



UNIVERSIDAD DE GUANAJUATO

CAMPUS IRAPUATO-SALAMANCA
DIVISIÓN DE INGENIERÍAS

“DISEÑO E IMPLEMENTACIÓN DE UNA PLATAFORMA
DE REALIDAD AUMENTADA COLABORATIVA”

T E S I S

QUE PARA OBTENER EL TITULO DE:
INGENIERÍA EN MECATRÓNICA

PRESENTA:

GUSTAVO ADOLFO MURILLO GUTIERREZ

DIRECTORES:

DR. URIEL HAILE HERNÁNDEZ BELMONTE
DR. JUAN PABLO IGNACIO RAMÍREZ PAREDES

Salamanca, Guanajuato.

Julio 2023

Resumen

El estudio de los sistemas colaborativos basados en realidad aumentada es limitado debido a su inherente complejidad. Estos sistemas requieren la interacción de numerosos elementos para lograr una colaboración efectiva. En la actualidad, se están llevando a cabo varios esfuerzos por parte de empresas para crear este tipo de experiencias. Sin embargo, estas soluciones suelen ser privativas o poseer un alto costo. Por lo tanto, el objetivo de este trabajo es proponer un sistema que facilite la colaboración entre múltiples usuarios en un entorno compartido de realidad aumentada.

La obra se centra en el diseño, implementación, pruebas y evaluación de un sistema de realidad aumentada colaborativa. El sistema desarrollado mostró tener una alta precisión y baja latencia en la visualización y manipulación de objetos aumentados entre distintos dispositivos de cómputo, lo cual contribuye a una experiencia de usuario fluida y natural. Los usuarios pueden interactuar simultáneamente en un mismo escenario, realizar acciones en tiempo real y compartir información de manera efectiva.

Además, se ha puesto énfasis en la creación de estructuras de usuarios intuitivas y eficientes, permitiendo una fácil identificación y comunicación entre los participantes. Finalmente, este trabajo representa un valioso aporte a la literatura actual en el campo de la realidad aumentada colaborativa. Se abordan los desafíos y limitaciones existentes en las tecnologías actuales, se presenta una alternativa viable y de código libre que servirá como base para futuras investigaciones y desarrollos en este campo. Así mismo, se contribuye al cuerpo de conocimiento proporcionando una documentación detallada del sistema y los resultados obtenidos en las pruebas, lo que permite a otros investigadores y profesionales profundizar en el tema y utilizar esta información como referencia en sus propios trabajos.

Dedicatoria

A todas aquellas personas que, de una u otra manera, han formado parte de mi vida y han contribuido a mi formación personal y académica.

A mis amigos: Víctor, Mauricio y Serafín, por su compañía, risas y momentos inolvidables durante este trayecto. Gracias por haberme adoptado y brindado su amistad.

A mi mejor amiga, futura socia y compañera de vida que llegó en el momento adecuado; Angy, por su comprensión, cariño, motivación, inspiración y regaños en esta última etapa. Gracias por tu apoyo en este proyecto.

A mis hermanos, con quienes he compartido toda mi vida; Carlos y Karla. Han sido una fuente de inspiración y apoyo durante toda mi trayectoria académica.

Y, por supuesto, a mis padres, por su amor, paciencia y apoyo incondicional en cada uno de mis proyectos. A mi mamá, Ma. Antonia, quien madrugaba conmigo todos los días para hacerme compañía y reducir mi carga emocional. Y a mi papá, Juan Carlos, quien siempre iniciaba sus temas de conversación con la frase -¿Cuánto te falta?-, pero sabiendo que en el fondo quería compartir este momento conmigo. Todo ese cansancio y sacrificio que hicieron durante todos estos años ha dado frutos y, por ello, son las personas a quién más les estoy agradecido.

Esta tesis está dedicada a ustedes, por su amor, su confianza y su apoyo incondicional. Gracias por hacer posible este logro y por ser parte de mi camino.

Agradecimientos

El desarrollo de esta tesis ha sido posible gracias a la ayuda y apoyo de numerosas personas, a quienes me gustaría expresar mi profundo agradecimiento.

En primer lugar, agradezco a mi director de tesis, el Dr. Uriel Haile Hernández Belmonte, por su orientación experta, su inestimable apoyo y tiempo invertido durante todo el proceso de investigación, desarrollo y evaluación. Sus valiosos comentarios y sugerencias han sido esenciales para la culminación de esta tesis.

También quiero agradecer al Dr. Juan Pablo Ignacio Ramírez Paredes, por su colaboración y asesoramiento en aspectos específicos de la redacción. Su ayuda, apoyo y comentarios fueron fundamentales para la mejora en la calidad de este trabajo, así como el haberme permitido hacer uso de las instalaciones y equipo de investigación.

Agradezco a mi casa de estudios, la Universidad de Guanajuato, por el conocimiento que me permitió adquirir.

Finalmente, agradezco a todas las personas que me apoyaron a evaluar los resultados de esta tesis dedicando su tiempo al participar en las pruebas. Su contribución ha sido esencial para la mejora de esta investigación.

Índice general

1	Introducción	1
1.1	Historia de la realidad aumentada	1
1.2	Panorama actual	3
1.3	Panorama actual de la realidad aumentada colaborativa	4
1.4	Objetivo general de la obra	6
1.5	Objetivos específicos	6
1.6	Justificación	7
1.7	Estructura del documento	7
2	Fundamentos teóricos	9
2.1	Realidad aumentada, realidad virtual y realidad mixta	9
2.1.1	Clases de realidad aumentada	11
2.2	Componentes de un sistema de realidad aumentada	12
2.3	Generación de un entorno aumentado	13
2.3.1	Navegación por el espacio aumentado	14
2.3.2	Descripción espacial	19
2.3.3	Marcadores de realidad aumentada	20
2.4	Proceso de interacción aumentada	22
2.5	Realidad aumentada móvil	23
2.6	Realidad aumentada colaborativa	23
2.6.1	Taxonomía de entornos colaborativos	24
2.6.2	Arquitectura de una colaboración síncrona	25
2.7	Resumen del capítulo	27
3	Metodología	28
3.1	Planteamiento del problema	28
3.2	Sistema de RA colaborativo	29
3.3	Aplicación de RA	30
3.3.1	Interacción gestual con Lean Touch	30
3.4	Framework de diseño RA	31
3.4.1	Framework de RA Vuforia	32
3.4.2	Reconocimiento de imágenes fiduciaras	32
3.4.3	Características de la imagen	33
3.4.4	Descripción de pose en los elementos aumentados	33
3.5	Servidor	34

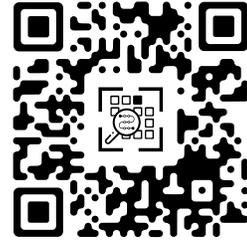
3.5.1	Plataforma de desarrollo .NET	35
3.5.2	Interfaz de comunicación	35
3.5.3	Servidor local	35
3.5.4	Envío de mensajes por Broadcast	38
3.5.5	NetCat	38
3.6	Base de datos	38
3.6.1	Envío de datos	39
3.6.2	MongoDB	39
3.6.3	Esquemas Mongoose	39
3.6.4	Serialización de mensajes	40
3.6.5	Descripción de usuarios	40
3.6.6	Descripción de interacción	41
3.7	Resumen del capítulo	43
4	Pruebas y resultados	44
4.1	Protocolo de pruebas	44
4.1.1	Descripción general de las pruebas	45
4.1.2	Configuración de las pruebas	45
4.2	Recursos utilizados	46
4.3	Despliegue de resultados	48
4.3.1	Aplicación de RA	49
4.4	Exactitud y latencia en la descripción de pose de los dispositivos de RA . .	50
4.4.1	Latencia en búsqueda de una sesión colaborativa	52
4.4.2	Latencia en creación de una sesión colaborativa	55
4.4.3	Latencia y precisión durante la creación de objetos de realidad au- mentada	57
4.5	Análisis de resultados	60
5	Conclusiones y perspectivas	61
5.1	Conclusiones	61
5.2	Perspectivas	62

Índice de figuras

1.1	Diseño industrial del Simulador Sensorama	2
2.1	Diagrama de Milgram	10
2.2	Ciclo de generación de un entorno aumentado	14
2.3	Sistema coordinado rectangular en el plano	16
2.4	Sistema coordinado rectangular tridimensional	17
2.5	Ciclo de interacción en entornos aumentados	22
2.6	Arquitectura cliente servidor	26
3.1	Arquitectura propuesta del entorno compartido de RA	29
3.2	Diagrama de flujo del servidor	37
4.1	Interfaz gráfica de la aplicación de RA colaborativa	49
4.2	Marcador fiduciario utilizado por la aplicación de RA para la definición del marco de referencia global	51
4.3	Latencia de los dispositivos de RA en la actualización de pose	52
4.4	Precisión en la descripción de pose de los dispositivos de RA	53
4.5	Latencias de los dispositivos utilizados en la búsqueda de espacio colaborativo	55
4.6	Comparativo de latencias entre el conjunto de dispositivos evaluados en la creación de un espacio colaborativo	56
4.7	Latencias de los dispositivos utilizados durante la operación de creación de objetos de RA	57
4.8	Precisión de los objetos aumentados instanciados durante la operación de creación de objetos de RA	59

Capítulo 1

Introducción



En el presente capítulo se realiza una revisión histórica de la realidad aumentada, las principales tecnologías que permiten una inmersión en la actualidad y los antecedentes que facilitan la justificación de la obra. Finalmente, se describe la estructura y plan de la tesis, así como la manera en que será desglosada durante el resto del documento.

1.1. Historia de la realidad aumentada

El término realidad aumentada, abreviado como RA, pertenece al grupo de nuevas tecnologías disruptivas destinadas a modificar la percepción de nuestra realidad empleando entidades sintéticas. Aunque su origen data de la década de 1950, época cuyos avances en el área de la electrónica y la computación fueron representativos, no fue sino hasta mediados de 1990 cuando tomó relevancia como potencial campo de investigación.

Carmigniani y Furht [1] describen que los primeros intentos de producir una experiencia sensorial con elementos digitales fueron propuestos por el filósofo, visionario y realizador de cine Morton Heilig. Heilig, en el año de 1955, había propuesto el diseño de un sistema de cine con la capacidad de involucrar los sentidos del espectador en el que se condensara la experiencia por medio de proyecciones 3D, visión periférica, sonido binaural, vibraciones, así como sensaciones táctiles y aromáticas, presentándolo bajo el nombre de “El cinema del futuro”. Sin embargo, fue hasta 1961 cuando presentó un primer prototipo funcional de su concepto; el Simulador Sensorama, ilustrado en la Figura 1.1, consiguiendo su patente hasta 1962.

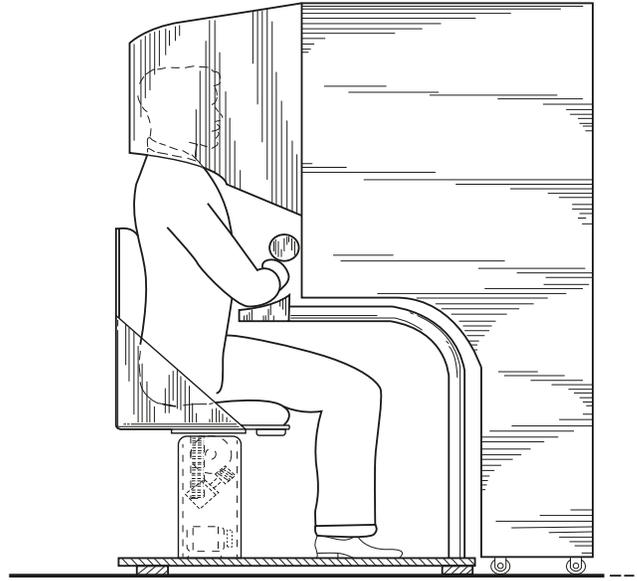


Figura 1.1: Diseño industrial del Simulador Sensorama
Obtenido del diseño original de la patente US-3050870-A [2]

Tres años después, en 1966, Ivan Sutherland continúa con el concepto de interactuar con elementos digitales y, como resultado, inventa el visualizador de montaje sobre la cabeza, HMD (Por sus siglas en inglés, *Head Mounted Display*), una maquinaria de extensas proporciones que se colocaba sobre la cabeza del usuario para conseguir la combinación del mundo virtual y digital. En el año de 1975, en otro intento por ampliar la experiencia virtual, Myron Krueger diseña el *Videoplace*, una habitación que permitía al usuario, por primera vez, interactuar con objetos virtuales.

No obstante, el término realidad aumentada apareció hasta el año de 1992, mientras Caudell y Mizell, investigadores de Boeing implicados en el proceso de desarrollo y mejora en los procesos de fabricación, empleaban un software para desplegar los planos de cableado sobre las piezas producidas. Con ello se comenzó la discusión acerca de las ventajas de la realidad aumentada en contraste a la realidad virtual, concepto con mayores estudios durante la época.

Después de 5 años, en 1997, Ronald Azuma [3], uno de los principales investigadores referidos en el campo de la realidad aumentada, escribe el primer estudio sobre los avances e investigaciones en el campo realizados hasta su fecha. La relevancia del estudio de Azuma marcó un punto de referencia para próximos investigadores, permitiendo considerar a la realidad aumentada como potencial herramienta de soporte en múltiples campos como la medicina, la manufactura o la robótica.

Sin embargo, sus principales aplicaciones comerciales habían mantenido un enfoque en el campo del desarrollo de videojuegos. Por ejemplo, para el año 2000, Bruce Thomas

consiguió desarrollar el primer videojuego de realidad aumentada al aire libre titulado ARQuake. En el 2010, Microsoft lanzó un periférico para su consola de videojuegos; Kinect para Xbox 360 [4], un controlador de juego innovador que permitía a los usuarios interactuar con diferentes videojuegos virtuales y aumentados desarrollados para la misma consola. En último término, Pokemon GO, siendo uno de los títulos mas recientes al ser lanzado en el año 2016.

Empresas pioneras en el desarrollo de *software* tal como Apple, Google y Microsoft, también han incursionado en el campo de la realidad aumentada. Mientras que Apple y Google han invertido recursos para el desarrollo de herramientas de diseño de RA como ARKit y ARCore respectivamente, Microsoft ha apostado por el desarrollo de periféricos que funcionan bajo su propia plataforma; los HoloLens.

Así mismo, desde el lanzamiento del iPhone 4 y el iPad 2, Apple ha contemplado la adición de recursos computacionales para ofrecer experiencias de realidad aumentada desde sus dispositivos. Mientras que, en su oposición y como resultado fallido de un intento por desarrollar dispositivos de RA, Google continua ofreciendo nuevas características y funcionalidades a su propio *toolkit* de desarrollo en el que facilita el diseño de aplicaciones de RA para su sistema operativo Android.

1.2. Panorama actual

Con el abaratamiento y ubicuidad de los dispositivos móviles, la realidad aumentada se ha masificado encontrando nuevas formas de aplicación. Hoy en día, la realidad aumentada ha probado su utilidad en campos como la medicina, la robótica, el mantenimiento de instalaciones industriales, el entrenamiento de personal, la visualización de datos, la educación, la militarización, entre otros [5].

En conjunto con el crecimiento en las áreas de aplicación de la RA, se han desarrollado diversos dispositivos especializados que permiten su despliegue más allá de los dispositivos móviles convencionales, por ejemplo, los Microsoft's HoloLens, Magic Leap One y Meta 2. Pero su demanda exige nuevas funcionalidades tal como la colaboración, esto es, la ejecución en entornos aumentados que permitan la integración síncrona de múltiples usuarios en el mismo ambiente físico.

Algunos modelos tradicionales de colaboración han sido desarrollados en el campo del desarrollo de videojuegos. Los avances, tecnologías y técnicas de comunicación aplicadas a la gamificación pueden ser implementadas en los entornos de realidad aumentada [6]. No obstante, a pesar de ser varios los intentos por conseguir modelos y aplicaciones de entornos colaborativos, sigue sin estandarizarse un modelo único de experiencias compartidas, lo que da lugar a un nuevo tópico de investigación dentro de la realidad aumentada.

A continuación, se presentan algunas aportaciones recientes y trabajos de investigación realizados en el campo de la colaboración en entornos de realidad aumentada.

1.3. Panorama actual de la realidad aumentada colaborativa

Como una propuesta de entornos aumentados compartidos, Zhang *et al.* [7], presentaron CARS (Por sus siglas en inglés, *Collaborative Augmented Reality for Socialization*), una plataforma de interacción inmersiva basada en el reconocimiento de objetos empleando cómputo en la nube. Con el modelo de operación CARS se consiguió generar el entorno de colaboración al separar el medio de procesamiento y el ambiente de proyección aumentada. Mientras que el cómputo en la nube se encarga del procesamiento de los datos y la detección de objetos de referencia, los dispositivos de inmersión se encomiendan al intercambio de la información asociada al cómputo de un objeto previamente proyectado. De esta manera, con el modelo CARS, el costo computacional de cada dispositivo de inmersión se ve reducido en gran escala.

Algunas otras investigaciones se han enfocado en proponer metodologías para la creación de aplicaciones colaborativas. Tal es el caso de la propuesta de Bhattacharyya *et al.* [8], en donde se propone un modelo para el diseño de videojuegos de realidad aumentada síncrona llamado Brick: un videojuego multijugador colocado a escala habitación. Con Brick, se mantuvo el objetivo de forzar la interacción entre los usuarios al asignar metas en común que requerían la exploración e interacción de elementos dispersos en un mapa colaborativo. Presentar un entorno ambientado con diversos elementos de inmersión les permitió sugerir como resultado la necesaria caracterización de cinco categorías en los sistemas de realidad aumentada de mundos compartidos: jugador único, intrapersonal, multijugador, interpersonal y ambiente.

Un modelo de experiencia aumentada colaborativa más general fue presentado por Benbelkacem *et al.* [9]. En su propuesta MVC-3DC (Por sus siglas en inglés, *Model-View-Controller 3D Collaborative*), un modelo de desarrollo 3D para sistemas colaborativos aumentados y virtuales, extienden la aplicación de modelos de realidad aumentada colaborativa a modelos de realidad virtual. La propuesta MVC-3DC presenta una metodología de colaboración basada en una plataforma en la nube. En esta se mantiene el registro de cada elemento y configuración del entorno aumentado. El proceso de colaboración se consigue por medio de sesiones. En cada sesión los usuarios poseen privilegios para realizar modificaciones del entorno. Cada cambio produce una actualización en la misma plataforma, consiguiendo ofrecer un entorno disponible y actualizado para el resto de los usuarios de la sesión durante cada modificación.

En un mismo enfoque, Boonbrahm *et al.* [10] realizaron una exploración de la realidad

aumentada como técnica para la generación de modelos 3D complejos, estables y remotos en entornos de diseño y manufactura colaborativa. En su estudio se emplearon diferentes marcadores para su posicionamiento. Confirman que el sistema propuesto puede ser empleado para una colaboración real en aplicaciones que requieran un herramienta de soporte, tal como asesorías o capacitaciones.

