

# Creación de Experiencias de Realidad Aumentada Realistas por Usuarios Finales

Álvaro Montero Montes

en cumplimiento parcial de los requisitos para el grado de  
Doctor en Ciencia y Tecnología Informática

Universidad Carlos III de Madrid

Directores:

Dra. Paloma Díaz Pérez

Dr. Telmo Zarranandia Ayo

Tutora:

Dra. Paloma Díaz Pérez

Leganés, enero de 2021

Esta tesis se distribuye bajo licencia “Creative Commons **Reconocimiento – No Comercial – Sin Obra Derivada**”.



A mi madre y a mi padre.

## AGRADECIMIENTOS

Cuando uno se halla en los últimos pasos del trayecto para la obtención del título de doctor y alza su mirada atrás puede recordar muchas personas que conoció y que, de alguna manera, directa o indirectamente, apoyaron la consecución de los propósitos de este libro.

Sin más demora, me gustaría expresar mi mayor agradecimiento a Telmo Zarraonandia y Paloma Díaz por los consejos, ayuda, comentarios y paciencia que me han brindado y que me ayudaron a guiarme en la senda de la investigación. Una vez que se llega hasta aquí uno es capaz de valorar aún más los consejos dados y cómo se deberían de haber seguido con una mayor consideración. También, agradecer al resto de compañeros del grupo de investigación de Sistemas Interactivos (DEI) por sus opiniones, comentarios y valoraciones. Puesto que me sería difícil no mencionarlos, los incluyo a continuación, Nacho Aedo, Andrea Bellucci, Lize de Coster, Judith Ley, Teresa Onorati, Ana Tajadura, Mónica Sánchez y Andrés Santos. Quiero hacer extensible estos agradecimientos a otros compañeros de la Universidad Carlos III de Madrid como Alejandro Baldominos, Silvina Caíno, German Dugarte, Sergio Jiménez, Sergio Martínez, Alejandro Quirós, Esther Salichs y Estefanía Serrano, entre otros.

Asimismo, por las innumerables veces que me han preguntado acerca del progreso de la tesis, me gustaría hacer mención en los agradecimientos a Irene Cuevas, Cristina Bernardo, Álvaro Dávila, Sara Carrión, Adrián Camino, Enrique García-Tizón, Pusp Lata, David Martín, Ylenia Ordóñez, Rubén Recio, María Rodríguez, Rocío Ruesga, Luis Miguel Sánchez, Raquel Sastre, Javier Silvero, Tomás Valdelomar, Virginie Veyret y Héctor Vidal. Por último, pero no por ello menos importante, quisiera incluir en los agradecimientos a Sara Ortiz con la que estoy en deuda por esos últimos empujones.

## CONTENIDOS PUBLICADOS Y PRESENTADOS

Algunas de las propuestas y resultados que se presentan en la siguiente tesis doctoral ha sido usados para la publicación de artículos en revistas y conferencias como parte de los requisitos exigidos por el programa de doctorado en Ciencia y Tecnología Informática. A continuación, se enumeran los artículos que fundamentan y presentan parte de los resultados de esta tesis:

- *Designing and Implementing Interactive and Realistic Augmented Reality Experiences*,  
Rol del doctorando: *primer autor*,  
Publicado en: *Universal Access in the Information Society*,  
Factor de Impacto: 1,176 (Q3),  
URL: <https://link.springer.com/content/pdf/10.1007/s10209-017-0584-2.pdf>,  
DOI: 10.1007/s10209-017-0584-2,  
Nota: *Este trabajo está parcialmente incluido en el capítulo 5 de la tesis*,  
Declaración: *El material de esta fuente incluido en la tesis podría no estar señalado por medios tipográficos ni referencias.*
- *An End-User Tool for Creating Augmented Reality Experiences*,  
Rol del doctorando: *primer autor*,  
Publicado en: *6<sup>th</sup> International Symposium on End-User Development; Working in Progress (IS-EUD 2017)*,  
URL: [https://pure.tue.nl/ws/files/69763287/IS\\_EUD2017\\_extended\\_abstracts.pdf#page=64](https://pure.tue.nl/ws/files/69763287/IS_EUD2017_extended_abstracts.pdf#page=64),  
Nota: *Este trabajo está parcialmente incluido en el capítulo 5 de la tesis*,  
Declaración: *El material de esta fuente incluido en la tesis podría no estar señalado por medios tipográficos ni referencias.*
- *Exploring the Benefits of Immersive End-User Development for Virtual Reality*,  
Rol del doctorando: *tercer autor*,  
Publicado en: *10<sup>th</sup> International Conference on Ubiquitous Computing & Ambient Intelligence (UCAMI 2016)*,  
URL: [https://doi.org/10.1007/978-3-319-48746-5\\_46](https://doi.org/10.1007/978-3-319-48746-5_46),  
DOI: 10.1007/978-3-319-48746-5,  
Nota: *Este trabajo fundamenta parte del capítulo 6 de la tesis*,  
Declaración: *El material de esta fuente incluido en la tesis podría no estar señalado por medios tipográficos ni referencias.*
- *Creating Interactive and Realistic Augmented Reality Experiences*,  
Rol del doctorando: *primer autor*,  
Publicado en: *Proceedings of the XVII International Conference on Human Computer Interaction (Interacción 2016)*,  
URL: <https://dl.acm.org/citation.cfm?id=2998660>,  
DOI: 10.1145/2998626.2998660,  
Nota: *Este trabajo está parcialmente incluido en el capítulo 5 de la tesis*,  
Declaración: *El material de esta fuente incluido en la tesis podría no estar señalado por medios tipográficos ni referencias.*

- *Immersive End User Development for Virtual Reality*,  
Rol del doctorando: *cuarto autor*,  
Publicado en: *Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI 2016)*,  
URL: <https://dl.acm.org/citation.cfm?id=2926067>,  
DOI: *10.1145/2909132.2926067*,  
Nota: *Este trabajo fundamenta parte del capítulo 6 de la tesis*,  
Declaración: *El material de esta fuente incluido en la tesis podría no estar señalado por medios tipográficos ni referencias.*
- *Agree or Perish: A Crowd Controlled Augmented Reality Pervasive Game*,  
Rol del doctorando: *primer autor*,  
Presentado en: *CHI 2016 Workshop on Pervasive Play*,  
Nota: *Este trabajo está parcialmente incluido en el capítulo 5 de la tesis*,  
Declaración: *El material de esta fuente incluido en la tesis podría no estar señalado por medios tipográficos ni referencias.*

## OTROS MÉRITOS DE INVESTIGACIÓN

Otras contribuciones realizadas durante el periodo de doctorado:

- *Using a Google Glass-based Classroom Feedback System to improve students to teacher communication*,  
Rol del doctorando: *tercer autor*,  
Publicando en: *IEEE Access*,  
Factor de Impacto (2018): 3,557 (Q1),  
URL: <https://doi.org/10.1109/ACCESS.2019.2893971>,  
DOI: 10.1109/ACCESS.2019.2893971
- *"Magic Flowerpot": An AR Game for Learning about Plants*,  
Rol del doctorando: *segundo autor*,  
Publicando en: *CHI Play 19'*,  
URL: <https://doi.org/10.1145/3341215.3356290>,  
DOI: 10.1145/3341215.3356290
- *A Toolkit for Creating Cross-Reality Serious Games*,  
Rol del doctorando: *cuarto autor*,  
Publicando en: *International Conference on Games and Learning Alliance (GALA 2018)*,  
URL: [https://doi.org/10.1007/978-3-030-11548-7\\_28](https://doi.org/10.1007/978-3-030-11548-7_28),  
DOI: 10.1007/978-3-030-11548-7\_28
- *A Platform for Creating Digital Educational Games as Combinations of Archetypical Games*,  
Rol del doctorando: *cuarto autor*,  
Publicando en: *5<sup>th</sup> International Symposium on End-User Development; Working in Progress (IS-EUD 2015)*,  
URL: [https://doi.org/10.1007/978-3-319-18425-8\\_34](https://doi.org/10.1007/978-3-319-18425-8_34),  
DOI: 10.1007/978-3-319-18425-8\_34
- *End User Development System for Adaptive Augmented Environments*,  
Rol del doctorando: *primer autor*,  
Publicando en: *5<sup>th</sup> International Symposium on End-User Development; Working in Progress (IS-EUD 2015)*,  
URL: [https://doi.org/10.1007/978-3-319-18425-8\\_27](https://doi.org/10.1007/978-3-319-18425-8_27),  
DOI: 10.1007/978-3-319-18425-8\_27

## ABSTRACT

The Augmented Reality (AR) technologies allow creating the illusions that virtual elements generated by computer, coexist with the rest of the real objects in the real world. In the last decades, distinct research studies have identified multiples benefits that this kind of technology can report in several areas, such as, medicine or education. However, despite of the appearance on the market of new AR devices more powerful and ergonomic, the use of AR applications is not as widespread as it might be expected.

There are many possible causes that might explain this situation. One of them is the lack of realism achieved by many current AR applications, which in many cases superimpose the virtual content on top of the users' vision of the environment without properly integrate it on it. Applications that deliver high quality AR experiences usually require specific hardware, not available to the general public due its high cost. Another problem that might conditions the adoption of this technology is the inherent complexity to the development of AR applications, which requires having advanced knowledge in programming, image processing and 3D modelling, among many others. In recent years, some toolkits have alleviated the problem allowing end users to create AR experiences on their own. However, these solutions are primarily focused on managing, configuring, and modifying the virtual content, and the level of realism achieved when integrating virtual objects into the real world is often very limited.

Taking these problems into account, the objective of this thesis is to facilitate the creation of realistic augmented reality experiences to users who might not have a high technical profile. In other words, to provide them with the means to create AR experiences in which virtual content appears integrated into the physical environment. More specifically, in the first part of this thesis a process for creating AR experiences and a system architecture for its implementation are proposed. The proposed approach allows to create AR experiences with a high level of realism but does not require specific AR hardware. Both the proposed process and the architecture were evaluated during the design and implementation of two AR experiences in which a large indoor space and an outside location were augmented realistically, respectively. These augmented environments served as scenarios for several collaborative AR games, which were tested by more than 100 people simultaneously in the first case and more than 60 in the second one.

In the second part of the thesis a method of authoring AR experiences for end users is presented. The proposed method uses Virtual Reality (VR) technology as the basis to provide users without a high technical knowledge with a simple way to augment a physical environment. The proposed method was evaluated through an experiment in which 10 participants designed the same AR scene using an authoring tool based on the proposed method (SimulAR), and another that implements the traditional mobile-based AR authorship approach (InsituAR). The results obtained suggest that the proposed method is equally effective when it comes to supporting the creation process and allows solving some limitations of the traditional AR-based approach.



## RESUMEN

La tecnología de realidad aumentada (RA) permite crear la ilusión de que elementos virtuales, generados por ordenador, coexisten con el resto de los objetos del mundo real. En los últimos años distintos estudios han identificado múltiples beneficios que este tipo de tecnología puede reportar en diversas áreas, por ejemplo, la medicina o la educación. Sin embargo, y a pesar de la aparición de nuevos dispositivos de RA cada vez más potentes y ergonómicos, la aplicación práctica de la RA no está tan extendida como cabría esperar.

Entre las posibles causas que pueden explicar la limitada implantación de este tipo de sistemas cabe destacar el bajo realismo conseguido por muchas de las aplicaciones de RA, que se limitan a la simple superposición de contenido virtual en la visión que el usuario tiene del entorno, sin integrarlo correctamente dentro del mismo. Si bien existen algunos sistemas que son capaces de conseguir experiencias de mayor calidad, estos suelen requerir del uso hardware específico que por su elevado coste no suele encontrarse disponible para el público general. Otro de los problemas que condicionan la adopción de esta tecnología es la complejidad inherente al desarrollo de aplicaciones RA, que requiere disponer de conocimientos avanzados en programación, procesamiento de imagen y modelado 3D, entre otros. En los últimos años han aparecido algunas herramientas que tratan de aliviar el problema y permiten a usuarios finales crear experiencias de RA por sí mismos. Sin embargo, estas soluciones se centran fundamentalmente en la gestión, configuración y modificación del contenido virtual, y el nivel de realismo alcanzado a la hora de integrar los objetos virtuales en el mundo real suele ser limitado.

Teniendo en cuenta estos problemas, el objetivo de esta tesis se centra en facilitar la creación de experiencias de realidad aumentada realistas a usuarios que no disponen de un perfil tecnológico alto. Es decir, se intentará proporcionar herramientas para que personas que tienen habilidades y conocimientos de tecnología a nivel de usuario puedan, bien por sí mismos o con una mínima asistencia técnica, crear experiencias de RA en las que el contenido virtual parezca integrado en el entorno físico.

De manera más concreta, en la primera parte de esta tesis se propone un proceso de creación de experiencias de RA que permite lograr un alto grado de realismo y no requiere de hardware específico, así como una arquitectura de sistema para su implementación. Tanto el proceso como la arquitectura propuestos fueron evaluados mediante el desarrollo de 2 experiencias de RA en las que se aumentaron de forma realista un espacio interior de grandes dimensiones y una localización exterior, respectivamente. Estos entornos aumentados sirvieron como escenarios a varios juegos de RA colaborativos, que fueron probados por más de 100 personas simultáneamente en el primer caso y más de 60 en el segundo.

En la segunda parte de la tesis se propone un método de autoría de experiencias de RA para usuarios finales. El método propuesto aprovecha las ventajas que ofrece la tecnología de Realidad Virtual (RV) con el fin de ofrecer a usuarios sin un alto grado de conocimiento técnico una forma sencilla de aumentar un entorno físico. El método propuesto fue evaluado por medio de un experimento en el que 10 participantes diseñaron una misma escena de RA empleando una herramienta de autoría basada en el método propuesto (SimulAR), y otra que implementa el tradicional enfoque de autoría basada en RA (InsituAR). Los resultados obtenidos sugieren que el método propuesto es igualmente eficaz a la hora de soportar el proceso de creación, y además permite solventar algunas limitaciones del tradicional enfoque basado en RA.

# TABLA DE CONTENIDO

Agradecimientos .....	iv
Contenidos Publicados y Presentados .....	v
Otros Méritos de Investigación .....	vii
Abstract .....	viii
Resumen .....	ix
Tabla de contenido .....	x
Índice de Tablas.....	xiii
Índice de Figuras.....	xiv
Lista de Abreviaturas .....	xvi
<b>1 Introducción.....</b>	<b>1</b>
1.1 Contexto.....	2
1.2 Objetivos de la Tesis.....	3
1.3 Metodología .....	4
1.4 Contribución.....	5
1.5 Estructura de la Tesis .....	6
<b>2 Estado de la Cuestión .....</b>	<b>7</b>
2.1 Algoritmo General de Ejecución de Realidad Aumentada.....	12
2.2 Evolución de las Tecnologías y Técnicas de Realidad Aumentada.....	13
2.3 Alineación de los elementos Virtuales.....	17
2.3.1 Tracking Mecánico .....	17
2.3.2 Tracking por Ultrasonidos .....	18
2.3.3 Tracking Electromagnético.....	19
2.3.4 Tracking Óptico .....	20
2.3.5 Tracking Inercial .....	22
2.3.6 Tracking Híbrido .....	23
2.4 Técnicas para la Mejora del Realismo en las Experiencias de Realidad Aumentada..	23
2.4.1 Efecto de Oclusión.....	23
2.4.2 Colisión .....	24
2.4.3 Gravedad .....	25
2.4.4 Luces y Sombras .....	25
2.4.5 Viento .....	30
2.4.6 Sonido.....	30
2.4.7 Adaptación de la Realidad.....	31
2.4.8 Tabla Resumen de los efectos y las técnicas.....	32
2.5 <i>End User Development</i> en el proceso de creación de Realidad Aumentada .....	32
2.6 Resumen.....	36

<b>3</b>	<b>Identificación del Problema y Objetivos .....</b>	<b>38</b>
3.1	Baja integración del contenido virtual en el mundo real en las experiencias de RA basadas en tecnología móvil .....	38
3.2	Dificultad para aumentar de forma realista entornos de grandes dimensiones o exteriores .....	38
3.3	Falta de herramientas de autoría adecuadas para la creación de experiencias de RA realistas por usuarios finales.....	39
3.4	Relación de los problemas identificados con los objetivos del trabajo de investigación	39
<b>4</b>	<b>Metodología de Investigación .....</b>	<b>41</b>
4.1	Ciclo de Relevancia.....	41
4.2	Ciclo de Rigor.....	43
4.3	Ciclo de Diseño .....	43
<b>5</b>	<b>Creación de Experiencias de Realidad Aumentada Realistas .....</b>	<b>45</b>
5.1	Método de creación de experiencias de RA realistas .....	45
5.1.1	Modelado del Entorno .....	46
5.1.2	Composición y alineación de la escena de RA.....	47
5.1.3	Simulación de las condiciones ambientales .....	49
5.1.4	Animación de Escena e Interacción con los Usuarios .....	50
5.2	Implementación del Prototipo de Motor de RA .....	51
5.2.1	GREP .....	51
5.2.2	AR GREP.....	52
5.3	Caso de Validación .....	52
5.3.1	Modelado del Entorno .....	53
5.3.2	Composición de la escena de RA.....	55
5.3.3	Simulación de las condiciones ambientales .....	55
5.3.4	Animación de Escena e Interacción con los Usuarios .....	56
5.3.5	Resultados Obtenidos .....	56
5.4	Propuesta de Arquitectura.....	57
5.5	Segundo Caso de Validación .....	58
5.6	Discusión .....	58
5.6.1	Caso Escenario Interior: Auditorio de la Universidad Carlos III de Madrid .....	58
5.6.2	Caso Escenario Exterior: Centro de Secundaria Colegio Virgen de Europa .....	60
<b>6</b>	<b>Autoría de Experiencias de Realidad Aumentada Basadas en Realidad Virtual .....</b>	<b>62</b>
6.1	Motivación .....	62
6.2	Objetivo y Preguntas de Investigación.....	62
6.3	Hipótesis.....	63
6.4	Herramientas de Autoría de Experiencias de RA Desarrollados para la Investigación	63
6.4.1	Herramienta de autoría de RA mediante tecnología realidad virtual (SimuLAR) .....	63
6.4.2	Herramienta de autoría RA mediante RA basada en móviles (InSituAR).....	68
6.4.3	Ejecutar la Experiencia de RA.....	70
6.5	Metodología .....	70
6.5.1	Estudio Piloto .....	70
6.5.2	Participantes y Dispositivos.....	73
6.5.3	Procedimiento .....	73

6.5.4	Instrumentos de Recolección de Datos.....	74
6.6	Resultados.....	75
6.7	Discusión.....	77
<b>7</b>	<b>Conclusiones, contribuciones y trabajos futuros.....</b>	<b>79</b>
7.1	Conclusiones.....	79
7.2	Contribuciones.....	80
7.3	Trabajos Futuros.....	81
<b>8</b>	<b>Bibliografía.....</b>	<b>82</b>
<b>Anexo A</b>	<b>.....</b>	<b>a</b>
<b>Anexo B</b>	<b>.....</b>	<b>b</b>
<b>Anexo C</b>	<b>.....</b>	<b>d</b>

## ÍNDICE DE TABLAS

Tabla 1.1 Problemas y Objetivos abordados en la tesis.....	3
Tabla 2.1 Ejemplos de sistemas optical tracking [43,55,76,111,197–201] .....	21
Tabla 2.2 Ventajas y Desventajas de las subcategorías de los optical tracking .....	22
Tabla 2.3 Tabla de los tipos de luces para entornos 3D.....	26
Tabla 2.4 Modelos de Iluminación sobre superficies.....	27
Tabla 2.5 Técnicas de simulación de sombras [46,146,163,205].....	27
Tabla 2.6 Proceso de Shadow Mapping [163].....	28
Tabla 2.7 Resumen de las diferentes técnicas empleadas para incluir cada uno de los efectos realistas .....	32
Tabla 5.1 Efectos realistas visibles en la experiencia de RA [110] .....	60
Tabla 6.1 Resumen de los parámetros de configuración de la cámara. ....	65
Tabla 6.2 Resumen de los gestos para la edición de los elementos virtuales .....	67
Tabla 6.3 Resumen de los gestos táctiles para la edición de un elemento virtual .....	70
Tabla 6.4 Resumen de resultados del NASA TLX.....	75
Tabla 6.5 Resumen de resultados del cuestionario con escala de Likert.....	76

## ÍNDICE DE FIGURAS

Figura 1.1 Intersección de las áreas de realidad aumentada y end-user development .....	2
Figura 1.2 DSISR contextualizado para en esta tesis doctoral .....	5
Figura 2.1 Adaptación del Virtual Continuum de Milgram [106] con ejemplos de cada uno de los niveles [5,24,52] . .....	7
Figura 2.2 Conjunto de tipo de aplicaciones o sistemas .....	8
Figura 2.3 Combinación del Continuum de Milgram y Weiser [113]. Combinación del Continuum de Milgram, Weiser y la multiplicidad de usuarios [22].....	9
Figura 2.4 Adaptación de la clasificación de espacios compartidos de acuerdo con las dimensiones de transportación and artificialidad [12] .....	9
Figura 2.5 Representación bidimensional de la Virtualidad-Mediada [101] y Jerarquía de conjuntos de la Realidad Mediada [101] .....	10
Figura 2.6 Ejemplo de combinación del uso de las tecnologías de realidad aumentada y realidad reducida [173] y adaptación del diagrama presentado por Mann [66].....	10
Figura 2.7 Adaptación de combinaciones de tecnologías mediadas basado en el diseño de Hillstead [66] .....	11
Figura 2.8 Diagrama de entrada y salida de un sistema de realidad aumentada.....	13
<i>FIGURA 2.9 LÍNEA TEMPORAL CON ALGUNOS DE LOS HITOS DE LA RA. EL GRÁFICO INCLUYE ALGUNAS TENDENCIAS DE LAS INVESTIGACIONES EN EL ÁREA DE LA RA.</i>	16
Figura 2.10 Adaptación de la taxonomía de Mark Billinghurst sobre Tracker de RA [102].....	17
Figura 2.11 La parte izquierda muestra el HMD de Sutherland suspendido del techo mediante un brazo [149]. En la parte derecha se presenta una instalación de unos prismáticos RA en el Canterbury Museum de NZ [196].....	18
Figura 2.12 La imagen izquierda muestra el HMD de Sutherland situado dentro de un espacio delimitado por cuatro receptores de ultrasonidos que están dispuestos en el techo [149]. La imagen de la derecha muestra los emisores-receptores de ultrasonidos del sistema KARMA que tenían una forma triangular [48].....	18
Figura 2.13 La imagen de la izquierda presente un diagrama de la distribución de generadores de ondas electromagnéticas y receptores del sistema construido por Caudell [28]. La fotografía central muestra el HMD de Caudell siendo utilizado para ensamblar cables[28]. La imagen de la derecha es un screenshot de una aplicación location-based RA [190].....	19
Figura 2.14 Fotografía de la aplicación AR Invaders [172] junto con un screenshot [202] .....	23
Figura 2.15 Renderización de una escena 3D en una imagen.....	26
Figura 2.16 Resumen de los pasos del proceso de Shadow Volumes [207] .....	28
Figura 2.17 Imagen de depuración. El toroide es un elemento virtual mientras que la lata es real. Sobre este último se ha incluido un phantom object para definir su forma [58].....	29
Figura 2.18 Recreación de la luz proveniente del cielo [80] .....	30
Figura 2.19 Estilización de la realidad para integrar la taza virtual en el entorno real [49] .....	31
Figura 2.20 Capturas del vídeo “Pokémon Go and Interactive Dynamic Video” sobre las que se incluyen flechas para indicar los cambios con la simulación de fuerza [210].....	32
Figura 2.21 Representación de la relación de cada rama dentro del área de end-user development .....	34
Figura 2.22 Línea Temporal con algunas herramientas de RA para usuarios finales.....	36
Figura 3.1 Relación de los problemas y los objetivos.....	40
Figura 4.1 Metodología de Hevner et al. contextualizado a la presente tesis.....	42
Figura 5.1 izquierda muestra una imagen de un entorno real a aumentar. La imagen derecha muestra la superposición de bloques para las zonas del entorno [110] .....	47
Figura 5.2 La imagen izquierda muestra la configuración básica de la escena virtual. La derecha muestra la composición de la escena con elementos virtuales y la réplica configurada [110]. ..	49
Figura 5.3 Visión que se tiene de la cámara virtual que graba la escena [110]. .....	49

Figura 5.4 Imágenes del proceso del proceso de incluir efectos realista en la escena. La imagen izquierda muestra la escena virtual mientras que la derecha ilustra el resultado que se puede observar desde la cámara virtual [110].....	50
Figura 5.5 Imagen capturada de la escena final de RA [110] .....	51
Figura 5.6 Planos del autorio.....	53
Figura 5.7 Imágenes del autorio a aumentar .....	54
Figura 5.8 Configuración del campo de visión de la cámara virtual .....	54
Figura 5.9 Bloques del pasillo mientras se ajustan al panel de fondo .....	54
Figura 5.10 Réplica del entorno construida utilizando bloques 3D .....	54
Figura 5.11 Escena RA una vez aplicada las texturas a los diferentes tipos de bloques y añadido los elementos virtuales de la escena .....	55
Figura 5.12 Escenario RA mientras se configura las luces para reproducir las del ambiente.....	56
Figura 5.13 Ejemplo del comportamiento definido de recolección de monedas en el fichero de XML.....	56
Figura 5.14 Instantánea del juego de RA de recolección de monedas [110].....	57
Figura 5.15 Instantánea del juego de RA del pillapilla [110].....	57
Figura 5.16 Arquitectura de sistema AR GREP Editor completa [109] .....	58
Figura 5.17 Captura de pantalla del proyector de la experiencia RA sobre un campo de fútbol	59
Figura 5.18 Captura de pantalla de la herramienta de medir distancias y superficies integrada en Google Maps .....	61
Figura 6.1 Captura de HMD mostrando la edición de un elemento virtual.....	66
Figura 6.2 Resumen de las funciones de los botones de los mandos.....	67
Figura 6.3 Captura de pantallas del dispositivo móvil .....	69
Figura 6.4 Resultados de las dimensiones de NASA TLX .....	75
Figura 6.5 Valoración de los usuarios en los diferentes enfoques.....	76
Figura 6.6 Comparación de los enfoques en función de los aspectos .....	77
Figura 7.1 Gráfico resumen de los problemas y su relación con los objetivos .....	80

## LISTA DE ABREVIATURAS

<b>AR/RA</b>	<i>Augmented Reality</i> (Realidad Aumentada)
<b>AV</b>	<i>Augmented Virtuality</i> (Virtualidad Aumentada)
<b>BRIEF</b>	<i>Binary Robust Independent Elementary Features</i>
<b>CSCW</b>	<i>Computer Supporting Collaborative Work</i> (Trabajo Cooperativo Asistido por Computadora)
<b>CPU</b>	<i>Central Processing Unit</i> (Unidad central de procesamiento)
<b>CRT</b>	<i>Cathode Ray Tube</i> (Tubo de rayos catódicos)
<b>DSISR</b>	<i>Design Science in Information System Research</i>
<b>EUD</b>	<i>End-User Development</i> (Desarrollo para usuarios finales)
<b>EUP</b>	<i>End-User Programming</i> (Programación para usuarios finales)
<b>EUSE</b>	<i>End-User Software Engineering</i> (Ingeniería de software para usuarios finales)
<b>FAST</b>	<i>Features from Accelerated Segment Test</i>
<b>FEM</b>	<i>Finite Element Method</i> (Método de Elementos Finitos)
<b>FoV</b>	<i>Field of View</i> (Campo de Visión)
<b>GPS</b>	<i>Global Position System</i> (Sistema de posicionamiento global)
<b>GPU</b>	<i>Graphic Processing Unit</i> (Unidad de procesamiento gráfico)
<b>ISO</b>	<i>International Organization for Standardization</i> (Organización Internacional de Normalización)
<b>LED</b>	<i>Light-emitting diode</i> (Diodo emisor de luz)
<b>HCI</b>	<i>Human-Computer Interaction</i> (Interacción Persona-Ordenador)
<b>HMD</b>	<i>Head-Mounted Display</i>
<b>HUD</b>	<i>Head-Up Display</i>
<b>MR</b>	<i>Mixed Reality</i>
<b>ORB</b>	Oriented FAST Rotative BRIEF
<b>PTAM</b>	<i>Parallel Tracking and Mapping</i> (Algoritmo de Tracking y Mapeo paralelo)
<b>PSP</b>	<i>PlayStation Portable</i>
<b>SIFT</b>	<i>Scale-invariant feature transform</i>
<b>SURF</b>	Speeded-Up Robust Features
<b>SLAM</b>	<i>Simultaneous Localisation and Mapping</i> (Localización y Modelado Simultáneos)
<b>VR/RV</b>	<i>Virtual Reality</i> (Realidad Virtual)
<b>XML</b>	<i>Extensible Markup Language</i> (Lenguaje de Marcado Extensible)



# 1 INTRODUCCIÓN

La presente tesis doctoral se enmarca en el campo multidisciplinar *human-computer interaction* (HCI), el cual estudia cómo los seres humanos hacen uso de la tecnología para el diseño de nuevas formas de interacción o la mejora de las presentes. El objetivo principal de HCI es lograr una mayor eficiencia en la interacción con la tecnología o conseguir un aumento del alcance de esta a un mayor número de personas. Para hacer frente a esta y otras metas es habitual que en esta área se haga uso de métodos y teorías de otros campos científicos, tanto relacionados con la tecnología como las ciencias de la computación, o centradas en las personas que hacen uso de ella, como la psicología o la sociología, por ejemplo.

Dentro del HCI, este trabajo se enfoca principalmente en dos áreas específicas: la realidad aumentada (RA) y el *end-user development* (EUD).

La RA se define como el conjunto de técnicas que tienen como objetivo lograr que los elementos virtuales, que son generados por ordenador, coexistan con el resto de elementos reales sin que pueda llegar a ser apreciable el origen de los elementos, real o virtual [7]. Hasta la fecha, la RA ha demostrado su potencial en un amplio rango de áreas como la educación [15,74,75,166,167], la medicina [68,114,137], la defensa [7,95], la industria [29,40,42,139], el turismo [36,82], el marketing [25,35,70] o el entretenimiento [120,175]. De manera más concreta, entre los ejemplos anteriormente citados, se pueden destacar el enriquecimiento del contenido y la interactividad de los libros tradiciones [15] en el campo de la educación, o el uso de la RA como instrumento de apoyo en las operaciones [114], entre otros muchos ejemplos. Sin embargo, el sector dónde mayor evolución e impacto ha tenido la RA en los últimos años es el área del entretenimiento, donde en apenas menos de quince años prototipos desarrollados con fines de investigación [120] se han visto reflejados en productos comerciales para el público en general [176–178], siendo el caso más destacado y significativo Pokémon Go [175], la versión móvil del popular videojuego Pokémon. Es de esperar que esta tendencia continúe, y que la presencia y uso de la tecnología de RA aumente y se expanda a otros sectores gracias a la aparición de dispositivos específicos de RA más baratos, potentes y ergonómicos como las gafas de realidad aumentada [179], así como la posibilidad de utilizar este tipo de aplicaciones en las nuevas generaciones de teléfonos móviles inteligentes [27].

Por su parte, *end-user development* (EUD) se define como “un conjunto de métodos, técnicas y herramientas que permiten a usuarios de sistemas software actuar como si fueran desarrolladores de software hasta el punto de ser capaces de crear, modificar y extender artefactos software por sí mismos” [93] (p.p. 2). Aunque esta definición menciona explícitamente artefactos software también puede ser aplicada al diseño de hardware. Como se mencionó en el primer párrafo, uno de los objetivos principales de HCI es lograr que los sistemas puedan ser caracterizados como *easy to use*, en otras palabras, que proporcionen interfaces válidas y fáciles de usar por un amplio número personas. La meta de EUD va más allá en este sentido, y pretende lograr que los sistemas sean *easy to develop* [93], es decir, que cualquier persona tenga la posibilidad de desarrollar sistemas por sí mismos. Más concretamente, EUD trata de facilitar a los usuarios de los sistemas los mecanismos necesarios que puedan ser capaces de llevar a cabo los procesos de creación, modificación y extensión de software y/o hardware, de tal forma que puedan disponer de nuevos artefactos o modificar y extender aquellos existentes sin depender para ello de expertos en desarrollo de software que no suelen contar con un conocimiento tan profundo como sobre el dominio de aplicación como ellos. La grabación de *macros* [180], *visual programming* [181], *trigger-action programming* [158], *programming by example* [57], *programming-by-specification* [94] son ejemplos de propuestas desarrolladas dentro del área de EUD con el fin de capacitar a los usuarios para poder describir el funcionamiento de algoritmos sin necesidad de aprender un lenguaje de programación.

## 1.1 Contexto

El trabajo de investigación descrito en este documento se centra en una intersección común de ambas áreas, más concretamente, en la creación y autoría de experiencias de RA por usuarios finales que pueden no disponer de conocimientos técnicos. La Figura 1.1 muestra una representación gráfica de la intersección donde se centrará la presente tesis.

Hasta ahora el uso práctico de la tecnología de RA no se encuentra tan extendido como cabría esperar atendiendo a sus potenciales beneficios de aplicación. Parte del problema radica en los altos costes y elevada complejidad que suelen ir asociados a este tipo de desarrollos, y que a su vez están relacionados con la amplia variedad de habilidades y conocimientos requeridos, entre los que se incluyen visión artificial, diseño 3D, tecnologías móviles y dispositivos *wearable*, entre otros. Existen algunas herramientas dirigidas a simplificar el proceso de autoría de experiencias de RA y permitir a usuario no técnicos su creación. Este es el caso de PowerSpace [59] que permite componer un escenario de RA usando la aplicación de ofimática PowerPoint, DART [98] que capacita a artistas a componer y simular experiencias de RA usando *scripts* predefinidos, y iaTAR [91] que permite la creación y configuración de los elementos virtuales de una experiencia de RA de manera inmersiva. Sin embargo, estas herramientas únicamente permiten crear experiencias muy sencillas, para un ámbito o tipo de experiencia muy específica y el resultado final suele ser bastante pobre.

En general, tanto las herramientas de EUD de RA como muchas de las aplicaciones de RA realizadas por profesionales adolecen del mismo problema: no logran generar en el usuario la ilusión de que los objetos virtuales pertenecen al mundo real que Azuma propugnaba en [7] como definitoria de una experiencia de RA. Únicamente cuando se emplean dispositivos específicos de RA de última generación, como las Microsoft Hololens, se llega a lograr satisfacer las expectativas de visualizar e interactuar con objetos virtuales en el mundo físico de manera similar a cómo lo hacemos con los objetos que nos rodean. Sin embargo, estos dispositivos aún resultan extremadamente caros y su uso se encuentra limitado a laboratorios y entornos profesionales específicos. La mayoría de las experiencias de RA disponibles al público general emplean tecnologías móviles, y a menudo se limitan a superponer los objetos virtuales sobre la visión del entorno, sin integrarlos correctamente en él ni reproducir los efectos ambientales que los objetos físicos manifiestan. Se han propuesto distintas técnicas que tratan de aliviar este problema y mejorar el realismo de las experiencias de RA [13,20,37,50,80,87], pero cada una de ellas únicamente es capaz de abarcar al mismo tiempo uno o dos efectos realistas (colisión, gravedad, oclusión...), lo que hace necesario la combinación de varias técnicas simultáneamente e incrementar la complejidad del desarrollo. Además, algunas de estas técnicas pueden ser difíciles o imposibles de aplicar dependiendo de las características del entorno, por ejemplo, el empleo de cámaras de profundidad en espacios grandes [129] o el empleo de proyectores para

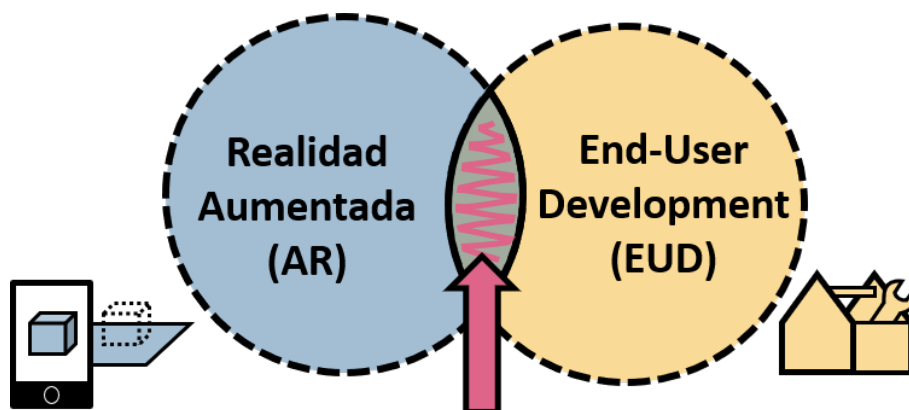


FIGURA 1.1 INTERSECCIÓN DE LAS ÁREAS DE REALIDAD AUMENTADA Y END-USER DEVELOPMENT

la RA en espacios luminosos [85]. Así, aumentar espacios abierto o interiores de grandes dimensiones resulta especialmente complicado.

## 1.2 Objetivos de la Tesis

El objetivo principal que persigue la presente tesis es el de **facilitar la creación de experiencias de realidad aumentada realistas a usuarios finales que no cuenten con un perfil tecnológico alto**. El término experiencias realistas en el contexto de este trabajo hace hincapié en que las experiencias creadas deberán de lograr la impresión de que los elementos virtuales den la impresión de pertenecer al entorno real. También, es relevante aclarar que las experiencias creadas no deberán requerir el empleo de dispositivos específicos de RA, como por ejemplo las Microsoft Hololens, sino que deberá ser posible desplegarlas utilizando los medios que el público en general habitualmente emplea como soporte de este tipo de aplicaciones, como por ejemplo teléfonos móviles.

Dada la amplitud que puede abarcar tal objetivo, se hace necesario delimitar los problemas que se abordarán en este trabajo, así como los objetivos específicos que se plantean a la hora de abordar estos problemas (Tabla 1.1).

TABLA 1.1 PROBLEMAS Y OBJETIVOS ABORDADOS EN LA TESIS

Problemática	Objetivos
P1.- Limitado realismo de las experiencias de RA.	O1.- Diseñar un método para aumentar un espacio de grandes dimensiones de forma realista.
P2.- Dificultad para aumentar entornos de grandes dimensiones o exteriores.	
P3.- Limitaciones del enfoque tradicional de autoría de RA empleando tecnología móvil.	O2.- Estudiar los beneficios y limitaciones del uso de la RV como parte del proceso de autoría de experiencias de RA.

Una vez fragmentado el objetivo principal general de la tesis en objetivos más específicos y abarcables, se enumera las tareas que se han de llevar a cabo para alcanzar el objetivo general de esta tesis:

- *Realización de una revisión de la literatura en el campo de la RA con especial hincapié en la creación y autoría de RA por usuarios de sistemas.* Esta revisión se encarga de realizar una revisión del área principal de la tesis para conocer las taxonomías existentes alrededor del término de RA, la evolución que ha tenido la RA y las diferentes técnicas empleadas. También, se estudia el área de EUD para mostrar los trabajos que combinan ambas áreas. Esta tarea permite conocer qué se ha hecho hasta el momento y las limitaciones existentes basadas en la literatura.
- *Diseño de un proceso de creación de experiencias de RA realistas para entornos de grandes dimensiones.* El resultado de esta tarea es un proceso que consta de una serie de pasos para aumentar un espacio de grandes dimensiones que permita combinar una variedad de efectos realistas al mismo tiempo. Los pasos del proceso deben de poder ser llevados por usuarios.
- *Validación de proceso de creación de experiencias de RA realistas para entornos de grandes dimensiones.* Esta tarea se centra en evaluar el proceso definido en la tarea anterior en casos reales. De tal modo que se pueda conocer si el proceso es válido, sus limitaciones y el nivel de complejidad que presenta.
- *Diseño de una arquitectura o flujo de trabajo para la elaboración de experiencias RA realistas en entornos de grandes dimensiones.* Con la información obtenida tras la ejecución de la

tarea anterior, el objetivo que se persigue es la definición de una arquitectura que permita llevar el proceso de desarrollo definido desde una herramienta.

- *Diseño e implementación de una herramienta de autoría de RA basada en el uso de un móvil RA y una que haga uso de la RV.* El resultado de esta tarea son dos herramientas de autoría implementadas que permitan llevar a cabo el proceso autoría haciendo uso de dos enfoques diferentes.
- *Evaluación de los dos enfoques de autoría de RA.* El objetivo de esta tarea será conocer los beneficios y limitaciones que presenta ambos enfoques durante el proceso de autoría de una misma experiencia de RA.
- *Diseño de una guía para el desarrollo de una herramienta para usuarios de sistemas para la creación de experiencias de realidad aumentada realistas.* Este objetivo tiene como meta recopilar todas las recomendaciones a seguir para la construcción de una herramienta que permita crear experiencias de RA realistas haciendo uso de los resultados obtenidos en los trabajos anteriores.

### **1.3 Metodología**

Para el alcance de los objetivos expuestos se hace necesario la utilización de una metodología válida que permita guiar este proceso de investigación. En la presente tesis doctoral se ha optado por el marco de trabajo de investigación definido por Hevner et al. [63,64] denominado *Design Science in Information System Research* (DSISR). Este marco de trabajo no solo guía la investigación, sino que también vela por la originalidad y utilidad de la contribución.

Para llevar a cabo el proceso de investigación, este marco de trabajo de investigación está constituido por tres ciclos que toman parte en diferentes momentos de la investigación: Ciclo de Relevancia, Ciclo de Diseño y Ciclo de Rigor. Estos ciclos son los encargados de relacionar la investigación en curso con el dominio de aplicación y la base de conocimiento de las áreas que intervienen. Desde una visión muy general, en la fase inicial de la investigación, el Ciclo de Relevancia y el Ciclo de Rigor se encargan de proporcionar a la investigación la información y el conocimiento del dominio de aplicación y de la base de conocimiento, respectivamente. Esta información y conocimientos son recogidos para el diseño y/o construcción de un artefacto o proceso que posteriormente es objetivo de una o varias evaluaciones a través del Ciclo de Diseño. Finalmente, el conocimiento que se pueda sintetizar tras las evaluaciones y análisis contribuirán en la base de conocimiento mediante el Ciclo de Rigor y aportará mejoras o avances en el entorno donde puedan ser aplicados estos artefactos o procesos por medio del Ciclo de Relevancia. La Figura 1.2 se muestra un esquema visual de los ciclos contextualizado para este trabajo.

## 1.4 Contribución

El objetivo final cualquier trabajo de investigación es realizar una aportación significativa al área que pertenece. En el caso concreto de esta tesis doctoral, las aportaciones estarán suscritas dentro de las áreas de la RA y el EUD. De manera más concreta, éstas se centran en la facilitación en la creación de experiencias de realidad aumentada por parte de usuarios de sistemas que no tiene por qué disponer de conocimientos técnicos altos.

El marco de trabajo de investigación de Hevner distingue dos tipos de contribuciones una vez que la evaluación del artefacto o proceso haya arrojado resultados interesantes: contribuciones en la base de conocimiento, por medio del Ciclo de Rigor que se materializan en artículos, congresos y revistas, y contribuciones en el dominio de aplicación, por medio del Ciclo de Relevancia que son los artefactos y procesos construidos y evaluados en el dominio de la investigación. Las contribuciones hacia la base de conocimiento de esta tesis son:

- *Un método de creación de experiencias de RA realistas para entornos de grandes dimensiones.* Esta contribución versa directamente sobre el primer objetivo O1 de la tesis. Una vez realizado la validación del método propuesto en un caso real, se realiza una discusión de los resultados obtenidos que se encuentran plasmados sobre varias publicaciones científicas ya publicadas [2,108,110].
- *Conocimiento de los beneficios y limitaciones del uso de la RV como parte del proceso de autoría de RA.* Esta contribución está estrechamente ligada al objetivo O2 de la tesis. Una vez hecha la discusión de los datos extraídos del experimento con usuarios. Las conclusiones que se pueden extraer formarán parte de la base de conocimiento por medio de la presente tesis doctoral.

Respecto a las contribuciones que se realizan al dominio de aplicación de la investigación se resumen en dos puntos:

- *Herramientas y scripts desarrollados para la creación de experiencias de RA realistas en entornos de grandes dimensiones:* Una vez hecha la evaluación del proceso y obtenidos unos resultados positivos en el estudio, las herramientas y *scripts* que fueron desarrollados forman parte de las contribuciones sobre el dominio de aplicación del estudio.
- *Herramientas de autoría de RA:* Las dos herramientas que han sido desarrolladas y evaluadas por usuarios son contribuciones que se realizan al dominio de aplicación estudiado en esta tesis.

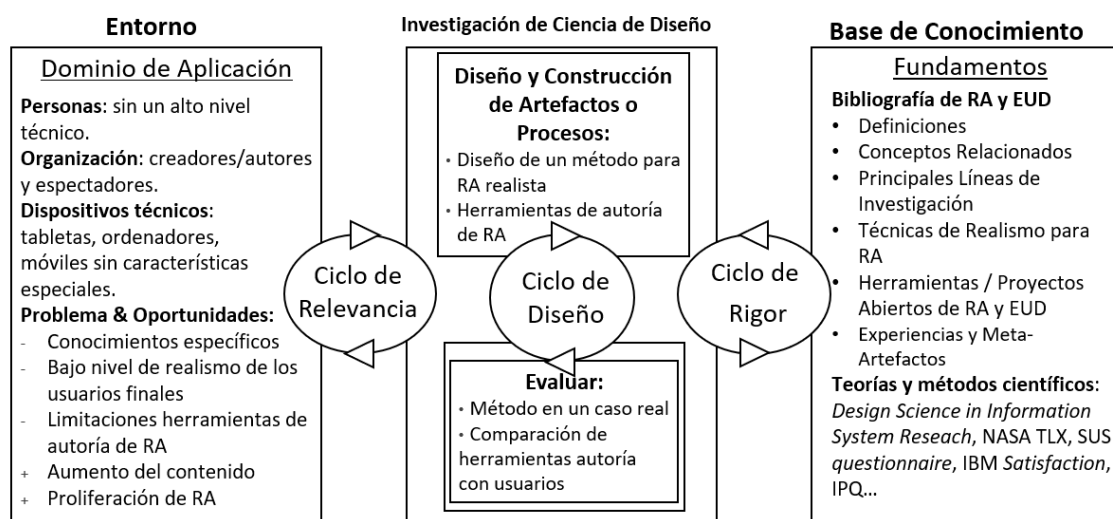


FIGURA 1.2 DSISR CONTEXTUALIZADO PARA EN ESTA TESIS DOCTORAL

## 1.5 Estructura de la Tesis

La memoria de tesis se encuentra dividida en siete capítulos, una sección de referencias y tres anexos. Sin considerar el presente capítulo de introducción, este libro se distribuye en los siguientes aparatos:

- Estado de la Cuestión: presenta una revisión de la literatura con especial interés en las áreas de RA y EUD. Como punto de partida se define el término RA y se presentan y explican conceptos relacionados a través de diferentes taxonomías de clasificación. A continuación, se detalla el algoritmo general que siguen los sistemas de RA para presentar la evolución que ha tenido la RA de manera cronológica. Posteriormente, las siguientes secciones se centran en presentar los diferentes tipos de tracking usados en la RA y técnicas desarrolladas para el incremento del realismo. Para finalizar se presenta el área de EUD y se explican los diferentes enfoques que se han ido proponiendo para las herramientas de creación y autoría de RA para ser usadas por usuarios de sistemas.
- Identificación de Problemas y Objetivos: expone las problemáticas que se abordan en esta tesis y el objetivo general de la misma. El cual, dada su amplitud, se divide en dos objetivos que son abordados por los siguientes capítulos.
- Metodología de Investigación: explica el marco de trabajo de investigación que utiliza esta tesis. Este capítulo además presenta una contextualización detallada del marco de investigación utilizado en esta tesis.
- Creación de experiencias de realidad aumentada realistas: presenta una propuesta de proceso de creación de experiencias de RA realistas en espacios de grandes dimensiones que no requiere de un perfil tecnológico alto para llevarse a cabo. Una vez explicado el proceso, se presenta un caso de uso real en un auditorio de grandes dimensiones y una prueba de concepto de una segunda iteración del proceso para una localización exterior. Al final de este capítulo, se discuten los resultados obtenidos en los casos reales.
- Autoría de Experiencias de Realidad Aumentada basada en Realidad Virtual: En este capítulo se presenta un método de autoría de RA en el que se hace uso de la RV durante parte del proceso y una herramienta considerada más tradicional que sigue la actual tendencia de herramientas de autoría de RA. Con el fin de conocer los beneficios y limitaciones de la propuesta se lleva a cabo un estudio experimental comparativo entre ambos métodos con usuarios. Para finalizar, este capítulo discute los resultados obtenidos durante la evaluación.
- Conclusiones, Contribuciones y Trabajos Futuros: Este capítulo detalla las conclusiones que se puede extraer de todos los trabajos realizados durante la tesis doctoral. Además, resume las contribuciones originadas durante este trabajo. Así mismo, también se presenta una serie de líneas de trabajo que potencialmente podrían ser continuadas.
- Bibliografía: recopila toda la información de citas incluidas a lo largo del documento de tesis.
- Anexo A: Consentimiento y encuesta: Este anexo muestra la documentación dada a los participantes en las evaluaciones.
- Anexo B: Respuestas de las encuestas: Este anexo incluye las respuestas de los participantes en tablas formateadas.
- Anexo C: Gráficos adicionales extraídos de los datos recogidos pero que no se encuentran alineados con los objetivos del estudio.

