska-Nawarycz, L., & Drygas, W. (2015). Anthropometric predictors and Artificial Neural Networks in the diagnosis of hypertension. *Federated Conference on Computer Science and Information Systems (FedCSIS)* (pp. 287-290). Lodz, Poland: IEEE.
- Ren, Z., Meng, J., Yuan, J., & Zhang, Z. (2011). Robust hand gesture recognition with kinect sensor. *Proceedings of the 19th ACM international conference on Multimedia* (pp. 759-760). Scottsdale, Arizona, USA: ACM.
- Ren, Z., Yuan, J., Meng, J., & Zhang, Z. (2013). Robust part-based hand gesture recognition using kinect sensor. *IEEE transactions on multimedia*, *15*(5), 1110-1120.
- Ruiz, J., Li, Y., & Lank, E. (2011). User-defined motion gestures for mobile interaction. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 197-206). Vancouver, BC, Canada: ACM.
- Rusu, C., Rusu, V., Roncagliolo, S., & González, C. (2015). Usability and user experience: what should we care about? *International Journal of Information Technologies and Systems Approach (IJITSA)*, *8*(2), 1-12.
- Schlömer, T., Poppinga, B., Henze, N., & Boll, S. (2008). Gesture recognition with a Wii controller. *Proceedings of the 2nd international conference on Tangible and embedded interaction* (págs. 11-14). Bonn, Germany: ACM.

- Schrepp, M., Hinderks, A., & Thomaschewski, J. (2014). Applying the user experience questionnaire (UEQ) in different evaluation scenarios. *International Conference of Design, User Experience, and Usability* (pp. 383-392). Heraklion, Crete, Greece: Springer, Cham.
- Sassi, I. B., Mellouli, S., & Yahia, S. B. (2017). Context-aware recommender systems in mobile environment: On the road of future research. *Information Systems*, 72, 27-61.
- Singh, V. (2019). Personas and Scenarios as a Methodology for Information Sciences. *Qualitative and Quantitative Methods in Libraries*, 7(1), 123-134.
- Smith, T. R., & Gilbert, J. E. (2018). Dancing to design: a gesture elicitation study. *Proceedings of the 17th ACM Conference on Interaction Design and Children* (pp. 638-643). New York, NY, USA: ACM.
- Starner, T., Weaver, J., & Pentland, A. (1998). Real-time American sign language recognition using desk and wearable computer-based video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12), 1371-1375.
- Sutton, R. S. (1992). Introduction: The challenge of reinforcement learning. *Reinforcement Learning*, Springer, Boston, MA., 1-3.
- Syam, N., & Sharma, A. (2018). Waiting for a sales renaissance in the fourth industrial revolution: Machine learning and artificial intelligence in sales research and practice. *Industrial Marketing Management*, 69, 135-146.
- Tan, D. S., Robertson, G. G., & Czerwinski, M. (2001). Exploring 3D navigation: combining speed-coupled flying with orbiting. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 418-425). Seattle, WA, USA: ACM.
- Tran, C., & Trivedi, M. M. (2011). 3-D posture and gesture recognition for interactivity in smart spaces. *IEEE Transactions on Industrial Informatics*, 8(1), 178-187.
- Valdes, C., Eastman, D., Grote, C., Thatte, S., Shaer, O., Mazalek, A., . . . Konkel, M. (2014). Exploring the design space of gestural interaction with active tokens through user-defined gestures. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 4107-4116). New York, NY, USA: ACM.

- Vallet, F., Puchinger, J., Millionig, A., Lamé, G., & Nicolai, I. (2020). Tangible futures: combining scenario thinking and personas—A pilot study on urban mobility. *Futures*, 102513.
- Vatavu, R. D., & Wobbrock, J. O. (2016). Between-Subjects Elicitation Studies: Formalization and Tool Support. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 3390-3402). San Jose, CA, USA: ACM.
- Vo, M. T. (1994). Incremental learning using the time delay neural network. *Proceedings of ICASSP'94. IEEE International Conference on Acoustics, Speech and Signal Processing (Vol. 2)*. (págs. II-629). Adelaide, South Australia, Australia: IEEE.
- Wüthrich, C. A. (1999). An analysis and a model of 3d interaction methods and devices for virtual reality. *Proceedings of the Eurographics Workshop* (pp. 18-29). Braga, Portugal: Springer, Vienna.
- Wan, Q., Li, Y., Li, C., & Pal, R. (2014). Gesture recognition for smart home applications using portable radar sensors. *Proceedings of the 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society EMBC* (págs. 6414-6417). Chicago, Illinois, USA: IEEE.
- Wang, I., Narayana, P., Smith, J., Draper, B., Beveridge, R., & Ruiz, J. (2018). EASEL: Easy Automatic Segmentation Event Labeler. *23rd International Conference on Intelligent User Interfaces* (pp. 595-599). Tokyo, Japan: ACM.
- White, R. W. (2018). Opportunities and challenges in search interaction. *Communications of the ACM*, 61(12), 36-38.
- Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. *5th International Conference on Language Resources and Evaluation (LREC 2006)* (pp. 1556-1559). Genoa, Italy: European Language Resources Association (ELRA).
- Wobbrock, J. O., Aung, H. H., Rothrock, B., & Myers, B. A. (2005). Maximizing the guessability of symbolic input. *CHI'05 extended abstracts on Human Factors in Computing Systems* (pp. 1869-1872). Portland, Oregon, USA: ACM.
- Wu, D., Pigou, L., Kindermans, P. J., Le, N. D., Shao, L., Dambre, J., & Odobez, J. M. (2016). Deep dynamic neural networks for multimodal gesture segmentation and

- recognition. *IEEE transactions on pattern analysis and machine intelligence*, 38(8), 1583-1597.
- Yujian, L., & Bo, L. (2007). A normalized Levenshtein distance metric. *IEEE transactions on pattern analysis and machine intelligence*, 29(6), 1091-1095.
- Zhang, L., Zhang, S., Jiang, F., Qi, Y., Zhang, J., Guo, Y., & Zhou, H. (2017). BoMW: Bag of Manifold Words for One-shot Learning Gesture Recognition from Kinect. *IEEE Transactions on Circuits and Systems for Video Technology*.
- Zhang, B., Horváth, I., Molenbroek, J. F., & Snijders, C. (2010). Using artificial neural networks for human body posture prediction. *International Journal of Industrial Ergonomics*, 40(4), 414-424.
- Zhang, Z., & Zhao, H. (2018). One-shot learning for question-answering in gaokao history challenge. *CoRR (Vol. abs/1806.09105)*, 449-461.
- Zhou, Z. H. (2018). Machine learning research: some recent progress in China and beyond. *National Science Review*.
- Zhu, X. (2005). Semi-Supervised Learning Literature Survey. *Computer Sciences, University of Wisconsin-Madison, No. 1530*.
- Zhu, X., & Goldberg, A. B. (2009). Introduction to Semi – Supervised Learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning, Vol. 3, No. 1*, 1-130.
- Zou, Y., Xiao, J., Han, J., Wu, K., Li, Y., & Ni, L. M. (2017). Grfid: A device-free rfid-based gesture recognition system. *IEEE Transactions on Mobile Computing*, 16(2), 381-393.