En enfoques con directriz práctica, Mourtzis *et al.* [11], presentaron un estudio experimental de realidad aumentada colaborativa bajo el contexto de *Learning Factory*; la aplicación directa de una plataforma de diseño en la nube que acomete facilitar el diseño 3D CAD colaborativo, en el que, de manera similar a la propuesta de Benbelkacem, integra los cambios realizados por un conjunto de usuarios adscritos a una misma sala o sesión de colaboración.

En cambio, en el campo de la educación, Villanueva *et al.* [12], propusieron Meta-AR-App, una plataforma de realidad aumentada orientada a la educación regida por una arquitectura de operación *Pull*, proveniente de los sistemas de gestión de versiones. Meta-AR-App, a diferencia de los modelos que realizan cambios o actualizaciones instantáneas en todos los usuarios, permite la integración de nuevas modificaciones una vez estos son admitidos por un administrador de sesión. La propuesta permite controlar de manera particular cualquier modificación y reintegrarla a todos los usuarios de la sesión una vez hayan sido aprobados el conjunto de cambios o actualizaciones.

Una aplicación directa de colaboración en entornos aumentados la realiza López-Faican y Jaen [6] al proponer EmoFindAR; un estudio para evaluar el uso de la realidad aumentada móvil (MAR) en un escenario de geolocalización multiusuario enfocado en niños de primaria. Con EmoFindAR se buscaba enseñar a los niños el reconocimiento de emociones básicas por medio de un juego colaborativo/competitivo de búsqueda de personajes virtuales. La aplicación reconocía su entorno sin el uso de marcadores y añadía funciones en las que se obligaba a interactuar a los usuarios de manera integral.

En un estudio experimental, Kerthyayana [13] describe una aplicación prototipo de colaboración simple en la que se utiliza Unity, Google Cloud Anchor y ARCore en dispositivos Android. Su estudio mantuvo su enfoque en la interconexión de dos dispositivos, de los cuales, uno fue el anfitrión de la comunicación y, el segundo, el usuario remoto. En el experimento ambos usuarios fueron capaces de agregar y manipular objetos virtuales desde la pantalla de su dispositivo sin la necesidad de elementos tangibles como referencia.

En un intento por crear una comunicación genérica en los entornos aumentados colaborativos, Kostov y Wolfartsberger [14] propusieron un método para expandir el proceso de colaboración en aplicaciones de realidad extendida (XR) con soporte multiplataforma, es decir, una comunicación capaz de ejecutarse entre entornos de escritorio, teléfonos, tabletas, Steam VR, HoloLens, etc. Como resultado propusieron una plataforma en la que, empleando una red local, la interacción entre los objetos virtuales fue posible sin necesidad de emplear los mismos dispositivos.

Finalmente, Rompapas *et al.* [15] presentaron los resultados de un estudio donde crean HoloRoyale, la primera experiencia colaborativa de larga escala y alta fidelidad. Rompapas consiguió crear e implementar una arquitectura de software que mejora la precisión de la posición y orientación sincronizada entre múltiples usuarios. Su estudio fue aplicado a la exploración de un espacio de 12,500 m^2 que, sin la necesidad del uso de marcadores de referencia, presentó resultados muy prometedores en el campo de la RA colaborativa.

1.4. Objetivo general de la obra

El objetivo general de esta obra es diseñar, implementar y evaluar una plataforma de propósito general que permita la colaboración de dos o más usuarios en el mismo espacio físico, de manera síncrona y compartiendo un mismo entorno de realidad aumentada.

1.5. Objetivos específicos

Los objetivos específicos de esta obra son:

1. Desarrollar una arquitectura multiplataforma, escalable y *open-source* que permita la colaboración entre dos usuarios, en tiempo real y en escala habitación en entornos aumentados compartidos.
2. Evaluar la efectividad del sistema de realidad aumentada colaborativa mediante el diseño e implementación de una interfaz gráfica, midiendo el nivel de precisión y latencia en su proceso de interacción, alcanzando un tiempo de respuesta de menos de 500 ms y un margen de error inferior al 10 %.
3. Desarrollar una estructura de datos serializable que permita la colaboración multiplataforma en el sistema de realidad aumentada colaborativa, garantizando la interoperabilidad entre diferentes dispositivos y plataformas, mediante la implementación de un formato de datos estandarizado y eficiente.
4. Identificar los desafíos técnicos y de implementación que surgen al utilizar la plataforma de realidad aumentada colaborativa, mediante un análisis cuantitativo, evaluando aspectos como la conectividad y el rendimiento del sistema.
5. Discutir la relevancia y el aporte de la plataforma de realidad aumentada colaborativa en la literatura actual, considerando aspectos como la infraestructura propuesta, la integración tecnológica y la accesibilidad.

1.6. Justificación

En años recientes, la realidad aumentada colaborativa ha sido tópico de discusión. López-Faicán [6] menciona que la colaboración ha sido identificada como una habilidad del siglo XXI que debe ser incluida en los actuales modelos educativos; a la par del surgimiento de nuevos avances tecnológicos, se requiere una incorporación integral de estos nuevos cambios.

Esta habilidad se ha visto reflejada en campos como la industria, cuya implementación de la realidad aumentada colaborativa ha tomado relevancia, por ejemplo, en tareas industriales donde los usuarios requieren coordinar e interactuar entre sus actividades del mismo modo que en el campo de la educación, cuyos modelos de enseñanza han optado por incluir espacios gamificados de aprendizaje. La colaboración ha hecho de la realidad aumentada una experiencia compartida, un espacio de inmersión donde los usuarios son capaces de experimentar una misma situación desde diversas perspectivas.

Por el contrario, mientras que la realidad aumentada como tópico ha sido extensamente estudiada, son pocos los trabajos que abordan y estudian la colaboración en este tipo de entornos. La variedad de intentos por desarrollar plataformas colaborativas emplean modelos ya existentes o experimentan con nuevas tecnologías emergentes, lo que produce un basto y complejo campo de exploración.

Hoy en día existen tecnologías y plataformas que permiten simplificar e implementar experiencias de RA colaborativas, no obstante, resultan ser costosas y/o restrictivas en sus funcionalidades. Ante un panorama en el que se dispone de herramientas, modelos y tecnologías bajo una filosofía *open-source* con la que es posible generar un contrapeso a las tecnologías privativas, el presente trabajo se justifica bajo la necesidad de desarrollar sistemas de RA colaborativos que permitan la experimentación e implementación de experiencias de inmersión compartidas, a bajo costo y de propósito general.

1.7. Estructura del documento

Con la finalidad de conseguir abordar satisfactoriamente los objetivos y planes de esta obra, este documento divide su contenido en cinco capítulos.

En el **Capítulo 1** se ha abordado una introducción sobre el panorama histórico y actual de la realidad aumentada, presentándose los objetivos y justificación del escrito.

En el **Capítulo 2** se describen los conceptos fundamentales de la realidad aumentada, los *frameworks* y las tecnologías que se pueden emplear para la creación de aplicaciones de RA. Posteriormente, se complementa con la descripción de algunos de los problemas

asociados a las elecciones de diseño, la arquitectura de un entorno de colaboración y los distintos modelos de comunicación y abstracción de información presente en entornos colaborativos.

En el **Capítulo 3** se ofrece un panorama de la arquitectura diseñada, las herramientas y *software* empleados, así como la descripción particular de los componentes que integran al proyecto para su respectiva réplica.

En el **Capítulo 4** se describen y presentan las condiciones y limitantes de operación, las especificaciones técnicas de los dispositivos bajo evaluación, así como las pruebas realizadas y resultados obtenidos.

Finalmente, en el **Capítulo 5**, se exterioriza un análisis y conclusión a los resultados de esta obra, compartiendo el alcance, valor y futuros aportes que este trabajo puede aportar a la literatura.

Capítulo 2

Fundamentos teóricos



Este capítulo aborda los sistemas de realidad aumentada, partiendo de las definiciones acuñadas por diferentes investigadores, hasta el grado de remarcar diferencias entre tecnologías como la realidad virtual. Posteriormente, se presentan los componentes que integran un sistema inmersivo, así como su respectivo diseño empleando herramientas de desarrollo actuales. Finalmente, se procede con la exposición sobre los entornos aumentados colaborativos, sus respectivas necesidades y las principales técnicas de colaboración y comunicación.

2.1. Realidad aumentada, realidad virtual y realidad mixta

La definición dada por Ronald T. Azuma [3] en 1995, plantea a la realidad aumentada como una variación de los entornos virtuales, también llamados entornos de realidad virtual. Surge como contraste de esta tecnología completamente inmersiva en la que es sintetizado un mundo completamente digital, por el de una tecnología que superpone objetos virtuales en el mundo real. Azuma sugiere que, idealmente en la realidad aumentada, los objetos del mundo real deben coexistir con los objetos del mundo digital al grado de mejorar la percepción y experiencia de interacción del usuario.

Investigadores recientes, tal como Furht, Linowes y Aukstakalnis, comparten los con-

ceptos base definidos por Azuma. Furht [16] describe a la realidad aumentada como una vista en vivo del entorno físico del mundo real, cuyos elementos se fusionan con imágenes aumentadas generadas por ordenador creando una realidad mixta. Linowes [17] define a la RA como la combinación de datos digitales y entradas del mundo real registradas por los sentidos del humano que aparentemente se mezclan en un espacio físico y en tiempo real. Y Aukstakalnis [18] puntualiza a la RA como una tecnología de visualización capaz de superponer o combinar información alfanumérica, simbólica o gráfica con la visión del usuario del mundo real.

No obstante, la definición empleada por Azuma involucra el concepto de realidad virtual como principal punto de diferenciación. Milgram y Kishino [19], en cambio, reconocen que la presencia de los objetos virtuales en el entorno real pueden exceder los límites del entorno físico al grado de no necesariamente regirse por las leyes físicas que gobiernan el espacio, tiempo, mecánica y propiedades de los materiales, por lo que puntualiza que ambas tecnologías, tanto realidad aumentada como realidad virtual, deben mantener una fina distinción en sus definiciones. El diagrama de Milgram, ilustrado en la Figura 2.1, representa la continuidad de la virtualidad, un concepto definido por Milgram y Kishino para referirse a la mezcla de clases de objetos en cualquier situación de visualización concreta.



Figura 2.1: Diagrama de Milgram

Se representa la continuidad de la virtualidad originalmente publicada por Milgram y Kishino [3].

El diagrama de Milgram permite visualizar esa pequeña distinción entre realidad virtual y realidad aumentada. Posiciona al entorno real en un extremo de la línea de continuidad de la virtualidad, y al entorno virtual en el extremo opuesto. El extremo del entorno real consiste en objetos exclusivamente reales, o elementos que pueden visualizarse a partir de una pantalla de video convencional que ha capturado una escena del mundo real. El extremo del entorno virtual, en contra parte, consiste en objetos exclusivamente virtuales.

Cualquier punto intermedio entre ambos extremos de la línea de continuidad de virtualidad se define como realidad mixta, esto es, la presentación conjunta de los objetos virtuales y reales. En cuanto a la realidad aumentada, esta se posiciona como una expe-

riencia más cercana al entorno real, y la virtualidad aumentada como una experiencia más próxima al entorno virtual. Sin embargo, la popularidad que el término de realidad aumentada adquirió durante la década de 1990, facilitó su proceso de adopción en la comunidad, aunque se mantiene el matiz distintivo con la realidad virtual.

2.1.1. Clases de realidad aumentada

Dentro del concepto de realidad aumentada se presentan características que sobresalen entre las diferentes aplicaciones de este tipo de entornos [16] [17]. Estas particularidades pueden ser aprovechadas para generar distintas clasificaciones en los entornos aumentados, listando dos de ellas a continuación:

Clasificación por el tipo de entorno

Una primer clasificación puede realizarse por las características del entorno bajo el cual el usuario es capaz de interactuar: un entorno fijo o un entorno móvil.

Fijo. En un entorno fijo, el usuario se adapta al alcance y campo de operación del sistema de realidad aumentada. Este último es quien permanece en una posición fija durante todo el proceso de inmersión. Esta clasificación presenta su variante para espacios en interiores y exteriores.

Móvil. Un entorno móvil también presenta una variante para espacios en interiores y exteriores pero, al contrario de un entorno fijo, el dispositivo de realidad aumentada tiene la capacidad de desplazarse en conjunto con el usuario y generar el proceso de superposición de información de manera simultánea.

Clasificación por el tipo de seguimiento

Una segunda clasificación puede centrarse por el tipo de seguimiento o rastreo de información que permite la generación del entorno aumentado. La generación de elementos digitales requiere un atenuante para la superposición de las entidades y las dos principales técnicas se describen a continuación:

Con marcadores. En esta técnica se emplean imágenes en alto contraste y con patrones asimétricos computacionalmente simples de reconocer. Los marcadores son empleados como lanzadores de eventos o marcos de referencia para la interacción del entorno

aumentado. Aunque idealmente se busca la implementación de marcadores sencillos, pueden ser utilizados diferentes variantes.

Imágenes: Como fotografías, ilustraciones o geometrías complejas.

Marcador codificado: Como códigos QR, códigos de barras o patrones geométricos.

Reconocimiento de texto: Como símbolos, letras o guarismos numéricos.

Sin marcadores. Esta segunda técnica requiere de características y componentes adicionales lo que incrementa su complejidad y costo de implementación. Consiste en utilizar elementos integrados al sistema de RA como sensores o instrumentos de medición, con los que se adquieran datos del entorno en tiempo real. El uso de estos componentes permiten escalar el espacio de interacción por medio de distintas funcionalidades como:

Ubicación: Disponible por el uso de sensores espaciales como acelerómetros, giroscopios, brújulas, unidades de medición inercial o con el uso de tecnología GPS.

Reconocimiento de contornos: Presente en entornos físicos con nulos o mínimos obstáculos en el que, con el uso de cámaras digitales con sensores de profundidad, se extraen características del espacio con los cuales se realiza la generación de planos virtuales.

Proyección: Enfocado en la presentación de elementos virtuales sobre superficies bien definidas, usualmente en un estado estacionario, como hologramas ilustrativos o instructivos.

2.2. Componentes de un sistema de realidad aumentada

La realidad aumentada es un medio en el que se proyectan entidades virtuales, generadas por ordenador, sobre la información presente en el mundo real. Esta superposición de información produce en el operador una experiencia de inmersión cuyos elementos no quedan limitados a la mejora de experiencias visuales, por el contrario, se extienden al grado de la generación de sensaciones hápticas, aromáticas o sonoras.

Azuma [5] sugiere que, como un sistema que engloba a un conjunto de componentes y tecnologías, cualquier aplicación de realidad aumentada debería satisfacer los siguientes tres requisitos:

1. Asociación y alineamiento de la información del mundo físico y virtual.
2. Registro y renderizado de los elementos virtuales en tiempo real a través de todos los canales sensoriales disponibles.

3. Entorno interactivo en tiempo real.

Cuando se considera la integración de un sistema que busca satisfacer estos tres requisitos, se contempla entonces la integración de un subsistema que presente la capacidad de capturar imágenes del mundo real, la asistencia de un subsistema electrónico que permita recuperar información del espacio físico, un elemento que facilite la estimación de la posición y orientación (pose) en el espacio, y un subsistema que reproduzca los elementos aumentados desde un computador. La integración de un sistema de realidad aumentada puede sintetizarse por los siguientes cinco elementos:

1. Dispositivo de visualización y captura de información. Pantalla e instrumentos adecuados para la combinación y coexistencia de elementos virtuales y reales en función de la información adquirida del medio físico.
2. Cámara. Dispositivo para la captura de imágenes.
3. Procesador. Dispositivo que permite la operación y procesamiento de la información capturada, en el que se implementa la funcionalidad del entorno y la generación de respuestas y salidas de información.
4. Software de RA. Conjunto de funcionalidades que describen el comportamiento e integración de todos los subsistemas.
5. Contenido digital. Sonido, imágenes y modelos digitales que se implementan sobre el mundo real.

2.3. Generación de un entorno aumentado

La generación de un entorno aumentado se refiere a la creación de escenarios digitales que pueden ser superpuestos sobre el mundo real empleando dispositivos de realidad aumentada.

Este proceso de coexistencia implica el reconocimiento del espacio de interacción; la combinación del entorno aumentado con las estructuras del entorno real, así como la descripción de la pose de los objetos aumentados en referencia a su entorno interactivo. De manera general, como queda ilustrado en la Figura 2.2, la generación de un entorno aumentado es un proceso cíclico de detección de elementos tangibles y generación o modificación de elementos intangibles.

Partiendo de la captura de una imagen obtenida por el dispositivo de realidad aumentada, se buscan patrones, marcadores o información de referencia que permitan describir

la pose del dispositivo con respecto a esa referencia. Posteriormente, se genera la información aumentada sobre la pantalla del dispositivo en función de la información de referencia obtenida para, de manera ininterrumpida, continuar con el ciclo de generación del entorno.

En lo particular, la localización del dispositivo se limita a una estimación. La extracción de características relevantes obtenidas de las imágenes del mundo real, en conjunto con información procedente de sensores incluidos en los dispositivos de realidad aumentada, imposibilitan definir trayectorias de movimiento exactas, por lo que emplear algoritmos de corrección se frecuente en aplicaciones de RA.

Esta descripción empleando métodos de aproximación, en donde se busca presentar varios elementos aumentados en conjunto con animaciones y sonidos, precisan de un alta capacidad computacional para conseguir la reproducción estática o dinámica de dichos elementos. Debido a ello, en los sistemas de realidad aumentada se contemplan los componentes adecuados que permitan satisfacer el proceso de interacción.

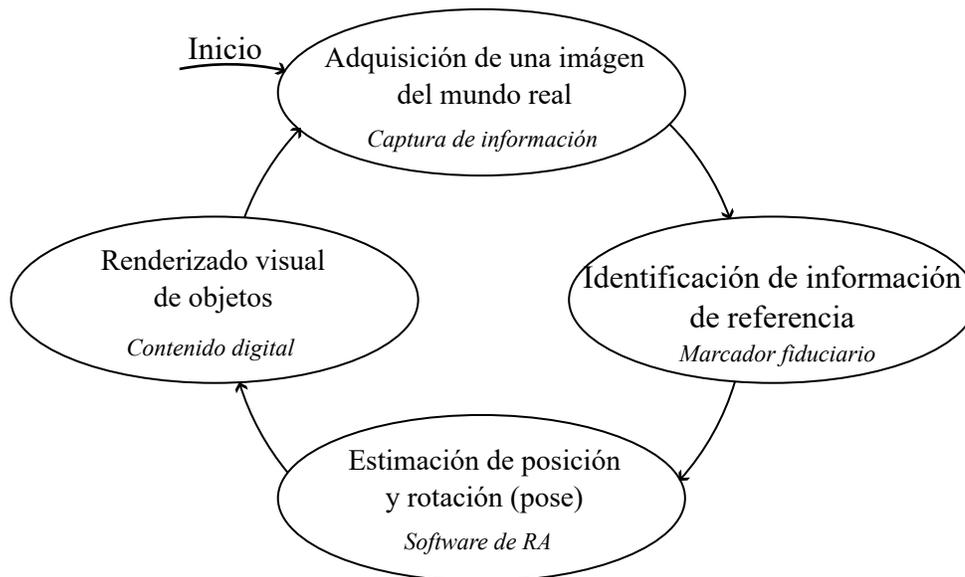


Figura 2.2: Ciclo de generación de un entorno aumentado

2.3.1. Navegación por el espacio aumentado

El espacio aumentado o espacio percibido puede ser descrito como el entorno bajo el que se desenvuelve un observador. Consiste íntegramente de todos los elementos, tanto reales como digitales, que pueden ser divisados por el usuario.