## 2 ESTADO DE LA CUESTIÓN

En los últimos años hemos podido ser testigos de cómo la tecnología de realidad aumentada (RA) ha alcanzado la suficiente madurez para dar el salto del ámbito de los laboratorios y entornos especializados a su uso en aplicaciones comerciales. Sin embargo, a pesar de que la experiencia que proporcionan los actuales dispositivos de RA ha mejorado considerablemente, todavía se encuentra alejada de la visión propuesta por algunos autores como Azuma [7]:

*“Augmented Reality (AR) [...] allows the user to see the real world, with virtual objects superimposed upon or composited with the real world. [...] Ideally, it would appear to the user that the virtual and real objects coexisted in the same space [...]”* [7] (pp. 2)

Esta definición de RA cuenta con una amplia aceptación en la comunidad científica<sup>1</sup>, y destaca por ser independiente de la tecnología de dispositivos y por la importancia que otorga a conseguir provocar en el usuario la ilusión de que los elementos virtuales se encuentran y forman parte del mundo real. Para lograr esta última ilusión, el mismo autor señala 3 propiedades que una experiencia de RA debe satisfacer:

1. Debe permitir combinar los elementos del mundo real y elementos virtuales generados por ordenador.
2. Debe ser interactiva en tiempo real.
3. El contenido virtual debe encontrarse registrado en el entorno real en una ubicación 3D.

Atendiendo a estas propiedades muchas de las experiencias a menudo etiquetadas como RA, como por ejemplo efectos especiales de películas o la superposición de un texto o imagen sobre un vídeo, no deberían ser consideradas como tales. Por otro lado, el autor también añade que la RA puede ser aplicada al resto de los sentidos, no solo a la vista.

De manera previa a esta definición, distintos autores han propuesto varias taxonomías que incluyen este concepto y términos relacionados [106,11,12,101,113,22]. Entre ellos destaca el Virtual Continuum de Milgram [106], que organiza las distintas experiencias de realidades mediadas a través de un segmento continuo en función del grado o nivel de combinación de la realidad y contenido generado por ordenador (Figura 2.1). Dentro de este continuo se pueden distinguir 5 tipos principales de experiencias:

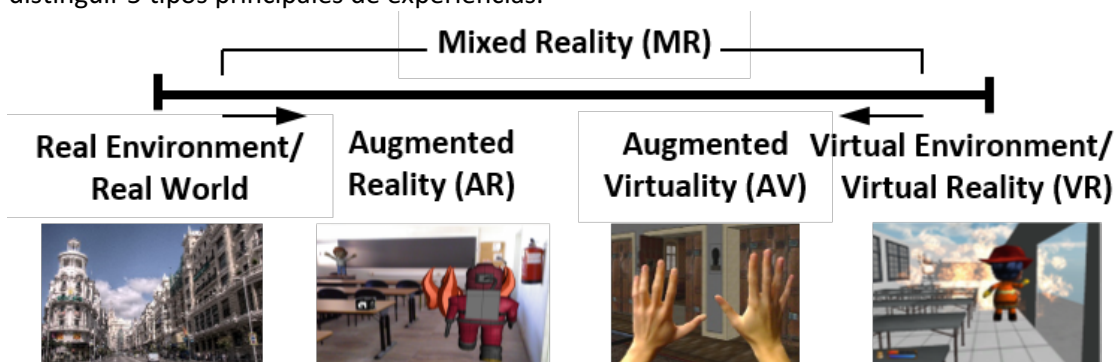


FIGURA 2.1 ADAPTACIÓN DEL VIRTUAL CONTINUUM DE MILGRAM [106] CON EJEMPLOS DE CADA UNO DE LOS NIVELES [5,24,52].

- *Real Environment* o *Real World*: En este tipo de experiencia el usuario percibe la realidad tal como es, sin ningún tipo de modificación realizada por ordenador.
- *Augmented Reality* (RA): Engloba a aquellas experiencias en las que el usuario percibe una realidad en la que se han introducido contenidos generados por ordenador que parecen coexistir con el resto de los elementos reales.

<sup>1</sup> Artículo referenciado con más de 9.000 citas (dato proporcionados por Google Scholar: <https://scholar.google.es/scholar?q=A+survey+of+augmented+reality>).

- *Augmented Virtuality (AV)*: En estas experiencias los usuarios perciben un entorno generado totalmente por ordenador en el que se incluyen algunos elementos reales.
- *Virtual Environment o Virtual Reality*: Se refiere a los casos en los que se percibe únicamente una realidad sintética completamente generada por ordenador.
- *Mixed Reality (MR)*: Agrupa el conjunto de realidades que combinan en distinto grado elementos del mundo real y elementos virtuales. En otras palabras, es un segmento o subconjunto del Virtual Continuum al que pertenecen la *Augmented Reality* y el *Augmented Virtuality*.

El continuo de Milgram permite clasificar a estos tipos de sistema en función del grado de contenido virtual que genera en la realidad que rodea al usuario, pero no establece umbrales o límites a partir de los cuales definir con claridad cuándo pertenece a una u otra categoría. Tal y como se muestra en la Figura 2.2, si bien los sistemas de realidad virtual se distinguen claramente del resto al estar basados únicamente en contenido virtual, aquellos que combinan elementos reales y elementos virtuales en una medida similar podrían ser considerados tanto como *Augmented Reality* como *Augmented Virtuality*.

Esta taxonomía ha servido como base a posteriores propuestas realizadas por otros autores que se han centrado en incorporar otras particularidades que también se encuentran presentes en muchos sistemas de realidad mediada. Así, varios autores [22,113] han ampliado el segmento continuo de Milgram por medio de nuevas dimensiones a través de las que representar el grado de integración de la tecnología en el entorno o la cuantía de usuarios que participan en la experiencia. La Figura 2.3 muestran dos ejemplos de estas representaciones. La Figura 2.3 izquierda muestra el Continuo de Milgram por medio de un segmento horizontal el cual es cortado perpendicularmente por el punto medio con una línea vertical que representa los diferentes grados de ubicuidad. Estos dos segmentos forman un plano bidimensional sobre el cual se sitúan diferentes categorías de sistemas mediados delimitados por un rango finito de ambos dominios. Con el objetivo de ejemplificar algunas de estas categorías se mencionan algunos sistemas construidos y documentados por la literatura. La Figura 2.3 derecha muestra una interpretación de una propuesta de taxonomía que incorpora el grado de virtualidad, ubicuidad y multiplicidad de usuarios concurrentes en la experiencia.

Por otro lado, Benford et al. [11,12] desde una perspectiva fuertemente orientada al área de *Computer Supported Cooperative Work (CSCW)*, desarrolló una taxonomía para los sistemas CSCW en función de las características que definen los espacios de trabajo compartidos haciendo uso de tres dimensiones: *transportación (transportation)*, *artificialidad (artificiality)* y *espacialidad (spatiality)*. La primera de estas dimensiones representa el nivel de “desplazamiento” de realidad que los usuarios y objetos experimentan desde el espacio que ocupan (local) a uno remoto (diferente del local). En un extremo de esta dimensión se encontrarían las aplicaciones de videoconferencia como Skype [182], que no producen en los usuarios la sensación de transportarse a un entorno remoto, y en el otro, sistemas como Mozilla Hubs [183] o Facebook Horizon [184], que generan la sensación a los usuarios de ser transportados a espacios distintos de en los que se encuentran físicamente. El término de

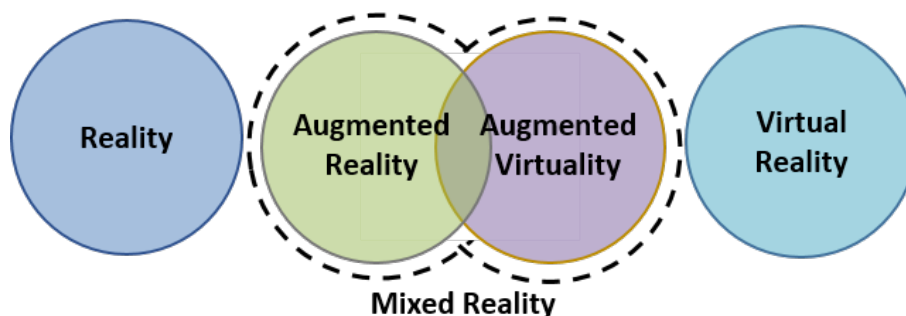


FIGURA 2.2 CONJUNTO DE TIPO DE APLICACIONES O SISTEMAS



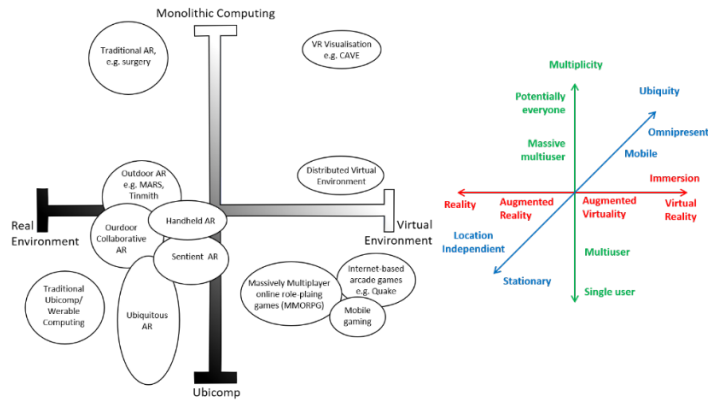


FIGURA 2.3 COMBINACIÓN DEL CONTINUUM DE MILGRAM Y WEISER [113]. COMBINACIÓN DEL CONTINUUM DE MILGRAM, WEISER Y LA MULTIPLICIDAD DE USUARIOS [22]

transportación presenta cierta similitud con el concepto de inmersión en sistemas de RV, distinguiéndose que la inmersión únicamente se aplica a la experiencia individual del usuario, mientras que la transportación puede referirse también a objetos. En cuanto a la segunda dimensión, artificialidad, representa el grado en que el contenido de la experiencia es obtenido de la realidad o generado por ordenador. En un extremo de esta dimensión nos encontraríamos los sistemas puramente físicos en los que todo el contenido es obtenido directamente de la realidad, como los sistemas de videoconferencia. En el otro extremo están representados los puramente sintéticos donde el contenido del entorno es virtual, como Facebook Horizon. Esta dimensión, en palabras del propio autor, se encuentra estrechamente ligada al Virtual Continuum propuesto por Milgram. Por último, la dimensión de espacialidad alude a qué grado y en qué cantidad el contenido es compartido en el espacio de trabajo. Una de las características más interesantes de esta propuesta es que tomando en consideración únicamente las dos primeras dimensiones, dado que la tercera está fuertemente ligada a los sistemas CSCW, es posible representar los espacios compartidos. La Figura 2.4 muestra una adaptación del gráfico realizado por Benford basándose en las dimensiones de artificialidad y transportación y sobre las que el mismo autor identificó cuatro cuadrantes: (1) *Augmented Reality*, (2) *Virtual Reality*, (3) *Physical Reality* y (4) *Telepresence*. De este modo, Benford logra incluir gran parte de los sistemas representados en el *Virtual Continuum*, y ampliar la clasificación incluyendo los sistemas de telepresencia. Este último tiene como objetivo lograr que los usuarios tengan la sensación de estar presentes en otro entorno real diferente al entorno real donde se encuentra. Para ello, sensores capturan el entorno remoto proporcionando la información sin alteración al usuario. En esta clasificación los sistemas de *Augmented Virtuality* serían un subconjunto de los sistemas de *Virtual Reality*, mientras que los de *Mixed Reality* se distribuirían a través de los cuatro cuadrantes ocupando las zonas más próximas entre los mismos.

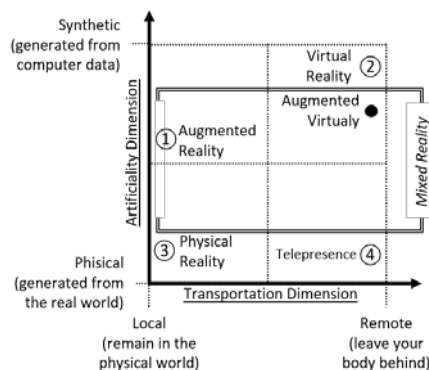


FIGURA 2.4 ADAPTACIÓN DE LA CLASIFICACIÓN DE ESPACIOS COMPARTIDOS DE ACUERDO CON LAS DIMENSIONES DE TRANSPORTACIÓN AND ARTIFICIALIDAD [12]

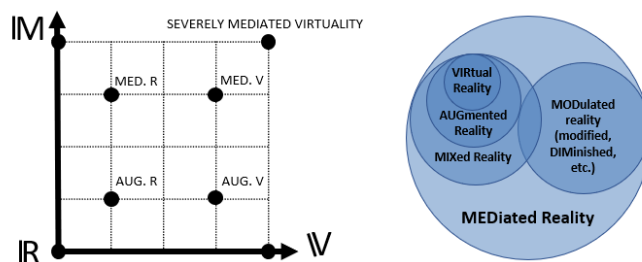


FIGURA 2.5 REPRESENTACIÓN BIDIMENSIONAL DE LA VIRTUALIDAD-MEDIADA [101] Y JERARQUÍA DE CONJUNTOS DE LA REALIDAD MEDIADA [101]

Desde una perspectiva diferente, Mann [101] propone una clasificación centrada en las diferentes formas y posibilidades de modificar o adaptar la realidad. Más específicamente, Mann tiene en cuenta, por una parte, el grado de modificación o eliminación de la realidad percibida por el usuario, y por otra, el nivel de “virtualidad” de la experiencia. La Figura 2.5 ilustra la clasificación de Mann por medio de un diagrama con dos ejes (izquierda), y un diagrama de Venn (derecha), en el que un mayor nivel de profundidad de los subconjuntos revelaría un mayor grado de virtualidad o modulación de la realidad.

Por medio de esta clasificación, el autor distingue entre realidades modificadas (*modified reality*) y realidades reducidas (*diminished reality*). Ambos tipos de realidades se encuentran englobadas dentro del conjunto de realidades moduladas (*modulated reality*) que tienen como fin proporcionar una experiencia en la cual se modifica o se elimina parte de la realidad. De manera más concreta, las realidades modificadas transforman la percepción de la realidad por medio de la modificación de las propiedades de los objetos como, por ejemplo, su tamaño, color, forma... Por su parte las realidades reducidas se centran en conseguir que ciertos elementos de la realidad no sean percibidos y con ello crear la ilusión de que no existen. Para lograr este objetivo estos sistemas provocan en los usuarios las sensaciones que ellos obtendrían si el objeto real objetivo no existiera. Si bien pudiera parecer que la experiencia proporcionada por este tipo de sistemas es la opuesta a la obtenida mediante RA, es posible combinar ambos tipos de realidades para crear una experiencia en la que objetos reales son completamente reemplazados por otros virtuales, tal y como se muestra en la parte izquierda de la Figura 2.6. Algunos autores como Hillstead [66] han propuesto versiones corregidas del diagrama original de Mann para capturar estas combinaciones de realidades, Figura 2.6 derecha, y distinguir a los sistemas de virtualidad modulada (*modulated virtuality*) que combinan RA y AV.

A modo ilustrativo, la Figura 2.7 representa por medio de imágenes algunas de las combinaciones de tecnologías mediadas. Así, la fila inferior de la Figura 2.7 muestra a través de

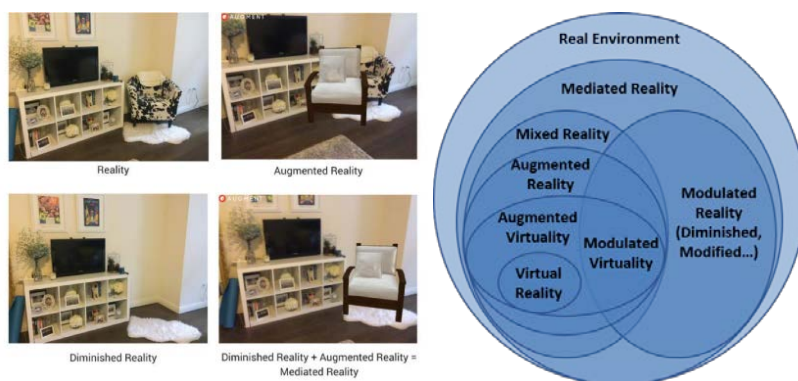


FIGURA 2.6 EJEMPLO DE COMBINACIÓN DEL USO DE LAS TECNOLOGÍAS DE REALIDAD AUMENTADA Y REALIDAD REDUCIDA [173] Y ADAPTACIÓN DEL DIAGRAMA PRESENTADO POR MANN [66]

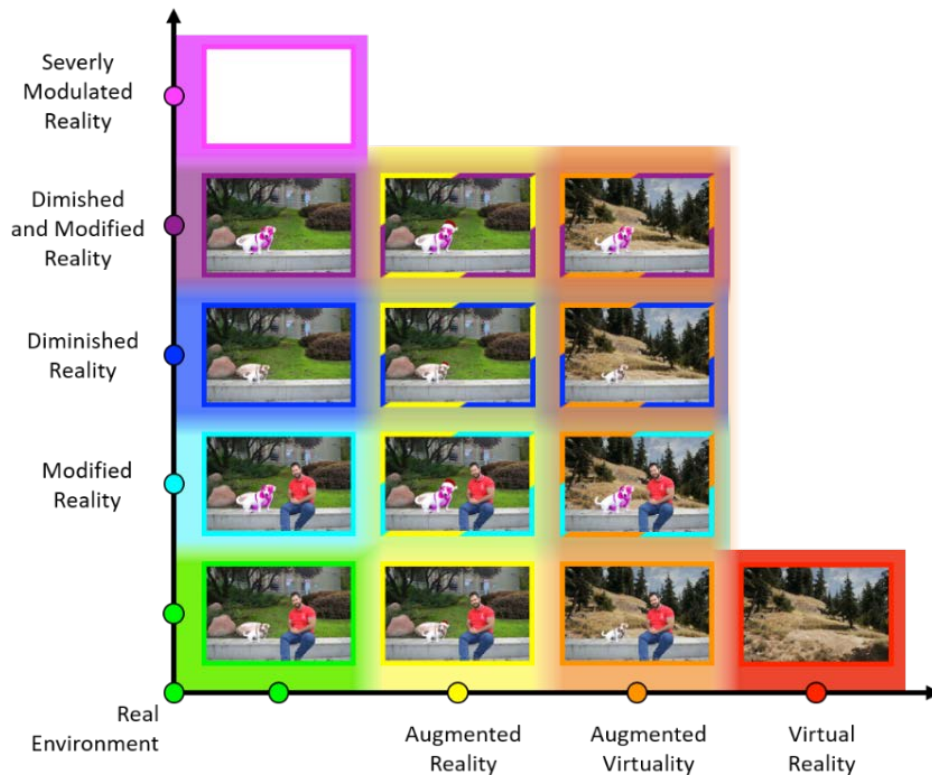


FIGURA 2.7 ADAPTACIÓN DE COMBINACIONES DE TECNOLOGÍAS MEDIADAS BASADO EN EL DISEÑO DE HILLSTEAD [66]

distintos ejemplos los diferentes niveles del *Virtual Continuum*. En la primera columna, desde la parte inferior hasta la parte superior, se ilustra cómo la realidad es modulada progresivamente. Por último, las 6 imágenes que no pertenecen ni a la primera columna ni a la última fila hacen referencia al término Virtualidad Modulada en las que se combinan ambas vertientes de las tecnologías mediadas. No se ha ilustrado la imagen superior izquierda que correspondería a una realidad que ha sido modificada de manera tan severa que no se parece a la realidad.

En términos generales, todas las definiciones y taxonomías desarrolladas alrededor de las tecnologías medidas son bastantes sólidas y se encuentran aceptadas por la comunidad científica. Sin embargo, en los últimos años hemos podido ser testigos de cómo algunos términos como realidad aumentada, han empezado a ser empleados de manera indiscriminada por cuestiones de marketing, o incluso, se ha intentado desligar en favor de otros términos relacionados por cuestiones comerciales buscando fundamentalmente la diferencia frente a la competencia. Ejemplo de estos hechos lo encontramos con las Google Glass [185] o el uso del término realidad mixta por parte de Microsoft para sus Microsoft Hololens [179]. En el caso de las Google Glass, el dispositivo fue presentado como unas gafas de realidad aumentada cuando únicamente es posible superponer objetos virtuales sobre la pequeña porción del campo de visión de los usuarios que cubre la proyección del dispositivo, por lo que no es posible conseguir su alineación con el resto de los elementos del entorno real. Por su parte, el uso que hace Microsoft del término realidad mixta para describir la experiencia proporcionada por las Microsoft Hololens resulta confuso al ligarlas directamente con los dispositivos empleados. De acuerdo con la definición propuesta por Milgram [106] en su descripción del *Continuum Virtual*, el término realidad mixta englobaría a todas aquellas realidades que ni son realidad virtual pura, ni percepción de realidad no modificada a través de tecnología. Esto incluye por tanto a la RA, pero también a otro tipo de realidades mediadas tecnológicamente como la virtualidad aumentada.

Una vez que se ha explicado en detalle qué es la RA y las tecnologías relacionadas, en el siguiente apartado se explica el flujo principal para lograr crear experiencias RA. Tras ello, se explica de manera cronológica la evolución que ha tenido la RA desde sus primeros pasos hasta la actualidad. Posteriormente, se listan y explican las principales técnicas que se emplean en las experiencias de RA para conseguir alinear el contenido virtual con los objetos del mundo real, así como las empleadas para incrementar su realismo e integración en el entorno. Por último, se trata en detalle las aplicaciones y herramientas enfocadas a la creación de RA por usuarios que no tienen un perfil técnico.

## 2.1 Algoritmo General de Ejecución de Realidad Aumentada

Como se ha descrito anteriormente, un sistema de RA tiene como fin proporcionar una experiencia en la que elementos virtuales generados por ordenador parecen estar presentes en la realidad. Para lograr esta sensación, el sistema debe ser capaz de alinear el contenido virtual adecuadamente en el espacio real y al mismo tiempo dotarlos de capacidad de interacción con los usuarios e incluso con otros objetos reales o virtuales. Para lograr alcanzar cada una de estas características, se hace indispensable que los sistemas de RA dispongan de sensores que permitan capturar información sobre el entorno a aumentar e identifiquen referencias en el mismo. Además, debe contar con interfaces de entrada que permita al usuario interactuar con los elementos de virtuales. Toda la información capturada deberá ser procesada en tiempo real para que el sistema genere los elementos virtuales y los proyecte sobre algún tipo de dispositivo de salida que permita superponerlos sobre la visión del usuario, o modificar la percepción de algún otro de sus sentidos.

Hoy en día es posible encontrar una gran variedad de dispositivos que tienen como fin dar soporte a las experiencias de RA por medio de distintas técnicas y tecnologías. A pesar de esta gran diversidad, el flujo de ejecución de un sistema RA suele ser similar en todos ellos y se puede resumir en tres pasos principales:

- **Reconocimiento (*Recognition*):** El objetivo de esta fase es recopilar información acerca del entorno en el que se encuentra actualmente el usuario para determinar si se debe aumentar su percepción de este. Para ello, el sistema debe procesar los datos obtenidos de los diferentes sensores de entrada. El funcionamiento interno de esta fase depende del tipo de los dispositivos y técnicas empleadas en cada caso. Así, por ejemplo, en un enfoque que haga uso de una cámara y marcadores, el algoritmo se centrará en ejecutar varios filtros sobre las imágenes obtenidas y procesarlas en búsqueda del patrón de los marcadores, mientras que los sistemas de RA que emplean sensores de ultrasonidos tratarán de detectar patrones de impulsos de sonidos en la información capturada por los mismos. Una vez evaluada la información extraída del medio, el resultado de la identificación podrá ser positivo, si reconoce una o varias referencias del entorno, o negativo, si no ha sido posible por falta de referencias o las referencias no están registradas. En caso favorable, el flujo de ejecución continua en la siguiente etapa. En el caso contrario, el sistema ejecuta esta fase de nuevo para conocer si el usuario ha cambiado de espacio.
- **Alineación (*Registration*):** Una vez que el entorno ha sido reconocido, se determina la posición que debe ocupar el contenido virtual en el entorno real. Para ello se parte de los datos proporcionados por la fase anterior y otros datos adicionales de referencia. De nuevo, el proceso depende de los dispositivos y técnicas empleadas. Para el caso del uso de marcadores, el algoritmo se centrará en calcular la distancia y rotación de la cámara haciendo uso para ello la proyección del marcador en la imagen capturada y otros datos como los parámetros intrínsecos de la cámara y el tamaño real del marcador. En cambio, en un enfoque que haga uso de sensores de ultrasonidos, el algoritmo se centra en determinar la posición por medio del cálculo de la diferencia de fases entre las señales recibidas teniendo en cuenta la distribución de las fuentes de emisión. El resultado final de esta etapa es una serie de vectores o matrices que determinan la posición relativa y orientación del sistema a las referencias reconocidas de la etapa anterior.

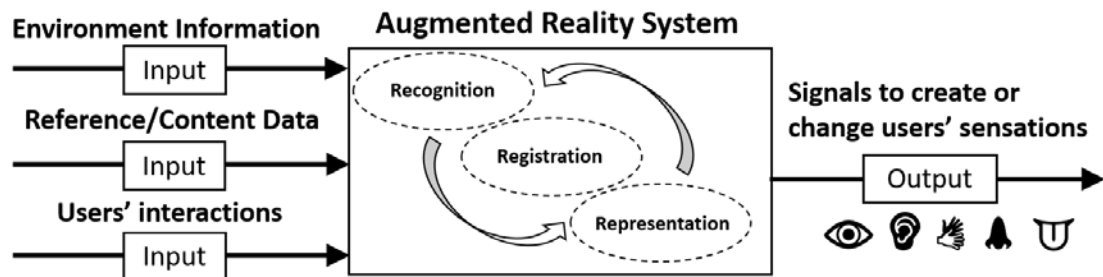


FIGURA 2.8 DIAGRAMA DE ENTRADA Y SALIDA DE UN SISTEMA DE REALIDAD AUMENTADA

- Representación del contenido (*Content Representation*): En esta fase el sistema genera la sensación de que el contenido virtual es parte del entorno percibido por el usuario a través de sus sentidos. Para ello, el sistema procesa los datos obtenidos en la fase anterior junto con la información relativa al contenido y las interacciones del usuario. El resultado de esta fase es el conjunto de señales que recibe las diferentes interfaces de salida para provocar la experiencia de RA al usuario. De nuevo, el algoritmo que gobierna esta fase depende en gran medida de las tecnologías e interfaces de salida empleadas. Si la visión de la experiencia es mostrada en un monitor o pantalla el algoritmo se centrará en combinar imágenes capturadas de la realidad junto con el contenido virtual. En cambio, si el contenido virtual es proyectado sobre la visión del usuario del entorno, el algoritmo debe representar únicamente los elementos virtuales teniendo en cuenta las deformaciones de la superficie sobre la que se proyectan. De igual modo, los algoritmos que gestionan los efectos acústicos serán distintos dependiendo de si el sonido es emitido por auriculares, que debe mezclar el sonido de la experiencia con el ruido del entorno, o por altavoces en el entorno, que deberán tener en cuenta la posición que ocupa el elemento virtual y el usuario de la experiencia.

En la Figura 2.8 se muestra un diagrama de un sistema RA. En él se presenta las tres entradas principales de datos: información del entorno, fuente de datos (relativa al contenido virtual o de referencia) y las acciones del usuario con el contenido virtual. La salida de este tipo de sistema es un conjunto de señales que son manejadas por diferentes interfaces de salida: pantallas, proyectores, auriculares, etcétera. También, se representa las tres fases que se pueden abstraer de la mayoría de los sistemas de RA y que dadas las entradas anteriormente mencionadas es capaz de generar el conjunto de señales de salida en tiempo real.

## 2.2 Evolución de las Tecnologías y Técnicas de Realidad Aumentada

El origen detrás del concepto de realidad aumentada se remota a la década de los años 1960, para ser más exactos entre 1965 y 1968. En 1965, Ivan Sutherland planteó el diseño de un sistema de visión que permitiría incluir gráficos generados por ordenador en la visión de los usuarios al que denominó inicialmente Ultimate Display [148]. En aquel momento, y debido a la potencia computacional de los ordenadores de la época, los gráficos generados por ordenador se limitaban a simples líneas rectas, pero Sutherland ya vislumbraba el potencial que podría llegar a alcanzar esta tecnología.

*“The ultimate display would, of course, be a room within which the computer can control the existence of matter. [...]. With appropriate programming such a display could literally be the Wonderland into which Alice walked.”* [148](pp. 3)

En 1968, dicha propuesta de sistema fue construida y finalmente fue apodado *The Sword of Damocles* debido a su peso y a que se encontraba suspendido del techo [149]. Este sistema estaba compuesto por unas lentes estereoscópicas que permitían ver la realidad a la vez se superponían sobre ellas gráficos generados por un ordenador mediante unos pequeños tubos de rayos catódicos (CRT). Para componer los gráficos con la perspectiva correcta el autor planteó

dos enfoques diferentes. El primero de ellos consistía en la colocación de unos sensores mecánicos a lo largo del brazo articulado que permitía calcular la posición y orientación del visor. Para ello, la unidad de cómputo hacía una lectura de cada uno de los sensores y calculaba estas dos componentes, posición y orientación, teniendo en cuenta la forma del brazo. Para el segundo enfoque el autor equipó al visor con tres generadores de ultrasonidos con diferentes frecuencias e instaló cuatro sensores en el techo de manera estratégica con los que podía conocer la posición y orientación del visor por medio de la diferencia de fases entre ondas que recibían los receptores.

Aunque el diseño y construcción del sistema de Sutherland es considerado por muchos como el primer *Head-Mounted Display* (HMD) de RV y RA, ya existían algunas patentes previas de HMDs de RV [62,144] y de sistemas que buscaban proporcionar la sensación de “telepresencia”, como el HMD de Philco Corporation [31].

En paralelo a la aparición de estos primeros HMD también surgieron los primeros *Head-Up Display* (HUD) en el ámbito militar. Estos dispositivos fueron considerados de gran utilidad para facilitar la conducción de aeronaves [100], y aunque todavía eran muy rudimentarios y únicamente constaban de unos cristales sobre los que se proyectaba distinta información, su evolución y desarrollo resultó decisivo para en la evolución de la tecnología de RA.

En los años posteriores a estos desarrollos, la mayoría de los esfuerzos de investigación relacionados con el desarrollo de HMD estuvieron centrados en la RV [30,51] y en el desarrollo de nuevos mecanismos de interacción con el contenido digital [83,84]. El único trabajo destacable del campo de RA de esta época es el teclado programable de Knowlton en 1977 [78]. Este sistema, denominado *Virtual Push Button System*, consistía un teclado en el que cada tecla podía ser reetiquetada en función del uso. Para ello, Knowlton dispuso un teclado sin etiquetas debajo de un cristal semitransparente que permitía ver el teclado al mismo tiempo que una pantalla CRT reflejaba su contenido alineado con el teclado. Esta pantalla era la encargada de etiquetar cada tecla de modo que hacía posible reconfigurar la plantilla del teclado en función de las tareas.

Décadas más tarde, en 1992, Thomas Caudell acuñó el término *Augmented Reality* y presentó cuatro ejemplos de uso de esta tecnología [28]. Caudell, conocedor de la complejidad del proceso de fabricación y mantenimiento de aeronaves, vio en esta tecnología una herramienta útil para facilitar la tarea de los operarios que tenían que seguir al pie de la letra extensos manuales mientras llevaban a cabo su trabajo. Su sistema permitía superponer información sobre su visión del entorno mientras llevaban a cabo las tareas de montaje del armazón de cableado, ensamblado o procesos de mantenimiento de componentes de una aeronave. El sistema consistía en una unidad generadora de ondas, que proporcionaba una referencia del entorno real, un HMD equipado con un receptor de ondas, para la obtención de datos de la distancia relativa a la referencia, y un ordenador estándar, que se encargaba del cálculo de la perspectiva y generación del contenido. A pesar de que el sistema presentaba algunas limitaciones, como un campo de visión de solo 7 grados y cierto retraso a la hora de actualizar el contenido, su desarrollo sirvió para demostrar el potencial de esta tecnología en este campo.

Siguiendo la misma línea, se presentó KARMA (*Knowledge-based Augmented Reality for Maintenance Assistance*) un prototipo que proporcionaba una asistencia para el mantenimiento y la reparación de dispositivos por medio de la RA [47,48]. Para ilustrar la utilidad del sistema, los autores presentaron un caso de ejemplo consistente en el soporte del proceso del mantenimiento de una impresora de oficina en la que los usuarios podían obtener información adicional mediante iconos (flechas, etiquetas y representaciones de los componentes de la impresora) que se superponían en la visión de los usuarios. En cuanto a la arquitectura, una característica a destacar fue la utilización de varias unidades de procesamiento en paralelo para el cálculo de las posiciones y orientaciones de diferentes partes de la impresora con el objetivo de reducir el tiempo de retardo o *lag* en la experiencia.

Al mismo tiempo que el término RA se iba asentando a la hora de designar este tipo de tecnología, la potencia de cálculo de los procesadores fue progresivamente mejorando, propiciando la aparición de sistemas que hacían uso de un enfoque de tipo *Video See-through*. Hasta ese momento, la mayoría de los sistemas seguían un enfoque de RA de tipo *Optical See-through*, es decir, conseguían aumentar la visión del usuario mediante algún tipo de cristal sobre los que se proyectaban únicamente los elementos virtuales, lo que requería una menor capacidad de cómputo a la hora de generar el contenido a representar. El aumento de potencia de los equipos hizo posible desarrollar sistemas capaces de componer la visión aumentada mediante la captura y modificación de video en tiempo real. Un ejemplo de ello lo podemos encontrar en el sistema presentado por Bajura et al. [9] en 1992, capaz de capturar ecografías de una mujer embarazada y superponerlas en un vídeo del vientre de la madre capturado en tiempo real.

A la vez que estos trabajos iban fijando los cimientos de la RA, diversos autores centraban sus esfuerzos en el estudio y análisis de esta tecnología desde un enfoque más teórico. Entre estos trabajos destaca la taxonomía Milgram [106] en 1994 descrita en al inicio del capítulo. Tras ello, empezaron a surgir también estudios que exploraban tecnologías y enfoques para mejorar el realismo de las experiencias de RA con la introducción de efectos como la oclusión y la colisión [20], además de trabajos que se enfocaban en describir y presentar arquitecturas y *frameworks* de sistemas de RA como Studierstube [150] y Tinmith [119]. El primero de ellos consistió en un *framework* orientado a las experiencias de RA colaborativas que proporcionaba un panel interactivo personal para interactuar con la experiencia. Uno de los objetivos de este entorno era permitir su uso en diferentes ámbitos o con distintos propósitos, dado que hasta ese momento la mayoría de los sistemas RA se desarrollaban *ad hoc* para una determinada tarea. El sistema propuesto permitía su uso de manera simultánea por varios usuarios empleado distintos HMDs, lo que hacía posible que cada uno de ellos pudiera tener una representación distinta de los datos, y que pudiesen existir contenidos privados o públicos. Por su parte, la arquitectura de Tinmith [119] buscaba dar soporte al desarrollo de experiencias de RA en el exterior. Para hacer frente a la complejidad que presenta implementar las tareas de reconocimiento en espacios exteriores el sistema combinaba el GPS y una cámara óptica para el cálculo de la posición del usuario en el entorno a aumentar.

Progresivamente, en la segunda mitad de la década de los 90 fueron apareciendo métodos y técnicas que trataban de dar solución al problema del *tracking* del usuario en la experiencias de RA haciendo uso de marcadores [125,127], hasta que a finales de esta década se publicó la librería ARToolkit [186]. Esta biblioteca de desarrollo daba soporte a la implementación de experiencias RA basadas en marcadores empleando únicamente un ordenador y una cámara Web. Esto impulsó considerable a la investigación en el área debido al bajo coste que representaba frente al desarrollo de hardware específico, y fomentó la aparición de las primeras aplicaciones RA para usuarios de sistemas en los siguientes años. Otro hito a destacar de esta misma época fue la publicación del *survey* de Azuma [7] en 1997 en el que además de resumir las limitaciones y principales posibilidades de uso de la RA del momento, se determinaban las características fundamentales de la tecnología y se articulaba la definición RA más ampliamente aceptada por la comunidad científica. Tras este hito la RA evolucionó notablemente en sus diferentes ámbitos. Muestra de ello es la publicación de un nuevo *survey* por parte de Azuma sobre los avances producidos dos años después [6]. En él, se presenta una mayor variedad de usos de la RA entre los que destaca la proliferación del uso comercial en el ámbito del contenido audiovisual.

Ya en la primera década del nuevo milenio, se pueden encontrar los primeros esfuerzos enfocados en conseguir que usuarios no expertos puedan crear experiencias de RA. Ejemplo de ello son el proyecto europeo AMIRE [54,187] y el surgimiento de herramientas como DART [97,98]. Posteriormente, a lo largo de esta década aparecieron más herramientas de este tipo [188,91,140,141,122,165] y productos comerciales accesibles al público en general a través de consolas, como el juego Invizimals para PlayStation Portable (PSP) [189]. Otro impulso más fue



la aparición a finales de esta década de los primeros *smartphones* con la suficiente potencia de cómputo para ejecutar las primeras aplicaciones de RA móviles comerciales como AR Invaders [99], la cual ya hacía uso de algunas bibliotecas para desarrolladores de RA disponibles en la época [190]. También cabe destacar la aparición en el mercado de modelos de gafas de RA comerciales [191], si bien aún a un precio muy elevado y dirigidos a su uso con fines de investigación, y la publicación de los primeros algoritmos de procesamiento de imagen que permitían dejar atrás la necesidad de usar marcadores, como el algoritmo *Parallel Tracking and Mapping* (PTAM) [76].

Ya en la segunda mitad de la siguiente década, la inversión en el mercado de RA aumentó considerablemente y las principales compañías tecnológicas comenzaron a presentar sus propios dispositivos para la RA de manera comercial, la mayoría empleando un enfoque *optical-see through* como la Google Glass [185], las Microsoft Hololens [179] y las Magic Leap [192]. Al mismo tiempo comenzaron a aparecer dispositivos móviles y tabletas con sensores hardware embebidos que facilitaban el desarrollo de experiencias de RA, como los dispositivos con tecnología Project Tango de Google [174]. Estos dispositivos cuentan con dos cámaras adicionales en la parte de atrás, una cámara de profundidad y una cámara de detección de movimiento. Las cuales unidas al sensor inercial logran un tracking más estable y liviano para el dispositivo y la incorporación del efecto de oclusión en las experiencias RA. En cualquier caso, el incremento en el precio del dispositivo que suponía la incorporación de las dos cámaras adicionales supuso un obstáculo considerable para la popularización del sistema, y en los años siguientes aparecieron bibliotecas como ARKit [193] de Apple y, posteriormente, ARCore [194] de Android, que hacían posible el empleo de el algoritmo SLAM en la mayoría de los teléfonos móviles de las marcas, sin necesidad de emplear cámaras adicionales.

La Figura 2.9 presenta de forma cronológica los principales hitos en el desarrollo científico y comercial de la tecnología de RA. Debajo de línea de tiempo se señalan las principales tendencias, que se resumen a continuación:

- Cambio de foco de importancia entre en Hardware y el Software en la investigación: En los primeros trabajos de investigación de RA gran parte del trabajo se centraba en desarrollar y describir en profundidad el hardware que hacía posible la creación de la experiencia RA. Esta tendencia ha cambiado debido a la existencia de dispositivos comerciales para ejecutar experiencias RA y, actualmente, se centra en el propio artefacto software y en la percepción del usuario.

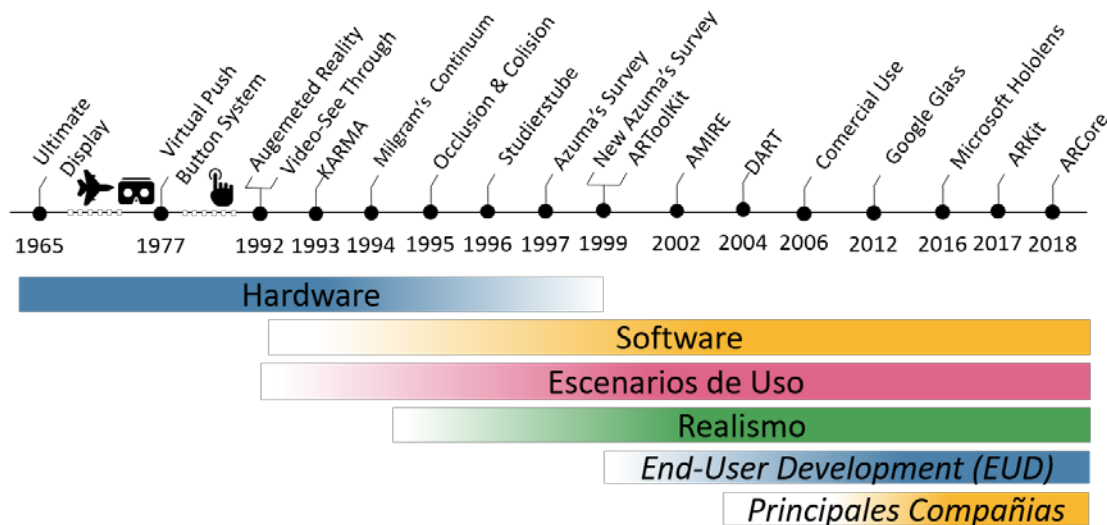


FIGURA 2.9 LÍNEA TEMPORAL CON ALGUNOS DE LOS HITOS DE LA RA. EL GRÁFICO INCLUYE ALGUNAS TENDENCIAS DE LAS INVESTIGACIONES EN EL ÁREA DE LA RA.



- Tendencia de exploración de las aplicaciones de la RA: Una vez que el término de RA fue acuñado a principios de la década de los 90 y surgiendo los primeros prototipos, los trabajos de investigación se centraron en la exploración de usos de la RA en diferentes áreas.
- Mejora del realismo: En paralelo a la exploración de nuevos usos de la RA a mediados de la década de los 90, empezaron a surgir trabajos de investigación enfocados en mitigar algunas carencias de las experiencias de RA en cuanto al realismo. Estos trabajos han estado proponiendo diferentes técnicas para distintos aspectos para incrementar el realismo de las experiencias de RA.
- Herramientas de creación y autoría de RA: Tras la aparición de la librería ARToolKit a finales de la década del 90 que facilitaba el desarrollo de la RA con un hardware bastante accesible, una webcam, empezaron a surgir trabajos de investigación centrados en poner en manos de los usuarios de sistemas la creación y autoría de estas experiencias de RA.
- Explotación comercial: Desde mediados de la década del 2000 han ido apareciendo dispositivos y software desarrollado por las principales empresas tecnológicas, que han visto en la RA un nicho de mercado con un alto potencial de crecimiento en las próximas décadas.

### 2.3 Alineación de los elementos Virtuales

Siguiendo la definición de RA de Azuma, los elementos virtuales de una experiencia de RA deben de estar alineados correctamente en el entorno real, de tal manera que evoquen la ilusión de que ocupan un espacio físico. Para lograr este objetivo es necesario procesar la información captura por los sensores de los sistemas de RA. Este proceso abarca en gran medida a las dos primeras etapas descritas en la sección 2.1.

Para abordar esta problemática a lo largo de los años se han desarrollado distintas propuestas que hacen uso de diferentes tecnologías y algoritmos. La Figura 2.10 muestra una adaptación de la taxonomía presentada por Mark Billinghurst [102] que divide en cinco grandes categorías las técnicas de *tracking*. En la mayoría de los contextos es posible utilizar exclusivamente una única técnica de *tracking*, a excepción de la técnica de *inertial tracking* que debe ser complementada, en la mayoría de los casos, con una o más técnicas dando lugar a un enfoque híbrido.

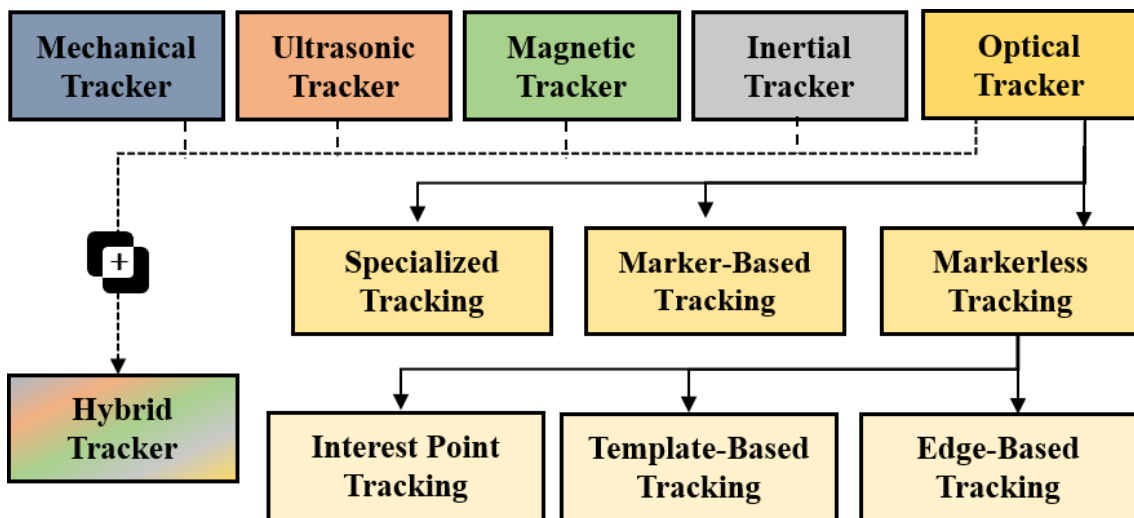


FIGURA 2.10 ADAPTACIÓN DE LA TAXONOMÍA DE MARK BILLINGHURST SOBRE TRACKER DE RA [102]

#### 2.3.1 Tracking Mecánico

Esta técnica se basa en el uso de una instalación en un punto fijo del que parten una serie de extensiones y articulaciones mecánicas equipadas con sensores, y que permiten conocer la posición relativa del visor de RA respecto al punto de instalación del artefacto. Esta técnica de

*tracking* fue empleada en el HMD de Sutherland [149], en el sistema de ecografías de RA de Bajura et al. [9] y es posible encontrar, a día de hoy, en sistemas comerciales situados en algunos museos [195]. La Figura 2.11 izquierda muestra el HMD de Sutherland. Como se puede apreciar en la imagen el HMD está suspendido del techo mediante un brazo que está equipado de sensores mecánicos. En la Figura 2.11 derecha se muestra un sistema en uso instalado en el Canterbury Museum de Christchurch, Nueva Zelanda. Este visor es un prismático de RA que a través del sensor situado en la base puede calcular la orientación para la correcta alineación del contenido virtual.

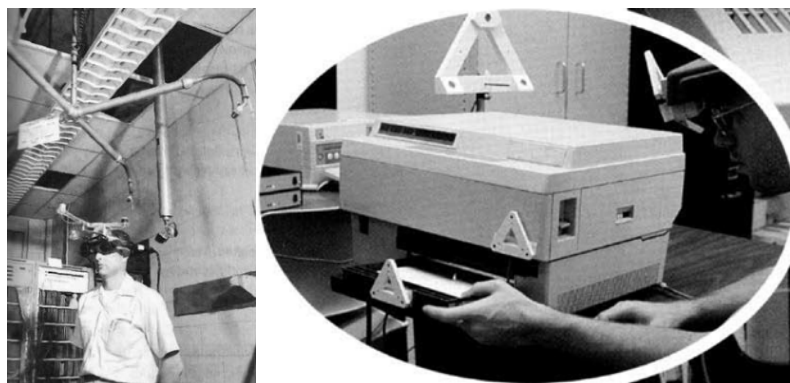
Entre las ventajas de este tipo de sistemas destacan la reducida dificultad de lectura de los datos de posición y la baja carga de procesamiento. Por el contrario, estos sistemas cuentan con serios inconvenientes en cuanto a la limitación de movilidad del usuario y el esfuerzo adicional que este debe realizar para mover o girar las estructuras mecánicas.



*FIGURA 2.11 LA PARTE IZQUIERDA MUESTRA EL HMD DE SUTHERLAND SUSPENDIDO DEL TECHO MEDIANTE UN BRAZO [149]. EN LA PARTE DERECHA SE PRESENTA UNA INSTALACIÓN DE UNOS PRISMÁTICOS RA EN EL CANTERBURY MUSEUM DE NZ [196]*

### 2.3.2 Tracking por Ultrasonidos

Este tipo de sistema se basa en el uso de generadores y receptores de ultrasonidos. Los primeros se configuran con distintas frecuencias y se adhieren a dispositivos u objetos móviles, mientras que los receptores se colocan en puntos fijos del entorno y disponen de filtros de frecuencia. De esta forma, conociendo todas las diferencias de fases de las señales de todos los receptores se puede obtener una aproximación de la posición relativa de los objetos frente a cada receptor. La configuración de los generadores de sonido y los micrófonos puede ser instalada con una distribución opuesta, colocando los generadores de sonido en el espacio de manera fija y equipando receptores en cada objeto físico sobre los que se va a hacer *tracking*.