Apéndice A. Estructura de datos para almacenamiento de información antropométrica descriptiva de un gesto

```
<?xml version="1.0" encoding="UTF-8" standalone="no" ?>
<capability>
  <motor>
    <face>
      <leftEye>
        <PointOfFixation></PointOfFixation>
        <verticalRotation></verticalRotation>
        <horizontalRotation></horizontalRotation>
        <eyeLidDistance></eyeLidDistance>
      </leftEye>
      <rightEye>
        <PointOfFixation></PointOfFixation>
        <verticalRotation></verticalRotation>
        <horizontalRotation></horizontalRotation>
        <eyeLidDistance></eyeLidDistance>
      </rightEye>
      <leftEyebrow>
        <browHeadPosition></browHeadPosition>
        <browArchPosition></browArchPosition>
        <browTailPosition></browTailPosition>
      </leftEyebrow>
      <rightEyebrow>
        <beginBrowPosition></beginBrowPosition>
        <pointOfArchPosition></pointOfArchPosition>
        <endBrowPosition></endBrowPosition>
      </rightEyebrow>
      <nose>
        <nasalTipPosition></nasalTipPosition>
        <nasalRidgeWrinkling></nasalRidgeWrinkling>
        <leftNostrilFlaring></leftNostrilFlaring>
        <rightNostrilFlaring></rightNostrilFlaring>
      </nose>
      <mouth>
        <lipDistance></lipDistance>
        <position></position>
      </mouth>
      <tongue>
        <tipPosition></tipPosition>
        <curling></curling>
        <turning></turning>
        <folding></folding>
      </tongue>
      <leftCheek>
        <inflation></inflation>
        <contraction></contraction>
      </leftCheek>
      <rightCheek>
        <inflation></inflation>
        <contraction></contraction>
      </rightCheek>
    </face>
  </motor>
</capability>
```

```
</rightCheek>
</face>
<neck>
  <flexion></flexion>
  <extension></extension>
  <rightLateralRotation></rightLateralRotation>
  <leftLateralRotation></leftLateralRotation>
  <rightLateralFlexion></rightLateralFlexion>
  <leftLateralFlexion></leftLateralFlexion>
</neck>
<leftUpperLimb>
  <shoulder>
    <flexion></flexion>
    <extension></extension>
    <adduction></adduction>
    <abduction></abduction>
    <externalRotation></externalRotation>
    <internalRotation></internalRotation>
  </shoulder>
  <elbow>
    <flexion></flexion>
    <extension></extension>
    <hyperExtension></hyperExtension>
  </elbow>
  <forearm>
    <supination></supination>
    <pronation></pronation>
  </forearm>
  <wrist>
    <flexion></flexion>
    <extension></extension>
    <radialDeviation></radialDeviation>
    <ulnarDeviation></ulnarDeviation>
  </wrist>
  <hand>
    <supination></supination>
    <pronation></pronation>
    <finger1>
      <flexionA></flexionA>
      <extensionA></extensionA>
      <adductionA></adductionA>
      <abductionA></abductionA>
      <hyperExtensionA></hyperExtensionA>
      <flexionB></flexionB>
      <extensionB></extensionB>
      <abductionB></abductionB>
      <hyperExtensionB></hyperExtensionB>
      <flexionC></flexionC>
      <extensionC></extensionC>
      <hyperExtensionC></hyperExtensionC>
    </finger1>
    <finger2>
      <flexionA></flexionA>
      <extensionA></extensionA>
      <adductionA></adductionA>
```

```
        <abductionA></abductionA>
        <hyperExtensionA></hyperExtensionA>
        <flexionB></flexionB>
        <extensionB></extensionB>
        <abductionB></abductionB>
        <hyperExtensionB></hyperExtensionB>
        <flexionC></flexionC>
        <extensionC></extensionC>
        <hyperExtensionC></hyperExtensionC>
    </finger2>
    <finger3>
        <flexionA></flexionA>
        <extensionA></extensionA>
        <adductionA></adductionA>
        <abductionA></abductionA>
        <hyperExtensionA></hyperExtensionA>
        <flexionB></flexionB>
        <extensionB></extensionB>
        <abductionB></abductionB>
        <hyperExtensionB></hyperExtensionB>
        <flexionC></flexionC>
        <extensionC></extensionC>
        <hyperExtensionC></hyperExtensionC>
    </finger3>
    <finger4>
        <flexionA></flexionA>
        <extensionA></extensionA>
        <adductionA></adductionA>
        <abductionA></abductionA>
        <hyperExtensionA></hyperExtensionA>
        <flexionB></flexionB>
        <extensionB></extensionB>
        <abductionB></abductionB>
        <hyperExtensionB></hyperExtensionB>
        <flexionC></flexionC>
        <extensionC></extensionC>
        <hyperExtensionC></hyperExtensionC>
    </finger4>
    <finger5>
        <flexionA></flexionA>
        <extensionA></extensionA>
        <adductionA></adductionA>
        <abductionA></abductionA>
        <hyperExtensionA></hyperExtensionA>
        <flexionB></flexionB>
        <extensionB></extensionB>
        <abductionB></abductionB>
        <hyperExtensionB></hyperExtensionB>
        <flexionC></flexionC>
        <extensionC></extensionC>
        <hyperExtensionC></hyperExtensionC>
    </finger5>
</hand>
</leftUpperLimb>
<rightUpperLimb>
```

```
<shoulder>
  <flexion></flexion>
  <extension></extension>
  <adduction></adduction>
  <abduction></abduction>
  <externalRotation></externalRotation>
  <internalRotation></internalRotation>
</shoulder>
<elbow>
  <flexion></flexion>
  <extension></extension>
  <hyperExtension></hyperExtension>
</elbow>
<forearm>
  <supination></supination>
  <pronation></pronation>
</forearm>
<wrist>
  <flexion></flexion>
  <extension></extension>
  <radialDeviation></radialDeviation>
  <ulnarDeviation></ulnarDeviation>
</wrist>
<hand>
  <supination></supination>
  <pronation></pronation>
  <finger1>
    <flexionA></flexionA>
    <extensionA></extensionA>
    <adductionA></adductionA>
    <abductionA></abductionA>
    <hyperExtensionA></hyperExtensionA>
    <flexionB></flexionB>
    <extensionB></extensionB>
    <hyperExtensionB></hyperExtensionB>
    <flexionC></flexionC>
    <extensionC></extensionC>
    <hyperExtensionC></hyperExtensionC>
  </finger1>
  <finger2>
    <flexionA></flexionA>
    <extensionA></extensionA>
    <adductionA></adductionA>
    <abductionA></abductionA>
    <hyperExtensionA></hyperExtensionA>
    <flexionB></flexionB>
    <extensionB></extensionB>
    <abductionB></abductionB>
    <hyperExtensionB></hyperExtensionB>
    <flexionC></flexionC>
    <extensionC></extensionC>
    <hyperExtensionC></hyperExtensionC>
  </finger2>
  <finger3>
    <flexionA></flexionA>
```

```
        <extensionA></extensionA>
        <adductionA></adductionA>
        <abductionA></abductionA>
        <hyperExtensionA></hyperExtensionA>
        <flexionB></flexionB>
        <extensionB></extensionB>
        <abductionB></abductionB>
        <hyperExtensionB></hyperExtensionB>
        <flexionC></flexionC>
        <extensionC></extensionC>
        <hyperExtensionC></hyperExtensionC>
    </finger3>
    <finger4>
        <flexionA></flexionA>
        <extensionA></extensionA>
        <adductionA></adductionA>
        <abductionA></abductionA>
        <hyperExtensionA></hyperExtensionA>
        <flexionB></flexionB>
        <extensionB></extensionB>
        <abductionB></abductionB>
        <hyperExtensionB></hyperExtensionB>
        <flexionC></flexionC>
        <extensionC></extensionC>
        <hyperExtensionC></hyperExtensionC>
    </finger4>
    <finger5>
        <flexionA></flexionA>
        <extensionA></extensionA>
        <adductionA></adductionA>
        <abductionA></abductionA>
        <hyperExtensionA></hyperExtensionA>
        <flexionB></flexionB>
        <extensionB></extensionB>
        <abductionB></abductionB>
        <hyperExtensionB></hyperExtensionB>
        <flexionC></flexionC>
        <extensionC></extensionC>
        <hyperExtensionC></hyperExtensionC>
    </finger5>
</hand>
</rightUpperLimb>
<leftLowerLimb>
    <hip>
        <flexion></flexion>
        <extension></extension>
        <adduction></adduction>
        <abduction></abduction>
        <externalRotation></externalRotation>
        <internalRotation></internalRotation>
    </hip>
    <knee>
        <flexion></flexion>
        <extension></extension>
    </knee>
```

```
<ankle>
  <dorsiFlexion></dorsiFlexion>
  <plantarFlexion></plantarFlexion>
  <eversion></eversion>
  <inversion></inversion>
  <pronation></pronation>
  <supination></supination>
  <lateralRotation></lateralRotation>
  <medialRotation></medialRotation>
</ankle>
<foot>
  <footToe1>
    <flexion></flexion>
    <extension></extension>
  </footToe1>
  <footToe2>
    <flexion></flexion>
    <extension></extension>
  </footToe2>
  <footToe3>
    <flexion></flexion>
    <extension></extension>
  </footToe3>
  <footToe4>
    <flexion></flexion>
    <extension></extension>
  </footToe4>
  <footToe5>
    <flexion></flexion>
    <extension></extension>
  </footToe5>
</foot>
</leftLowerLimb>
<rightLowerLimb>
  <hip>
    <flexion></flexion>
    <extension></extension>
    <adduction></adduction>
    <abduction></abduction>
    <externalRotation></externalRotation>
    <internalRotation></internalRotation>
  </hip>
  <knee>
    <flexion></flexion>
    <extension></extension>
  </knee>
  <ankle>
    <dorsiFlexion></dorsiFlexion>
    <plantarFlexion></plantarFlexion>
    <eversion></eversion>
    <inversion></inversion>
    <pronation></pronation>
    <supination></supination>
    <lateralRotation></lateralRotation>
    <medialRotation></medialRotation>
```

```
</ankle>
<foot>
  <footToe1>
    <flexion></flexion>
    <extension></extension>
  </footToe1>
  <footToe2>
    <flexion></flexion>
    <extension></extension>
  </footToe2>
  <footToe3>
    <flexion></flexion>
    <extension></extension>
  </footToe3>
  <footToe4>
    <flexion></flexion>
    <extension></extension>
  </footToe4>
  <footToe5>
    <flexion></flexion>
    <extension></extension>
  </footToe5>
</foot>
</rightLowerLimb>
</motor>
</capability>
```