Dentro de este entorno, los usuarios son capaces de desplazarse por un espacio finito rodeado de distintas entidades, que pueden adoptar un estado estático o dinámico, con las que es posible interactuar física o virtualmente. Sin embargo, la exposición de todo

elemento presente en el espacio es subjetivo.

Cada observador presenta una posición, orientación y perspectiva diferente que afecta la descripción de las entidades en el espacio, por lo que el uso de sistemas coordenados, marcos de referencia y modelos físicos y matemáticos con los cuales se compare el estado de cada entidad con el paso del tiempo, es acostumbrado como solución a esta problemática.

Sistema de referencia

En términos simples, un sistema de referencia es un observador en el espacio que dispone de una regla y un reloj para realizar mediciones. Dichas mediciones se efectúan bajo estándares o convenciones con las cuales la descripción de movimiento de los objetos en el espacio pueda ser definido.

Los sistemas de referencia pueden ser fijos o móviles. Son fijos cuando el observador permanece en un punto sin presencia o nula rotación, y son móviles cuando se encuentran sujetos a un movimiento de: traslación, rotación o una combinación de ambos, por lo que el movimiento que presenta un cuerpo es fundamental para la definición de un sistema de referencia.

Tipos de movimiento

Partiendo de un análisis general a un análisis particular, el movimiento espacial general lista al conjunto de movimientos presentes en los entornos de realidad aumentada. En este tipo de movimiento un cuerpo tiene la capacidad de moverse libremente en el espacio. Corresponde a un caso muy particular del cual derivan tipos de movimiento más específicos como:

Movimiento de rotación alrededor de un punto fijo. También conocido como movimiento esférico, se caracteriza por la presencia de un centro de rotación, es decir, un punto que permanece fijo, mismo que forma parte de trayectorias esféricas concéntricas de los cuerpos que orbitan sobre este punto, manteniendo como constante la distancia con cada cuerpo en movimiento, por ejemplo, un giroscopio.

Movimiento de rotación alrededor de un eje fijo. En este tipo de movimiento todas las partículas de una línea del cuerpo bajo análisis permanecen con la misma posición y orientación, mientras que el resto de partículas del mismo cuerpo quedan definidas por un movimiento circular y perpendicular al eje fijo. Un globo terráqueo, una noria o el giro de un carrusel son algunos ejemplos de este tipo de movimiento.

Movimiento de traslación. En un movimiento de traslación todas las líneas de un cuerpo permanecen con una misma orientación, pero son capaces de desplazarse en distintas direcciones. De este tipo de movimiento sobresale la traslación rectilínea, traslación plana y la traslación espacial.

Movimiento plano general. El movimiento plano general se caracteriza por involucrar el movimiento de todas las partículas del cuerpo en planos paralelos, por ejemplo, el movimiento de un borrador sobre la superficie de un pizarrón.

Sistema coordenado

Un sistema coordenado surge como herramienta para la correspondencia de puntos en el espacio. Es el observador quien lo emplea para describir qué, cómo y dónde sucede un fenómeno. Existen diferentes sistemas coordenados, cada uno destinado a facilitar la descripción de fenómenos particulares, siendo los más empleados el sistema coordenado rectangular, cilíndrico y esférico para la descripción de movimiento. A continuación revisaremos algunos de los sistemas coordenados que se utilizan para describir el espacio físico y de como son utilizado en la RA.

Sistema coordenado rectangular

El sistema coordenado rectangular, también llamado sistema coordenado cartesiano, es una entidad matemática conformada por dos rectas ortogonales entre sí llamadas ejes del sistema e identificadas como eje x , llamado eje de las abscisas, y eje y , llamado eje de las ordenadas. El punto de intersección de ambas rectas recibe el nombre de origen, representado por O en la Figura 2.3, y las cuatro regiones creadas al dividir el plano son llamadas cuadrantes.

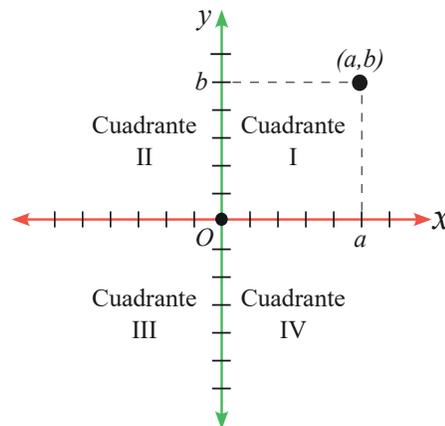


Figura 2.3: Sistema coordenado rectangular en el plano

En un espacio bidimensional, la descripción de posición se realiza al definir un par ordenado de coordenadas con referencia a la distancia que debe recorrerse sobre los ejes x e y respectivamente, de la forma (x, y) . Cada eje presenta una graduación y es en base a ella con la que es descrita la proyección de posición.

Esta representación, originalmente propuesta por René Descartes, puede ser extendida a un espacio tridimensional. El grado adicional de descripción se consigue al incluir un eje z ortogonal al plano formado por los ejes x e y , de la forma (x, y, z) . Por convención, este eje adicional satisface la regla de la mano derecha para la definición de la dirección del eje z , como queda ilustrado en la Figura 2.4.

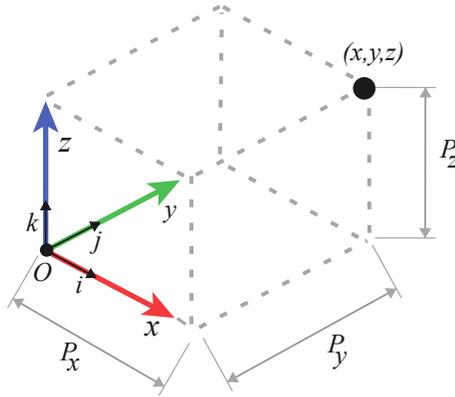


Figura 2.4: Sistema coordenado rectangular tridimensional

Adicionalmente, esta definición puede emplear la convención de descripción vectorial si a cada eje del sistema coordenado se le asocia un vector unitario. Un vector \hat{i} en dirección del eje x , un vector \hat{j} en dirección al eje y , y un vector \hat{k} en dirección del eje z . Esta representación facilita la descripción de posición en el espacio de manera vectorial de la forma $P_x\hat{i} + P_y\hat{j} + P_z\hat{k}$.

Orientación

En un entorno aumentado, además de existir traslación, las entidades digitales presentan rotación. Debido a que la rotación no es un vector y que cada entidad del entorno aumentado puede estar sujeta a un marco de referencia distinto, la descripción de la orientación debe mantener una convención que facilite su análisis cinemático en función a un marco de referencia global.

Las matrices de rotación canónicas \mathbf{R} , cuyo eje de rotación es alguno de los vectores unitarios $\hat{i}, \hat{j}, \hat{k}$, tal como como quedan representadas en la ecuación 2.1, satisfacen esta necesidad.

$$\begin{aligned}
\mathbf{R}_x &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \text{ Cuando :} \\
&\quad \hat{u} = \hat{i} \\
\mathbf{R}_y &= \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix} \text{ Cuando :} \\
&\quad \hat{u} = \hat{j} \\
\mathbf{R}_z &= \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ Cuando :} \\
&\quad \hat{u} = \hat{k}
\end{aligned} \tag{2.1}$$

Cuaterniones

Los cuaterniones son un concepto algebraico isomorfo, una extensión del campo de los números reales e imaginarios, que permiten describir la rotación en el espacio tridimensional. Dentro de la matemática, forma parte del cálculo vectorial.

Su elemento básico es un vector de cuatro componentes: una componente real, expresada como q_w en la ecuación 2.2, y tres componentes imaginarias, q_x, q_y, q_z , todas ellas ortogonales a la componente real.

$$\mathbf{q} = q_w + q_x \hat{i} + q_y \hat{j} + q_z \hat{k} \tag{2.2}$$

La norma de un cuaternion es 1, lo que resulta ser una ventaja con respecto a las matrices de rotación, cuya normalización de una matriz 3x3 y verificación de su ortogonalidad genera mayor complejidad, tanto analítica como computacionalmente.

Las rotaciones básicas, como se expresan en la ecuación 2.3, son la representación de las rotaciones alrededor de los ejes x, y, z .

$$\begin{aligned}
\mathbf{q}_x(\alpha) &= \left[\cos\left(\frac{\alpha}{2}\right) \quad \sin\left(\frac{\alpha}{2}\right) \quad 0 \quad 0 \right] \\
\mathbf{q}_y(\phi) &= \left[\cos\left(\frac{\phi}{2}\right) \quad 0 \quad \sin\left(\frac{\phi}{2}\right) \quad 0 \right] \\
\mathbf{q}_z(\theta) &= \left[\cos\left(\frac{\theta}{2}\right) \quad 0 \quad 0 \quad \sin\left(\frac{\theta}{2}\right) \right]
\end{aligned} \tag{2.3}$$

2.3.2. Descripción espacial

En un entorno de realidad aumentada, los objetos virtuales pueden estar sometidos a diferentes tipos de movimiento o permanecer de manera estática en algún punto del espacio. Para la correcta navegación e interacción por el entorno de manera efectiva, la descripción espacial de cada entidad aumentada es necesaria para la superposición de información digital.

Se tiene una descripción completa cuando se describe la posición y orientación del cuerpo bajo análisis en el entorno. Un primer acercamiento se consigue suponiendo a todas las entidades del espacio aumentado como cuerpos rígidos.

Un cuerpo rígido es un modelo que supone a un sistema de partículas cuya distancia y orientación entre dos o más cualesquiera de ellas permanecen constantes, es decir, no sufren deformación, pero es posible presentar una transformación en traslación y rotación.

Se entiende por transformación de un cuerpo rígido a la invariabilidad en la distancia y orientación entre partículas del cuerpo cuando éste se somete a la traslación y rotación. Además, de mantener el orden cíclico.

Transformación homogénea

La pose de un objeto del entorno aumentado puede definirse con un vector de posición cuyo origen es un marco de referencia global, generalmente identificado como $\{W\}$, y por una matriz de rotación \mathbf{R} . Sin embargo, las matrices de rotación no siempre son coincidentes o proyectadas sobre marcos de referencia fijos.

Si la entidad aumentada presenta un marco de referencia móvil, usualmente presenta una rotación y un desfase con respecto a $\{W\}$, debido a ello, es necesario considerar el desfase a la par de la rotación \mathbf{R} .

Una solución que facilita la integración de estos dos componentes descriptivos es el uso de una matriz de transformación homogénea \mathbf{H} . La matriz de transformación homogénea es una representación matricial del movimiento de un cuerpo rígido del grupo $SE(3)$ (Por sus siglas en inglés, *Special Euclidian Group Order 3x3*). Esta matriz expande el orden de una matriz de rotación 3×3 , a una matriz de orden 4×4 que tiene la estructura expresada en la ecuación 2.4.

$$\mathbf{H} = \begin{bmatrix} \mathbf{R} & d \\ 0 & 1 \end{bmatrix} \quad (2.4)$$

Donde:

- R es la matriz de rotación 3×3 .
- d es el vector de desplazamiento 3×1 .
- 0 vector de ceros de orden 1×3 que representa la perspectiva.
- 1 es el factor de escala que, al suponerse un modelo de cuerpo rígido, es invariable debido a su nula deformación.

2.3.3. Marcadores de realidad aumentada

En el mundo real, la asignación de un marco de referencia puede delegarse a cualquier estructura, cuerpo o punto en el espacio. En el entorno aumentado, esta referencia puede realizarse empleando marcadores o contenido sobre imágenes planas, modelos tridimensionales o seguimiento de características que desencadenen el contenido aumentado.

Imágenes

Son representaciones visuales, impresas en papel o presentados en pantallas, empleadas como referencias estáticas para la presentación de elementos digitales. Sus principales variantes se describen a continuación:

Códigos de barras

El código de barras, originalmente presentado en 1952 en una versión circular muy primitiva, se integra de varias líneas paralelas y verticales de color negro con espacios en blanco, ambos de distinto grosor.

Cada código es leído por un sistema óptico que emite una luz sobre el código. Las líneas color negro absorben la luz incidente, mientras que los espacios en blanco reflejan la luz hacia el lector óptico, quien la captura e interpreta en un formato digital.

El código de barras, por sí mismo, no almacena información. Es una abstracción numérica que puede ser empleada como una llave de acceso que permite obtener información relacionada de, por ejemplo, una base de datos.

Para su lectura se requiere disponer de un dispositivo especializado, lo que reduce su uso en aplicaciones de RA.

Códigos QR

Los códigos QR, o códigos de respuesta rápida (Por sus siglas en inglés, *Quick Response Code*), son gráficos adaptados al uso de dispositivos digitales para su rápida conversión a un formato comprensible para el humano.

Se integra por una matriz cuadrículada, usualmente en colores negro y blanco, en la que se generan patrones en función de los caracteres que se deseen codificar, siendo posible codificar hasta 7,089 caracteres ASCII.

El código QR presenta una mayor capacidad que un código de barras. Puede ser leído a alta velocidad en cualquier dirección y tiene la capacidad de corrección de errores, siendo posible corregir hasta el 30% de información dañada o corrupta.

Reconocimiento de texto

El reconocimiento óptico de caracteres OCR (Por sus siglas en inglés, *Optical Character Recognition*), es un sistema de reconocimiento de texto no digitalizado.

Pertenece al reconocimiento de patrones por inteligencia artificial en el que, para su funcionamiento, se requiere la presencia de una base de datos. Opera capturando imágenes o escaneando completamente un escrito y compara la simbología capturada con la disponible en la base de datos.

El proceso de comparación parte de un reconocimiento del espacio entre cada carácter, en el que se discierne entre símbolos y letras, para acondicionar la captura y analizar la alineación del contenido.

Modelos 3D

Los modelos 3D son empleados para realizar un reconocimiento y seguimiento de objetos presentes en el entorno real analizando la forma o estructura del cuerpo. Juguetes, impresiones 3D, herramientas, vehículos, edificaciones o cualquier objeto rígido con superficies estables puede ser empleado como modelo 3D.

Para el uso de modelos 3D como marcadores de referencia, se requiere disponer previamente de los datos de un modelo CAD o un escaneo 3D. Se verifica la capacidad de ser rastreable y se configura para la definición de las vistas o elementos que se desean proyectar sobre el modelo. Debido a la cantidad de información que debe ser procesada, este tipo de marcadores requieren una mayor capacidad de procesamiento.

2.4. Proceso de interacción aumentada

Retomando la propuesta de Azuma para la consideración de un sistema de realidad aumentada, ofrecer un entorno de interacción entre usuario y objetos aumentados en tiempo real, es uno de los requisitos fundamentales que deben ser satisfechos por todo sistema de RA.

Linowes [17] resume el proceso de interacción como un ciclo de eventos, esto es, un conjunto de procesos u operaciones de interacción generados por el usuario, tal y como se ilustra en la Figura 2.5. En este ciclo se parte del registro de un evento, como puede ser: la interacción del usuario con la pantalla del dispositivo, el contacto con un botón virtual, la detección de un gesto táctil o el cambio de magnitud en algún sensor implementado.

Cuando se tiene registro de una interacción, el sistema correlaciona las entidades aumentadas con el entorno real. Se hace uso de funciones de trazado de rayos para poder determinar si un objeto aumentado es seleccionado o sufre una modificación desde las interfaces del entorno real. El conjunto de funcionalidades disponibles, definidas por el desarrollador del sistema de RA, producirán las modificaciones geométricas adecuadas al evento generado.

Finalmente, si se trata de una rotación, adición, eliminación, escalamiento o modificación del cuerpo digital, se verá reflejado en ambos entornos al renderizar los cambios ejecutados.

El proceso de generación de eventos y correlación pantalla-mundo puede presentar variantes de acuerdo a la aplicación de RA destino y del dispositivo de realidad aumentada empleado.

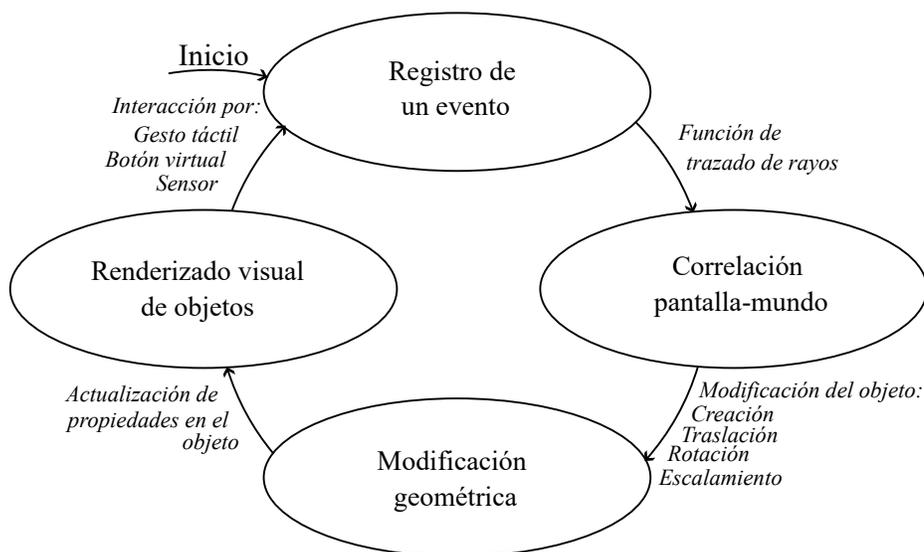


Figura 2.5: Ciclo de interacción en entornos aumentados

2.5. Realidad aumentada móvil

Craig define a la realidad aumentada móvil, MAR, (Por sus siglas en inglés, *Mobile Augmented Reality*), como cualquier sistema de realidad aumentada que puede trasladarse a todos lugares e integre los componentes suficientes para ejecutarse de manera práctica, tal como teléfonos inteligentes o tabletas.

Al ejecutarse sobre dispositivos de mayor accesibilidad y menor costo, la realidad aumentada móvil ha presentado un crecimiento notable desarrollándose en áreas que fueron exclusivas de la RA convencional. Presentan un menor coste económico, son dispositivos de uso general, aunque las aplicaciones de realidad aumentada son limitadas tecnológica y ambientalmente por las especificaciones de los dispositivos móviles.

2.6. Realidad aumentada colaborativa

La realidad aumentada colaborativa, CAR, (Por sus siglas en inglés, *Collaborative Augmented Reality*), es una variante de la realidad aumentada que permite a múltiples usuarios interactuar y colaborar en un entorno compartido en tiempo real. Esto se logra mediante la superposición de información digital sobre la realidad física a través de dispositivos móviles o gafas de RA.