*FIGURA 2.12 LA IMAGEN IZQUIERDA MUESTRA EL HMD DE SUTHERLAND SITUADO DENTRO DE UN ESPACIO DELIMITADO POR CUATRO RECEPTORES DE ULTRASONIDOS QUE ESTÁN DISPUESTOS EN EL TECHO [149]. LA IMAGEN DE LA DERECHA MUESTRA LOS EMISORES-RECEPTORES DE ULTRASONIDOS DEL SISTEMA KARMA QUE TENÍAN UNA FORMA TRIANGULAR [48]*

Este enfoque fue desarrollado también en el HMD de Sutherland [149] y es usado en una de las implementaciones del sistema KARMA [48]. La Figura 2.12 muestra el HMD de Sutherland dotado de tres generadores de ondas situados encima de visor en forma de triángulo equilátero. En la ilustración también se observan dos de los cuatro receptores de ultrasonidos colocados en el techo. La Figura 2.12 derecha muestra los sensores-receptores de ultrasonidos del sistema KARMA, los cuales estaban colocados en el HMD y en las diferentes partes móviles de la impresora. Esto no solo permite conocer dónde se encuentra la impresora respecto al visor, sino que también dónde se encuentra la bandeja y sus otras partes móviles.

Entre las ventajas de este enfoque cabe destacar la posibilidad de poder hacer *tracking* de varios elementos físicos a la vez y una mayor libertad de movimiento frente a un *tracking* mecánico. En cuanto a los inconvenientes cabe mencionar la necesidad de realizar una preparación previa del entorno a aumentar consistente en la instalación, configuración y calibración de los sensores.

### 2.3.3 Tracking Electromagnético

La base de funcionamiento de este tipo de *tracking* es muy similar al *tracking* por ultrasonidos. Sin embargo, en vez de utilizar generadores de ondas de ultrasonidos y micrófonos de altas frecuencias se hace uso de generadores de ondas electromagnéticas y sensores. Para este caso, los generadores de ondas electromagnéticas generan frecuencias o impulsos de ondas que son recibidas por los sensores para el cálculo de distancias a partir de la diferencia de fases o tiempos.

Este enfoque fue usado por el sistema de Caudell [28] e implementado en KARMA como *tracking* alternativo al basado en ultrasonidos. En la actualidad, se puede encontrar este tipo de *tracking* para aplicaciones móviles basadas en RA de *location-based*, que hacen uso del GPS y el magnetómetro (o brújula) incluida en la inmensa mayoría de los móviles. En la Figura 2.13 izquierda se muestra un esquema de la disposición de generadores de ondas y receptores utilizado por el sistema desarrollado por Caudell. En la imagen se puede apreciar como los diferentes sensores dispuestos en el panel de trabajo cubrían las diferentes áreas donde podía ser detectado el HMD. La Figura 2.13 central muestra el HMD de Caudell siendo utilizado para el ensamblado de cables de un almacén. La Figura 2.13 derecha presenta un screenshot de una aplicación de RA basada en localización GPS y brújula que muestra los locales cercanos [190].

La principal ventaja de este enfoque es que no requiere que el receptor esté en visión directa con la fuente emisora como sucedía en el anterior *tracking*. Sin embargo, este enfoque presenta algunos inconvenientes como la interferencia por campos magnéticos o la necesidad de una preparación previa del entorno como el anterior enfoque.

Hay que destacar que el uso de la tecnología GPS no requiere una instalación previa, sin embargo, solo puede ser empleada en exteriores de manera factible cuando existe un amplio margen de error de metros sin romper la experiencia de RA. La problemática alrededor del margen de error de metros se puede mitigar en parte con el uso de GPS diferencial [159].



FIGURA 2.13 LA IMAGEN DE LA IZQUIERDA PRESENTE UN DIAGRAMA DE LA DISTRIBUCIÓN DE GENERADORES DE ONDAS ELECTROMAGNÉTICAS Y RECEPTORES DEL SISTEMA CONSTRUIDO POR CAUDELL [28]. LA FOTOGRAFÍA CENTRAL MUESTRA EL HMD DE CAUDELL SIENDO UTILIZADO PARA ENSAMBLAR CABLES[28]. LA IMAGEN DE LA DERECHA ES UN SCREENSHOT DE UNA APLICACIÓN LOCATION-BASED RA [190].

### 2.3.4 Tracking Óptico

El incremento progresivo de la potencia en los procesadores de los dispositivos de mano como las tabletas y móviles favoreció el desarrollo y uso de algoritmos de *tracking* basadas en procesamiento de imagen. Estos algoritmos se centran en procesar las imágenes de dos dimensiones capturadas del entorno para identificar patrones y así calcular la orientación y posición de la cámara respecto a la referencia del mundo real. Dentro de los *Optical Tracker* podemos llegar a diferenciar varios tipos de enfoques que han ido apareciendo según iban siendo las unidades de procesamiento más potentes [102]: *Specialized Tracking*, *Marker-Based Tracking* y *Markerless Tracking*.

Los sistemas de *Specialized Tracking* se basan en la detección y reconocimiento de elementos añadidos deliberadamente en el entorno a aumentar. Estos elementos por lo general tienen alguna característica que los hacen fácilmente identificables frente al resto de partes del entorno. De tal modo que el objetivo principal del algoritmo es llevar a cabo esta identificación por medio de diferentes filtros y operaciones sobre las imágenes. Un ejemplo de esta propuesta fue llevada por el sistema presentado por Baruja [8]. Este sistema funcionaba empleando una serie de LEDs rojos situados en el entorno y un algoritmo que aislaba estas fuentes de luces en una imagen por medio de filtros y umbrales de brillo. El resultado final de este proceso es un mapa de píxeles que determinaban una posición dónde se encontraba las diferentes fuentes de luces. Con esta información unida a la distribución de las luces se podía estimar la posición y rotación de la cámara.

Por su parte, los sistemas de *Marker-Based Tracking* también requieren de la preparación del entorno, si bien en este caso suelen hacer uso para ello de algún tipo de etiqueta o marcador más o menos sencillos. Entre los primeros sistemas documentados en la literatura podemos destacar NaviCam [127] en 1995. Este sistema permitía superponer texto e imágenes en la visión del usuario a través de una *Palmtop*. Para ello, este sistema reconocía e identificaba etiquetas de colores parecidos a códigos de barras con las que podría estimar la posición relativa y superponer el contenido. Tres años más tarde, el autor principal de este trabajo presenta varias mejoras a través de la utilización marcadores cuadrados monocolor (en blanco y negro). Esta decisión de diseño, en palabras del autor, buscaba decrementar la complejidad en el reconocimiento gracias a la forma cuadrada y a reducir los fallos gracias a un alto contraste entre estos dos colores. Adicionalmente a estas ventajas, este diseño permite incluir en los marcadores patrones para una identificación entre ellos [125]. Estas características están presentes en la versión publicada de la biblioteca ARToolkit [186], cuya versión inicial se publicó un año más tarde.

Por último, al contrario de los dos enfoques anteriores, el enfoque *Markerless Tracking* persigue lograr calcular la posición relativa de la cámara al entorno sin necesidad de haber llevado a cabo una preparación o modificación previa del mismo. Los algoritmos empleados por este tipo de sistemas son más complejos que los anteriores y su ejecución requiere de información previa del entorno y una mayor capacidad de cómputo al llevar un procesamiento más intenso de las imágenes captadas en busca de puntos de interés o características que sirvan de referencia. Según la taxonomía de Billingham, esta subcategoría se puede clasificar a su vez en *Interest Point Tracking*, *Template-based Tracking* y *Edge-based tracking*.




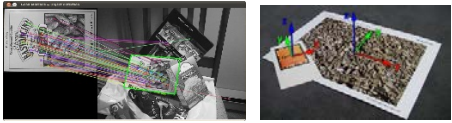
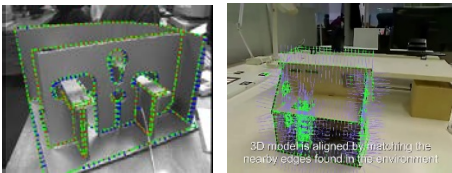
Los sistemas basados en algoritmos de *Interest Point Tracking* se centran en identificar una serie de puntos característicos a través de algoritmos conocidos de extracción de *features* como SIFT [96], SURF [10], ORB [135] o FAST [132] entre otros. Para ello, estos sistemas son entrenados de manera previa en el entorno a aumentar para la obtención de las características del entorno y su almacenamiento. De este modo, al volver a ejecutar el sistema desde una posición o ángulo diferente en el entorno, estos algoritmos son capaces de identificar de nuevo gran parte de estas *features* y extrapolar la posición y orientación en la que se encuentra la cámara del sistema [160]. Respecto a los sistemas basados en *Template-based Tracking*, se fundamentan en los mismos principios y algoritmos, sin embargo, el sistema no tiene por qué ser entrenado en el espacio a aumentar, sino que puede reconocer una referencia de entorno u objeto a partir de

una imagen de textura previamente proporcionada (la imagen debe presentar *features*, por lo que una imagen con un color plano no sería válida) [151]. Ya, por último, la subcategoría *Edge-based tracking* se encarga de reconocer los bordes y formas de los objetos reales y así poder generar una referencia del entorno. Estos algoritmos requieren información de manera previa de la forma o distribución de bordes de los objetos para poder identificarlos e inferir la posición y rotación desde la cual se captura el objeto [43].

La ventaja principal de los *trackers* ópticos es que no necesitan de ningún sensor especial, salvo una cámara, lo que los hace bastante accesible puesto que la inmensa mayoría de móviles y tabletas en la actualidad disponen de al menos una. Sin embargo, hay que tener presente que estos algoritmos presentan una gran carga computacional por lo que el funcionamiento fluido podría estar limitado a la potencia del dispositivo y al tipo de algoritmo. Como es evidente, esta limitación se irá mitigando según vayan apareciendo dispositivos más potentes. Sin embargo, una problemática presente que se puede encontrar en la mayoría de estos enfoques es una dependencia parcial de la luz ambiental que puede dificultar el reconocimiento del entorno por exceso de luz o su defecto.

La Tabla 2.1 que ilustra ejemplos de sistemas en cada una de las categorías en las que se divide los *trackings* ópticos. La primera categoría de la Tabla 2.1 muestra un ejemplo de *Specialized Tracking* donde aparece un jugador que porta una barita mágica en su mano [197]. Esta barita en realidad es mando PlayStation 3 Move [198] el cual incorpora un globo en su parte superior que puede ser iluminado con diferentes colores. Este color es filtrado por la cámara permitiendo conocer la ubicación del mando y así como una aproximación de la distancia en función del tamaño. La orientación de barita es proporcionada por un giroscopio que se incluyen en el mando. La categoría *Marker-based Tracking* presenta un ejemplo de inclusión de contenido virtual utilizando como referencia un marcador impreso en papel. La imagen de la izquierda presenta un sistema que hace uso de la librería de ARToolkit mientras que la imagen

TABLA 2.1 EJEMPLOS DE SISTEMAS OPTICAL TRACKING [43,55,76,111,197–201]

Optical Tracking		Ejemplos de Sistemas	
<i>Specialized Tracking</i>			
<i>Marker-based Tracking</i>			
<i>Markerless Tracking</i>	<i>Interest Point Tracking</i>		
	<i>Template-based Tracking</i>		
	<i>Edge-based Tracking</i>		



de la derecha ilustra diferentes *fiducial markers* de esta y otras bibliotecas de RA basadas en marcadores [55]. La subcategoría *interest point tracking*, dentro de la categoría *Markerless Tracking*, muestra dos capturas de pantalla de dos sistemas [76,111] que hacen uso de algoritmos de detección de puntos característicos para calcular la posición y orientación en cada *frame*. Los puntos de colores que se muestran en la imagen izquierda corresponden a *features* detectadas en la imagen capturada. Con relación al segundo tipo, *Template-based Tracking* muestra una imagen de cómo funcionaría un algoritmo basado en reconocimiento de texturas [199] y un ejemplo de la utilización de la biblioteca Vuforia [200]. La última fila muestra dos ejemplos de sistemas de RA [43,201] basados en *Edge-based Tracking*.

La Tabla 2.2 enumera las ventajas e inconvenientes entre las diferentes categorías de esta familia de *tracking* ópticos. Como se puede observar, la diferencia entre los sistemas *Specialized Tracking* y *Marker-based Tracking* frente a la subcategoría de *Markerless tracking* es la necesidad de modificar el entorno y el aumento de la carga computacional. El algoritmo *Edge-based Tracking* presenta mejores características, sin embargo, la utilización de estos enfoques depende en gran medida de la capacidad de cómputo del sistema y los datos del entorno que se disponen.

TABLA 2.2 VENTAJAS Y DESVENTAJAS DE LAS SUBCATEGORÍAS DE LOS OPTICAL TRACKING

<b>Optical Tracking</b>		<b>(+) Ventajas / (-) Inconvenientes</b>
<b>Specialized Tracking</b>		(+) Ventaja en entornos específicos (ej.: oscuros o baja luz) (-) Modificar el entorno (-) Requiere de hardware específico como balizas visuales u otro tipo de artefacto
<b>Marker-based Tracking</b>		(+) Marcadores baratos (-) No recomendable para entornos grandes
<b>Markerless Tracking</b>	<b>Interest Point Tracking</b>	(+) No modifica el entorno (-) Alta carga computacional
	<b>Template-based Tracking</b>	(+) No modifica el entorno (+) Referencias desde imágenes de texturas (-) Alta carga computacional
	<b>Edge-based Tracking</b>	(+) No modifica el entorno (+) Robusto a la luz (-) Alta carga computacional

### 2.3.5 Tracking Inercial

El enfoque de *tracking* inercial hace uso de sensores como el acelerómetro y el giroscopio que pueden estimar la posición y rotación desde un punto inicial que corresponde generalmente al inicio del sistema. Si bien este tipo de sensores pueden proporcionar una estimación inicial válida y a un bajo coste computacional, a medida que transcurre el tiempo de empleo de este enfoque dicha estimación se suele deteriorar y aparecen desviaciones y errores. Esto hace el sistema muy proclive a romper la correcta alineación del contenido virtual con el entorno real. Por lo general, este tipo de *tracking* no puede ser usado de manera exclusiva para crear experiencias de RA siendo necesario integrándolo con algún otro tipo de *tracking*.

A pesar de esta limitación, existen aplicaciones de móviles de RA como AR Invaders [99] que hacen uso exclusivo de este tipo de sensor debido al contexto específico de la experiencia. El juego tiene como objetivo abatir los ovnis que van rodeando al usuario del dispositivo móvil. Para ello el usuario tiene que rotar sobre sí mismo y pulsar un botón habilitado para abatirlos. En la Figura 2.14 se muestra un *screenshot* de este juego.



FIGURA 2.14 FOTOGRAFÍA DE LA APLICACIÓN AR INVADERS [172] JUNTO CON UN SCREENSHOT [202]

Tal y como se ha mencionado, las principales ventajas de este enfoque son su reducido coste y la pequeña carga computacional que requiere. Sin embargo, presenta la desventaja de que puede ser empleada en un número muy reducido de contextos de manera aislada y tiene que ser combinados con otros enfoques para resolver esta problemática.

### 2.3.6 Tracking Híbrido

Los sistemas de tracking híbridos combinan diferentes categorías *tracking* para resolver o mitigar las problemáticas que presentarían cada enfoque por separado. Es posible encontrar una variedad de sistemas RA que hacen uso de este enfoque híbrido [67,72,145] con la meta de alcanzar una mejor precisión o con la intención de no comprometer el tiempo real de la experiencia de RA. Así, los autores de [145] presentaban en 1996 un sistema de RA que utilizaba marcadores de colores como referencia (*optical tracking*) que fue combinado con un sensor magnético con la intención de reducir el tiempo de procesamiento de los marcadores. Por su parte, el sistema descrito en [72] combina el uso de una cámara para realizar *model-based tracking* de las líneas que forman los edificios y un giroscopio. En este otro caso el giroscopio complementa los datos de la cámara para predecir la dirección que tenía las líneas de los edificios, reduciendo así el procesamiento de la imagen. Además, los datos de este sensor eran corregidos gracias al sensor óptico. Por último, el sistema presentado en [67] basado en *Markerless tracking* óptico lograba reducir el ruido de las imágenes por medio de un sensor inercial logrando obtener buenos resultados incluso cuando la cámara estaba en movimiento.

La ventaja principal de este tipo de *tracking* es que resuelve o mitiga los problemas que presentan cada tracking por separado. El inconveniente mayor de este enfoque es que tiende a aumentar la complejidad del sistema de reconocimiento, pudiendo llegar a requerir de hardware adicional. En cualquier caso, esta última desventaja está desapareciendo debido a la tendencia actual de incorporar más sensores en los dispositivos móviles.

## 2.4 Técnicas para la Mejora del Realismo en las Experiencias de Realidad Aumentada

Una experiencia de RA no solo consiste en lograr una correcta alineación del contenido virtual en un entorno real, sino que además debe conseguir que este contenido parezca pertenecer al entorno real. Para lograr este fin se han desarrollado diferentes técnicas que tienen como meta replicar los efectos que tiene todos los elementos reales de un entorno como las sombras, la oclusión, el brillo, etcétera. Este proceso se lleva en la última fase del flujo de ejecución de un sistema RA que describe en la sección 2.1, es decir, la etapa de representación.

### 2.4.1 Efecto de Oclusión

Entre los distintos obstáculos a los que hay que hacer frente a la hora de tratar de dotar de realismo a una experiencia de RA el problema de la oclusión puede considerarse de los más relevante siendo de los primeros en ser tratados por la literatura [20]. El efecto de la oclusión se produce cuando dentro del campo de visión una semirrecta imaginaria que tiene como origen

el punto de mira del usuario es capaz de atravesar dos objetos reales. Este suceso cotidiano en un entorno real provoca que el elemento más lejano quede oculto de manera total o parcial por la silueta que ocupa el más cercano. Este caso concreto de dos elementos reales no genera ningún problema en una experiencia de RA, ni tampoco una en el que elemento más cercano sea virtual y el lejano real o directamente los dos elementos sean virtuales, gracias en este último caso a mecánicas de rasterización 3D las GPU como el *z-buffering*. Sin embargo, cuando no se aplica ningún tipo de mecanismo, en el caso de que el elemento real está más cercano y el virtual quede situado detrás de este, se creará la sensación de que este último esté flotando en el aire y sea más pequeño de lo que debería ser. Para solucionar esta problemática el sistema ha de conocer cuál es el orden de cercanía de los elementos o de cada una de las partes de ellos. Esta información se obtiene en la mayoría de los casos empleados cámaras de profundidad, cámaras estéreo [20], modelos tridimensionales [20] o mediante una ordenación manual de los contornos de los elementos [13].

El uso de cámaras de profundidad permite obtener un mapa de distancia para cada píxel de la imagen que se captura. Esto se consigue por medio una fuente de luz infrarroja que ilumina el entorno, y un sensor que captura dicha luz que ha rebotado previamente sobre los objetos físicos. En función de la intensidad de la luz reflejada es posible estimar la distancia de cada objeto a la cámara. En el caso de las cámaras estereoscópicas, la distancia de los objetos se estima mediante cálculos por el desplazamiento en píxeles que tiene un mismo objeto en las dos imágenes capturas junto con una serie de parámetros como la distancia entre cámaras y el campo de visión de cada cámara (FoV). Por medio de alguno de estos dos enfoques el software puede deducir si los elementos virtuales se encuentran o no ocluidos por los objetos reales y que partes de los objetos virtuales deben ser representadas. De tal modo que se consiga proporcionar en la experiencia la sensación de que el elemento virtual se encuentra detrás de uno real. Tanto las cámaras de profundidad como las estereoscópicas suelen ofrecer buenos resultados, si bien presentan limitaciones a la hora de ser utilizados en espacios grandes debido al restringido rango de alcance del sensor y el incremento de error en grandes distancias para las cámaras estéreo. Además, las cámaras de profundidad pueden presentar dificultades para representar la distancia a la que se encuentran objetos muy oscuros, puesto que absorben gran parte de la luz infrarroja, o reflectantes, como cristales, espejos u objetos con acabados metálicos dado que tienden a refractar la luz de manera excesiva. Estos problemas pueden ser resueltos con un enfoque que haga uso de un modelo 3D del entorno real [20]. Esta técnica se basa en la utilización de una representación 3D precisa de la forma del entorno que determine qué elementos virtuales o qué partes de ellos deben de ser ocluidas. En este caso la dificultad se deriva de la necesidad de realizar un modelado, y de previo del entorno la tarea de mantener dicho modelo correctamente alineado con la visión del usuario. Además, este enfoque no puede tener en cuenta cambios que se realicen en el entorno o la presencia de nuevos objetos en él. Por último, las técnicas basadas en la delimitación del contorno de los objetos [13] pueden evitar tener que realizar el modelado completo, si bien requieren especificar manualmente el orden de profundidad de las siluetas trazadas. Una vez que se conoce dónde se encuentra cada uno de los elementos reales del entorno ya se puede renderizar las partes de los elementos virtuales que deberán de ser visibles desde la posición del usuario.

#### **2.4.2 Colisión**

Otro efecto que puede afectar en gran medida al realismo de la experiencia de RA es la ausencia de una correcta colisión entre los objetos virtuales y los reales. A menos que sea su objetivo, cuando un objeto virtual atraviesa u ocupa el espacio de un objeto físico en lugar de colisionar, la ilusión de que forman parte del entorno decae. Para solucionar este problema se suele delimitar el espacio que los objetos virtuales no deben atravesar por medio de un mapa o maya con la forma del entorno [20]. Como en el caso de la oclusión, este mapa puede ser obtenido por medio de cámaras de profundidad, cámaras estereoscópicas o modelos del entorno 3D, si bien en este caso, las soluciones basadas en cámaras presentan la limitación



frente a los modelos 3D de no permitir conocer la forma de las partes ocluidas por elementos reales. Es decir, no se puede conocer con qué elementos podría colisionar los elementos virtuales cuando existe un elemento real que ocluye esas superficies.

### 2.4.3 Gravedad

La gravedad es una fuerza con la que se convive a diario lo que lleva al ser humano a predecir de manera intuitiva y casi instantánea cómo se comportarán los cuerpos que no se encuentra sobre una superficie horizontal. Por ello, con el objetivo de que un usuario de una experiencia de RA no perciba que un elemento no pertenece a la realidad, éste deberá de simular sufrir dicha fuerza cayendo de forma análoga a como lo haría un objeto real hasta colisionar con una superficie plana. Para lograr este efecto, algunos autores determinan de manera explícita esta fuerza en las experiencias a través de un vector [20], de tal modo que el sistema lleve a cabo este movimiento píxel a píxel hasta detectar una colisión. En cambio, otros autores como Kurz [87,88] han propuesto abordar este problema por medio de sensores inerciales que se encuentran presente en la mayoría de los dispositivos móviles. Estos componentes son capaces de calcular el módulo, dirección y sentido de la gravedad a partir de la diferencia de potencial medido en unas cavidades que contienen un fluido dieléctrico. De ese modo, el sistema puede calcular la dirección y sentido del vector de gravedad y aplicarlos a los elementos virtuales para replicar un comportamiento de caída o incluso de ascenso si fuera preciso.

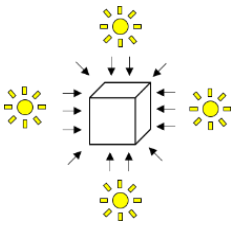
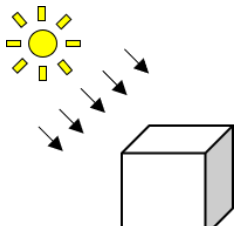
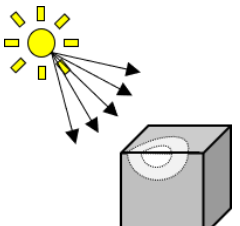
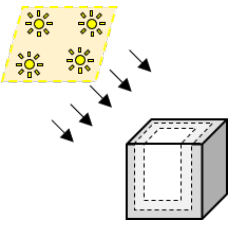
### 2.4.4 Luces y Sombras

El diseño de técnicas que mejoran el realismo a través de la simulación de los efectos de luces y sombras ha sido estudiado en profundidad en el área de la computación gráfica para la implementación de videojuegos, que deben soportar la interacción con el usuario en tiempo real, así como para la renderización de películas de animación. Todas estas técnicas siempre han intentado lograr un equilibrio para alcanzar el mayor nivel de realismo posible de luces y sombras teniendo en cuenta la potencia de computación del momento [118]. Antes de profundizar en estas técnicas en el área de la RA, la primera parte de esta sección realiza un recorrido por los conceptos básicos y técnicas empleadas en películas de animación y videojuegos para la iluminación y sombras.

La renderización o rasterización 3D es una técnica que permite definir una imagen a partir de la proyección de una escena 3D. En otras palabras, este proceso se encarga de definir el color de cada píxel de la imagen que se visualiza, la cual puede considerarse como un mapa de píxeles, a partir de la distribución de los elementos que forman la escena 3D. Los tipos de elementos que componen una escena 3D se pueden agrupar en 3 tipos: modelos tridimensionales, fuentes de luces y cámaras. Los modelos representan el contenido a mostrar. Estas entidades, en su más bajo nivel, son realmente un conjunto de vértices con coordenadas 3D que definen triángulos y cuadrados. Este conjunto de polígonos forma una maya que representa la forma del modelo. Sobre esta maya se suele incluir una imagen que actúa como textura definiendo colores y detalles adicionales de la forma como su normal, su difusión, su brillo. Con respecto a las fuentes de luces, estas se encargan de definir el origen y dirección de la luz simulada en la escena tridimensional. La Tabla 2.3 muestra los cuatro tipos de luces más extendidos en motores de videojuegos y software de edición 3D [153,203,204]:

- La Luz Ambiental u Omnidireccional es una fuente de luz que simula incidir sobre todas las superficies de los modelos con la misma intensidad. Este tipo de luz no tiene correspondencia en la realidad y por sí sola y no logra un alto nivel de realismo puesto que los modelos se tienden a mostrar con colores vivos. Respecto a sus ventajas cabe destacar la mínima carga computacional que representa y a su potencial utilidad para recrear la luz indirecta en una escena.
- La Luz Direccional intenta imitar los rayos de luz que provienen del sol. Para configurar este tipo de luz en la escena es necesario definir una dirección y una intensidad. En la escena este

TABLA 2.3 TABLA DE LOS TIPOS DE LUCES PARA ENTORNOS 3D

Tipos de Fuentes de Luces			
Ambiental / Omnidireccional	Luz Direccional	Luz Puntual	Luz de Área
			

tipo de luz se lleva a cabo creando rayos en paralelo que incidirán en todas las superficies con la misma intensidad y provenientes de la misma dirección.

- La Luz Puntual es una recreación bastante cercana a fuentes de luces artificiales como una bombilla. El origen de este tipo de luz es un punto del que surgen diferentes rayos de luz en todas de las direcciones o un rango limitado. La intensidad de estos rayos disminuye según se aleja del punto de origen, lo que aumenta la carga computacional al tener que calcularse para cada superficie.
- La Luz de Área es capaz de simular un panel de luz a través de la combinación de las características de los dos anteriores tipos. El origen de los rayos de luz es un área del que surgen diferente rayo en paralelo, sin embargo, a diferencia de la luz direccional este decremanta su intensidad según se aleja de la fuente. La carga computacional de este últimos es mayor a la luz direccional debido al cálculo de la intensidad.

En cuanto al último elemento de una escena 3D, la cámara, esta se encarga de definir la dirección, orientación y el campo de visión desde la cual se observa el entorno virtual. Su objetivo es calcular el color de cada uno de los píxeles de la imagen final en función de su posición y la distribución de los anteriores tipos de elementos. La Figura 2.15 muestra un esquema general del proceso.

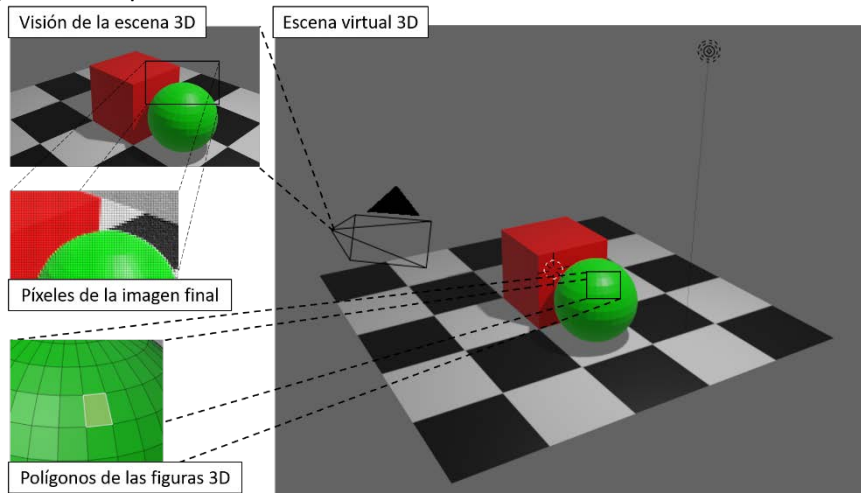
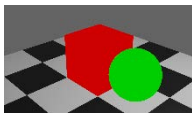
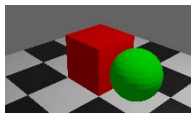
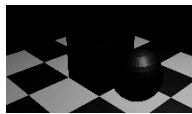
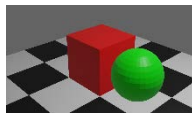
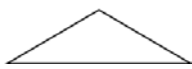
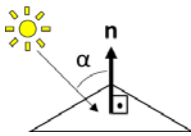
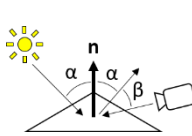
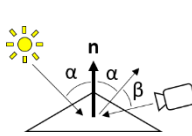


FIGURA 2.15 RENDERIZACIÓN DE UNA ESCENA 3D EN UNA IMAGEN

El resultado que se muestra en la imagen renderizada de la cámara de la Figura 2.15 no es la única representación posible de la escena 3D definida. Esto es debido a que no existe una única técnica para el cálculo del brillo de los polígonos que forman los modelos y el resultado depende de las técnicas empleadas. Las cuales son escogidas generalmente en función de la cantidad de elementos de la escena y la potencia del dispositivo. La Tabla 2.4 presenta varios tipos de técnicas para la iluminación de objetos.

TABLA 2.4 MODELOS DE ILUMINACIÓN SOBRE SUPERFICIES

Modelos de Iluminación			
Ambiental	Difusa / Lambert	Brillo Especular	Especular / Phong
			
			
$I = I_a * k_a$	$I = I_p * k_d * \cos \alpha$	$I = I_p * k_s * \cos^n \beta$	$I = I_a * k_a + I_p * [k_d * \cos \alpha + k_s * \cos^n \beta]$
<p><math>I</math> = Intensidad de luz para la superficie.  <math>I_a</math> = Intensidad de la luz ambiental.  <math>k_a</math> = Reflexión.  <math>I_p</math> = Intensidad fuente de luz.  <math>k_d</math> = Reflexión difusa.  <math>\alpha</math> = Ángulo entre la normal de la superficie y el rayo de luz.  <math>k_s</math> = Reflexión especular.  <math>\beta</math> = Ángulo entre el rayo de reflexión y la cámara.</p>			

El modelo de luz ambiental es un algoritmo que determina el color de las superficies de los objetos haciendo uso del color del propio objeto y la intensidad de la luz ambiental. Este algoritmo tiene una baja carga computacional y no depende de la posición de la cámara ni de la dirección de la superficie. Por otro lado, el algoritmo de luz difusa, también llamado iluminación de Lambert, en honor al científico Johann Heinrich Lambert que introdujo el concepto de difusión de la luz en 1760, tiene en cuenta la intensidad de las fuentes de luz, el color de la superficie, así como el ángulo de la dirección de la luz y la normal de la superficie. Este tipo de modelo de iluminación presenta unos reseñables resultados en superficies mate, es decir, en superficies que no reflejan la luz. Su funcionamiento se basa en el cálculo del producto escalar del vector normal con el vector de dirección de la luz cuando este llega a la superficie. La tercera columna de la tabla muestra el cálculo del resaltado especular. Su cálculo solamente tiene en cuenta el ángulo de diferencia entre vector de reflexión de la luz y la cámara siguiendo la fórmula descrita en la tabla. Por último, el modelo de luces Phong, denominado *Specular* en muchas herramientas 3D, fue introducido por Bui Tuong Phong [118]. Este algoritmo calcula el color de cada polígono que forma la superficie combinando el modelo de iluminación ambiental, difuso y el resalto especular. Su fórmula se presenta en tabla y consiste en el sumatorio de todas las fórmulas anteriores.

Otro de los elementos que juega un papel importante en el realismo son las sombras. Este tipo de efectos no solo es capaz de incrementar la sensación de realismo con la simple representación de una silueta proyectada sobre el suelo, sino que corrobora la percepción que

TABLA 2.5 TÉCNICAS DE SIMULACIÓN DE SOMBRAS [46,146,163,205]

Técnicas de sombras			
Projected Geometry	Shadow Mapping	Shadow Volumes	RayTracing
			

TABLA 2.6 PROCESO DE SHADOW MAPPING [163]

Proceso del <i>Shadow Mapping</i>				
				
Escena 3D	Proyección desde la fuente de luz	Mapa de profundidad desde la fuente de luz	Combinación de los mapas de profundidad	Resultado final

se tiene de la posición y tamaño de los objetos de la escena. La Tabla 2.5 muestra cuatro de las técnicas más usadas [104]: *Projection*, *Shadow Mapping*, *Shadow Volume* and *Ray Casting*.

La técnica denominada *Projection* o *Projected Geometry* fue elaborada por Blinn [16] en 1988 y se basa en el trazo de una proyección plana de la silueta del objeto sobre el suelo o un muro. Este algoritmo presenta una baja carga computacional, sin embargo, solo puede ser utilizado si la superficie sobre la que se proyecta la sombra es plana. En la primera columna de la Tabla 2.5 se muestra un ejemplo de este algoritmo en el que la silueta del flamenco ha sido proyectada sobre el suelo plano.

Por otra parte, *Shadow Mapping*, es un algoritmo que se encarga de determinar si una superficie debe ser sombreada a partir de los cálculos de un mapa de profundidad generado desde la fuente de luz. Este algoritmo fue publicado en 1978 por Lance Williams [164] y se basa en el uso de la imagen de profundidad o *z-buffer* de las cámaras. Para ello, cuando una escena va a ser renderizada (primera columna de la Tabla 2.6) el algoritmo sustituye internamente las fuentes de luces por cámaras para obtener un mapa de profundidad desde las diferentes perspectivas (segunda y tercera columna de la Tabla 2.6). Estos mapas de profundidad son combinados con el mapa de profundidad de la cámara a renderizar (cuarta columna de la Tabla 2.6) para posteriormente llevar a cabo un filtrado que determina qué píxel en la imagen final debe ser sombreado o cuáles no. El resultado final se muestra en última columna de la Tabla 2.6.

En cambio, la técnica de *Shadow Volumes*, introducida por Franklin Crow [33] en 1977, se centra en la separación del volumen del mundo virtual en dos áreas: áreas que están en la sombra y áreas que no están en la sombra. Para ello, esta técnica se encarga de crear por cada fuente de luz direccional y por cada elemento virtual un prisma que tiene como base la silueta del elemento virtual y el cual se extiende desde la posición del elemento virtual a hacia el infinito continuando en la dirección de los rayos de la luz direccional. En el caso de hacer uso de un *Spotlight* en la escena, en vez de generar un prisma detrás de cada elemento virtual, se hace construir un tronco de pirámide que tiene como base pequeña la silueta del elemento virtual y un ángulo de amplitud indirectamente proporcional a la distancia del objeto virtual y la fuente de luz. La Figura 2.16 izquierda muestra el trazado de estos volúmenes de sombra creados desde una fuente de luz puntual. Una vez trazado todos los volúmenes de sombras se calcula cuantas veces una superficie es atravesada por estos volúmenes para definir un mapa (*stencil buffer*)

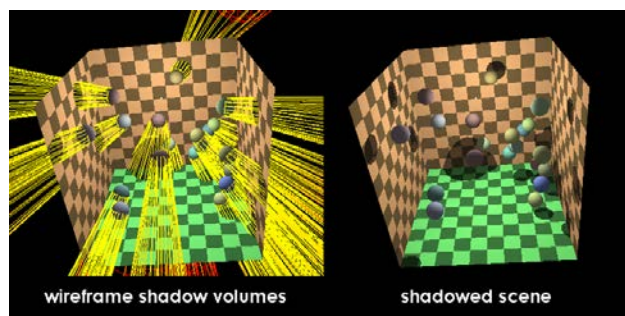


FIGURA 2.16 RESUMEN DE LOS PASOS DEL PROCESO DE SHADOW VOLUMES [207]

que funciona como una máscara para determinar las superficies a brillar y las sombras. La Figura 2.16 derecha muestra el resultado tras ejecutar esta técnica.

Por último, el *Ray Tracing* es la técnica con mayor nivel de realismo para recrear las sombras y luces. Este algoritmo fue inicialmente diseñado por Arthur Appel en 1968 [4] y mejorado en años posteriores por otros autores como Turner Whitted [162]. Este algoritmo se basa en la proyección de un rayo desde la cámara virtual por cada píxel de la imagen final hasta colisionar con una superficie. Cuando esto ocurre, este algoritmo lanza diferentes rayos (reflexión, refracción y sombra) para determinar el color de esa superficie y por tanto el del píxel. La carga computacional que presenta este algoritmo es bastante elevada dado que en una pantalla de una resolución de 1080p, este algoritmo lanzaría más de 2 millones de rayos inicialmente y estos al colisionar desencadenaría una progresión geométrica. Hoy en día, las últimas generaciones de tarjetas gráficas pueden llegar a ejecutar este algoritmo con una fluidez medianamente aceptable en algunos pocos juegos gracias a la incorporación de hardware específico que soporta esta tarea [206]. Pese a la elevada carga computacional, este tipo de algoritmo se ha utilizado en el pasado para la renderización de películas de animación o de imágenes fotorrealistas, las cuales no exigen requisitos de tiempo. En el caso de los videojuegos, se pudo incluir en la década pasada gracias a una técnica llamada *Lightmaps* que consiste en el precálculo, durante el desarrollo del juego, de las luces y sombras estáticas de la escena. La última columna de la Tabla 2.5 muestra un ejemplo de imagen fotorrealistas haciendo uso de este algoritmo.

Una de las conclusiones que se puede extraer de las técnicas mostradas anteriormente es que cuanto más información se tiene de la escena 3D más operaciones se pueden llegar a realizar para alcanzar un mayor realismo. Debido a esto, pese a que las escenas de RA cuentan con menor número de elementos virtuales y por tanto se podría deducir que son menos cálculos, presenta un reto mayor dado que la información del origen de las fuentes de luz o la forma del entorno debe de ser calculada o estimada por el sistema. Por ello, muchas de estas técnicas no pueden ser aplicadas directamente en una experiencia de RA.

Para hacer frente a esta problemática, varios autores han creado nuevas técnicas adaptado las existentes. Así pues, centrado únicamente en el efecto de sombras Haller et al [58], propone una adaptación de la técnica de *Shadow Volumes* [33]. Para ello, el autor desarrolla un nuevo concepto llamado *phantom objects* que representa algunas de las formas de los objetos reales. De tal modo, que el motor de renderización pueda proyectar de manera correcta las sombras de los elementos virtuales. La Figura 2.17 muestra un ejemplo de depuración en el que se ha definido la forma de la lata con un *phantom object*.

Por otro lado, los autores de [161] proponen un método que calcula la posición e intensidad de múltiples fuentes de luz en una escena a partir una única imagen conociendo la geometría de algún objeto de dicha imagen y su textura difusa. Para ello, los autores combinan dos algoritmos centrados en la sombra y el brillo del objeto conocido. Posteriormente, con los resultados de ubicaciones e intensidad de las fuentes de luz son capaces de replicar las mismas condiciones sobre los elementos virtuales logrando que brillen y genere sombras de forma

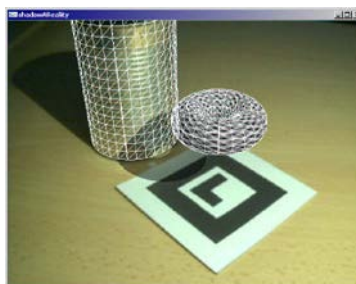


FIGURA 2.17 IMAGEN DE DEPURACIÓN. EL TOROIDE ES UN ELEMENTO VIRTUAL MIENTRAS QUE LA LATA ES REAL. SOBRE ESTE ÚLTIMO SE HA INCLUIDO UN PHANTOM OBJECT PARA DEFINIR SU FORMA [58]

adecuada. Otro uso que los autores describen en la publicación es la “reiluminación” de la escena de manera virtual pudiendo recrear una foto de los mismos objetos, pero con la ubicación de las fuentes de luz cambiadas, uso que se encuadra para la sección 2.4.7. Por su parte, los autores en [80] presentan un método para el cálculo de la iluminación de objetos virtuales en espacios exteriores teniendo en cuenta el color del cielo. Dicho método hace uso de modelos de estimación y recreación del color del cielo Perez Sky Diffuse Model [117] y Preetham Model [123] y el cálculo de la posición del Sol basada en la latitud, longitud, hora y fecha. Con ello, los autores logran estimar la dirección y color de la luz proveniente del cielo y recrearla en los elementos virtuales. En la Figura 2.18 se muestran dos imágenes publicadas por los autores en las que se destaca el cambio de luces y sombras.



*FIGURA 2.18 RECREACIÓN DE LA LUZ PROVENIENTE DEL CIELO [80]*

#### **2.4.5 Viento**

Reproducir en los objetos virtuales los efectos que las condiciones atmosféricas producen en los reales es otra forma dotar a las experiencias de RA de un mayor grado de realismo. Entre los fenómenos atmosféricos que el desarrollador puede considerar implementar destaca el del viento. La ilusión de que un objeto virtual pertenece al entorno real decaerá rápidamente si este no se mueve u oscila en la dirección del viento como el resto de los objetos que le rodean.

La simulación más realistas del comportamiento de este tipo de condiciones, o fluidos como el agua, ha sido estudiada en profundidad en campos como la física con el objetivo de simular las dinámicas de medios continuos [107]. Esto ha dado lugar a complejos modelos matemáticos que debido a que logran una fiel recreación de cómo se comportan los objetos a unas determinadas condiciones requieren una enorme carga computacional. Es por ello, que algunos investigadores han tratado de simplificar esta simulación buscando un equilibrio entre un realismo y carga computacional aceptable [39,121].

Dentro del campo de la RA, y dada la carga computacional inherente a este tipo de aplicaciones, los trabajos se han limitado a recrear estas animaciones por medio de configuraciones de vectores de fuerza y magnitudes de ruido [26], consiguiendo lograr animaciones realistas, pero totalmente independientes de la realidad. Otros trabajos de RA han hecho uso de simulación de viento pero en este caso de manera que el viento era una fuerza controlada virtualmente sin relación con las condiciones actuales del entorno aumentado [71,157].

#### **2.4.6 Sonido**

El oído junto al de la vista son los dos sentidos que más información pueden proporcionar a una persona [3]. Por este motivo, muchos sistemas de RV suelen necesitar implementar un sonido acorde a la experiencia con el fin de aumentar la sensación de presencia del usuario en el entorno virtual. De igual forma, el nivel de realismo de las experiencias de RA variará dependiendo de la forma en que el desarrollador haya implementado el sonido en las mismas. De manera ideal, no solo sería deseable que los elementos virtuales generen el mismo sonido que producirían si fuera reales, sino que este sonido debería tener el nivel de intensidad adecuado y ser percibido como proveniente de la posición correcta.



Desde la primera década del nuevo siglo, distintos investigadores han construido y evaluado *headsets* que se han centrado en mezclar el sonido del entorno real con sonido sintético generado por ordenador [124,154,155]. Todos estos dispositivos han tendido a incorporar micrófonos para capturar el sonido del entorno para su posterior mezcla en tiempo real con sonido generados por ordenador. En cuanto a nivel de implementación software, los sistemas RA han tendido a utilizar bibliotecas de audio [143,171] como OpenAL [208] o FMOD [209].

#### 2.4.7 Adaptación de la Realidad

Todas las técnicas descritas en los anteriores apartados se centran en representar adecuadamente el contenido virtual para lograr la ilusión de integración real de estos elementos en el espacio real. Sin embargo, algunos autores han seguido un rumbo completamente diferente para alcanzar dicha ilusión realizando modificaciones sobre la propia realidad [37,49,50]. Así, algunos autores como Ficher et al. [49,50] ha propuesto balancear el realismo de los elementos virtuales y la propia realidad por medio de diferentes filtros sobre esta última. Para ello, los autores hacen uso de una *tablet* con cámara sobre la que aplican diferentes filtros a las imágenes capturas de la realidad para que este se asemeje al estilo del elemento virtual a añadir. De esta forma resulta difícil distinguir qué representaciones corresponden a objetos físicos y cuáles a objetos virtuales aumentados. La Figura 2.19 muestra una instantánea en el que la realidad se ha caricaturizado como si fuera dibujos animados siguiendo el mismo estilo que la taza que es el elemento virtual introducido. Como se ve en la imagen esta propuesta de enfoque combinan los dos ejes de la taxonomía definida por Benford al combinar RA y la Modified Reality. Si bien es cierto que la apariencia final de la imagen no es del todo realista, sí se logra integrar correctamente los elementos virtuales a la visión de la realidad.



FIGURA 2.19 ESTILIZACIÓN DE LA REALIDAD PARA INTEGRAR LA TAZA VIRTUAL EN EL ENTORNO REAL [49]

Por su parte, Davis et al. [37] desarrollaron un algoritmo de procesamiento de vídeo capaz de extraer una representación imagen-espacial de la estructura de un objeto sobre el cual se ha ejercido una pequeña fuerza. En otras palabras, el algoritmo es capaz de diseñar un modelo matemático para determinar los movimientos u oscilaciones de un objeto cuando una fuerza es aplicada sobre él. Para ello, como se ha mencionado, este algoritmo se entrena a partir de pequeños fragmentos en el que el objeto aparece siendo empujado de una fuerza. Adicionalmente, el autor presenta un algoritmo capaz de simular el modelo matemático generado para crear una especie de vídeo interactivo en el que el objeto puede responder a diferentes fuerzas que incluso no fueron grabadas. Hasta la fecha, este proceso se puede llegar a conseguir de manera manual modelando el sistema estructural del objeto mediante el método de los elementos finitos (FEM) que es una técnica utilizada en el área de mecánicas de estructuras. Sin embargo, el tiempo requerido para llevar este proceso no lo hace adecuado para su utilización en RA. Las imágenes de la Figura 2.20 [210] muestran los cambios que puede ser ejercido en la representación de una instalación de parque infantil cuando se simula diferentes fuerzas aplicadas por el peso de un elemento virtual, en este caso cuando un Pokémon salta y se apoya en un punto de la instalación.



FIGURA 2.20 CAPTURAS DEL VÍDEO “POKÉMON GO AND INTERACTIVE DYNAMIC VIDEO” SOBRE LAS QUE SE INCLUYEN FLECHAS PARA INDICAR LOS CAMBIOS CON LA SIMULACIÓN DE FUERZA [210]

#### 2.4.8 Tabla Resumen de los efectos y las técnicas

La Tabla 2.7 muestra en resumen de las técnicas y soluciones propuestas para implementar los distintos efectos realistas comentados en esta sección. Tal y como se observa en la tabla, la mayoría de los efectos pueden ser conseguidos por medio de una o dos técnicas distintas. Por ejemplo, el efecto de oclusión puede ser logrado con el uso de un modelo del entorno [20] o la delimitación de bordes y profundidad [13]. Por otro lado, cada una de estas técnicas tratan de solucionar uno o dos efectos distintos. Ejemplo de ello es el empleo del uso de mapas de profundidad [20] para determinar la colisión y la oclusión. Esto quiere decir que los desarrolladores de experiencias de RA que deseen dotar a sus creaciones de una gran variedad de efectos realistas tendrán que ir escogiendo entre las distintas técnicas disponibles aquellas que mejor se ajusten a sus necesidades, examinando cuidadosamente las posibles incompatibilidades que se puedan producir a la hora de implementar varias de ellas simultáneamente o la tecnología sobre las que se va a construir la RA. Así, por ejemplo, las técnicas de *physical modelling* no debería ser usada bajo un enfoque de *Optical See-through*.

Para simplificar el proceso de desarrollo de experiencias de RA realistas sería más recomendable poder disponer de una técnica que tratase la integración de todos los efectos de forma combinada y simultánea, de tal forma que el diseñador de la experiencia únicamente necesitase seleccionar aquellos que desease activar en una determinada escena.