Apéndice B. Reporte de caracterización de gestos del estudio sobre identificación de un vocabulario gestual para interacción con múltiples aplicaciones

Etiqueta (tasa de acuerdo)	Descripción textual	Caracterización de acuerdo con el modelo de capacidad asociado al GVUM
G001 (5/14)	Inclinar la cabeza a la derecha.	Flexión lateral derecho del cuello.
G002 (3/14)	Girar la cabeza a la derecha.	Rotación lateral derecho del cuello.
G003 (3/14)	Inclinar la cabeza a la izquierda.	Flexión lateral izquierda del cuello.
G004 (5/14)	Levantar ambos hombros.	Flexión de ambos hombros.
G005 (2/14)	Levantar ambos hombros dos veces.	Flexión de ambos hombros dos veces.
G006 (3/14)	Inclinar la cabeza a la derecha dos veces.	Flexión lateral derecha del cuello dos veces.
G007 (3/14)	Mirar hacia abajo.	Flexión del cuello.
G008 (4/14)	Asentir dos veces.	Flexión y extensión del cuello dos veces.
G009 (4/14)	Mirar hacia arriba.	Extensión del cuello.
G010 (3/14)	Girar la cabeza a la izquierda.	Rotación lateral izquierda del cuello.
G011 (2/14)	Levantar el hombro izquierdo.	Flexión del hombro izquierdo.
G012 (2/14)	Mover ambos hombros hacia delante.	Aducción de ambos hombros.
G013 (2/14)	Levantar el hombro izquierdo dos veces.	Flexión del hombro izquierdo dos veces.
G014 (2/14)	Levantar el hombro izquierdo e inclinar la cabeza a la izquierda.	Flexión del hombro izquierdo y flexión lateral izquierda del cuello.
G015 (6/14)	Mover el brazo izquierdo a la izquierda y el brazo derecho a la derecha.	Extensión de ambos hombros, extensión de ambos codos.
G016 (6/14)	Iniciando con ambos brazos extendidos, llevarlos hacia el centro.	Iniciando con la extensión de ambos hombros y ambos codos, flexionar ambos hombros.
G017 (6/14)	Levantar el brazo derecho.	Abducción del hombro derecho, extensión del codo izquierdo.
G018 (4/14)	Bajar el brazo derecho.	Aducción del hombro derecho, extensión del codo derecho.
G019 (1/14)	Mover la muñeca derecha hacia la derecha.	Flexión del codo derecho y rotación interna del hombro derecho.
G020 (1/14)	Iniciando con el brazo derecho extendido, mover la muñeca derecha hacia delante.	Iniciando con el codo derecho extendido, flexionar el codo derecho.
G021 (4/14)	Mover la muñeca izquierda a la izquierda.	Flexión del codo izquierdo, rotación interna del hombro izquierdo.
G022 (4/14)	Iniciando con el brazo derecho flexionado, mover la muñeca derecha hacia delante.	Iniciando con el codo izquierdo flexionado, rotar el hombro derecho.
G023 (6/14)	Levantar el brazo izquierdo.	Abducción del hombro izquierdo, extensión del codo izquierdo.
G024 (6/14)	Bajar el brazo izquierdo.	Aducción del hombro derecho, extensión del codo derecho.
G025 (4/14)	Iniciando con el codo izquierdo flexionado, rotar el hombro izquierdo.	Iniciando con el codo flexionado, rotar el hombro izquierdo.