La realidad aumentada colaborativa tiene aplicaciones en diversos campos, como la educación, los videojuegos, la publicidad y el entretenimiento. Por ejemplo, en el ámbito educativo, puede ser utilizada para facilitar la enseñanza y el aprendizaje de conceptos abstractos de manera lúdica y atractiva. En el ámbito de los videojuegos, puede ofrecer experiencias de juego inmersivas y enriquecedoras. En el ámbito publicitario, puede utilizarse para crear campañas publicitarias interactivas y personalizadas y, en el ámbito del entretenimiento, complementar las experiencias de manera integral.

Aunque la realidad aumentada colaborativa tiene un gran potencial, también presenta algunos desafíos y limitaciones. Uno de los principales problemas es la necesidad de una conexión a internet de alta velocidad para poder utilizarse de manera efectiva. Además, la interoperabilidad entre diferentes dispositivos y sistemas de RA todavía es un problema que necesita ser abordado.

Para su comprensión, los espacios compartidos de entornos aumentados forman parte de la realidad aumentada en el continuo de la virtualidad, en los que se mantienen las propiedades y propósitos de la RA. Sin embargo, los entornos aumentados compartidos consideran características adicionales que se deben satisfacer para la creación de espacios interactivos, por ejemplo, el tiempo y el espacio de ejecución.

2.6.1. Taxonomía de entornos colaborativos

Al igual que un sistema de realidad aumentada puede clasificarse por el enfoque de la aplicación, los entornos colaborativos pueden categorizarse por la tecnología o técnica que permite la colaboración.

Una primera clasificación puede centrarse por el tiempo de respuesta del entorno, dividiéndose en colaboración síncrona y asíncrona.

Síncrono. En un entorno de respuesta síncrono, la interacción entre los usuarios se realiza de manera simultánea. Los cambios, modificaciones o actualizaciones se reflejan en tiempo real en cada uno de los usuarios presentes en el entorno aumentado. Presenta como variante la colaboración distribuida o remota y la colaboración en el mismo lugar.

Asíncrono. En contraparte, un entorno asíncrono refleja la creación y visualización del contenido aumentado cuyas solicitudes no requieren la operación simultánea de los usuarios. Es decir, cada usuario trabaja de manera local y, posteriormente, se comparten sus operaciones con el resto de los usuarios en algún momento en el futuro.

Otra clasificación puede realizarse al considerar el espacio donde se desenvuelve el proceso de colaboración, bien sea de manera remota o en el mismo lugar.

En el mismo lugar. La colaboración en el mismo lugar hace referencia a un entorno de colaboración cuyos usuarios se desenvuelven en un mismo espacio físico, haciendo indispensable su presencia en el entorno. Así mismo, este tipo de colaboración puede variar en la escala en que se realiza la interacción, pudiendo ser:

Escala habitación. Para espacios internos.

Gran escala. Para espacios de exteriores.

Remoto. En cambio, en una colaboración remota, los usuarios prescinden de una interacción en el mismo espacio, es posible una integración desde distintos lugares sin ser indispensable operar bajo un mismo ambiente.

Finalmente, una tercer categoría caracteriza a los entornos colaborativos por su simetría de operación, es decir, el rol que desempeña cada usuario o el tipo de dispositivo que se requiere implementar para la colaboración.

Simetría de roles. Se refiere al tipo de tareas o funciones que los usuarios desempeñan. En un rol simétrico, todos los usuarios ejecutan el mismo tipo de actividades, mientras que en un rol asimétrico los usuarios asisten de manera independiente.

Simetría tecnológica. Implica las especificaciones de *software* y *hardware* de los dispositivos de realidad aumentada. Donde, en el caso tecnológicamente simétrico, es indispensable emplear los mismos dispositivos, mientras que, en un caso tecnológicamente asimétrico, se extiende el uso de dispositivos con distintas especificaciones, es decir, se considera una colaboración multiplataforma.

2.6.2. Arquitectura de una colaboración síncrona

Los entornos de realidad aumentada colaborativa pueden ser desarrollados bajo distintas arquitecturas de diseño basados en sistemas distribuidos, siendo las arquitecturas centralizadas y descentralizadas las más frecuentes en uso.

Bikramaditya [20] define como sistema distribuido al conjunto de computadoras autónomas comunicadas por una red gestionada por *software*. Se caracteriza por su alta capacidad de integración entre computadoras y usuarios sin afectar su dispersión geográfica, la autonomía de las computadoras, la seguridad y robustez de la red, la calidad de servicio que proporciona y su movilidad.

Arquitecturas centralizadas

Una arquitectura centralizada es un sistema distribuido definido por un patrón de jerarquía. Implica la presencia de un componente lógico principal y varios componentes que se ramifican. El flujo de información se gestiona por un módulo principal. Si el módulo principal sufre una avería, el resto de los módulos subsecuentes pierden su comunicación hasta ser restaurado.

La arquitectura cliente-servidor, como la que se ilustra en la Figura 2.6, es uno de los sistemas distribuidos centralizados más empleados en la actualidad. En este tipo de arquitectura los clientes realizan las peticiones de un servicio, normalmente siendo estos mismos quienes inician la comunicación, y los servidores proveen los servicios solicitados por el cliente. El servidor se mantiene entonces en calidad de lectura y a la espera de peticiones, por lo que su tiempo de respuesta se mantiene sujeto a la cantidad de peticiones y a su velocidad de procesarlas.

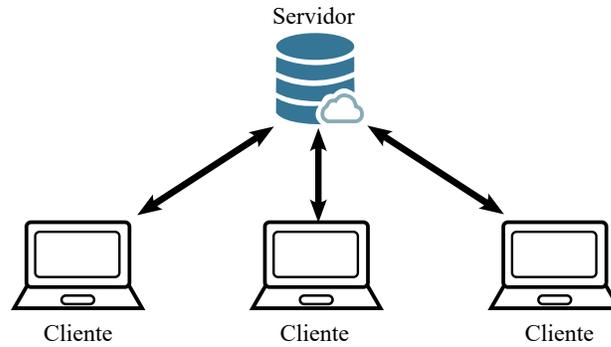


Figura 2.6: Arquitectura cliente servidor

Arquitecturas descentralizadas

Las arquitecturas descentralizadas, en contraste, son sistemas distribuidos que balancean la carga y distribución de tareas entre todos los módulos lógicos del sistema; las unidades lógicas o nodos del sistema operan de manera autónoma.

En este tipo de arquitectura, los componentes del sistema actúan como clientes y servidores de manera simultánea, por lo que se elimina la dependencia de un módulo principal. De esta forma, ante la presencia de una falla, el sistema es capaz de mantener su funcionamiento.

2.7. Resumen del capítulo

En este capítulo fueron abordados los fundamentos teóricos de los sistemas de realidad aumentada colaborativos, partiendo de los principios de un sistema de realidad aumentada, tal como: definiciones, clasificación y componentes, hasta la explicación del proceso de generación y descripción del entorno aumentado.

Así mismo, fueron presentadas algunas generalidades sobre el proceso de navegación en un entorno aumentado considerando el modelo matemático de un cuerpo rígido como herramienta de descripción cinemática del espacio aumentado.

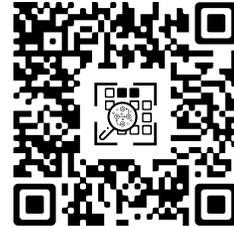
Posteriormente, se procedió con la exposición de las diferentes técnicas de reconocimiento espacial utilizadas en los entornos aumentados colaborativos, sus respectivas necesidades que deben satisfacerse, así como las principales técnicas de colaboración y comunicación.

Por último, se presentaron las dos principales arquitecturas de comunicación empleadas actualmente caracterizando su aplicación en función al enfoque deseado en un entorno aumentado colaborativo.

En el siguiente capítulo se presentará la metodología empleada para el desarrollo de un sistema de realidad aumentado colaborativo, mismo que se sustenta de los principios teóricos del presente capítulo.

Capítulo 3

Metodología



En este capítulo se presenta la metodología empleada para abordar el problema de la realidad aumentada colaborativa. Se comienza con el planteamiento del problema, posteriormente se procede con la explicación del sistema de solución propuesto, incluyendo una breve reseña de los módulos que lo integran, y se prosigue con una explicación a detalle del sistema integrador exponiendo las características, funciones y técnicas empleadas en su fase de evaluación. Finalmente, se explica el conjunto de consideraciones y restricciones por contemplar en los procesos involucrados para su respectivo funcionamiento.

3.1. Planteamiento del problema

Un ambiente de realidad aumentada colaborativa es donde dos o más usuarios comparten un mismo entorno aumentado de manera síncrona. Debido a que dos o más usuarios experimentan la inmersión de manera simultánea y en el mismo espacio físico, esta nueva característica aumenta la complejidad en esta clase de sistemas.

El primer problema en un entorno colaborativo se presenta en la definición y/o establecimiento de un marco de referencia global o absoluto común, tanto para objetos aumentados como usuarios, con los que se describa la pose de cada entidad sobre el entorno. El conocer la pose de cada elemento es indispensable para conseguir una adecuada

interacción.

En segunda instancia, el tener un medio de comunicación que supervise el estado de cada objeto y usuario, así como actualizar los cambios en las propiedades de los objetos aumentados, tal como: pose, tamaño, visibilidad, interacción, entre otras, es fundamental. Además, se requiere que dicho sistema de comunicación tenga una baja latencia.

Finalmente, un tercer problema surge de la necesidad de presentar el conjunto de objetos aumentados con la perspectiva y escala métrica adecuada a cada usuario del sistema. El no disponer de un marco de referencia absoluto implica que cada usuario defina un marco de referencia relativo; cada objeto presente en el entorno se especifica en función del marco de referencia del usuario que realiza la modificación. Una vez que la modificación es realizada, el resto de los usuarios deben ser capaces de reflejar los cambios en sus propios marcos de referencia.

Considerando estas tres problemáticas, el presente trabajo aborda como objetivo principal el desarrollo de una arquitectura de uso general, código libre, y multiplataforma que permita implementar un entorno colaborativo síncrono de realidad aumentada.

3.2. Sistema de RA colaborativo

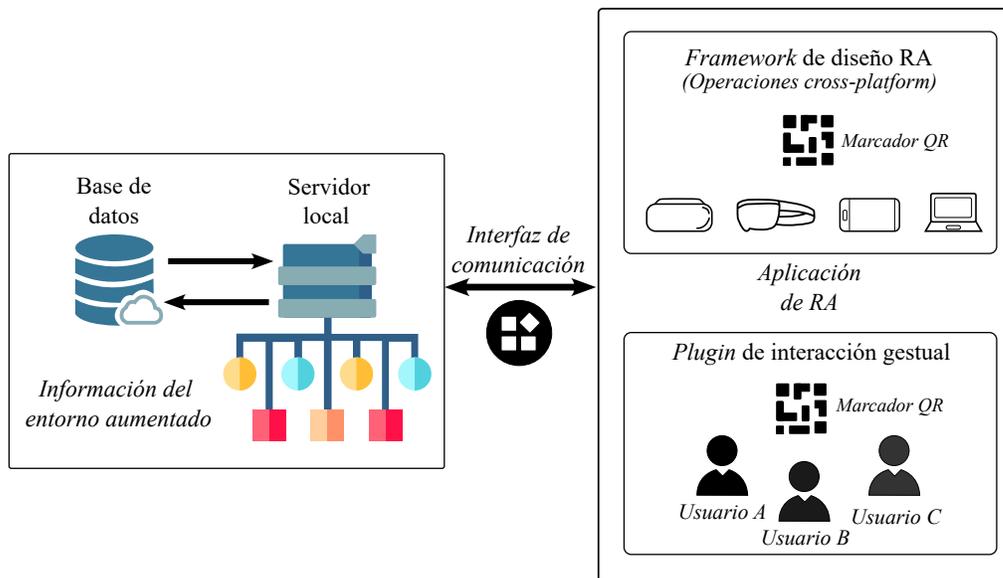


Figura 3.1: Arquitectura propuesta del entorno compartido de RA

Como sistema de solución propuesto, se construyó un sistema de RA colaborativo bajo una arquitectura cliente-servidor en el que se incorpora el diseño de una aplicación de realidad aumentada, un *framework* de diseño RA, un *plugin* de interacción gestual, un

servidor local, y una base de datos como elementos principales, tal y como se ilustra en la Figura 3.1.

El sistema aborda una división entre la interacción de los usuarios con el entorno y su proceso de colaboración, siendo la aplicación de RA el módulo integrador entre ambas divisiones. La integración de un sistema con estos cuatro elementos permite obtener una solución modular a los problemas previamente planteados.

Por un lado, la aplicación de RA implementa un *framework* de diseño RA, mismo que permite reconocer marcadores para establecer marcos de referencia en el espacio aumentado, y un módulo de comunicación multiplataforma para compartir la información disponible en la aplicación con el resto de usuarios. En cambio, el servidor local funciona como intermediario y controlador en la difusión de información entre todos los usuarios del sistema que, al limitar el acceso a una cantidad finita y reducida de clientes, mejora su capacidad de respuesta.

3.3. Aplicación de RA

El sistema de RA colaborativo propuesto parte del diseño de una aplicación de realidad aumentada, es decir, el desarrollo de una aplicación que facilita la comunicación e interacción entre los usuarios y el entorno aumentado.

La aplicación de RA, diseñada utilizando el motor de videojuegos Unity, se compone de una interfaz de usuario que incluye botones digitales y funciones de reconocimiento de gestos táctiles, ambos necesarios para la manipulación, modificación e interacción dentro del entorno de RA. Funciona bajo un paradigma orientado a eventos, por lo que su diseño lógico ejecuta acciones cuando el usuario interactúa con algún elemento de la aplicación o el entorno aumentado.

Por su función en el envío de datos y procesamiento de los mensajes recibidos desde un servidor, además de la visualización gráfica del entorno bajo exploración, la aplicación cumple el rol de cliente dentro del sistema colaborativo y es pieza elemental de la plataforma propuesta.

3.3.1. Interacción gestual con Lean Touch

Para satisfacer la necesidad de interacción con el entorno aumentado, la aplicación de RA hace uso de Lean Touch, un *plugin* para el motor de videojuegos Unity. Lean Touch ofrece una extensa variedad de funcionalidades relacionadas a la integración de control con interfaces UI empleando gestos táctiles.

El *plugin* incluye todas las funciones, librerías y dependencias necesarias para la correcta manipulación e interacción de los gestos y la pantalla de los dispositivos. Se incluye, además, un repertorio de funcionalidades que son aprovechados para el desarrollo de la aplicación, listando las más representativas en la Tabla 3.1.

Funcionalidad	Descripción
Finger Tap	Contabiliza la cantidad de pulsaciones realizadas sobre la pantalla del dispositivo, así como la respectiva posición en pantalla desde donde se realizan.
Double	Permite la creación de objetos virtuales al presionar consecutivamente un mismo punto en la pantalla.
Finger Old	Muestra en pantalla la posición de la entidad que interactúa con el dispositivo, así como el estado y tiempo durante el cual se realizó el contacto del mismo.
Transform	Permite trasladar, rotar o escalar objetos 2D y 3D empleando gestos como <i>drag</i> , <i>twist</i> y <i>pinch</i> .
Tap to select	Permite realizar la selección de objetos al presionar sobre ellos y deseccionarlos cuando se presiona un elemento distinto al seleccionado dentro de la pantalla.
Destroy	Funcionalidad que permite destruir objetos virtuales en la escena cuando se presiona sobre ellos.
Finger trail	Dibuja la trayectoria que el puntero o entidad de contacto genera al interactuar con la pantalla del dispositivo. Después de un breve periodo de tiempo es destruido.

Tabla 3.1: Funcionalidades de Lean Touch empleadas en el sistema RA colaborativo

3.4. Framework de diseño RA

Aunque un motor de videojuegos facilita el diseño de una aplicación, por sí mismo no incluye las características y funcionalidades que una aplicación de RA requiere. En consecuencia, se hace indispensable el uso de un *framework* de diseño RA, tal como Vuforia o ARCore, para poder realizar la aumentación de la escena.

El *framework* de diseño RA cumple la función de dotar a la aplicación de características de reconocimiento espacial y puntos de referencia, como la identificación de marcadores fiduciarios o el sobre posicionamiento métrico de objetos virtuales en el entorno, por ejemplo.

3.4.1. Framework de RA Vuforia

Vuforia [21] es una plataforma utilizada para el desarrollo de RA, con soporte para la mayoría de teléfonos, tabletas y gafas existentes en el mercado. Permite a los desarrolladores añadir funciones avanzadas de visión por ordenador a aplicaciones Android, iOS y UWP (Por sus siglas en inglés, *Universal Windows Platform*) para la creación de experiencias de realidad aumentada que interactúan de forma realista con los objetos y el entorno físico.

Para el reconocimiento de elementos de referencia ofrece el seguimiento de una variedad de elementos y espacios como imágenes, objetos y entornos. Sin embargo, en el proyecto se emplea un elemento fiduciario como entidad de reconocimiento.

3.4.2. Reconocimiento de imágenes fiduciarias

World Center Mode, una de las distintas características que el *framework* de Vuforia proporciona, permite especificar el comportamiento de reconocimiento espacial. Esta funcionalidad permite especificar qué elemento se pretende considerar como marco de referencia global $\{W\}$, pudiendo ser: el dispositivo de inmersión, el primer objetivo detectado del entorno o un objetivo fiduciario específico.

Debido a que se utiliza una imagen fiduciaria como único objetivo de referencia en el sistema de RA, es posible simplificar el problema inicialmente planteado al considerarse como el origen del espacio aumentado para todos los usuarios del sistema colaborativo, sin embargo, se consideran algunas características importantes.

Dimensión real del objetivo: Se trata de un parámetro de configuración y relaciona la dimensión real que tendrá el objetivo con el espacio virtual. En otras palabras, permite especificar la dimensión del marcador desde el cual se realiza la descripción de pose global.

Posición estática del objetivo: El sistema se configura para trabajar con una imagen fiduciaria que se asume mantiene una posición estática, por lo que cualquier entorno con condiciones distintas a las previstas queda fuera del modelo de plataforma propuesto.

Cámara del dispositivo: No todas las cámaras son iguales. Algunas de ellas presentan diferentes tipos de lente o un zoom integrado que proporciona una captura distinta, por lo que su posición inicial en el espacio o la perspectiva del entorno puede no ser la misma que el resto de usuarios, incluso aunque todas ellas partan del reconocimiento del mismo objetivo fiduciario.

3.4.3. Características de la imagen

Para mantener un adecuado reconocimiento del espacio, los marcadores fiduciaros se diseñan utilizando el *software* Inkscape considerando distintas técnicas de optimización recomendadas por el *framework* de Vuforia, mismas que se enuncian a continuación:

Escala de grises: Las imágenes fiduciaras utilizadas presentan una escala en grises, es decir, una escala que se extiende entre blanco y negro. Esta escala permite identificar las características que se pueden utilizar para el reconocimiento y el seguimiento. El *framework* Vuforia menciona que, si la imagen posee un contraste general bajo, es probable que el marcador no sea una buena imagen objetivo al no presentar características aprovechables.