TABLA 2.7 RESUMEN DE LAS DIFERENTES TÉCNICAS EMPLEADAS PARA INCLUIR CADA UNO DE LOS EFECTOS REALISTAS

Techniques	Occlusion	Collision	Gravity	Light Reflection	Shadows	Wind	Reality Adjustment	Sounds
Model-based	[20]	[20]						
Depth Maps	[20]	[20]						
Contour Model	[13]							
Manual Direction			[20]					
Inertial Sensor			[87,88]					
Light Estimation				[80,161]	[58,80,161]			
Simulation of Forces						[71,157]		
Stylization Algorithms							[49,50]	
Physical Modelling							[37]	
Spatial Sound through Libraries								[143,171]

#### 2.5 End User Development en el proceso de creación de Realidad Aumentada

El término *End User Development* (EUD) se define como el “un conjunto de métodos, técnicas y herramientas que permiten a usuarios de sistemas software actuar como si fueran desarrolladores de software hasta el punto de ser capaces de crear, modificar y extender artefactos software por sí mismos” [93] (pp. 2). Aunque esta definición se centra en el software también puede ser aplicada a la creación, modificación y extensión de artefactos hardware. Esta



área de investigación, que se enmarca dentro del campo multidisciplinar del HCI, tiene como objetivo ir más allá de conseguir que los sistemas sean *easy to use*, siendo su verdadera meta la construcción de métodos, técnicas y herramientas que permitan que los sistemas sean *easy to develop* por usuario de sistemas [93]. En otras palabras, tratan de ir más lejos del objetivo de desarrollar interfaces válidas y fáciles de usar por un amplio número de persona, a uno más ambiciosos en el que cualquier persona tenga la posibilidad de desarrollar sistemas por sí mismo. A lo largo de las últimas décadas se ha podido observar que la diversidad y números de usuarios de sistemas se ha ido incrementando [138] lo que ha dado lugar a una mayor variedad de sistemas, necesidades y usos. En cambio, el desarrollo software continua siendo un proceso guiado en su mayor parte por analistas y programadores que en la mayoría de los casos no disponen de un conocimiento profundo del dominio para el cuál desarrollan, lo que se traduce en que los procesos software puedan llegar a ser largos y tediosos tan solo a la hora de comprender las necesidades del usuario en el sistema y al mismo tiempo complejos para lograr satisfacer de manera plena las necesidades de cada uno de los usuarios finales. Esta área se enfoca en reducir y eliminar esta problemática facilitando a los usuarios de estos sistemas los mecanismos necesarios para que puedan ser capaces de crear o hacer cambios, por sí solos, en los artefactos. El objetivo último es por tanto que puedan llevar a cabo los procesos de creación, adaptación, modificación y extensión de software y/o hardware de manera continua de sistemas del campo/área que conocen y dominan.

El ejemplo más representativo de esta área coincide con la proliferación de los primeros equipos personales en la década de los 70. Con anterioridad a ese momento la arquitectura de computadoras predominante se basaba en *mainframes* centralizado que tenían conectadas varias terminales. Esto hacía que fuera realmente difícil la adaptación del software por el personal de gestión del sistema a petición de los usuarios [19]. Con la aparición de los ordenadores personales esta tendencia cambió dando la posibilidad a los propios usuarios de sistemas realizar pequeñas adaptaciones puesto que los equipos tenían la suficiente potencia para compilar e interpretar código. Sin embargo, el verdadero hito fue la aparición del software VisiCalc [21] de la mano Dan Bricklin y Bob Frankston en 1978, que es el predecesor de los actuales software de hojas de cálculo. Esta aplicación no solo fue un éxito que fomentó la proliferación de los ordenadores personales [69], sino también cambió la visión que se tenía de los equipos personales que pasaron a ser vistos como instrumentos de hobby a herramientas de negocio. Los sistemas basados en el paradigma de las hojas de cálculo permiten definir modelos o plantillas de relaciones numéricas y lógicas entre las celdas de la rejilla. Aunque esta tarea pueda percibirse alejada de la programación, realmente, la formulación de esta relación es considerada programación funcional de primer orden [73]. Aunque es preciso matizar que el concepto de las hojas de cálculo ya había sido descrito por el economista Richard Mattessich [103] y que realmente el primer sistema en implementarlo fue LANPAR (LANguage for Programming Arrays at Random) en 1969 [65]. Sin embargo, sin duda es incuestionable, que VisiCalc impulsó la proliferación de los equipos personales permitiendo la elaboración de algoritmos para el cálculo numérico a un gran número de expertos de diferentes áreas como ingeniería, diseño, economistas, educadores, etcétera.

Tras este hito, los trabajos de investigación en EUD se han centrado en la capacitación de usuarios finales para la creación de artefactos en distintas áreas como el diseño de páginas web, automatización de tareas de ofimática, animación de contenido 3D, etcétera. Esto ha dado lugar a la propuesta y evaluación de diferentes estilos de interacción para elaborar algoritmos de forma sencilla como *Macros* por *Scripts* [180], *Visual Programming* [181], *Trigger-Action Programming* [158], *Programming by Example* [57], *Programming-by-specification* [94], etcétera.

Al mismo tiempo, dentro de esta área han surgido áreas más específicas o subconjuntos como el *End-User Programming* (EUP) o *End-User Software Engineering* (EUSE). La Figura 2.21 ilustra un diagrama de Venn con la relación entre estas áreas.

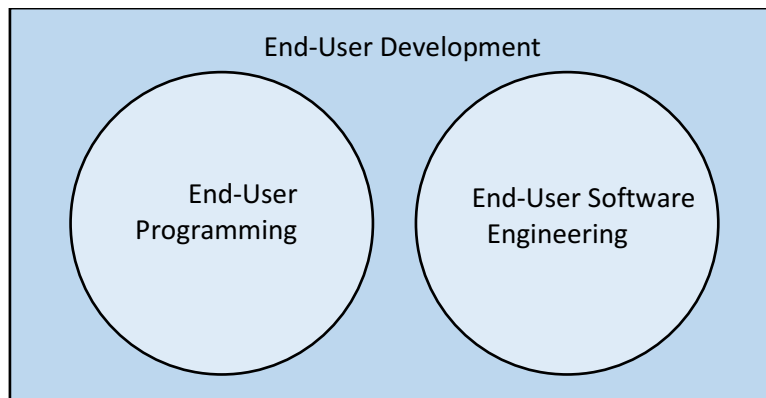


FIGURA 2.21 REPRESENTACIÓN DE LA RELACIÓN DE CADA RAMA DENTRO DEL ÁREA DE END-USER DEVELOPMENT

- *End-User Programming* (EUP): Esta rama de EUD se centra conseguir que usuarios sin conocimientos técnicos puedan crear artefactos software. Ko [79] define el *End-User Programming* como “programar para alcanzar el resultado de un programa, en lugar del programa en sí”. Como se menciona en la definición, el desarrollo del programa es para alcanzar un resultado o resolver un problema y no tanto en crear una solución software que posteriormente pueda ser modificado o extendido por otros usuarios que no son expertos.
- *End-User Software Engineering*: Esta rama del EUD se centra en resolver o mitigar los problemas que existen alrededor de los artefactos creados, modificados o extendidos por usuarios no expertos, como su seguridad, eficiencia, etcétera. Muchos de estos problemas suelen derivarse de la no aplicación de un enfoque metódico a la hora de crear el artefacto software. Por este motivo, gran parte de los esfuerzos del EUSE se encuentran dirigidos a conseguir que los usuarios finales dispongan de un proceso de ingeniería de software adaptado a sus necesidades, que puedan aplicar por sí mismos y que tenga en cuenta todas las perspectivas del producto a desarrollar.

Como se han mencionado antes el conjunto de todos los trabajos relacionados con el EUD ha proporcionado útiles métodos, técnicas y herramientas en numerosas áreas, siendo una de ellas la RA.

Dentro de la RA uno de los trabajos más importantes en relación con el EUD fue la aparición de ARToolkit [186] en 1999 que propició el aumento de desarrollos de herramientas y aplicaciones de RA para usuarios finales. Esto es debido a que como se ha explicado anteriormente, los desarrollos anteriores de RA necesitaban construir hardware específico. Si bien se trabaja de una biblioteca dirigida a su uso por desarrolladores su importancia radica en conseguir mitigar el problema de implementación de sistemas de RA al reducir el hardware necesario a un ordenador y una simple webcam.

De cualquier manera, antes de la publicación de esta biblioteca ya existían algunas herramienta de autoría enfocada a usuarios finales [125]. que permitía disponer etiquetas de texto sobre un espacio real haciendo uso de dos imágenes diferentes en las que salían un mismo marcador. El usuario tan solo tenía que marcar el mismo punto en las dos imágenes e introducir el texto a mostrar de tal manera que la aplicación podía calcular el punto de la realidad donde debía de aparecer el texto cuando se ejecutara la aplicación.

Ya en el año 2000, cabe mencionar el sistema Here&There [133], que tenía la capacidad de aumentar espacios físicos mediante audio. Para llevar a cabo la definición de los audios que se iban a escuchar en el entorno final esta herramienta proporcionaba dos estilos de interacción. Por un lado, la aplicación permitía al usuario definir el sonido y el lugar mediante un mapa 2D. Respecto a la segunda opción, el usuario se podía desplazar al lugar a aumentar y seleccionar el audio con el que quería aumentar el entorno sin hacer uso de una interfaz de usuario gráfica compleja.

Dos años más tarde, fue presentada la plataforma PowerSpace [59] un conjunto de herramientas y *plug-in* para el desarrollo de guías o manuales técnicos de RA basada en marcadores. Dicho conjunto estaba orientado para ser usado por editores de documentación técnica. Los autores de este trabajo subrayan que la RA tienen un enorme potencial en proporcionar información de montaje y mantenimiento pero que como en otros sectores la generación de contenidos de RA es bastante costosa. Para resolver este problema, los autores presentan un flujo de trabajo que hace uso de la aplicación de ofimática Microsoft PowerPoint en su primera fase para la composición del contenido de RA. Una vez diseñado el contenido a mostrar en la experiencia, éste se exporta a una herramienta que se encarga de transformar el contenido bidimensional a 3D y configurarlo espacialmente. El resultado final de este proceso es un fichero XML que describe el contenido generado y su presentación para la visualización del manual técnico en RA. Ese mismo año también se presentó AMIRE (Authoring Mixed REality) [41,54], un *framework* que propone una arquitectura de componentes para los desarrollo de herramientas de RA. Si bien el ciclo de desarrollo que se presenta junto con el *framework* no podía ser llevado por completo por usuarios no expertos, tenía como base que los expertos de otros dominios puedan proporcionar las necesidades, requisitos y restricciones para que componentes reutilizables fueran construidos por expertos programadores para ese y futuros usos. Ejemplos de herramientas de autoría diseñadas siguiendo el enfoque de este *framework* se pueden encontrar en las propuestas [169,1]. En [169] se presenta una herramienta de creación de manuales de montaje de muebles junto con una pieza software para su posterior visualización en RA. Por otro lado, los autores de [1] se centran en estudiar el diseño de una herramienta para el diseño de guías por los museos con RA. Entre las características a destacar en estas dos últimas herramientas respecto al desarrollo por parte de usuarios finales cabe mencionar la estructuración del sistema por componentes, el apoyo de interfaces de usuario tangibles y la simulación de la experiencia final mediante vídeos.

Otro enfoque centrado en la reutilización de código de las aplicaciones de RA fue presentado por la plataforma ImageTclAR [116] basada en la plataforma de edición de imágenes ImageTcl en 2004. Como describen los autores de la plataforma, existen varias partes que son comunes en la mayoría de las aplicaciones de RA y los potenciales creadores de aplicaciones RA suelen tener diferentes niveles de conocimientos de programación. Para responder a estas diferencias esta plataforma presentaba diferentes capas que podían ser usadas según el alcance y conocimientos del usuario.

De forma paralela a trabajos centrados en investigar cómo debería ser las arquitecturas de las herramientas de autoría, algunos autores centraron sus esfuerzos en la composición del contenido de RA. El trabajo de Güven y Feiner [56] presenta una herramienta de autoría de “documentales situacionales” para RA que no se centraba en ubicar contenidos multimedia individuales en un entorno real como se había hecho hasta el momento [126,133]. Los autores proponían una jerarquía de contenidos multimedia que podían ser configurados en una barra temporal. Siguiendo una idea parecida pero para experiencias de RA más genéricas, MacIntyre et al. presentan DART [97,98] en 2004. Esta herramienta se ejecutaba sobre la plataforma Macromedia Director y permitía realizar un diseño rápido de experiencias de RA por medio de una barra temporal y el uso de componente/script predefinidos que pueden ser equipados a los elementos virtuales para su comportamiento.

Poco después se presenta aiTAR [90,91] una herramienta de autoría que permite la creación y configuración de los elementos virtuales de manera inmersiva utilizando interfaces tangibles con marcadores. De esta manera, los usuarios pueden ver el resultado de cómo aparecerán los elementos virtuales en los marcadores al mismo tiempo que están componiéndolos.

Ya en el 2008, podemos encontrar la herramienta ComposAR [140], que trata de conseguir que usuarios sin conocimientos de programación puedan definir la interacción de los elementos virtuales. Para ello, esta herramienta proporcionaba 3 niveles de desarrollo en función de las habilidades o conocimientos de los usuarios. Fuera del ámbito científico, en el marco del software libre, podemos encontrar la herramienta ATOMIC Authoring Tool [188], que ofrece una

interfaz sencilla a través de la cual asignar elementos virtuales o modelos 3D a marcadores planos.

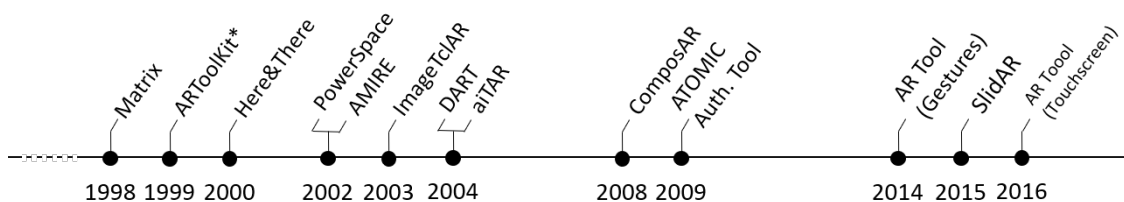
En los siguientes años los trabajos sobre autoría y creación de experiencias de RA se focalizan en estudiar nuevos enfoques que permitan disponer los elementos virtuales de una manera más natural, como los propuestos por [122,141,165]. Así, los autores de [141] presentan un sistema que permite la configuración de los elementos virtuales realizando gestos con las manos. Por su parte, los autores de [122] proponen un sistema que permite ubicar etiquetas haciendo el gestos de *slide* sobre una pantalla táctil de un móvil sin requerir ningún tipo de marcado. Por otro lado, y con un dominio de gestos más amplios pero manteniendo el uso de marcadores, los autores de [165] proponen un sistema que permite crear y editar elementos virtuales a través de gestos sobre una pantalla táctil de un dispositivo móvil.

La Figura 2.22 organiza una línea temporal las diferentes herramientas de EUD para experiencias de RA descritas en este apartado, identificando las principales tendencias y líneas de investigación, las cuales se enumeran a continuación:

- Elementos multimedia a elementos virtuales: Algunos de los primeros trabajos descritos permitían al usuario la introducción de texto, audio o imágenes como contenido para aumentar el entorno. Esta tendencia ha ido cambiando desde principios de la década del 2000 con un enfoque más centrado en contenido virtual 3D.
- Uso de módulos/componentes: A medida que fueron apareciendo más aplicaciones de RA las propuestas de herramientas de desarrollo para usuarios de sistemas se fueron enfocando a permitir el ensamblaje o composición por medio de módulos o componentes comúnmente utilizados en este tipo de artefactos.
- Interacción más natural: Al mismo tiempo que los sistemas de información han ido desarrollando interfaces más modernas, estas se han sido incluyendo o usando en las herramientas de RA. Esto ha sido debido principalmente a las ventajas y las metáforas más naturales que se pueden desarrollar.
- Uso de enfoques Markerless: En el ámbito general de la RA, la aparición de sistemas móviles más potentes y accesibles al público en general ha permitido el desarrollo de tracking ópticos sin marcadores. De este hecho se han beneficiado también las herramientas centradas enfocadas a usuarios de sistemas.

## 2.6 Resumen

Este capítulo lleva a cabo una revisión de la literatura de las dos áreas principales en las que está enmarcada la presente tesis: realidad aumentada y *end-user development*. Al inicio del



### Tendencias:

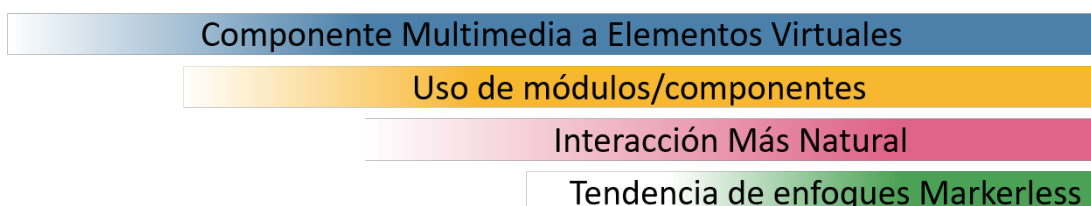


FIGURA 2.22 LÍNEA TEMPORAL CON ALGUNAS HERRAMIENTAS DE RA PARA USUARIOS FINALES

capítulo se expone la definición del término realidad aumentada que dispone de mayor aceptación de la comunidad científica. Tras ello, se presenta varias taxonomías, entre las que destaca el *Virtual Continuum* de Milgram por su sencillez y nivel de aceptación, para presentar y describir diferentes términos relacionados con la RA. Además, se incorpora algunos problemas identificados fuera del ámbito académico relacionados con el uso del término RA. En el primer apartado del capítulo, se describe el flujo general de ejecución de los sistemas de RA para presentar sus entradas y salidas. Posteriormente, se presenta la evolución cronológica que ha tenido la RA que se puede resumir en los siguientes acontecimientos. Los primeros sistemas tendían a ser pesados y voluminosos. Además, tenían limitaciones a nivel de cómputo y gráficos de modo que hacían uso de enfoques *Optical See-through* y los sistemas tracking se basaban en mecánicos, magnéticos y ultrasonidos. Ya avanzada la década los 90, se empezaron a ver los primeros sistemas *Video See-through* y enfoques *Optical tracking*. Respecto a los trabajos de investigación se enfocaron en el desarrollo de arquitecturas, exploración de casos de uso y mejora del realismo debido al incremento de la potencia computacional y el uso de PDAs y webcams. Al inicio del nuevo siglo, surgieron las primeras aplicaciones comerciales y la investigación de RA se enfocó en la creación de estas experiencias por usuarios finales. Ya en la última década, bibliotecas, aplicaciones y dispositivos de RA de diferentes empresas han empezado a consolidar el mercado. Tras esta sección, el capítulo describe los diferentes tipos de técnicas *tracking* y una clasificación basada en los sensores que se emplean. Posteriormente, se presentan las técnicas empleadas para aumentar el realismo de las experiencias de RA. Por último, se presenta la evolución y características de las herramientas de creación y autoría de RA para usuarios de sistemas que se puede resumir en los siguientes sucesos. El surgimiento de los primeros sistemas se puede situar a finales de la década de los 90, con la aparición de sistemas que permiten ubicar elementos multimedia como texto, sonido o vídeo en el entorno real. Estos enfoques fueron cambiando a lo largo de la siguiente década con herramientas con tendencias a incorporar elementos 3D virtuales, llevar a cabo una separación de componentes y proporcionar diferentes capas en función de las habilidades de los usuarios. Ya en la última década, los enfoques se han empezado a centrar en hacer uso de una interacción más natural para la edición del contenido y el empleo de las técnicas *tracking* más avanzadas como *optical tracking* basados en *markerless*.

### 3 IDENTIFICACIÓN DEL PROBLEMA Y OBJETIVOS

Este trabajo de investigación se basa en la premisa de que para liberar el potencial que la tecnología de RA es necesario facilitar el desarrollo e implementación de experiencias de RA realistas y de calidad, de tal manera que usuarios que no posean un alto perfil técnico, pero sí sean expertos en cierto dominio puedan crear experiencias y aplicaciones de RA para su campo por sí mismos, o con mínima asistencia técnica. En este capítulo se analizan los principales obstáculos y problemas que se presentan para lograr esta meta, atendiendo a la revisión de trabajos relacionados presentada en el capítulo anterior.

#### 3.1 Baja integración del contenido virtual en el mundo real en las experiencias de RA basadas en tecnología móvil

Según Azuma [7], las experiencias de RA deben ser capaces de crear en el usuario la ilusión de que el contenido virtual parezca coexistir con el resto de los elementos físicos del entorno aumentado. Tal y como se vio en el capítulo anterior, en la actualidad la mayoría de las aplicaciones de RA basadas en tecnología móvil y disponibles para el público general no logran satisfacer este objetivo, y se limitan a superponer objetos virtuales sobre la visión del usuario. En la mayoría de los casos reconocen en la imagen capturada por la cámara del móvil algún tipo de marcador o referencia a partir de la cual establecen la posición y tamaño de los objetos virtuales. Sin embargo, estos objetos no suelen comportarse de forma realista y pueden aparecer flotando en el aire y sin experimentar la fuerza de la gravedad, superpuestos sobre otros objetos más próximos al usuario en lugar de ocluidos por los mismo, y sin reaccionar a las condiciones ambientales, por ejemplo.

Tal y como se describió en el capítulo anterior, existen en la literatura distintas propuestas de técnicas para aumentar el realismo de las experiencias de RA [13,20,37,49,50,58,71,80,87,88,143,157,161,171]. Sin embargo, cada una de estas técnicas se centra en conseguir algún tipo de efecto concreto, por lo que la aplicación de varios efectos en una misma experiencia requiere combinar distintas técnicas, que pueden emplear enfoques de RA radicalmente distintos. Por ejemplo, un sistema *optical see-through* como las Microsoft HoloLens que haga uso de su sensor de distancia para la oclusión [179] no funcionaría adecuadamente con una estilización de la realidad [49,50] o simulación de la realidad basada en modelos [37]. Se echa en falta un enfoque integrador que permita seleccionar y aplicar distintos efectos para una determinada RA de forma sencilla y sin que implique aumentar la complejidad del desarrollo.

Teniendo en cuenta estos problemas se puede afirmar que *[P1] las actuales experiencias de RA disponibles para el público general ofrecen un bajo nivel de realismo y no logran generar en el usuario la ilusión de que el contenido virtual se encuentra integrado con el resto de los elementos del entorno.*

#### 3.2 Dificultad para aumentar de forma realista entornos de grandes dimensiones o exteriores

Otro de los problemas que nos encontramos a la hora de implementar experiencias de RA es la dificultad adicional que supone aumentar cierto tipo de entornos, como interiores de grandes dimensiones o localizaciones exteriores. Por una parte, las aplicaciones de RA que hacen uso de referencias o marcadores como medio para establecer la posición y tamaño de los elementos virtuales suelen tener problemas para reconocer en la imagen capturada por la cámara referencias que se encuentren alejadas de la misma, que aparecen a un menor tamaño en la imagen según la distancia respecto del usuario. Incluso dispositivos específicos para RA que utilizan cámaras de profundidad para crear modelos virtuales del espacio a aumentar resultan inadecuados para espacios de grandes dimensiones debido al limitado alcance de la profundidad que pueden llegar a percibir. De igual manera, aumentar ciertas localizaciones puede resultar complicado debido a la dificultad de reconocer referencias en condiciones de luz variable, como

suele pasar en exteriores. Por último, en algunos entornos resulta difícil encontrar particularidades u objetos suficientemente distintivos como para servir como referencias identificables, y llevar a cabo las modificaciones necesarias para introducir esas referencias en los mismo no siempre es posible.

Una posible solución podría consistir en establecer el tamaño y posición de los objetos virtuales empleados modelos 3D del entorno a aumentar en lugar de referencias. Sin embargo, no se han encontrado en la literatura propuestas que permitan emplear este tipo de enfoque de RA basada en tecnología móvil, de tal forma que permita al público general aumentar espacios de estas características.

*[P2] Es por tanto necesario disponer de un enfoque que permita la creación de experiencias aumentadas realistas en entornos interiores de grandes dimensiones y localizaciones exteriores.*

### **3.3 Falta de herramientas de autoría adecuadas para la creación de experiencias de RA realistas por usuarios finales**

Por último, se observa cómo las actuales herramientas de autoría dirigidas al público general, es decir a usuarios con un conocimiento no experto, únicamente permiten implementar experiencias muy sencillas, y en muchos casos específicas para un dominio o tipo de aplicación concreta. Así estas herramientas [45,70] suelen emplear un enfoque de autoría similar, consistente en permitir al usuario asociar elementos virtuales a marcadores o referencias en la imagen capturada por el dispositivo móvil. Como consecuencia, las experiencias que permiten implementar adolecen del problema presentado en la sección 3.1, es decir, son poco realistas y los objetos virtuales no aparecen integrados con el resto de los objetos físicos.

Además, si bien estos elementos virtuales pueden incluir animaciones o sonidos, la interacción del usuario con ellas suele ser muy limitada, de tal forma que, en la mayoría de los casos, el usuario que accede a la experiencia actúa como un mero observador de la misma. Este problema se debe en parte a que el proceso de creación se lleva a cabo empleado el mismo tipo de dispositivo que será utilizado para desplegar la aplicación final: un móvil o tableta. Así, a la dificultad de proporcionar metáforas adecuadas para permitir que un usuario sin conocimientos de programación pueda describir experiencias RA con cierto nivel de interacción o narrativa, se suma el limitado espacio disponible en la pantalla de este tipo de dispositivos para proporcionar un interfaz para llevar a cabo estas operaciones de forma clara y sencilla, y el restringido campo de visión que ofrecen la cámara en ellos.

Por último, estas herramientas de autoría emplean un enfoque “WYSIWYG” (lo que ves es lo que obtienes), de tal manera que permiten al creador de la experiencia observar directamente y en tiempo de edición el aspecto final que tendrá la misma al ser desplegada por el usuario a quien va dirigida. El inconveniente que presenta este tipo de enfoque es que es necesario presente en el propio espacio a aumentar para poder editar y modificar la experiencia, lo cual supone una limitación a la hora de aumentar espacios públicos o de acceso restringido, que pueden no encontrarse disponibles en todo momento al creador de la experiencia.

Teniendo en cuenta estos problemas se puede afirmar que *[P3] las herramientas de autoría de RA disponibles para el público general sufren ciertas limitaciones que dificultan su uso a la hora de crear experiencias de RA interactivas y de propósito general.*

### **3.4 Relación de los problemas identificados con los objetivos del trabajo de investigación**

Para dar solución a estas problemáticas se desarrollará un enfoque que integre un amplio número de efectos realistas pero que no requiera de un perfil tecnológico alto. Por otro lado, se llevará a cabo la exploración y diseño de un método alternativo del proceso de autoría mitigue los problemas identificados, pero al mismo tiempo aúne una interacción natural y la posibilidad de conocer cómo es la experiencia mientras se lleva a cabo el proceso de autoría. De manera

más concreta, se explorará el uso de la realidad virtual para conocer los beneficios y limitaciones que puede presentar en el proceso de autoría.

En este contexto, el objetivo principal que plantea esta tesis es la de facilitar la creación de experiencias de realidad aumentada realistas por personas que no disponen de un perfil tecnológico alto. De tal manera que usuarios finales puedan realizar por sí mismo el proceso de creación de experiencias RA y al mismo tiempo puedan incluir un abanico de efectos realistas con los que incrementar el realismo y, en consecuencia, aumentar la sensación de que el contenido virtual pertenece al entorno real.

Puesto que este objetivo es demasiado extenso para los diferentes dominios de aplicación y las casi infinitas posibilidades de los entornos reales, se decidió centrar los trabajos en los siguientes objetivos:

O1.- Diseño de un método para la creación de experiencias de realidad aumentada en espacios de grandes dimensiones con una variedad de efectos realistas.

O2.- Estudio de los beneficios y limitaciones del uso de la realidad virtual como parte del proceso de autoría de realidad aumentada.

La Figura 3.1 muestra un gráfico resumen de los objetivos que se abordan en esta tesis y las problemáticas que abordan cada uno de ellos.

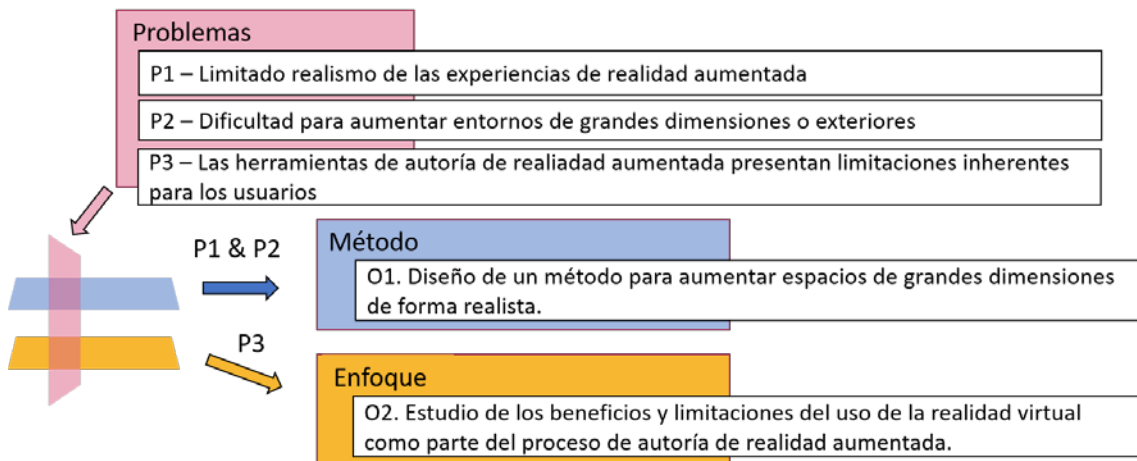


FIGURA 3.1 RELACIÓN DE LOS PROBLEMAS Y LOS OBJETIVOS



## 4 METODOLOGÍA DE INVESTIGACIÓN

Como se ha comentado anteriormente, las áreas que abarca esta tesis son la RA y EUD, las cuales a su vez están englobadas simultáneamente por las ciencias de la computación y el campo multidisciplinar *human-computer interaction* (HCI). Siendo el fin último del presente libro plasmar una contribución original del autor como parte de los requisitos exigidos en el programa de doctorado, se hace crucial e ineludible la utilización de una metodología válida con los campos que se engloban. Una metodología de investigación se puede definir como un conjunto de técnicas y métodos que son usados de manera sistemática y ordenada que guían la investigación científica con el fin de la elaboración y alcance de los objetivos [86]. Una metodología de investigación no se puede considerar como un elemento externo a la investigación puesto que la elección del método repercute en los pasos de la investigación, así como trascendencia de los resultados.

Para el caso concreto de la presente tesis doctoral, se ha hecho uso del marco de trabajo de investigación definido por Hevner et al. [63,64] denominado *Design Science in Information System Research* (DSISR). Este marco de trabajo no sólo se centra en guiar la investigación, sino que también vela por la originalidad e interés de la contribución.

Esta metodología está fundamentada en el hecho que una investigación en los sistemas de información, campo que se trata en este libro, han de combinar dos paradigmas que podrían considerarse distintos: *behavioral-science* (o *natural science*) y *design science*. El primero de ellos, *behavioral-science* es un paradigma que busca predecir y explicar el comportamiento de los individuos y organizaciones mediante el desarrollo de teorías y la verificación de estas. En cambio, *design science* es un paradigma de *problem-solving* que tiene como objetivo extender los límites de las capacidades de los humanos y las organizaciones para lograr resolver nuevas metas o lograr las actuales de una manera más eficiente. La justificación que subyace de la necesidad de la combinación de ambos paradigmas es debido su retroalimentación mutua. Así pues, en el escenario de una organización, el diseño de un nuevo artefacto de sistema de información (*design science*) debe de tener en cuenta los comportamientos que tienen y tendrán los individuos (*behavioral-science*). Posteriormente, que una vez introducido el nuevo artefacto en la organización los comportamientos resultantes deberán de ser estudiados para futuros diseños.

Para lograr compaginar ambos paradigmas, este marco de investigación está constituido por de tres ciclos: Ciclo de Relevancia, Ciclo de Diseño y Ciclo de Rigor, los cuales participan en las diferentes etapas de la investigación y están definidos sobre tres elementos definidos en la metodología: Entorno del dominio de aplicación, Base de conocimiento e Investigación de Ciencias de Diseño. En la fase inicial de la metodología, desde una visión muy general, el Ciclo de Relevancia y el Ciclo de Rigor proveen de información del dominio de aplicación y conocimiento de corpus de conocimiento respectivamente. Esta información es recogida en el diseño y construcción de un artefacto o proceso que posteriormente es evaluado a través de una o varias iteraciones del Ciclo de Diseño. Finalmente, el conocimiento que se pueda sintetizar tras las evaluaciones y análisis contribuirán en la base de conocimiento mediante el Ciclo de Rigor y aportará mejoras o avances en el entorno donde puedan ser aplicados estos artefactos o procesos por medio del Ciclo de Relevancia. La Figura 4.1 muestra de diagrama de marco de investigación definido por Alan Hevner contextualizado para la presente tesis. Las siguientes secciones describirá el detalle más información del flujo de este marco de trabajo.

### 4.1 Ciclo de Relevancia

El Ciclo de Relevancia relaciona el entorno de investigación y la investigación en curso. En su primera fase este ciclo proporciona a la investigación información acerca del entorno, las personas, el sistema organizativo y la tecnología que actualmente se usa en el dominio sobre el cual se quiere investigar. Esta información es la base sobre la cual se podrán identificar los problemas existentes y las oportunidades sobre las que se puede enfocar la investigación.

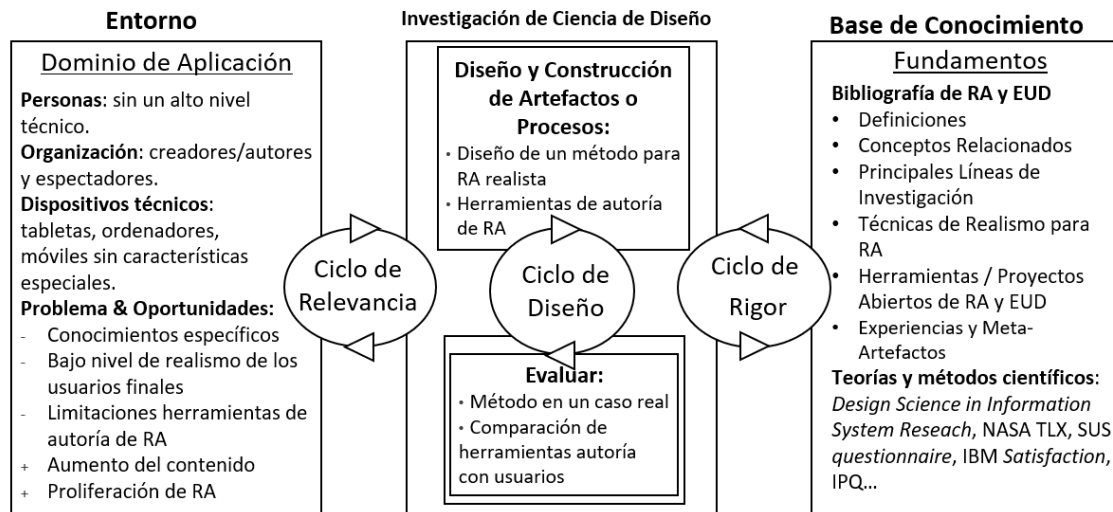


FIGURA 4.1 METODOLOGÍA DE HEVNER ET AL. CONTEXTUALIZADO A LA PRESENTE TESIS

Posteriormente, en una etapa más avanzada de este ciclo proporcionará un entorno para la realización de casos de prueba o evaluaciones. Una vez haya finalizado la investigación y haya logrado resolver o mitigar algunas problemáticas las personas de estos entornos se podrán beneficiar conocimientos y artefactos desarrollados.

En el caso concreto de la presente tesis, el espacio físico del dominio de aplicación cubre los espacios de interior de edificios como salas o aulas dónde podría ser útil una experiencia de RA para la aprehensión de conocimientos o la práctica de ellos; y los entornos de grandes dimensiones como auditorios o espacios exteriores dónde podría ser usados para la creación de experiencias de RA con fines exploratorios, principalmente lúdicos en esta tesis, y en los que la literatura no se ha centrado especialmente para lograr incluir un alto grado de realismo. En cuanto a las características más relevantes a la creación de experiencias de RA en estos entornos cabe mencionar que por lo general los entornos de interior cuentan con fuentes de luces artificiales como bombillas incandescentes, fluorescentes, lámparas, etcétera y además cuenta con la posibilidad de luz natural proveniente del sol a través de las ventanas. Estos espacios cuentan con mobiliario fijo que contiene o sobre los que se colocan una variedad de objetos físico que pueden llegar a cambiar de posición fácilmente. Respecto a los espacios amplios como los auditorios, su iluminación por lo general es artificial y cuenta con mobiliario fijo. Con relación a los espacios abierto o de exterior, su fuente de luz principal por el día es el sol que depende de la meteorología y por la noche en las áreas pobladas es artificial proveniente de farolas u otros aparatos. En cuanto a los elementos que se pueden encontrar en las áreas urbanas, por lo general se pueden hablar de elementos fijos como mobiliario urbano, árboles y edificios.

Dada la diversidad de características de los entornos y la variedad de usos que se podría dar a las experiencias de RA, esta tesis centrará el dominio de uso a la divulgación y el esparcimiento en los entornos anteriormente descritos con el propósito de permitir la simulación la aprehensión de conocimientos y la promoción del entretenimiento entre las personas. En relación con el sistema organizativo de estos contextos podemos identificar dos roles principales en todos ellos: las personas que crean experiencias RA y que además tiene un conocimiento centrado en un dominio específico como divulgadores o diseñadores de videojuegos y las personas que consumen estas experiencias como personal en visitantes o jugadores de RA.

Los usuarios finales que desean crear experiencias AR tendrán conocimientos específicos para el campo que dominan. Por ejemplo, expertos de prevención de emergencias conocerán en detalle los protocolos de emergencias que las personas deberán de seguir en caso de un suceso junto con las casuísticas para las diferentes situaciones. Además, serán los primeros en disponer acceso de primera mano a los cambios de normativa. Los diseñadores de videojuegos sí que podrían tener un conocimiento en el desarrollo de las experiencias de RA, aunque su

principal especialización sea el arte o desarrollo de la lógica de videojuegos. Por esta razón, y dado que el dominio específico de las personas que podrían crear experiencias RA puede ser más amplio solo se podrá asumir que al menos tendrán habilidades y conocimientos del manejo de la tecnología actual como móviles, tabletas, ordenadores, proyectores, etcétera a nivel de usuario.

En cuanto a las características y metas de las personas que consumirán las experiencias de RA se puede deducir lo siguiente. Las personas en formación de están interesadas en aprender y reforzar conocimientos de manera práctica. Por ejemplo, una persona en prevención de emergencias poder poner en práctica una situación de emergencia. Los visitantes de un museo podrán estar interesados, por ejemplo, en tener una visión cómo eran los hogares hace décadas o poder interactuar con réplicas virtuales halladas en yacimientos arqueológicos. En cuanto a las personas que están interesadas en el entretenimiento con la tecnología su principal objetivo será su recreación y espaciamento por lo que se debería de proporcionar experiencias que estén bien integradas en la realidad. Teniendo en cuenta estos perfiles en conjunto, de nuevo, solo se podrá asumir que dispondrán de conocimientos de la tecnología actual a nivel de usuario.

Por ende, una visión en conjunto sin entrar en profundidad sobre los requisitos y características del conjunto de espacios del entorno sobre el cual se investigará se puede resumir en:

- Usuarios sin conocimientos técnicos. Presunción de conocimientos de la tecnología actual a nivel de usuario.
- Dispositivos de móviles sin características especiales.
- Dispositivos de difusión de contenido como proyectores y pantallas.
- Dos tipos de roles de usuario en el entorno: usuarios que desean crear experiencias RA y usuarios que van a ejecutar las experiencias.
- Espacios de interior con control sobre la luz artificial como salas.
- Espacios amplios cerrados como salas de museo y auditorios dónde se podría tener control sobre la fuente de luz.
- Espacios abiertos como sitios arqueológicos y parques donde no se puede controlar la luz.

## 4.2 Ciclo de Rigor

Este ciclo es el encargado de relacionar la base de conocimiento actual con la investigación en curso. Su principal propósito es proveer los conocimientos que ya se conoce y han sido demostrados en investigaciones científicas anteriores. Este ciclo facilitará a la investigación teorías y métodos de investigación que proporcione una base científica sólida para la investigación y una base de conocimientos sobre experiencias previas sobre el área que se va a tratar. Una vez que la investigación haya dado sus frutos, el conocimiento generado será sintetizado para ser volcada a la base de conocimiento para futuras investigaciones.

En el caso concreto de la presente tesis, esta base de conocimiento proporciona el marco de trabajo que se está describiendo en este capítulo, los trabajos previos de investigadores en el campo de la realidad aumentada y el *end-user development*, los métodos y fundamentos de evaluación de prototipos como *System Usability Scale* [23], *IBM Computer Usability Satisfaction Questionnaires* [92], *NASA Task Load Index* [61], etcétera y conocimientos acerca de diseño de procesos y diseño de productos. El conocimiento proporcionado por la base de conocimiento se presenta principalmente en el capítulo de Estado de la cuestión.

En cuanto a la fase de contribución a la base de conocimiento en esta tesis se llevará a cabo mediante el presente documento y el conjunto de artículos ya publicados.

## 4.3 Ciclo de Diseño

El Ciclo de Diseño tiene como objetivo, en su primera fase, la construcción, diseño o adaptación de un prototipo o proceso teniendo en cuenta todas las restricciones y oportunidades identificadas en el dominio de aplicación y los conocimientos extraídos de la base de conocimiento, haciendo uso de los dos ciclos anteriores. Una vez realizado, en la siguiente

fase el artefacto o proceso se evalúa con usuarios de tal modo que se puedan obtener datos que son analizados para conocer si existe una contribución que pueda ser considerada como aportación. En caso afirmativo, los dos ciclos anteriores se encargarán de realizar contribuciones tanto en el dominio de aplicación como en la base de conocimiento. En cualquier caso, los resultados obtenidos en las evaluaciones pueden ser empleados para refinar el artefacto o proceso, llevando a cabo nuevas iteraciones en el ciclo de diseño en búsqueda de una contribución.

En el caso concreto de la presente tesis, una vez que se analizaron en profundidad los requisitos y oportunidades sobre el rango de características del entorno y se sintetizó la base de conocimientos se llevará a cabo dos ramas principales de investigación. Por un lado, se diseñó un método de creación de experiencias realistas en entornos de grandes dimensiones que no requiera de conocimientos técnicos altos. Por otro lado, se diseñó un enfoque de autoría de RA que haga uso de la RV inmersiva durante parte del proceso. Posteriormente, cada artefacto de diseño se evaluó haciendo uso del ciclo de diseño. La mayoría de las investigaciones centradas en usuarios, campo que se trata en este libro, tiene una orientación o carácter: descriptivo, relacional, experimental o, incluso, cubren varias de estas categorías a la vez [89]. Dado los objetivos que se plantean, el primero trabajo tiene un carácter principalmente descriptivo que describirá cómo fue la ejecución del método y los resultados del proceso. Mientras que el segundo será principalmente experimental y se centrará en identificar las causas de una situación o conjuntos de eventos para poder llegar a proclamar con alta probabilidad de que “X es responsable de Y” [131] por medio de experimentos controlados.

## 5 CREACIÓN DE EXPERIENCIAS DE REALIDAD AUMENTADA REALISTAS

El objetivo de este capítulo es proponer **un método para el desarrollo de experiencias de RA** que permita aumentar de forma realista espacios de grandes dimensiones. Este método se basa en la combinación de las técnicas de RA basadas en el modelado del entorno y las posibilidades que proporcionan los actuales motores de videojuegos para la creación de entornos virtuales realistas.

Asimismo, se presenta **un prototipo de sistema de RA** que valida el método propuesto en un caso real donde un espacio interior de grandes dimensiones es aumentado. Los resultados del uso de este prototipo han sido empleados para la definición de **una arquitectura de sistema RA** que permite generar experiencias de este tipo aplicando el método propuesto. Posteriormente, se presenta un segundo caso real donde un espacio exterior de grandes dimensiones es aumentado validando la posibilidad de flexibilizar parte del proceso para llevar a cabo algunas operaciones del proceso como es la alineación en el tiempo de ejecución de la experiencia RA.

### 5.1 Método de creación de experiencias de RA realistas

Tal y como se ha explicado en los capítulos anteriores, muchos de los actuales sistemas de realidad aumentada no consiguen producir en el usuario la ilusión de que los objetos virtuales generados por el ordenador pertenecen al entorno físico. No solo los objetos físicos y virtuales pueden no aparecer alineados entre sí, sino que estos últimos no suelen experimentar ni responder a las condiciones y cambios en el entorno de la misma manera que los primeros. Con el fin de solucionar estos problemas y permitir el desarrollo de experiencias aumentadas en espacios de grandes dimensiones se propone hacer un uso combinado de las técnicas de RA basadas en modelado de entorno, con las posibilidades que ofrecen los actuales motores de videojuegos para dotar de un alto grado de realismo a los elementos virtuales. Por una parte, las técnicas de RA basadas en modelos permiten alinear objetos virtuales y físicos, e implementar efectos de oclusión y colisión en función de la geometría del entorno a aumentar y la posición actual de los objetos virtuales. Por otra parte, los actuales motores de videojuegos proporcionan extensas bibliotecas de componentes que permiten simular de manera realista tanto los movimientos y respuestas de objetos físicos en un determinado entorno como el comportamiento y los efectos de la luz sobre los mismos.

Siguiendo este enfoque, la experiencia de RA se configura empleando una réplica tridimensional virtual del entorno a aumentar. A partir de esta réplica se compone una escena en la que se sitúan los objetos virtuales con los que se desea aumentar el entorno real, y sobre la que se aplica un conjunto de efectos que simulen de forma realista las condiciones del entorno replicado. Una vez configurada la escena, la réplica virtual se superpone sobre la visión del usuario alineando los componentes virtuales que representan los objetos fijos del entorno sobre la visión que el usuario tiene de los mismos, y aplicando sobre los mismos una textura de forma que se vuelvan invisibles. De esta forma los objetos virtuales aparecerán correctamente integrados en el entorno real, de tal forma que su tamaño y proporción se corresponderá con la distancia a la cual se encontrarían del usuario, y parecerán comportarse y responder a las condiciones del entorno previamente simuladas en la réplica.

El método propuesto se organiza en cuatro fases:

1. Modelado del Entorno: Esta etapa se centra en la creación de una réplica tridimensional del entorno a aumentar. El resultado final es un modelo que define la forma o geometría del entorno real.
2. Composición de la escena de RA: Este conjunto de tareas incluye la alineación de la réplica 3D y la configuración de los elementos virtuales.

3. Simulación de las condiciones ambientales: Este conjunto de tareas se encarga de simular y recrear las condiciones ambientales del entorno dentro del espacio virtual.
4. Definición de las animaciones en la escena e interacciones con el usuario: Esta etapa se enfoca en definir las animaciones que tendrán los elementos virtuales y las reglas que gobernarán las interacciones entre ellos y las acciones que hagan los usuarios.