G026 (4/14)	Iniciando con el codo derecho flexionado, rotar el hombro derecho.	Iniciando con el codo derecho flexionado, rotar el hombro derecho.
G027 (4/14)	Iniciando con el brazo flexionado y la muñeca apuntando hacia arriba, bajar el codo.	Iniciando con el codo flexionado, aducir el hombro.
G028 (4/14)	Iniciando con el brazo flexionado y la muñeca apuntando hacia arriba, mover la muñeca para apuntar hacia abajo.	Iniciando con el codo flexionado, flexionar el hombro.
G029 (5/14)	Apuntar hacia delante usando el dedo índice.	Flexión de todos los dedos excepto el segundo.
G030 (2/14)	Girar la muñeca derecha de la derecha a la izquierda con los dedos índice y medio extendidos.	Flexión de todos los dedos excepto el segundo y el tercero, flexión de la muñeca.
G031 (1/14)	Apuntando hacia arriba con el dedo índice solo, flexionar la punta del dedo dos veces.	Iniciando con todos los dedos flexionados excepto el segundo, flexionar el segundo dedo dos veces.
G032 (2/14)	Usando ambas manos, alcanzar el dedo índice de cada una con el pulgar de la otra creando una forma de 'marco'.	Extensión del primer y segundo dedo de cada mano, alcanzando el segundo dedo de la mano izquierda con el primero de la mano derecha y el segundo dedo de la mano derecha con el primero de la izquierda.
G033 (13/14)	Usar solo el pulgar para apuntar hacia arriba.	Flexionar todos los dedos y extender el primero.
G034 (8/14)	Juntar cada dedo de una mano con los de la otra para formar una figura similar a un corazón.	Flexionar todos los dedos y alcanzar cada uno de los dedos de una mano con el mismo de la otra.
G035 (3/14)	Usar ambas manos para 'teclear' en el aire.	Iniciando con todos los dedos extendidos, rápidamente flexionarlos y extenderlos en múltiples ocasiones.
G036 (4/14)	Usar los dedos índices de ambas manos para crear un 'signo de suma'.	Extendiendo solo el segundo dedo de cada mano, usar uno de ellos para apuntar hacia arriba y poner el otro de manera perpendicular para formar un 'signo de suma'.
G037 (2/14)	Con el pulgar y el índice extendidos, mover el dedo índice hacia abajo para alcanzar el pulgar sin llegar a tocarlo.	Iniciando solo con el primer y segundo dedo extendido, flexionar ligeramente el segundo dedo hacia el primero sin llegar a alcanzarlo.
G038 (2/14)	Con la palma de la mano hacia Adelante, tocar el pulgar con el meñique.	Iniciando con todos los dedos extendidos, flexionar el primero y el quinto para que puedan tocarse.
G039 (7/14)	Usando los índices de ambas manos, hacer un 'símbolo de x'.	Extendiendo solo el segundo dedo e cada mano, usar uno de ellos para señalar arriba y colocar el otro perpendicularmente para hacer un signo de 'x'.
G040 (1/14)	Usando solo el dedo índice, dibujar en el aire un triángulo.	Extender solo el Segundo dedo, dibujar un triángulo en el aire ya sea rotando el codo o el hombro.
G041 (2/14)	Usando solo el dedo índice, dibujar un círculo en el aire.	Extendiendo solo el Segundo dedo, dibujar un círculo en el aire ya sea rotando el codo o el hombro.
G042 (9/14)	Flexionar todos los dedos excepto el meñique y el pulgar.	Flexionar todos los dedos y extender solo el primero y el quinto.
G043 (3/14)	Deslizar el pie hacia atrás.	Extender la cadera sobre la pierna izquierda.
G044 (6/14)	Deslizar el pie hacia adelante.	Flexionar la cadera sobre la pierna derecha.
G045 (5/14)	Pisar fuertemente.	Flexionar una rodilla, luego, rápidamente extenderla.
G046 (4/14)	Dar un paso adelante.	Flexionar la cadera, extender la rodilla, extender la cadera.

G047 (3/14)	Deslizar el pie izquierdo hacia la izquierda.	Abducción de la cadera sobre la pierna izquierda.
G048 (5/14)	Dar una patada.	Flexionar la cadera, flexionar una rodilla, luego extenderla.
G049 (5/14)	Brincar usando ambas piernas.	Flexionar la cadera, flexionar las rodillas, flexionar los tobillos, luego extender los tobillos y las rodillas lanzando el cuerpo hacia arriba.
G050 (3/14)	Levantar una pierna y después pisar fuertemente.	Flexionar la cadera, flexionar una rodilla, luego rápidamente extender la cadera y la rodilla.
G051 (3/14)	Caminar hacia adelante.	Repetir con cada extremidad inferior: Flexionar la cadera, extender la rodilla, extender la cadera.
G052 (7/14)	Rotar el tobillo izquierdo.	Flexión plantar, aducción, dorsiflexión, abducción de un tobillo.
G053 (5/14)	Saltar usando ambas piernas dos veces.	Flexionar la cadera, flexionar las rodillas, flexionar los tobillos, luego extender los tobillos y las rodillas lanzando el cuerpo hacia arriba dos veces.
G054 (5/14)	Agacharse.	Flexionar la cadera, flexionar las rodillas, flexión plantar de los tobillos.
G055 (3/14)	Deslizar el pie izquierdo hacia la izquierda.	Aducción de la cadera sobre la pierna izquierda.
G056 (4/14)	Deslizar el pie derecho hacia la derecha.	Aducción de la cadera sobre la pierna derecha.
G057 (6/14)	Pararse de puntas.	Flexión plantar.
G058 (4/14)	Pararse de puntas y caminar hacia adelante.	Flexión plantar, extender la cadera.
G059 (3/14)	Pararse de puntas / levantar un pie.	Flexión plantar, flexión de la rodilla / extender la rodilla.
G060 (6/14)	Girar hacia arriba el pie izquierdo.	Eversión del pie izquierdo.
G061 (4/14)	Hacer un símbolo de 'x' con un pie.	Flexión plantar, dibujar un signo de 'x' en el aire usando la punta del pie.
G062 (4/14)	Girar ambos pies hacia adentro.	Eversión de ambos pies.
G063 (5/14)	Separar los pies.	Aducción en ambos tobillos.
G064 (7/14)	Rotar el pie izquierdo hacia afuera.	Inversión del pie izquierdo.
G065 (5/14)	Mover el pie izquierdo hacia adelante.	Extensión de la rodilla izquierda.
G066 (8/14)	Rotar el pie izquierdo hacia la izquierda.	Eversión, dorsiflexión, inversión, flexión plantar.
G067 (7/14)	Rotar el pie derecho hacia la izquierda.	Inversión, dorsiflexión, eversión, flexión plantar.
G068 (5/14)	Separar los dedos del pie.	Abducción de todos los dedos del pie.

Apéndice C. Reporte de caracterización de gestos del estudio sobre identificación de un vocabulario gestual para interacción con tareas abstractas

Gesto	Tareas asociadas	Descripción	Caracterización
E001	1, 2, 3, 6, 7, 8, 11, 13, 17, 21, 22	Mover lengua hacia arriba	Mover lengua hacia arriba
E002	1, 2, 3, 6, 7, 11, 13, 18, 21, 22	Mover lengua hacia abajo	Mover lengua hacia abajo
E003	1, 2, 6, 7, 11, 13, 19, 21, 22	Mover lengua a la izquierda	Mover lengua a la izquierda
E004	1, 2, 6, 7, 11, 13, 20, 21, 22	Mover lengua a la derecha	Mover lengua a la derecha
E005	1, 2, 6, 7, 11, 13, 21, 22	Mantener lengua afuera	Mantener lengua afuera
E006	2	Guiñar ojo derecho	Minimizar la distancia entre los párpados
E007	4, 12	Guiñar dos veces con el ojo derecho.	Minimizar la distancia entre los párpados dos veces
E008	5, 11	Inflar la mejilla derecha	Inflar la mejilla derecha
E009	6, 7, 13	Abrir la boca	Incrementar la distancia entre los labios
E010	8	Inflar y desinflar la mejilla derecha	Inflar mejilla derecha, luego contraer mejilla derecha
E011	9	Inflar y desinflar la mejilla izquierda	Inflar mejilla derecha, luego contraer mejilla izquierda
E012	10, 14	Sonreír	-
E013	11	Inflar la mejilla izquierda	Inflar la mejilla izquierda
E014	15	Cerrar la boca ocultando los labios	Minimizar la distancia entre los labios
E015	16	Fruncir el ceño	-
E016	23	Lanzar un beso	-
N001	1, 6	Levantar el brazo derecho hacia el frente	Flexionar el hombro derecho
N002	2, 5	Levantar el hombro derecho dos veces	Mover el hombro derecho hacia arriba dos veces
N003	2	Levantar el hombro izquierdo dos veces	Mover el hombro izquierdo hacia arriba dos veces
N004	4	Asentir dos veces	Repetir dos veces: Flexión del cuello, luego extensión del cuello
N005	7	Rotar el cuello a la derecha	Rotación lateral derecha del cuello
N006	8	Levantar el hombro derecho	Mover hombro derecho hacia arriba
N007	8	Bajar el hombro izquierdo	Mover el hombro izquierdo hacia abajo
N008	9	Levantar el hombro izquierdo	Mover hombro izquierdo hacia arriba
N009	9	Bajar el hombro derecho	Mover el hombro derecho hacia abajo
N010	10, 21	Inclinar la cabeza hacia atrás dos veces	Dos veces: extensión del cuello
N011	11, 16, 23	Mover el hombro derecho hacia atrás	Mover hombro derecho hacia atrás