Contraste local: Las imágenes mantienen el diseño y estructura de un código QR con el que se mejora el contraste de la imagen al realzar características generales como: una asimetría geométrica, misma que permite identificar los marcadores desde cualquier dirección de observación, y la nitidez en las imágenes, con la que se acentúa el marcador al utilizar bordes rectangulares.

Enfoque de la cámara: La imagen fiduciaria se ha diseñado para mantener una posición estática con la dimensión y el espacio adecuado. Esto con la finalidad de garantizar las mejores condiciones de enfoque de la cámara que, de acuerdo a las recomendaciones del *framework* Vuforia, entre mejor enfoque por la vista de la cámara presente el marcador, los detalles del objetivo serán más sencillos de detectar. De lo contrario, la detección y el seguimiento pueden verse afectados negativamente.

Condiciones de iluminación: Las condiciones de iluminación pueden afectar significativamente a la detección y seguimiento de marcadores. Como solución, se establece un ambiente en interiores con suficiente luz para que los detalles de la escena y las características del marcador sean visibles en la vista de la cámara.

3.4.4. Descripción de pose en los elementos aumentados

La descripción de la pose implica la descripción de la posición y orientación de todos los elementos presentes en el espacio aumentado. En principio, para resolver el problema de perspectiva y escala métrica, se requiere conocer la pose de cada uno de los usuarios con respecto al origen del espacio $\{W\}$, es decir, con respecto al marcador fiduciario. Cuando se tiene su respectiva descripción, entonces se procede al envío de dicho elemento y a su representación como un objeto presente en el espacio colaborativo.

En el *framework* de Vuforia, los cálculos necesarios para su obtención se realizan de manera implícita empleando métodos de visión por computadora; métodos que se desconocen por la restricción privativa del *framework*. A pesar de ello, es posible obtener el

valor final de cada elemento empleando la lectura del componente de transformación, es decir, una entidad intangible a nivel de *software* del elemento deseado. El componente de transformación proporciona tres bloques de información: posición, rotación y escala.

Posición: La posición se almacena como un vector tridimensional (`Vector3`), es decir, una entidad matemática de la forma $\langle x, y, z \rangle$, y sus unidades se expresan en metros.

Rotación: La rotación se maneja por medio de cuaterniones, expresados en radianes, bajo un formato de punto flotante. Sin embargo, por comodidad y mayor comprensión, son transformados a ángulos de Euler, definidos en grados.

Escala: La escala, al igual que la posición, se almacena como un vector tridimensional (`Vector3`) y sus unidades se expresan en metros. Para la descripción de pose, no implica mayor relevancia en su modificación. No obstante, es un elemento clave para la funcionalidad interactiva dentro del espacio aumentado con el que se facilita al usuario ampliar o reducir la vista de los objetos virtuales.

3.5. Servidor

Como módulo de comunicación del sistema colaborativo de RA propuesto se diseña un servidor. El servidor actúa como controlador lógico del flujo de información entre cada uno de los dispositivos que se desenvuelve en el entorno aumentado; es un módulo intermediario que facilita la comunicación entre los usuarios. Dentro de sus funciones se encuentra la gestión en la creación de salas virtuales, la asignación de identificadores únicos a cada conexión y la modificación de registros sobre una base de datos.

Se rige por un paradigma orientado a eventos donde evalúa acciones definidas en mensajes serializados, por lo que su funcionamiento reconoce estructuras de información bien definidas. El servidor es capaz de procesar cuatro operaciones: ingresar a una sala colaborativa, actualizar la pose de un jugador, actualizar la pose de un objeto, desconectarse.

El servidor queda a la escucha de cualquiera de las instrucciones y realiza alguna acción en función a ella manteniendo la misma metodología durante cada operación: deserializar un mensaje, ejecutar la instrucción definida en el mensaje, registrar los cambios correspondientes en una base de datos, enviar la instrucción al resto de usuarios presentes en el entorno colaborativo.

3.5.1. Plataforma de desarrollo .NET

.NET [22] es una plataforma de desarrollo de software de código abierto, multiplataforma y gratuita utilizada para la creación de aplicaciones de software en una variedad de plataformas, incluyendo Windows, macOS y Linux.

La plataforma .NET, desarrollada por Microsoft, proporciona un entorno de ejecución común para los lenguajes de programación de Microsoft, incluyendo C#, el lenguaje de programación empleado para el desarrollo de la aplicación de RA. Sin embargo, su principal razón de uso dentro de este sistema se debe a la disponibilidad de su librería .NET Sockets, misma que proporciona una API (Por sus siglas en inglés, *Application Programming Interfaces*) para la comunicación en red.

Los sockets son una forma de comunicación en la que dos dispositivos se conectan para intercambiar datos a través de una red compatible en una amplia variedad de protocolos. Su facilidad de uso, y sus características que permiten el control de flujo y la detección de errores durante la comunicación, permite la creación de aplicaciones de red, como servidores web, clientes de correo electrónico, sistemas de mensajería o juegos multijugador.

3.5.2. Interfaz de comunicación

El envío de instrucciones desde el servidor hacia cada uno de los clientes del sistema de colaboración involucra dos elementos: la plataforma de desarrollo .NET y la lógica de control y gestión de eventos, incluida en la aplicación de RA. Esta interfaz de comunicación envía y recibe los mensajes serializados por la aplicación de RA, mientras comprueba y corrige la información corrupta o con presencia de errores en los mensajes. Así mismo, define la lógica que posibilita la comunicación multiplataforma al estructurar los mensajes en un formato y notación genérico y abierto JSON.

3.5.3. Servidor local

Como servicio de control, administración y comunicación de información de baja latencia entre los usuarios, se diseñó un servidor local, adecuado para la escala habitación en que se desenvuelve el sistema. El servidor fue diseñado empleando Node.js (motor de ejecución JavaScript asíncrono basado en eventos) [23] y programado en el lenguaje Typescript [24], para la creación de aplicaciones de red escalables.

La selección de Node.js se remite a facilitar el diseño lógico requerido para la construcción de un servidor local. Al funcionar de manera asíncrona y por eventos, mantiene el modelo de diseño empleado en la aplicación de RA; procesar los mensajes que se reciben

por un canal de comunicación abierto, ejecutar las instrucciones y proporcionar respuesta al resto de los usuarios del entorno colaborativo.

El servidor local utiliza como complemento herramientas no privativas tales como Mongoose, que permite una solución basada en esquemas para el modelado de datos, y MongoDB, que permite la gestión de datos NoSQL.

La información se almacena por medio de estructuras JSON, es decir, estructuras de la forma:

```
{
  commad: Operación,
  playerID: ID,
  ...
  atributos: Atributos
}
```

donde se almacenan todos los atributos generados por la aplicación de RA. Con la estructura JSON, se construyen objetos de *software* que el servidor es capaz de operar con la información que contiene en su estructura de datos, siendo aquéllas las asociadas al atributo `command` de la estructura de mensaje.

Como se ilustra en la Figura 3.2, la primera función del servidor es mantenerse a la escucha de las peticiones entrantes por un puerto definido por un administrador, siendo el puerto 27017 el asignado por defecto. La primera petición que se recibe es un intento de conexión donde se reconoce el identificador de red del dispositivo (IP). El servidor genera un identificador único (`_id`) y lo envía al dispositivo que realiza su conexión. Así mismo, el servidor añade la información respectiva del dispositivo que se conecta a un espacio virtual denominado `Lista de usuarios`.

Cuando el dispositivo recibe el `_id` asignado por el servidor, el dispositivo genera y envía un nuevo mensaje con la instrucción: `SEARCH_ROOM`. El servidor, por su parte, al procesar el nuevo mensaje entrante `SEARCH_ROOM`, asocia la información del dispositivo con la ya existente en la `Lista de usuarios` y lo incluye en un espacio virtual definido como *lobby*, cuya capacidad está limitada a dos usuarios.

El servidor, una vez que finaliza de procesar una instrucción, se mantiene a la espera de una nueva petición de manera iterativa. Durante la espera de creación de *lobby*, es decir, cuando un segundo dispositivo solicita unirse a la `Lista de usuarios`, ninguno de los usuarios ya presentes en la lista es capaz de enviar ninguna otra petición. Es el servidor quien controla el momento de recepción de nuevas peticiones de los usuarios al enviar el mensaje `ROOM_CREATED` a cada uno de los usuarios en la lista.

Si la petición que recibe el servidor se conforma por el mensaje UPDATE POSE en alguna de sus dos variantes: UPDATE_PLAYER o UPDATE_OBJECT, el servidor hace uso de la herramienta Mongoose para la escritura de información en una base de datos NoSQL. Una vez almacenada dicha información, se comparte la operación al resto de usuarios registrados en la Lista de usuarios.

Por el contrario, si la instrucción que se recibe no coincide con alguna de las operaciones registradas, el servidor realiza una limpieza en el *buffer* de mensajes e ignora la última operación hasta recibir una nueva por parte de los dispositivos conectados. Finalmente, si la instrucción que se recibe es CLOSE, el servidor finaliza la comunicación del espacio colaborativo.

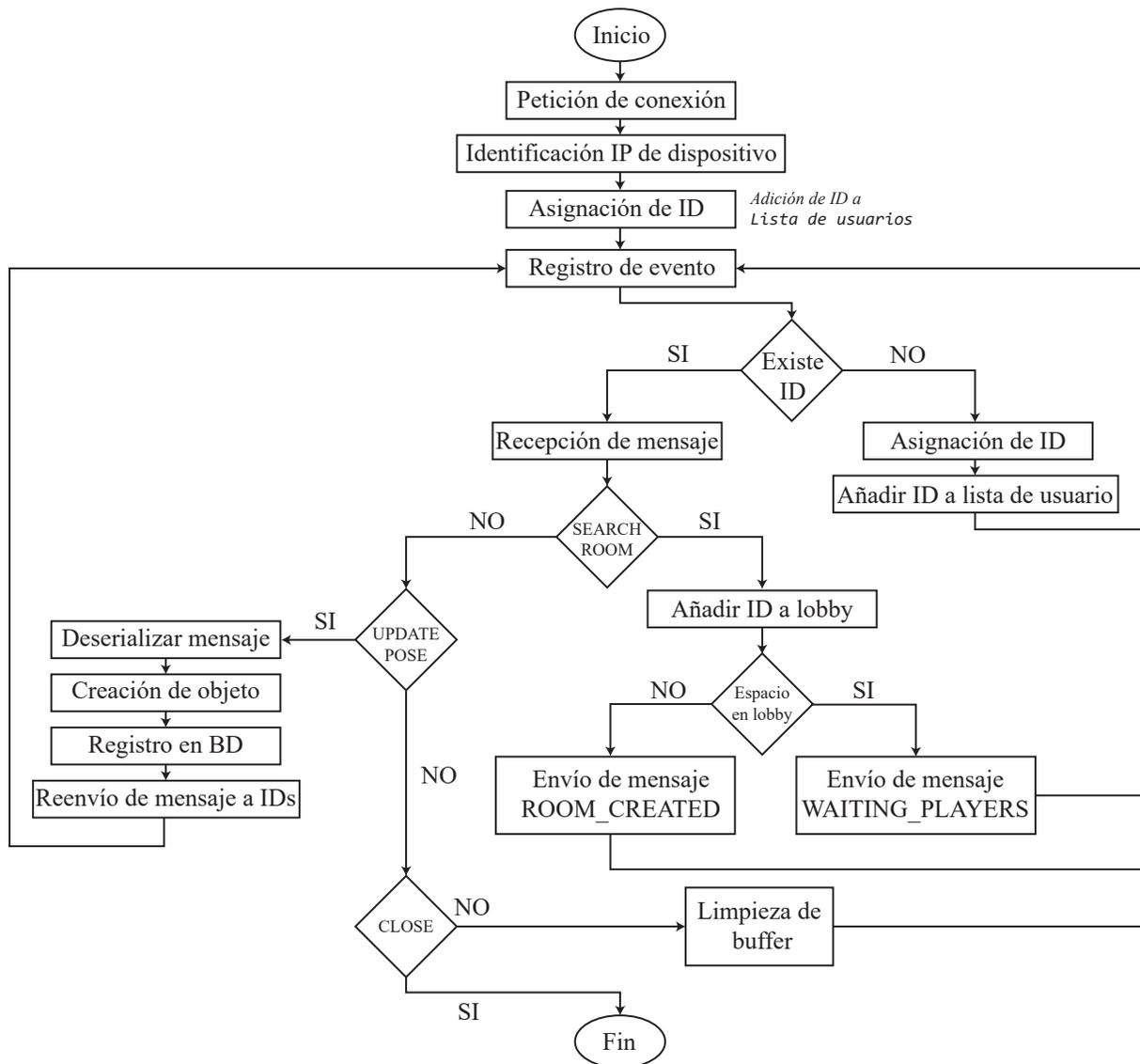


Figura 3.2: Diagrama de flujo del servidor

3.5.4. Envío de mensajes por Broadcast

El servidor crea una sala colaborativa cuando dos usuarios con `_id` diferentes se conectan empleando el mensaje `SEARCH_ROOM`. Como respuesta, el servidor envía una réplica a cada usuario de la sala con el mensaje `ROOM_CREATED`, habilitando a los usuarios la interfaz gráfica, la interacción con el entorno aumentado y la capacidad de enviar nuevos mensajes al servidor, es decir, funciona como iniciador del espacio aumentado colaborativo.

Sin embargo, el envío de este mensaje no se realiza de manera simultánea a todos los usuarios, por el contrario, se envía de manera secuencial en un proceso de *Broadcast*. Node.js facilita el envío de mensajes a distintos usuarios conociendo el identificador de red del dispositivo. Empleando esta funcionalidad y asociando la **Lista de usuarios**, el envío de mensajes se realiza en función del orden en que fue ingresado el primer usuario, es decir, con una estructura FIFO (Por sus siglas en inglés, *First In, First Out*).

3.5.5. NetCat

NetCat es un software desarrollado en la década de 1990 que proporciona una extensa variedad de funciones relacionadas a la comunicación TCP/IP. Con la finalidad de reducir los tiempos de respuesta en el proceso de desarrollo del servidor, fue empleado para asociar el envío de estructuras de datos serializados sin necesidad de ejecutar en tiempo real la simulación en el motor de videojuegos Unity o la aplicación desde el dispositivo móvil, en otras palabras, se pretendía ahorrar tiempo al enviar mensajes personalizados con los que fuese posible representar la comunicación entre los dispositivos.

3.6. Base de datos

Una base de datos es un conjunto de información almacenada y consultada sistemáticamente. Por el tipo de estructura en que organiza la información puede clasificarse como base de datos relacional (SQL) o base de datos no relacional (NoSQL).

El sistema de colaboración de RA propuesto emplea una base de datos no relacional que, integrada por 4 componentes, ofrece una mayor rapidez y descripción en el acceso a la información:

Datos: Palabras y números que por sí mismos no tienen significado, pero que en las estructuras de datos JSON de las instrucciones representan las propiedades de los objetos aumentados.

Campos: Área de almacenamiento de los datos con un tipo de dato específico. Se emplea para describir a las variables del entorno.

Registros: Colección de datos relacionados con igual o diferente tipo de dato. En el sistema existen dos registros: el registro de usuarios y el registro de objetos, cada uno corresponde a la descripción de las entidades presentes en el espacio colaborativo.

Archivos: Se trata de una colección de registros relacionados con estructura homogénea. Sólo existe un archivo y representa al espacio colaborativo.

3.6.1. Envío de datos

Los datos asociados a la interacción del entorno, tal como la creación y modificación de elementos aumentados, o la interacción entre el usuario con la aplicación de RA, se envían empleando *sockets* proporcionados por la plataforma .NET. No obstante, previo a su envío, la información es registrada bajo estructuras de datos de la forma `nombre:valor`, con la finalidad de facilitar su envío.

3.6.2. MongoDB

De acuerdo a la documentación oficial de MongoDB [25], MongoDB es una base de datos NoSQL de código abierto, escalable y de alto rendimiento diseñado para trabajar con datos no estructurados, por lo que es capaz de manejar datos complejos y dinámicos sin necesidad de esquemas predefinidos. Se caracteriza por almacenar datos en documentos JSON que pueden incluir campos y valores anidados.

MongoDB se utiliza en una amplia variedad de aplicaciones web, móviles y empresariales para almacenar y recuperar datos. Algunos ejemplos incluyen aplicaciones de comercio electrónico, sistemas de gestión de contenidos, aplicaciones de seguimiento de activos y aplicaciones de análisis de datos.

3.6.3. Esquemas Mongoose

Mongoose [26] es una herramienta que permite realizar operaciones a una base de datos de MongoDB a través de Node.js como controlador. Su entidad elemental son los esquemas, es decir, la representación de una colección de una base de datos NoSQL. Su implementación se debe a la facilidad que otorga en la ejecución de operaciones de consulta básicas como: crear, leer, eliminar y actualizar documentos de una base de datos.

Los esquemas son implementados como plantillas de estructuras de datos en las que se define cada uno de los atributos presentes en los mensajes que ingresan como peticiones al servidor. Cualquier atributo presente en un mensaje y no descrito en un esquema, el servidor se encarga de descartar el atributo no considerado.

Dentro del conjunto de atributos disponibles, es de principal interés el atributo `_id`, mismo que se emplea como llave de una colección. El `_id` de un mensaje, además de asociar el identificador único que asigna el servidor a un usuario, es un identificador único utilizado por MongoDB para el registro y actualización de documentos.

3.6.4. Serialización de mensajes

La estructura de un mensaje generado por la aplicación se define por un formato de texto JSON como estructura de intercambio de datos. JSON se basa en texto plano con codificación UTF-8, es ligero y tiene un formato de datos fácil de analizar sin necesidad de la creación de código adicional.

Al tratarse de una estructura universal e independiente al lenguaje de programación, permite la compatibilidad entre los distintos dispositivos empleados para el proceso de interacción aumentada. Así mismo, al desenvolverse en una aplicación con funcionalidades demostrativas de un entorno aumentado colaborativo elemental, la estructura de los mensajes se generaliza a dos estructuras de datos: estructura para la descripción de usuarios y estructura para la descripción de objetos.

3.6.5. Descripción de usuarios

La descripción de usuarios se resume a una estructura de datos para la descripción de pose, listada en la Tabla 3.2. Se integra por cuatro atributos: `command`, `_id`, `position` y `rotation`.

command: Especifica la operación que el usuario desea ejecutar sobre el espacio colaborativo, estando disponible las operaciones en cadena de texto: búsqueda de espacio colaborativo, asignación de estado de espera, creación de espacio colaborativo, actualización de pose y desconexión del espacio colaborativo.

_id: Identificador único universal UUID (Por sus siglas en inglés, *Universally Unique Identifier*). Se integra por un número entero pseudo-aleatorio de 128 bits (16 bytes) con el que se define y diferencia a cada usuario del espacio colaborativo. Debido a esta característica pseudo-aleatoria, este identificador se implementa como clave de

acceso en la base de datos del entorno colaborativo. Sin embargo, se trabaja como una cadena de texto para su proceso de operación.

position: Describe la posición del dispositivo con respecto al marcador fiduciario, siendo este el marco de referencia global. Se expresa en unidades del sistema internacional (SI), siendo el metro la escala base.

rotation: Describe la rotación del dispositivo con respecto al marcador fiduciario, siendo este el marco de referencia global. Se expresa en cuaterniones.

Nombre	Valor
<code>command</code>	Cadena de texto
<code>_id</code>	Cadena de texto
<code>position</code>	Vector3
<code>rotation</code>	Cuaternion

Tabla 3.2: Estructura de datos de pose

3.6.6. Descripción de interacción

La descripción de interacción, de manera similar a la descripción de usuarios, se resume a una estructura de datos para la descripción de pose y modificación de los objetos aumentados, mismos que quedan listados en la Tabla 3.3. Existe, empero, una integración de nueve atributos: `command`, `_id`, `position`, `rotation`, `scale`, `objectMesh`, `isSelected`, `playerCreator`, `playerEditor`.

command: Cumple la misma función que el atributo del mismo nombre presente en la estructura de pose presente en la estructura de usuarios, con la diferencia en que especifica la operación que el usuario desea ejecutar sobre un objeto del espacio colaborativo, estando disponible únicamente la operación **actualización de objeto** como una cadena de texto.

_id: Identificador único y global con el que se referencia a cada uno de los objetos colocados en el espacio colaborativo.

position: Describe la posición del objeto aumentado con respecto al marcador fiduciario, siendo este el marco de referencia global. Se expresa en unidades del sistema internacional (SI), siendo el metro la escala base.

rotation: Describe la rotación del objeto aumentado con respecto al marcador fiduciario, siendo este el marco de referencia global. Se expresa en cuaterniones.

scale: Describe la escala, tanto ampliación como reducción, del objeto aumentado. Se expresa en metros.

objectMesh: Define en una cadena de caracteres el nombre de la malla de un objeto aumentado, es decir, el conjunto de objetos tridimensionales, conformados por vértices que pueden unirse o dividirse, para proporcionar un aspecto de suavizado o arrugado a la perspectiva del usuario. Todos los objetos virtuales posibles de manipular en el entorno presentan las mismas funcionalidades interactivas, aunque visualmente el usuario los diferencia por la malla o forma que estos objetos presentan.

isSelected: Este atributo, delimitado por un estado lógico (activo o inactivo), describe el estado lógico de selección de los objetos del entorno aumentado. Sólo es posible la selección unitaria por jugador de cada objeto disponible en el espacio.

playerCreator: Almacena como cadena de texto el identificador único (`_id`) del usuario que crea un objeto dentro del entorno aumentado.

playerEditor: Describe como cadena de texto el identificador único (`_id`) del usuario que manipula un objeto existente dentro del entorno aumentado. Cumple la función de restringir una modificación simultánea entre un mismo objeto.

Nombre	Valor
<code>command</code>	Cadena de texto
<code>_id</code>	Cadena de texto
<code>position</code>	Vector3
<code>rotation</code>	Cuaternion
<code>scale</code>	Vector3
<code>playerCreator</code>	Cadena de texto
<code>playerEditor</code>	Cadena de texto
<code>objectMesh</code>	Cadena de texto
<code>isSelected</code>	Booleano

Tabla 3.3: Estructura de datos de interacción

3.7. Resumen del capítulo

En este capítulo se presentaron los tres principales problemas presentes en los entornos aumentados colaborativos: definición y/o establecimiento de un marco de referencia global, protocolos de comunicación, perspectiva y escala métrica. Así mismo, se describió la plataforma de colaboración aumentada propuesta.

Fueron presentadas las partes modulares del sistema colaborativo y su respectiva función dentro del mismo. Se describió a la aplicación de RA como elemento integrador del sistema al ser la unidad encargada de registrar los eventos interactivos entre el usuario y el entorno aumentado, así como su respectiva tarea de compartir las operaciones realizadas por el usuario que lo manipula.

Posteriormente, se procedió con la exposición del diseño de un servidor y el modelo de comunicación utilizado, sus respectivas tecnologías empleadas para su funcionamiento y la metodología de operación bajo la cual se rige.

Por último, se presentaron las estructuras de datos que permiten la comunicación multiplataforma del sistema y aquellas herramientas que permiten mantener el registro de las operaciones dentro del entorno aumentado colaborativo.

En el siguiente capítulo se presentarán los resultados obtenidos del sistema de realidad aumentado colaborativo implementado, mismo que se construye empleando la metodología y tecnologías descritas en el presente capítulo.

Capítulo 4

Pruebas y resultados



En este capítulo se describe el protocolo de pruebas realizado, los recursos disponibles para evaluar la plataforma de realidad aumentada colaborativa propuesta y se presentan los resultados obtenidos en el proceso de evaluación. Finalmente, se concluye con un análisis de resultados, así como los criterios de aceptación para determinar si la plataforma satisface los objetivos propuestos.

4.1. Protocolo de pruebas

El presente protocolo de pruebas fue diseñado con el objetivo de evaluar el desempeño de la plataforma de realidad aumentada colaborativa propuesta. La plataforma permite a los usuarios interactuar en un ambiente de trabajo compartido, combinando elementos virtuales y reales, de manera síncrona y entre distintos dispositivos.

Para evaluar la efectividad y eficiencia de la plataforma, se eligió un escenario estático de escala habitación. Los aspectos que fueron seleccionados para la evaluación fue la precisión de seguimiento de los objetos y la latencia de los dispositivos en cada una de las operaciones disponibles por el servidor local. La relevancia de estos aspectos se fundamenta en la usabilidad de la aplicación. Como describe Kostov y Wolfartsberger [14], la colaboración en espacios compartidos con baja latencia y responsivos promueve una mejor aceptación de la misma.

En las siguientes secciones, se detallarán los procedimientos, objetivos y métricas de evaluación para, finalmente, presentar un análisis de resultados.

4.1.1. Descripción general de las pruebas

La aplicación de RA es el elemento integrador de la plataforma de realidad aumentada colaborativa propuesta. Como objetivo, se mantuvo determinar su latencia en la experiencia de colaboración en todas las operaciones disponibles en el entorno de RA.

Las pruebas consistieron en evaluar de manera iterativa el funcionamiento y funciones de la aplicación de RA. Su evaluación se realizó entre dispositivos móviles con sistema operativo Android y la plataforma universal de Windows.

Las acciones que se realizan en el sistema son: el ingreso a una sala colaborativa, la interacción con el entorno o la exploración de los usuarios, así como el asegurar un correcto funcionamiento entre los diferentes dispositivos y entornos. Para ello, se realizaron pruebas en un escenario de escala habitación recopilando los resultados y observaciones en una bitácora de resultados para su posterior análisis.

4.1.2. Configuración de las pruebas

Previo a la realización de pruebas, se efectuó una configuración del ambiente y los equipos utilizados para garantizar la validez y confiabilidad de los resultados. En ella, se ajustaron las configuraciones de red y se posicionaron en el espacio físico todos los dispositivos involucrados, con la finalidad de asegurar una adecuada sincronización y descripción espacial en tiempo real. Se realizó una revisión de las tareas que se evaluaron, mismas de las que se proporciona una explicación detallada a continuación.

1. Verificación de especificaciones técnicas: La aplicación requiere un espacio de almacenamiento de 82 Mb para su versión móvil y 250 Mb para su versión de escritorio. Un sistema operativo igual o superior a Android 8 o Windows 10, ambos con procesadores ARM de 32/64 bits en la versión móvil, o Intel de 32/64 bits o ARM de 64 bits para su versión de escritorio.
2. Verificación de la cámara: Se hizo indispensable el uso de una cámara de alta calidad con una resolución mínima de 720 px y soporte de gráficos OpenGL ES 3.x, para la versión móvil, y DirectX 11, para la versión de escritorio. Siendo utilizada en la versión de escritorio una cámara web Logitech C920 HD Pro, en contraste a las cámaras digitales integradas a los dispositivos móviles.

3. Instalación de la aplicación: Para la versión móvil se hizo indispensable garantizar los permisos de acceso a la cámara y al almacenamiento del dispositivo, mientras que, para la versión de escritorio, no se requirió mayores permisos de uso.
4. Acondicionamiento del espacio de prueba: Se aseguró una buena iluminación y un espacio suficientemente amplio para desplazarse e interactuar con el entorno. El espacio acondicionado fue una superficie sin obstáculos de $5m^2$ con incidencia de luz natural.
5. Configuración de una red inalámbrica: Se creó y configuró una red local de acceso cableado e inalámbrico que permitió mantener una conexión entre los dispositivos involucrados en el entorno colaborativo.

4.2. Recursos utilizados

El conjunto de recursos utilizados durante el proceso de evaluación son listados a continuación. Los recursos son organizados en dos tipos: recursos de *hardware*, donde se abordan las especificaciones técnicas de aquellos dispositivos empleados para el proceso de inmersión, mismos que son agrupados en las Tablas 4.1 y 4.2, y los recursos de *software*, listados en la Tabla 4.3, que presentan el conjunto de herramientas digitales empleadas para el desarrollo de *software* de la plataforma propuesta.

La Tabla 4.1 lista las especificaciones técnicas correspondientes a las computadoras de escritorio utilizadas durante el proceso de evaluación. Se utilizaron tres equipos: una computadora personal empleada para ejecutar los servicios del servidor local (Computadora Tipo A), y dos computadoras de escritorio con las mismas especificaciones técnicas para evaluar la plataforma de realidad aumentada colaborativa (Computadora Tipo B).

Característica	Computadora Tipo A	Computadora Tipo B
Marca	Sin marca	HP
Modelo	Sin modelo	Alienware
CPU	Intel Core i3-8145 (2.1 GHz)	Intel Core i7-8700 (3.20 GHz)
GPU	Intel graphics	NVIDIA GeForce GTX 1080
Memoria RAM	8 Gb	16 Gb
Almacenamiento	128 Gb	600 Gb
Resolución de pantalla	1920 x 1080 px	2560 x 1440 px
Cámara	N/A	2.1 MP, f/2.0
Sensores	N/A	N/A
Sistema operativo	Manjaro Linux	Windows 10

Tabla 4.1: Características de las computadoras de escritorio utilizadas

En la Tabla 4.2, se listan las especificaciones técnicas de los dos dispositivos móviles empleados durante el proceso de evaluación del sistema colaborativo. Ambos dispositivos,

cada uno con distinta gama y recursos computacionales. El objetivo de esta selección fue tener una representación de aquellos dispositivos de mayor alcance comercial: un dispositivo de gama media (*smartphone A*) y un dispositivo de gama baja (*smartphone B*).

Característica	Smartphone A	Smartphone B
Marca	Xiaomi	Motorola
Modelo	Redmi Note 10	Moto E6
CPU	Snapdragon 678 (2.2 GHz)	Mediatek MT6739 (1.5 GHz)
GPU	Adreno 612	PowerVR GE8100
Memoria RAM	6 Gb	2 Gb
Almacenamiento	128 Gb	32 Gb
Resolución de pantalla	1080 x 2400 px	1440 x 720 px
Cámara	48 MP, f/1.8	13 MP, f/2.2
Sensores	Acelerómetro, giroscopio	N/A
Sistema operativo	Android 12	Android 9

Tabla 4.2: Características de los dispositivos móviles utilizados

Herramienta	Descripción
Unity	Motor de videojuegos empleado para el desarrollo de la aplicación de RA y comunicación con el servidor local. Implementado en su versión 2020.3.29f1 con los <i>plugins</i> : LeanTouch, ARCore y Unity UI para añadir funcionalidades de interacción, realidad aumentada e interfaz gráfica respectivamente.
Vuforia SDK	<i>Software</i> de desarrollo y diseño de RA implementado para añadir características de reconocimiento espacial y puntos de referencia, como la identificación de marcadores fiduciaros o el sobre posicionamiento métrico de objetos virtuales en el entorno, utilizado en su versión 10.7.2.
Node.js	Se emplea con la finalidad de facilitar la construcción y programación de funciones del servidor local, tal como: leer y enviar mensajes. Utilizado en su versión LTS 18.12.1.
Visual Studio	Editor de código que facilita la integración de las herramientas de desarrollo y evaluación.
MongoDB	Sistema de bases de datos NoSQL en su versión <i>community</i> empleado para almacenar los objetos creados en la aplicación de RA.
Compass	Entorno gráfico de administración de las bases de datos diseñadas en MongoDB.
Mongoose	Librería para Node.js que facilita la realización de consultas a una base de datos de MongoDB.
Typescript	Lenguaje de programación seleccionado para programar las funcionalidades y operaciones del servidor local.
JSON	Notación de objetos empleada para el envío de mensajes.
.NET	Plataforma de desarrollo empleada para el control de mensajes entre el servidor local y la aplicación de RA.
Inkscape	<i>Software</i> de diseño vectorial empleado para el diseño de iconos, gráficos y marcadores de la aplicación de RA.

Tabla 4.3: Descripción de herramientas utilizadas durante el desarrollo del sistema propuesto

4.3. Despliegue de resultados

En esta sección se describen los resultados obtenidos durante el conjunto de pruebas realizadas. Se puntualizan los pasos específicos realizados durante el proceso de evaluación, así como la configuración y las acciones llevadas a cabo durante su respectiva ejecución.

4.3.1. Aplicación de RA

La aplicación de RA, tanto en su versión móvil como de escritorio, queda ilustrado en la Figura 4.1. Se divide en tres secciones principales: la zona interactiva, el menú de opciones y la barra de objetos de RA, también llamado *dock* de objetos RA.



Figura 4.1: Interfaz gráfica de la aplicación de RA colaborativa. Se muestra el diseño final de la aplicación RA vista desde un dispositivo móvil.

Zona interactiva La zona interactiva de la aplicación de RA es el área específica, reservada y adaptativa en la pantalla de la aplicación. Se muestra el contenido de realidad aumentada que es extraído del mundo físico por la cámara del dispositivo. El usuario puede interactuar con el contenido aumentado a través de toques, *clicks* o gestos táctiles como deslizar, pellizcar, y arrastrar objetos dentro de esta zona interactiva. De la misma manera, permite la exploración del usuario y la manipulación de los objetos de realidad aumentada, tal como rotar o desplazar los objetos aumentados. Es la zona de la aplicación que permite realizar la visualización del entorno.

Menú de opciones

El menú de opciones de la aplicación de RA proporciona al usuario una guía sobre el uso de la aplicación, así como la capacidad de abandonar una sala colaborativa o cerrar la aplicación. El menú de opciones es un menú desplegable que se extiende de manera vertical sobre la pantalla del dispositivo sin importar su orientación al ser pulsado, y se retrae al volverse a pulsar. Su acceso se mantiene a la vista del usuario en la parte superior izquierda de la pantalla y se superpone de manera general a cualquier objeto del entorno. Se buscó un diseño minimalista, así como mantener un mayor espacio libre en la zona interactiva.

Dock de objetos AR

El *dock* de objetos de la aplicación de RA es una barra de objetos plana y flotante que proporciona al usuario el acceso a los objetos disponibles de la aplicación. El *dock* se

mantiene fijo en la parte inferior central de la pantalla, tanto en su orientación vertical y horizontal, en el que se ofrece al usuario cinco objetos de realidad aumentada con los cuales interactuar o posicionar en el entorno aumentado: pirámide, esfera, cubo, cilindro y diamante. El usuario es capaz de seleccionar el objeto que se desea colocar en el entorno y modificarlo desde la zona interactiva con las operaciones listadas en la Tabla 4.4.

Gesto	Descripción
Doble pulsación	Creación del objeto seleccionado en la posición donde fue presionada la pantalla.
Pulsación	Selección o deselección del objeto aumentado. Si se pulsa sobre un objeto existente, el objeto se selecciona, siendo esta operación reconocible por el cambio de color del objeto y bloqueando cualquier otra interacción con el resto del entorno. En cambio, si se selecciona un objeto aumentado previamente seleccionado, satisface la operación de deselección.
Deslizar	Permite trasladar un objeto aumentado si previamente ha sido seleccionado, utilizando el dedo del usuario, en el caso de la versión móvil, o el <i>mouse</i> , en la versión de escritorio.
Pellizcar	Permite escalar el objeto aumentado utilizando un gesto de pellizco empleando dos de los dedos del usuario, en el caso de la versión móvil, o el <i>mouse</i> y la tecla CTRL de manera simultánea, en la versión de escritorio, sólo si previamente ha sido seleccionado el objeto aumentado.
Giro	Permite rotar un objeto aumentado utilizando un gesto de giro en cualquier sentido empleando dos de los dedos del usuario, en el caso de la versión móvil, o el <i>mouse</i> y la tecla CTRL de manera simultánea, en la versión de escritorio, sólo si previamente ha sido seleccionado el objeto aumentado.

Tabla 4.4: Gestos de interacción disponibles en la aplicación de RA

4.4. Exactitud y latencia en la descripción de pose de los dispositivos de RA

Con la finalidad de analizar la capacidad de reconocimiento espacial del sistema de RA colaborativo, se realizó una evaluación inicial colocando un marcador fiduciario de 17.5×17.5 centímetros al centro del espacio físico. Este marcador se presenta en la Figura 4.2, actuando como origen del entorno aumentado.

En esta prueba, se hizo uso de cuatro dispositivos: dos computadoras de escritorio con

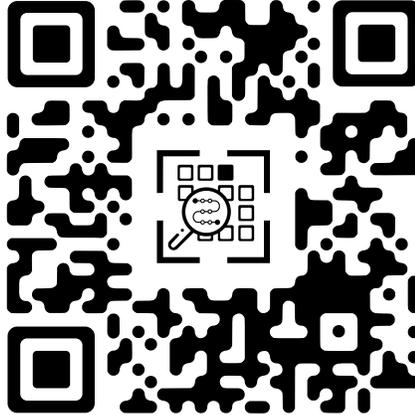


Figura 4.2: Marcador fiduciario utilizado por la aplicación de RA para la definición del marco de referencia global

iguales condiciones y características y dos dispositivos móviles inteligentes (*smartphone*) de diferente gama. Cada uno de los dispositivos evaluados, cuya posición y orientación fue conocida, se mantuvieron ausentes de movimiento dentro del entorno. Posteriormente, con la finalidad de dar validez estadística al conjunto de resultados, se ejecutó la prueba 50 veces. Los resultados obtenidos permitieron evaluar el funcionamiento de la aplicación y su respectiva latencia durante la comunicación utilizando la red local con el servidor local previamente configurado.

La evaluación, cuyos datos se obtuvieron empleando la consola de desarrollo de la aplicación, consistió en comparar la pose real de los dispositivos con aquella obtenida empleando el *framework* Vuforia. Así mismo, se registró la latencia correspondiente a la actualización de pose que cada dispositivo presentó durante el conjunto de iteraciones.

La comparativa entre la descripción de pose de los dispositivos en el mundo físico y el calculado con el *framework* Vuforia queda ilustrado en la Figura 4.3 y 4.4. Por su parte, la Figura 4.3 corresponde a la latencia presente entre los dispositivos al actualizar la descripción de su respectiva pose, siendo las computadoras de escritorio aquellos dispositivos que presentaron una latencia promedio menor al resto de dispositivos.

En cambio, en la Figura 4.4, se ilustran dos conjuntos de resultados asociados a la precisión que tuvo cada dispositivo al aproximar su pose con el *framework* Vuforia comparada con la pose real. El gráfico 4.4a corresponde a los resultados obtenidos por la computadora personal A y el dispositivo móvil Moto E6 play (*smartphone* de gama baja), cuya posición real fue -0.25 m en dirección x , 0.55 m en dirección y , y 0 m en dirección z . En contraste, en el gráfico 4.4b se ilustran los resultados obtenidos por la computadora personal B y el dispositivo móvil Xiaomi Redmi Note 10 (*smartphone* de gama media), cuya posición real fue 0.45 m en dirección x , 0.55 m en dirección y , y 0 m en dirección z .

Finalmente, en la Tabla 4.5 y 4.6, se resumen los principales resultados obtenidos del

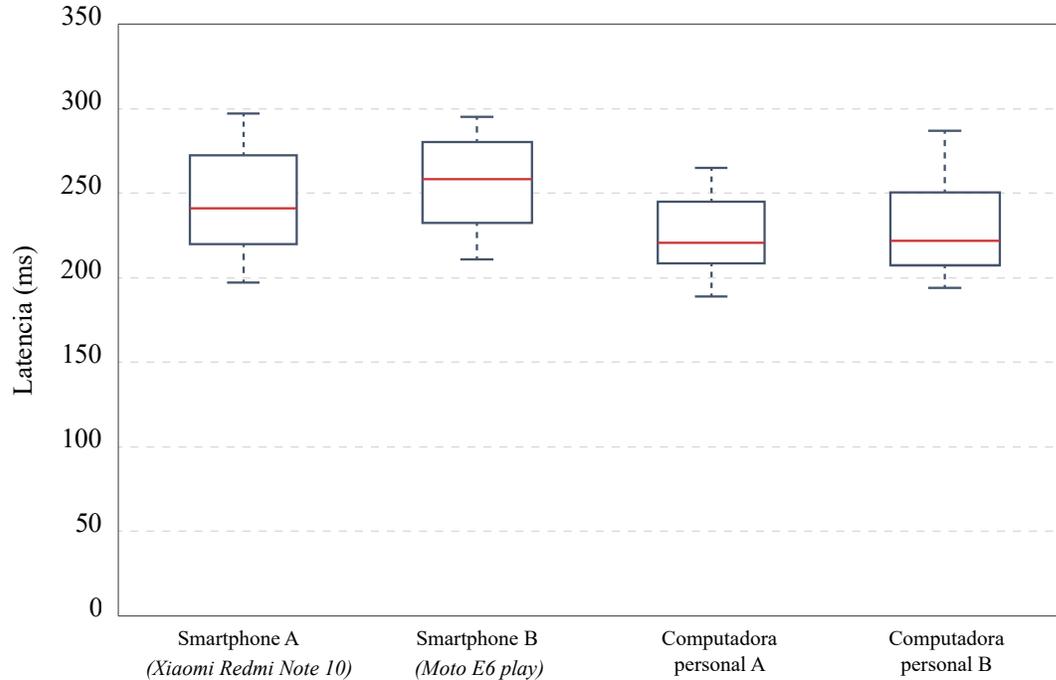


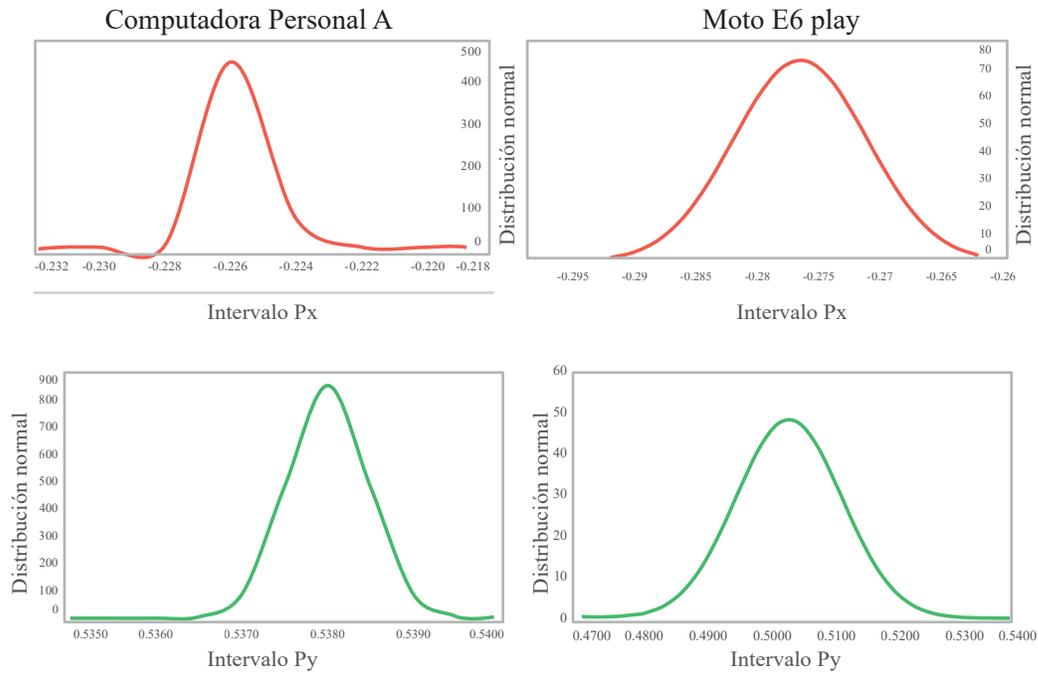
Figura 4.3: Latencia de los dispositivos de RA en la actualización de pose

conjunto de dispositivos evaluados en tiempo de respuesta y descripción de posición. En esta tabla, la columna \bar{x} representa la media aritmética de las pruebas realizadas respecto a posición. La columna Ex refleja el porcentaje de precisión en los resultados, mientras que la columna σ^2 representa la desviación estándar del conjunto de resultados obtenidos.

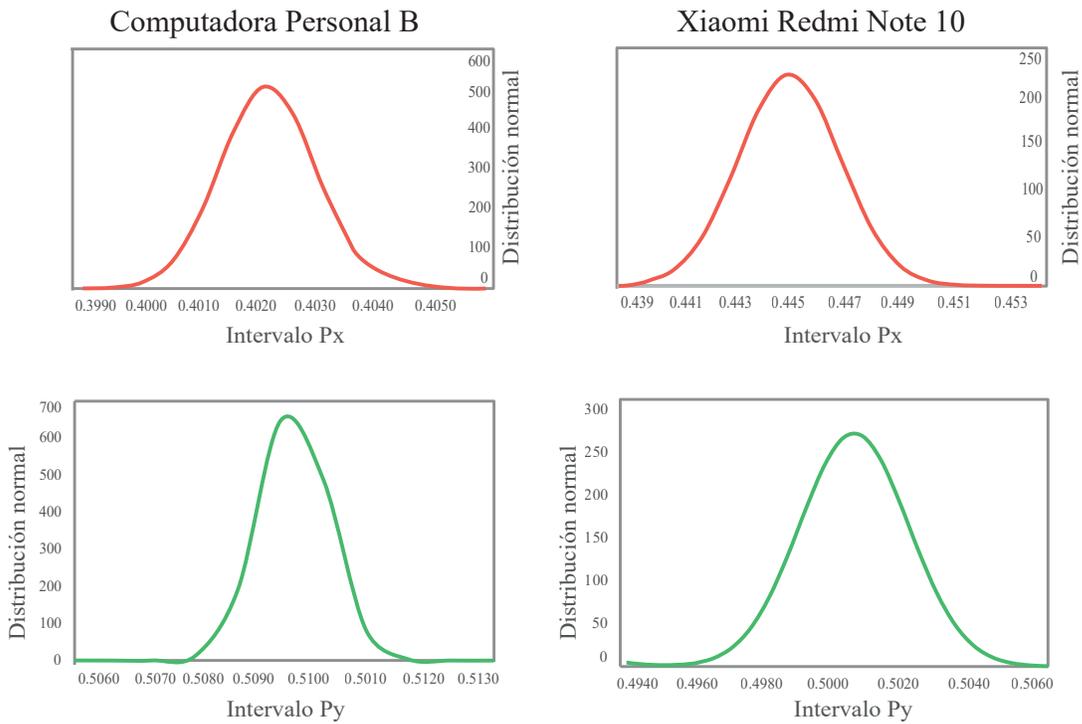
4.4.1. Latencia en búsqueda de una sesión colaborativa

Para la evaluación de latencia en operaciones asociadas al usuario, la colaboración se evaluó entre pares de la forma: computadora personal con computadora personal, computadora personal con *smartphone* y *smartphone* con *smartphone*. Cada par de dispositivos fueron conectados a una misma red local, sin embargo, los dispositivos móviles fueron conectados de manera inalámbrica a la red, mientras que las computadoras de escritorio se conectaron de manera cableada.

El análisis de latencia en búsqueda de una sesión colaborativa, es decir, la latencia que cada dispositivo presentó durante el proceso de conexión a una sala colaborativa, se realizó utilizando el servidor local desarrollado con Node.js. Cada dispositivo tuvo previamente configurado la dirección IP y el puerto de comunicación del servidor local, por lo que fue posible observar y registrar el tráfico de información y el respectivo tiempo que se demoró el servidor local en proporcionar una respuesta al dispositivo que efectuó la petición.



(a) Precisión del primer par de dispositivos RA en la descripción de pose



(b) Precisión del segundo par de dispositivos RA en la descripción de pose

Figura 4.4: Precisión en la descripción de pose de los dispositivos de RA

Latencia		
	Smartphone A	Computadora B
\bar{t} (ms)	245.32	229.7
σ^2 (ms)	29.804	26.905
Pose		
$x : 0.45$ m		
$y : 0.55$ m		
$z : 0$ m		
\bar{x} (m)	0.445	0.402
\bar{y} (m)	0.500	0.509
\bar{z} (m)	0.026	-0.015
μ_x (m)	0.0017	0.0007
μ_y (m)	0.0014	0.0005
μ_z (m)	0.0006	0.0006
Ex_x (%)	98.32	88.62
Ex_y (%)	90.59	92.41

Tabla 4.5: Resultados de latencia, precisión y exactitud en descripción de pose del primer par de dispositivos implementados

Latencia		
	Smartphone B	Computadora A
\bar{t} (ms)	254.94	225.64
σ^2 (ms)	26.108	22.142
Pose		
$x : -0.25$ m		
$y : 0.55$ m		
$z : 0$ m		
\bar{x} (m)	-0.276	-0.225
\bar{y} (m)	0.502	0.537
\bar{z} (m)	-0.003	-0.0018
μ_x (m)	0.0054	0.0007
μ_y (m)	0.0082	0.0004
μ_z (m)	0.0061	0.0008
Ex_x (%)	83.38	89.67
Ex_y (%)	87.82	97.66

Tabla 4.6: Resultados de latencia, precisión y exactitud en descripción de pose del segundo par de dispositivos implementados

En la Figura 4.5, se presenta una comparativa entre las latencias promedio, expresadas en milisegundos, correspondientes a cada uno de los dispositivos utilizados en las distintas pruebas realizadas.

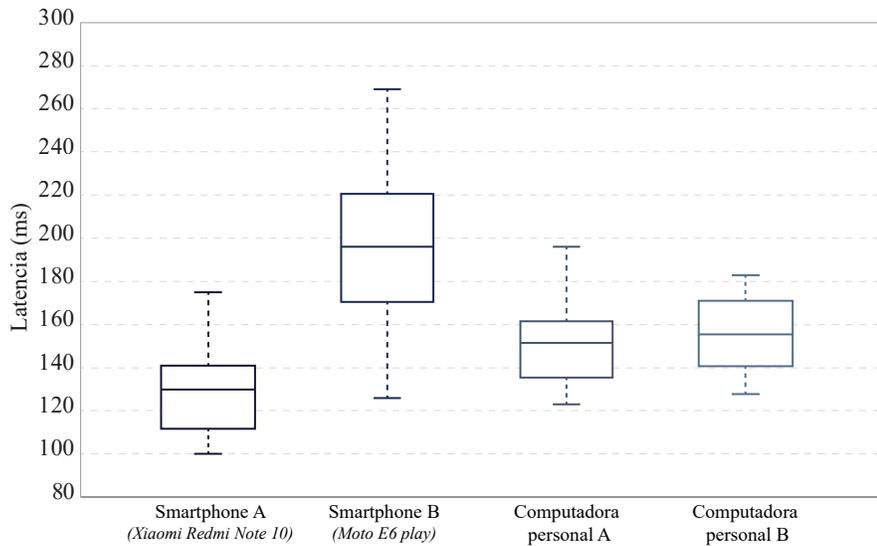


Figura 4.5: Latencias de los dispositivos utilizados en la búsqueda de espacio colaborativo

En contraste, en la Tabla 4.7, se puede observar que los resultados promedio de latencia oscilan entre 129 ms y 198 ms, donde el dispositivo Xiaomi Readmi Note 10 (*smartphone A*) presenta la latencia más baja del conjunto de dispositivos, mientras que el dispositivo Moto E6 play (*smartphone B*) presenta la latencia más alta del conjunto. Variación contemplada considerando la diferencia computacional requerida para el procesamiento de información entre los dispositivos móviles.

	Smartphone A	Smartphone B	Computadora A	Computadora B
\bar{t} (ms)	129.36	197.22	150.28	162.64
σ^2 (ms)	19.572	33.367	17.710	16.610

Tabla 4.7: Resumen de resultados en la búsqueda de sesión colaborativa

4.4.2. Latencia en creación de una sesión colaborativa

El objetivo de esta prueba fue analizar la latencia presente en el servidor local durante la creación de una sesión colaborativa ante las posibles combinaciones de los dispositivos disponibles, esto es, el tiempo de respuesta transcurrido desde el momento en que un segundo dispositivo solicita ser integrado a una sesión, hasta que el servidor local procesa la instrucción y envía un mensaje de creación de sesión a todos los usuarios presentes en el entorno.

Los resultados arrojan, entonces, una comparativa de los tiempos de respuesta del conjunto de pares de dispositivos evaluados, siendo que la latencia en este proceso queda definida en función de los recursos computacionales de los dispositivos, así como del orden en que los dispositivos solicitaron la conexión al servidor local.

En la Figura 4.6, se presentan los resultados de latencia durante la creación de una sesión colaborativa entre los tres pares de dispositivos evaluados.

En el primer par de dispositivos evaluados, correspondiente a una computadora personal Alienware y el *smartphone* Xiaomi Redmi Note 10, el dispositivo móvil presenta una latencia promedio de 151.02 ms, en contraste a la computadora personal, cuyo promedio fue de 158.1 ms, resultados promedio de cincuenta pruebas realizadas con este mismo par de dispositivos.

En el segundo par de dispositivos evaluados, correspondientes a los dos dispositivos móviles, el *smartphone* A ilustra la latencia presente en el *smartphone* Xiaomi Redmi Note 10, teniendo una latencia promedio de 166.06 ms, en cambio, el gráfico correspondiente a la latencia que presentó el *smartphone* B (Moto E6), tuvo una latencia promedio de 224.36 ms, siendo esta la más alta del conjunto de pruebas realizadas con este mismo par de dispositivos.

Finalmente, el último par evaluado corresponde a dos computadoras de escritorio Alienware. El gráfico asociado a la computadora personal A, presenta una latencia promedio de 152.74 ms. En cambio, el gráfico correspondiente a la computadora personal B, mantuvo una latencia promedio de 151.74 ms. Conjunto de resultados esperados en ambos dispositivos por presentar características y condiciones similares.

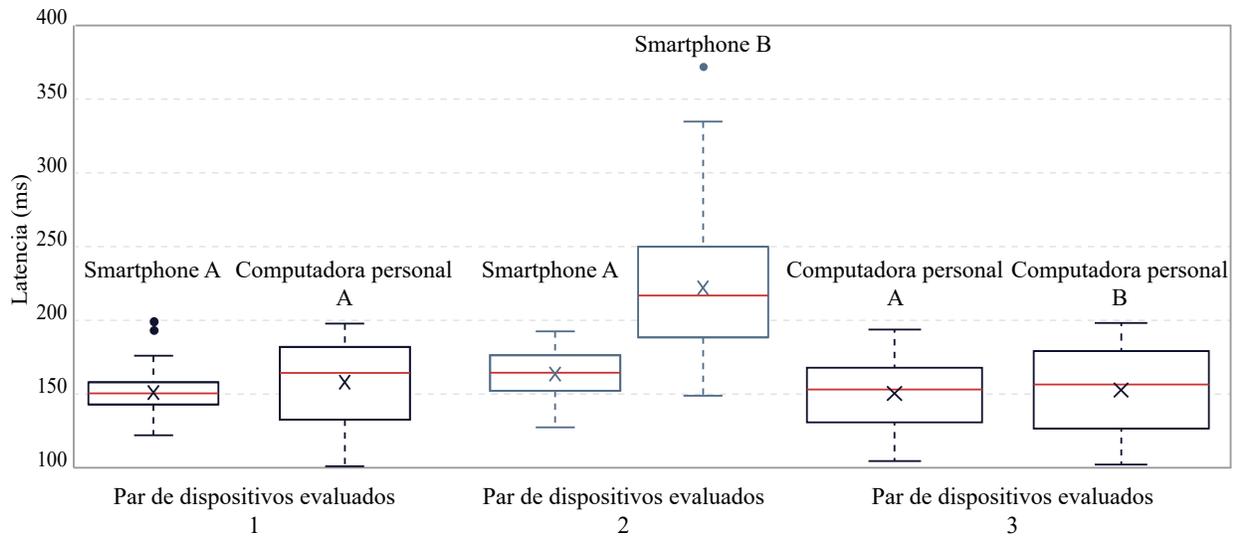


Figura 4.6: Comparativo de latencias entre el conjunto de dispositivos evaluados en la creación de un espacio colaborativo

4.4.3. Latencia y precisión durante la creación de objetos de realidad aumentada

En esta última prueba, utilizando la aplicación de RA, se realizó un registro de la latencia y precisión que presentó cada dispositivo durante el proceso de creación de objetos aumentados. Para ello, durante 50 pruebas realizadas en cada dispositivo que compartía un entorno aumentado, se colocó un objeto virtual en la esquina superior derecha del marcador fiduciario, esto es, una posición estratégica conocida del marcador. La evaluación, cuyos datos se obtuvieron empleando la consola de desarrollo de la aplicación de RA, consistió en comparar la pose real que deberían ocupar los objetos digitales con respecto a aquella obtenida empleando el *framework* Vuforia, así como el registro de la latencia correspondiente a la creación del objeto aumentado durante el conjunto de iteraciones.

La comparativa entre la latencia que presentó cada dispositivo se ilustra en la Figura 4.7. Así mismo, en la Tabla 4.8, se resumen los resultados cuantitativos correspondiente al tiempo de respuesta que se demoró cada uno de los distintos dispositivos durante la misma operación de creación de objetos de RA.

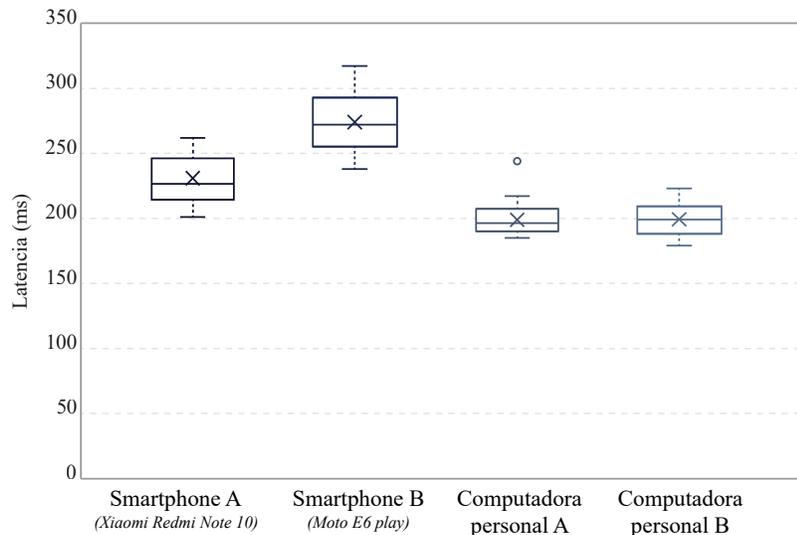


Figura 4.7: Latencias de los dispositivos utilizados durante la operación de creación de objetos de RA

	Smartphone A	Smartphone B	Computadora A	Computadora B
\bar{t} (ms)	230.88	274.06	198.94	199.4
σ^2 (ms)	19.231	23.218	11.777	12.833

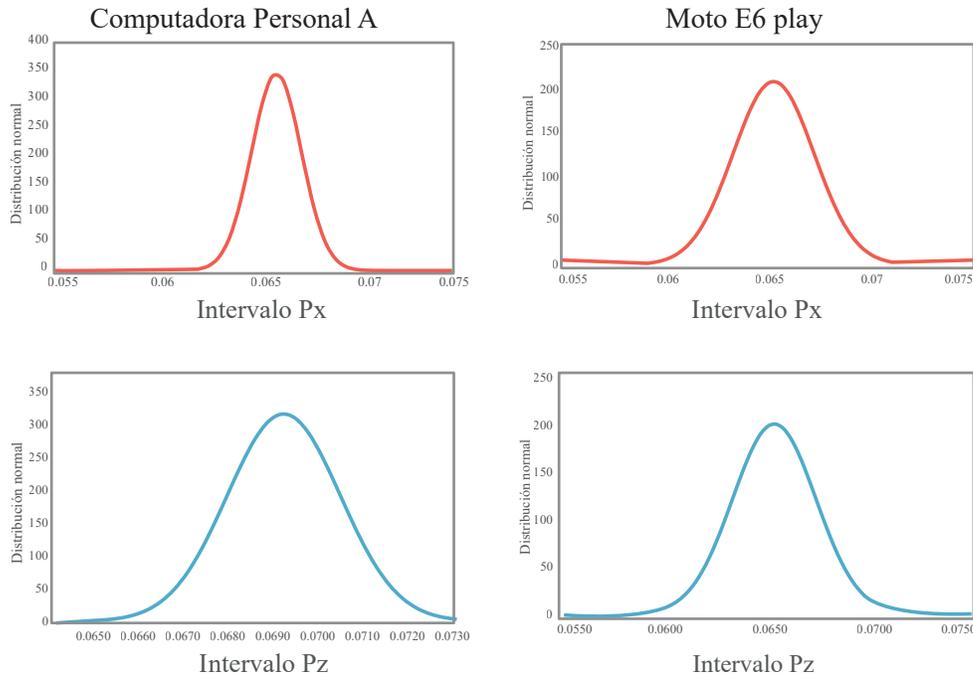
Tabla 4.8: Resumen de resultados de latencia durante la operación de creación de objetos de RA

En cambio, la comparativa entre la descripción de pose de los objetos aumentados calculado con el *framework* Vuforia queda ilustrado en la Figura 4.8 y, en la Tabla 4.9,

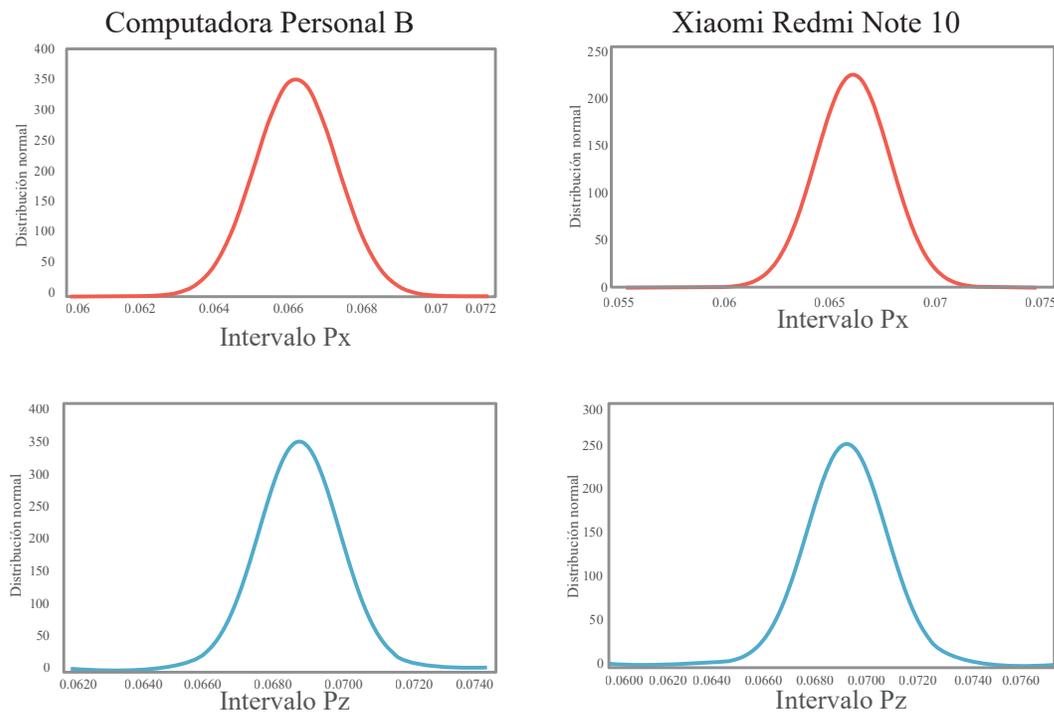
se resume de manera cuantitativa los resultados obtenidos por el conjunto de dispositivos bajo evaluación.

Computadora personal				
	\bar{x} (m)	Ex (%)	σ^2 (cm)	Real (m)
x	0.0657	90.078	0.0011	0.07
z	0.6924	95.047	0.0012	0.07
Xiaomi Redmi Note 10				
	\bar{x} (m)	Ex (%)	σ^2 (cm)	Real (m)
x	0.0661	87.664	0.0017	0.07
z	0.0697	94.221	0.0015	0.07
Moto E6 play				
	\bar{x} (m)	Ex (%)	σ^2 (cm)	Real (m)
x	0.0659	84.604	0.0024	0.07
z	0.0651	87.124	0.0019	0.07

Tabla 4.9: Resumen de resultados en la precisión de pose de objetos AR



(a) Precisión de los objetos aumentados instanciados por un primer par de dispositivos de RA



(b) Precisión de los objetos aumentados instanciados por un segundo par de dispositivos de RA

Figura 4.8: Precisión de los objetos aumentados instanciados durante la operación de creación de objetos de RA

4.5. Análisis de resultados

Como análisis de resultados de esta tesis, se evaluaron dos métricas fundamentales para el rendimiento de la plataforma; la latencia y la precisión de los objetos aumentados en el conjunto de operaciones disponibles. Los resultados mostraron que la plataforma tiene una latencia promedio de 238.9 ms en operaciones interactivas, en contraste a los 500 ms inicialmente propuestos, así como una exactitud del 83.389 % con una precisión de ± 0.0394 m en la descripción métrica de la superposición de objetos virtuales con respecto al mundo real.

Así mismo, el análisis de los resultados obtenidos revela una serie de hallazgos significativos. En primera instancia, el conjunto de resultados obtenidos en las pruebas realizadas indican que la plataforma desarrollada es capaz de proporcionar una experiencia de realidad aumentada fluida en tiempo real, permitiendo a los usuarios colaborar de manera efectiva en el mismo entorno aumentado sin retrasos relevantes.

En segundo lugar, se ha concluido que la plataforma es adaptable a diferentes entornos y requisitos computacionales, lo que la convierte en una herramienta versátil y dinámica entre los dos principales sistemas operativos en el mercado; Windows y Android.

De la misma manera, la estructura de datos serializable propuesta permite considerar a la plataforma como un sistema altamente compatible en una amplia gama de dispositivos, lo que la hace accesible para un gran número de usuarios al reducir las barreras de entrada para su implementación en diferentes contextos.

Finalmente, en tercer lugar, se ha demostrado que la arquitectura flexible y modular de la plataforma ofrece las herramientas básicas que permiten a los desarrolladores personalizar y ajustarla a sus necesidades específicas.

Capítulo 5

Conclusiones y perspectivas



En este capítulo, se presentan las conclusiones clave que se han obtenido de la investigación, se aborda la contribución que se realiza a la literatura, se presentan los hallazgos para el diseño y desarrollo de aplicaciones de RA y se concluye con una discusión sobre futuras investigaciones desde su implementación en distintos campos.

5.1. Conclusiones

En el presente trabajo se desarrolló un sistema de realidad aumentada colaborativa que permite la interacción y colaboración entre dos usuarios en entornos compartidos de realidad aumentada.

Se comenzó con una revisión exhaustiva de la literatura existente sobre realidad aumentada colaborativa, explorando los conceptos clave, las tecnologías utilizadas y los desafíos asociados a este campo. A partir de esta revisión, se identificaron las oportunidades y características requeridas para el desarrollo de un sistema que proporcione una experiencia de usuario fluida y una colaboración efectiva. El diseño del sistema prestó especial atención a la creación de una interfaz gráfica intuitiva, que permitió una fácil identificación y comunicación entre los usuarios, así como la manipulación de objetos aumentados de manera natural y precisa.

La implementación del sistema se llevó a cabo utilizando herramientas y tecnologías de desarrollo de realidad aumentada ampliamente utilizadas, como Unity y Vuforia. Se estableció una estructura de datos serializable que facilitó la colaboración multiplataforma reduciendo la latencia para garantizar una experiencia de usuario sin interrupciones. El sistema se sometió a pruebas rigurosas para evaluar su rendimiento y eficacia.

Los resultados de la investigación muestran que la plataforma se desenvuelve fluidamente en entornos de trabajo local, mismos que permitirían mejorar la productividad en los proyectos colaborativos dentro de esta misma escala.

Mediante la implementación de esta plataforma basada en herramientas de código abierto, se ha logrado proporcionar una experiencia colaborativa en tiempo real. La plataforma consiguió ofrecer experiencias de realidad aumentada interactivas que facilitan la comprensión conceptual de los principios de un entorno colaborativo aumentado.

La interfaz gráfica intuitiva que facilita la interacción natural entre los usuarios, mejorando así la inmersión y la colaboración en el entorno aumentado, es un resultado clave. Además, su arquitectura cliente-servidor proporciona la ventaja de permitir la conexión y colaboración entre diferentes dispositivos, ampliando las posibilidades de uso y la accesibilidad para un mayor número de usuarios.

Otro aspecto destacado es la compatibilidad de la plataforma, misma que amplía las oportunidades de implementación en diversos contextos y aplicaciones. Así mismo, se resalta la capacidad de personalización y ajuste de la plataforma mediante su arquitectura modular. Los desarrolladores tienen la posibilidad de adaptarla a sus necesidades específicas, fomentando así la innovación y la creación de experiencias de realidad aumentada personalizadas. Esto ofrece un potencial significativo para el avance y la evolución continua de la plataforma, a medida que se satisfacen las necesidades cambiantes de los usuarios y los requerimientos de los diferentes dominios de aplicación.

Finalmente, la plataforma presentada demostró un gran potencial, identificando puntos de mejora y desarrollo para futuras investigaciones, como la mejora en la reducción de latencia, el aumento de la capacidad de usuarios interconectados simultáneamente en un mismo espacio y la optimización en la gestión de los espacios colaborativos. Ante ello, se consiguió la satisfacción de los objetivos propuestos demostrando el valor del sistema de realidad aumentada colaborativa desarrollado.

5.2. Perspectivas

El campo de la realidad aumentada sigue en constante evolución y mejora, por lo que la plataforma desarrollada en esta tesis podría ser considerada como una de muchas otras referencias en futuros trabajos e investigaciones dentro del campo de la colaboración

en entornos aumentados. No obstante, se mantiene como expectativa que la plataforma propuesta pueda ofrecer una forma más natural, efectiva y generalizada de interactuar con objetos aumentados en tiempo real y colaborar con otros usuarios o miembros de un equipo de trabajo.

Se espera que esta investigación sea capaz de extenderse al desarrollo de aplicaciones en diferentes escenarios, lo que podría ayudar a evaluar y sugerir la viabilidad y utilidad de la plataforma para diferentes situaciones y áreas como en el campo de la educación, donde los estudiantes utilicen la plataforma para trabajar y explorar conceptos y proyectos complejos en un entorno inmersivo, con la finalidad de mejorar la comprensión y retención del contenido de aprendizaje.

Se contempla que esta plataforma se complemente con características que ofrezcan herramientas para el desarrollo y personalización de experiencias de realidad aumentada como:

Aumento en capacidad de usuarios: Esto considera una evaluación de la eficacia al aumentar la capacidad límite de usuarios interconectados, así como una evaluación de su experiencia de uso.

Prácticas de colaboración: Exploración de aquellos factores y características que hacen que la plataforma de realidad aumentada colaborativa sea efectiva, sin obstaculizar o solicitar mayores capacidades de cómputo.

Reutilización de la plataforma: Revisión y modificación del diseño de la plataforma propuesta que permita su reutilización en proyectos con distintos enfoques.

Finalmente, esta investigación de tesis ha producido como resultado una plataforma de colaboración en entornos de realidad aumentada multiplataforma, por lo que se espera que la metodología y los resultados presentados aporten valor a la literatura actual al ser considerada como un modelo de colaboración simple, efectivo y replicable.

Bibliografía

- [1] J. Carmigniani, B. Furht, M. Anisetti, P. Ceravolo, E. Damiani, and M. Ivkovic, “Augmented reality technologies, systems and applications,” *Multimedia Tools and Applications*, vol. 51, pp. 341–377, 1 2011.
- [2] M. Heilig, “Sensorama simulator.” Accedido: 2023-05-15, 8, 1962. Disponible: <https://image-ppubs.uspto.gov/dirsearch-public/print/downloadPdf/3050870>.
- [3] R. T. Azuma, “A survey of augmented reality,” *Caudell*, 1995.
- [4] Z. Zhang, “Microsoft kinect sensor and its effect,” *IEEE MultiMedia*, vol. 19, no. 2, pp. 4–10, 2012.
- [5] M. Sereno, X. Wang, L. Besancon, M. J. McGuffin, and T. Isenberg, “Collaborative work in augmented reality: A survey,” *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 10 2020.
- [6] L. López-Faicán and J. Jaen, “Emofindar: Evaluation of a mobile multiplayer augmented reality game for primary school children,” *Computers and Education*, vol. 149, 5 2020.
- [7] W. Zhang, B. Han, P. Hui, V. Gopalakrishnan, E. Zavesky, and F. Qian, “Cars: Collaborative augmented reality for socialization,” *HotMobile 2018 - Proceedings of the 19th International Workshop on Mobile Computing Systems and Applications*, vol. 2018-February, pp. 25–30, 2 2018.
- [8] P. Bhattacharyya, R. Nath, Y. Jo, K. Jadhav, and J. Hammer, “Brick: Toward a model for designing synchronous colocated augmented reality games,” *Conference on Human Factors in Computing Systems - Proceedings*, 5 2019.
- [9] S. Benbelkacem, N. Zenati-Henda, D. Aouam, Y. Izountar, and S. Otmane, “Mvc-3dc: Software architecture model for designing collaborative augmented reality and virtual reality systems,” *Journal of King Saud University - Computer and Information Sciences*, vol. 32, pp. 433–446, 5 2020.
- [10] P. Boonbrahm, C. Kaewrat, and S. Boonbrahm, “Effective collaborative design of large virtual 3d model using multiple ar markers,” *Procedia Manufacturing*, 2020.

- [11] D. Mourtzis, V. Siatras, J. Angelopoulos, and N. Panopoulos, “An augmented reality collaborative product design cloud-based platform in the context of learning factory,” *Procedia Manufacturing*, vol. 45, pp. 546–551, 1 2020.
- [12] A. Villanueva, Z. Zhu, Z. Liu, K. Pepler, T. Redick, and K. Ramani, “Meta-ar-app: An authoring platform for collaborative augmented reality in stem classrooms,” *Conference on Human Factors in Computing Systems - Proceedings*, 4 2020.
- [13] I. B. K. Manuaba, “Mobile based augmented reality application prototype for remote collaboration scenario using arcore cloud anchor,” *Procedia Computer Science*, vol. 179, pp. 289–296, 2021.
- [14] G. Kostov and J. Wolfartsberger, “Designing a framework for collaborative mixed reality training,” *Procedia Computer Science*, 2022.
- [15] D. C. Rompapas, C. Sandor, A. Plopski, D. Saakes, J. Shin, T. Taketomi, and H. Kato, “Towards large scale high fidelity collaborative augmented reality,” *Computers and Graphics (Pergamon)*, vol. 84, pp. 24–41, 11 2019.
- [16] B. Furht, ed., *Handbook of Augmented Reality*. Springer New York, NY, 1 ed., 2011.
- [17] J. Linowes and K. Babilinski, *Augmented reality for developers : build practical augmented reality applications with Unity, ARCore, ARKit, and Vuforia*. Packt Publishing, 2017.
- [18] S. Aukstakalnis, *Practical Augmented Reality*. Crawfordsville - Indiana: Mark L. Taub, 2016.
- [19] P. Milgram and F. Kishino, “A taxonomy of mixed reality visual displays,” *IEICE Transactions on Information Systems*, 1994.
- [20] B. Singhal, G. Dhameja, and P. S. Panda, *Beginning Blockchain*. Apress Berkeley, CA, 2018.
- [21] Vuforia, “Vuforia developer portal.” Accedido: 2023-01-15, 2021. Disponible: <https://developer.vuforia.com/>.
- [22] Microsoft Corporation, “.NET documentation.” Accedido: 2023-01-17, 2021. Disponible: <https://docs.microsoft.com/en-us/dotnet/>.
- [23] N. Foundation, “Node.js documentation.” Accedido: 2023-01-17, 2021. Disponible: <https://nodejs.org/en/docs/>.
- [24] M. Corporation, “Typescript.” Accedido: 2023-01-17, 2021. Disponible: <https://www.typescriptlang.org/>.
- [25] MongoDB, “Mongodb documentation.” Accedido: 2023-01-21, 2021. Disponible: <https://mongoosejs.com/>.
- [26] Mongoose, “Mongoose.” Accedido: 2023-01-21, 2021. Disponible: <https://mongoosejs.com/>.