### 5.1.1 Modelado del Entorno

La primera fase del método propuesto se centra en la construcción de una réplica tridimensional del entorno que posteriormente es introducida y usada a un entorno virtual. El proceso de modelado de entorno ha sido estudiado en varias áreas como la robótica [115] y la visión artificial [130]. Estas áreas han diseñado diferentes técnicas haciendo uso de una variedad de sensores como escáneres láser [34], cámaras de profundidad [44], cámaras estéreo [53] y sistemas SLAM (*Simultaneous Localization and Mapping*) [38], etcétera. Si bien estas soluciones presentan algunas limitaciones la mayoría de ellas permiten llevar a cabo el proceso de manera casi automática. Sin embargo, para la propuesta de este capítulo se ha de tener presente el dominio de usuarios al cual está dirigido. Estos podrían no disponer de acceso a sistemas tan específicos o necesitar conocimientos adicionales. Manteniendo esta idea se toman en consideración los siguientes tres enfoques para capturar el entorno real:

- Modelado usando técnicas de SLAM: Los sistemas SLAM principalmente han sido diseñados y utilizados en el ámbito de la robótica y la conducción de vehículos autónomos. Estos sistemas permiten la creación de modelos o mapas de entornos desconocidos mientras se estima la trayectoria realizada hasta el momento durante la exploración. Hasta la fecha, expertos de estas áreas han publicado numerosas variantes de esta técnica [38,111,112,147,152] en función de los tipos de sensores a utilizar, las limitaciones de tiempo de cálculo, la precisión y el tipo de resultado del entorno (bidimensional o tridimensional). Dentro de este repertorio, los Visual SLAM podrían ser considerados los más factibles para ser usados por un usuario de sistema puesto que precisa un dispositivo que pueda disponer de una o varias cámaras para capturar el entorno. Condición que se dan en los actuales teléfonos móviles. La ventaja principal de este enfoque es que los creadores tan solo tienen que recorrer el entorno para capturarlo. Sin embargo, esta técnica cuenta con algunas limitaciones. El principal inconveniente es que las características del entorno pueden tener una gran dependencia en la precisión del resultado final. Los algoritmos SLAM tienen como objetivo detectar e identificar marcas que caracterizan el entorno y calcular diferencias de estas marcas entre fotogramas para obtener la geometría que forman el entorno. Por ello, espacios con una intensidad lumínica muy alta o baja podría dificultar la detección de estas marcas, así como espacios donde pueda haber una falta de estas marcas como espacios abiertos. Los entornos con espejos y superficies con alta reflexión también presentan dificultades puesto que estos algoritmos pueden identificar incorrectamente algunos reflejos como marcas y por ello emplearlos en los cálculos. Además, otro de los problemas que plantea esta técnica es la disminución de la precisión cuando los entornos es de gran tamaño debido a la acumulación del error en la estimación del espacio recorrido. Por ello, esta técnica puede presentar mejores resultados en espacios pequeños que en entornos de grandes dimensiones, aunque también hay que recalcar, que esta problemática se puede mitigar con la incorporación de correctores de *Loop Closure* en el algoritmo para la reevaluar estimaciones cuando se atraviesa una zona ya conocida. Por último, otro de los problemas que se presenta con este proceso automatizado es que el resultado final suele ser un modelo de una única pieza. Por lo sería necesario trocearlo en bloques para especificar cuáles de ellos representan el suelo y cuales superficies que podrían generar oclusión.
- Reconstrucción utilizando la información de planos o mapas: Consiste en el modelado del entorno a partir de planos o mapas del entorno a aumentar, mediante la utilización de

herramientas y software de modelado. Este enfoque tiene la ventaja de poder modelar un entorno sin necesidad de desplazarse a él y no requerir dispositivos especiales. Sin embargo, este proceso puede ser largo y complejo debido a la necesidad de interpretación de los planos y necesidad de conocimientos previos para el manejo del software. Además, en algunos casos el acceso a los planos del entorno a aumentar puede ser una dificultad añadida.

- Modelado del entorno en un espacio virtual: Consiste en la construcción de un modelo 3D del entorno a aumentar dentro de un espacio virtual mediante el uso de bloques que representan la geometría de entorno. Para su creación, una imagen o vídeo del entorno real es situada en el fondo del espacio virtual. Dicha representación del entorno real es utilizada como soporte para la superposición y alineamiento de distintos bloques hasta su correcto solapamiento. La utilización de distintos bloques permite identificar las diferentes superficies del entorno, como el suelo por el que puede moverse un personaje virtual o los objetos que potencialmente pueden generar oclusión.

Teniendo en cuenta las limitaciones de cada enfoque y tomando en consideración que debe de ser posible de utilizar en un amplio rango de entornos, el enfoque de modelado del entorno en un espacio virtual puede considerarse como una opción válida. Sin embargo, el resto de las opciones podrían ser consideradas en condiciones determinadas.

La Figura 5.1 muestra un ejemplo del proceso de modelado por bloques de un entorno que tiene dos zonas de espacios distintos. La figura de la izquierda muestra un entorno real en el que se diferencia una zona interior (una cafetería) y un espacio abierto (una calle arbolada). En la imagen derecha se muestra el resultado de la superposición bloques en cada uno de los entornos. En la zona interior, se ha colocado un bloque para modelar la geometría de la mesa y otro para el servilletero. La zona exterior se compone de dos bloques para representar los diferentes niveles del suelo y cinco bloques verticales para las columnas y árboles. Adicionalmente, se ha utilizado un bloque que representa la ventana que separa ambas zonas.



*FIGURA 5.1 IZQUIERDA MUESTRA UNA IMAGEN DE UN ENTORNO REAL A AUMENTAR. LA IMAGEN DERECHA MUESTRA LA SUPERPOSICIÓN DE BLOQUES PARA LAS ZONAS DEL ENTORNO [110]*

### **5.1.2 Composición y alineación de la escena de RA**

La segunda fase del método propuesto se centra en definir la escena virtual que será usada durante la ejecución de la experiencia de RA. Como se ha explicado en la visión general de la propuesta, el objetivo es diseñar una escena virtual sobre la réplica del entorno a aumentar, alinearla sobre la visión que el usuario tiene actualmente del entorno y aplicar distintas texturas transparentes o semitransparentes sobre sus elementos para conseguir que los objetos virtuales aparezcan correctamente integrados en el mundo. Este proceso se llevará a cabo empleando un motor de videojuego o plataforma de desarrollo de mundos virtuales como Unity o Unreal Engine, que proveen editores para construir entornos virtuales 3D a partir de un conjunto de modelos 3D importados a la herramienta.

Además de los objetos virtuales con los que se quiere aumentar el entorno real la composición de la escena se llevará a cabo mediante 3 elementos principales:

- Panel de captura de vídeo: Al fondo de la escena se sitúa un plano sobre cuya textura se proyecta un video en tiempo real del entorno a aumentar capturado por la cámara del sistema de RA, y que sirve de referencia de la visión que el usuario tiene del espacio que le rodea.
- Cámara virtual: La cámara representa el punto de vista del usuario del sistema de RA. Este elemento se sitúa en frente del panel de captura de video, de tal forma que el campo de visión de la cámara se corresponde con el de la cámara del sistema de RA.
- Modelo del entorno: Entre la cámara virtual y el panel de video se sitúa el modelo del entorno obtenido del paso anterior.

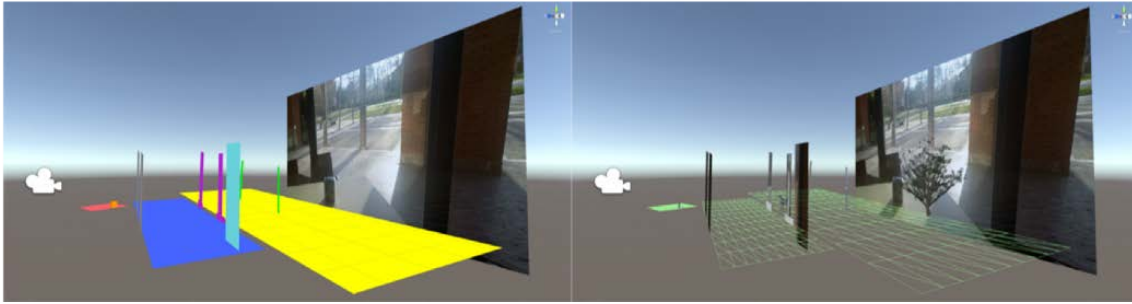
Una vez dispuestos estos 3 elementos en la escena, se deben configurar su tamaño, posición y distancia de tal forma que desde la perspectiva de la cámara virtual los elementos del modelo aparezcan superpuestos sobre el video, coincidiendo sobre los contornos de las partes del entorno que cada uno de ellos representa. Cuando estos 3 elementos se encuentran alineados, se realiza la modificación de texturas de los bloques. Se distinguen 3 tipos de textura en función de la capacidad de oclusión del bloque:

- Bloques que no generan oclusión: Son las partes del modelo que se corresponden con superficies de apoyo sobre las que se situarán los elementos virtuales con los que aumentar la escena. Se incluyen en esta categoría el suelo, las escaleras, los muros de delimitación del entorno virtual, etcétera. Estos bloques resultan invisibles a la cámara de manera que se observa la zona del vídeo correspondiente a la parte del entorno que representan.
- Bloques que generan oclusión: Sobre los bloques con capacidad de oclusión se aplica una textura con la zona del vídeo de la parte del entorno que representan. Esto permite que permanezcan visibles desde la perspectiva de la cámara y que oculten parcial o completamente aquellos objetos virtuales situados entre ellos y el plano del vídeo.
- Bloques semitransparentes: Algunos entornos pueden contener elementos translucidos como cristales o superficies semitransparentes. En el entorno real estas superficies pueden generar cambios de forma o color sobre los objetos situados tras ellas. Por tanto, los bloques que las representan deben ser capaces de aplicar filtros que simulen dichas transformaciones sobre los objetos virtuales situados tras ellos.

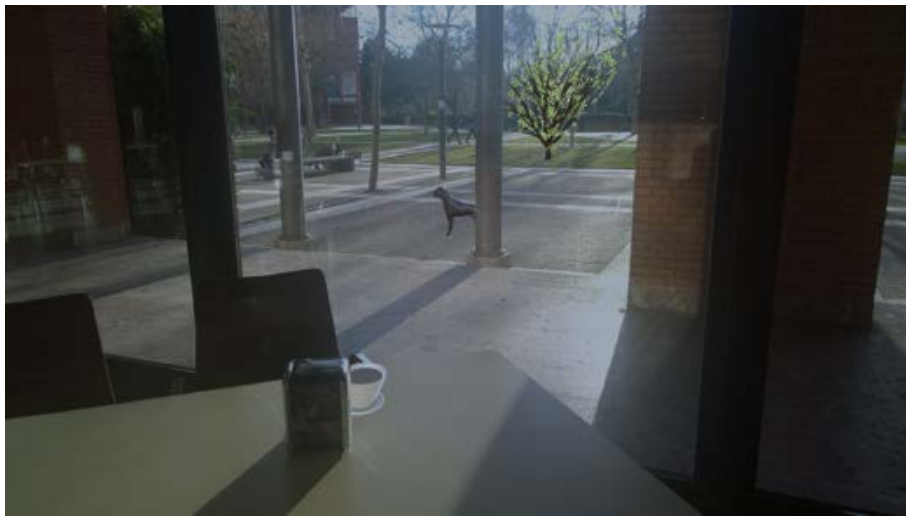
Una vez aplicadas las texturas correspondientes, el último paso de la composición consiste en situar sobre la réplica virtual del entorno los objetos o personajes virtuales con lo que se quiere aumentar el espacio real. Los elementos virtuales se irán situando sobre el modelo de réplica del entorno. Ambos entornos mantienen la misma perspectiva por lo que se crea la apariencia de que el elemento virtual está presente en la correspondiente ubicación del entorno real. Si se dota de movimiento a los objetos o personajes virtuales, estos pueden colisionar con los bloques que conforman la réplica del entorno, independientemente de la textura que se haya aplicado a los mismos.

La Figura 5.2 izquierda muestra una configuración de escena básica en la que se muestra el panel que video, la cámara virtual y la réplica del entorno a aumentar situada entre ambos. Esta réplica se compone de tres bloques horizontales que representan dos superficies de suelo (amarillo y azul en la imagen) y una mesa (rojo en la imagen), 6 bloques verticales que representan distintas columnas, paredes y tronco de árbol (morado, azul claro y verde en la imagen) y un servilletero (naranja en la imagen), y dos bloques semitransparentes para representar los cristales (gris oscuro). La Figura 5.2 derecha muestra la escena una vez que se han aplicado las texturas transparentes para las partes del modelo que no generan oclusión (suelos), texturas de videos para las que sí la generan (columnas y paredes) y se han introducido diferentes elementos virtuales como un árbol, una taza de café y un perro. La Figura 5.3 muestra la visión que obtiene desde la cámara virtual y donde se aprecia la correcta integración de estos elementos en el video del entorno real.





*FIGURA 5.2 LA IMAGEN IZQUIERDA MUESTRA LA CONFIGURACIÓN BÁSICA DE LA ESCENA VIRTUAL. LA DERECHA MUESTRA LA COMPOSICIÓN DE LA ESCENA CON ELEMENTO VIRTUALES Y LA RÉPLICA CONFIGURADA [110].*



*FIGURA 5.3 VISIÓN QUE SE TIENE DE LA CÁMARA VIRTUAL QUE GRABA LA ESCENA [110].*

### **5.1.3 Simulación de las condiciones ambientales**

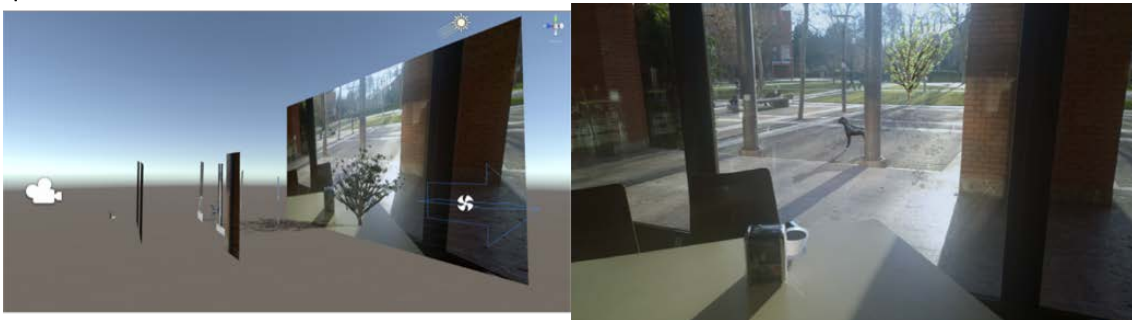
Una vez la escena virtual ha sido compuesta el siguiente paso consiste en incluir en la misma distintos elementos que permitan simular sobre los objetos virtuales los efectos que las actuales condiciones de luz y ambiente en el entorno tendrían sobre los mismos. Para implementarlo, se añaden a la escena diferentes efectos y componentes disponibles entre las bibliotecas del motor de videojuegos utilizado para el desarrollo. A modo de ejemplo, de entre la amplia variedad de efectos actualmente disponibles en este tipo de librerías el creador de la escena virtual puede añadir:

- Fuentes de luces: los actuales motores de videojuegos disponen de una amplia gama de efectos de luz mediante los que simular y reproducir las condiciones lumínicas y ambientales de una amplia gama de situaciones (atardecer, niebla, sol intenso...). Aplicando estos efectos a la escena RA es posible variar el brillo y reflejo de los objetos virtuales de tal manera que estos se asemejen a los que se producirían en el entorno real.
- Viento: Los últimos motores de videojuegos disponen de algoritmos de comportamientos de partículas que pueden recrear efectos de viento con un bajo consumo computacional. Añadiendo a la escena distintas “fuentes de viento” y variando su intensidad y posición en la escena se pueden reproducir los efectos que el viento tendría sobre los objetos virtuales de encontrarse estos situados en el entorno real.
- Gravedad: La mayoría de los motores de videojuegos disponen de implementaciones para recrear los efectos que la fuerza de gravedad tiene sobre los objetos en un entorno virtual. Aplicando este efecto los objetos y personajes de la escena se caerán y moverán sobre la

misma siguiendo trayectorias parecidas a como lo harían en el mundo real, contribuyendo a dotar de realismo a la escena.

- Sonido espacial: El sonido es otra característica que puede incrementar el realismo de la experiencia de RA [170,171]. Asociando clips de audio a los objetos virtuales en la escena y variando su intensidad se puede incrementar la sensación de que se encuentran físicamente integrados en la misma.

Con el fin de ilustrar esta etapa de la composición la Figura 5.4 izquierda muestra la escena del paso anterior una vez que se ha introducido sobre la misma algunos efectos de ambiente. Concretamente se ha añadido en la parte superior una fuente de luz que simula la intensidad lumínica del sol durante la tarde, y una fuente de viento situada en la parte inferior derecha. En la Figura 5.4 derecha se aprecian los efectos que estos elementos tendrían sobre los objetos virtuales. Tal y como se observa tanto el perro como la taza generarían sombras con una dirección e intensidad semejante a las generadas por las columnas y paredes reales que les rodean. De igual forma las hojas del árbol virtual se moverían ligeramente en la misma dirección que los árboles reales cercanos.



*FIGURA 5.4 IMÁGENES DEL PROCESO DEL PROCESO DE INCLUIR EFECTOS REALISTA EN LA ESCENA. LA IMAGEN IZQUIERDA MUESTRA LA ESCENA VIRTUAL MIENTRAS QUE LA DERECHA ILUSTR EL RESULTADO QUE SE PUEDE OBSERVAR DESDE LA CÁMARA VIRTUAL [110].*

#### **5.1.4 Animación de Escena e Interacción con los Usuarios**

Una vez que las condiciones del entorno han sido recreadas en la réplica del mundo virtual, se puede llevar a cabo la fase de animación e interacción con los usuarios. Los actuales motores de videojuegos también incluyen sofisticados sistemas de control de animación. Por lo que, haciendo uso de ellos, es posible animar la escena RA para mejorar el realismo de las experiencias. Este proceso consiste en dos tareas principales:

- Definición de un conjunto de clips de animación para representar los diferentes estados de un elemento virtual en la escena.
- Especificar la lógica que dirige la activación de los clips de animación.

Los motores de videojuegos dan soporte a la segunda tarea permitiendo a los diseñadores definir una máquina de estados que controla cuando activar una animación, repetirla, pararla o cambiar de una a otra. En función de los tipos de condiciones que desencadenan su activación, se puede distinguir:

- Animaciones no interactivas. En las cuales la animación es automáticamente activada sin interacción del usuario. Este tipo de animaciones mejora el realismo de la escena y está activo, aunque el usuario no realice una interacción. Un ejemplo de este tipo podría ser incluir un movimiento periódico en la cola de un perro virtual.
- Animación interactiva con la escena RA. En este tipo de animaciones son producidas por la interacción que realiza el usuario en la escena RA. Para ello, es necesario proporcionar al diseño mecanismos que permitan la asociación entre diferentes estilos de interacción del usuario con desencadenantes de ejecución de animaciones ya predefinidas. Por ejemplo, sistema de reconocimiento de voz y gestos son usados con frecuencia en sistemas *wearable*

de RA [81]. En este caso, el diseñador debería de especificar los comandos y gestos que el sistema reconoce y enlazarlos con el desencadenante de transiciones de animaciones de la máquina de estados.

La imagen de la Figura 5.5 muestra la visión tras animar la escena como se ha indicado en esta sección. En este caso hemos definido cuatro animaciones para que el perro pueda pasearse en las cuatro direcciones y una cuando el perro se encuentra quieto, las cuales se activan desde un *gamepad*.



*FIGURA 5.5 IMAGEN CAPTURADA DE LA ESCENA FINAL DE RA [110]*

## **5.2 Implementación del Prototipo de Motor de RA**

Con el objetivo de validar la propuesta extendimos una plataforma de juegos llamada GREP (Game Rules scEnario Platform) [168] con un módulo que diera soporte al desarrollo de experiencias de realidad aumentada siguiendo el método propuesto. La plataforma GREP está enfocada a la creación de videojuegos educativos y está implementada mediante el motor de videojuegos de Unity. El módulo de RA (AR GREP) expande las funcionalidades de GREP para soportar ejecutar juegos de RA. En esta sección, introduciremos brevemente el sistema GREP para posteriormente describir la implementación del módulo de RA.

### **5.2.1 GREP**

El sistema GREP proporciona un conjunto de herramientas para facilitar la definición y ejecución de juegos digitales educativos. Más específicamente, GREP proporciona dos tipos de aplicaciones: GREP Editor, que da soporte al diseño de escenas de juegos y definición de reglas para el juego, y GREP Player, que permite interpretar los diseños generados en el editor y generar entornos virtuales 3D para los juegos.

GREP Editor incluye una herramienta de interfaz gráfica que permite a los diseñadores crear escenas estáticas añadiendo, modificando o borrando recursos gráficos desde diferentes repositorios de modelos 3D. El usuario puede empezar definiendo una escena desde cero o seleccionar un modelo de entorno predefinido del repositorio de escenarios de la plataforma. Esta escena puede ser enriquecida mediante la inclusión y modificación de sus elementos. Una vez definida la escena, el diseñador pueden determinar acciones, comportamientos de los elementos y reglas de juego. Finalmente, el diseño puede ser exportado como fichero XML.

GREP Player que permite procesar las especificaciones creadas usando GREP Editor para generar el juego 3D. En primer lugar, procesa los ficheros XML que contiene los diseños de los juegos, recupera de los repositorios de la plataforma los modelos 3D en ellos referenciados y los instancia, posiciona y escala según las indicaciones de los ficheros XML para construir la escena virtual. Una vez que la escena virtual es reconstruida, GREP Player procesa las definiciones de las reglas del juego, configuraciones adicionales y activa las animaciones definidas.

### **5.2.2 AR GREP**

El módulo RA desarrollado para la plataforma GREP permite ejecutar experiencias de RA creadas usando GREP Editor y un modelo 3D del entorno a aumentar.

Siguiendo el enfoque descrito en la sección anterior, AR GREP soporta este tipo de experiencias usando un tipo especial de escena RV (AR Scene) que incluye una cámara virtual fija y un panel de fondo que renderiza el vídeo capturado por el ordenador o dispositivo que ejecuta el sistema. La distancia del panel y la cámara está especificada en las descripciones contenidas en los ficheros XML. Dentro de este fichero también se especifica el modelo de bloques 3D que recrea el entorno a aumentar. Cada uno de los bloques que componen la réplica presenta propiedades específicas para su nivel de transparencia u oclusión.

El módulo AR GREP procesa esta información por cada bloque en el modelo y provee la correspondiente textura para producir el efecto deseado. Por último, para dar soporte a la recreación del ambiente con la configuración real, AR GREP puede interpretar etiquetas en el fichero de XML especificando la posición, orientación e intensidad de las fuentes de luz, cajas de viento y activar la fuerza de gravedad por cada elemento virtual en la escena. La definición para enriquecer las experiencias con animaciones y sus estados también son definidos.

Actualmente, AR GREP editor no proporciona soporte al diseño de modelos 3D que represente el entorno. Este proceso necesitó llevarse a cabo usándose Unity Editor añadiendo y modificando bloque directamente en la escena RA, siguiendo el enfoque descrito en la sección (Composición de escena de RA). Para dar soporte a este proceso, la plataforma fue provista de bloques 3D que incluían un script con propiedades específicas para la construcción de la réplica. Una vez construido el modelo, el diseñador puede añadir el entorno definido como escenario predefinido para el entorno, así más tarde puede ser recuperado desde el GREP Editor. Hasta ahora, la posición del fondo y la distancia a la cámara requiere superponer el modelo 3D. El fichero de ambiente debe ser incluido como fichero XML manualmente.

### **5.3 Caso de Validación**

El objetivo principal de la evaluación fue probar el realismo de la experiencia implementado usando una configuración real. Se desarrollaron dos juegos mediante AR GREP Player que fueron ejecutados en el auditorio de la Universidad Carlos III de Madrid, en el contexto de un evento organizado para la Noche de los Investigadores. Al tratarse de un evento de acceso libre, el público estaba compuesto por participantes de diferentes edades interesados en la ciencia. Un total 232 personas asistieron al evento. El juego se ejecutó en el propio auditorio. Para integración de los elementos virtuales del juego en el mundo real, se situó una cámara en el centro del escenario apuntando a las gradas. La imagen capturada por la cámara fue procesada por el módulo de AR GREP y añadida a la representación virtual del entorno. Los participantes vieron el vídeo proyectado en una gran pantalla colocada frente al escenario.

Con el objetivo de involucrar a la audiencia, se diseñaron dos juegos para ser jugados por la multitud del público. Se organizó al público en 2 equipos, y cada uno de ellos se encargaba de controlar un personaje en tiempo real. Los personajes de los videojuegos se movían a lo largo de los pasillos alrededor del público. Los participantes tenían que hacer uso de sus teléfonos móviles para votar el siguiente movimiento del personaje que su equipo controlaba. Cada segundo, el movimiento más votado se llevaba a cabo. En el primer juego, los personajes tenían que recolectar monedas repartidas por los pasillos del auditorio. El personaje que recolectaba más monedas antes de que terminase el tiempo ganaba la partida. El segundo juego consistió

en una adaptación del “juego del pillapilla” en el que un personaje tenía que huir y recolectar monedas, mientras que el segundo tenía como objetivo capturarlo antes de que se acabase el tiempo. Para aumentar la dificultad del juego, algunas piedras virtuales aparecían y desaparecían en mitad del pasillo impidiendo el paso de los personajes. Ambos juegos requerían que la audiencia se coordinase para decidir el siguiente movimiento.

### 5.3.1 Modelado del Entorno

La implementación del juego requirió modelar todo el auditorio que tiene capacidad para 1052 personas. El área de butacas tiene unas dimensiones máximas de 29 metros de ancho y 24 metro de profundidad y tiene una inclinación hacia abajo mirando hacia el escenario. La Figura 5.6 muestra un plano de las instalaciones. Las filas de las butacas estaban divididas en 5 bloques separados por dos pasillos horizontales y 5 pasillos de escaleras verticales. Para realizar el proceso de modelado, en primer lugar, se tomó una foto con la cámara desde el centro del escenario a las gradas. La Figura 5.7 muestra dos imágenes para contextualizar la forma del auditorio, la imagen de la izquierda muestra el escenario desde un punto de las gradas. La imagen de la derecha muestra las gradas del auditorio desde el centro del escenario. Esta imagen fue renderizada como panel de fondo de la escena RA mediante parte del módulo de AR GREP en Unity Editor. Se situó una cámara virtual dentro de la escena y fue posicionada y orientada a la imagen virtual. Se ajustó el *Field of View* (FoV) de la cámara virtual para que tuviera el mismo valor que la imagen capturada. La Figura 5.8 muestra la cámara virtual y el panel de fondo y cómo en función del FoV el panel de fondo siempre cubre la visión de la cámara. Una vez configurada la cámara y el fondo, se comenzó a modelar los pasillos y las escaleras mediante bloques. El modelado de esta parte necesitó de 9 bloques que fueron conectados replicando las direcciones de los pasillos. Tras esto, los bloques se inclinaron haciendo uso de un script prediseñado para que se ajustasen correctamente a la inclinación de las escaleras. La Figura 5.9 muestra cómo se distribuyeron los bloques para representar los pasillos y cómo fueron ajustados posteriormente para que coincidieran con el panel de fondo. Para evitar que los personajes de la escena se cayeran de la escena, se colocaron unos muros invisibles delimitando los pasillos. Finalmente, se añadió un conjunto último conjunto de bloques entre las filas de asientos para ocultar parcialmente las partes no visibles. En total se emplearon 28 bloques para modelar el auditorio. La Figura 5.10 muestra los bloques que forman la réplica del entorno.

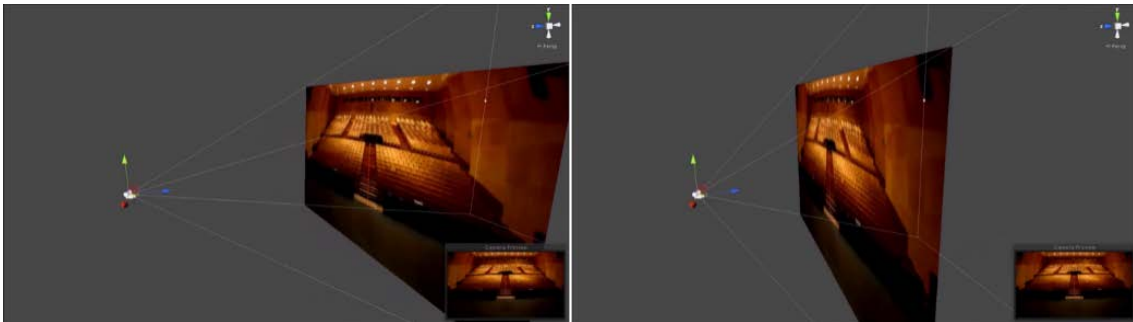


FIGURA 5.6 PLANOS DEL AUTORIO

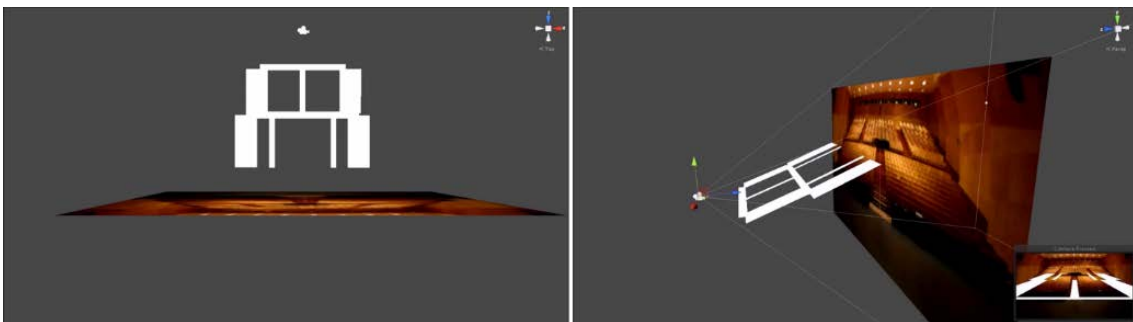




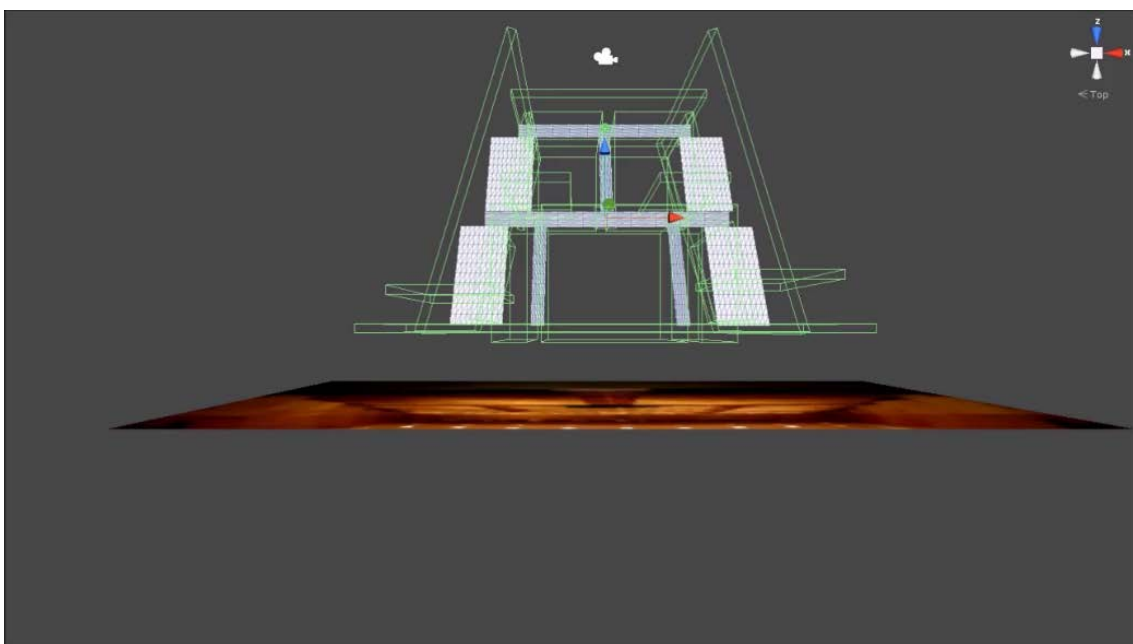
*FIGURA 5.7 IMÁGENES DEL AUTORIO A AUMENTAR*



*FIGURA 5.8 CONFIGURACIÓN DEL CAMPO DE VISIÓN DE LA CÁMARA VIRTUAL*



*FIGURA 5.9 BLOQUES DEL PASILLO MIENTRAS SE AJUSTAN AL PANEL DE FONDO*

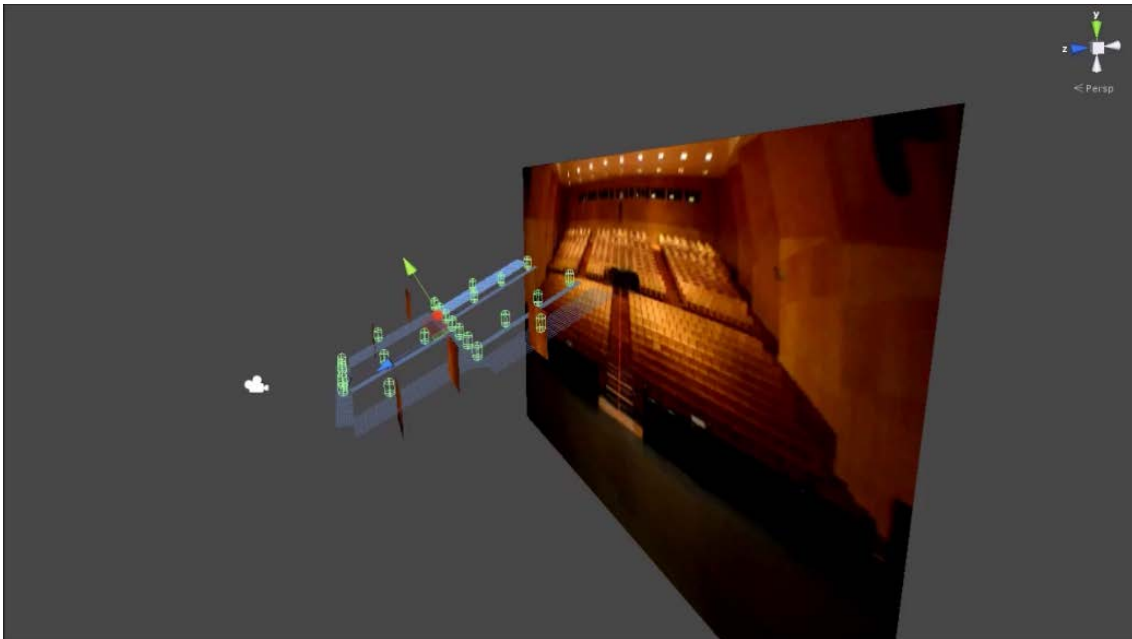


*FIGURA 5.10 RÉPLICA DEL ENTORNO CONSTRUIDA UTILIZANDO BLOQUES 3D*

### 5.3.2 Composición de la escena de RA

Siguiendo el método propuesto una vez llevado a cabo el proceso de modelado, se aplicaron sobre los bloques las texturas correspondientes en función de su capacidad para generar oclusión en la visión del usuario. De esta forma, a los bloques que representaban el suelo, las escaleras y las delimitaciones de los pasillos se les aplicó una textura transparente, mientras que a los bloques que representaban las paredes se les aplicó una textura con fragmento del escenario de fondo. En esta experiencia, no fue necesario configurar ningún bloque como semitransparente. Una vez completada la configuración del modelo este fue exportado como *prefab* y almacenado en el repositorio de escenarios de GREP Editor.

Debido a que cada juego tenía diferentes personajes y algunos elementos como las piedras o el rol de cada equipo cambiaba, se desarrolló dos escenarios de RA basados en el modelo exportado en el repositorio. Para el juego de recolección se colocaron varias monedas de virtuales por los pasillos y se añadieron los dos personajes virtuales que se eran controlados por el público. Para el segundo juego además de colocar las monedas, se cambiaron los roles de los personajes que controlaba el público y se añadió las piedras como obstáculos en los pasillos. La Figura 5.11 muestra la escena RA para el primero de los juegos.



*FIGURA 5.11 ESCENA RA UNA VEZ APLICADA LAS TEXTURAS A LOS DIFERENTES TIPOS DE BLOQUES Y AÑADIDO LOS ELEMENTOS VIRTUALES DE LA ESCENA*

### 5.3.3 Simulación de las condiciones ambientales

Una vez terminada la composición de la escena RA virtual, la siguiente fase del proceso se centró en replicar las condiciones del entorno aumentado. En este caso concreto, debido a que el entorno era un espacio cerrado con focos, la única condición ambiental a replicar fue las fuentes de luces para generar el efecto de brillo y sombras sobre los elementos virtuales. Para lograr este fin se colocaron antorchas de luces en lo alto del entorno virtual enfocando hacia abajo. La intensidad de las fuentes de luz fue regulada por un script hasta que coincidieran con los focos del público del auditorio.

Al tratarse de un espacio cerrado, no fue necesario incluir cajas de viento o modificar la dirección de la gravedad. La Figura 5.12 muestra cómo se comporta diferentes configuraciones de la intensidad de las antorchas colocadas en la escena RA desde la visión de la cámara virtual.



FIGURA 5.12 ESCENARIO RA MIENTRAS SE CONFIGURA LAS LUCES PARA REPRODUCIR LAS DEL AMBIENTE

### 5.3.4 Animación de Escena e Interacción con los Usuarios

En el último paso del proceso haciendo uso de las posibilidades que ofrece el motor de GREP se implementó en un fichero XML la lógica de comportamiento. En el primer juego se indicó que las monedas podían ser recogidas por los jugadores y que la partida acababa cuando los personajes recogían todas las monedas o se expiraba el tiempo definido. En el segundo juego, se configuró para que un personaje pudiera recoger monedas, mientras que el otro tuviera que capturar el primer personaje. Además, se configuró para que unos obstáculos en los pasillos aparecieran y desaparecieran de manera aleatoria. La partida acababa cuando el primer personaje recogía todas las monedas, cuando el segundo personaje atrapaba al primero o cuando expiraba el tiempo definido. La Figura 5.13 muestra un fragmento del fichero XML que contenía la lógica de comportamientos y las condiciones para que se acabara la partida.

Los personajes eran controlados mediante una conexión de *Websocket*. Para ello, de manera externa se desarrolló una página web que permitía recoger datos y en cada segundo notificaba el movimiento a GREP Player cuando estaba ejecutando la experiencia.

```

187
188 <!-- *** Events Treasure hunt *** -->
189 <event id="Ev_1_1_1_ObPick1" type="near" distance="3.2">
190   <entities>
191     <entity-name name="Player"/>
192     <entity-name name="Coin"/>
193   </entities>
194   <consequences>
195     <consequence type="feedback" feedback-type="script" feedback-id="gainpoints" numlayer="1"
196     <consequence type="feedback" feedback-type="audio" value="coin_sounds/coin7" />
197     <consequence type="remove" entity-name="Coin" />
198     <consequence type="recover" entity-name="Player" attribute="gennum2" value="1" />
199   </consequences>
200 </event>
201
202 <!-- *** Events Treasure hunt *** -->
203 <event id="Ev_1_1_1_ObPick2" type="near" distance="3.2">
204   <entities>
205     <entity-name name="bombero"/>
206     <entity-name name="Coin"/>
207   </entities>
208   <consequences>
209     <consequence type="feedback" feedback-type="script" feedback-id="gainpoints" numlayer="2" \
210     <consequence type="feedback" feedback-type="audio" value="coin_sounds/coin7" />
211     <consequence type="remove" entity-name="Coin" />
212     <consequence type="recover" entity-name="Player" attribute="gennum2" value="1" />

```

FIGURA 5.13 EJEMPLO DEL COMPORTAMIENTO DEFINIDO DE RECOLECCIÓN DE MONEDAS EN EL FICHERO DE XML

### 5.3.5 Resultados Obtenidos

En las Figura 5.14 y Figura 5.15 se muestra una representación del resultado final del proceso de ambas experiencias de juego. En ellas aparece un panel superior central con el tiempo de juego disponible en cada momento. Además, aparecen dos paneles superiores laterales que muestran los votos emitidos en el último segundo, así como las monedas recolectadas hasta el momento por cada equipo. La información de *log* de las conexiones indicó que se conectaron 121 dispositivos diferentes en ambas experiencias y se registró una media de 235,7 votos emitidos.



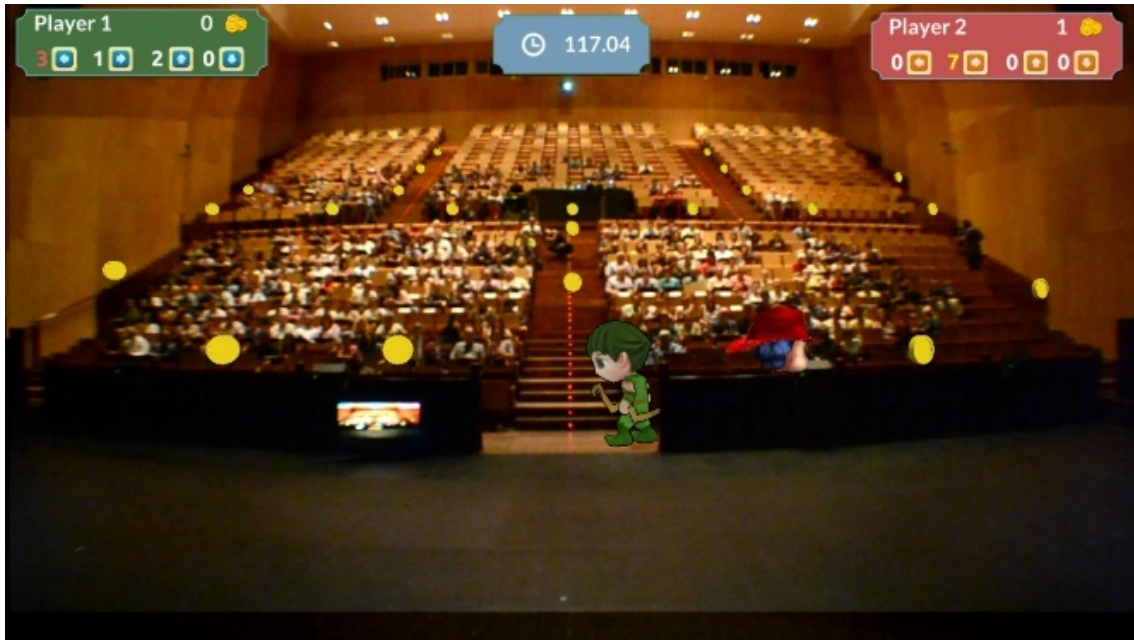


FIGURA 5.14 INSTANTÁNEA DEL JUEGO DE RA DE RECOLECCIÓN DE MONEDAS [110]

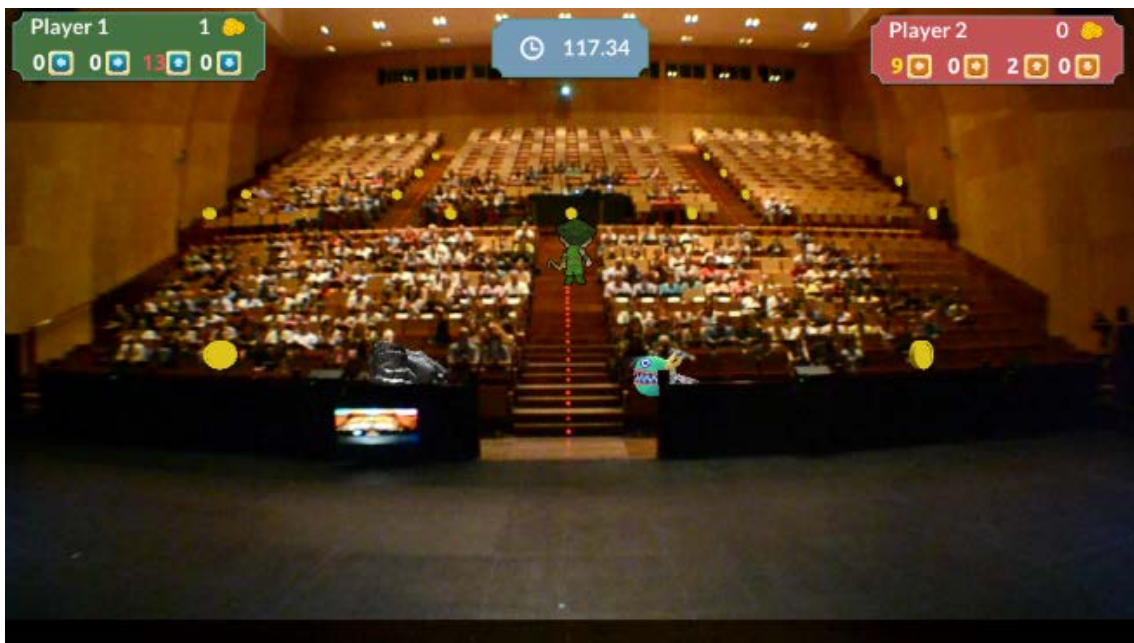


FIGURA 5.15 INSTANTÁNEA DEL JUEGO DE RA DEL PILLAPILLA [110]

#### 5.4 Propuesta de Arquitectura

Basada en el desarrollo del prototipo y los resultados de la validación, se definió una arquitectura de sistema que pudiera llevar a cabo el método definido de manera secuencial. La Figura 5.16 muestra de izquierda a derecha el flujo de información desde la cámara que captura el entorno hasta la experiencia ejecutándose. La figura representa las cuatro fases del método: modelado del entorno, composición de la escena junto con la inclusión de efecto realistas y la definición de los comportamientos.

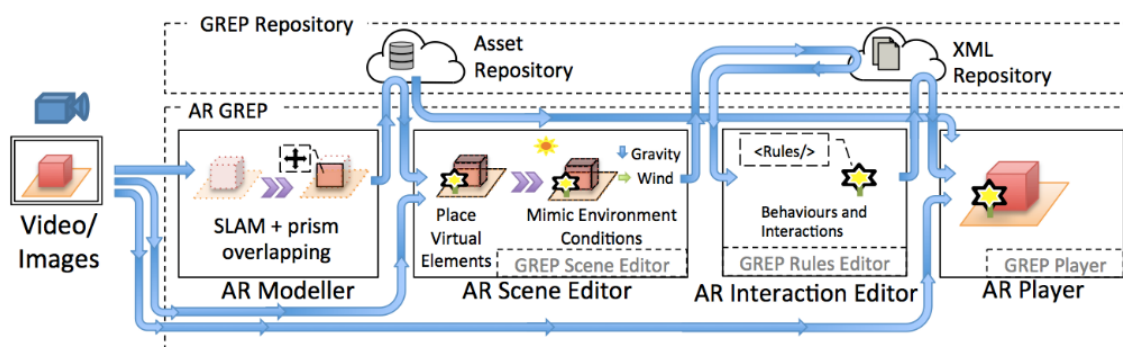


FIGURA 5.16 ARQUITECTURA DE SISTEMA AR GREP EDITOR COMPLETA [109]

## 5.5 Segundo Caso de Validación

Además del caso de validación realizado en el auditorio de la Universidad Carlos III de Madrid, se llevó a cabo un segundo caso de validación para aumentar un espacio de grandes dimensiones en el exterior. En este segundo caso se decidió aumentar un campo de fútbol con motivo de las jornadas de la ciencia de un instituto de secundaria de la Comunidad de Madrid. El juego definido se basó en la misma mecánica de competición de recolección de monedas con dos populares personajes de videojuegos. A los alumnos de secundaria se les proporcionó teléfonos móviles y se les agrupó en dos equipos.

En este caso, los alumnos sentados en sus pupitres podían ver a través de un proyector la visión que había desde una de las ventanas del aula que daba a un campo de fútbol. Sobre él, se diseñó un laberinto donde se encontraban las monedas y los dos personajes.

En este caso, el modelado del entorno consistió en la utilización de un plano basado en el tamaño aproximado del campo de fútbol. Una de las ideas para conocer el tamaño del campo fue la utilización del medidor de distancias de *Google Earth*, sin embargo, el campo de fútbol había sido construido hacía poco y no había sido capturado por los satélites. Por ello, se optó por medir el campo utilizando pasos como medida aproximada. Durante la composición de la escena se aplicó al plano una textura transparente al tratarse de un bloque que no genera oclusión. A continuación, se colocaron diferentes elementos virtuales como arbustos en forma de laberinto y monedas. A diferencia de la versión anterior, en este caso la alineación de la réplica se llevó a cabo en tiempo de ejecución de la experiencia, de tal modo que hubiera mayor flexibilidad en el momento de colocar la cámara cuando se fuera a ejecutar la experiencia. Para ello, se definieron varios puntos relevantes que en este caso fueron las cuatro esquinas del campo de fútbol. Estas posiciones tenían que ser seleccionadas sobre la pantalla táctil de móvil cuando la experiencia fuera a iniciarse. Estos puntos definidos junto con las conocidas medidas del campo permiten la correcta alineación de la réplica en el entorno real. Una vez definidos los efectos que simulaban los efectos ambientales en la escena, se implementó la misma lógica de recolección de monedas que en el caso de validación anterior. La Figura 5.17 muestra una instantánea de la ejecución de la experiencia RA.

## 5.6 Discusión

Tal y como se ha expuesto, el método de creación de experiencias de RA realistas fue validado por medio de dos casos de uso consistente en el desarrollo de dos juegos de RA, que fueron implementados sobre un escenario interior (el auditorio de la Universidad Carlos III de Madrid) y un escenario exterior (el Colegio Virgen de Europa de Boadilla del Monte), respectivamente, y probados con público real (121 y 60 jugadores simultáneos, respectivamente).