N012	11, 16, 23	Mover el hombro izquierdo hacia atrás	Mover hombro izquierdo hacia atrás
N013	12	Asentir	Flexión de cuello, luego extensión de cuello
N014	13	Levantar brazo derecho hacia el frente dos veces	Dos veces: flexión del hombro derecho
N015	14	Inclinar la cabeza a la derecha dos veces	Dos veces: flexión lateral derecha del cuello
N016	15	Mover la cabeza de izquierda a derecha dos veces	Dos veces: flexión lateral izquierda, luego flexión lateral derecha
N017	1, 2, 3, 6, 7, 17, 21, 22	Mover la cabeza hacia atrás	Extensión de cuello
N018	1, 2, 3, 6, 7, 18, 21, 22	Mover la cabeza hacia adelante	Flexión de cuello
N019	1, 2, 6, 7, 19, 21, 22	Mover la cabeza a la izquierda	Flexión lateral izquierda de cuello
N020	1, 2, 6, 7, 20, 21, 22	Mover la cabeza a la derecha	Flexión lateral derecha de cuello
N021	22	Mover la cabeza de izquierda a derecha	Flexión lateral izquierda de cuello, luego flexión lateral derecha de cuello
A001	1, 4, 12, 21	Flexionar el codo derecho	Flexión de codo derecho
A002	1, 2, 3, 4, 7, 8, 9, 10, 12, 14, 15, 16, 17, 18, 19, 20, 21, 23	Extender el brazo derecho al frente	Flexión del hombro derecho y extensión del codo derecho
A003	2, 16	Usando el brazo derecho, dibujar un círculo en el aire hacia la derecha	Rotación externa del hombro
A004	3	Mover brazo derecho hacia arriba	Flexión del hombro derecho
A005	3, 15	Mover brazo derecho hacia abajo	Extensión del hombro derecho
A006	5	Extender el brazo derecho al frente flexionando el codo	Flexión de hombro derecho y flexión de codo derecho
A007	5	Rotar brazo derecho a la izquierda	Rotación interna del hombro derecho
A008	13, 22	Extender brazo derecho a la derecha flexionando el codo	Abducción de hombro derecho y flexión de codo derecho
A009	6, 13	Rotar el brazo derecho al frente	Rotación interna del hombro derecho
A010	7	Rotar la muñeca derecha al frente dos veces	Dos veces: Flexión de la muñeca derecha
A011	8, 14	Rotar la muñeca derecha a la derecha dos veces	Dos veces: Rotación hacia la derecha de la muñeca derecha
A012	9	Rotar la muñeca derecha a la izquierda dos veces	Dos veces: Rotación hacia la izquierda de la muñeca derecha
A013	10	Flexionar y extender el codo derecho dos veces	Dos veces: flexión de codo derecho, luego extensión de codo derecho
A014	11	Poner ambos brazos juntos formando una 'X'	Flexión de hombro derecho, flexión de hombro izquierdo, flexión de codo derecho, flexión de codo izquierdo
A015	15, 18, 21	Flexionar la muñeca derecha hacia abajo	Flexión de muñeca derecha
A016	17	Flexionar la muñeca derecha hacia abajo	Extensión de muñeca derecha
A017	19	Flexionar la muñeca derecha hacia la izquierda	Desviación ulnar de la muñeca derecha
A018	20	Flexionar la muñeca derecha hacia la derecha	Desviación radial de la muñeca derecha

A019	22	Rotar el brazo izquierdo hacia el frente	Rotación interna del hombro izquierdo
A020	23	Extender el brazo izquierdo al frente	Flexión del hombro izquierdo y extensión del codo izquierdo
A021	23	Rotar brazo derecho a la derecha	Rotación externa de hombro derecho
A022	23	Rotar el brazo izquierdo a la izquierda	Rotación externa del hombro izquierdo
H001	1, 11, 12, 13, 15, 22	Abrir la mano derecha con la palma hacia el frente	Extensión de la muñeca derecha y extensión de todas las articulaciones de los dedos de la mano derecha
H002	1	Juntar todas las puntas de los dedos de la mano derecha	Flexionar las articulaciones A de todos los dedos de la mano derecha
H003	2	Hacer símbolo de 'OK' hacia el frente con la mano derecha	Con la mano derecha, extensión de todas las articulaciones de los dedos 1, 2 y 3, y flexión de todas las articulaciones de los dedos 4 y 5 haciendo que sus puntas se toquen
H004	3, 17	Abrir la mano derecha con la palma hacia arriba	Extensión de todas las articulaciones de los dedos de la mano derecha, supinación del antebrazo derecho
H005	3, 18	Abrir la mano derecha con la palma hacia abajo	Extensión de todas las articulaciones de la mano derecha, pronación del antebrazo derecho
H006	3, 17	Mover la mano derecha hacia arriba	Mover la mano derecha hacia arriba
H007	3, 18	Mover la mano derecha hacia abajo	Mover la mano derecha hacia abajo
H008	4	Cerrar el puño izquierdo dejando el dedo índice y el dedo medio juntos y extendidos	Flexión de todas las articulaciones de los dedos 1, 2 5 5 de la mano izquierda, y extensión de todas las articulaciones de los dedos 3 y 4 de la mano izquierda
H009	5, 19	Con la palma derecha abierta, mover la mano derecha hacia la izquierda	Extensión de todas las articulaciones de los dedos de la mano derecha, supinación/pronación del antebrazo derecho y mover la mano derecha hacia la izquierda
H010	6	Usando la mano derecha, extender el dedo índice al frente	Flexionar todas las articulaciones de los dedos 1, 2, 3 y 5 de la mano derecha, extensión de todas las articulaciones del dedo 4
H011	6	Usando la mano izquierda, extender el dedo índice al frente	Flexionar todas las articulaciones de los dedos 1, 2, 3 y 5 de la mano izquierda, extensión de todas las articulaciones del dedo 4
H012	7	Con la mano derecha, hacer una 'L'	Flexionar todas las articulaciones de los dedos 1, 2 y 3 de la mano derecha, y extensión de todas las articulaciones de los dedos 4 y 5 de la mano derecha
H013	8	Cerrar el puño derecho dejando los dedos índice y pulgar extendidos, tocar dos veces la punta del dedo pulgar usando la punta del índice	Flexión de todas las articulaciones de los dedos 1, 2 y 3 de la mano derecha, y extensión de todas las articulaciones de los dedos 4 y 5 de la mano derecha. Luego, flexión de las articulaciones de los dedos 4 y 5 haciendo que sus puntas se toquen dos veces.
H014	9, 21	Cerrar puño derecho dejando el dedo índice y el medio juntos y extendidos	Flexión de todas las articulaciones de los dedos 1, 2 y 5 de la mano derecha, extensión de todas

			las articulaciones de los dedos 3 y 4 de la mano derecha
H015	10	Usando la mano derecha, hacer una señal de pulgar arriba	Flexión de todas las articulaciones de los dedos 1, 2, 3 y 4 de la mano derecha, extensión de todas las articulaciones del dedo 5 de la mano derecha
H016	11, 13, 15	Cerrar el puño derecho	Flexión de todas las articulaciones de los dedos de la mano derecha.
H017	14, 20	Con la palma derecha abierta, mover la mano a la derecha	Extensión de todas las articulaciones de los dedos de la mano derecha, supinación/pronación del antebrazo derecho y mover la mano a la derecha
H018	16	Cerrar el puño derecho con la palma hacia el frente	Flexión de todas las articulaciones de los dedos de la mano derecha y extensión de la muñeca derecha
H019	16	Mover la mano derecha hacia abajo	Mover la mano derecha hacia abajo
H020	22	Bajar las puntas de los dedos de la mano derecha	Flexionar las articulaciones A y B de todos los dedos de la mano derecha
H021	23	Abrir la mano izquierda con la palma hacia el frente	Extensión de la muñeca izquierda y extensión de todas las articulaciones de los dedos de la mano izquierda
H022	23	Mover la mano derecha hacia la derecha	Mover la mano derecha hacia la derecha
H023	23	Mover la mano izquierda hacia la izquierda	Mover la mano izquierda hacia la izquierda
L001	1, 6	Flexionar la rodilla derecha hacia atrás	Extensión de la cadera derecha y flexión de la rodilla derecha
L002	2, 10	Flexionar la rodilla derecha hacia el frente	Flexión de la cadera derecha y flexión de la rodilla derecha
L003	2	Rotar la rodilla derecha a la izquierda	Rotación interna de la cadera derecha
L004	3, 13, 14	Flexionar la rodilla izquierda hacia el frente	Flexión de la cadera izquierda y flexión de la rodilla izquierda
L005	3	Levantar la rodilla izquierda	Flexionar la cadera izquierda
L006	3	Bajar la rodilla izquierda	Extender la cadera izquierda
L007	4	Pisar dos veces con el pie izquierdo	Flexión de la cadera izquierda, extensión de la rodilla izquierda y dos veces flexión plantar del tobillo izquierdo
L008	5, 16, 19, 22	Mover la pierna izquierda hacia la izquierda	Abducción de la cadera izquierda y extensión de la rodilla izquierda
L009	7	Flexionar la rodilla izquierda hacia atrás	Extensión de la cadera izquierda y flexión de la rodilla izquierda
L010	8	Dibujar un círculo hacia la derecha en el piso usando la pierna derecha extendida	Rotación externa de la cadera derecha
L011	9	Dibujar un círculo hacia la izquierda en el piso usando la pierna izquierda extendida	Rotación externa de la cadera izquierda
L012	10	Rotar la rodilla izquierda hacia la derecha	Rotación interna de la cadera izquierda

L013	11, 20	Mover la pierna derecha hacia la derecha	Abducción de la cadera derecha y extensión de la rodilla derecha
L014	11	Mover la pierna derecha hacia la izquierda	Aducción de la cadera derecha y extensión de la rodilla derecha
L015	12	Levantar la punta del pie derecho	Dorsiflexión del tobillo derecho
L016	12	Levantar la punta del pie izquierdo	Dorsiflexión del tobillo izquierdo
L017	12	Bajar la punta del pie derecho	Flexión plantar del tobillo derecho
L018	12	Bajar la punta del pie izquierdo	Flexión plantar del tobillo izquierdo
L019	13	Rotar la rodilla izquierda hacia la izquierda	Rotación externa de la cadera izquierda
L020	15	Mover la punta del pie izquierdo hacia la derecha	Inversión del tobillo izquierdo
L021	15	Mover la punta del pie derecho hacia la izquierda	Inversión del tobillo derecho
L022	16, 22	Extender la pierna izquierda hacia la derecha	Aducción de la cadera izquierda y extensión de la rodilla izquierda
L023	17	Mover la pierna izquierda hacia el frente	Flexión de la cadera izquierda y extensión de la rodilla izquierda
L024	18	Mover la pierna izquierda hacia atrás	Extensión de la cadera izquierda y extensión de la rodilla izquierda
L025	21	Pisar dos veces usando el pie derecho	Flexión de la cadera izquierda, extensión de la rodilla derecha y flexión plantar del tobillo derecho dos veces.
L026	23	Sentado, mover la pierna derecha hacia la derecha	Sentado (flexión de la cadera derecha, flexión de la rodilla derecho, flexión de la cadera izquierda, flexión de la rodilla izquierda), abducción de la cadera derecha
L027	23	Sentado, mover la pierna izquierda hacia la izquierda	Sentado (flexión de la cadera derecha, flexión de la rodilla derecho, flexión de la cadera izquierda, flexión de la rodilla izquierda), abducción de la cadera izquierda
F001	1, 2, 6, 7, 17, 21, 22	Mover el pie derecho al frente	Mover el pie derecho al frente
F002	1, 2, 6, 7, 18, 21, 22	Mover el pie derecho hacia atrás	Mover el pie derecho hacia atrás
F003	1, 2, 6, 7, 9, 19, 21, 22	Mover el pie derecho hacia la izquierda	Mover el pie derecho hacia la izquierda
F004	1, 2, 6, 7, 8, 14, 20, 21, 22, 23	Mover el pie derecho hacia la derecha	Mover el pie derecho hacia la derecha
F005	23	Mover el pie izquierdo hacia la izquierda	Mover el pie izquierdo hacia la izquierda
F006	1, 2, 6, 7, 21, 22	Pisar dos veces	Dos veces: flexión plantar del tobillo derecho
F007	3, 14	Mover la punta del pie derecho hacia arriba	Dorsiflexión del tobillo derecho
F008	3	Mover la punta del pie derecho hacia abajo	Flexión plantar del tobillo derecho
F009	4, 12, 16	Mover la punta del pie derecho hacia arriba dos veces	Dos veces: Dorsiflexión del tobillo derecho
F010	5, 13	Inclinar el pie derecho hacia la derecha	Supinación del tobillo derecho
F011	6, 7	Inclinar el pie derecho hacia la izquierda	Pronación del pie derecho

F012	10	Mover el dedo gordo del pie derecho hacia arriba	Extensión del primer dedo del pie derecho
F013	11	Mover la punta del pie derecho hacia la izquierda	Inversión del tobillo izquierdo
F014	11	Mover la punta del pie izquierdo hacia la derecha	Inversión del tobillo izquierdo
F015	15	Flexionar todos los dedos del pie derecho	Flexión de las articulaciones de todos los dedos del pie derecho

Apéndice D. Fragmentos de los archivos JSON y ARF generados para el experimento reportado en la sección 5.4 sobre entrenamiento de un clasificador gestual siguiendo un enfoque de aprendizaje one-shot