### 5.6.1 Caso Escenario Interior: Auditorio de la Universidad Carlos III de Madrid

Para lograr integrar de manera realista los elementos del juego en el escenario real, el diseño del juego incluyó una amplia variedad de efectos como perspectiva, luz y sombras,



FIGURA 5.17 CAPTURA DE PANTALLA DEL PROYECTOR DE LA EXPERIENCIA RA SOBRE UN CAMPO DE FÚTBOL

colisiones y gravedad. La Tabla 5.1 muestra un resumen de los efectos incluidos, ilustrando cada uno de ellos con varias imágenes del juego. Tal y como se puede ver en la primera columna de la tabla, los personajes principales se podían desplazar por los pasillos y escaleras, variando su tamaño dependiendo del lugar que se encontrasen respecto a la cámara que capturaba el entorno. Por otro lado, los personajes eran parcialmente ocluidos cuando estos se encontraban detrás de un obstáculo físico como las gradas o los muros de separación entre el foso para orquesta y el público. Además, los personajes y otros elementos como las monedas presentaban un nivel de brillo parecido al resto de elementos del entorno y disponían de su propia sombra. Debido a que la mayoría de las fuentes de luces del auditorio se encontraban en el techo, al replicar una configuración parecida con las fuentes de luces virtuales las sombras de los elementos virtuales se proyectaron debajo de estos. Los jugadores pudieron observar además cómo los personajes del juego colisionaban con las paredes del auditorio y los límites de los pasillos. De igual forma, también pudieron ver cómo los personajes sufrían los efectos de una fuerza de la gravedad simulada al subir y bajar las escaleras del auditorio, mientras que las monedas a recoger no la experimentaban y aparecían flotando y girando en el aire. Para esta experiencia concreta no fue necesario incluir efectos de viento, ya que tenía lugar en un espacio interior, o sonido, ya que la propia acción del juego era acompañada de una melodía que dificultaría su escucha. De cualquier manera, no hubiera sido complicado introducir estos efectos si las características de la experiencia las hubieran requerido.

En general, la experiencia demostró que fue posible aumentar un espacio grande interior mediante el método propuesto, integrando un gran número de efectos y obteniendo un resultado final altamente realista. La puesta en práctica del método permitió comprobar la facilidad para añadir y quitar efectos sobre una escena de RA realista, permitiendo que los diseñadores prueben sus efectos individualmente o por separado hasta conseguir la combinación adecuada para cada caso. Otra posibilidad sería activar los efectos en función de valores obtenidos a partir de sensores situados sobre el entorno a aumentar, de manera que se aplicasen sobre la escena como respuesta a ciertas condiciones lumínicas, climatológicas, etc.

La experiencia también permitió identificar algunas limitaciones del método propuesto. Así, al igual que como otros sistemas de RA video *see-through*, el realismo de la experiencia puede verse comprometido por un retraso en el vídeo capturado con respecto al resultado final. Para el caso específico de la experiencia probada, hubo un retraso escasamente perceptible por la



TABLA 5.1 EFECTOS REALISTAS VISIBLES EN LA EXPERIENCIA DE RA [110]

Efectos					
Perspectiva	Oclusión	Brillo	Sombra	Colisión	Gravedad
 Cerca	 No ocluido	 Iluminado			 Afectado
 Lejos	 Ocluido	 No iluminado			 No afectado

conversión de la señal de vídeo. La cámara encargada de capturar las gradas proporcionaba una señal de vídeo mediante la interfaz HDMI que tuvo que ser conectado a un convertidor que transformó la señal a USB 2.0 mediante el códec H.264 [128]. Este problema puede ser resuelto fácilmente utilizando un dispositivo de captura que transmita el vídeo con un códec más liviano y haciendo uso de un ancho de banda mayor como un USB 3.0 o directamente capturando el vídeo con una tarjeta conectada directamente a la placa base del ordenador. Otra de las limitaciones que se observó durante el desarrollo de la experiencia fue la necesidad de mantener fija la posición de la cámara que captura la realidad debe estar fija. Cualquier pequeño desplazamiento o rotación podía romper la alineación del modelo 3D con la imagen capturada del entorno provocando que los objetos aparecen proyectados sobre la imagen en posiciones incorrectas. Para evitar este problema y permitir cierto desplazamiento de la posición y orientación de la cámara sería necesario contar con sensores que capturasen dichas variaciones y un proceso que reajustase la posición del modelo 3D del entorno en el mundo virtual a partir de esa información.

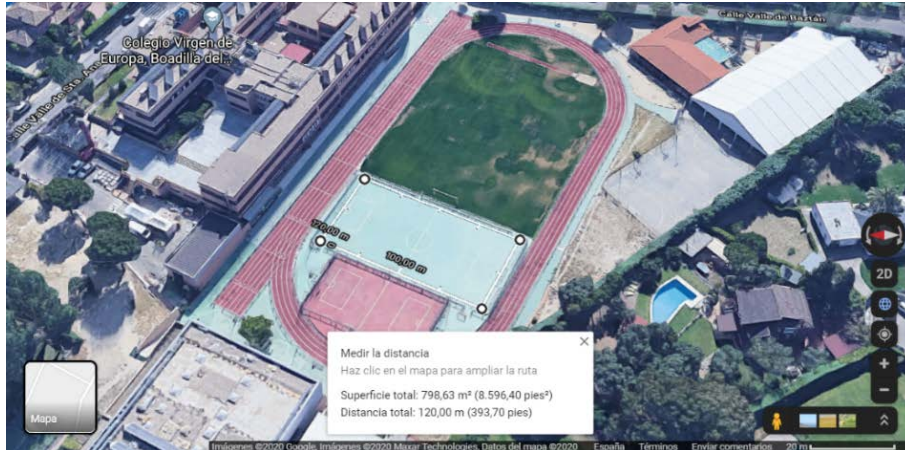
### 5.6.2 Caso Escenario Exterior: Centro de Secundaria Colegio Virgen de Europa

La segunda experiencia llevada a cabo en un centro de secundaria validó el método propuesto en un caso real en el que un espacio exterior de grandes dimensiones era aumentado. Una característica importante que destacar en esta validación fue la mayor distancia existente entre la cámara real y el entorno a aumentar. Esto se traduce en que cualquier pequeño movimiento o rotación de la cámara provocaba que los elementos virtuales estuviesen totalmente desalineados con la realidad. Para hacer frente a este problema se flexibilizó el proceso de alineación de la réplica del entorno desde la segunda fase del método propuesto al momento de la ejecución de la experiencia. De tal modo, que la ubicación entre ejecuciones de esta experiencia pudiera ser movida y no tuviera que colocarse en el mismo punto exacto. En este caso concreto, se utilizó como referencia del mundo real los córneres del campo de fútbol, por lo que antes de empezar a ejecutar la experiencia tuvieron que ser seleccionados en la pantalla táctil del móvil que apuntaba al campo.

El resultado final de la experiencia fue satisfactorio, apareciendo correctamente alineado el laberinto virtual en la mitad del campo de fútbol con el tamaño previsto. Esta adaptación del método unida a la elección de las referencias en la pantalla táctil fue adecuada para una correcta alineación. Aunque en la actualidad existen algoritmos de procesamiento de imagen que puede llevar a cabo el reconocimiento de estas referencias de manera automática, la gran distancia a la que se encontraban junto con el ruido del ISO, la luz o la resolución de la cámara hacen que su aplicación en este escenario concreto no sea viable. Por ello, pese a requerir de una mayor participación del usuario se puede lograr una mayor precisión. De cualquier manera, pese a que este proceso de identificación se realice de manera manual, es posible combinarlo con

algoritmos de odometría que permitan incluso pequeños desplazamientos y rotaciones de la cámara durante la ejecución de la experiencia.

En relación con el proceso de modelado de exteriores, como se ha comentado durante la elaboración de la experiencia, puede facilitarse por el empleo de la plataforma Google Maps para medir distancias entre dos o más puntos. En el caso del ejemplo, no se pudo utilizar dado que el campo de fútbol era relativamente nuevo y no aparecía en las últimas imágenes capturadas por el satélite de Google. Aun así, la Figura 5.18 muestra cómo se podría realizar el proceso actualmente para conocer el tamaño del campo de fútbol y por ende del plano de la réplica del entorno real.



*FIGURA 5.18 CAPTURA DE PANTALLA DE LA HERRAMIENTA DE MEDIR DISTANCIAS Y SUPERFICIES INTEGRADA EN GOOGLE MAPS*

Aunque como se ha comentado la experiencia fue satisfactoria, consiguiéndose aumentar de forma realista un escenario exterior, el diseño y desarrollo de este caso permitió identificar algunas limitaciones. Así, al final de la experiencia se pudo observar algo de lluvia, un efecto que no se encontraba implementado pero que podría incluirse con sistemas de partículas o filtros que recreen el vaho de los cristales de las ventanas. Además de ello, la maya metálica presentó una forma compleja de considerar para la oclusión. Para este caso se tendría que considerar la inclusión filtros especiales que permitan una selección inteligente por colores o proximidad de píxeles (como los de algunos editores de imágenes) para aplicar efectos de oclusión.

Por último, una cuestión a considerar tras la ejecución de la experiencia es que debido a las grandes distancias y pequeños tamaños de los elementos virtuales algunos efectos realistas como las sombras podrían no percibirse, por lo que, debería de estudiarse la relación de tamaño y distancia a la que se debe considerar obligatorio este efecto concreto.

## 6 AUTORÍA DE EXPERIENCIAS DE REALIDAD AUMENTADA BASADAS EN REALIDAD VIRTUAL

Las herramientas de autoría de RA tienen como objetivo permitir que usuarios tengan la capacidad de registrar por sí mismo contenido virtual en el mundo real de manera sencilla e intuitiva [14]. Para abordar esta problemática se han identificado una serie de retos para diferentes dominios de aplicación [156] y se han explorado diferentes soluciones o enfoques centradas en hacer frente a uno o varios de estos desafíos [59,77,91].

### 6.1 Motivación

Actualmente, la mayoría de las aplicaciones de autoría de RA siguen un enfoque WYSIWYG ("What You See Is What You Get", en español "lo que ves es lo que obtienes"), permitiendo que el usuario añada el contenido digital superponiéndolo sobre la propia visión del entorno obtenida a través de la pantalla del mismo dispositivo de RA similar al que se utilizará en la experiencia. La principal ventaja de este enfoque es que permite al autor observar en tiempo de creación cómo la experiencia de RA será percibida por el usuario final. Además, la mayoría de los sistemas de autoría de RA de este tipo utilizan dispositivos móviles como tabletas o *smartphones* como plataforma de la propia experiencia de RA, por lo que a los autores les resulta natural interactuar con la aplicación manipulando imágenes sobre la pantalla, de igual manera que hacen habitualmente al utilizar sus móviles. Sin embargo, este enfoque también presenta ciertos inconvenientes, como la necesidad de encontrarse físicamente en el espacio a aumentar, el limitado espacio de trabajo que proporcionan las pantallas de los dispositivos móviles (o en su defecto, el limitado campo de visión de las actuales gafas de RA) o la dificultad de situar objetos virtuales a grandes distancias del observador.

En este capítulo se propone un enfoque de autoría de experiencias de RA que busca explotar las características y propiedades de la tecnología de RV para proporcionar a los usuarios un mecanismo sencillo, intuitivo y eficiente de creación de experiencias de RA. Entre estas propiedades destaca la capacidad de los dispositivos de RV para permitir al usuario acceder a recreaciones 3D del entorno a aumentar posibilitando que puedan explorarlas y trabajar en ellas añadiendo nuevo contenido digital, aunque no se encuentren físicamente situados en el propio entorno a aumentar. Por otra parte, la interacción que este tipo de tecnología soporta se asemeja más a la forma en que normalmente manipulamos y modificamos de nuestro entorno. Así, en contraste a la interacción táctil proporcionada por las pantallas utilizados en autoría mediante RA basada en dispositivos móviles, los usuarios de tecnología RV inmersiva realizan gestos para coger o modificar los objetos virtuales a su alcance, bien empleando sus propias manos o mediante dispositivos como los Oculus Touch, por lo que el uso de RV como medio para el proceso de autoría les puede resultar más intuitivo y fácil de usar. Además, las pantallas de los dispositivos de RV ofrecen un amplio campo de visión del entorno a aumentar, y los usuarios pueden navegar y explorar estas réplicas virtuales fácilmente, salvando obstáculos que en el entorno físico podrían resultar difíciles de sobrepasar. Por último, la sensación de presencia proporcionado por este tipo de tecnología suele asociarse a un alto grado de *engagement* [105,136], lo que a su vez puede contribuir a que el usuario se encuentre más centrado en la tarea de creación y encuentre la experiencia más placentera.

### 6.2 Objetivo y Preguntas de Investigación

El trabajo que se presenta en este capítulo tiene como objetivo explorar los **beneficios y limitaciones que proporciona la Realidad Virtual inmersiva durante parte del proceso de autoría de experiencia de Realidad Aumentada**, donde el término inmersivo hace referencia a la tecnología que el usuario hará uso y no a su percepción de presencia [142]. De manera más concreta, esta investigación busca obtener respuesta a las siguientes preguntas:

*PR1. ¿Llevar a cabo las tareas básicas asociadas al proceso de autoría de realidad aumentada (crear, mover, rotar y redimensionar los elementos virtuales) a través de la realidad*

*virtual inmersiva requiere un esfuerzo (mental y físico) similar al realizado al utilizar una herramienta de RA basada en dispositivos móviles?*

*PR2. ¿La satisfacción de los usuarios durante el proceso de autoría cuando usan la realidad virtual inmersiva es similar a cuando utilizan una herramienta móvil de RA para aumentar el entorno?*

*PR3. ¿El tiempo requerido para completar la tarea de autoría en un entorno de RV inmersiva es menor que el necesario para completar dicha tarea mediante RA basada en móviles?*

La pregunta de investigación PR1 surge de la posibilidad de que el trabajo sobre una réplica virtual del entorno implique una mayor carga cognitiva al usuario. Esto puede producirse debido a que las formas del espacio virtual no resulten tan inmediatamente reconocibles como las del entorno real a aumentar, por ejemplo. Además, el desplazamiento a través del entorno virtual y la interacción con sus elementos a través de gestos puede suponer un mayor esfuerzo físico que el realizado cuando la tarea se efectúa mediante el teléfono móvil. La segunda pregunta de investigación PR2 se centra en la propia experiencia subjetiva que el usuario tiene al realizar la tarea en el entorno virtual que independientemente de conllevar o no un esfuerzo mayor puede resultar en menor o mayor medida satisfactorio y placentera, por ejemplo, debido a la posible incomodidad de ser realizada con un casco que aísla al usuario del entorno. Por último, la tercera pregunta de investigación PR3, trata de esclarecer si la mayor libertad de movimiento en RV se traduce en una mayor velocidad en la realización de tareas de autoría.

### **6.3 Hipótesis**

Para dar respuesta a las preguntas planteadas se hace necesario definir un conjunto de hipótesis que deben ser refutadas o rechazadas a partir de datos extraídos en la investigación.

*H1. La cantidad de esfuerzo que el usuario tiene que ejercer cuando hace uso de la técnica propuesta de autoría con RV inmersiva no es superior a la requerida cuando usa una técnica de autoría basada en RA.*

*H2. El grado de satisfacción de un usuario cuando usa la técnica propuesta de autoría con RV inmersiva no es inferior a cuando se hace uso de una técnica de autoría basada en RA.*

*H3. La eficiencia en términos de tiempo de un usuario cuando usa la técnica propuesta de autoría con RV inmersiva no es inferior a cuando usa una técnica de autoría basada en RA.*

### **6.4 Herramientas de Autoría de Experiencias de RA Desarrollados para la Investigación**

Para poder llevar a cabo esta investigación se desarrollaron dos herramientas de autoría de escenarios RA: uno empleado tecnología de RV inmersiva (SimuAR) y otro un enfoque de RA basado en móviles (InSituAR).

#### **6.4.1 Herramienta de autoría de RA mediante tecnología realidad virtual (SimuAR)**

La principal característica de esta herramienta es que el proceso de autoría de la escena RA se desarrolla empleado una réplica virtual 3D del espacio real a aumentar a la que se accede y se manipula por medio de la tecnología RV inmersiva, y más concretamente un HMD y dos controles de RV (Oculus Touch). El proceso de autoría propuesto está dividido en tres pasos:

- **Modelado del Entorno:** El objetivo de esta fase es crear una réplica tridimensional digital del entorno a aumentar. Como resultado de esta etapa se generará un fichero de modelo 3D.
- **Creación y manipulación de los elementos virtuales:** En esta fase el autor accede a la réplica y posiciona los elementos virtuales con los que aumentar el entorno que son configurados con el tamaño y orientación de desea. Como resultado se obtiene la información de

distribución de los elementos virtuales, así como sus propiedades, dentro de la réplica virtual.

Además, antes de desplegar el diseño de la escena sobre el espacio a aumentar y poder disfrutar de la experiencia RA será necesario efectuar un tercer paso:

- Alineación de la réplica con el espacio real: Para que los objetos virtuales aparezcan correctamente situados sobre el entorno real a aumentar el usuario deberá previamente a la ejecución de la experiencia alinear la visión que tiene del entorno con la réplica virtual del mismo utilizada durante el proceso de autoría.

#### **6.4.1.1 Modelado del entorno**

El proceso de modelado de entorno tiene como objetivo crear una representación digital de un espacio físico. Para esta tarea se seleccionaron y estudiaron las ventajas e inconvenientes de cuatro técnicas distintas mediante las que se puede abordar este proceso:

- Modelado manual del entorno: Una de las opciones para construir el modelo 3D es hacer uso de una herramienta de edición 3D como Maya, Blender o AutoCAD. Estas herramientas permiten a usuarios expertos crear manualmente y desde cero modelos 3D digitales de entornos u objetos para su uso en videojuegos, películas de animación o procesos industriales. Para poder llevar a cabo esta tarea es necesario disponer de información precisa sobre la forma y tamaño del objeto a modelar. En el caso del modelado del entorno físico esto implica conocer la distancia exacta entre los principales puntos de este, una información que bien se puede obtener a partir de planos, en el caso de que el entorno se corresponda con un edificio, o bien puede requerir realizar mediciones en el propio lugar a aumentar, por ejemplo, utilizando dispositivos láser. Es importante destacar que la correcta elaboración del modelo 3D depende en gran parte de la correcta extracción e interpretación de las mediciones. Una vez generado la maya del modelo 3D el siguiente paso consiste en realizar un proceso de texturización del modelo para aplicar colores al volumen creado. Debido a los requisitos de tiempo y conocimientos que requiere esta técnica de modelado, esta opción fue descartada.
- Modelado del entorno mediante prismas: Otra posibilidad es llevar a cabo el proceso de modelado directamente sobre un entorno virtual, utilizando una imagen de referencia del entorno a aumentar y utilizando un conjunto de bloques o prismas y scripts predefinidos. Esta técnica ha sido aplicada al modelado de espacios grandes con bastante éxito como en el capítulo anterior [110], aunque cuenta con la desventaja de requerir que la cámara mantenga siempre la misma perspectiva, lo cual impide que el autor pueda moverse dentro de la réplica para obtener diferentes puntos de vista para la siguiente fase del proceso. Por este motivo esta opción fue descartada.
- Modelado del entorno empleando un dispositivo móvil con cámara estéreo y sensores de unidades de medición inercial: La tercera técnica estudiada se basa en el empleo de un dispositivo móvil equipado con varias cámaras. Los actuales móviles tienen la suficiente potencia para capturar el entorno desde sus cámaras y procesar las imágenes obtenidas empleando algoritmos de odometría que permiten reconstruir el espacio capturado mediante un modelo 3D. Esta técnica de modelado fue evaluada un dispositivo móvil equipado con tecnología Google Tango que cuenta con dos cámaras traseras, una cámara normal y un sensor de profundidad, y la aplicación de código abierto Open Constructor [211], a su vez basada en la plataforma Google Constructor Developer Tool [212]. Los resultados obtenidos a través de diversas pruebas indicaron que el sistema funciona mejor para el modelado de salas pequeñas que para espacios grandes. A pesar de hacer uso de un algoritmo de odometría basado en sensores visuales e iniciales con corrección *loop closure* en la aplicación, un largo periodo usando la aplicación causaba una desviación de posición y rotación posiblemente a imágenes desenfocadas, caras demasiado oscuras y superficies reflectantes. Esto causaba que el modelo 3D tuviera varias paredes, y que algunas curvas



suaves aparecieran entre las paredes con el techo y el suelo. Además, la herramienta de código abierto no podía terminar de procesar correctamente espacios grandes y el modelo no era texturizado. Respecto al uso de la herramienta en espacios abiertos, una luz extremadamente alta o baja hacía imposible llevar el proceso como puede ser un día soleado o por la noche. La mejor condición para llevar el modelado en exteriores son los días nublados, aunque esto no garantiza una calidad suficientemente buena. La falta de textura de UV hacía que la réplica del entorno generada fuese a veces difícil de identificar como el espacio a aumentar, por lo que finalmente esta técnica de modelado se descartó.

- **Modelado por fotogrametría:** La fotogrametría es una técnica que permite reconstruir tanto objetos como entornos (incorporando tanto la forma como la textura) haciendo uso de fotografías o fragmentos de vídeo. Las imágenes son procesadas en un flujo de ejecución que hace uso de diferentes algoritmos para determinar la relación entre las imágenes y calcular el volumen de los elementos que se muestran. Por lo general, este proceso tiene a consumir bastante tiempo, sin embargo, tiene la ventaja de que no tiene que ser supervisado por el usuario. Este tiempo de ejecución depende del número de imágenes, su resolución y la potencia de cómputo del equipo. Después de probar distintas configuraciones se obtuvieron los mejores resultados para esta técnica empleado una cámara réflex con los parámetros mostrados en la Tabla 6.1. Las pruebas realizadas indicaron que un mayor número de imágenes no implica necesariamente una mayor calidad, pero que es necesario que las utilizadas no estén movidas, quemadas o incorporen mucho ruido, y que es recomendable que el entorno no presente muchas superficies reflectantes porque tienden a falsear referencias para el modelo final. Para modelar el entorno en el que finalmente se llevó a cabo el experimento que sirvió de base a la investigación, una sala de reuniones de aproximadamente 37 m<sup>2</sup>, fueron capturadas 104 imágenes. El mobiliario de la habitación consistió en 3 mesas, 5 librerías vacías y un par de sillas colocadas en las mesas. Todas estas imágenes fueron procesadas posteriormente por una herramienta comercial llamada Agisoft Metashape haciendo uso de un proceso por lote predefinido [213]. Al final del proceso se obtuvo una réplica 3D de la sala bastante realista. El único defecto destacable fue una pizarra blanca magnética que no fue modelada debido a la reflexión de la superficie. Esta superficie fue arreglada usando un simple plano de Unity.

*TABLA 6.1 RESUMEN DE LOS PARÁMETROS DE CONFIGURACIÓN DE LA CÁMARA.*

Propiedad	Valor	Nota
Sensibilidad ISO	100 (o más bajo)	Este valor configura la sensibilidad del sensor que captura la imagen. Un valor bajo permite reducir el nivel de ruido de la imagen, pero hace que las imágenes aparezcan más oscuras.
Flash	No flash	El flash permite aumentar la luz del entorno, pero incrementa el efecto de reflexión de la luz sobre las superficies reflectantes.
Obturador de la Camera	f/8 (o más bajo)	Este valor indica el tiempo que el sensor de la cámara estará expuesto a la luz. Esta configuración permite capturar más luz para contrarrestar un bajo ISO.
Trípode	Altamente recomendado	Poder mantener la cámara lo más fija posible permite contrarrestar el efecto borroso debido a una configuración baja de obturador.
Distancia Focal	Manual 18mm (tan bajo como sea posible)	Este valor define el campo de visión capturado en cada fotografía. Es altamente recomendado mantener el mismo valor en todas las imágenes que se vayan a utilizar en el proceso. Un valor bajo permite capturar más entorno, pero es posible que un valor demasiado bajo produzca deformaciones en la imagen como una lente ojo de pez

#### **6.4.1.2 Creación y manipulación de los elementos virtuales**

Una vez se dispone de una réplica virtual del entorno a aumentar se puede comentar el proceso de autoría de la escena RA. Tal y como se ha explicado anteriormente, empleado un HMD y unos controles de RV como las Oculus Touch, el usuario situará en la réplica los objetivos virtuales con los que aumentar el entorno con la posición, tamaño y orientación deseados. El diseño de la aplicación para llevar a cabo esta tarea se elaboró buscando proporcionar una

manipulación intuitiva y directa basada en gestos con las manos. Para una mejor comprensión en detalle del proceso de autoría, esta sección describe las operaciones básicas necesarias para aumentar una sala de reuniones colocando una butaca virtual y modificando sus propiedades por medio de la aplicación desarrollada (SimulAR).

En primer lugar, una vez que el usuario se coloca el equipo de RV puede acceder a la réplica de la sala y observar una representación de sus manos agarrando unos mandos virtuales. El usuario se podrá desplazarse a través del entorno virtual bien dando pasos con sus propios pies, en caso de que el desplazamiento a realizar sea corto, o bien utilizando el joystick del mando izquierdo. Ambas técnicas de navegación en el entorno virtuales fueron seleccionadas e implementadas en la herramienta de autoría debido a su fácil utilización, que no requiere de un entrenamiento previo del usuario. Para poder añadir elementos virtuales a la escena, el usuario debe desplegar un “Menú de Creación de Elementos Virtuales” mediante el botón físico del mando izquierdo. Tal y como se muestra en la Figura 6.1 este menú permite navegar a través de los objetos virtuales incluidos en el repositorio de la herramienta y seleccionar aquel que se desee añadir a la escena. Para ello, el usuario tiene que apuntar con el puntero láser virtual al icono 2D y, manteniendo el pulsado el botón de operación (que se corresponde con el botón A, B o el gatillo del mando derecho) arrastrar el icono hasta la posición de la réplica donde desea que aparezca. Finalmente, el usuario suelta el botón de operaciones para acabar el proceso. Cuando el elemento 2D seleccionable sale de los bordes del menú, este se convierte en un modelo 3D. En el caso concreto de la butaca, el autor tendría que buscar el elemento en el menú y arrastrarlo sobre una superficie de la réplica virtual de la sala (Figura 6.1 derecha). Si tras crear el elemento virtual el usuario desea cambiar su posición será necesario que lo señale con el láser y pulse un botón de operación. En ese momento el elemento virtual empezará a brillar para indicar que se encuentra seleccionado y aparecerá un menú de acciones entre el usuario y el elemento virtual. El usuario pulsa el botón “Edit” y el menú de acción pasa a ser de edición como el que se muestra en la Figura 6.1 izquierda. En este menú, el usuario pulsa el botón “Move” y, a continuación, pulsado y manteniendo pulsado el botón de operación el usuario puede arrastrar el elemento virtual desde sitio donde se encuentra al lugar deseado. Una vez que se suelta el botón de operación, el movimiento termina y vuelve a aparecer el menú de edición. En el caso de que se quiera rotar el elemento virtual, el autor pulsa sobre el botón “Rotate”. Entonces el usuario puede iniciar el proceso y manteniendo pulsado un botón de operaciones y rotando el mando derecho sobre el eje Y. Si el usuario rota el mando hacia un lado, el elemento virtual empieza a rotar sobre su eje perpendicular a la superficie imitando el giro. Si el usuario vuelve a colocar el mando en la posición inicial, el elemento virtual parará de girar. Si el mando es girado en el otro sentido, entonces el elemento virtual empieza a girar en el sentido contrario. Por último, si el botón de operaciones es soltado, la acción de rotación termina y vuelve a aparecer el menú de edición. Si el usuario pulsa sobre el botón “Scale”, la acción de redimensionamiento puede realizarse. Para ello, el usuario necesita pulsar sobre los botones de operaciones de ambos mandos (botón X, Y o gatillo en el mando izquierdo) y empezar a separar o juntar los mandos. Si los mandos son separados, el elemento virtual empieza a crecer. Si la distancia vuelve a ser igual a la inicial, con un margen de error de un  $\pm 10\%$ , el elemento virtual para de cambiar su tamaño.

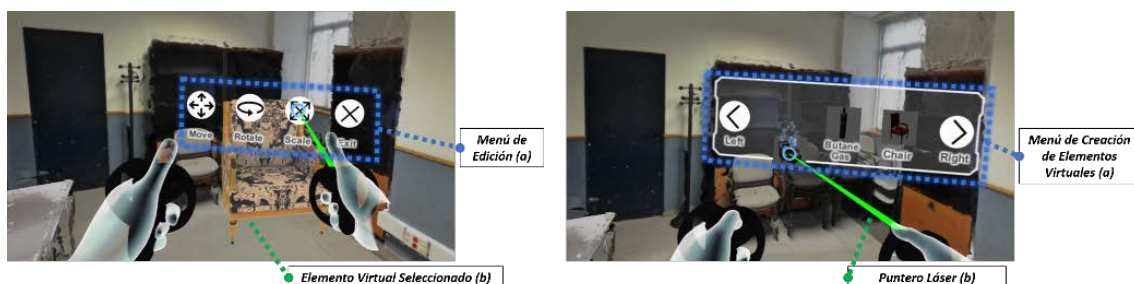


FIGURA 6.1 CAPTURA DE HMD MOSTRANDO LA EDICIÓN DE UN ELEMENTO VIRTUAL

Por otro lado, si los mandos se arriman, el elemento virtual empieza a ser cada vez más pequeño. Por último, si el usuario suelta los botones de acciones de los mandos, la operación finaliza y vuelve a aparecer el menú de edición. Al contrario que las anteriores operaciones, esta operación ha sido limitada con un tamaño máximo y mínimo que no se puede exceder. Por otro lado, dado que la rotación de la muñeca es limitada y la máxima separación de los mandos depende de la altura y anatomía del usuario, estos gestos fueron diseñados para indicar el inicio, fin y dirección de las modificaciones. Para lograr mayor precisión en las modificaciones se incorporó una aceleración y velocidad máxima de modificación. Finalmente, si se desea retirar el elemento virtual de la escena, el usuario puede hacerlo pulsando sobre el elemento virtual y seleccionado el botón "Remove" del menú de acciones. La herramienta también fue provista de la posibilidad de definir lógicas de comportamiento. La Figura 6.2 muestra un resumen de los botones de los mandos y la Tabla 6.2 ilustra todos los gestos que se hacen uso en la herramienta e información adicional.

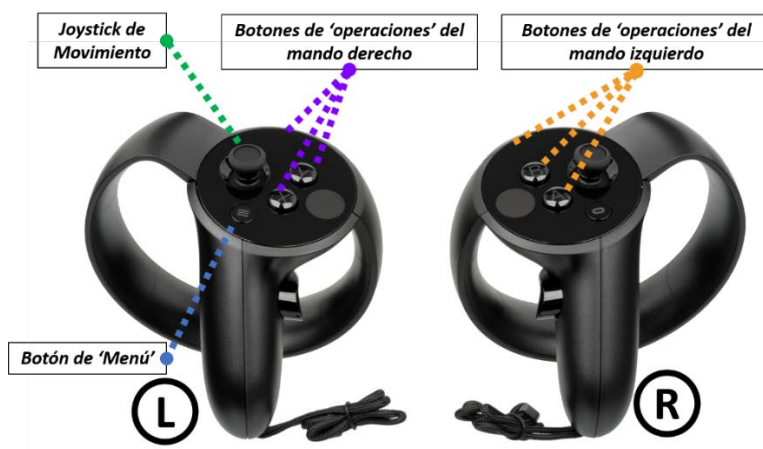


FIGURA 6.2 RESUMEN DE LAS FUNCIONES DE LOS BOTONES DE LOS MANDOS

TABLA 6.2 RESUMEN DE LOS GESTOS PARA LA EDICIÓN DE LOS ELEMENTOS VIRTUALES

Edición	Gesto	Nota
Desplazamiento del usuario		Adicionalmente, el autor puede moverse por sí solo andando para desplazarse en distancias cortas.
Mover un elemento virtual		El punto de colisión del láser con una superficie determina la ubicación.
Rotar un elemento virtual		La rotación del mando derecho indica el inicio, parada y sentido de la rotación del elemento virtual.
Redimensionar un elemento virtual		La distancia entre los mandos determina el inicio, parada y crecimiento/reducción del tamaño del elemento virtual.

### 6.4.1.3 Alineación de la réplica con el espacio real

Una vez que los elementos virtuales han sido colocados y editados dentro de la réplica del entorno virtual, es necesario definir la correspondencia con la realidad para que los elementos virtuales registrados en la réplica puedan estar correctamente registrados en la realidad. Este paso, a diferencia de los anteriores, depende de la tecnología empleada para ejecutar la experiencia de RA y, por tanto, depende de la técnica de tracking que vaya a ser utilizada por el sistema de ejecución de experiencia de RA. En el caso que se presenta para el estudio, y dado que se hizo uso de un dispositivo móvil con tecnología Google Tango, la réplica y el entorno real fueron alineados haciendo uso de dos de las tres capacidades principales del núcleo de Google Tango: el *motion tracking* y el *area learning*. *Tango motion tracking* permite al dispositivo determinar por sí solo la trayectoria y rotación desde que esta funcionalidad fue activada. Por tanto, cuando dicha función es iniciada el dispositivo aplica como posición inicial las coordenadas cartesianas de origen ( $x: 0, y: 0, z: 0$ ) y como rotación el cuaternión identidad ( $w: 1, x: 0, y: 0, z: 0$ ). Esta capacidad por sí sola no es suficiente para llevar el registro de la réplica, es necesario combinarla con el *area learning* que permite al dispositivo, aprender cómo es un entorno, para luego, poder reconocerlo. Para lograr esto, esta funcionalidad necesita extraer características claves del entorno que son guardadas en un fichero de formato privativo denominado *Area Description File* (ADF). Este fichero contiene características extraídas del entorno y su relación con la posición y rotación del *motion tracking*. De este modo el dispositivo puede reconocer su posición y rotación cuando captura una parte del entorno anteriormente capturada. Una vez explicado el funcionamiento, en este paso es necesario alinear el sistema de coordenadas de la réplica con el sistema de coordenadas arbitrario del *motion tracking* de Tango.

Para lograr la correspondencia, se situó el móvil Tango en un punto de referencia fácilmente reproducible en posición y rotación y se iniciaron las funcionalidades *motion tracking* y *area learning* al mismo tiempo. Una vez hecho esto, se movió el dispositivo para que aprendiera sobre el entorno con el objetivo de generar un fichero de ADF. A continuación, se modificó en SimulAR la posición y rotación de la réplica, haciéndola coincidir con el punto de referencia elegido. Para aumentar la precisión del proceso, se realizó un boceto del modelado de baja precisión usando el dispositivo móvil e iniciando el proceso en el mismo punto de referencia. Posteriormente, se introdujo este boceto en el SimulAR y con la ayuda de Unity Editor se alineó con la réplica 3D.

## 6.4.2 Herramienta de autoría RA mediante RA basada en móviles (InSituAR)

La herramienta de autoría RA mediante RA para móviles también utilizó la tecnología de Google Tango debido a que sigue siendo considerada una de las tecnologías más avanzadas para RA respecto a dispositivos móviles en cuanto estabilidad y precisión [32]. Esta tecnología se basa en hacer uso de *tracking markesless*, lo que evita tener que modificar el entorno con elementos adicionales que sirvan de referencia, como *fiducial markers*. Sin embargo, sigue siendo necesario realizar un proceso previo de “aprendizaje del entorno” que dota al sistema de la capacidad de reconocer el espacio a aumentar. Una vez se ha completado este proceso, los usuarios pueden aumentar el entorno realizando gestos desde el dispositivo móvil. Los gestos de los dedos fueron diseñados para tener cierta similitud con los gestos usados en la herramienta de RV.

### 6.4.2.1 Aprendizaje del Entorno

Cuando se abre la herramienta por primera vez o se aumenta un nuevo entorno, es necesario extraer los puntos característicos del mismo, es decir, la aplicación debe aprender del entorno (*area learning*). Para ello, el autor necesita seleccionar en el menú principal de la aplicación que va a aumentar un nuevo entorno y moverse por él llevando consigo el móvil y apuntado a las texturas y formas que sean especialmente significativas o reconocibles y puedan servir de referencia. Durante este proceso, la aplicación muestra un panel minimalista que indica que el autor debe moverse por el entorno para poder extraer puntos característicos del espacio. Al final del proceso, los puntos obtenidos y la posición relativa del dispositivo desde el inicio del

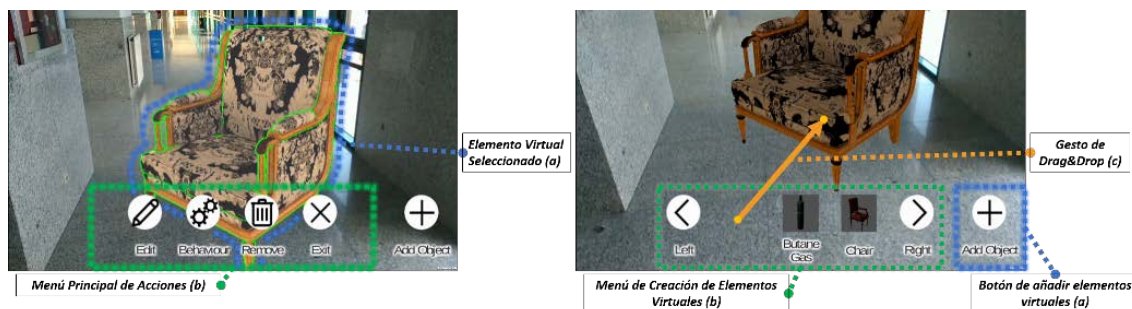


FIGURA 6.3 CAPTURA DE PANTALLAS DEL DISPOSITIVO MÓVIL





proceso es almacenada en el fichero ADF. En un ejemplo en el que un pasillo fuera a ser aumentado, el usuario tendría que pasear con el dispositivo en la mano apuntando a los diferentes elementos del entorno.

#### 6.4.2.2 Creación y Manipulación de Elementos Virtuales

Una vez completado el proceso de *area learning*, la aplicación está preparada para añadir elementos virtuales al entorno. Con el objetivo de proporcionar en detalle el funcionamiento de esta herramienta y la interacción necesaria, en las siguientes subsecciones se detalla el proceso a seguir colocando una butaca virtual y modificando sus propiedades como rotación y tamaño.

Para añadir un elemento virtual es necesario que el usuario pulse sobre el botón 'Add Object' situado en la parte inferior derecha de la pantalla (Figura 6.3). Entonces, un menú horizontal aparece en la parte inferior de la pantalla mostrando 3 elementos seleccionables y 2 botones de navegación. Para añadir un elemento virtual, el usuario arrastra el elemento del menú hasta una superficie como el suelo, la pared o el techo. Cuando el elemento 2D arrastrado sale del menú este se convierte en un elemento 3D como se muestra en la Figura 6.3 derecha. En el caso del ejemplo, el usuario tendría que buscar el icono de butaca y arrastrarlo el icono hasta una superficie horizontal. Para manipular las propiedades del elemento, el usuario selecciona el elemento virtual pulsando sobre él con el dedo. Entonces, este empezará a brillar para indicar que ha sido seleccionado y un menú de acción aparece en la parte inferior de la pantalla. El usuario pulsa el botón "Edit" y el menú pasa a mostrar operaciones de edición. Si el usuario quiere mover el elemento, pulsa el botón "Move" del menú de edición y a continuación arrastra el elemento virtual con el dedo hasta otro punto de una superficie. Una vez que el dedo es levantado de la pantalla la operación acaba y vuelve a aparecer el menú de edición. Si el elemento virtual quiere ser rotado, el usuario pulsa sobre el botón "Rotate". A continuación, el autor coloca el dedo sobre la pantalla y lo desliza sin levantarlo hacia la izquierda o la derecha. Eso hace que el elemento gire en dicho sentido, sobre el eje perpendicular a la superficie en la que se encuentra. Si el dedo vuelve a la posición inicial, el elemento virtual para de girar. En cambio, si el dedo es desplazado hacia el otro lado, el elemento virtual empezará a girar en el sentido contrario. Cuando el dedo es levantado de la pantalla, el proceso de rotación termina y vuelve a aparecer el menú de edición. Si el usuario pulsa sobre el botón "Scale", se podrá redimensionar el elemento virtual. Para ello el usuario, posa los dos dedos sobre la pantalla y realizar un gesto de *pitch* o *spread*. Si el usuario separa los dedos, el elemento virtual empezará a crecer. Si la distancia entre los dedos se vuelve a ser parecida a la inicial, el elemento virtual parará de redimensionarse. En cambio, si los dedos se aproximan, el elemento virtual empezará a hacerse más pequeño. Cuando el usuario levanta los dedos de la pantalla, la operación termina y el menú vuelve a aparecer. Como en el enfoque anterior un clon semitransparente muestra la posición iniciado del elemento virtual mientras este es editado. Si un elemento quiere ser borrado, el usuario selecciona dicho elemento y pulsa el botón "Remove" el menú de acción. De manera similar, esta herramienta proporciona la posibilidad de definir lógicas de comportamiento. A continuación, la Tabla 6.3 muestra un resumen de los gestos sobre la pantalla del dispositivo móvil.

TABLA 6.3 RESUMEN DE LOS GESTOS TÁCTILES PARA LA EDICIÓN DE UN ELEMENTO VIRTUAL

Acción de Edición	Gestos	Nota
Seleccionar elementos virtuales y pulsar botones virtuales	 Tap sobre un botón o elemento virtual.	La posición del dedo determina qué se está pulsando en la pantalla
Mover Elementos Virtuales	 Arrastrar y soltar un elemento virtual.	La posición final del dedo determina la posición final del elemento que se arrastra.
Rotación de Elementos Virtuales	 Deslizamiento a la izquierda o derecha para iniciar la modificación de la rotación de un elemento virtual.	La posición de los dedos determina el sentido de la modificación, su inicio y final.
Redimensionar Elementos Virtuales	 Gestos de <i>pinch and spread</i> para iniciar la modificación del tamaño de un elemento virtual.	

### 6.4.3 Ejecutar la Experiencia de RA

Una vez completado el proceso de autoría para definir una escena RA, en ambos enfoques (RV y RA), el modo ejecución de la aplicación móvil permite desplegar la experiencia RA. Para ello, el usuario se debe desplazar al entorno aumentado para que el móvil reconozca el entorno y a continuación los objetos virtuales, añadidos con anterioridad a la escena, se muestren en la pantalla del dispositivo. Si en el proceso de creación se hubiera definido alguna lógica, estas podrían ser desencadenadas si se dieran las condiciones durante la fase de ejecución de RA.

## 6.5 Metodología

Para estudiar las tres hipótesis enunciadas en la sección 6.3 se llevó a cabo un estudio experimental con usuarios a través de un diseño *within-subjects* de medidas repetidas que involucró un total de 10 participantes. Este tipo de experimentos permite detectar diferencias entre las condiciones que experimentan los mismos participantes y proporcionar retroalimentación sobre cada una de ellas [89]. Este tipo de experimento fue seleccionado dado que en el fondo esta investigación tiene como fin comparar y evaluar si la autoría inmersiva en RV proporciona un soporte al desarrollo de escenarios de RA adecuado como lo hace las herramientas de autoría de RA basadas en dispositivos móviles.

Dado el objetivo del estudio, como variable independiente controlable se estableció el método de autoría que se concretó en la herramienta que utiliza el participante: “Herramienta de Autoría de realidad aumentada basada en RV inmersiva” o “Herramienta de Autoría de realidad aumentada basada en móvil RA”. Las variables que actúan como variables dependientes en el estudio se detallan en profundidad en el apartado 6.5.4 Instrumentos de Recolección de Datos, pero se puede avanzar que son datos que permiten estudiar cada una de hipótesis expuestas.

### 6.5.1 Estudio Piloto

Antes de llevar a cabo el experimento, se realizó un estudio piloto que tuvo como objetivo validar y garantizar la usabilidad de los dos prototipos empleados, y de refinar el proceso de evaluación. Un estudio piloto es una herramienta científica en la que se lleva a cabo un experimento con las mismas condiciones propuestas para el estudio real, pero con una población más pequeña [89]. El objetivo de este instrumento es la depuración del proceso que se llevará en el estudio genuino para identificar, por ejemplo, preguntas ambiguas, fuentes de datos adicionales que se han observado como necesarias, etcétera. Como cualquier estudio, este consume recursos como tiempo, dinero, participantes, pero su ventaja puede ser mayor al poderse identificar problemas que podrían invalidar los resultados del estudio final. Aunque



siempre debería llevarse a cabo uno, es posible que, debido a las características del estudio o limitaciones de recursos como el número de participantes, no pueda seguirse esta recomendación.

El estudio piloto de esta investigación se realizó con un total de 6 participantes. Como paso previo al inicio del experimento piloto, los todos participantes firmaron un formulario de consentimiento e hicieron una lectura de la hoja de información del experimento. Tras ello se les solicitó rellenar un cuestionario sobre preguntas demográficas como la edad y conocimientos en RV, RA y decoración. Posteriormente, los participantes utilizaron las correspondientes herramientas de autoría siguiendo el orden asignado por un diseño *Latin Square Design* de 2x2. Los participantes impares hicieron uso de la herramienta basada en RA antes de utilizar el enfoque de RV, mientras que los pares lo hicieron en orden inverso. La herramienta de RA se ejecutaba en un móvil en posición horizontal. Previamente, los participantes vieron un vídeo tutorial de aproximadamente 3 minutos en el que se explicaba el funcionamiento de dicha herramienta. Tras ello, un investigador solicitaba a los participantes llevar a cabo un conjunto de tareas en orden con indicaciones verbales que seguían la siguiente estructura: “Colocar el elemento X sobre una mesa / en una pared / en techo”. Donde X era un elemento virtual del repositorio. Además, se solicitó añadir ciertas lógicas de comportamiento como “El elemento X debe explotar/incendiarse/volcar al acercarse al elemento en Y metros/pulsar sobre el elemento/ transcurrir Z segundos”. Donde de nuevo X era un elemento virtual del repositorio mientras que Z e Y un número de 1 a 15. Una vez completadas las tareas iniciales, los participantes observaron dos imágenes en un dispositivo móvil de 5,1 pulgadas donde aparecía un mismo elemento virtual en el entorno desde diferentes perspectivas. Mientras estas eran observadas, se les indicó a los participantes que su siguiente actividad era la reproducción de la ubicación, tamaño y rotación del elemento virtual mostrado. En ambos casos, el uso de menús y *slides* situados en la parte inferior de la pantalla permitía al usuario la realización de operaciones de rotación y redimensionamiento del elemento, y la definición de lógicas de comportamiento. En cuanto a las tareas de ubicación de un elemento virtual o movimiento a otra localización, los participantes tenían que indicarlo sobre la pantalla táctil del dispositivo. Una vez que el usuario finalizaba todas las tareas, se accedía al modo de ejecución de la experiencia RA con el móvil para probar la experiencia que había creado. Posteriormente, los participantes rellenaron un cuestionario acerca de la usabilidad y satisfacción general de la herramienta.

En cuanto a la herramienta de RV inmersiva, los participantes eran sentados en una silla y se les colocaba el HMD de Samsung Gear equipado con un móvil Samsung Galaxy S6. Una vez que el casco estaba bien ajustado, se les colocaba en la mano derecha el Samsung Gear Controller, que aparecía representado en el mundo virtual como un puntero laser. No se llevó a cabo ninguna adaptación en función de la mano dominante de los participantes. Dado que el experimento se realizó en la misma sala que iba a ser aumentada, se decidió que el usuario apareciera en el mismo punto donde se encontraba sentado en una réplica del entorno real. Como formación de la herramienta, los participantes realizaron un tutorial en RV para el aprendizaje sobre las funciones del mando y las operaciones realizables sobre los elementos virtuales de la escena. Una vez el tutorial era completado, el investigador empezaba a solicitar a los participantes un conjunto de tareas en orden para crear elementos virtuales e implementar lógicas. Cuando la lista fue completada, dos paneles curvos virtuales delante de la visión del usuario aparecían mostrando cada uno una imagen del mismo elemento virtual desde dos perspectivas diferentes. El tipo de elemento virtual, su ubicación, rotación y tamaño era diferente al que se solicitaba en la herramienta basada en RA. Una vez examinado, los participantes tenían que reproducir esa configuración de ubicación, tamaño y rotación del elemento virtual dentro del mundo virtual. Una vez terminada la tarea, los participantes quitaron el HMD y se les proporcionó un móvil para probar y ejecutar la experiencia de RA construida. Después, como en el anterior caso, rellenaron dos cuestionarios acerca de la usabilidad y satisfacción. Posteriormente tras la realización del cuestionario de la segunda

herramienta, se les solicitó responder varias preguntas abiertas sobre su opinión acerca del contexto podrían ser utilizadas cada herramienta.

Tras analizar los datos obtenidos y junto a la percepción de los investigadores de cómo fue desarrollándose el experimento piloto se identificaron los siguientes sucesos:

- Baja de la cantidad de *frames* por segundo en la ejecución de la herramienta de RV cuando era grabada la visión del usuario. Esta limitación era debida en parte a la potencia del Samsung Galaxy S6 y la optimización del prototipo. Este problema a su vez provocaba una pérdida de la calidad de experiencia de usuario dificultado su uso y una mayor probabilidad para inducir mareos a los participantes.
- Imposibilidad de resolver dudas de los participantes durante el tutorial RV. Los participantes comentaron problemas para leer las indicaciones del tutorial y algunos hicieron preguntas que fueron difíciles de resolver dado que el investigador no tenía acceso a lo que observaba el participante en tiempo real. Esto alargaba el tiempo del tutorial y dificultaba su seguimiento.
- Pobre interacción basada en interfaces virtuales 2D dentro de RV e botones superpuestos en la pantalla RA. Ambas herramientas fueron diseñadas lo más parecidas posibles, sin embargo, al final la diferencia se basaba en la utilización de interfaces virtuales 2D dentro del mundo virtual e interfaces superpuestas en la pantalla del móvil. Por tanto, ninguna de las herramientas podía hacer uso del potencial intrínseco de la propia tecnología.
- Algunos participantes no entendieron la utilidad de cuando utilizar cada tecnología. Esto podría ser debido a que al final ambos procesos de autoría se realizaban en el mismo espacio a aumentar o no tenían unos claros conocimientos de la diferencia entre RA y RV. Esto provocaba que los participantes no pudieran ser conscientes de la utilidad de cada enfoque.
- Diferente dificultad para interpretar las imágenes de la tarea de recrear la ubicación, rotación y tamaño de un elemento virtual por los participantes. El soporte y las imágenes mostradas cuando se empleaba cada herramienta era diferente por lo que podría existir un nivel de dificultad diferente. Por ejemplo, debido a un menor número elementos físicos reconocibles de referencia o simplemente el tamaño. Esto podría penalizar la precisión en uno de los dos enfoques.
- La tarea guiada por el investigador limitaba el dinamismo del participante a probar la herramienta. Cuando la herramienta era utilizada por los participantes se limitaban a seguir las indicaciones y preguntar al investigador si estaba bien lo que acababan de hacer. Incluso cuando se les había indicado desde el principio que únicamente se encargaban de dictar las indicaciones y ser observadores de la tarea. Esto invalidaba los datos de tiempo empleado debido a las interrupciones.
- La definición de una lógica de comportamiento en ambas herramientas no presentaba suficiente valor para ser estudiada. Esta tarea consistía en navegar por los menús de ambos enfoques. Por lo que al final, solo se estaría dado evaluando el tiempo y esfuerzo en la navegación de menús.

Las modificaciones llevadas a cabo para solventar los problemas identificados de cara al estudio final fueron:

- Adaptación de la herramienta de RV a un HMD que no dependiera de la potencia de un móvil y que al mismo tiempo permitiera mostrar al investigador la visión del participante en tiempo real. De este modo se podía grabar la sesión sin perjudicar el rendimiento de la visión del participante y poder identificar y actuar en consecuencia antes una duda o un problema técnico si lo hubiera.
- Explicación del funcionamiento de ambas herramientas de manera verbal por parte del investigador. De este modo no existe una diferencia en la manera de aprender a usarlas y el investigador puede recibir un *feedback* del participante y poder llegar a responder dudas, si las hubiera.



- Hacer uso de una interacción más natural de cada tecnología permitiendo así una valoración el potencial que proporcionan cada una de ellas. En la herramienta de RV el participante puede hacer gestos 3D, mientras que, en la herramienta móvil, los participantes pueden hacer gestos sobre la pantalla táctil.
- Llevar a cabo el experimento en dos salas diferentes para mostrar cuál es la ventaja de la herramienta basada en RV. Además, esta decisión pone a prueba el enfoque de RV cuando el usuario no ha estado nunca físicamente en el entorno a aumentar.
- Hacer uso del mismo soporte para mostrar las imágenes para proporcionar la misma precisión por el tamaño y emplear las dos mismas imágenes para evitar una diferencia de dificultad en cuanto a las referencias del resto de elementos reales que aparecen en la imagen.
- Realización de una tarea no guiada de colocar un conjunto de elementos virtuales libremente en la primera fase del experimento. De este modo el participante puede llevar a cabo de manera más dinámica y obtener una percepción de las herramientas menos limitada.
- No incluir la definición de lógica durante el experimento. Dado que no el objetivo del estudio no es comparar la navegación de menús 2D en RV con botones en una pantalla táctil, fueron excluidas esta funcionalidad en el estudio.

### **6.5.2 Participantes y Dispositivos**

Un total de 10 voluntarios participaron en el estudio, de las cuales 3 fueron mujeres. La media de edades de los participantes fue 24,8 (20 - 37 años). Entre la información demográfica de conocimientos previos casi la mitad de los usuarios declararon que tenían poco o nulas experiencias previas de realidad virtual (6 participantes) y realidad aumentada (5 participantes). Solo dos sujetos destacaron tener conocimientos a nivel de experto en estas tecnologías. Todos estos participantes de manera aleatoria para reducir la variabilidad interindividual del estudio fueron agrupados en dos grupos: G1 y G2. Que únicamente determinaban el orden de los valores del dominio de la variable independiente, es decir, el orden de utilización de las herramientas de autoría.

Para el estudio de la técnica de RV inmersiva se utilizó el HMD de Oculus Rift CV1, utilizando los controles Oculus Touch como mecanismo de interacción con los elementos del mundo virtual. La aplicación se ejecutó en un portátil Intel® Core™ i7-7700HQ CPU (2.8 Ghz y 16 GB de RAM) con una tarjeta gráfica NVIDIA® GeForce® GTX 1060 6GB GDDR5.

Con relación al aparato empleado para el estudio de la técnica de autoría mediante RA, los participantes hicieron uso de un móvil Asus Zenfone™ AR que incluye la tecnología Google Tango.

Respecto a la reproducción del resultado final, en ambos casos se hizo uso del dispositivo móvil Tango.

### **6.5.3 Procedimiento**

El procedimiento que se llevó a cabo en el estudio final constó de los siguientes pasos. En primer lugar, y antes de comenzar el experimento, el investigador proporcionó a los participantes un formulario de consentimiento y un documento informativo, incluido en el Anexo A, en el que se describía de forma breve el contexto del experimento y un resumen de las tareas a realizar. A continuación, se solicitaba al participante rellenar un cuestionario demográfico. Una vez completado el cuestionario, el participante era asignado de forma aleatoria a un grupo, G1 o G2. Ambos grupos realizan las mismas tareas, pero en orden de las herramientas a probar era diferente con el objetivo de reducir la variabilidad interindividual y disminuir el impacto del efecto de aprendizaje. A continuación, se les solicitaba a los participantes probar una de las dos herramientas (la herramienta basada en RV en el grupo G1, y la herramienta basada en RA en el grupo G2). Una vez completadas las tareas de la primera herramienta, se les proporcionó un cuestionario sobre la experiencia. Después de esto, los

participantes repitieron las mismas tareas con la segunda herramienta y volvieron a rellenar un nuevo cuestionario sobre la experiencia de autoría. El segundo cuestionario incluía preguntas adicionales específicas sobre comparación entre ambos enfoques y una pregunta abierta en la que podían argumentar alguna respuesta dada, incluir más comentarios o cualquier información que considerasen relevante. Tras completar este último cuestionario, fueron entrevistado brevemente.

La tarea de composición de la escena de RA consta de dos fases. En la primera fase los participantes debían situar libremente 9 objetos virtuales, con una posición, tamaño y orientación a su elección. La colección de objetos a colocar incluía muebles y electrodomésticos de diverso tamaño, como sillas, televisiones, alarmas de incendio o carteles de salida de emergencia. En la segunda fase, los participantes debían colocar un único objeto, una silla, pero de tal forma que su posición y aspecto replicase de la forma más exacta posible la mostrada previamente en varias fotografías. En el caso de hacer uso de la herramienta de RV, el investigador ayudó a retirar el HMD del participante y le ayudó a ajustárselo de nuevo una vez revisadas las imágenes. Tras terminar esta segunda fase, se les proporcionó a los participantes un móvil para probar la experiencia que habían creado. En el caso de usar la herramienta de RV, tanto el participante como el investigador se desplazaron a la sala que había sido aumentada para que el participante probara la experiencia que había elaborado.

#### **6.5.4 Instrumentos de Recolección de Datos**

Para los dos valores que tomó la única variable independiente del estudio, herramienta de autoría de RA basada en uso de la RV Inmersiva y herramienta de autoría de RA basada en el uso de un móvil RA, varias variables dependientes fueron recogidas. Las variables dependientes enfocadas a estudiar el esfuerzo físico y mental fueron recogidas mediante el cuestionario NASA-Task Load Index (NASA RTLX) [60]. La evaluación de la carga de trabajo se considera un factor clave para el diseño de interfaces que logren un alto nivel de satisfacción, confort, seguridad y eficiencia en un lugar de trabajo [134]. El cuestionario de NASA-TLX es uno de los más comúnmente empleados y se basa en la medición de 6 dimensiones: esfuerzo mental, esfuerzo físico, tiempo requerido, grado de fracaso, esfuerzo, nivel de estrés. Cada una de estas dimensiones se califica en uno de los 21 niveles equidistante en un rango delimitado por 0 y 100 en los extremos. El promedio de todas las dimensiones se define como el coeficiente Raw Task Load Index (RTLX) que cuantifica la carga general del participante.

La recogida de la variable de satisfacción se realizó mediante un cuestionario diseñado para evaluar la satisfacción de realización de 4 operaciones con las herramientas. Estas operaciones fueron derivadas de las tareas básicas en entornos virtuales [17,18]. Estas 4 operaciones básicas abarcan seleccionar un elemento virtual y su manipulación (mover, rotar y redimensionar). Además, los participantes valoraron la dificultad, la precisión y la facilidad de aprendizaje para usar la herramienta. Todas estas preguntas fueron respondidas mediante escalas de Likert de 7 valores. Los valores extremos del rango de cada pregunta fueron etiquetados para una mejor comprensión de la graduación.

Adicionalmente, los participantes respondieron un cuestionario comparativo de ambas herramientas. Más específicamente, los participantes fueron preguntados su grado de predilección entre las dos herramientas respecto a la satisfacción general, la precisión, la facilidad para aprender, el uso cuando se manipulan objetos pequeños, el uso cuando se manipulan elementos grandes, la preferencia general y el mejor resultado. Por último, fue incluida en el cuestionario una pregunta abierta donde los participantes podían justificar alguna respuesta dada o incluir sugerencias y opiniones adicionales.

Además de los datos de los cuestionarios, las herramientas recolectaron y almacenaron la posición, la rotación y el factor de escala de los elementos virtuales. Adicionalmente, los dispositivos fueron habilitados para grabar la pantalla y las voces de los participantes utilizando aplicaciones de terceros (Game Bar Recorder y AZ Screen Recorder, para el ordenador y dispositivo móvil respectivamente). A pesar de que la ejecución de estas aplicaciones

representaba un consumo adicional de recursos de los dispositivos el rendimiento de las herramientas no se vio afectado.

## 6.6 Resultados

En cuanto a los datos obtenidos acerca de la sobrecarga de trabajo mental y física del NASA TLX están resumidos en la Tabla 6.4. Ambas herramientas fueron valoradas positivamente por los participantes. La valoración más baja se obtuvo en el Esfuerzo Físico (EF) de la herramienta de RV y la herramienta de RA fue la que mostró un mayor Tiempo Requerido (TR). Los valores bajos en Grado de Fracaso (GF) mostrados en ambas herramientas se traduce en un alto grado de perfección, de cualquier manera, no se observaron diferencias significativas entre ellas. Es importante recordar que valores bajos en un cuestionario NASA TLX tienden a indicar que el sistema no sobrecarga al usuario. RV mostró mejores resultados en esfuerzo físico (EF), Tiempo Requerido (TR), Frustración (F) y Sobrecarga General (RTLX) con una diferencia de 8,5; 6; 4,5 y 1,5 puntos sobre 100 en el valor medio respectivamente. Por otro lado, RA mostró mejores resultados en las variables de Esfuerzo Mental (EM), Grado de Fracaso (GF) y Esfuerzo General (EG) con una diferencia de 6; 2,5 y 2 puntos respectivamente. Sin embargo, no se encontraron diferencias estadísticamente significativas entre las dos herramientas en ninguna de las variables anteriormente descritas. La Figura 6.4 muestra los resultados gráficamente.

TABLA 6.4 RESUMEN DE RESULTADOS DEL NASA TLX

	Herramienta de Autoría basada en RV	Herramienta de Autoría basada en RA	p-value
Esfuerzo Mental (EM)	33.5 ± 17,5	27.5 ± 14,8	0,296
Esfuerzo Físico (EF)	22.5 ± 17,2	31.0 ± 17,6	0,319
Tiempo Requerido (TR)	31.5 ± 13,8	37.5 ± 14,2	0,300
Grado de Fracaso (GF)	22.5 ± 19,9	20.0 ± 13,1	0,585
Esfuerzo General (EG)	32.0 ± 18,6	30.0 ± 16,3	0,702
Frustración (F)	26.5 ± 22,9	31.0 ± 20,7	0,567
Sobrecarga General (RTLX)	28.0 ± 18,3	29.5 ± 16,5	0,572

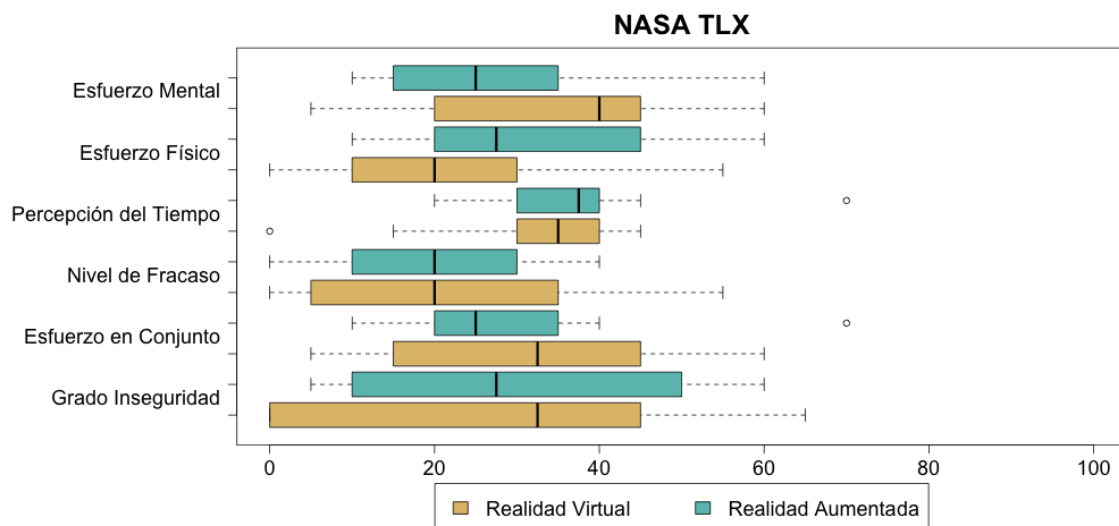


FIGURA 6.4 RESULTADOS DE LAS DIMENSIONES DE NASA TLX

En cuanto a las respuestas de la experiencia de los usuarios con la herramienta se muestran en la Tabla 6.5. En la primera parte del cuestionario, 4 tareas básicas (Q1 a Q4) fueron calificadas en una escala de 1 a 7. Como se muestra en la Tabla 6.5 la media de respuestas sobre el nivel de satisfacción se encuentra en un rango alto entre 5 y 6,3. La valoración más baja de satisfacción fue para Mover Objetos (Q2) en la herramienta de RA, y la más alta para seleccionar elementos

en RV. Las valoraciones para RV fueron mejores en Q1, Q2, Q3 e iguales en Q4 con respecto a RA. La satisfacción con la interacción de Mover Objetos (Q2) fue la única variable que mostró una diferencia estadísticamente significativa entre ambas herramientas.

En la segunda parte de cuestionario, los participantes valoraron su nivel de percepción respecto a la Dificultad de Uso (Q5), la Precisión (Q6) y la Facilidad de Aprender a Usar la Herramienta (Q7). Dicha valoración se realizó en una escala de 1 a 7. Las valoraciones fueron muy positivas respecto a la dificultad y a la facilidad de utilización. Algo más baja fue registrada en cuanto a la precisión. No hubo ninguna diferencia significativa en esta parte del cuestionario. La Figura 6.5 muestra una representación gráfica de los resultados.

TABLA 6.5 RESUMEN DE RESULTADOS DEL CUESTIONARIO CON ESCALA DE LIKERT

		Autoría RV	Autoría RA	p-value
Por favor, valora el nivel de satisfacción de las siguientes tareas desde 1 (nada satisfecho) a 7 (muy satisfecho).				
Q1	Seleccionar Objetos	6,2 ± 1,03	5,7 ± 1,34	0,322
Q2	Mover Objetos	6,0 ± 1,15	5,0 ± 1,33	0,041*
Q3	Rotar Objetos	6,0 ± 0,67	5,4 ± 1,84	0,357
Q4	Redimensionar Objetos	5,6 ± 1,58	5,6 ± 1,35	1,000
Por favor, valora en una escala de 1 a 7 los siguientes aspectos de la herramienta.				
Q5	Dificultad de Uso	2,3 ± 1,25	2,8 ± 0,63	0,343
Q6	Precisión	4,7 ± 2,05	5,0 ± 1,76	0,691
Q7	Facilidad para aprender a Usar	6,3 ± 0,95	6,0 ± 0,82	0,434

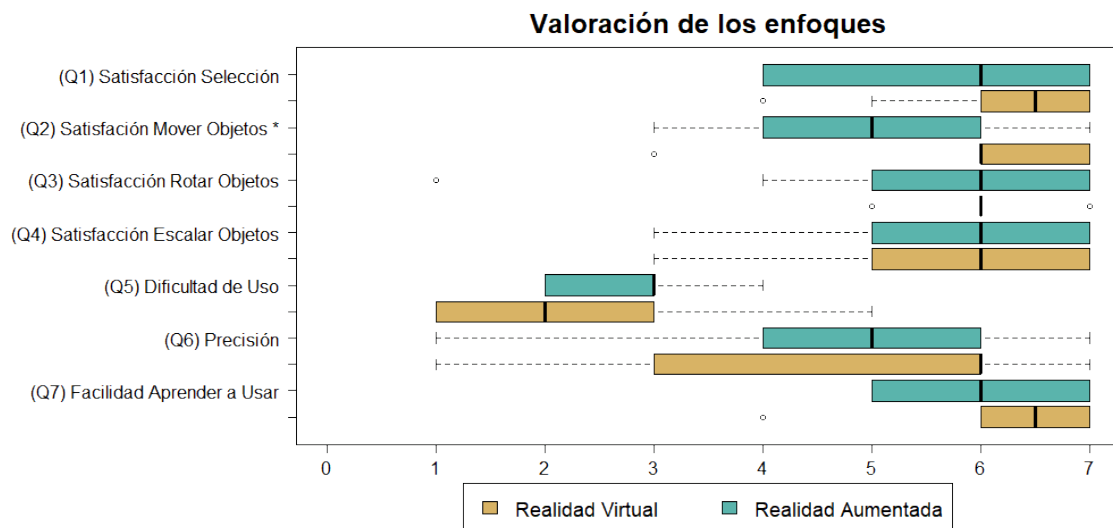


FIGURA 6.5 VALORACIÓN DE LOS USUARIOS EN LOS DIFERENTES ENFOQUES

Finalmente, los participantes realizaron un tercer cuestionario en el que se comparaban 11 variables de predilección de las herramientas. Los resultados se muestran en Figura 6.6 Comparación de los enfoques en función de los aspectos. En general, no indicaron una predilección por ninguna de las herramientas. Solamente “Colocar objetos grades” en RV fue marcado por la mayoría de los participantes (n=8).

La pregunta abierta no permitió destacar una gran diferencia en términos de experiencia y satisfacción del usuario entre las dos herramientas. Algunos participantes describieron pequeños problemas cuando manipulaban objetos (“Es difícil ser precisos cuando rotas un objeto en RA”) y proporcionaron sugerencias (“En RV podría ser más fácil de rotar usado la muñeca gira en el eje z”), pero ninguno reportó ningún problema serio u obstáculo para la realización de las tareas. Solo dos participantes expresaron una clara preferencia por una de las dos herramientas, la de RV indicando que les parecía más entretenida. Adicionalmente, dos

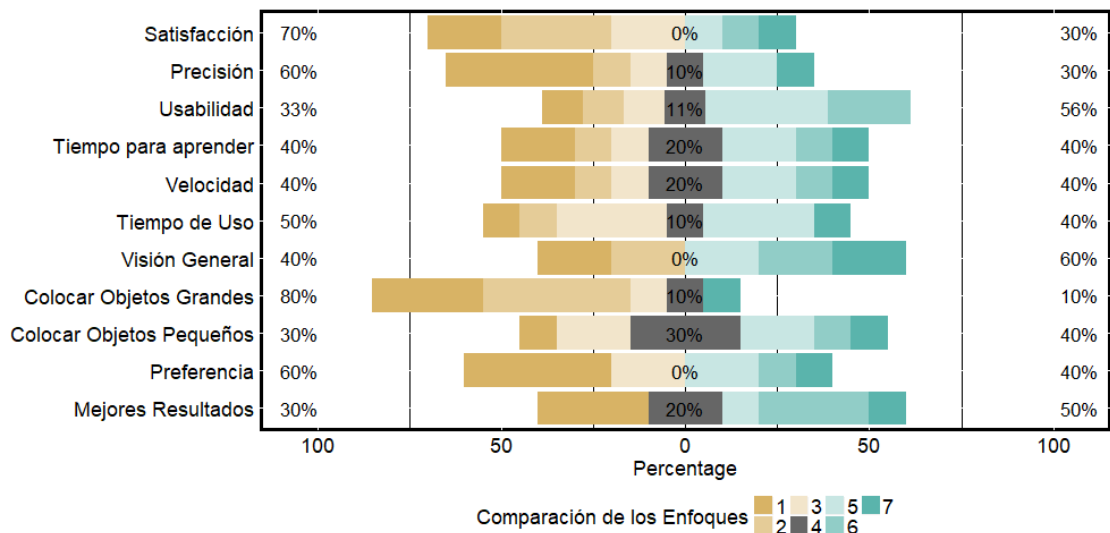


FIGURA 6.6 COMPARACIÓN DE LOS ENFOQUES EN FUNCIÓN DE LOS ASPECTOS

sujetos mencionaron que se sintieron un poco mareados y les costó tiempo hacerse con los controles para desplazarse en RV.

En cuanto al tiempo requerido para completar las tareas de autoría, únicamente se consideró a estudiar el tiempo empleado para realizar la segunda tarea, tomando como el tiempo transcurrido desde que el participante activaba el menú hasta que el propio participante indicaba que el proceso se había concluido. El tiempo medio empleado fue menor en la herramienta de RV inmersiva ( $54,5 \pm 30,2$  segundos frente a  $75,1 \pm 39,6$  segundos) mostrándose una diferencia estadísticamente significativa entre ambas herramientas ( $p=0,03$ ).

## 6.7 Discusión

El objetivo de este trabajo fue conocer los beneficios y limitaciones del uso de la RV inmersiva durante parte del proceso de autoría de RA en comparación con un enfoque más tradicional basado en el uso de un teléfono móvil con RA. En la primera parte de esta sección se discute los principales hallazgos sobre los tres principales objetivos iniciales: evaluación del esfuerzo requerido en las herramientas, satisfacción del usuario y efectividad en términos de tiempo durante el proceso de autoría. Posteriormente, esta sección menciona y discute otros beneficios y limitaciones que se han podido identificar durante el estudio.

*La cantidad de esfuerzo requerido para crear una escena RA en RV inmersiva es similar al uso de un móvil con RA.* De acuerdo con los resultados obtenidos a través del cuestionario NASA TLX no hay una evidencia que suponga que la creación de la escena RA usando una herramienta inmersiva RV requiera de un esfuerzo mental o físico mayor respecto al uso de un móvil RA. Los valores obtenidos a través del cuestionario fueron parecidos y bajos en ambos casos, tanto en lo referente a demanda mental, física y temporal, como en cuanto a niveles de esfuerzo, frustración y grado de fracaso. Estos resultados son especialmente relevantes considerando que la mayoría de los usuarios expresaron que tenían una baja experiencia en tecnologías RV, como en HMD y controles de RV, mientras que están acostumbrados a utilizar aplicaciones de teléfonos móviles de manera diaria. Estos resultados pueden verse justificados en parte debido al uso de una interacción natural que no requiere una atención alta por parte de los participantes para la realización de las tareas solicitadas.

*El nivel de satisfacción de los participantes cuando creaban una escena RA in RV inmersiva es similar a cuando se hacía uso de un móvil RA.* De manera similar, los resultados en cuanto a nivel de satisfacción de los participantes al realizar la composición de la escena RA mediante RV o RA fueron también similares. En términos de satisfacción (preguntas Q1-Q4 en el cuestionario) los usuarios no mostraron una clara inclinación por ninguna de las dos técnicas. En cuanto a interacción, los participantes parecen valorar algo más positivamente (pregunta Q2) la

herramienta de RV inmersiva que la RA, pero en cualquier caso la diferencia era pequeña. De nuevo es importante señalar que la mayoría de los participantes no había usado tecnología de RV con anterioridad, por lo que resulta especialmente revelador que tras una corta sesión fuesen capaces de utilizar la herramienta de forma RV de forma tan satisfactoria a como utilizan una herramienta móvil a la que están acostumbrados. De igual forma, la alta valoración obtenida en las preguntas sobre la facilidad de aprendizaje, y la baja valoración en términos de dificultad de uso, sugiere que la herramienta de RV inmersiva puede constituir una alternativa viable al tradicional enfoque basado autoría móvil con RA.

*La efectividad en términos de tiempo necesario para manipular un elemento virtual parece ser más alto en la herramienta inmersiva de RV que en la herramienta móvil.* Los datos extraídos de los vídeos tomados durante el estudio indican que hay una diferencia significativa en cuanto a velocidad para llevar a cabo la segunda tarea. A pesar de estos excepcionales resultados en el estudio, es necesario ser cautos debido al limitado número de muestras y duración de la tarea. Sin embargo, de cualquier manera, estos resultados sugieren una tendencia de mejora en la efectividad en cuanto al tiempo. Los resultados obtenidos en el estudio pueden ser entendidos debido a uno de los siguientes tres factores o por la combinación de varios. El primero de ellos es la manera de desplazarse por el entorno a aumentar. Con la herramienta móvil cuando un usuario se quería acercarse a un elemento virtual para interactuar con él, tenía que moverse esquivando los obstáculos del mundo real mientras miraba la pantalla para ver cuánto de cerca estaba. Sin embargo, dentro de la réplica virtual del mundo el usuario podía desplazarse en línea recta con un joystick atravesando todos los obstáculos y pudiendo fijar su mirada en el destino final. El segundo factor que puede estar relacionado con el resultado obtenido es el campo de visión que se obtiene de la escena RA. Mientras que el FoV de RA es limitado debido al tamaño de la pantalla, las Oculus Rift proporcionan una perspectiva de la escena más grande para el elemento virtual con el que se trabaja y su resultado en entorno. De este modo, el usuario en RV podría serle más fácil o necesitar menos puntos de vista diferentes para saber si la silla estaba colocada de acuerdo con sus expectativas. Además, algunas de las perspectivas que se pueden lograr dentro de la réplica no son accesibles en el mundo real. El tercero de los factores podría estar relacionado con el limitado alcance que tiene la cámara de profundidad frente al puntero laser disponible en el entorno RV, el cual podía seleccionar y mover elementos a mucha más distancia. De esta forma, los participantes no necesitaban acercarse tanto al elemento virtual para realizar el *drag and drop* del elemento virtual. Por último, habría que añadir que estos buenos resultados podrían ser incluso potenciados si los botones de los mandos en RV inmersiva fueran configurados como atajos para operaciones de edición determinadas o se complementasen los mecanismos de locomoción en RV con teletransporte en el entorno, que no son posibles su aplicación en la aplicación móvil debido al tamaño limitado de la pantalla del dispositivo y a la naturaleza del enfoque.

En los comentarios de los participantes obtenidos a través de las preguntas abiertas y las entrevistas parecen confirmar que ambas técnicas pueden ser utilizadas de forma igualmente efectiva, y que no existen grandes diferencias en términos de esfuerzo físico, mental y satisfacción entre ambas. Además, ninguno de los participantes del estudio reportó cansancio o sentir fatiga en los brazos durante el experimento como consecuencia de los gestos necesarios para colocar los objetos virtuales en el entorno RV, o de tener que sostener el móvil y manejar la aplicación con las manos durante tiempo prolongado, en el caso de la aplicación de RA. En cuanto a los gestos de interacción de los elementos virtuales, un participante comentó que le sería más fácil rotar la muñeca cuando quería rotar un elemento virtual colgado en la pared. Tras observar los vídeos de los participantes se pudo observar que otros participantes inconscientemente intentaron realizar este gesto, lo que da pie a pensar que los participantes estaban tan centrados en la tarea que inconscientemente pudieron pensar que lo estaban sosteniendo. Lo que demuestra el potencial que tiene la interacción natural en general para el proceso de autoría de RA y el *engagement* en una tarea concreta si se hace de manera inmersiva.

## 7 CONCLUSIONES, CONTRIBUCIONES Y TRABAJOS FUTUROS

Este capítulo se centrará en resumir las conclusiones de los trabajos de la tesis realizada, las contribuciones generales de la tesis y los trabajos futuros que podrían ser llevados a cabo posteriormente.

### 7.1 Conclusiones

Este trabajo de investigación se ha centrado en el proceso de creación y autoría de experiencias de realidad aumentada para usuarios sin un alto grado de conocimiento técnico.

La primera parte de la investigación estuvo enfocada al desarrollo de un proceso de creación de experiencias de realidad aumentada que pudiese ser empleado por usuarios sin un perfil tecnológico alto, y que al mismo tiempo permitiese dotar de un alto realismo a las experiencias desarrolladas, aun cuando estas tuviesen lugar en escenarios de grandes dimensiones o exteriores. Con este fin se diseñó un proceso de desarrollo de experiencias de RA realistas que combina el uso de técnicas basadas en modelos del entorno con los beneficios que los actuales motores de videojuegos ofrecen para el diseño y gestión de los elementos virtuales 3D. Con el fin de poder evaluar el enfoque propuesto se desarrolló un prototipo de un sistema de RA que implementaba el enfoque propuesto y extendía las funcionalidades de una plataforma de videojuegos existente llamada GREP. Este prototipo fue evaluado por medio de una experiencia durante la que se aumentó el auditorio de la Universidad Carlos III de Madrid con distintos elementos virtuales que fueron integrados en el entorno empleando distintos efectos de oclusión, perspectiva, colisión, luces y sombras, y con los que el público asistente interactuó durante dos juegos AR colaborativos. Con los resultados obtenidos y el conocimiento extraído durante la elaboración de la experiencia se diseñó una arquitectura de sistema que diera soporte al proceso. Posteriormente, se llevó a cabo un segundo caso de uso que tuvo como fin evaluar la capacidad del enfoque propuesto a la hora de aumentar espacios exteriores abiertos. Durante el desarrollo de este caso la arquitectura inicial del sistema se extendió con el fin de disponer de mecanismos para flexibilizar el proceso de alineación entre el espacio real a aumentar y la réplica de este en el mundo virtual.

La segunda parte del trabajo de investigación se centró en el proceso de autoría, y más concretamente estuvo dirigida a proponer un enfoque de creación que pudiese superar algunas de las limitaciones del tradicional método basado en dispositivos móviles, y que al mismo tiempo fuese apto para su utilización por usuarios no expertos. En la segunda parte de la tesis se desarrollaron dos herramientas de autoría de RA: una de ellas implementa el enfoque tradicional de autoría de RA, mientras que la segunda fue diseñada para llevar parte del proceso de autoría de RA de manera remota haciendo uso de la tecnología de RV. Los beneficios y limitaciones de cada enfoque de autoría fueron evaluados por medio de un experimento en el que 10 usuarios participaron creando escenas de AR empleando ambas herramientas. En general, los resultados del experimento fueron positivos para ambos enfoques de autoría, siendo similares en términos tanto de esfuerzo requerido como de satisfacción del usuario. En relación con el tiempo necesario para colocar elementos virtuales, los resultados indicaron una diferencia estadísticamente significativa que sugiere que la RV inmersiva podría tener una ventaja real. De cualquier manera, los resultados indican que el método RV inmersiva podría constituir una alternativa real al método tradicional de autoría de RA basada en tecnología móvil, permitiendo además crear experiencias de RA de manera remota, algo que no resulta posible con las habituales herramientas actuales.

## 7.2 Contribuciones

Tal y como se explicó en el capítulo 3, a la hora de tratar de dotar a usuarios sin un alto perfil tecnológico de los medios para lograr crear experiencias de AR realistas se identificaron 3 problemas principales:

1. Falta de efectos realistas en las experiencias de realidad aumentada
2. Dificultad para aumentar entornos de grandes dimensiones o exteriores
3. Limitaciones en el enfoque de la actual tendencia de herramientas de autoría para RA

Para abordar estas problemáticas se centró la tesis doctoral en llevar a cabo dos objetivos, por un lado, el diseño de un método para aumentar espacios de grandes dimensiones y hacerlo de manera realistas y, por otro, el estudio de los beneficios y limitaciones del uso de la Realidad Virtual como parte del proceso de autoría de RA. La Figura 7.1 muestra un gráfico con los problemas y su relación con los objetivos. Los trabajos realizados alrededor de estos objetivos dieron lugar a las siguientes contribuciones a la base de conocimiento y en dominio de aplicación por medio del ciclo de rigor y relevancia definidos en el marco de trabajo de Alan Hervner [63,64].

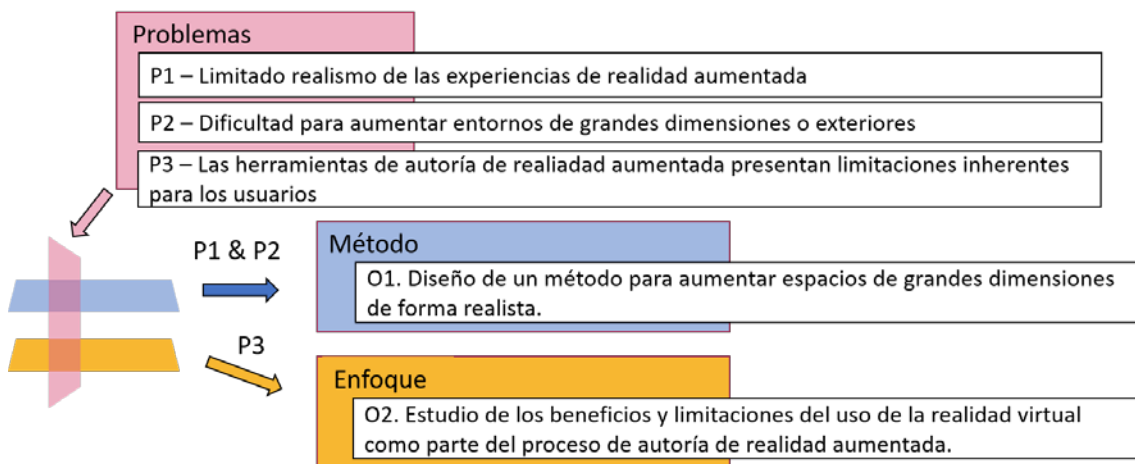


FIGURA 7.1 GRÁFICO RESUMEN DE LOS PROBLEMAS Y SU RELACIÓN CON LOS OBJETIVOS

- *Un método de creación de experiencias de RA realistas para entornos de grandes dimensiones* (Fundamento teórico): Esta contribución versa directamente sobre el objetivo O1 de la tesis. Una vez diseñado el método y evaluado el resultado de proceso en un caso de contexto real, este proceso ha podido pasar a la base de conocimiento gracias a varias publicaciones [2,108,110].
- *Herramientas y scripts desarrollados para la creación de experiencias de RA realistas en entornos de grandes dimensiones* (Artefacto para el dominio de Aplicación): Una vez hecha la evaluación del proceso y obtenidos unos resultados positivos en el estudio, las herramientas y *scripts* que fueron desarrollados forman parte de las contribuciones sobre el dominio de aplicación del estudio. En este caso un dominio de aplicación que requiera aumentar un espacio de grandes dimensiones de manera realistas.
- *Conocimiento de los beneficios y limitaciones del uso de la RV como parte del proceso de autoría de RA* (Fundamento teórico): Esta contribución está estrechamente ligada al objetivo O2 de la tesis. Una vez hecha la discusión de los datos extraídos del experimento con usuarios. Las conclusiones que se pueden extraer formarán parte de la base de conocimiento por medio de la presente tesis doctoral.
- Herramienta de autoría inmersiva de realidad aumentada para entornos inaccesibles, acceso remoto o por preferencia del usuario y herramienta de autoría móvil de RA



(Artefactos de Diseño): Las dos herramientas que han sido desarrolladas y evaluadas por usuarios son contribuciones que se realizan al dominio de aplicación estudiado en esta tesis.

### 7.3 Trabajos Futuros

Tras los resultados obtenidos en los dos estudios principales desarrollados a lo largo de esta tesis, se plantea como trabajo futuro el desarrollo de un *toolkit* que permita llevar a cabo el proceso completo de creación y autoría de realidad aumentada realista propuesto en esta tesis, y hacerlo disponible de manera libre para que usuarios y expertos en distintas áreas como educación, el patrimonio cultural o entretenimiento puedan emplearlo para desarrollar experiencias de AR por sí mismos, y explotar los beneficios que esta tecnología puede ofrecer en sus respectivas áreas.

Además, del desarrollo de esta herramienta sería conveniente la adaptación del primer método propuesto para exploración de un enfoque que permita mover la cámara que capta el entorno. Por otro lado, como se ha comentado en la discusión del apartado 5.6.1, si bien la experiencia en el auditorio permitió evaluar la capacidad del sistema desarrollado para aumentar de forma realista grandes espacios interiores, sería necesario desarrollar más experiencias mediante las que profundizar y conocer mejor en la opinión de los usuarios, y así mejorar la propuesta.

De igual forma, si bien los resultados obtenidos en el estudio comparativo entre los dos enfoques de autoría de RA presentado en esta tesis sugieren que los dos son adecuados para ser usados por usuarios sin un alto conocimiento técnico, nuevos estudios podrían servir para profundizar en los beneficios y diferencias entre ambos métodos de autoría. Así, estos experimentos podrían servir para analizar la precisión lograda durante la tarea al emplear uno u otro enfoque. En el estudio fueron recolectados la posición, rotación y factor de escala, pero no se pudieron analizar dada la existencia de una dependencia entre las coordenadas X y Z con el tamaño del elemento virtual dispuesto. Esto es debido a que según la percepción de los autores podría haber hecho el elemento virtual más pequeño de lo normal y al mismo tiempo para compensarlo acercarlo. Aun así, los datos parecen ser similares como se muestran en el Anexo C.

Adicionalmente a las evaluaciones, sería interesante poder ampliar los tipos de dispositivos con los que poder ejecutar la experiencia elaborada en el proceso, como por ejemplo las Microsoft HoloLens. En este caso sería necesario obtener la correspondencia entre el fichero de descripción de entornos de HoloLens con el modelo creado por fotogrametría. Este proceso puede ser similar al realizado, e incluso, podría ser automatizado en tiempo de ejecución haciendo uso de un apropiado algoritmo que minimice la distancia, por ejemplo, con mínimos cuadrados, entre los vértices del modelo y la figura que se forma al obtener la forma del entorno que rodea a las Microsoft HoloLens.

## 8 BIBLIOGRAFÍA

- [1] Daniel F. Abawi, Jose Luis Los Arcos, Michael Haller, Werner Hartmann, Kalle Huhtala, and Marjaana Träskbäck. 2004. A mixed reality museum guide: The challenges and its realization. In *Proceedings of the 10th International Conference on Virtual Systems and Multimedia (VSMM 2004)*.
- [2] Alvaro Montero, Telmo Zarranandia, Ignacio Aedo, and Paloma Díaz. 2016. Agree Or Perish: A Crowd Controlled Augmented Reality Pervasive Game. In *CHI Pervasive Play Workshop*.
- [3] D.B. Anderson and M.A. Casey. 1997. The sound dimension. *IEEE Spectr.* 34, 3 (March 1997), 46–50. DOI:<https://doi.org/10.1109/6.576008>
- [4] Arthur Appel. 1968. Some techniques for shading machine renderings of solids. In *Proceedings of the April 30–May 2, 1968, spring joint computer conference*, 37–45.
- [5] Luis Arenas, Telmo Zarranandia, Paloma Díaz, and Ignacio Aedo. 2015. A Platform for Supporting the Development of Mixed Reality Environments for Educational Games. In *Learning and Collaboration Technologies (Lecture Notes in Computer Science)*, Springer International Publishing, 537–548.
- [6] R. Azuma, Y. Baillet, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. 2001. Recent advances in augmented reality. *IEEE Comput. Graph. Appl.* 21, 6 (November 2001), 34–47. DOI:<https://doi.org/10.1109/38.963459>
- [7] Ronald T. Azuma and others. 1997. A survey of augmented reality. *Presence* 6, 4 (1997), 355–385.
- [8] M. Bajura and U. Neumann. 1995. Dynamic registration correction in augmented-reality systems. In *Virtual Reality Annual International Symposium, 1995. Proceedings.*, 189–196. DOI:<https://doi.org/10.1109/VRAIS.1995.512495>
- [9] Michael Bajura, Henry Fuchs, and Ryutarou Ohbuchi. 1992. Merging virtual objects with the real world: Seeing ultrasound imagery within the patient. *ACM SIGGRAPH Comput. Graph.* 26, 2 (1992), 203–210.
- [10] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. 2006. Surf: Speeded up robust features. In *European conference on computer vision*, Springer, 404–417.
- [11] Steve Benford, Chris Brown, Gail Reynard, and Chris Greenhalgh. 1996. Shared spaces: transportation, artificiality, and spatiality. In *Proceedings of the 1996 ACM conference on Computer supported cooperative work*, 77–86.
- [12] Steve Benford, Chris Greenhalgh, Gail Reynard, Chris Brown, and Boriana Koleva. 1998. Understanding and constructing shared spaces with mixed-reality boundaries. *ACM Trans. Comput.-Hum. Interact. TOCHI* 5, 3 (1998), 185–223.
- [13] M. O. Berger. 1997. Resolving occlusion in augmented reality: a contour based approach without 3D reconstruction. In *, 1997 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997. Proceedings,* 91–96. DOI:<https://doi.org/10.1109/CVPR.1997.609304>
- [14] Bhaskar Bhattacharya and Eliot H. Winer. 2019. Augmented reality via expert demonstration authoring (AREDA). *Comput. Ind.* 105, (February 2019), 61–79. DOI:<https://doi.org/10.1016/j.compind.2018.04.021>
- [15] M. Billinghurst, H. Kato, and I. Poupyrev. 2001. The MagicBook - moving seamlessly between reality and virtuality. *IEEE Comput. Graph. Appl.* 21, 3 (May 2001), 6–8. DOI:<https://doi.org/10.1109/38.920621>
- [16] Jim Blinn. 1988. Me and my (fake) shadow. *IEEE Comput. Graph. Appl.* 8, 1 (1988), 82–86.
- [17] Doug A. Bowman. 1998. Interaction techniques for immersive virtual environments: Design, evaluation, and application. *J. Vis. Lang. Comput.* 10, (1998), 37–53.
- [18] DOUG A. Bowman and LARRY F. Hodges. 1999. Formalizing the Design, Evaluation, and Application of Interaction Techniques for Immersive Virtual Environments. *J. Vis. Lang. Comput.* 10, 1 (February 1999), 37–53. DOI:<https://doi.org/10.1006/jvlc.1998.0111>

- [19] James C. Brancheau and James C. Wetherbe. 1987. Key issues in information systems management. *MIS Q.* (1987), 23–45.
- [20] David E. Breen, Eric Rose, and Ross T. Whitaker. 1995. Interactive occlusion and collision of real and virtual objects in augmented reality. *Munich Ger. Eur. Comput. Ind. Res. Cent.* (1995). Retrieved October 27, 2015 from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.38.113&rep=rep1&type=pdf>
- [21] Bricklin, D., Frankston, B. and Fylstra, D. 1979. VisiCalc, Software Arts. Retrieved September 12, 2017 from <http://www.bricklin.com/history/intro.htm>
- [22] Wolfgang Broll, Irma Lindt, Iris Herbst, Jan Ohlenburg, Anne-Kathrin Braun, and Richard Wetzel. 2008. Toward next-gen mobile AR games. *IEEE Comput. Graph. Appl.* 28, 4 (2008), 40–48.
- [23] John Brooke. 1996. SUS-A quick and dirty usability scale. *Usability Eval. Ind.* 189, 194 (1996), 4–7.
- [24] G. Bruder, F. Steinicke, K. Rothaus, and K. Hinrichs. 2009. Enhancing Presence in Head-Mounted Display Environments by Visual Body Feedback Using Head-Mounted Cameras. In *2009 International Conference on CyberWorlds*, 43–50. DOI:<https://doi.org/10.1109/CW.2009.39>
- [25] Marius Bulearca and Daniel Tamarjan. 2010. Augmented reality: A sustainable marketing tool. *Glob. Bus. Manag. Res. Int. J.* 2, 2 (2010), 237–252.
- [26] Alberto Candussi, Tobias Hollerer, and Nicola Candussi. 2005. Real-time rendering of realistic trees in mixed reality. In *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR'05)*, IEEE, 204–205.
- [27] Michelle Castillo. 2017. Watch Apple demonstrate iPhone 8 AR augmented reality. Retrieved September 18, 2017 from <https://www.cnn.com/2017/09/12/watch-apple-demonstrate-iphone-8-ar-augmented-reality.html>
- [28] Thomas P. Caudell and David W. Mizell. 1992. Augmented reality: An application of heads-up display technology to manual manufacturing processes. In *System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on*, IEEE, 659–669. Retrieved June 10, 2017 from <http://ieeexplore.ieee.org/abstract/document/183317/>
- [29] J. W. S. Chong, S. K. Ong, A. Y. C. Nee, and K. Youcef-Youmi. 2009. Robot programming using augmented reality: An interactive method for planning collision-free paths. *Robot. Comput.-Integr. Manuf.* 25, 3 (June 2009), 689–701. DOI:<https://doi.org/10.1016/j.rcim.2008.05.002>
- [30] James C. Chung, Mark R. Harris, Fredrick P. Brooks, Henry Fuchs, Michael T. Kelley, John Hughes, Ming Ouh-Young, Clement Cheung, Richard L. Holloway, and Michael Pique. 1989. Exploring virtual worlds with head-mounted displays. In *Three-Dimensional Visualization and Display Technologies*, International Society for Optics and Photonics, 42–53.
- [31] Comeau, C.P. and Bryan, J.S. 1961. Headsight television system provides remote surveillance - Google Académico. *Electronics*, 86–90.
- [32] Santiago Cortés, Arno Solin, Esa Rahtu, and Juho Kannala. 2018. ADVIO: An authentic dataset for visual-inertial odometry. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 419–434.
- [33] Franklin C. Crow. 1977. Shadow algorithms for computer graphics. *Acm Siggraph Comput. Graph.* 11, 2 (1977), 242–248.
- [34] Brian Curless and Marc Levoy. 1996. A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, 303–312.
- [35] Sumesh S. Dadwal and Azizul Hassan. 2015. The augmented reality marketing: A merger of marketing and technology in tourism. In *Emerging innovative marketing strategies in*

- the tourism industry*. IGI Global, 78–96. Retrieved September 18, 2017 from <https://books.google.es/books?hl=es&lr=&id=hXtRCgAAQBAJ&oi=fnd&pg=PA78&dq=augmented+reality+marketing&ots=WWVG22cYOM&sig=mmO34gDxBLi6EHKRra1R9-LiLiSA>
- [36] Areti Damala, Tobias Schuchert, Isabel Rodriguez, Jorge Moragues, Kiel Gilleade, and Nenad Stojanovic. 2013. Exploring the affective museum visiting experience: adaptive augmented reality (A2R) and cultural heritage. *Int. J. Herit. Digit. Era* 2, 1 (2013), 117–142.
- [37] Abe Davis, Justin G. Chen, and Frédo Durand. 2015. Image-space modal bases for plausible manipulation of objects in video. *ACM Trans. Graph. TOG* 34, 6 (October 2015), 239:1–239:7. DOI:<https://doi.org/10.1145/2816795.2818095>
- [38] Andrew J. Davison, Ian D. Reid, Nicholas D. Molton, and Olivier Stasse. 2007. MonoSLAM: Real-time single camera SLAM. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 6 (2007), 1052–1067.
- [39] Tina LM Derzaph and Howard J. Hamilton. 2013. Effects of wind on virtual plants in animation. *Int. J. Comput. Games Technol.* 2013, (2013).
- [40] Fabian Doil, W. Schreiber, T. Alt, and C. Patron. 2003. Augmented reality for manufacturing planning. In *Proceedings of the workshop on Virtual environments 2003*, ACM, 71–76. Retrieved September 18, 2017 from <http://dl.acm.org/citation.cfm?id=769962>
- [41] Ralf Dörner, Christian Geiger, Michael Haller, and Volker Paelke. 2003. Authoring mixed reality—a component and framework-based approach. In *Entertainment Computing*. Springer, 405–413.
- [42] Ashish Doshi, Ross T. Smith, Bruce H. Thomas, and Con Bouras. 2017. Use of projector based augmented reality to improve manual spot-welding precision and accuracy for automotive manufacturing. *Int. J. Adv. Manuf. Technol.* 89, 5–8 (2017), 1279–1293.
- [43] Tom Drummond and Roberto Cipolla. 2002. Real-time visual tracking of complex structures. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 7 (2002), 932–946.
- [44] Hao Du, Peter Henry, Xiaofeng Ren, Marvin Cheng, Dan B. Goldman, Steven M. Seitz, and Dieter Fox. 2011. Interactive 3D modeling of indoor environments with a consumer depth camera. In *Proceedings of the 13th international conference on Ubiquitous computing*, 75–84.
- [45] John Ahmet Erkoyuncu, Iñigo Fernández del Amo, Michela Dalle Mura, Rajkumar Roy, and Gino Dini. 2017. Improving efficiency of industrial maintenance with context aware adaptive authoring in augmented reality. *CIRP Ann.* 66, 1 (January 2017), 465–468. DOI:<https://doi.org/10.1016/j.cirp.2017.04.006>
- [46] Cass Everitt and Mark J. Kilgard. 2003. Practical and Robust Stenciled Shadow Volumes for Hardware-Accelerated Rendering. *arXiv:cs/0301002* (January 2003). Retrieved September 1, 2020 from <http://arxiv.org/abs/cs/0301002>
- [47] Steven Feiner, Blair MacIntyre, and Dorée Seligmann. 1992. Annotating the Real World with Knowledge-based Graphics on a See-through Head-mounted Display. In *Proceedings of the Conference on Graphics Interface '92*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 78–85. Retrieved September 14, 2018 from <http://dl.acm.org/citation.cfm?id=155294.155304>
- [48] Steven Feiner, Blair Macintyre, and Dorée Seligmann. 1993. Knowledge-based augmented reality. *Commun. ACM* 36, 7 (1993), 53–62.
- [49] J. Fischer, D. Bartz, and W. Straber. 2005. Stylized augmented reality for improved immersion. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, 195–202. DOI:<https://doi.org/10.1109/VR.2005.1492774>
- [50] Jan Fischer, Dirk Bartz, and Wolfgang Straßer. 2005. Artistic Reality: Fast Brush Stroke Stylization for Augmented Reality. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '05)*, ACM, New York, NY, USA, 155–158. DOI:<https://doi.org/10.1145/1101616.1101649>

- [51] Scott S. Fisher, Micheal McGreevy, James Humphries, and Warren Robinett. 1987. Virtual environment display system. In *Proceedings of the 1986 workshop on Interactive 3D graphics*, ACM, 77–87.
- [52] Felipe Gabaldón. 2010. *Gran Vía (Madrid)*. Retrieved September 13, 2018 from [https://www.flickr.com/photos/felipe\\_gabaldon/4789306741/](https://www.flickr.com/photos/felipe_gabaldon/4789306741/)
- [53] Andreas Geiger, Julius Ziegler, and Christoph Stiller. 2011. Stereoscan: Dense 3d reconstruction in real-time. In *2011 IEEE intelligent vehicles symposium (IV)*, IEEE, 963–968.
- [54] Paul Grimm, Michael Haller, Volker Paelke, Silvan Reinhold, Christian Reimann, and R. Zauner. 2002. AMIRE-authoring mixed reality. In *Augmented Reality Toolkit, The First IEEE International Workshop*, IEEE, 2–pp.
- [55] Raju Gupta. 2019. Fiducial Marker-Based Motion Capture System for Biomechanics and Lower Limb Prosthesis Alignment. (2019).
- [56] S. Guven and S. Feiner. 2003. Authoring 3D hypermedia for wearable augmented and virtual reality. *IEEE*, 118–126. DOI:<https://doi.org/10.1109/ISWC.2003.1241401>
- [57] Daniel Conrad Halbert. 1984. Programming by example. University of California, Berkeley. Retrieved September 11, 2017 from <https://danhalt.org/pbe-html.htm>
- [58] Michael Haller, Stephan Drab, and Werner Hartmann. 2003. A Real-time Shadow Approach for an Augmented Reality Application Using Shadow Volumes. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '03)*, ACM, New York, NY, USA, 56–65. DOI:<https://doi.org/10.1145/1008653.1008665>
- [59] Matthias Haringer and Holger T. Regenbrecht. 2002. A pragmatic approach to augmented reality authoring. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, 237. Retrieved February 24, 2017 from <http://dl.acm.org/citation.cfm?id=854991>
- [60] Sandra G. Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Sage Publications Sage CA: Los Angeles, CA, 904–908.
- [61] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Elsevier, 139–183.
- [62] Morton L. Heilig. 1960. Stereoscopic-television apparatus for individual use. Retrieved September 14, 2018 from <https://patents.google.com/patent/US2955156A/en>
- [63] Alan R. Hevner. 2007. A three cycle view of design science research. *Scand. J. Inf. Syst.* 19, 2 (2007), 4.
- [64] Alan R. Hevner, Salvatore T. March, Jinsoo Park, and Sudha Ram. 2004. Design Science in Information Systems Research. *MIS Q* 28, 1 (March 2004), 75–105.
- [65] Hannah Higgins. 2009. *The grid book*. MIT Press.
- [66] Andrew Hillstead. 2017. Simulating Psychedelic Therapy Through Mediated Reality.
- [67] Jeroen D. Hol, T. B. Schon, Fredrik Gustafsson, and Per J. Slycke. 2006. Sensor fusion for augmented reality. In *Information Fusion, 2006 9th International Conference on*, IEEE, 1–6.
- [68] Hossein Mousavi Hondori, Maryam Khademi, Lucy Dodakian, Steven C. Cramer, and Cristina Videira Lopes. 2013. A spatial augmented reality rehab system for post-stroke hand rehabilitation. In *MMVR*, 279–285. Retrieved September 12, 2017 from <http://ai2-s2-pdfs.s3.amazonaws.com/e183/7e399d3b8ecb2d8e7d13078c4a8c2c51743b.pdf>
- [69] Daniel Ichbiah and Susan L. Knepper. 1991. *Making of Microsoft; How Bill Gates and His Team Created the World's Most Successful Software...* Prima Communications, Inc.
- [70] IKEA. 2017. *IKEA Store*. IKEA. Retrieved September 18, 2017 from <https://play.google.com/store/apps/details?id=com.ikea.kompis&hl=en>
- [71] Hiroshi Ishii, Eran Ben-Joseph, John Underkoffler, Luke Yeung, Dan Chak, Zahra Kanji, and Ben Piper. 2002. Augmented urban planning workbench: overlaying drawings, physical

- models and digital simulation. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, 203.
- [72] B. Jiang, U. Neumann, and Suya You. 2004. A robust hybrid tracking system for outdoor augmented reality. In *IEEE Virtual Reality 2004*, 3–275. DOI:<https://doi.org/10.1109/VR.2004.1310049>
- [73] Simon Peyton Jones, Alan Blackwell, and Margaret Burnett. 2003. A user-centred approach to functions in Excel. *Acm Sigplan Not.* 38, 9 (2003), 165–176.
- [74] MCFD Juan, David Furió, Leila Alem, Peta Ashworth, Juan Cano, and others. 2011. ARGreenet and BasicGreenet: Two mobile games for learning how to recycle. (2011). Retrieved September 13, 2017 from <https://dspace5.zcu.cz/handle/11025/10807>
- [75] Hannes Kaufmann and Dieter Schmalstieg. 2003. Mathematics and geometry education with collaborative augmented reality. *Comput. Graph.* 27, 3 (June 2003), 339–345. DOI:[https://doi.org/10.1016/S0097-8493\(03\)00028-1](https://doi.org/10.1016/S0097-8493(03)00028-1)
- [76] Georg Klein and David Murray. 2007. Parallel tracking and mapping for small AR workspaces. In *Proceedings of the 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, 1–10.
- [77] Christian Knopfle, Jens Weidenhausen, Laurent Chauvigne, and Ingo Stock. 2005. Template based authoring for AR based service scenarios. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.*, IEEE, 237–240.
- [78] K. C. Knowlton. 1977. Computer displays optically superimposed on input devices. *Bell Syst. Tech. J.* 56, 3 (March 1977), 367–383. DOI:<https://doi.org/10.1002/j.1538-7305.1977.tb00514.x>
- [79] Andrew J. Ko, Robin Abraham, Laura Beckwith, Alan Blackwell, Margaret Burnett, Martin Erwig, Chris Scaffidi, Joseph Lawrance, Henry Lieberman, Brad Myers, Mary Beth Rosson, Gregg Rothermel, Mary Shaw, and Susan Wiedenbeck. 2011. The State of the Art in End-user Software Engineering. *ACM Comput Surv* 43, 3 (April 2011), 21:1–21:44. DOI:<https://doi.org/10.1145/1922649.1922658>
- [80] Hoshang Kolivand and Mohd Shahrizal Sunar. 2014. Realistic Real-Time Outdoor Rendering in Augmented Reality. *PLOS ONE* 9, 9 (September 2014), e108334. DOI:<https://doi.org/10.1371/journal.pone.0108334>
- [81] M. Kolsch, R. Bane, T. Hollerer, and M. Turk. 2006. Multimodal interaction with a wearable augmented reality system. *IEEE Comput. Graph. Appl.* 26, 3 (May 2006), 62–71. DOI:<https://doi.org/10.1109/MCG.2006.66>
- [82] Chris D. Kounavis, Anna E. Kasimati, and Efpraxia D. Zamani. 2012. Enhancing the Tourism Experience through Mobile Augmented Reality: Challenges and Prospects. *Int. J. Eng. Bus. Manag.* 4, (January 2012), 10. DOI:<https://doi.org/10.5772/51644>
- [83] Myron W. Krueger. 1977. Responsive environments. In *Proceedings of the June 13-16, 1977, national computer conference*, ACM, 423–433.
- [84] Myron W. Krueger, Thomas Gionfriddo, and Katrin Hinrichsen. 1985. VIDEOPLACE—an artificial reality. In *ACM SIGCHI Bulletin*, ACM, 35–40.
- [85] E. Kruijff, J. E. Swan, and S. Feiner. 2010. Perceptual issues in augmented reality revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, 3–12. DOI:<https://doi.org/10.1109/ISMAR.2010.5643530>
- [86] Ranjit Kumar. 2019. *Research methodology: A step-by-step guide for beginners*. Sage Publications Limited.
- [87] D. Kurz and S. Benhimane. 2011. Gravity-aware handheld Augmented Reality. In *2011 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 111–120. DOI:<https://doi.org/10.1109/ISMAR.2011.6092376>
- [88] Daniel Kurz and Selim Benhimane. 2012. Handheld augmented reality involving gravity measurements. *Comput. Graph.* 36, 7 (2012), 866–883.
- [89] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. 2017. *Research methods in human-computer interaction*. Morgan Kaufmann.

- [90] Gun A. Lee and Gerard J. Kim. 2009. Immersive authoring of tangible augmented reality content: a user study. *J. Vis. Lang. Comput.* 20, 2 (2009), 61–79.
- [91] Gun A. Lee, Claudia Nelles, Mark Billingham, and Gerard Jounghyun Kim. 2004. Immersive Authoring of Tangible Augmented Reality Applications. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '04)*, IEEE Computer Society, Washington, DC, USA, 172–181. DOI:<https://doi.org/10.1109/ISMAR.2004.34>
- [92] James R. Lewis. 1995. IBM computer usability satisfaction questionnaires: psychometric evaluation and instructions for use. *Int. J. Hum.-Comput. Interact.* 7, 1 (1995), 57–78.
- [93] Henry Lieberman, Fabio Paternò, Markus Klann, and Volker Wulf. 2006. End-User Development: An Emerging Paradigm. In *End User Development*, Henry Lieberman, Fabio Paternò and Volker Wulf (eds.). Springer Netherlands, 1–8. Retrieved January 20, 2015 from [http://link.springer.com/chapter/10.1007/1-4020-5386-X\\_1](http://link.springer.com/chapter/10.1007/1-4020-5386-X_1)
- [94] Hugo Liu and Henry Lieberman. 2005. Metafor: visualizing stories as code. In *Proceedings of the 10th international conference on Intelligent user interfaces*, ACM, 305–307. Retrieved September 12, 2017 from <http://dl.acm.org/citation.cfm?id=1040908>
- [95] Mark A. Livingston, Lawrence J. Rosenblum, Simon J. Julier, Dennis Brown, Yohan Baillot, I. I. Swan, Joseph L. Gabbard, Deborah Hix, and others. 2002. *An augmented reality system for military operations in urban terrain*. NAVAL RESEARCH LAB WASHINGTON DC ADVANCED INFORMATION TECHNOLOGY BRANCH. Retrieved September 12, 2017 from <http://www.dtic.mil/docs/citations/ADA499032>
- [96] David G. Lowe. 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* 60, 2 (2004), 91–110.
- [97] Blair Macintyre, Maribeth G, Steven Dow, and Jay David Bolter. 2004. DART: A Toolkit for Rapid Design Exploration of Augmented Reality Experiences. In *ACM Symp. on User Interface Software and Technology (UIST'04)*, 197–206.
- [98] Blair Macintyre, Maribeth Gandy, J. Bolter, S. Dow, and B. Hannigan. 2003. DART: the Designer's Augmented Reality Toolkit. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings*, 329–330. DOI:<https://doi.org/10.1109/ISMAR.2003.1240744>
- [99] Lester Madden. 2011. *Professional Augmented Reality Browsers for Smartphones: Programming for junaio, Layar and Wikitude*. John Wiley & Sons.
- [100] Hearst Magazines. 1955. Popular Mechanics. *Windshield TV Screen To Aid Blind Flying*, 100.
- [101] Steve Mann. 2002. Mediated reality with implementations for everyday life. *Presence Connect* 1, (2002).
- [102] Mark Billingham. 01:20:25 UTC. A Survey of Augmented Reality. Retrieved September 14, 2018 from <https://www.slideshare.net/marknb00/a-survey-of-augmented-reality>
- [103] Richard Mattessich and Giuseppe Galassi. 2000. History of the spreadsheet: From matrix accounting to budget simulation and computerization. In *Accounting and history: a selection of papers presented at the 8th World Congress of Accounting Historians: Madrid-Spain, 19-21 July 2000*, Asociación Española de Contabilidad y Administración de Empresas, AECA, 203–232.
- [104] Michael D. McCool. 2000. Shadow volume reconstruction from depth maps. *ACM Trans. Graph. TOG* 19, 1 (2000), 1–26.
- [105] Alison McMahan. 2003. Immersion, engagement and presence. *Video Game Theory Read.* 67, (2003), 86.
- [106] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE Trans. Inf. Syst.* 77, 12 (1994), 1321–1329.
- [107] Joe J. Monaghan. 2005. Smoothed particle hydrodynamics. *Rep. Prog. Phys.* 68, 8 (2005), 1703.

- [108] Álvaro Montero, Telmo Zarraonandia, Paloma Díaz, and Ignacio Aedo. 2016. Creating Interactive and Realistic Augmented Reality Experiences. In *Proceedings of the XVII International Conference on Human Computer Interaction*, ACM, 17. Retrieved February 24, 2017 from <http://dl.acm.org/citation.cfm?id=2998660>
- [109] Álvaro Montero, Telmo Zarraonandia, Paloma Díaz, and Ignacio Aedo. 2017. An End-User Tool for Creating Augmented Reality Experiences. *-EUD 2017* (2017), 63.
- [110] Alvaro Montero, Telmo Zarraonandia, Paloma Diaz, and Ignacio Aedo. 2017. Designing and implementing interactive and realistic augmented reality experiences. *Univers. Access Inf. Soc.* (November 2017). DOI:<https://doi.org/10.1007/s10209-017-0584-2>
- [111] Raúl Mur-Artal, J. M. M. Montiel, and Juan D. Tardós. 2015. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Trans. Robot.* 31, 5 (October 2015), 1147–1163. DOI:<https://doi.org/10.1109/TRO.2015.2463671>
- [112] Raul Mur-Artal and Juan D. Tardós. 2017. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE Trans. Robot.* 33, 5 (2017), 1255–1262.
- [113] Joseph Newman, Alexander Bornik, Daniel Pustka, Florian Echtler, Manuel Huber, Dieter Schmalstieg, and Gudrun Klinker. 2007. Tracking for distributed mixed reality environments. In *Workshop on Trends and Issues in Tracking for Virtual Environments at the IEEE Virtual Reality Conference (VR'07)*, Citeseer.
- [114] Stéphane Nicolau, Luc Soler, Didier Mutter, and Jacques Marescaux. 2011. Augmented reality in laparoscopic surgical oncology. *Surg. Oncol. Oxf.* 20, 3 (2011), 189–201. DOI:<http://dx.doi.org/10.1016/j.suronc.2011.07.002>
- [115] Edwin B. Olson and Seth Teller. 2008. Robust and efficient robotic mapping. (2008).
- [116] Charles Owen, Arthur Tang, and Fan Xiao. 2003. ImageTclAR: A blended script and compiled code development system for augmented reality. In *Proceedings of the International Workshop on Software Technology for Augmented Reality Systems*, Citeseer, 537–544. Retrieved April 7, 2016 from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.3.5081&rep=rep1&type=pdf>
- [117] Richard Perez, Robert Seals, and Joseph Michalsky. 1993. All-weather model for sky luminance distribution—preliminary configuration and validation. *Sol. Energy* 50, 3 (1993), 235–245.
- [118] Bui Tuong Phong. 1975. Illumination for computer generated pictures. *Commun. ACM* 18, 6 (1975), 311–317.
- [119] Wayne Piekarski Piekarski and Bruce H. Thomas. 2001. Tinmith-evo5—An Architecture for Supporting Mobile Augmented Reality Environments. In *Augmented Reality, International Symposium on*, IEEE Computer Society, 177–177. Retrieved January 30, 2015 from <http://www.computer.org/csdl/proceedings/isar/2001/1375/00/13750177.pdf>
- [120] Wayne Piekarski and Bruce Thomas. 2002. ARQuake: The Outdoor Augmented Reality Gaming System. *Commun ACM* 45, 1 (January 2002), 36–38. DOI:<https://doi.org/10.1145/502269.502291>
- [121] Sören Pirk, Till Niese, Torsten Hädrich, Bedrich Benes, and Oliver Deussen. 2014. Windy trees: computing stress response for developmental tree models. *ACM Trans. Graph. TOG* 33, 6 (2014), 1–11.
- [122] Jarkko Polvi, Takafumi Taketomi, Goshiro Yamamoto, Arindam Dey, Christian Sandor, and Hirokazu Kato. 2016. SlidAR: A 3D positioning method for SLAM-based handheld augmented reality. *Comput. Graph.* 55, (2016), 33–43.
- [123] Arcot J. Preetham, Peter Shirley, and Brian Smits. 1999. A practical analytic model for daylight. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 91–100.
- [124] Jussi Rämö and Vesa Välimäki. 2012. Digital augmented reality audio headset. *J. Electr. Comput. Eng.* 2012, (2012).



- [125] J. Rekimoto. 1998. Matrix: a realtime object identification and registration method for augmented reality. In *Proceedings. 3rd Asia Pacific Computer Human Interaction (Cat. No.98EX110)*, 63–68. DOI:<https://doi.org/10.1109/APCHI.1998.704151>
- [126] Jun Rekimoto, Yuji Ayatsuka, and Kazuteru Hayashi. 1998. Augment-able reality: Situated communication through physical and digital spaces. In *Wearable Computers, 1998. Digest of Papers. Second International Symposium on*, IEEE, 68–75.
- [127] Jun Rekimoto and Katashi Nagao. 1995. The world through the computer: Computer augmented interaction with real world environments. In *Proceedings of the 8th annual ACM symposium on User interface and software technology*, ACM, 29–36.
- [128] Iain E. Richardson. 2011. *The H. 264 advanced video compression standard*. John Wiley & Sons.
- [129] R. Roberto, J. P. Lima, T. Araújo, and V. Teichrieb. 2016. Evaluation of Motion Tracking and Depth Sensing Accuracy of the Tango Tablet. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, 231–234. DOI:<https://doi.org/10.1109/ISMAR-Adjunct.2016.0082>
- [130] Marcos A. Rodrigues, Alan Robinson, and Willie Brink. 2008. Fast 3D reconstruction and recognition. *New Asp. Signal Process. Comput. Geom. Artif. Vis. 8th WSEAS ISCGAV Rhodes (2008)*, p15–21.
- [131] Robert Rosenthal and Ralph L. Rosnow. 2008. *Essentials of behavioral research: Methods and data analysis*.
- [132] Edward Rosten and Tom Drummond. 2006. Machine learning for high-speed corner detection. In *European conference on computer vision*, Springer, 430–443.
- [133] Joseph Michael Rozier. 2000. Here&There: an augmented reality system of linked audio. PhD Thesis. Massachusetts Institute of Technology.
- [134] Susana Rubio, Eva Díaz, Jesús Martín, and José M. Puente. 2004. Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Appl. Psychol.* 53, 1 (2004), 61–86.
- [135] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary R. Bradski. 2011. ORB: An efficient alternative to SIFT or SURF. In *ICCV*, Citeseer, 2.
- [136] Andrés Santos-Torres, Telmo Zarranandia, Paloma Díaz, Teresa Onorati, and Ignacio Aedo. 2020. An empirical comparison of interaction styles for map interfaces in immersive virtual environments. *Multimed. Tools Appl.* (2020), 1–22.
- [137] Yoshinobu Sato, Masahiko Nakamoto, Yasuhiro Tamaki, Toshihiko Sasama, Isao Sakita, Yoshikazu Nakajima, Morito Monden, and Shinichi Tamura. 1998. Image guidance of breast cancer surgery using 3-D ultrasound images and augmented reality visualization. *IEEE Trans. Med. Imaging* 17, 5 (1998), 681–693.
- [138] Christopher Scaffidi, Mary Shaw, and Brad Myers. 2005. Estimating the numbers of end users and end user programmers. In *Visual Languages and Human-Centric Computing, 2005 IEEE Symposium on*, IEEE, 207–214. Retrieved September 5, 2017 from <http://ieeexplore.ieee.org/abstract/document/1509505/>
- [139] Bernd Schwald and Blandine De Laval. 2003. An augmented reality system for training and assistance to maintenance in the industrial context. (2003). Retrieved September 18, 2017 from <https://dspace5.zcu.cz/handle/11025/1662>
- [140] Hartmut Seichter, Julian Looser, and Mark Billinghurst. 2008. ComposAR: An intuitive tool for authoring AR applications. In *Proceedings of the 7th IEEE/ACM international symposium on mixed and augmented reality*, IEEE Computer Society, 177–178. Retrieved April 7, 2016 from <http://dl.acm.org/citation.cfm?id=1605325>
- [141] Jinwook Shim, Minje Kong, Yoonsik Yang, Jonghoon Seo, and Tack-Don Han. 2014. Interactive features based augmented reality authoring tool. In *Consumer Electronics (ICCE), 2014 IEEE International Conference on*, IEEE, 47–50. Retrieved February 24, 2017 from <http://ieeexplore.ieee.org/abstract/document/6775902/>
- [142] Mel Slater. 2003. A note on presence terminology. *Presence Connect* 3, 3 (2003), 1–5.

- [143] Jaka Sodnik, Saso Tomazic, Raphael Grasset, Andreas Duenser, and Mark Billinghurst. 2006. Spatial sound localization in an augmented reality environment. In *Proceedings of the 18th Australia conference on computer-human interaction: design: activities, artefacts and environments*, ACM, 111–118. Retrieved January 3, 2017 from <http://dl.acm.org/citation.cfm?id=1228197>
- [144] Austin N. Stanton. 1962. Headgear mounted cathode ray tube and binocular viewing device. Retrieved September 14, 2018 from <https://patents.google.com/patent/US3059519A/en>
- [145] Andrei State, Gentaro Hirota, David T. Chen, William F. Garrett, and Mark A. Livingston. 1996. Superior Augmented Reality Registration by Integrating Landmark Tracking and Magnetic Tracking. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96)*, ACM, New York, NY, USA, 429–438. DOI:<https://doi.org/10.1145/237170.237282>
- [146] stefan\_b. 03:05:32 UTC. Advanced Lighting for Interactive Applications. Retrieved September 1, 2020 from [https://www.slideshare.net/stefan\\_b/advanced-lighting-for-interactive-applications](https://www.slideshare.net/stefan_b/advanced-lighting-for-interactive-applications)
- [147] Bruno Steux and Oussama El Hamzaoui. 2010. tinySLAM: A SLAM algorithm in less than 200 lines C-language program. In *2010 11th International Conference on Control Automation Robotics & Vision*, IEEE, 1975–1979.
- [148] Ivan E. Sutherland. 1965. The ultimate display. *Multimed. Wagner Virtual Real.* (1965), 506–508.
- [149] Ivan E. Sutherland. 1968. A head-mounted three dimensional display. In *Proceedings of the December 9-11, 1968, fall joint computer conference, part I*, ACM, 757–764. Retrieved June 10, 2017 from <http://dl.acm.org/citation.cfm?id=1476686>
- [150] Z. Szalavári, D. Schmalstieg, A. Fuhrmann, and M. Gervautz. 1998. “Studierstube”: An environment for collaboration in augmented reality. *Virtual Real.* 3, 1 (March 1998), 37–48. DOI:<https://doi.org/10.1007/BF01409796>
- [151] Nobuko Taketa, Kenichi Hayashi, Hirokazu Kato, and Shogo Noshida. 2007. Virtual Pop-Up Book Based on Augmented Reality. In *Human Interface and the Management of Information. Interacting in Information Environments* (Lecture Notes in Computer Science), Springer Berlin Heidelberg, 475–484.
- [152] Keisuke Tateno, Federico Tombari, Iro Laina, and Nassir Navab. 2017. Cnn-slam: Real-time dense monocular slam with learned depth prediction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 6243–6252.
- [153] Unity Technologies. Unity Real-Time Development Platform | 3D, 2D VR & AR Visualizations. Retrieved February 18, 2020 from <https://unity.com/>
- [154] Miiikka Tikander, Matti Karjalainen, and Ville Riikonen. 2008. An augmented reality audio headset. In *Proc. of the 11th Int. Conf. on Digital Audio Effects (DAFx-08), Espoo, Finland.*
- [155] Miiikka Tikander and V. Riikonen. 2005. Sound quality of an augmented reality audio headset. In *Proc. of the 8th Int. Conf. on Digital Audio Effects (DAFx-05), Madrid, Spain.*
- [156] Martha Tsapatori, Jorge A. Soler Díaz, Nessa O’Connor, and Luis Salgado. 2005. The ORION Project: A European Union Thematic Network. *Marq Arqueol. Mus.* (2005), 103–112.
- [157] Christiane Ulbricht and Dieter Schmalstieg. 2003. *Tangible augmented reality for computer games*. Citeseer.
- [158] Blase Ur, Elyse McManus, Melwyn Pak Yong Ho, and Michael L. Littman. 2014. Practical trigger-action programming in the smart home. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 803–812. Retrieved September 11, 2017 from <http://dl.acm.org/citation.cfm?id=2557420>
- [159] Vassilios Vlahakis, Nikolaos Ioannidis, John Karigiannis, Manolis Tsotros, Michael Gounaris, Didier Stricker, Tim Gleue, Patrick Daehne, and Luís Almeida. 2002.

- Archeoguide: an augmented reality guide for archaeological sites. *IEEE Comput. Graph. Appl.* 5 (2002), 52–60.
- [160] Daniel Wagner, Gerhard Reitmayr, Alessandro Mulloni, Tom Drummond, and Dieter Schmalstieg. 2009. Real-time detection and tracking for augmented reality on mobile phones. *IEEE Trans. Vis. Comput. Graph.* 16, 3 (2009), 355–368.
- [161] Yang Wang and Dimitris Samaras. 2003. Estimation of multiple directional light sources for synthesis of augmented reality images. *Graph. Models* 65, 4 (2003), 185–205.
- [162] Turner Whitted. 2005. An improved illumination model for shaded display. In *ACM Siggraph 2005 Courses*. 4–es.
- [163] Praetor alpha at English Wikipedia. 2005. *English: praetor alpha*. Retrieved September 1, 2020 from <https://commons.wikimedia.org/wiki/File:7fin.png>
- [164] Lance Williams. 1978. Casting curved shadows on curved surfaces. In *Proceedings of the 5th annual conference on Computer graphics and interactive techniques*, 270–274.
- [165] Yoonsik Yang, Jiwook Shim, Seungho Chae, and Tack-Don Han. 2016. Mobile augmented reality authoring tool. In *Semantic Computing (ICSC), 2016 IEEE Tenth International Conference on*, IEEE, 358–361. Retrieved February 24, 2017 from <http://ieeexplore.ieee.org/abstract/document/7439364/>
- [166] Telmo Zarraonandia, Ignacio Aedo, Paloma Díaz, and Alvaro Montero. 2013. An augmented lecture feedback system to support learner and teacher communication. *Br. J. Educ. Technol.* 44, 4 (July 2013), 616–628. DOI:<https://doi.org/10.1111/bjet.12047>
- [167] Telmo Zarraonandia, Ignacio Aedo, Paloma Díaz, and Alvaro Montero Montes. 2014. Augmented Presentations: Supporting the Communication in Presentations by Means of Augmented Reality. *Int. J. Hum.-Comput. Interact.* 30, 10 (October 2014), 829–838. DOI:<https://doi.org/10.1080/10447318.2014.927283>
- [168] Telmo Zarraonandia, Paloma Diaz, and Ignacio Aedo. 2016. Using combinatorial creativity to support end-user design of digital games. *Multimed. Tools Appl.* (2016), 1–26.
- [169] Jürgen Zauner, Michael Haller, Alexander Brandl, and Werner Hartmann. 2003. Authoring of a mixed reality assembly instructor for hierarchical structures. In *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, IEEE Computer Society, 237.
- [170] ZhiYing Zhou, Adrian David Cheok, Yan Qiu, and Xubo Yang. 2007. The role of 3-D sound in human reaction and performance in augmented reality environments. *IEEE Trans. Syst. Man Cybern.-Part Syst. Hum.* 37, 2 (2007), 262–272.
- [171] Zhiying Zhou, Adrian David Cheok, Xubo Yang, and Yan Qiu. 2004. An experimental study on the role of 3D sound in augmented reality environment. *Interact. Comput.* 16, 6 (2004), 1043–1068.
- [172] 2012. *parranormal AR Invaders.MOV*. Retrieved August 17, 2020 from [https://www.youtube.com/watch?v=aIO2-jV5ZBY&feature=emb\\_logo](https://www.youtube.com/watch?v=aIO2-jV5ZBY&feature=emb_logo)
- [173] 2017. Don't Miss the Importance of Diminished Reality. *Loup Ventures*. Retrieved August 16, 2020 from <https://loupventures.com/dont-miss-the-importance-of-diminished-reality/>
- [174] 2017. Tango | Google Developers. Retrieved October 17, 2019 from <https://web.archive.org/web/20170815231050/developers.google.com/tango/>
- [175] Pokémon Go Brings Augmented Reality to a Mass Audience - The New York Times. Retrieved September 12, 2017 from [https://www.nytimes.com/2016/07/12/technology/pokemon-go-brings-augmented-reality-to-a-mass-audience.html?\\_r=0](https://www.nytimes.com/2016/07/12/technology/pokemon-go-brings-augmented-reality-to-a-mass-audience.html?_r=0)
- [176] Zombies Everywhere! Augmented Reality Apocalypse (Halloween Edition) on the App Store. *App Store*. Retrieved September 18, 2017 from <https://itunes.apple.com/us/app/zombies-everywhere-augmented-reality-apocalypse-halloween/id530292213?mt=8>

- [177] EyePet™. *Playstation*. Retrieved September 18, 2017 from <https://www.playstation.com/en-us/games/eyepet-ps3/>
- [178] Nintendo AR Cards. Retrieved September 18, 2017 from <http://www.nintendo.com/3ds/ar-cards/>
- [179] The leader in Mixed Reality Technology | HoloLens. Retrieved September 18, 2017 from <https://www.microsoft.com/en-us/hololens>
- [180] Create or run a macro. Retrieved September 11, 2017 from <https://support.office.com/en-us/article/Create-or-run-a-macro-c6b99036-905c-49a6-818a-dfb98b7c3c9c>
- [181] Blockly. *Google Developers*. Retrieved September 11, 2017 from <https://developers.google.com/blockly/>
- [182] Skype | Communication tool for free calls and chat. Retrieved February 3, 2020 from <https://www.skype.com/en//>
- [183] Hubs by Mozilla. Retrieved February 3, 2020 from /
- [184] Facebook Horizon. Retrieved February 3, 2020 from <https://www.oculus.com/facebookhorizon/>
- [185] Google Developers. *Google Developers*. Retrieved May 31, 2015 from <https://developers.google.com/glass/design/index>
- [186] ARToolKit Home Page. *ARToolkit*. Retrieved from <http://www.hitl.washington.edu/artoolkit/>
- [187] Authoring Mixed Reality | AMIRE Project | FP5 | CORDIS | European Commission. Retrieved October 1, 2019 from <https://cordis.europa.eu/project/rcn/61830/factsheet/en>
- [188] ATOMIC Authoring Tool download | SourceForge.net. Retrieved April 25, 2016 from <https://sourceforge.net/projects/atomic-project/>
- [189] Invizimals™. *Playstation*. Retrieved October 1, 2019 from <https://www.playstation.com/es-es/games/invizimals-psp/>
- [190] Wikitude - The World's leading Augmented Reality SDK. Retrieved April 25, 2016 from <http://www.wikitude.com/>
- [191] STAR 1200XLD Vuzix. *Vuzix*. Retrieved from [http://www.vuzix.com/augmented-reality/products\\_star1200xld/](http://www.vuzix.com/augmented-reality/products_star1200xld/)
- [192] Home | Magic Leap. Retrieved October 20, 2019 from <https://www.magicleap.com/>
- [193] ARKit - Apple Developer. Retrieved April 14, 2019 from <https://developer.apple.com/arkit/>
- [194] ARCore de Google - Aplicaciones en Google Play. Retrieved April 14, 2019 from <https://play.google.com/store/apps/details?id=com.google.ar.core&hl=es>
- [195] Digital Binocular Station | Digital Binocular Station information. Retrieved September 14, 2018 from <http://digitalbinocularstation.com/>
- [196] Augmented reality technology brings artefacts to life | infonews.co.nz New Zealand News. Retrieved August 7, 2020 from <https://www.infonews.co.nz/news.cfm?id=42664>
- [197] Wonderbook™: El Libro de los Hechizos. *Playstation*. Retrieved October 23, 2019 from <https://www.playstation.com/es-es/games/wonderbook-book-of-spells-ps3/>
- [198] PlayStation Move motion controller. *Playstation*. Retrieved October 23, 2019 from <https://www.playstation.com/en-gb/explore/accessories/playstation-move-motion-controller/>
- [199] Features2D + Homography to find a known object — OpenCV 2.4.13.7 documentation. Retrieved October 23, 2019 from [https://docs.opencv.org/2.4/doc/tutorials/features2d/feature\\_homography/feature\\_homography.html#feature-homography](https://docs.opencv.org/2.4/doc/tutorials/features2d/feature_homography/feature_homography.html#feature-homography)
- [200] Vuforia | Augmented Reality. Retrieved February 27, 2017 from <https://www.vuforia.com/>

- [201] Augmented Reality / 3D Tracking | Real-time 3D model-based tracking using edges. Retrieved October 23, 2019 from <http://virtual.vtt.fi/virtual/proj2/multimedia/videos/2018/007-Edge-tracking.html>
- [202] AR Invaders Augmented Reality iPhone App | AR Invaders iPhone Game by Soulbit7. Retrieved August 17, 2020 from <http://invaders.soulbit7.com/>
- [203] blender.org - Home of the Blender project - Free and Open 3D Creation Software. *Blender*. Retrieved from <http://www.blender.org/>
- [204] Unreal Engine | The most powerful real-time 3D creation platform. Retrieved February 18, 2020 from <https://www.unrealengine.com/en-US/>
- [205] Ray Casting - online presentation. Retrieved September 1, 2020 from [//en.ppt-online.org/703109](http://en.ppt-online.org/703109)
- [206] Ray Tracing, Your Questions Answered: Types of Ray Tracing, Performance On GeForce GPUs, and More. Retrieved September 1, 2020 from <https://www.nvidia.com/en-us/geforce/news/geforce-gtx-dxr-ray-tracing-available-now/>
- [207] File:Shadow volume illustration.png - Wikimedia Commons. Retrieved September 1, 2020 from [https://commons.wikimedia.org/wiki/File:Shadow\\_volume\\_illustration.png](https://commons.wikimedia.org/wiki/File:Shadow_volume_illustration.png)
- [208] OpenAL: Cross Platform 3D Audio. Retrieved February 24, 2020 from <https://www.openal.org/>
- [209] FMOD. Retrieved February 24, 2020 from <https://www.fmod.com/>
- [210] *Pokemon GO and Interactive Dynamic Video*. Retrieved February 12, 2020 from [https://www.youtube.com/watch?time\\_continue=79&v=9f1fCCb3hVg&feature=emb\\_logo](https://www.youtube.com/watch?time_continue=79&v=9f1fCCb3hVg&feature=emb_logo)
- [211] Open Constructor for Tango - Apps en Google Play. Retrieved April 2, 2020 from [https://play.google.com/store/apps/details?id=com.lvonasek.openconstructor&hl=es\\_419](https://play.google.com/store/apps/details?id=com.lvonasek.openconstructor&hl=es_419)
- [212] Constructor Developer Tool - Apps en Google Play. Retrieved April 2, 2020 from [https://play.google.com/store/apps/details?id=com.projecttango.constructor&hl=es\\_419](https://play.google.com/store/apps/details?id=com.projecttango.constructor&hl=es_419)
- [213] Agisoft Metashape. Retrieved May 12, 2020 from <https://www.agisoft.com/>

# ANEXO A



uc3m | Universidad Carlos III de Madrid

## Formulario de Consentimiento para un Experimento sobre Herramientas de Autoría de Creación de Experiencias de Realidad Aumentada

### Propósito y descripción del experimento

El propósito de este experimento se centra en estudiar los beneficios y las limitaciones del uso de diferentes enfoques para la creación de experiencias de Realidad Aumentada. El experimento consistirá en la realización de una serie de tareas haciendo uso de dos enfoques/herramientas. Antes, durante y después del experimento, se le pedirá que rellene unos cuestionarios.

### Confidencialidad

Durante el experimento sus movimientos serán monitorizados y guardados para su posterior procesamiento. Además, podría ser tomadas algunas capturas de pantalla de los dispositivos que utilizará. De cualquier manera, todos estos datos junto a las encuestas serán almacenados y cifrados/protegidos con una contraseña para protegerlos y codificados con el fin de proteger su anonimidad en cualquier documento o presentación resultante de este trabajo.

### Registro del consentimiento

Al firmar este documento, indica que ha comprendido la información al respecto del consentimiento y consiente su participación. La participación es de carácter voluntario y le permite negarse a responder ciertas preguntas del cuestionario o retirarse del estudio en cualquier momento sin ningún tipo de penalización. Usted recibirá una copia del formulario de consentimiento.

Nombre del participante: \_\_\_\_\_

Firma del participante: \_\_\_\_\_ Fecha: \_\_\_\_\_

Firma del investigador: \_\_\_\_\_ Fecha: \_\_\_\_\_

Si tiene alguna duda o pregunta al respecto, por favor contacte con el investigador:

Álvaro Montero Montes  
Departamento de Informática, Universidad Carlos III de Madrid  
ammontes@inf.uc3m.es

## ANEXO B

	T01VE	T02AE	T03VE	T04AE	T05VS	T06AS	T07VE	E08AX	E09VS	T10AS
<b>Edad</b>	21	21	37	21	24	20	21	30	27	26
<b>Genero</b>	Hombre	Hombre	Hombre	Hombre	Hombre	Mujer	Hombre	Mujer	Hombre	Mujer
<b>Estudios/Profesión/Campo de Investigación</b>	Ingeniería Informática	Ingeniería Informática	Desarrollador	Student	Msc student	Estudiante de Ingeniería Biomédica	Estudiante de Ingeniería Informática	Grado Ingeniería Informática	Responsable de Almacén	Doble Graduada
<b>1.- ¿Tienes experiencia con los videojuegos (Play Station, Xbox, Nintendo, Android, iOS, PC...)?</b>	Tengo mucha experiencia	Tengo mucha experiencia	Tengo bastante experiencia	Tengo bastante experiencia	Tengo experiencia	Tengo poca experiencia	Tengo mucha experiencia	Tengo muy poca experiencia	Tengo mucha experiencia	Tengo experiencia
<b>2.- ¿Tienes experiencia con el uso de Realidad Virtual?</b>	Tengo experiencia	Tengo poca experiencia	Tengo experiencia	Tengo mucha experiencia	Tengo experiencia	No tengo experiencia	Tengo poca experiencia	Tengo muy poca experiencia	Tengo muy poca experiencia	Tengo poca experiencia
<b>3.- ¿Tienes experiencia en el desarrollo de Realidad Virtual?</b>	Tengo experiencia	No tengo experiencia	Tengo poca experiencia	Tengo bastante experiencia	No tengo experiencia	No tengo experiencia	No tengo experiencia	No tengo experiencia	No tengo experiencia	No tengo experiencia
<b>4.- ¿Tienes experiencia con el uso de Realidad Aumentada?</b>	Tengo bastante experiencia	Tengo poca experiencia	Tengo experiencia	Tengo mucha experiencia	Tengo bastante experiencia	No tengo experiencia	Tengo poca experiencia	Tengo poca experiencia	Tengo experiencia	No tengo experiencia
<b>5.- ¿Tienes experiencia en el desarrollo de Realidad Aumentada?</b>	Tengo bastante experiencia	No tengo experiencia	Tengo experiencia	Tengo bastante experiencia	No tengo experiencia	No tengo experiencia	No tengo experiencia	No tengo experiencia	No tengo experiencia	No tengo experiencia

		T01VE	T02AE	T03VE	T04AE	T05VS	T06AS	T07VE	E08AX	E09VS	T10AS
<b>NASA TLX</b>	¿Qué esfuerzo mental has tenido que realizar?	VR 40	50	20	15	5	40	60	45	20	40
		AR 20	30	35	35	10	35	60	15	20	15
	¿Cuánto esfuerzo físico fue necesario?	VR 20	15	20	5	0	55	45	30	10	25
		AR 45	20	55	35	30	10	60	10	25	20
	¿Cuánto tiempo percibiste que fue necesario para llevar a cabo todas las tareas?	VR 45	40	0	30	15	35	40	35	35	40
		AR 45	35	40	35	20	40	70	40	30	20
	¿Qué nivel de éxito crees que has conseguido al realizar las tareas?	VR 50	25	0	5	15	35	10	30	0	55
		AR 25	15	10	30	5	35	15	25	0	40
	¿Cuánto esfuerzo (mental y físico) has realizado para llevar a cabo las tareas?	VR 40	25	15	30	5	55	60	35	10	45
		AR 30	25	40	25	10	35	70	20	25	20
	¿Qué grado de inseguridad, desánimo, irritación, estrés o molestia has experimentado al realizar las tareas?	VR 45	35	0	65	10	35	0	30	0	45
		AR 5	60	10	25	30	50	10	40	20	60
<b>Valoración del Usuario</b>	Q1 ¿Cómo de satisfecho estás con la forma de Seleccionar Objetos?	VR 5	7	7	6	7	4	7	7	6	6
		AR 5	7	7	4	7	7	6	6	4	4
	Q2 ¿Cómo de satisfecho estás con la forma de Mover Objetos?	VR 6	6	7	3	7	6	7	6	6	6
		AR 3	5	6	4	7	6	5	3	6	5
	Q3 ¿Cómo de satisfecho estás con la forma de Rotar Objetos?	VR 6	5	5	6	6	7	6	6	7	6
		AR 6	7	5	6	6	7	4	7	5	1
	Q4 ¿Cómo de satisfecho estás con la forma de Escalar Objetos?	VR 3	5	3	5	7	7	6	7	7	6
		AR 3	5	4	6	7	7	6	6	7	5
	Q5 ¿Cómo de difícil te ha parecido usar el sistema?	VR 3	3	2	5	1	3	1	2	2	1
		AR 3	3	3	2	3	3	4	2	3	2
	Q6 ¿Cómo de preciso te ha parecido el sistema?	VR 2	4	6	3	7	1	6	6	6	6
		AR 4	6	7	5	7	4	5	5	6	1
	Q7 ¿Cómo de fácil te ha resultado aprender a usar el sistema?	VR 7	6	6	4	7	7	7	6	7	6
		AR 5	6	5	6	7	5	7	6	7	6

		T01VE	T02AE	T03VE	T04AE	T05VS	T06AS	T07VE	E08AX	E09VS	T10AS
<b>¿Qué sistema elegirías si tuvieras en cuenta...</b>	la satisfacción?	6	2	2	7	3	2	1	5	3	1
	la precisión?	4	5	1	1	2	7	1	5	3	1
	la usabilidad?	5	6	3	5	6	2		5	4	1
	la curva de aprendizaje (el tiempo para aprender a usarlo)?	6	5	2	5	7	1	4	3	4	1
	la velocidad?	5	6	2	5	7	4	1	4	3	1
	el tiempo de uso?	5	5	1	5	7	3	3	4	3	2
	la visión general del entorno?	6	1	1	7	5	2	5	7	6	2
	colocar objetos de gran tamaño?	2	2	1	3	2	1	1	4	2	7
	colocar objetos de pequeño tamaño?	3	4	5	5	6	7	3	4	4	1
	tu preferencia?	7	3	1	6	5	3	1	5	1	1
mejores resultados?	4	6	1	6	5	7	1	4	6	1	

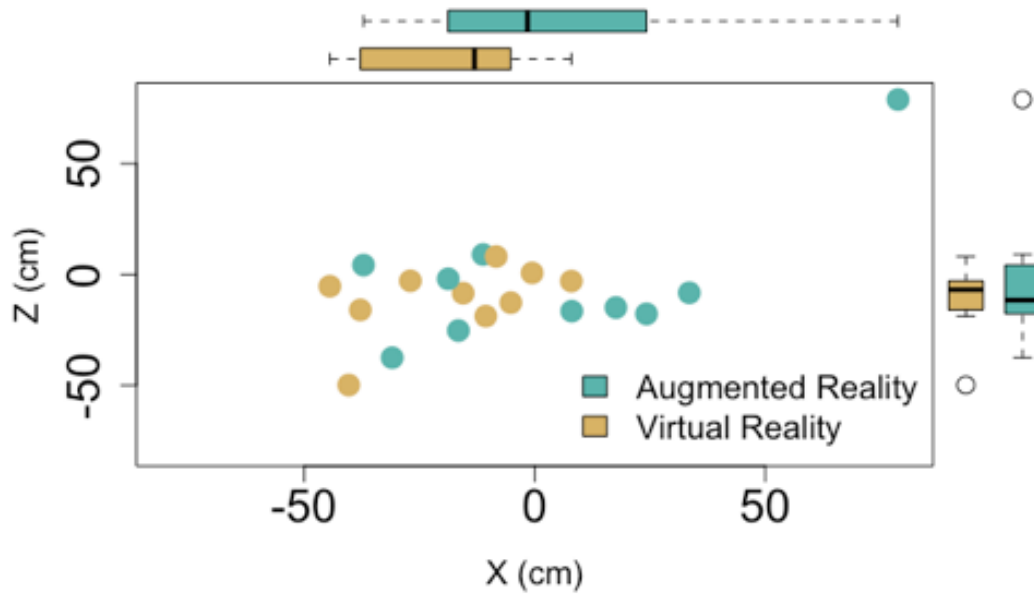
Comentarios adicionales	
T01VE	El unico problema relativamente grande que he notado es la sensacion de movimiento con la herramienta de VR, puede marear un poco hasta que te acostumbras y el tiempo de uso no es lo suficientemente largo como para acostumbrarse rápido.
T02AE	El sistema de Realidad Virtual es más satisfactorio para el usuario ya que los mandos y el estar viendo el entorno constantemente devuelven mejor sensación que hacerlo a través de un móvil que carece de feedback táctil (que sí tienen los mandos de la Realidad virtual). Por otro lado, yo no he tenido mucha experiencia usando dispositivos de Realidad Virtual y cuesta mucho acostumbrarse a desplazarse por un entorno que a la vista se puede percibir real pero sin desplazarte realmente, inconscientemente me inclinaba hacia delante al ver las mesas y las paredes moverse hacia mí. Supongo que este problema desaparecería al coger más experiencia con este tipo de dispositivos.
T03VE	el gesto de escalar es un poco difícil da aprender al principio. En la realidad virtual, el gesto de ejemplo tiene que estar mas cerca de la vision del usuario
T04AE	nunca
T05VS	When you use the AR system and you edit the virtual objects, you can only edit one time. I'd like to have the possibility to scale it two or more times instead of pressing the edit button twice. Confirm button would be great instead
T06AS	Ha habido unos pequeños problemas técnicos pero la experiencia en general ha sido muy buena. La realidad virtual es más divertida e interactiva pero personalmente me he mareado un par de veces, los resultados con las gafas han sido menos precisos pero su uso ha sido más fácil. Muchas gracias!! :D
T07VE	En el sistema de rotar objetos en realidad virtual, en los carteles sería mas intuitivo usar el giro de muñeca en el "Eje Roll"
E08AX	Al principio me ha costado entender como mover los objetos.
E09VS	A la realidad aumentada parece que le costaba más reaccionar a los intentos de cambiar el tamaño de los objetos. En la realidad virtual perdía realismo el no chocar contra la mesa o paredes y la forma de girar la cámara con el joystick.
T10AS	No me ha gustado la forma de rotar los objetos en la realidad aumentada, me ha parecido muy poco preciso.



## ANEXO C

Las siguientes gráficas muestran las ubicaciones, rotaciones y factores de escala de los elementos virtuales colocados en la segunda tarea del estudio.

### Distancia de X y Z en centímetros



### Rotación y Factor de Escala