Fragmento del archivo JSON

```
Gesture = [  
  {"HeadPositionX": 0.04842854, "HeadPositionY": 0.5392805, "HeadPositionZ": 1.510944,  
"HeadLeftAngle": 127.135871844375, "HeadRightAngle": 233.878314649505, "LeftShoulderPositionX": -  
0.1070735, "LeftShoulderPositionY": 0.4016083, "LeftShoulderPositionZ": 1.449027,  
"ShoulderLeftAngle": 152.480854889827, "LeftElbowPositionX": -0.284839, "LeftElbowPositionY":  
0.294411, "LeftElbowPositionZ": 1.481576, "ElbowLeftAngle": 127.898005481302, "LeftWristPositionX": -  
0.4771417, "LeftWristPositionY": 0.339039, "LeftWristPositionZ": 1.410418, "WristLeftAngle":  
168.611456796049, "RightShoulderPositionX": 0.182371, "RightShoulderPositionY": 0.406402,  
"RightShoulderPositionZ": 1.427629, "ShoulderRightAngle": 216.454778898308, "RightElbowPositionX":  
0.3963644, "RightElbowPositionY": 0.3044159, "RightElbowPositionZ": 1.458361, "ElbowLeftAngle":  
229.770919470685, "RightWristPositionX": 0.5844809, "RightWristPositionY": 0.3619376,  
"RightWristPositionZ": 1.388317, "WristLeftAngle": 194.70084298245, "LeftHipPositionX": -0.03172928,  
"LeftHipPositionY": -0.05449117, "LeftHipPositionZ": 1.417431, "HipLeftAngle": 271.341192314241,  
"LeftKneePositionX": -0.03345026, "LeftKneePositionY": -0.3775761, "LeftKneePositionZ": 1.44474,  
"KneeLeftAngle": 171.318733193601, "LeftAnklePositionX": -0.03877562, "LeftAnklePositionY": -  
0.6719708, "LeftAnklePositionZ": 1.425085, "AnkleLeftAngle": 105.319282260664, "RightHipPositionX":  
0.120883, "RightHipPositionY": -0.05103945, "RightHipPositionZ": 1.413195, "HipRightAngle":  
87.9088655249611, "RightKneePositionX": 0.1297719, "RightKneePositionY": -0.3630721,  
"RightKneePositionZ": 1.431291, "KneeRightAngle": 172.690240386478, "RightAnklePositionX": 0.1166367,  
"RightAnklePositionY": -0.6669061, "RightAnklePositionZ": 1.416788, "AnkleRightAngle":  
105.040105726055, "GestureLabel": "G01"},  
  {"HeadPositionX": 0.0384285354614258, "HeadPositionY": 0.5392805, "HeadPositionZ": 1.510944,  
"HeadLeftAngle": 127.135871844375, "HeadRightAngle": 233.878314649505, "LeftShoulderPositionX": -  
0.117073470950127, "LeftShoulderPositionY": 0.4016083, "LeftShoulderPositionZ": 1.449027,  
"ShoulderLeftAngle": 152.480854889827, "LeftElbowPositionX": -0.294838974475861,  
"LeftElbowPositionY": 0.294411, "LeftElbowPositionZ": 1.481576, "ElbowLeftAngle": 127.898005481302,  
"LeftWristPositionX": -0.487141708135605, "LeftWristPositionY": 0.339039, "LeftWristPositionZ":  
1.410418, "WristLeftAngle": 168.611456796049, "RightShoulderPositionX": 0.172371005415916,  
"RightShoulderPositionY": 0.406402, "RightShoulderPositionZ": 1.427629, "ShoulderRightAngle":  
216.454778898308, "RightElbowPositionX": 0.38636442065239, "RightElbowPositionY": 0.3044159,  
"RightElbowPositionZ": 1.458361, "ElbowLeftAngle": 229.770919470685, "RightWristPositionX":  
0.574480941295624, "RightWristPositionY": 0.3619376, "RightWristPositionZ": 1.388317,  
"WristLeftAngle": 194.70084298245, "LeftHipPositionX": -0.041729272233486, "LeftHipPositionY": -  
0.05449117, "LeftHipPositionZ": 1.417431, "HipLeftAngle": 271.341192314241, "LeftKneePositionX": -  
0.0434502644836903, "LeftKneePositionY": -0.3775761, "LeftKneePositionZ": 1.44474, "KneeLeftAngle":
```

171.318733193601, "LeftAnklePositionX": -0.048775622844696, "LeftAnklePositionY": -0.6719708, "LeftAnklePositionZ": 1.425085, "AnkleLeftAngle": 105.319282260664, "RightHipPositionX": 0.110883010327816, "RightHipPositionY": -0.05103945, "RightHipPositionZ": 1.413195, "HipRightAngle": 87.9088655249611, "RightKneePositionX": 0.119771932959557, "RightKneePositionY": -0.3630721, "RightKneePositionZ": 1.431291, "KneeRightAngle": 172.690240386478, "RightAnklePositionX": 0.106636678576469, "RightAnklePositionY": -0.6669061, "RightAnklePositionZ": 1.416788, "AnkleRightAngle": 105.040105726055, "GestureLabel": "G01"},

{ "HeadPositionX": 0.0584285354614258, "HeadPositionY": 0.5392805, "HeadPositionZ": 1.510944, "HeadLeftAngle": 127.135871844375, "HeadRightAngle": 233.878314649505, "LeftShoulderPositionX": -0.0970734709501267, "LeftShoulderPositionY": 0.4016083, "LeftShoulderPositionZ": 1.449027, "ShoulderLeftAngle": 152.480854889827, "LeftElbowPositionX": -0.274838974475861, "LeftElbowPositionY": 0.294411, "LeftElbowPositionZ": 1.481576, "ElbowLeftAngle": 127.898005481302, "LeftWristPositionX": -0.467141708135605, "LeftWristPositionY": 0.339039, "LeftWristPositionZ": 1.410418, "WristLeftAngle": 168.611456796049, "RightShoulderPositionX": 0.192371005415916, "RightShoulderPositionY": 0.406402, "RightShoulderPositionZ": 1.427629, "ShoulderRightAngle": 216.454778898308, "RightElbowPositionX": 0.40636442065239, "RightElbowPositionY": 0.3044159, "RightElbowPositionZ": 1.458361, "ElbowLeftAngle": 229.770919470685, "RightWristPositionX": 0.594480941295624, "RightWristPositionY": 0.3619376, "RightWristPositionZ": 1.388317, "WristLeftAngle": 194.70084298245, "LeftHipPositionX": -0.0217292772233486, "LeftHipPositionY": -0.05449117, "LeftHipPositionZ": 1.417431, "HipLeftAngle": 271.341192314241, "LeftKneePositionX": -0.0234502644836903, "LeftKneePositionY": -0.3775761, "LeftKneePositionZ": 1.44474, "KneeLeftAngle": 171.318733193601, "LeftAnklePositionX": -0.028775622844696, "LeftAnklePositionY": -0.6719708, "LeftAnklePositionZ": 1.425085, "AnkleLeftAngle": 105.319282260664, "RightHipPositionX": 0.130883010327816, "RightHipPositionY": -0.05103945, "RightHipPositionZ": 1.413195, "HipRightAngle": 87.9088655249611, "RightKneePositionX": 0.139771932959557, "RightKneePositionY": -0.3630721, "RightKneePositionZ": 1.431291, "KneeRightAngle": 172.690240386478, "RightAnklePositionX": 0.126636678576469, "RightAnklePositionY": -0.6669061, "RightAnklePositionZ": 1.416788, "AnkleRightAngle": 105.040105726055, "GestureLabel": "G01"},

{ "HeadPositionX": 0.0284285354614258, "HeadPositionY": 0.5392805, "HeadPositionZ": 1.510944, "HeadLeftAngle": 127.135871844375, "HeadRightAngle": 233.878314649505, "LeftShoulderPositionX": -0.127073470950127, "LeftShoulderPositionY": 0.4016083, "LeftShoulderPositionZ": 1.449027, "ShoulderLeftAngle": 152.480854889827, "LeftElbowPositionX": -0.304838974475861, "LeftElbowPositionY": 0.294411, "LeftElbowPositionZ": 1.481576, "ElbowLeftAngle": 127.898005481302, "LeftWristPositionX": -0.497141708135605, "LeftWristPositionY": 0.339039, "LeftWristPositionZ": 1.410418, "WristLeftAngle": 168.611456796049, "RightShoulderPositionX": 0.162371005415916, "RightShoulderPositionY": 0.406402, "RightShoulderPositionZ": 1.427629, "ShoulderRightAngle": 216.454778898308, "RightElbowPositionX": 0.37636442065239, "RightElbowPositionY": 0.3044159, "RightElbowPositionZ": 1.458361, "ElbowLeftAngle": 229.770919470685, "RightWristPositionX": 0.564480941295624, "RightWristPositionY": 0.3619376, "RightWristPositionZ": 1.388317, "WristLeftAngle": 194.70084298245, "LeftHipPositionX": -0.0517292772233486, "LeftHipPositionY": -0.05449117, "LeftHipPositionZ": 1.417431, "HipLeftAngle": 271.341192314241, "LeftKneePositionX": -0.0534502644836903, "LeftKneePositionY": -0.3775761, "LeftKneePositionZ": 1.44474, "KneeLeftAngle": 171.318733193601, "LeftAnklePositionX": -0.058775622844696, "LeftAnklePositionY": -0.6719708, "LeftAnklePositionZ": 1.425085, "AnkleLeftAngle": 105.319282260664, "RightHipPositionX": 0.100883010327816, "RightHipPositionY": -0.05103945, "RightHipPositionZ": 1.413195, "HipRightAngle": 87.9088655249611, "RightKneePositionX": 0.109771932959557, "RightKneePositionY": -0.3630721,

"RightKneePositionZ": 1.431291, "KneeRightAngle": 172.690240386478, "RightAnklePositionX":
0.0966366785764694, "RightAnklePositionY": -0.6669061, "RightAnklePositionZ": 1.416788,
"AnkleRightAngle": 105.040105726055, "GestureLabel": "G01"}];

Fragmento del archivo ARFF

```
@RELATION gestures
@ATTRIBUTE HeadPositionX NUMERIC
@ATTRIBUTE HeadPositionY NUMERIC
@ATTRIBUTE HeadPositionZ NUMERIC
@ATTRIBUTE HeadLeftAngle NUMERIC
@ATTRIBUTE HeadRightAngle NUMERIC
@ATTRIBUTE LeftShoulderPositionX NUMERIC
@ATTRIBUTE LeftShoulderPositionY NUMERIC
@ATTRIBUTE LeftShoulderPositionZ NUMERIC
@ATTRIBUTE ShoulderLeftAngle NUMERIC
@ATTRIBUTE LeftElbowPositionX NUMERIC
@ATTRIBUTE LeftElbowPositionY NUMERIC
@ATTRIBUTE LeftElbowPositionZ NUMERIC
@ATTRIBUTE ElbowLeftAngle NUMERIC
@ATTRIBUTE LeftWristPositionX NUMERIC
@ATTRIBUTE LeftWristPositionY NUMERIC
@ATTRIBUTE LeftWristPositionZ NUMERIC
@ATTRIBUTE WristLeftAngle NUMERIC
@ATTRIBUTE RightShoulderPositionX NUMERIC
@ATTRIBUTE RightShoulderPositionY NUMERIC
@ATTRIBUTE RightShoulderPositionZ NUMERIC
@ATTRIBUTE ShoulderRightAngle NUMERIC
@ATTRIBUTE RightElbowPositionX NUMERIC
@ATTRIBUTE RightElbowPositionY NUMERIC
@ATTRIBUTE RightElbowPositionZ NUMERIC
@ATTRIBUTE ElbowRightAngle NUMERIC
@ATTRIBUTE RightWristPositionX NUMERIC
@ATTRIBUTE RightWristPositionY NUMERIC
@ATTRIBUTE RightWristPositionZ NUMERIC
@ATTRIBUTE WristRightAngle NUMERIC
@ATTRIBUTE LeftHipPositionX NUMERIC
@ATTRIBUTE LeftHipPositionY NUMERIC
@ATTRIBUTE LeftHipPositionZ NUMERIC
@ATTRIBUTE HipLeftAngle NUMERIC
@ATTRIBUTE LeftKneePositionX NUMERIC
```

```
@ATTRIBUTE LeftKneePositionY NUMERIC
@ATTRIBUTE LeftKneePositionZ NUMERIC
@ATTRIBUTE KneeLeftAngle NUMERIC
@ATTRIBUTE LeftAnklePositionX NUMERIC
@ATTRIBUTE LeftAnklePositionY NUMERIC
@ATTRIBUTE LeftAnklePositionZ NUMERIC
@ATTRIBUTE AnkleLeftAngle NUMERIC
@ATTRIBUTE RightHipPositionX NUMERIC
@ATTRIBUTE RightHipPositionY NUMERIC
@ATTRIBUTE RightHipPositionZ NUMERIC
@ATTRIBUTE HipRightAngle NUMERIC
@ATTRIBUTE RightKneePositionX NUMERIC
@ATTRIBUTE RightKneePositionY NUMERIC
@ATTRIBUTE RightKneePositionZ NUMERIC
@ATTRIBUTE KneeRightAngle NUMERIC
@ATTRIBUTE RightAnklePositionX NUMERIC
@ATTRIBUTE RightAnklePositionY NUMERIC
@ATTRIBUTE RightAnklePositionZ NUMERIC
@ATTRIBUTE AnkleRightAngle NUMERIC
@ATTRIBUTE class {G01,G02,G03,G04,G05,G06,G07,G08}
@DATA
0.04765142,0.5395979,1.507303,128.112348931689,234.301220024309,-
0.1064216,0.4018644,1.445606,152.996343119462,-
0.287439,0.2968765,1.474209,129.022807325042,-
0.4793063,0.3423018,1.402354,166.291260031454,0.1826199,0.4058643,1.426285,215.7068807794
64,0.3943899,0.3012697,1.458287,228.939403207702,0.5849862,0.3547944,1.392525,197.4433962
36623,-0.03229206,-0.05413827,1.412009,271.694978675488,-0.03465796,-
0.3777019,1.442454,171.133250547673,-0.03903793,-
0.6720661,1.424546,105.209520723264,0.1206161,-
0.05125931,1.409772,87.6483304010621,0.1267694,-
0.3608835,1.428322,173.823916768231,0.1182888,-0.6661595,1.417055,104.854296118712,G01
0.0376514212787151,0.5395979,1.507303,128.112348931689,234.301220024309,-
0.116421604752541,0.4018644,1.445606,152.996343119462,-
0.297438958883286,0.2968765,1.474209,129.022807325042,-
0.489306280612946,0.3423018,1.402354,166.291260031454,0.172619869709015,0.4058643,1.42628
5,215.706880779464,0.384389927387238,0.3012697,1.458287,228.939403207702,0.57498615026474
,0.3547944,1.392525,197.443396236623,-0.0422920605540276,-
0.05413827,1.412009,271.694978675488,-0.0446579626202583,-
0.3777019,1.442454,171.133250547673,-0.0490379317104816,-
0.6720661,1.424546,105.209520723264,0.110616115629673,-
0.05125931,1.409772,87.6483304010621,0.116769408583641,-
```

0.3608835,1.428322,173.823916768231,0.108288844823837, -
0.6661595,1.417055,104.854296118712,G01
0.0576514212787151,0.5395979,1.507303,128.112348931689,234.301220024309, -
0.0964216047525406,0.4018644,1.445606,152.996343119462, -
0.277438958883286,0.2968765,1.474209,129.022807325042, -
0.469306280612946,0.3423018,1.402354,166.291260031454,0.192619869709015,0.4058643,1.42628
5,215.706880779464,0.404389927387238,0.3012697,1.458287,228.939403207702,0.59498615026474
,0.3547944,1.392525,197.443396236623, -0.0222920605540276, -
0.05413827,1.412009,271.694978675488, -0.0246579626202583, -
0.3777019,1.442454,171.133250547673, -0.0290379317104816, -
0.6720661,1.424546,105.209520723264,0.130616115629673, -
0.05125931,1.409772,87.6483304010621,0.136769408583641, -
0.3608835,1.428322,173.823916768231,0.128288844823837, -
0.6661595,1.417055,104.854296118712,G01

Apéndice E. Formato de consentimiento para participación en los estudios preliminares realizados

Yo <Nombre del participante>, declaro que he sido informado e invitado a participar en un estudio de investigación científica que cuenta con el respaldo de la <institución>. Mi participación se llevará a cabo en < lugar>, el < fecha y hora>) y consistirá en <detalle de actividades y duración>.

Me han explicado que la información registrada será confidencial, y que solamente datos demográficos y resultados cuantitativos y cualitativos serán recopilados y probablemente publicados.

Estoy en conocimiento que los datos resultantes no me serán entregados y que no habrá retribución por la participación en este estudio incluso si estos resultados resultan en un beneficio para la sociedad que está llevando a cabo la investigación. Asimismo, sé que puedo negar la participación o retirarme en cualquier momento.

Por esto, sí acepto voluntariamente participar en este estudio y he recibido una copia del presente documento.

Firma participante:

Fecha: