



UNIVERSIDAD NACIONAL DEL SUR

TESIS DE MAGISTER EN CIENCIAS DE LA COMPUTACIÓN

Interacción Humano Computadora en Ambientes Virtuales

Matias Nicolás Selzer

BAHÍA BLANCA

ARGENTINA

2018



UNIVERSIDAD NACIONAL DEL SUR

TESIS DE MAGISTER EN CIENCIAS DE LA COMPUTACIÓN

Interacción Humano Computadora en Ambientes Virtuales

Matias Nicolás Selzer

BAHÍA BLANCA

ARGENTINA

2018

Prefacio

Esta Tesis se presenta como parte de los requisitos para optar al grado Académico de Magíster en Ciencias de la Computación, de la Universidad Nacional del Sur y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otra. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el ámbito del Departamento de Ciencias e Ingeniería de la Computación durante el período comprendido entre el 14/07/2015 y el 27/08/2017, bajo la dirección del Dr. Martín Leonardo Larrea.

Matias Nicolás Selzer



UNIVERSIDAD NACIONAL DEL SUR
Secretaría General de Posgrado y Educación Continua

La presente tesis ha sido aprobada el / / , mereciendo la calificación de (.....)

Resumen

La Interacción Humano Computadora es la rama de las Ciencias de la Computación que busca conectar a los humanos con los sistemas computacionales de forma natural e intuitiva. Los humanos interactúan con la realidad a través de sus sentidos y la Realidad Virtual conecta a los humanos con los mundos virtuales.

Gracias a los avances tecnológicos de las últimas décadas, cientos de dispositivos e interfaces han surgido para intentar lograr esta conexión. Sin embargo, la mayoría de estos dispositivos fracasan al no estar diseñados en torno a la usabilidad. Aunque existen libros y trabajos en la literatura que realizan un relevamiento sobre los distintos dispositivos e interfaces utilizados en Realidad Virtual, ninguno lo hace desde el punto de vista de la Interacción Humano Computadora.

Esta tesis se enfoca en los sistemas de Realidad Virtual existentes desde el punto de vista de la Interacción Humano Computadora y la usabilidad. Específicamente, esta tesis presenta un relevamiento exhaustivo de las técnicas y dispositivos existentes para interconectar al humano a través de sus sentidos con la Realidad Virtual.

Las discusiones y conclusiones presentadas en cada capítulo analizan las características, ventajas y desventajas de aquellas técnicas y dispositivos asociados a cada sentido. A partir de esto, esta tesis no solo sirve como guía a la hora de seleccionar el dispositivo más adecuado para ser utilizado en una aplicación de Realidad Virtual particular, sino que también, presenta las características a tener en cuenta a la hora de diseñar y desarrollar nuevos dispositivos.

Abstract

Human Computer Interaction is the area of Computer Science which main goal is to connect humans with computer systems in a natural and intuitive way. Humans interact with reality through their senses and Virtual Reality connects humans with virtual worlds.

Thanks to the technological advances of the last decades, hundreds of devices and interfaces have emerged to try to achieve this connection. However, most of these devices fail because they are not designed based on usability. Although there are books and works in the literature that perform a survey on the different devices and interfaces used in Virtual Reality, none do so from the point of view of Human Computer Interaction.

This thesis focuses on the existing Virtual Reality systems from the point of view of Human Computer Interaction and usability. Specifically, this thesis presents a survey of the existing techniques and devices to interconnect humans with the Virtual Reality through their senses.

The discussions and conclusions presented in each chapter analyze the features, advantages and disadvantages of those techniques and devices associated to each sense. Hereby, this thesis not only provides a guide for selecting the most suitable device to be used in a particular Virtual Reality application, but also presents the characteristics to be taken into account when designing and developing new devices.

Índice general

1.	INTRODUCCIÓN	1
1.	Motivación	3
2.	Contribución de la Tesis	4
3.	Estructura de la Tesis	5
2.	EL MODELO DE INTERACCIÓN DE REALIDAD VIRTUAL	7
1.	Introducción	7
2.	Percepción de la Realidad	8
3.	La Interfaz Perfecta	11
4.	Interfaces de Usuario (IU)	13
5.	El Modelo de Entrada-Salida	15
6.	Conclusión	19
3.	VISIÓN EN REALIDAD VIRTUAL	21
1.	Introducción	22
2.	Visión Humana y Señales de Profundidad	24
2.1.	Señales de Profundidad Monoscópicas	24
2.2.	Señales de Profundidad Estereoscópicas	25
2.2.1.	Señales de Profundidad en Movimiento	26
2.2.2.	Señales de Profundidad Fisiológicas	26
3.	Propiedades de los Displays Visuales	26
3.1.	Propiedades Visuales	27
3.2.	Propiedades Logísticas	35

4.	Tipos de Displays Visuales	39
4.1.	Displays Estacionarios con Monitores	39
4.2.	Displays de RV Proyectados	40
4.3.	Displays para la Cabeza (HBDs)	42
4.4.	Otros Tipos de Display	43
5.	Revisión de la Literatura	46
5.1.	Displays Estacionarios con Monitores	46
5.2.	Displays Proyectados	48
5.3.	Displays Para la Cabeza (HBDs)	51
6.	Discusión y Conclusiones	54
6.1.	Propiedades de la Visión y Señales de Profundidad . . .	54
6.2.	Tipos de Display	55
4.	AUDIO EN REALIDAD VIRTUAL	59
1.	Introducción	59
2.	Audición Humana	62
2.1.	Presencia e Inmersión	62
2.2.	Intensidad	62
2.3.	Enmascaramiento y Atención	63
2.4.	Tono y Timbre	64
2.5.	Resolución Temporal	65
3.	Audición Espacial	65
3.1.	Señales Binaurales	67
3.2.	Señales de Distancia Anecóicas	69
3.3.	Reverberación o Eco	69
3.4.	Señales Dinámicas	70
3.5.	Experiencia	70
3.6.	Adaptación	72
3.7.	Información Inter-Sensorial	72
4.	Simulación Espacial	74
4.1.	Propiedades de los Dispositivos de Audio	74
4.1.1.	Contaminación Acústica o Ruido	75
4.1.2.	Movilidad de los Usuarios	75

4.2.	Simulación Utilizando Auriculares	76
4.2.1.	Displays Dióticos	77
4.2.2.	Displays Dicóticos	77
4.2.3.	Funciones de Transferencias en Relación a la Cabeza	78
4.3.	Simulación Utilizando Altavoces	81
4.3.1.	Display No Espacial	82
4.3.2.	Display Estéreo	83
4.3.3.	Altavoces Multicanal	84
5.	Revisión de la Literatura	85
6.	Discusión y Conclusiones	91
6.1.	Percepción y Estado Previo	91
6.2.	Diseño y Software	92
6.3.	Hardware	93
6.4.	Relación con los otros Sentidos	95
6.5.	Conclusión	95
5.	OLFATO Y GUSTO EN REALIDAD VIRTUAL	97
1.	Introducción	98
2.	Olfato Humano	100
2.1.	Sistema Psicofísico Olfativo	100
2.2.	Percepción del Olor	100
3.	Tipos de Display	101
3.1.	Displays Olfativos Ubicados en el Ambiente	101
3.2.	Displays Olfativos Usables	102
4.	Revisión de la Literatura	104
5.	Discusión y Conclusiones	116
6.	TACTO EN REALIDAD VIRTUAL	119
1.	Introducción	120
2.	Sistema Táctil Humano	121
3.	Tipos de Displays	125
3.1.	Clasificación Según la Referencia Física	125

3.1.1.	Dispositivos Hápticos Referenciados al Mundo	126
3.1.2.	Dispositivos Hápticos referenciados al cuerpo	126
3.2.	Clasificación Según el Estímulo	127
3.2.1.	Dispositivos Hápticos con Feedback Activo	127
3.3.	Dispositivos Hápticos Pasivos	131
4.	Revisión de la Literatura	133
5.	Discusión y Conclusiones	145
5.1.	Cuerpo Humano y Desafíos de la Tecnología	145
5.2.	Propiedades Físicas	147
5.3.	Seguridad	148
5.4.	Logística	149
5.5.	Conclusiones	150
7.	TRACKING EN REALIDAD VIRTUAL	153
1.	Introducción	154
2.	Tipos de Tracking	156
2.1.	Tracking Magnético	156
2.2.	Tracking Mecánico	158
2.3.	Tracking Óptico	159
2.4.	Tracking Videométrico	162
2.5.	Tracking Ultrasónico	165
2.6.	Tracking Inercial	166
2.7.	Tracking Neural	168
2.8.	Tracking Híbrido	169
3.	Tracking del Cuerpo	171
3.1.	Postura Corporal y Gestos	171
3.2.	Tracking de la Cabeza	172
3.3.	Tracking del Torso	173
3.4.	Manos y Dedos	174
3.5.	Tracking de Ojos	176
3.6.	Reconocimiento de Voz	177
3.7.	Otros	179
4.	Dispositivos de Tracking	181

4.1.	Controles y Props	181
4.2.	Plataformas	183
5.	Revisión de la Literatura	185
5.1.	Tipos de <i>tracking</i>	185
5.2.	Tracking del Cuerpo	194
5.3.	Dispositivos	197
6.	Discusión y Conclusiones	213
6.1.	Tipos de <i>Tracking</i>	213
6.2.	Tracking del Cuerpo	215
6.3.	Dispositivos	217
6.4.	Conclusión	218
8.	CONCLUSIONES	221
1.	Características de Hardware	223
2.	Logística	225
3.	Presencia e Inmersión	226
4.	Multi-modalidad	228
5.	Publicaciones	229
6.	Trabajo a Futuro	231
	BIBLIOGRAFÍA	280

1

Introducción

La Realidad Virtual (RV) es una rama de las Ciencias de la Computación que desde sus orígenes ha cautivado la mente de todas aquellas personas que han podido experimentarla. La realidad puede ser definida como todo aquello que es percibido por nuestros sentidos. Una realidad virtual es una forma de realidad distinta a la realidad que experimentamos todos los días. Más específicamente, es una realidad sintética, creada por computadora, que podemos percibir a través de nuestros sentidos.

Las investigaciones en RV avanzaron muy lentamente desde sus orígenes debido principalmente a la complejidad y al costo de los dispositivos. Por este motivo, hasta hace pocos años, los sistemas de RV solo podían verse en instituciones especializadas, en su mayoría como sistemas de entrenamiento militar en países del primer mundo. Actualmente los avances tecnológicos llevaron a sistemas más económicos, compactos y cómodos, permitiendo que la RV llegase a los hogares de todos aquellos que dispongan de un teléfono móvil. Es más, incluso los sistemas de RV más avanzados, que permiten interacciones con más de uno o dos sentidos al mismo tiempo, son actualmente accesibles para el público general.

Estamos entrando en una era en la que la RV podría llegar a ser parte de nuestra vida cotidiana. Una interfaz perfecta debería permitirnos ver, oír y sentir el mundo virtual como si estuviésemos en el mundo real. Sin embargo, esto no es así en la práctica.

Nuestras interfaces sensoriales, como nuestros ojos u oídos, no son perfectos, y agregan ruido a nuestra percepción del mundo virtual. Aunque la tecnología actual no nos permita alcanzar un nivel de inmersión completo, se pueden desarrollar aproximaciones muy reales.

Todo el tiempo se está desarrollando una nueva interfaz para conectarnos con la RV a través de nuestros sentidos. Sin embargo, la mayoría de éstos no se crea desde el punto de vista de la usabilidad. La Interacción Humano Computadora (IHC) es la rama de las Ciencias de la Computación que busca maximizar la usabilidad de las interfaces que conectan a los usuarios y a los sistemas de computación, generando interfaces lo más naturales posibles.

1. MOTIVACIÓN

Para poder realizar nuevas interfaces que permitan conectar a los usuarios con los sistemas de RV de la forma más natural y realista posible, es fundamental trabajar con los principios de la IHC. Aunque la tecnología siga avanzando y el hardware sea más pequeño y económico, nunca se podrá desarrollar un sistema de RV exitoso si no se dispone de los conocimientos adecuados. Por un lado, es importante conocer cómo se comporta el cuerpo humano y especialmente sus sentidos; y por el otro, es fundamental conocer las tecnologías actuales y cómo éstas hacen de intermediario entre el usuario y la computadora. Aunque existen libros y trabajos científicos que tratan estos temas en profundidad, ninguno de ellos analiza los sistemas de RV desde el punto de vista de cada uno de los sentidos y la IHC.

2. CONTRIBUCIÓN DE LA TESIS

El objetivo principal de este trabajo es el estudio y recopilación de los sistemas de RV que estimulan los sentidos del usuario, proveyendo una guía a seguir, acorde con los principios de la IHC, a la hora de querer desarrollar un nuevo sistema de RV, o querer seleccionar un sistema existente para alguna aplicación particular. Específicamente, la contribución de esta tesis se enfoca en:

- Realizar un estudio sobre el funcionamiento de cada uno de los sentidos del cuerpo humano para así comprender mejor cómo éstos perciben la realidad y por ende, la RV.
- Realizar un relevamiento exhaustivo del estado del arte sobre los sistemas de RV existentes, enfocado en cada uno de los sentidos.
- Realizar un análisis sobre las distintas ventajas y desventajas de cada sistema o técnica relevado.

3. ESTRUCTURA DE LA TESIS

Específicamente, la estructura de la tesis está organizada de la siguiente manera.

- El Capítulo 2 presenta el concepto de RV, haciendo hincapié en la percepción del humano dentro de la realidad. Además presenta el modelo de entrada-salida que servirá de fundamento base para el resto de la tesis.
- El Capítulo 3 presenta los conceptos relevantes para comprender cómo el sentido de la vista influye en el desarrollo de los sistemas de RV. Es el sentido más estudiado en cuanto a RV se refiere, dado que se presenta una amplia Sección sobre todos los tipos de displays existentes y un extenso relevamiento sobre los trabajos existentes en la literatura. Finalmente se analiza este relevamiento, teniendo en cuenta los tipos de displays y las aplicaciones particulares.
- El audio es uno de los sentidos más inmersivos y de los más fáciles de estimular. El Capítulo 4 presenta todas aquellas señales auditivas que componen el sonido, y cómo éste debe simularse correctamente para ser presentado en sistemas de RV. Además, se presentan los distintos tipos de displays auditivos existentes y las distintas técnicas para implementar el sonido virtual. La revisión de la literatura se enfoca en aquellos trabajos que utilizan audio en RV, ya sea para mejorar la presentación del sonido, o para aplicaciones particulares.
- El Capítulo 5 presenta cómo los sentidos del gusto y el olfato influyen en los sistemas de RV. Estos sentidos son los que aún presentan más incógnitas y los menos estudiados, debido en gran parte a la complejidad que requiere el poder sintetizar un aroma o un gusto particular. Asimismo, existen algunos trabajos que intentan simular estos sentidos, principalmente mediante la utilización de esencias. Este Capítulo presenta además un relevamiento sobre los trabajos existentes y una discusión al respecto.

- El Capítulo 6 presenta la relación entre el sentido del tacto y los sistemas de RV. A lo largo del Capítulo se hace una distinción sobre la parte tangible y la parte háptica, haciendo referencia a la información que “sale” y “entra” del usuario, respectivamente. Luego de analizar las características del sistema táctil humano, se presenta un relevamiento de la literatura y una discusión al respecto.
- Los Capítulos anteriores presentan las formas en que los usuarios perciben el mundo virtual a través de sus sentidos. El Capítulo 7 analiza el otro lado del sistema, es decir, aquellas formas en que el sistema de RV percibe el cuerpo del usuario. En este Capítulo se analizan todas las técnicas y tecnologías que se pueden utilizar para trackear al usuario, así como también aquellas partes del cuerpo que pueden ser traqueadas y cómo. Se presenta también un relevamiento de la literatura sobre distintas técnicas y sistemas desarrollados para hacer tracking del usuario, y más específicamente en sistemas de RV. El Capítulo concluye con una discusión sobre los trabajos analizados, presentando ventajas y desventajas de cada tipo de sistema y técnica de tracking.
- Finalmente, el Capítulo 8 concluye con un análisis general de la presente tesis respecto a las características que se consideraron más relevantes. Por último presenta las publicaciones que se llevaron a cabo durante el transcurso del presente Magíster, así como también el trabajo planificado a desarrollarse en el futuro.

2

El Modelo de Interacción de Realidad Virtual

1. INTRODUCCIÓN

En este Capítulo, exploraremos todos aquellos niveles en los que el usuario interactúa con el mundo virtual utilizando el medio de la RV. Nos enfocamos en las formas en las que el usuario se involucra con la experiencia de RV, desde las interfaces físicas que se usan para interactuar, hasta los métodos en que la experiencia se le presenta al usuario.

2. PERCEPCIÓN DE LA REALIDAD

El cerebro humano tiene diferentes zonas que se encargan de procesar la información sobre cada uno de los sentidos. Por ejemplo, la corteza visual procesa información proveniente del nervio óptico para así poder crear una representación del mundo real en base a la percepción visual. De esta manera, el cerebro procesa mucha información sobre el mundo externo. Pero, sin importar qué, no es capaz de sentir el mundo real por sí mismo, ya que no dispone de “órganos sensoriales”.

Imaginemos una persona sin órganos sensoriales viviendo en el mundo real. No hay duda de que esta persona existe en el mundo real, sin embargo, poco se puede decir sobre lo que esta persona siente y sobre sus conocimientos sobre el mundo externo. El cerebro está preparado para procesar mucha información sobre el mundo externo y en base a ésta, poder crear una representación del mismo. Sin embargo, necesita de los órganos sensoriales para transmitir la información necesaria sobre el mundo externo para así crear una percepción del mismo. La Figura 2.1 muestra una representación del cerebro dentro de la realidad.

El “Cerebro en un Frasco” es un escenario utilizado en una variedad de experimentos filosóficos que discuten cuestiones sobre conocimiento, realidad, mente y conciencia, entre otros ([1-3]). En este escenario, un científico loco, máquina, u otra entidad, remueve el cerebro de una persona de su cuerpo, lo coloca en un frasco con nutrientes que lo mantienen vivo y conecta las neuronas con cables a una supercomputadora que puede proveer los impulsos eléctricos idénticos a aquellos que el cerebro normalmente recibe. De esta forma, la computadora simularía la realidad (incluyendo las respuestas apropiadas a la entrada/salida del cerebro) y el cerebro sin cuerpo continuaría teniendo experiencias conscientes perfectamente normales, así como lo haría una persona con cuerpo, pero sin tener ninguna relación con objetos o eventos del mundo real (Figura 2.2).

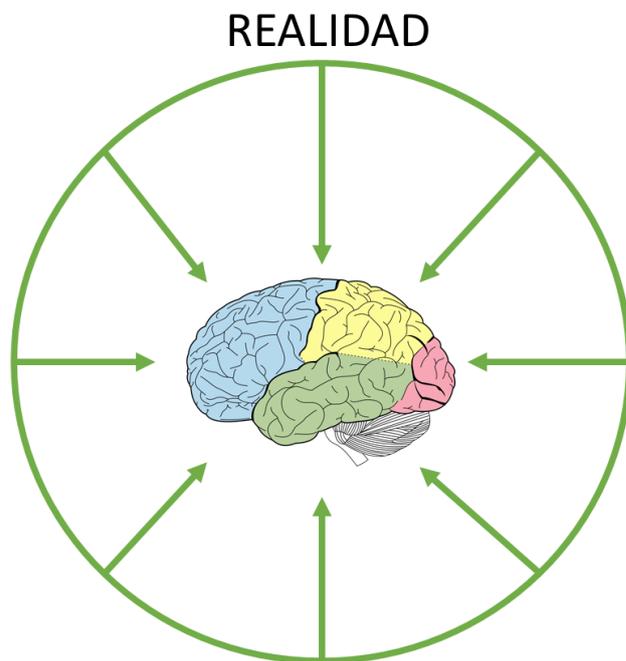


Figura 2.1: Representación del cerebro dentro de la realidad.

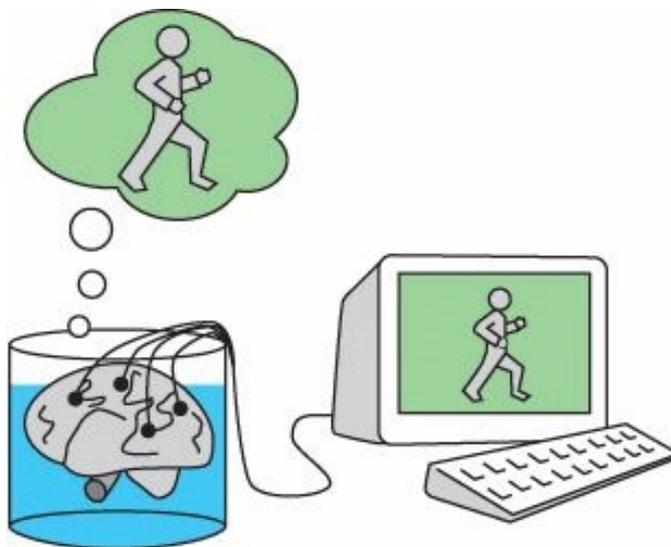


Figura 2.2: Cerebro "pensando" que está corriendo.

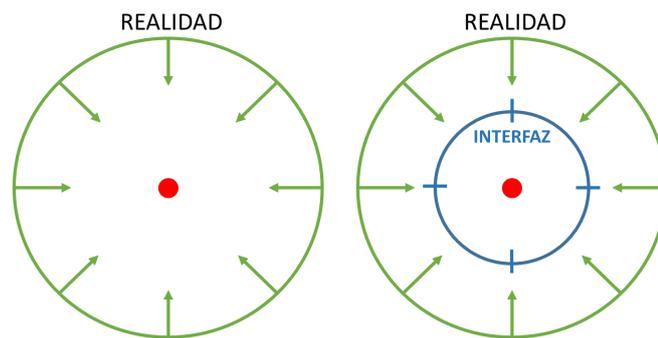


Figura 2.3: CDP dentro de la realidad (izquierda). Una capa de interfaz se encuentra entre el CDP y la realidad para poder percibirla (derecha).

Por un momento imaginemos que el cerebro humano es solo un simple punto en el espacio. Entonces, aunque la realidad exista y el punto exista dentro de la realidad, el punto no es capaz de sentir la realidad debido a que no tiene ningún “órgano sensorial”. Llamemos a este punto el Centro de Percepción (CDP) y podemos verlo en la Figura 2.3. Para que el CDP pueda percibir el mundo externo, necesita una interfaz sensorial que le provea el “hardware” necesario para sentir el mundo externo.

En el caso del ser humano, esta interfaz está definida por los órganos sensoriales, que establecen una conexión entre el cerebro y el mundo externo. Así, el cerebro puede percibir la realidad a través de esta interfaz, representada en la Figura 2.4.

Es importante notar que la realidad percibida a través de los sentidos humanos solo es una aproximación de la “realidad real”. Esta aproximación depende exclusivamente de qué puede ser sentido a través de los sentidos. Existen animales, por ejemplo, que pueden ver u oír un rango mayor al de los humanos. Aunque la realidad sea la misma para todos, dependemos de nuestra interfaz (nuestros sentidos) para determinar “qué tanta realidad” podemos percibir.

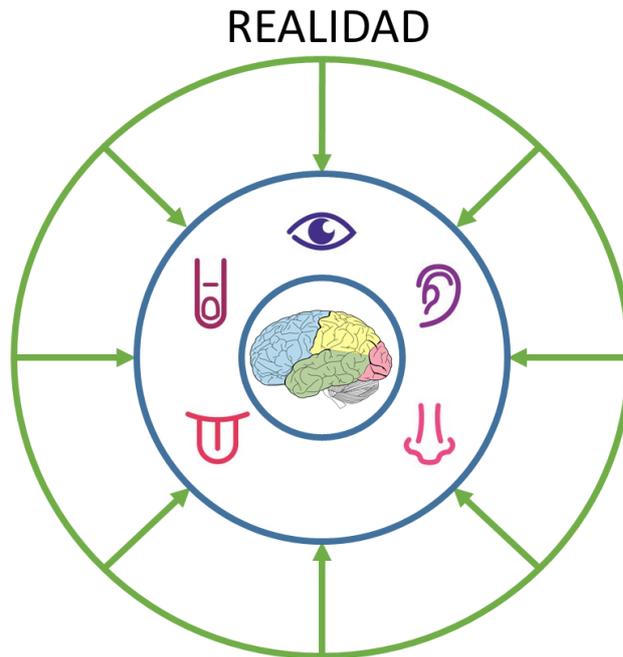


Figura 2.4: Los órganos sensoriales componen la interfaz entre el cerebro y la realidad externa.

3. LA INTERFAZ PERFECTA

El diseño de los mundos virtuales y de los sistemas de RV está afectado por la interfaz de usuario. Varias teorías consideran como el objetivo máximo de la RV el que sea un medio sin interfaces notables, una interfaz completamente ubicua. Esto es, una experiencia de RV debería estar tan bien diseñada que la frontera entre el usuario y el mundo virtual pareciera inexistente. La interfaz de usuario imitaría precisamente la forma en la que el usuario experimenta el mundo real. Esto se considera por muchos como la interfaz perfecta. Debemos tener en cuenta que aunque esta interfaz sea invisible, todavía existe.

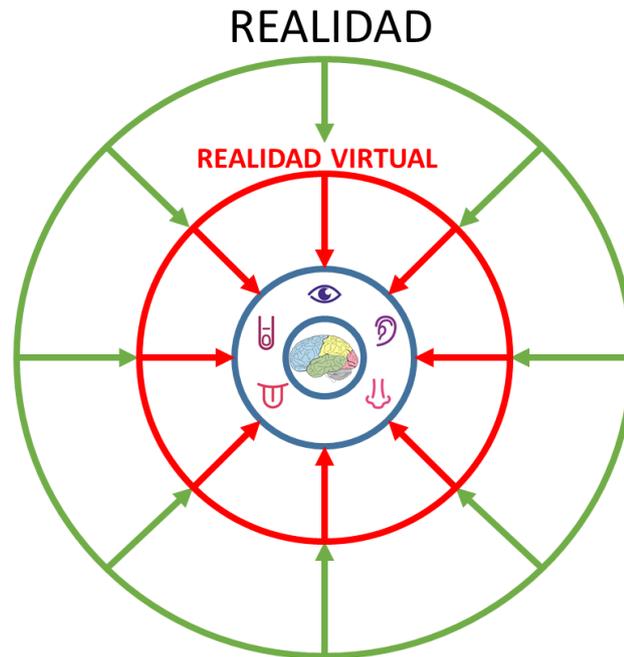


Figura 2.5: La RV presenta una nueva capa entre los organos sensoriales y la realidad.

Como vemos en la Figura 2.5, la RV agrega una nueva capa entre el mundo real y la interfaz humana. La interfaz perfecta debería ser capaz de establecer una comunicación entre el mundo virtual y el cerebro humano de tal forma que el usuario no pueda distinguir el mundo virtual del mundo real. Un ejemplo de este tipo de interfaz se ve representado en la película Matrix, en donde la RV es presentada en los cerebros de las personas desde su nacimiento, por lo que nunca se dan cuenta que están viviendo en un mundo virtual. Esto también representa una forma de “Cerebro en el Frasco”.

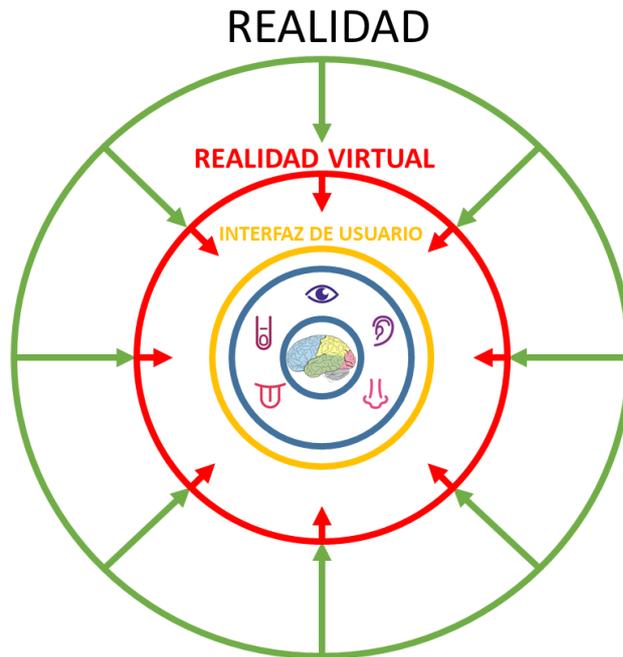


Figura 2.6: Capa de interfaz de usuario entre los órganos sensoriales y la RV.

4. INTERFACES DE USUARIO (IU)

Hoy en día, la interfaz perfecta no puede recrearse con la tecnología actual. Sin embargo, existen muchos dispositivos que permiten una comunicación entre los usuarios y el mundo virtual. Estas interfaces establecen una nueva capa entre el usuario y la RV, ya que conectan la RV con la interfaz humana, es decir, los órganos humanos. Esta capa se muestra en la Figura 2.6. Por un lado, la interfaz de usuario representa la forma en la que los usuarios se comunican con el mundo físico, y por el otro lado, representa cómo el mundo se manifiesta de tal forma que puede ser percibido por los sentidos del usuario.

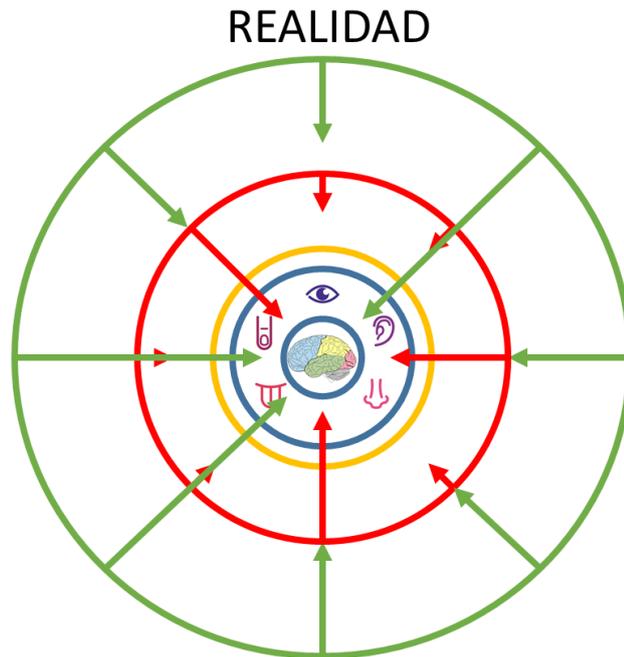


Figura 2.7: Los sentidos no utilizados para RV perciben la realidad y crean un conflicto sensorial.

La IU como una forma de conexión entre el cerebro y el mundo virtual se presenta en la siguiente sección. Sin embargo, es importante saber que si la IU no abarca todos los sentidos, existirá una competencia por la atención del usuario entre la realidad y la RV. Por ejemplo, en un sistema de RV que solo utiliza la visión, el usuario será capaz de ver el mundo virtual, pero al mismo tiempo percibirá el mundo real a través de sus otros sentidos, lo que producirá un conflicto sensorial.

Aunque la RV llegue al cerebro a través de alguno de los sentidos mediante la IU, la realidad también llegará al cerebro a través de los sentidos que no se estén utilizando para RV, mediante los órganos sensoriales. Esto se muestra en la Figura 2.7. En un sistema de este tipo, el usuario podría distinguir fácilmente entre el mundo real y el mundo virtual, cosa que se trata de evitar ya que reduce considerablemente la inmersión del sistema.

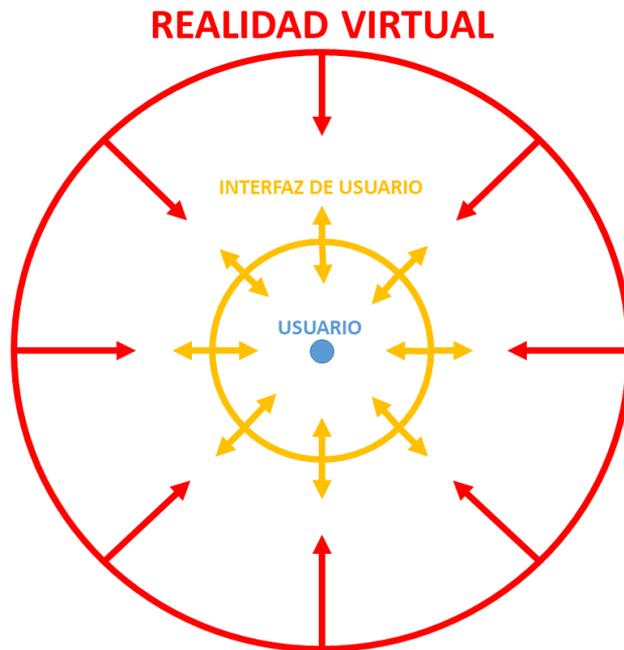


Figura 2.8: La interfaz de usuario entre el cerebro y la realidad virtual es bidireccional.

5. EL MODELO DE ENTRADA-SALIDA

Desde ahora, consideraremos solamente la interacción entre el usuario y el mundo virtual. Mientras el usuario percibe información del sistema de RV, ese sistema también recibe información del usuario. Así, la IU mostrada en la Figura 2.6 puede ser dividida en dos partes: la entrada al sistema de RV (o salida del usuario) y la salida del sistema de RV (o la entrada al usuario). La Figura 2.8 muestra cómo la interacción del usuario trabaja en una capa de comunicación bidireccional. Por conveniencia, usaremos un punto para representar tanto al cerebro humano como al sistema sensorial humano.

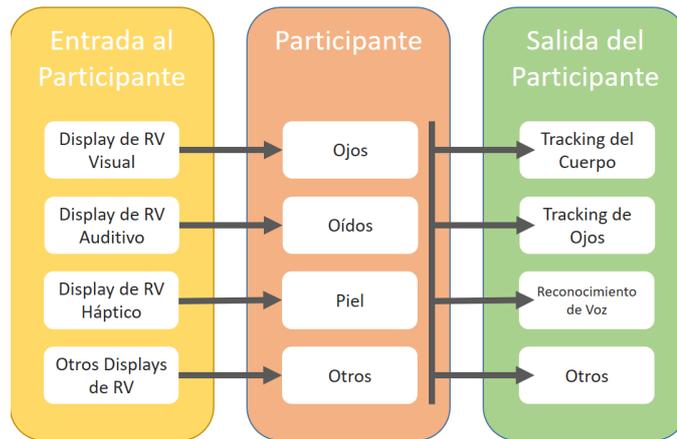


Figura 2.9: Información de entrada-salida desde el punto de vista del participante.

Un participante que está experimentando una aplicación de RV utiliza una IU para comunicarse con el mundo virtual. El participante visualizará y sentirá el mundo virtual a través de interfaces físicas de usuario o dispositivos, como un *Head Mounted Display*¹, auriculares, o guantes hápticos especiales, por ejemplo. Estas interfaces son la entrada con la que el participante adquiere la información proveniente del mundo virtual. Por otro lado, la información del cuerpo del participante es trackeada, de tal forma que la representación del usuario puede ser actualizada de forma precisa. Esa información trackeada representa la información que sale del usuario y que luego arribará como la información de entrada al mundo virtual. Este modelo es representado en la Figura 2.9.

¹ Pantalla Montada en la Cabeza

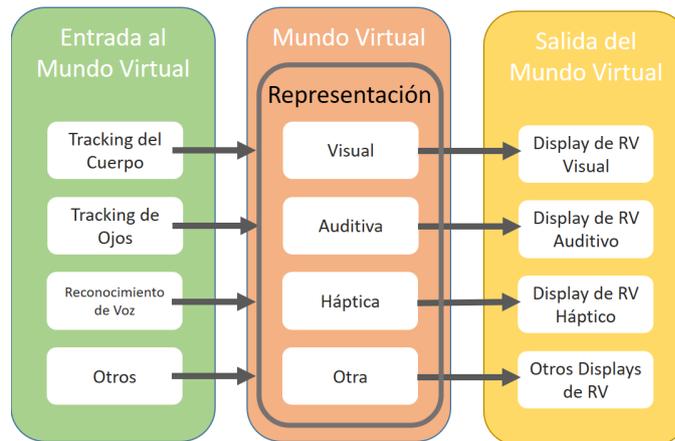


Figura 2.10: Información de entrada-salida desde el punto de vista del sistema de RV.

Por el otro lado, el mundo virtual debe tener alguna clase de representación, un medio por el que será expuesto al mundo. Si hablamos del medio de la RV, un sistema debe proveer una representación visual, una representación auditiva, una representación háptica, entre otras. Esa representación del mundo virtual debe ser mostrada a los participantes de alguna forma. La salida del mundo virtual se conforma por aquellas interfaces que son capaces de presentar dicho mundo virtual a los participantes. Lo que es más, el mundo virtual debe recibir información de los usuario para poder actualizarse acordeamente. El modelo del mundo virtual se ve en la Figura 2.10.

Es importante notar que la salida del mundo virtual es la entrada al participante y viceversa. Así, un modelo unificado puede crearse a partir de esto. La Figura 2.11 muestra los diagramas de la Figura 2.9 y la Figura 2.10 combinados en uno solo. La Figura 2.12 muestra una versión simplificada del modelo.

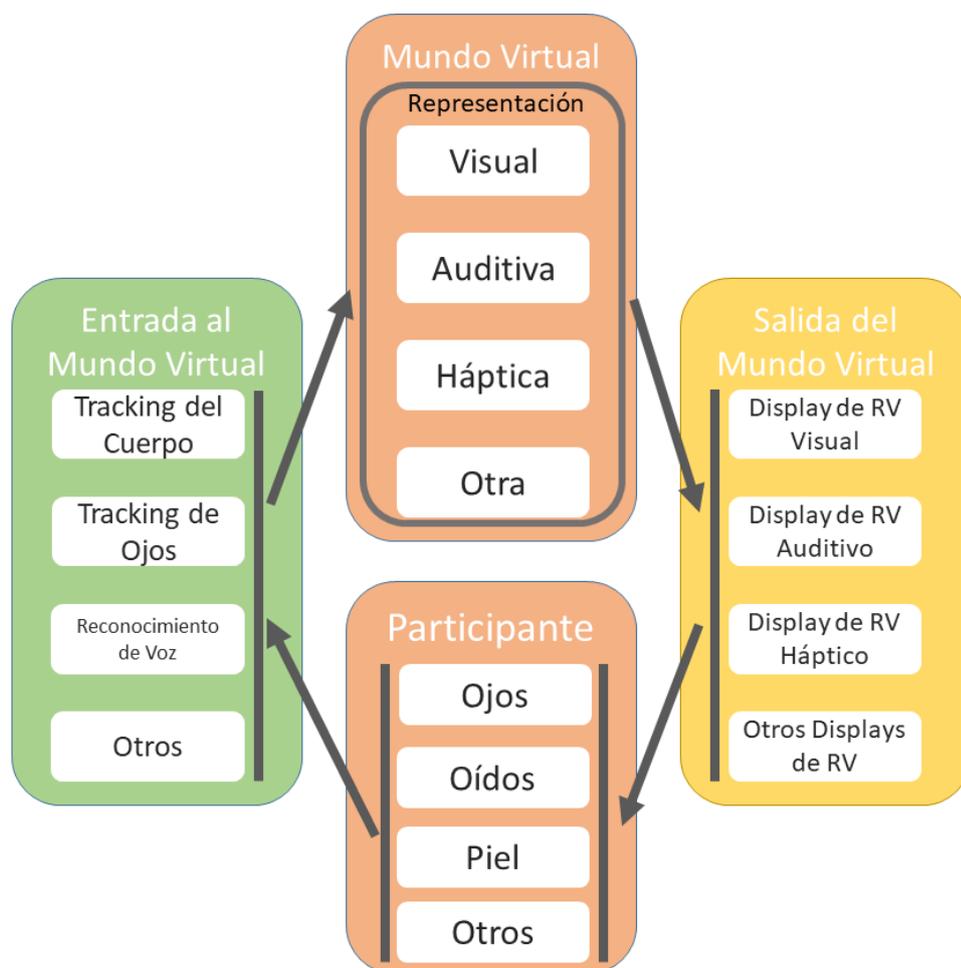


Figura 2.11: Modelo completo de entrada-salida entre el participante y el mundo virtual.

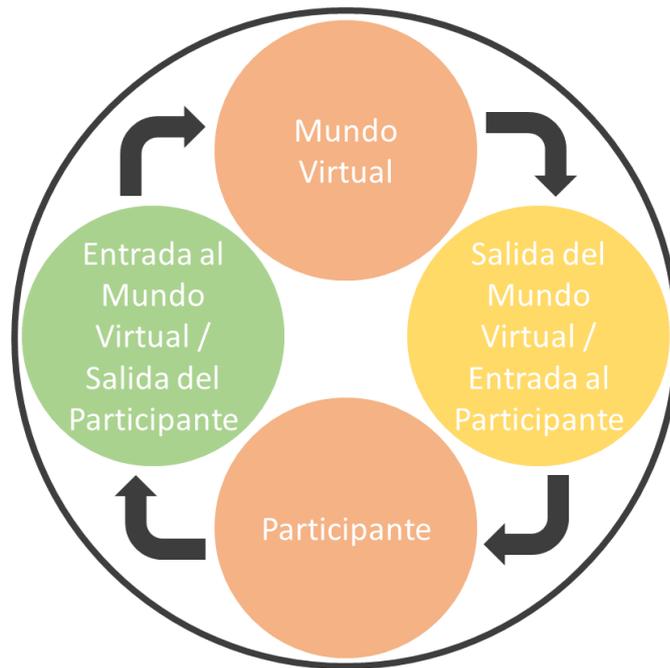


Figura 2.12: Modelo de entrada-salida reducido.

Podemos apreciar de la Figura 2.12 que la comunicación del participante con el mundo virtual forma un ciclo. El mundo virtual, con todas sus leyes y representaciones, debe comunicarse con el usuario para permitirle experimentar el mundo virtual. Al mismo tiempo, el usuario debe comunicarse con el mundo virtual para que éste pueda actualizar su representación según las interacciones del usuario. Ese medio de comunicación compone la IU, la entrada y la salida al mundo virtual.

6. CONCLUSIÓN

Hemos visto que el ser humano percibe la realidad puramente a través de sus sentidos. El objetivo fundamental de la RV es el de “suplantar” la realidad con otra realidad completamente ficticia, generada por computadora, en donde el usuario pueda experimentarla en su totalidad, a través de todos sus sentidos.

La mayoría de los sistemas de RV existentes reproducen contenido virtual que es percibido solo por uno o dos sentidos a la vez. Esto genera conflictos sensoriales ya que dicho usuario seguirá percibiendo la realidad a través de los sentidos restantes. Aunque actualmente la tecnología no nos permita lograr la interfaz perfecta en donde el usuario pueda utilizar todos sus sentidos de forma natural y cómoda, existen muchas interfaces que conectan al usuario con el mundo virtual de forma eficiente.

En los Capítulos siguientes se presenta una breve descripción del funcionamiento de cada uno de los sentidos, haciendo énfasis en aquellas características que deben ser consideradas para el diseño de interfaces de RV. Además, se realiza una revisión de la literatura y un análisis sobre todos aquellos sistemas e interfaces existentes que permiten conectar al usuario con la RV, a través de cada uno de los sentidos.

3

Visión en Realidad Virtual

1. INTRODUCCIÓN

La visión es uno de los sentidos más utilizados por los seres humanos. Con él podemos ver no solo la realidad que está a nuestro alcance, sino también la realidad que se encuentra fuera de nuestro alcance. Sin embargo, nuestros ojos son capaces de percibir solo una porción de todo el espectro electromagnético, que comprende lo que se denomina luz visible.

Es por esto mismo que desde los comienzos de la RV se ha tratado de conseguir una visualización del mundo virtual lo más real e inmersiva posible. Para lograrlo, se deben tener en cuenta todas aquellas características de la percepción visual humana que puedan aplicarse a los sistemas existentes de RV. La Sección 2 presenta un análisis de estas características en relación a los sistemas de RV.

Los avances tecnológicos han posibilitado la creación de dispositivos de hardware que nos permiten visualizar mundos virtuales de manera eficiente e inmersiva. Estos dispositivos constituyen distintos tipos de displays como, por ejemplo, los sistemas de proyección o aquellos displays que van sujetos a la cabeza del usuario.

Cada uno de estos tipos de displays posee sus ventajas y desventajas y su utilización dependerá de cada aplicación particular de RV. No es lo mismo si un sistema de RV está pensado para una o para varias personas, por ejemplo. Los distintos tipos de displays para RV se presentan en la Sección 4, junto a sus características, ventajas y desventajas.

La Sección 5 presenta la revisión de la literatura. Muchos trabajos han investigado la utilización de distintos tipos de displays para distintas aplicaciones de RV. Gracias a esto, los displays han ido evolucionando en sistemas inmersivos, cómodos de utilizar y algunos masivamente comercializados. En esta sección también se hace énfasis en aquellos displays y sistemas que se consideran fuera de lo común, como sistemas de proyección sobre humo o sistemas de proyección sobre la retina del observador.

Finalmente, el Capítulo concluye con una discusión sobre las características de la percepción visual que ayudan a la creación de contenidos de RV más inmersivos. Además, se analizan los tipos de displays investigados, relacionándolos entre sí, y concluyendo las ventajas y desventajas de cada uno para cada ocasión particular.

2. VISIÓN HUMANA Y SEÑALES DE PROFUNDIDAD

La vista es uno de los sentidos más importantes del ser humano. Cada día, los ojos trabajan como una máquina, recibiendo información visual del mundo exterior y enviando dicha información al cerebro para que sea procesada.

Para poder movernos en el mundo de forma segura, debemos poder percibir y visualizar nuestro entorno en 3 dimensiones. La distancia relativa entre los objetos es una de las cosas más importantes que nuestro sistema visual tiene que percibir. Esta información es procesada de varias formas y a los indicadores de distancia se los llama señales de profundidad. Esas señales de profundidad serán presentadas y explicadas a continuación.

2.1. SEÑALES DE PROFUNDIDAD MONOSCÓPICAS

Las señales de profundidad monoscópicas son aquéllas que pueden verse en una vista estática de una escena, como en una fotografía o pintura. Nuestro cerebro analiza estas señales a partir de imágenes planas y trata de determinar las características físicas de los objetos.

La **interposición** es la señal que recibimos cuando un objeto ocluye a otro de nuestra vista. Si un objeto cubre a otro, sabremos que el primero está más cerca.

El **sombreado** es la forma en que el objeto cambia de acuerdo a la iluminación, aportando información acerca de la forma del objeto. Las sombras indican la relación posicional entre dos objetos.

Al comparar el **tamaño de los objetos** respecto a otros objetos del mismo tipo, se puede determinar la distancia relativa entre los objetos. Para aproximar la distancia a la que se encuentran los objetos de nosotros, comparamos su tamaño con nuestro conocimiento previo de objetos similares.

La **perspectiva lineal** se observa cuando las líneas paralelas convergen en un mismo punto de fuga a la distancia. Esta señal se basa en la suposición de que los objetos vistos están hechos de líneas paralelas, como en la mayoría de los edificios, por ejemplo. De esta forma, la posición, orientación y tamaño de este tipo de objetos es estimado por nuestro cerebro.

Los **gradientes de textura** de superficie aparecen cuando nuestra retina no puede discernir el detalle de una textura desde lejos comparada con la misma textura desde cerca. Por ejemplo, si nos encontramos en un descampado cubierto por pasto, el gran detalle de la textura del pasto a nuestros pies cambiará a un verde difuso a la distancia.

El **brillo** provee una señal de profundidad moderada y cuando se ven muchos objetos juntos, los que sean más brillantes parecerán estar más cerca.

Finalmente, existen efectos atmosféricos, como la niebla, que afectan la forma en que los objetos son distinguidos por nuestra vista.

2.2. SEÑALES DE PROFUNDIDAD ESTEREOSCÓPICAS

La estereopsis es la percepción de la profundidad producida por la combinación de los estímulos visuales de ambos ojos. Esta deriva del paralaje entre las diferentes imágenes recibidas por la retina en cada ojo, es decir, por la disparidad binocular.

La estereopsis funciona muy bien para percibir la profundidad de objetos dentro de un rango de 5m y más aún cuando se manipulan objetos dentro del rango alcanzado por los brazos. Por el contrario, como la estereopsis se basa en las diferencias entre las imágenes percibidas por ambos ojos, ésta no funciona muy bien cuando los objetos están muy lejos. Esto se debe a que la diferencia entre las imágenes percibidas por cada ojo es muy pequeña.

2.2.1. SEÑALES DE PROFUNDIDAD EN MOVIMIENTO

Las señales de profundidad en movimiento surgen del paralaje generado al cambiar la posición relativa entre la cabeza y los objetos que están siendo observados. Los objetos que estén más cerca del ojo serán percibidos como si se moviesen más rápido que los objetos distantes. Cuando el cuerpo del observador se mueve, esto genera una respuesta propioceptiva que indica cuánto se ha movido, ayudando en la determinación de la distancia.

Cuando el observador no se mueve, el paralaje generado no aporta tanta información acerca del movimiento del objeto. Esto ocurre cuando solo los objetos se mueven o incluso cuando el observador se mueve pero no por sus propios medios, como por ejemplo desde un tren en movimiento.

2.2.2. SEÑALES DE PROFUNDIDAD FISIOLÓGICAS

Las señales de profundidad fisiológicas son generadas por los movimientos musculares del ojo cuando éste enfoca sobre un objeto particular. La acomodación es el ajuste de enfoque hecho por los ojos para cambiar la forma de los lentes. De esta forma, la cantidad de cambios musculares proveen información sobre la distancia para objetos que se encuentren entre dos y tres metros del observador. La convergencia es el movimiento coordinado de los ojos en el que sus ejes se desvían simultáneamente hacia el punto de visión. Los movimientos oculares usados para la convergencia proveen información al cerebro sobre la distancia a la que se encuentran los objetos observados.

3. PROPIEDADES DE LOS DISPLAYS VISUALES

Como no todos los displays visuales están hechos para ser utilizados en cualquier situación, cada uno de ellos tiene diferentes propiedades visuales y propiedades logísticas asociadas. Asimismo, todos ellos tienen métodos para transmitir la imagen visual al participante.

3.1. PROPIEDADES VISUALES

Las propiedades visuales son uno de los elementos más importantes a la hora de definir una experiencia de RV. Cada aplicación tendrá distintos requerimientos y no todos los dispositivos pueden adecuarse.

Alguno de esos dispositivos tendrá un gran campo de visión, por ejemplo, al disminuir su resolución. De esta forma, estos dispositivos no serán útiles cuando se necesite mucho nivel de detalle. Se debe tener especial consideración sobre estas propiedades ópticas, siempre teniendo en cuenta los requerimientos de la aplicación en cuestión.

El color es una característica importante cuando se consideran los displays visuales y esto depende directamente de la tecnología del hardware. El tipo de display puede variar desde tubos de rayos catódicos (CRT por sus siglas en inglés), displays de cristal líquido (LCD), cristal líquido sobre silicio (LCos) o diodo orgánico de emisión de luz (OLED).

Los monitores CRT son una de las primeras tecnologías de monitores color utilizados para PC. Estos contienen tres cañones de electrones, cada una apuntando a un tipo diferente de fósforo. Así, en una proximidad cercana en la pantalla, las señales rojas, verdes, y azules, pueden ser producidas. Por el otro lado, los displays LCD y OLED tienen elementos individuales que pueden cambiar de encendido a apagado para mostrar el color en una posición particular. Estas tecnologías también influirán en la resolución espacial de los píxeles en la pantalla.

La resolución espacial de cualquier display es dada por la cantidad de píxeles a través de las direcciones verticales y horizontales de la pantalla. Una forma de medir la resolución es a través del número de puntos por pulgadas (dpi por sus siglas en inglés). También es importante considerar el tamaño de la pantalla porque, dado un número de píxeles, una pantalla chica generará una mejor imagen que una pantalla grande.

La distancia del ojo a la pantalla también afecta la resolución percibida. Cuando se usa un *Head Based Display* (HBD) ¹, generalmente la pantalla se encuentra muy cerca de los ojos; por lo que los píxeles deben ser muy densos para resultar indistinguibles. En displays estacionarios, por el otro lado, la resolución aparente varía según la posición y la distancia del observador.

La cantidad de canales son importantes para poder lograr los efectos 3D característicos de RV. Un canal visual es una presentación de la información visual mostrada a cada ojo. Para sistemas 3D, se necesitan dos canales visuales con vistas separadas para lograr estereopsis, o efectos 3D. Esas imágenes se fusionan en el cerebro en una sola imagen estereoscópica.

Existen varias técnicas para poder transmitir señales simultáneas a través del mismo canal. Los métodos de multiplexado pueden transmitir dos señales de imagen simultáneamente para crear un display estereoscópico.

El multiplexado espacial consiste en posicionar imágenes separadas en frente de cada ojo, ya sea al utilizar dos pequeñas pantallas separadas o utilizar algún filtro para separar las vistas. Este técnica es la más utilizada en los HMDs actuales ya que al dividir en dos la pantalla se puede presentar una imagen diferente a cada ojo directamente.

El multiplexado temporal, por otro lado, presenta diferentes imágenes a cada ojo mediante el uso de gafas con obturador shutter glasses, que activan y ocultan la vista para cada ojo para prevenir que éste vea la vista del otro ojo. Esto lo realiza de forma coordinada y secuenciada. De este modo, los lentes opacos se vuelven transparentes para un ojo mientras que la vista apropiada para ese ojo es presentada.

El multiplexado polarizado es otra técnica que consiste en superponer dos imágenes separadas a través de filtros opuestamente polarizados. El observador usa un par de lentes con un filtro polarizado horizontal en un ojo y un filtro polarizado vertical en el otro; así cada ojo ve solo la información que le corresponde.

¹Display Para la Cabeza

El multiplexado espectral, por el otro lado, filtra la vista de cada ojo en un color diferente, mediante el uso de lentes estéreo anaglíficos por ejemplo. Entre estos se encuentran los típicos lentes rojos-azules que han sido muy populares debido a su bajo costo. Estos lentes neutralizan la vista para el ojo incorrecto, porque cada lente se correlaciona con una de las vistas coloreadas.

Los displays autoestéreo utilizan filtros para que cada ojo vea una imagen diferente sin necesidad de utilizar ningún dispositivo adicional. La técnica más común utiliza lentes lenticulares sobre la pantalla para asegurar que cada ojo vea la imagen correcta. Como resultado, el observador puede ver la imagen estereoscópica desde el display sin utilizar ningún dispositivo incómodo.

La distancia focal de un display es la distancia aparente de las imágenes desde los ojos del observador. Con la tecnología actual, la pantalla siempre está a la misma distancia del rostro. Por esto, todas las imágenes en la escena siempre están en el mismo plano focal sin importar su distancia virtual desde el observador. Los músculos de los ojos se ajustan para fijar la distancia focal a los objetos.

Generalmente, la habilidad de juzgar la profundidad es mucho más débil para objetos lejanos que para objetos cercanos. Por ejemplo, si una persona está sentada en un escritorio sobre el cual se hallan un par de lápices, es muy fácil determinar cuál está más cerca. Por el otro lado, si una persona observa algunas nubes en el cielo, será muy difícil determinar cuál está más cerca si ninguna de ellas pasa por delante de la otra.

La opacidad de los displays visuales se refiere a si el display permite ver el mundo físico o no. Las pantallas de los displays estacionarios no enmascaran el mundo real por lo que son no-opacas, es decir, permiten ver el resto del mundo. La mayoría de los HBDs son opacos y ocuyen el mundo exterior del usuario.

La opacidad juega un rol importante tanto en la seguridad como en el potencial colaborativo de un sistema de RV. Los displays oclusivos son peligrosos en el sentido de que los usuarios pueden chocar con objetos del mundo real que no pueden ver. Algunas aplicaciones, por el otro lado, requieren que los usuarios se comuniquen con el resto del grupo. Por lo tanto, el nivel de aislamiento también debería ser considerado.

Algunos HMDs proveen una visión del mundo real combinada con el mundo virtual. Estos displays se llaman see-through HMD y son generalmente utilizados en aplicaciones de Realidad Aumentada.

Como se ve en la Figura 3.1, el *Field of View* (FOV) ² normal para un humano es aproximadamente entre 200 y 220 grados en sentido horizontal y 130 grados en sentido vertical, con 114 grados de superposición binocular ([4]). El FOV del display es una medida del ancho angular de la visión del usuario que es cubierta por el display en cada momento. La medida puede ser un porcentaje del FOV del usuario o el ángulo de cobertura horizontal y vertical del display. Por ejemplo, un display con unos pocos grados de FOV horizontal (por ejemplo 60) provee una visión de túnel. Esto sucede en algunos HBDs que proveen una alta resolución a expensas del FOV.

La superposición estéreo del FOV también es importante. Ésta aparece en los HMDs donde las pantallas para cada ojo se encuentran lo suficientemente juntas. De esta forma, puede ocurrir que un ojo vea parte de la pantalla correspondiente al otro ojo. Con una superposición del FOV tan pequeña como de 30 grados, será difícil percibir la estereopsis.

²Campo de Visión

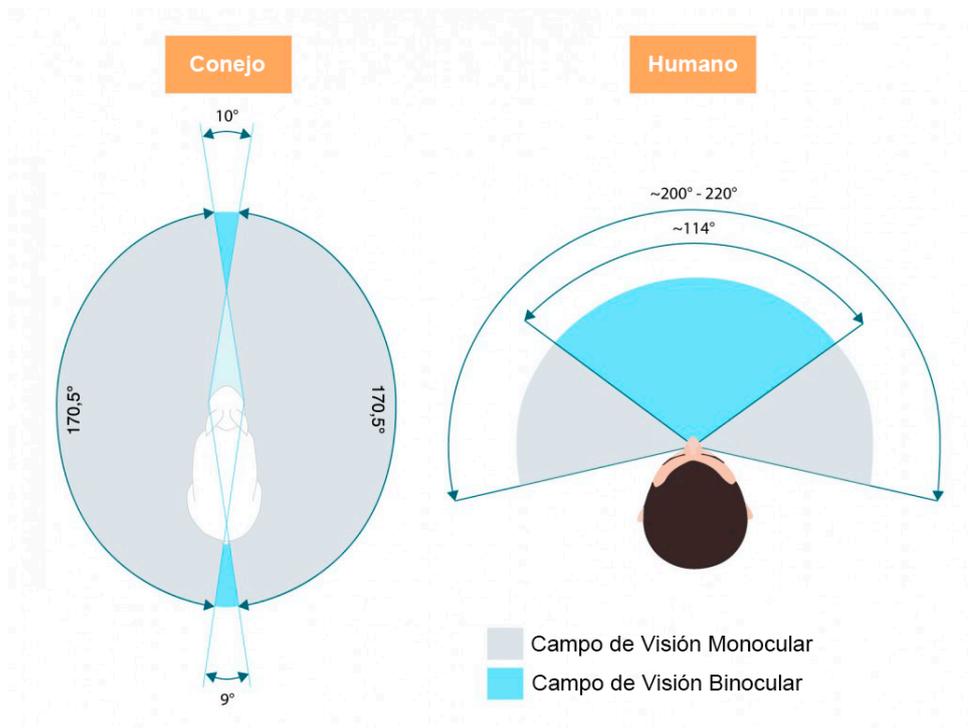


Figura 3.1: Representación del FOV de un humano comparado con el de un conejo. En azul se muestra la superposición binocular.

El *Field of Regard* (FOR) ³ de un display es la cantidad de espacio que es llenado con el mundo virtual alrededor del usuario, o en otras palabras, qué tanto el observador se encuentra envuelto por la representación visual. Por ejemplo, en un HMDs con rango de movimiento ilimitado, el FOR es de 100 % porque la pantalla siempre está en frente de los ojos del usuario. Sin importar qué dirección el observador mire, el mundo virtual siempre se muestra frente a sus ojos.

Para displays estacionarios, sin embargo, el FOR usualmente es menor al 100 %, ya que el mundo virtual no puede ser mostrado en el espacio que no tiene pantallas. Por ejemplo, una CAVE tendría un 100 % de FOR si rodea completamente al observador. Esto puede verse en la Figura 3.2.

El FOR es independiente del FOV. Se puede tener un FOV muy estrecho en un HMD e igualmente tener 100 % de FOR, ya que al ver en todas direcciones se puede ver el mundo virtual. Por otro lado, se puede tener una pantalla muy grande que genere un gran FOV, pero un FOR muy limitado ya que al mirar fuera de la pantalla, ya no se puede ver el mundo virtual.

En tanto que la información sobre la posición del usuario es beneficiosa para producir vistas precisas del mundo virtual, los distintos tipos de displays visuales confían en distinta información sobre la posición. En el caso de displays visuales estacionarios, los cálculos para producir una vista precisa se basan en la posición de los ojos respecto de la pantalla.

Si el usuario se encuentra en movimiento, la representación visual debería adecuarse a dichos movimientos. Para esto se debe conocer la posición de los ojos del usuario. En lugar de trackear los ojos del usuario, resulta más conveniente y fácil trackear la posición de la cabeza. A partir esto, se puede aproximar la posición de los ojos. Para pequeñas pantallas, se puede asumir que la cabeza del usuario está apuntando al centro de la pantalla. Sin embargo, no es posible asumir esto cuando se utilizan pantallas grandes.

³Campo de Consideración

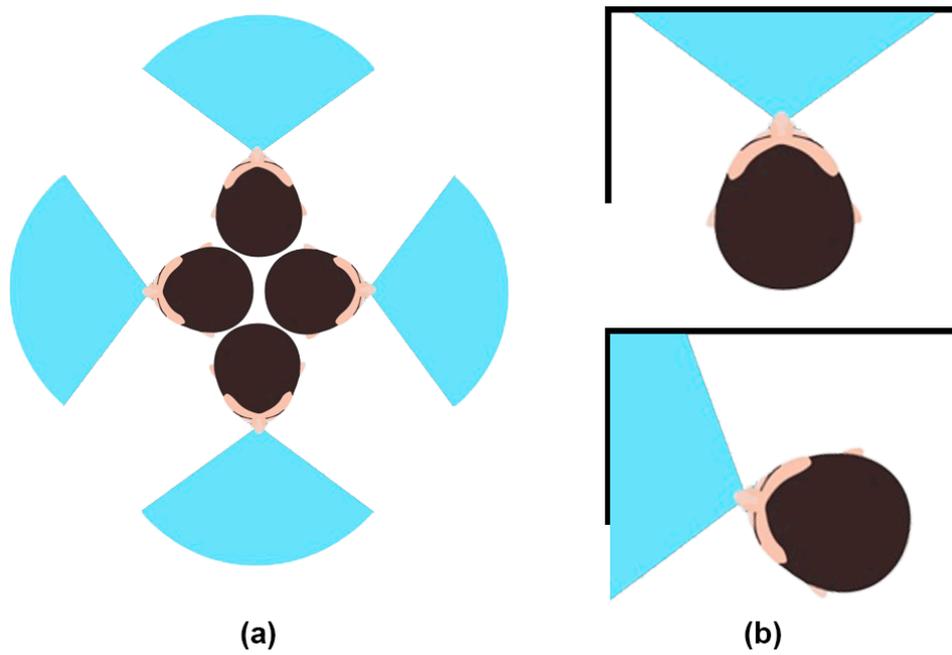


Figura 3.2: Campo de consideración (Field of Regard). (a) Un usuario con un HMD que le permite ver en todas direcciones tiene un FOR del 100 %. (b) Un usuario observando una pantalla fija tiene un FOR reducido debido a la superficie limitada de la pantalla.

En el caso de los HBDs, la orientación de la cabeza también es muy importante para el renderizado adecuado de la escena. Por ejemplo, cuando los usuarios giran su cabeza a la izquierda, ellos esperan ver lo que está a su izquierda. El tracking de la posición será útil para representar los desplazamientos laterales del usuario. Así, cuando los usuarios mueven su cabeza horizontal o verticalmente, el sistema renderizará la imagen acorde a ese movimiento. Más detalles sobre los sistemas de tracking se presentan en el Capítulo 7.

En instalaciones de RV y videojuegos, la latencia del sistema siempre introduce retardos entre las interacciones del usuario y la respuesta visual ([5]). En tanto que el cerebro tolera retrasos pequeños pero perceptibles, e incluso recalibraciones temporales ([6]), retrasos más grandes pueden afectar la percepción y el comportamiento.

Dependiendo del sistema de tracking, en la mayoría de los HBDs, los usuarios deben esperar a que la computadora dibuje la nueva vista cuando ellos mueven su cabeza. Cuando esto sucede, la computadora debe computar la nueva vista del mundo según la nueva dirección a la que el usuario está mirando. El tiempo transcurrido entre que el usuario mueve la cabeza y la actualización de la imagen debe ser tan corto como sea posible para no causar ninguna inconformidad en el usuario. Esto es la principal causa del efecto conocido como *Simulator Sickness*⁴. Existen varios efectos negativos que hacen que las personas se sientan incómodas mientras utilizan sistemas de RV. En [7] investigan los efectos de la fatiga visual mientras se utilizan HMD y *smartphones*. Para un *review* sobre la *Simulator Sickness* en relación a los displays en sistemas de RV, ver [8].

⁴Malestar de Simulación

La mayoría de las personas no son conscientes de los retardos por debajo de los 120ms ([9]). Este valor puede extenderse hasta por encima de los 150ms cuando, por ejemplo, se controlan personajes en videojuegos ([10]). Incluso, retardos de alrededor de 200ms entre el movimiento del ratón y el cursor, o entre presionar un botón y el estímulo visual, pueden ser inadvertidos ([11]). Por el otro lado, otros estudios reportan que algunos usuarios son capaces de notar una latencia tan pequeña como de 50ms en el contexto del movimiento entre el mouse y el cursor ([11]) e incluso la latencia de alrededor de 15ms mientras usan un HMD ([12]).

Held and Durlach ([13]), por ejemplo, reportaron que los retardos tan pequeños como de 60ms interfieren significativamente con la adaptación motriz. La estabilidad perceptiva de los usuarios a través de diferentes ambientes virtuales mientras usan un HMD requeriría latencias de menos de 16ms ([14]).

El ritmo con el que las imágenes son mostradas se denomina *frame rate*⁵ y está representado por el número de frames por segundo (FPS). El *frame rate* generalmente no es una función del tipo del display visual utilizado, sino de la habilidad de renderizado del hardware y el software, y de la complejidad visual del mundo virtual. El *frame rate* está directamente relacionado con la latencia y así también es la causa del *Simulator Sickness*.

3.2. PROPIEDADES LOGÍSTICAS

Existen varios factores logísticos relacionados a los displays visuales. El espacio del piso es también un requerimiento para aquellas aplicaciones donde los usuarios deben caminar. Generalmente, los sistemas proyectados requieren una gran cantidad de espacio físico para los proyectores, espejos o pantallas. Al contrario, un HBD tradicional requiere no suele requerir un espacio físico muy grande. Sin embargo, el peso, la opacidad, los cables conectados y otros elementos suelen ser un problema para este tipo de displays. Todas estas cuestiones logísticas que no se relacionan directamente con la salida visual del display, afectan igualmente a la experiencia del usuario.

⁵resolución temporal

La mayoría de los displays visuales imponen ciertas restricciones físicas al usuario, como por ejemplo cables, sistemas de tracking con rango limitado o displays estacionarios que no permiten que el usuario se mueva más allá de una cierta área. Todos estos aspectos afectan la movilidad del usuario, afectando directamente el nivel de inmersión.

Algunos sistemas, como por ejemplo el CAVE ([15]), requieren estrictamente que el observador se encuentre dentro de un área restringida. Otros también requieren que el observador se encuentre dentro de un área para visualizar la pantalla de forma correcta en relación a la dirección y distancia de observación.

La mayoría de los HBDs están alimentados por una PC de escritorio. En estos casos, el HMD generalmente se conecta a la computadora con un cable HDMI y USB, restringiendo así los movimientos del usuario. Por el otro lado, los nuevos HMDs que funcionan con smartphones no presentan esta restricción, pero se consideran otras propiedades, como la duración de la batería o la calidad de los gráficos.

Los métodos de tracking también se ven influenciados por el dispositivo visual. El sistema de tracking puede tener un rango de operación pequeño o grande, por lo que debe ser seleccionado acordemente. Aún más, algunos métodos agregan hardware extra que es prohibitivo para ciertas aplicaciones. En general, los displays de RV móviles requieren un método de tracking móvil, y para los displays de RV estacionarios suele bastar con un sistema de tracking fijo.

El ambiente en donde el sistema de RV es utilizado puede afectar la elección del display visual. Por ejemplo, un display de proyección requiere un ambiente con poca luz y sería muy ineficiente en exteriores. Por el otro lado, los HBDs no se ven afectados por la luz ya que rodean completamente la cabeza del usuario.

El tamaño del ambiente también es importante. Si la habitación es muy pequeña, tal vez un display de proyección no entraría. Por otro lado, para habitaciones muy grandes con muchas personas, un display simple de escritorio con una pequeña pantalla podría no ser suficiente para que todos lo vean. Para aplicaciones de RV con HBDs con muchos usuarios, se necesitarían muchos HMDs.

La portabilidad del sistema es otra propiedad importante a considerar. La importancia de la portabilidad depende de quién va a utilizar el sistema de RV y qué tan lejos el sistema se debe transportar. Generalmente, los sistemas de RV de proyección se diseñan para ser estacionarios, pero algunas veces pueden ser requeridos para grandes presentaciones y por eso, requerirían ser transportados.

Por otro lado, los HBDs usualmente dependen de una computadora para renderizar el mundo virtual y procesar la información de tracking. Esas computadoras deben tenerse en cuenta ya que serán transportadas con el sistema.

Uno de los requerimientos más importantes de los sistemas de RV, y que se encuentra directamente relacionado con la HCI, es la comodidad. Sin considerar los sistemas de tracking, la mayoría de los displays estacionarios no tienen muchos problemas de incomodidad.

Para lograr el efecto estereoscópico, la mayoría de estos sistemas utilizan gafas especiales, pero éstas suelen ser pequeñas y cómodas. Por otro lado, los HMDs suelen ser más incómodos de utilizar que los otros displays visuales; éstos pesan más que un par de gafas, por lo que los usuarios se cansan luego de un corto período de tiempo. Además, la mayoría de estos sistemas tienen cables, haciendo que los usuarios puedan caminar solo una corta distancia o rotar no más de 360°. También es muy común que estos dispositivos no se acomoden bien a la cabeza.

Existen muchos problemas de seguridad relacionados con los displays visuales. Los HBDs pueden provocar que los usuarios choquen con los objetos reales. La fatiga y el dolor de los ojos también son un problema; cuanto más se utilice el sistema sin descansar los ojos, mayor es el efecto. El peso del display también puede generar fatiga de cuello y dolor.

La *Simulation Sickness* es un potencial efecto a corto plazo generado por los displays de RV, especialmente cuando se utilizan los HMDs. La disparidad en la percepción de la profundidad o el retraso entre los movimientos de la cabeza y la actualización de la escena pueden causar náusea.

El costo de los displays visuales varía enormemente según su tipo. Los HBDs suelen ser más económicos que los sistemas de proyección. Cuando se está usando RV estacionaria, se pueden utilizar muchas pantallas y muchos proyectores. Sin embargo, existe un gran rango de precios y calidades en HBDs.

La mayoría de los *trade-offs* entre los displays surge de la resolución, el FOV, el contraste, y la durabilidad. Los displays de proyección suelen ser más costosos porque requieren muchos proyectores para crear un display que rodee al observador y para producir las vistas para el ojo izquierdo y derecho en displays estereoscópicos usando lentes polarizados.

4. TIPOS DE DISPLAYS VISUALES

Existen diferentes tipos de displays visuales y se pueden clasificar de varias formas. Las características principales que los diferencian son sus propiedades físicas, su disposición y el efecto visual que generan. No todos los dispositivos funcionan bien para todas las aplicaciones de RV imaginables. En esta sección se detallan las propiedades de estos tipos de displays visuales, sus componentes, sus características y los inconvenientes más comunes asociados a ellos.

4.1. DISPLAYS ESTACIONARIOS CON MONITORES

Este es el más simple y el primero de los displays de RV creados. Solo necesita un monitor de computadora para mostrar el mundo virtual, de tal forma que los observadores puedan moverse para verlo. Una gran desventaja de este tipo de display es que no es completamente inmersivo ya que los usuarios podrán ver el mundo real.

A diferencia de los videojuegos típicos donde no importa la posición real del usuario, en estos casos la cabeza del usuario debe ser trackeada para que el sistema pueda actualizar la representación visual de manera acorde. Como el sistema generalmente no se mueve mientras se está ejecutando la simulación, este tipo de display es llamado un display de RV estacionario.

Los displays de RV con monitores requieren solo unos pocos componentes. Por un lado, lo más importante es una computadora para renderizar el mundo virtual e interactuar con el usuario en caso de ser necesario. Además, se necesita una pantalla para mostrar la visualización. Cualquier monitor estándar de PC es suficiente. Sin embargo, es importante considerar las especificaciones de la aplicación particular. Por ejemplo, podría necesitarse una pantalla con mucha resolución, varios monitores o técnicas estereoscópicas.

El tracking también es una parte importante de este tipo de display. En estos casos, se puede asumir que el observador se encontrará frente al monitor. Así, la posición de la cabeza del observador sería trackeada y al asumir que éste observa siempre al monitor, no es necesario registrar la orientación de la misma. Como se verá en el Capítulo 7, la mayoría de estos sistemas utilizan una cámara situada en el monitor y técnicas de procesamiento de imagen para trackear la posición de la cabeza del observador. De esta forma, la visualización puede ser actualizada acordeamente.

Otro componente importante es el display binocular estereoscópico para generar una sensación de 3D o profundidad de la imagen. Existen muchos métodos para hacer esto y puede utilizarse cualquiera de los mencionados anteriormente. Los más comunes utilizan gafas especiales, como por ejemplo las anaglíficas o polarizadas. También pueden usarse filtros especiales sobre el monitor para crear efectos autoestereoscópicos. Sin embargo, esto requiere que el observador se encuentre directamente en frente de la pantalla, restringiendo así el rango de movimiento.

4.2. DISPLAYS DE RV PROYECTADOS

Los displays de RV que utilizan proyectores usualmente son dispositivos estacionarios debido a su tamaño y componentes. Presentan una gran pantalla comparados con los displays de monitor, generando así FOV y FOR más grandes. También permiten que el usuario camine más libremente. Este tipo de displays generalmente son proyectados por detrás de la pantalla porque sino los usuarios generarían sombras en las pantallas.

Como los sistemas de RV proyectados típicamente se utilizan en lugares grandes, suele requerirse la posibilidad de trabajar con usuarios que puedan moverse. En esos casos, el sistema de tracking juega un rol importante ya que la imagen debe actualizarse según la posición del observador, y a veces también la orientación del mismo. Además, en lugares grandes, el rango de tracking debe ser grande también.

En sistemas grandes, serán necesarios muchos proyectores. En esos casos, el proyector no solo debe estar alineado con la pantalla, sino que los proyectores deben estar alineados entre ellos. Por el otro lado, varios proyectores requieren también varios renderizados de imagen. Entonces, se necesitaría una computadora capaz de generar varias salidas de imagen, o varias computadoras que deberán estar sincronizadas para obtener la vista apropiada.

Existen varias técnicas para generar una vista 3D estereoscópica en displays de RV proyectados. Los lentes polarizados requerirán ya sea un solo proyector especial o dos proyectores equipados con filtros para polarizar las imágenes para el ojo izquierdo y el derecho. El participante entonces tendrá que utilizar un par de lentes equipados con filtros polarizados. Los lentes polarizados son más baratos y necesitan menos mantenimiento que los *shutter glasses*, pero requieren un segundo proyector por pantalla.

Este tipo de displays emplean una porción del FOV del espectador más grande que los monitores o la mayoría de los HBDs. Sin embargo, el utilizar más pantallas y pantallas más grandes requerirá más procesamiento, por lo que existe entonces un *trade-off* entre las pantallas y el FOV.

Si el sistema está pensado para usarse con varias personas al mismo tiempo y solo se trackea una persona, las otras verán la imagen con distorsiones. Cuando más cerca estén de la persona trackeada, menos distorsión verán. Algunos sistemas tienen filtros especiales para que varios espectadores puedan ver diferentes cosas en la misma pantalla. Sin embargo, esto suele ocasionar problemas de brillo y para lograr este efecto se requiere más hardware.

Al contrario de la mayoría de los HBDs, en un sistema de proyección, el espectador no está aislado del mundo real. Dependiendo de la aplicación, esto puede ser una ventaja o una desventaja. Si la aplicación requiere que varios usuarios se vean entre sí, esto es una ventaja. Sin embargo, estos serán capaces de ver el mundo real y todos los elementos en él. Esto podría no ser algo bueno y la inmersión se vería afectada. Por otro lado, esto elimina por ejemplo la necesidad de renderizar avatares ya que los usuarios se verían entre ellos y a los elementos de la sala.

Una de las desventajas principales en estos sistemas son los errores de oclusión. Por ejemplo, cuando un objeto virtual debe mostrarse más cerca que algún objeto real, la oclusión del objeto virtual por el objeto real bloquearía las señales de profundidad presentadas. Esto puede ser muy confuso para el sistema visual del usuario.

4.3. DISPLAYS PARA LA CABEZA (HBDs)

A diferencia de los displays de escritorio y proyectados, las pantallas de un HBD no son estacionarias. Como su nombre lo sugiere, se mueven junto con la cabeza del usuario. Por esto, las pantallas deben ser lo más pequeñas y livianas posibles.

La mayoría de los HBDs generan señales de imagen estereoscópicas mediante dos salidas visuales, una para cada ojo. Varios de estos sistemas se conectan a la computadora como si fuesen otra pantalla más. Este es el caso del Oculus Rift ([16]) o el HTC Vive ([17]), dos de los HMDs más populares de hoy en día. La ventaja de esto es que la calidad gráfica de la imagen de salida dependerá de la computadora utilizada. Por otro lado, como el sistema estará conectado a la computadora, los cables evitarían que el usuario pueda moverse libremente.

Existen sistemas que no necesitan una computadora para funcionar. Estos sistemas contienen todos los elementos necesarios para funcionar (CPU, GPU, pantalla, sensores, etc.) y usualmente funcionan con un *smartphone*. La ventaja de estos sistemas es la portabilidad. Los usuarios podrán caminar libremente sin ningún tipo de cable.

Una de las principales desventajas de los HBDs es que tiene mayor probabilidad de generar *Simualtor Sickness*. Esto es causado principalmente por el retardo en la imagen. Como tienen mecanismos y pantallas cerca de la cara de los usuarios, también suelen ser incómodos de utilizar. Los HMDs no se acomodan bien a los anteojos y el peso de muchos HMDs también causa fatiga y dolor de cuello. También puede ser difícil de utilizar por largos periodos de tiempo debido al conflicto entre las señales de profundidad de estereopsis, acomodación y convergencia. Este conflicto puede causar dolores de ojos, entre otros efectos negativos.

Estos displays enmascaran el mundo real y esto puede ser necesario para ciertas aplicaciones, ya que conservan la inmersión del sistema. Estos displays pueden operar en un gran rango de lugares a comparación de los sistemas de proyección, que requieren condiciones especiales de iluminación para operar. Sin embargo, debido a que el usuario no podrá ver el mundo real, se debe tener cuidado de no chocar con los objetos reales.

Los HBDs son buenos para generar una vista en primera persona y muestran el mundo virtual directamente a ojos del usuario. Esto provee una sensación muy intuitiva ya que, por ejemplo, los usuarios podrán “mirar a la izquierda” o “mirar a la derecha” si desean. Esto funciona mediante el tracking de la orientación de la cabeza de los usuarios. Además, si la posición también es trackeada, los usuarios podrán moverse o caminar en el mundo virtual. Muchos HBDs solo trackean la orientación de los usuarios porque esto puede realizarse fácilmente con sensores en el HMD. Para trackear la posición de los usuarios, se deben utilizar otras técnicas más complejas.

4.4. OTROS TIPOS DE DISPLAY

Existen otros tipos de displays que, aunque no se utilicen específicamente para RV, vale la pena mencionar. Por un lado, existen displays de uso masivo para aplicaciones no exclusivas de RV. También existen displays especiales y únicos, cuyo uso no ha sido masivamente estudiado.

Existen displays con los que se puede ver a través de ellos, denominados displays “*see-through*”. Este tipo de displays se utilizan para Realidad Aumentada permitiendo combinar la vista del mundo real con objetos y mundos virtuales.

Estos displays se pueden clasificar en dos tipos. Por un lado, aquellos en donde la imagen se proyecta sobre alguna superficie especial frente a los ojos del usuario, generando así el efecto de superponer objetos virtuales en el mundo real. El principal problema de este enfoque es que la proyección suele ser sobre alguna superficie transparente, por lo que los objetos virtuales no se verán perfectamente opacos. Este es el caso de los Google Glass([18]) y los HoloLens de Microsoft([19]).

Por otro lado, están aquellos displays donde la vista del mundo real que ve el usuario es capturada por una cámara. De esta forma los objetos virtuales se pueden superponer directamente a la imagen capturada por dicha cámara, presentando al usuario la imagen integrada. El problema principal de este enfoque es que el hardware utilizado debe permitir una reproducción veloz de la imagen capturada por la cámara. De lo contrario, el usuario verá la imagen con retraso, causando *Simulator Sickness*. Una ventaja de estos displays es que pueden utilizar los *smartphones* actuales, volviéndose un sistema portátil y de uso masivo.

Otro tipo de display que suele mencionar la bibliografía es el display “handheld” o “palm-VR”. Este tipo de display surge con la aparición de los dispositivos móviles, aprovechando su pantalla para utilizar como display de RV. El enfoque es muy similar a los displays de RV con monitores, solo que en este caso, el display no está pensado para ser estacionario sino que está en movimiento. Así, el observador mirará al dispositivo como si se tratase de una pequeña ventana a un mundo virtual.

El problema es que, como se mencionó anteriormente, se deben tener en cuenta ciertos aspectos para poder lograr el efecto de estereopsis 3D. Por un lado, si se trabaja con teléfonos celulares modernos, se pueden utilizar sus sensores para conocer la orientación del dispositivo y ajustar la imagen acorde. Por otro lado, se debe poder trackear la posición del observador y esto no es tan fácil de realizar. Para esto se puede aprovechar la cámara frontal del teléfono, pero el costo computacional puede ser excesivo. Además, para que el observador pueda ver efectivamente en 3D, debería utilizar alguna de las técnicas vistas, como por ejemplo la utilización de lentes especiales. Esto agrega más componentes de hardware al sistema. Finalmente, el FOV logrado es tan pequeño que por esta razón estos dispositivos han dejado de utilizarse para sistemas de RV inmersivos.

Okumura et al. ([20]) propusieron y desarrollaron un display con forma de domo que se proyecta en la cabeza, llamado Head-Dome Projector. Este display funciona tanto para RV como para Realidad Aumentada y puede verse en la Figura 3.3.

Existen displays que no utilizan pantallas, sino que proyectan la imagen directamente en la retina del observador. Estos displays presentan imágenes con alto brillo, contraste y resolución. Además, sugieren reducir el dolor de ojos, ya que en lugar de ver hacia la fuente de luz directamente, el observador mirará luz reflejada. Viirre et al. ([21]) presentaron uno de los primeros modelos. En su trabajo, ellos describen cómo funcionan estos displays y las posibles consecuencias de este mecanismo en aplicaciones médicas, incluyendo displays para realizar cirugías y displays para personas con visión disminuida.

Una versión más moderna de este tipo de displays es el Avegant Glyph ([22]). Éste presenta una matriz de 2 millones de micro espejos que reflejan la luz en las retinas. Esta luz proviene de un diodo emisor de luz de bajo poder. Así se produce un efecto similar a cómo nuestros ojos reciben la luz reflejada desde los objetos reales. Este sistema no necesita una pantalla ni ningún dispositivo externo.



Figura 3.3: Head-Dome Projector ([20]).

5. REVISIÓN DE LA LITERATURA

5.1. DISPLAYS ESTACIONARIOS CON MONITORES

Con respecto a los sistemas de RV basados en monitores, no se suelen utilizar debido a su pequeño tamaño y lo que esto representa en cuanto a la generación de inmersión. Sin embargo, estas limitaciones de tamaño de los displays pueden superarse al combinar varios monitores uno al lado del otro. Esta solución forma una matriz de monitores que pueden posicionarse sobre una pared o una mesa, por ejemplo.

Un ejemplo de esto es el display Lambda ([23]), un display de 100 mega píxeles que puede ser utilizado tanto vertical como horizontalmente. El Hyperwall es un display desarrollado por la NASA que consiste en 49 paneles LCD dispuestos en una matriz de 7x7. Cada panel de 18 pulgadas está controlado por nodos CPU, cada uno con su propia placa gráfica ([24]). Este display se utiliza para mostrar, analizar y estudiar conjuntos de datos de muchas dimensiones de varias formas. Dispone de diferentes herramientas, puntos de vista y parámetros para observar los datos de formas distintas. Actualmente, la NASA lo actualizó al Hyperwall-2 que consiste de 128 paneles LCD en una matriz de 8x16 ([25, 26]).

Esta distribución en mosaico de paneles LCD tiene la ventaja de que los paneles se pueden alinear y ajustar más fácilmente que los proyectores, quienes requieren una estructura compleja para fusionarse en un gran display. Sin embargo, la imagen resultante tendrá bordes visibles entre los monitores adyacentes ([27]), incluso en soluciones comerciales como el MultiTaction, donde cada pantalla cuenta con bordes ultra finos ([28]).

En [29] se presenta un display llamado CommunityWall, donde los usuarios pueden escribir sus mensajes con un lápiz especial o manipular objetos virtuales con el tacto. Peltonen et al. ([30]) presentaron CitiWall, un sistema con un gran display *multitouch*. Este sistema permitía que las personas buscaran fotos y videos descargados de las redes sociales. Otros sistemas se han utilizado en las calles para que las personas puedan buscar información, socializar con otras personas, y hasta jugar videojuegos ([31–33]).

Existen muchos otros sistemas de este estilo en donde los usuarios pueden interactuar con los displays. Como se mencionó anteriormente, el sistema debería trackear la posición del usuario para presentar las visuales adecuadas. Sin embargo, cuando se presentan múltiples usuarios, no se puede trackear a todos y mostrar las visuales para cada uno de forma adecuada. Es por esto que este tipo de sistemas no suele utilizarse para aplicaciones de RV en donde participen muchas personas.

Siguiendo con los displays estacionarios con monitores, también hay investigaciones sobre displays autoestereoscópicos. Por ejemplo, Peterka et al. ([34]) crearon un display autoestereoscópico para RV para visualizaciones científicas distribuidas. Este enfoque solo funciona con displays de escritorio con un rango de muestreo limitado y puede generar errores de trackeo. Sandit et al. ([35]) crearon The Varrier, una barrera de paralaje pasivo con 35 paneles que producían un gran campo de visión y una experiencia de RV con head-tracking inclusive. Para una revisión más extensa de displays autoestereoscópicos, Aritno et al. ([36]) presentaron un estudio comparativo utilizando displays autoestereoscópicos para RV y Realidad Aumentada.

5.2. DISPLAYS PROYECTADOS

Los sistemas de proyección han sido uno de los displays más investigados para sistemas de RV. Entre los primeros sistemas de proyección para RV encontramos, por ejemplo, un sistema de proyección sobre una mesa, el Responsive Workbench ([37]), otros con estilo de mesa de dibujo, como el ImmersaDesk ([38]) y otros sistemas de alta resolución, como el Infinity Wall ([38]).

A partir del surgimiento del sistema CAVE ([15]), el primer display de proyección pensado para RV, han surgido diversos sistemas en el ámbito académico e instituciones de investigación. Su uso masivo ha sido limitado, principalmente debido al alto costo para adquirirlos y al costo de mantenimiento. Los costos de adquisición típicamente son una combinación de diseño personalizado para cada instalación y el uso de proyectores especiales que puedan ser sincronizados y presenten display estereoscópico. Un sistema típico con tres pantallas proyectadas por detrás y otra para el piso o el techo, puede costar hasta medio millón de dólares. Esto se debe al diseño personalizado para soportar la estructura de las pantallas, más el costo de una sala especial para el sistema debido a los requerimientos de espacio para instalar las pantallas, el piso y los espejos.

Luego de que la CAVE fuera desarrollada, se crearon varias alternativas para sistemas inmersivos basados en proyección. Una de estas alternativas es un sistema con una o dos mesas, como el Responsive Workbench ([37]) y el ImmersaDesk ([38]), donde las imágenes se proyectan por detrás en una mesa y una pantalla vertical ortogonal en frente del usuario.

Los avances en la tecnología de proyección, especialmente la proyección en superficies no planas, llevaron al desarrollo de displays con pantalla curva como el i-Cone ([39]) o el Xphere ([40]). Este último utiliza varios proyectores para crear una imagen continua en una media esfera. El proyecto Allosphere ([41]) introdujo el concepto de usuarios caminando dentro de una gran esfera. Algunos sistemas como La Cueva Grande ([42]) o el StarCAVE ([43]) mejoraron la resolución y la calidad general del sistema CAVE, aunque sin dejar de ser costosos.

Los displays de mediana escala proveen colaboración multi-usuario ([44]), e incluyen displays planos inmersivos ([45]), displays con pantalla curva ([40]) y “mega pantallas” multi canales ([46]), todas éstas permitiendo un número pequeño de usuarios colaborativos. Los displays inmersivos de gran escala utilizan pantallas de gran FOV, como por ejemplo domos, para proveer un alto grado de inmersión para docenas o hasta cientos de espectadores ([47]). Pantallas cilíndricas o domos son preferidos sobre pantallas planas en aplicaciones cinemáticas ya que proveen una mejor apariencia sobre una gran rango de ángulos de visión y condiciones ambientales ([48]).

La tecnología de renderizado de video de gran escala se originó para sistemas de simuladores de entrenamiento ([49]), migrando luego por ejemplo, a los planetarios ([50-52]). Los domos digitales suelen encontrarse en ferias o parques, planetarios y centros científicos ([48, 53-55]), universidades, eventos corporativos u otros eventos.

Además, estos displays presentan un gran rango de aplicaciones incluyendo producciones de cine ([56]), simulaciones de RV en tiempo real ([57]), actuaciones de arte en vivo ([58]) y aplicaciones interactivas con audiencia ([59]). Una revisión de displays inmersivos de gran escala hasta el 2007 es presentada en [60].

En [61] dispusieron 15 proyectores en una matriz de 5x3, generando así un display de 6.400 x 3.072 píxeles sobre 5x2.5 metros. Uno de sus objetivos fue lograr unir las imágenes de cada proyector con sus vecinos sin que se note la línea divisoria entre ellos. Otros displays similares fueron presentados en [62], [30] y [63]. Estos displays requieren una estructura compleja y una tecnología específica para alinear correctamente las salidas de los proyectores. Por esto, el costo de dichos displays sigue siendo todavía muy alto.

Como el costo de estos displays de proyección sigue siendo muy elevado, algunas investigaciones se han enfocado en construir prototipos de bajo costo. En [64] presentan un diseño de pantalla de bajo costo y fácil de mantener, basado en componentes accesibles. La calidad del sistema se compara significativamente con las soluciones disponibles comercialmente que son más caras. Además, los usuarios con conocimientos promedio pueden implementar este diseño, que tiene la ventaja adicional que cualquiera de los componentes puede ser mejorado individualmente según las necesidades y el presupuesto disponible. Cruz-Neira et al. ([65]), por otro lado, presentaron un display de RV modular y accesible basado en componentes comerciales.

Algunas tecnologías de proyección permiten la visualización de la salida en una superficie para la cual no ha sido pensada, es decir, que se adapta a diferentes situaciones ([66–68]). De esta forma, la salida se puede proyectar en una pared, el piso o cualquier superficie, como el interior de un paraguas o un ventilador ([67]). Este tipo de instalación se usa en conjunto con tecnologías especiales de tracking que puedan detectar los gestos de los usuarios sobre las superficies proyectadas ([68, 69]).

Otros sistemas presentan displays esféricos o cilíndricos ([70–72]). En [73] presentan OrbeVR, un display esféricamente cóncavo que corrige la perspectiva. Este display combina imágenes proyectadas por múltiples pico-proyectores de alto rendimiento posicionados dentro de una esfera traslúcida. La posición de los usuarios es trackeada de tal forma que el display renderiza las imágenes teniendo en cuenta la perspectiva y las señales de profundidad estereoscópicas. Para una comparación entre displays esféricos, cilíndricos, verticales, y horizontales, se recomienda ver Benko et al. ([74]).

Steed et al. ([75]) proponen una nueva clase de display reconfigurable para soportar tanto pequeños grupos de observadores como así también usuarios individuales. En particular, desarrollaron un display robótico que puede cambiar de forma para soportar distintas condiciones de uso. Además, puede moverse rápidamente para generar un gran FOV.

Una técnica inusual de proyección es presentada en [76], en donde los autores experimentan con un display inmaterial, llamado FogScreen. Este display proyecta la imagen en un plano inmaterial en el aire. Así, el observador puede incluso pasar a través del display. Además, este display puede utilizarse de ambos lados. Las interacciones con objetos gráficos inmatrimales pueden realizarse al “tocar” estos objetos con la mano, ya que la aplicación también trackea al usuario.

Estos displays pueden extenderse para convertirse en displays de RV, pero la vista libre y los ángulos de proyección oblicuos imponen grandes problemas para que la tecnología pueda usarse como una pantalla de RV. En [77] presentan algunas medidas y experimentos de esta tecnología para obtener algunas limitaciones que puedan servir para generar mejores FogScreen pseudo 3D.

5.3. DISPLAYS PARA LA CABEZA (HBDs)

Los HMDs han sido utilizados desde los inicios de la RV. Debido a los límites tecnológicos de los HMDs disponibles hasta hace algunos años, uno de los tópicos de investigación más popular ha sido la evaluación de la presentación apropiada de la información estereoscópica.

Existen muchos enfoques hoy en día y en particular, en los últimos años se han propuesto varios modelos de HMDs accesibles, como por ejemplo Oculus Rift([16]), HTC Vive([17]), Playstation VR([78]), Samsung Gear VR([79]) y Google Cardboard([80]). Los modelos más recientes de HMD se caracterizan por poca latencia, displays de alta resolución y FOV más grandes en comparación de los dispositivos de la anterior generación. Además, algunos HMDs portátiles utilizan *smartphones* como unidad principal de procesamiento y visualización, permitiendo un uso fácil, gran difusión y una integración eficiente entre RV y las tecnologías y aplicaciones móviles.

En [81] propusieron un sistema para construir un HMD de alta resolución basado en una matriz de agujeros utilizando dispositivos móviles. Realizaron distintas pruebas de concepto y distintos prototipos que pueden ser utilizados con las pantallas de los dispositivos móviles, como teléfonos celulares o *smartwatches*. Mediante esta matriz de agujeros no es necesaria la utilización de ningún tipo de lentes, haciéndolo uno de los displays más simples de construir.

Varios sistemas han sido investigados para mejorar la resolución y el FOV de los HMDs ([82]). Cheng et al. ([83]) presentan cuatro diseños distintos de HMDs de gran FOV y alta resolución. Cada uno de ellos provee un FOV mayor a 80 grados y una resolución angular mejor que 4 arcominutos. En [84] diseñaron un sistema de HMD óptico monocular con un gran FOV y alta resolución. Gou et al. ([85]) desarrollaron un nuevo sistema de display LCD con tasas de refresco más rápidas que lo normal, lo que puede resultar muy beneficioso en aplicaciones de RV.

En [86] presentan un diseño novedoso de HMD con un gran FOV que puede cubrir hasta el FOV humano. Este sistema se basa en lentes curvados alrededor de los ojos, y presenta ventajas como el ser compacto, liviano, de bajo costo y poseer alta resolución.

La Realidad Virtual Cinemática se ha hecho muy popular en los últimos años y existen estudios que comparan el uso de HMDs contra pantallas tradicionales. MacQuarrie et al. ([87]) hicieron un experimento donde concluyeron que el uso de HMDs se disfruta más que las pantallas tradicionales. Además provee una mejor percepción espacial.

Existen investigaciones que analizan casos particulares utilizando distintos tipos de display de RV para definir cuál se adecúa más a sus necesidades particulares. Buttussi et al. ([88]) analizaron los procedimientos de seguridad en un juego serio de entrenamiento en aviación. Para esto utilizaron displays de escritorio con monitores y distintos tipos de HMDs. Sus resultados sugieren que el tipo de display es significativo en cuanto a la presencia e inmersión lograda.

Cordeil et al. ([89]) utilizaron un sistema estilo CAVE y un HMD para un estudio sobre el análisis de información abstracta en RV de forma colaborativa. Sus resultados sugieren que los HMDs actuales funcionan muy bien para realizar tareas colaborativas, por lo que en casos particulares puede evitarse invertir en costosos sistemas como el CAVE.

6. DISCUSIÓN Y CONCLUSIONES

6.1. PROPIEDADES DE LA VISIÓN Y SEÑALES DE PROFUNDIDAD

Hemos visto que las señales de profundidad son de vital importancia a la hora de desarrollar un sistema de RV que nos genere la sensación de percibir la profundidad 3D de los objetos en el ambiente virtual. Por esto es importante mencionar que no todas las señales de profundidad tienen la misma prioridad. La estereopsis es una señal de profundidad muy fuerte que suele dominar cuando está en conflicto con otras señales de profundidad. El movimiento relativo es tal vez la señal de profundidad que puede superar a la estereopsis.

En cuanto a las señales de profundidad estáticas monoscópicas, la interposición es la más poderosa. Algunas señales de profundidad solo funcionan en determinado rango. Por ejemplo, la estereopsis y la acomodación dejan de ser eficientes prácticamente luego de los 5m.

No todas las señales de profundidad son necesarias a la hora de implementar una aplicación de RV. Por ejemplo, tal vez no sea necesario implementar las sombras cuando el personaje se encuentra caminando por el suelo. Sin embargo, si el personaje comenzara a volar, las sombras proveerían información sobre la distancia del personaje al piso. La estereopsis puede no ser importante para todas las aplicaciones de RV. Generalmente, para aquellas aplicaciones en donde la vista del mundo se encuentra a una distancia mayor a 5m, la estereopsis no es requerida. Sin embargo, es importante para aquellas aplicaciones con acercamientos o manipulación de objetos con las manos.

La presencia de dos canales no siempre indica que una imagen estereoscópica se está mostrando. Una imagen estereoscópica consiste de dos vistas distintivas del mundo virtual, cada una específica para un ojo particular. Algunos sistemas muestran una sola imagen a ambos ojos. Estos sistemas se llaman displays binoculares monoscópicos. Pueden controlar las señales de profundidad focales y de convergencia a través del uso de la óptica y efectos de imagen.

No siempre es fácil crear displays estereoscópicos buenos. Además, si se hacen mal pueden causar incomodidad en el observador. Una revisión de la incomodidad visual y fatiga visual de los displays estereoscópicos puede verse en [90]. Además, representar imágenes estereoscópicas requiere alrededor del doble de recursos computacionales y hardware gráfico.

En aquellos displays sin un FOR completo, la estereopsis puede perderse cuando un objeto cercano solo aparece parcialmente en el display. El problema ocurre cuando los objetos virtuales localizados entre la pantalla y el observador se cortan por el borde de la pantalla. Las señales de profundidad de la estereopsis y el paralaje de movimiento están en conflicto con la señal de profundidad de la oclusión, que indica que el objeto debería estar detrás de la pantalla.

El impacto de la latencia usualmente se considera con respecto a un solo factor ([11, 12]). En otros casos, solo una simple tarea como presionar un botón son examinados sin enfocarse en RV ([91]). Los efectos de retardos de realimentación dependen del display utilizado, los movimientos realizados y la tarea motriz o perceptiva realizada. La mayoría de los estudios se han enfocado en solo una tarea y un movimiento, lo que hace difícil comparar resultados de diferentes estudios y derivar niveles aceptables de retardos para RV para diversas aplicaciones.

6.2. TIPOS DE DISPLAY

En cuanto a los diferentes tipos de display, los displays de escritorio con monitores suelen ser los más económicos y prácticos a la hora de desarrollar una nueva aplicación, siempre y cuando ya se disponga de una computadora para hacerlo. El hardware necesario es producirlo en masa y se encuentra disponible fácilmente. Incluso los componentes para lograr una vista estereoscópica no son muy costosos y están disponibles hoy en día. Puede utilizarse una cámara web para el tracking, que es un componente muy económico. Además, este tipo de display es fácil de instalar y testear, por lo que es recomendable utilizarlo cuando se genera nuevo contenido.

Como los displays estacionarios que utilizan monitores usan las mismas interfaces que las computadoras de escritorio y el mundo real es visible, algunas interfaces como los teclados y mouse son prácticos. Sin embargo, este tipo de interfaces reducen la sensación de inmersión ya que los usuarios no están interactuando con elementos que están “dentro” del mundo virtual.

Por otro lado, los monitores pueden juntarse uno al lado del otro para formar una matriz, creando así una pantalla mucho más grande. Esto tiene la ventaja de ser más fácil de alinear y ajustar que los proyectores, los que requieren una estructura compleja para combinarse en un solo gran display. Además, una gran pantalla generada por varios monitores suele ser más económica que una formada por proyectores. Sin embargo, como se mencionó, esto generaría un borde visible entre los monitores adyacentes y es un problema que aún no se ha solucionado.

Este tipo de displays presenta algunas otras desventajas. Como el observador debe ver en una dirección particular y el mundo real ocupa la mayoría del FOV, la sensación de inmersión es muy pobre. Por esta razón, este tipo de display es de los menos inmersivos. Sin embargo, si se implementa el tracking de la cabeza correctamente, se puede generar un buen efecto de que se está mirando a través de una ventana o dentro de una caja.

En cuanto a los sistemas de RV de proyección, al requerir componentes especiales para lograr una sincronización adecuada de las imágenes estereoscópicas, el costo de suele ser muy elevado. Además, estos displays requieren mucho espacio físico debido a las grandes pantallas y al espacio requerido entre los proyectores y las pantallas. Además, si los espectadores deben caminar, se requerirá más espacio. Por este motivo no se han comercializado masivamente y solo suelen encontrarse en institutos de investigación, grandes corporaciones o eventos especiales.

Otro de los problemas fundamentales de los display de proyección para RV es el poder lograr el efecto de estereopsis 3D. Para esto, se debe poder trackear al observador y se debe poder presentar imágenes distintas a ambos ojos, utilizando cualquiera de las técnicas anteriormente mencionadas. El problema es que, además de ser costoso, resulta casi imposible poder realizar esto para varios observadores al mismo tiempo. Por esto, este tipo de display no es práctico a la hora de presentar una experiencia de RV completa a muchos observadores.

Por el contrario, los displays de proyección siguen siendo uno de los displays más utilizados a la hora de mostrar mundos virtuales a varios observadores, pero sin generar efecto de estereopsis 3D. Esto se debe a que el gran tamaño de las pantallas permite una mejor visualización en lugares de gran tamaño y con muchas personas.

Existen también muchos otros tipos de superficies donde se puede proyectar. Los domos son más comunes hoy en día y mediante su utilización se puede lograr una visualización que rodee completamente al observador. De mismo modo se puede utilizar algún sistema de proyección estilo CAVE. Así, además de generar un gran FOV en el usuario, se generaría también un gran FOR.

Aunque estos sistemas sean considerados como los más costosos, se pueden crear versiones de bajo costo que, aunque no posean la misma calidad que los sistemas convencionales, su bajo costo hace que resulten muy convenientes. Otra ventaja de este tipo de display es que, como la pantalla usualmente está al menos a 50 cm del observador, producen menos dolor de ojos que los HBDs. Esto permite que los usuarios permanezcan inmersos por más tiempo.

Hoy en día, los HBDs son los displays más populares para utilizar en RV. El costo de los componentes de hardware se ha reducido, llevando a la comercialización masiva de varios de estos sistemas. Su bajo costo en comparación con los otros tipos de displays, hace que sean muy accesibles. Además, aquellos sistemas que funcionan solamente con un *smartphone* son todavía más accesibles.

Como sucede con los otros tipos de display, los HBDs no siempre logran generar un FOV igual al FOV humano. Esto genera una visión de túnel que reduce considerablemente la inmersión. Sin embargo, como se mencionó, existen proyectos orientados a investigar nuevas técnicas de construcción de displays para lograr un FOV igual o superior al humano. Además, en comparación a los otros tipos de displays, los HBDs suelen generar un FOR del 100 %, incrementando más aún la inmersión del sistema.

Los HBDs suelen estar equipados con sensores incorporados para trackear la rotación y la posición de la cabeza del usuario. Esto es necesario para lograr una buena representación del mundo virtual, incrementando también la inmersión. Los otros displays generalmente deben recurrir a sensores externos para poder hacer esto mismo.

Los HBDs también presentan varias desventajas. Por un lado, al estar siempre fijos a la cabeza del usuario, pueden generar mucha incomodidad. Además, aquellos sistemas con cables imposibilitan el libre movimiento del usuario. Por el otro lado, como la pantalla se encuentra muy cerca de los ojos del usuario, estos sistemas son los que tienen más probabilidad de generar *Simulator Sickness*. Por esto, el hardware debe ser lo suficiente veloz para evitar problemas de imagen, como por ejemplo, latencia.

Finalmente, estos sistemas son muy útiles solo cuando es una única persona la que va a utilizar el sistema de RV. Si se requiere que varias personas utilicen el sistema, éstas deberán tener que pasarse el HBD entre ellas, o de lo contrario se debe utilizar algún otro tipo de display.

4

Audio en Realidad Virtual

1. INTRODUCCIÓN

En nuestra vida diaria, recibimos una gran cantidad de información auditiva que contiene diversas señales sobre el ambiente que nos rodea. Escuchar es un sentido vital que define la percepción espacial mediante la distancia, dirección, posición y dimensión de las fuentes de sonido.

La industria del entretenimiento siempre ha reconocido la importancia del sonido para crear emoción. Colocar a una persona en un mundo virtual con una interfaz auditiva mal diseñada es equivalente a crear una discapacidad auditiva en el usuario. La percepción auditiva, especialmente la localización, es un fenómeno complejo que se ve afectado por la fisiología, la expectación e incluso la interfaz visual.

Las personas solo tienen dos oídos, pero pueden localizar sonidos en tres dimensiones. Esto es posible gracias a que el oído interno, medio y externo trabajan en conjunto para que el cerebro pueda generar inferencias sobre la posición del sonido. Aunque los ojos solo pueden ver una fracción del mundo alrededor del espectador, la habilidad para localizar fuentes de sonido funciona en todas direcciones, sin importar las condiciones de luz.

El sistema auditivo humano decodifica las señales de audio recibidas en los tímpanos para poder describir el ambiente acústico. Las señales espectrales del sonido, como la amplitud, frecuencia, forma, dirección y tiempo de arribo, generan una experiencia auditiva única y distintiva para cada individuo.

La habilidad de los humanos para poder localizar el sonido en el espacio ha sido un tema de gran importancia e interés en el ámbito académico. Las señales auditivas incrementan la conciencia del entorno, intensifican la atención visual y generan información sin sobrecargar el sistema visual. Así, cuando se diseña audio para Realidad Virtual (RV), las señales espectrales del sonido deben ser consideradas. La Sección 2 se enfoca en la audición humana y la Sección 3 describe aquellas señales que ayudan a crear una representación espacial del sonido.

Los avances tecnológicos recientes han permitido a los investigadores medir la precisión con la que las personas localizan el sonido. Para esto, es muy importante considerar los dispositivos utilizados para la reproducción del sonido. En la Sección 4 se presentan las diferencias entre estos dispositivos, así como sus propiedades y técnicas de implementación.

La Sección 5 presenta el trabajo previo relacionado con las técnicas de implementación de audio existentes, así como también a los dispositivos físicos utilizados para reproducir el sonido. Además, se hace especial hincapié en el trabajo relacionado con sistemas de RV.

La Sección 6 presenta una discusión sobre qué elementos deben tenerse en cuenta a la hora de diseñar audio para un sistema de RV. Cuando cualquier sistema de RV es desarrollado, se debe decidir qué técnica de audio utilizar. Las ventajas y desventajas de cada una también son presentadas. Finalmente, se sugiere qué sistemas de hardware conviene utilizar según el tipo de aplicación particular.

2. AUDICIÓN HUMANA

Esta sección provee una descripción de los aspectos más relevantes del sistema auditivo humano. Estos aspectos son importantes y deben ser considerados a la hora de diseñar el audio en los sistemas de RV.

2.1. PRESENCIA E INMERSIÓN

La presencia puede ser definida como "la sensación de estar inmerso en una simulación o ambiente virtual"([92]) y es un concepto muy difícil de medir. Se cree generalmente que la sensación de presencia es dependiente de la fidelidad auditiva, visual y táctil ([93]). Sin embargo, aunque el audio virtual no necesariamente incrementa el realismo percibido en RV, sí incrementa la sensación de presencia ([94]).

Los estudios muestran que el utilizar dispositivos auditivos de alta calidad puede mejorar la calidad percibida de la parte visual, e inversamente, utilizar dispositivos auditivos de baja calidad reduce la calidad visual percibida. ([95]). Por lo tanto, la implementación de audio apropiadamente diseñado incrementa la sensación general de presencia en un ambiente virtual o simulación.

2.2. INTENSIDAD

Las personas son sensibles a cambios en la intensidad del sonido en una escala logarítmica ([96]). El volumen de sonido percibido es básicamente proporcional al logaritmo del nivel de sonido transmitido a través del aire. Esta sensibilidad logarítmica brinda al sistema auditivo un gran rango dinámico.



Figura 4.1: Intensidad en decibeles de algunos sonidos representativos de nuestra vida diaria.

En la Figura 4.1 presenta una comparación de la intensidad en decibeles de algunos sonidos representativos encontrados en nuestra vida diaria. Por ejemplo, el rango comprendido entre el sonido apenas detectable y el sonido tan fuerte que puede causar dolor es aproximadamente de 110 a 120 dB. Los sistemas de reproducción de sonido típicos utilizan 16 bits para representar la presión de la señal acústica, lo que es suficiente para la mayoría de las simulaciones.

2.3. ENMASCARAMIENTO Y ATENCIÓN

Cuando múltiples fuentes de sonido son presentadas simultáneamente a un oyente, suelen interferir entre ellas. Por ejemplo, un tono que es audible cuando se reproduce solo, puede ser inaudible cuando otro sonido más fuerte es presentado al mismo tiempo. Estos efectos se conocen como efectos de enmascaramiento y se relacionan con problemas como la identificación de un sonido, varios sonidos compitiendo entre ellos, la identificación de una fuente de sonido dentro de una mezcla o alguna combinación de lo mencionado ([97]).

Estas fallas de atención auditiva selectiva pueden generar un gran impacto en la percepción, incluso con sonidos que son claramente audibles. Por ejemplo, la sensibilidad perceptiva en tareas de discriminación y detección usualmente se degrada cuando existe incertidumbre sobre las características de la fuente de sonido ([98]).

En un ambiente con múltiples fuentes de sonido, es más fácil prestar atención a un sonido particular si éstos se encuentran separados en el espacio. Esto se conoce como el *cocktail party effect* ([97-99]). Esta habilidad de dirigir la atención auditiva, que permite al oyente procesar qué fuente de sonido es más importante en cada momento, es crítica en muchas situaciones comunes como en teleconferencias ([100]) o comunicaciones de radio multicanal ([101-103]).

2.4. TONO Y TIMBRE

El tono es una característica perceptiva prominente y robusta que puede ser usada para brindar información al oyente. Así como la intensidad del sonido se relaciona con la percepción del volumen, la frecuencia del sonido se relaciona con la percepción del tono. El tono es también una señal muy importante para segregar fuentes de sonido que compiten y permitir al oyente enfocar su atención en la fuente objetiva ([104]).

La percepción del timbre es la propiedad del sonido que permite que un oyente distinga una flauta de una trompeta, por ejemplo. Al igual que el tono, la percepción del timbre depende de un número de parámetros físicos del sonido, incluyendo el contenido espectral y la envoltura temporal. El timbre es también una propiedad importante para que el oyente pueda identificar el sonido objetivo y puede ser utilizado para transmitir información a través de dispositivos auditivos ([105]). Sin embargo, los sonidos con distintos timbres tienen distintos pesos perceptivos, un factor que debería ser considerado a la hora de diseñar sonidos discretos para dispositivos de audio ([106]).

De la misma forma que el tono, el timbre es una característica que permite a los oyentes dirigir su atención auditiva selectiva al objetivo deseado entre varios sonidos. Esto permite extraer la información deseada desde la fuente de sonido sin importar la presencia de interferencias ([107]).

2.5. RESOLUCIÓN TEMPORAL

El canal auditivo es mucho más sensible a fluctuaciones temporales que el canal visual o propioceptivo ([108–110]). Por ejemplo, el sistema auditivo puede detectar fluctuaciones de amplitud en las señales de entrada de hasta 50Hz de manera fácil ([98]). Los oyentes no solo pueden detectar fluctuaciones rápidamente en un estímulo de entrada, sino que también pueden reaccionar rápidamente a estímulos auditivos. Por ejemplo, los tiempos de reacción a estímulos auditivos son entre 30 y 40 ms más rápidos que los correspondientes a estímulos visuales ([111]).

3. AUDICIÓN ESPACIAL

La audición espacial es muy importante en RV ya que permite que los usuarios estimen la posición de las fuentes de sonido, generando una experiencia más inmersiva y realista. La percepción espacial del audio es diferente a la de los otros sentidos. Los foto-receptores de la retina por ejemplo, se organizan topográficamente de tal forma que una fuente en una posición particular excita un grupo distinto de receptores ([108]). En cambio, la información espacial que llega a los oídos se computa desde las representaciones neuronales periféricas.

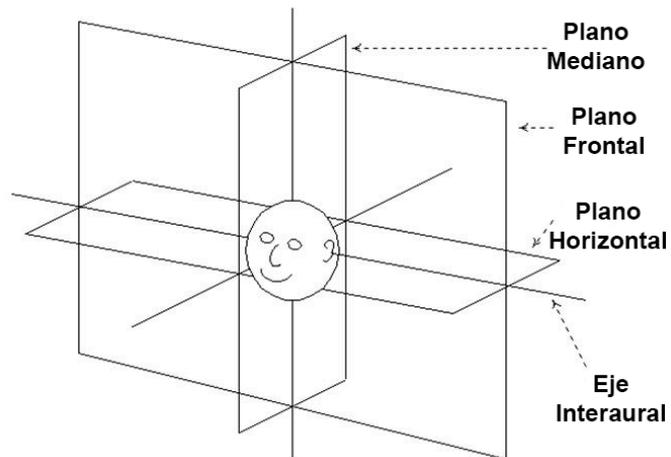


Figura 4.2: Representación del plano mediano, frontal, horizontal y el eje interaural, respecto a un oyente.

Para que un oyente pueda detectar un desplazamiento angular de una fuente de sonido a través del plano mediano (Figura 4.2), esta fuente debe desplazarse lateralmente al menos 1° . Además, esta fuente debe desplazarse al menos 15° para que el oyente pueda detectar un cambio en la dirección del sonido. Por otro lado, para una fuente de sonido directamente al costado, los oyentes no siempre detectan cambios laterales de menos de 10° . Aunque las personas puedan estimar cambios relativos de las fuentes de sonido, las distancias absolutas suelen estimarse erróneamente incluso bajo las mejores condiciones ([112]).

A continuación se presentan las propiedades auditivas que las personas utilizan para construir la representación espacial del sonido. Además, se discuten las formas en que la posición del sonido es computada por el cerebro.

3.1. SEÑALES BINAURALES

Las señales más robustas para identificar la posición de una fuente de sonido dependen de las diferencias entre las señales que llegan a los oídos. Estas señales interaurales o binaurales son robustas porque pueden ser computadas al comparar las señales que llegan a cada uno de los oídos. Como resultado, estas señales permiten que los oyentes estimen los atributos acústicos que surgen de dicha fuente.

Dependiendo del ángulo entre el eje interaural y la fuente de sonido, un oído recibe el sonido antes que el otro. Las **ITDs** (Interaural Time Differences)¹ son las señales principales que indican la lateralidad del sonido. La ITD crece según el ángulo del sonido respecto al plano mediano. La mayoría de las personas pueden detectar ITDs de 10 a 100 us, lo que se asemeja a una fuente de sonido posicionada entre 1° y 10° del plano mediano.

Cuando una fuente de sonido está a menos de un metro del oyente, surgen las **ILDs** (Interaural Level Differences)² (ILDs) extra grandes debido a las diferencias entre las distancias relativas entre el sonido y el oído izquierdo y derecho ([113, 114]). Estas ILDs ayudan a generar información sobre la distancia relativa y la dirección del sonido entre la fuente de sonido y el oyente ([115]). Otras ILDs de baja frecuencia que surgen desde el torso parecen ayudar a determinar la elevación de la fuente de sonido ([116]).

Nielsen et al. ([117]) realizaron un experimento para evaluar la precisión de la percepción de fuentes de sonido en el plano mediano. Reprodujeron un sonido a través de una serie de altavoces y se le pidió a los participantes que identificaran la posición de dicho sonido dentro de la sala. Los resultados mostraron que existe una gran diferencia entre los individuos y su habilidad para escuchar la profundidad auditiva absoluta. El límite observado en la dinámica de distancias percibidas es evidencia de un horizonte acústico, que fue explorado posteriormente por Bronkhorst y Houtgast ([118]).

¹Diferencias de Tiempo Interaurales

²Diferencias de Nivel Interaurales

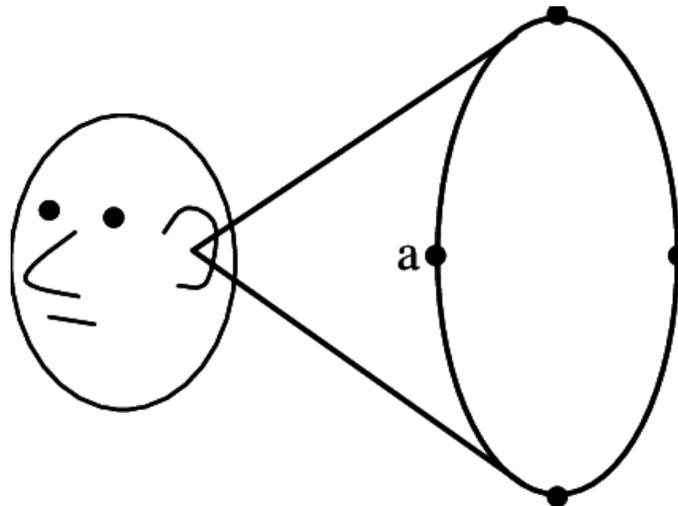


Figura 4.3: Cono de confusión. Para un oyente, las fuentes de sonido ubicadas en toda la base del cono, como por ejemplo el punto (a), tienen la misma ITD e ILD.

Aunque la posición percibida de una fuente de sonido generalmente es consistente con el ITD y el ILD, existen múltiples posiciones que generan aproximadamente las mismas señales ITD e ILD. En general, las posiciones que generan las mismas señales binaurales forman un círculo centrado en el eje interaural. Esto es un problema grave ya que el oyente no podría distinguir la posición de varias señales ubicadas en dicho círculo. Entre la posición del oyente y el círculo mencionado se forma un cono conocido como el cono de confusión (Figura 4.3).

Este problema es solucionado automáticamente mediante las propiedades físicas del oído y movimientos involuntarios de la cabeza. Sin embargo, no es simple de solucionar a la hora de implementar una aplicación de RV mediante la utilización del hardware disponible. Resulta difícil simular las características físicas del oído.

3.2. SEÑALES DE DISTANCIA ANECÓICAS

La intensidad del sonido directo que alcanza al oyente (es decir, sonido que no proviene de haber rebotado en las paredes o en cualquier otra superficie) se decrementa con la distancia a la fuente de sonido. La atmósfera absorbe energía mientras el sonido se propaga, causando pequeños cambios en la señal recibida cuando cambia la distancia a la fuente de sonido.

Si el sonido no es familiar para el oyente, la intensidad no será entonces una señal robusta para la distancia. Esto se debe a que puede confundirse con cambios en la intensidad de la señal emitida por la fuente. Sin embargo, incluso para fuentes desconocidas, la intensidad promedio y el contenido espectral proveen información sobre la distancia relativa a la fuente ([119]).

3.3. REVERBERACIÓN O ECO

Reverberación o eco es la energía que llega al oyente desde caminos indirectos como paredes, pisos, etc. y generalmente degrada la percepción de la dirección de la fuente de sonido ([96, 120, 121]). Sin embargo, esto ayuda a percibir mejor la distancia de dicha fuente ([122, 123]).

Una pequeña parte de la intensidad de la energía reflejada recibida en los oídos es independiente de la posición de la fuente de sonido relativa al oyente. Como resultado, para un dado ambiente, el sonido directo y el que rebota proveerán una medida absoluta de la distancia a la fuente de sonido.

El eco no solo provee señales robustas para distancias a la fuente de sonido, sino que también provee información sobre el tamaño y la configuración del ambiente. Mientras que muchos estudios psicofísicos sobre la localización del sonido se realizan en ambientes anecóicos, el eco está presente todo el tiempo en nuestra vida diaria. Por lo tanto, agregar eco a una simulación genera que todas las fuentes de sonido parezcan más realistas y provee información robusta sobre la distancia relativa a dichas fuentes ([124, 125]).

Mientras que el eco puede mejorar la distancia percibida y el realismo, puede disminuir la exactitud en la percepción direccional y puede interferir con la habilidad de extraer información al escuchar múltiples fuentes de sonido al mismo tiempo.

3.4. SEÑALES DINÁMICAS

Los cambios en las señales espaciales cuando el oyente se está moviendo también influyen en la percepción de la posición de la fuente de sonido ([126, 127]). Sin embargo, esto puede ayudar al oyente a resolver ambigüedades. Por ejemplo, un sonido que se encuentre ya sea directamente en frente o detrás del oyente, generará un ITD y ILD casi nulo. La Figura 4.4 muestra lo que ocurre en estos casos. Una rotación inconsciente hacia la izquierda generará un ITD y un ILD que favorezca ya sea al oído derecho (para un sonido al frente) o al oído izquierdo (para un sonido detrás).

Generalmente, la resolución temporal del sistema auditivo binaural es muy pobre. De hecho, existen investigaciones que sugieren que la percepción de una fuente de sonido en movimiento demora aproximadamente entre 100 y 200 ms en nivelarse. ([128, 129]).

3.5. EXPERIENCIA

La experiencia previa sobre la acústica de un ambiente particular también afecta la localización auditiva. Esta experiencia implícita afecta el rendimiento de la localización ([96, 130]). En otras palabras, la percepción auditiva espacial también es determinada por el estado del oyente. Aunque dichos efectos no se deben a decisiones conscientes, influyen directamente en la localización auditiva y la percepción espacial ([131]).

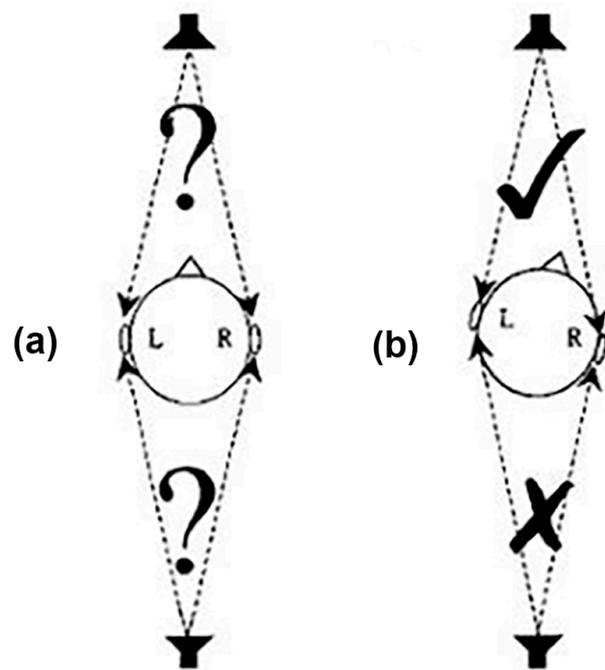


Figura 4.4: Ejemplo de cómo los movimientos de la cabeza permiten resolver ambigüedades de posición de sonido. Un sonido en frente del oyente (a) produce el mismo ITD y ILD que un sonido justo detrás. En (b), al inclinar la cabeza se revela la posición actual del sonido.

3.6. ADAPTACIÓN

Mientras que los oyentes responden a la ITD, a la ILD y a las señales espectrales basados en su experiencia diaria, también pueden aprender a interpretar señales que no son exactamente como las que ocurren naturalmente. Por ejemplo, los oyentes pueden aprender a adaptarse a señales espectrales que no son naturales, siempre y cuando se expongan a éstas por un período de tiempo suficiente ([132]).

Los entrenamientos cortos también permiten a los oyentes aprender a mapear las respuestas de señales espaciales fuera de lo normal a diferentes posiciones espaciales ([133]). Estos estudios sugieren que para aplicaciones en las que los usuarios pueden ser entrenados, no es necesario lograr simulaciones perfectas. Sin embargo, existen límites para los tipos de distorsiones de señales espaciales a los cuales los oyentes pueden adaptarse ([134]).

3.7. INFORMACIÓN INTER-SENSORIAL

La importancia de las interacciones multimodales relacionadas con el sistema auditivo no pueden ser ignoradas. Una gran cantidad de estudios señalan que los criterios de los oyentes sobre la modalidad de un sentido son influenciadas por la información de las modalidades de los otros sentidos ([135]). Por ejemplo, señales auditivas localizadas reducen el tiempo de respuesta de los objetivos visuales ([136–138]). Del mismo modo, la cantidad de eventos auditivos afectan la cantidad de eventos visuales percibidos que ocurren al mismo tiempo. Hasta las señales auditivas no informativas pueden mejorar la precisión del movimiento visual ([139]), demostrando el poder de los efectos perceptuales inter-sensoriales.

Las señales auditivas también amplifican o hasta sustituyen la información táctil y/o visual fuera de un limitado campo de visión. A través de dichas interacciones inter-sensoriales, las señales auditivas pueden transmitir información que pareciera comunicarse más naturalmente a través de otro canal sensorial.

Otros canales sensoriales, como la visión, se integran con la información espacial acústica para generar la percepción espacial ([111]). En particular, la información espacial auditiva se combina con la información espacial visual para conformar la percepción de un solo evento multisensorial, especialmente cuando diferentes entradas se correlacionan en el tiempo ([135, 140]). Cuando esto ocurre, la información espacial visual es mucho más fuerte que la auditiva, dominando así la posición percibida del evento ([140-142]).

El sistema espacial auditivo computa la posición de la fuente al combinar toda la información espacial acústica disponible. Sin embargo, conocimientos previos e información de otros canales sensoriales pueden tener un efecto directo en la percepción espacial de eventos auditivos y multisensoriales.

4. SIMULACIÓN ESPACIAL

Aunque que los displays gráficos son una elección obvia para mostrar información espacial a un espectador, el canal visual está generalmente sobrecargado. En estos casos, las señales de audio espacial pueden proveer información invaluable ([120, 143, 144]).

Se están desarrollando dispositivos de audio espacial para utilizarse en aplicaciones en donde la información visual no provee ningún beneficio, como por ejemplo, en aplicación con un campo de visión limitado ([145, 146]) o aplicaciones para presentar información a personas con disminución visual ([147–150]).

Como se mencionó, las señales espaciales permiten a los oyentes predecir la posición de una fuente de sonido. La forma más usual de simular estas señales de audio es mediante el uso de auriculares o altavoces. Las simulaciones que usan tanto auriculares como altavoces pueden variar en complejidad, desde no proveer información espacial hasta proveer casi todas las señales espaciales que ocurren naturalmente.

Los auriculares generalmente permiten un control más preciso de las señales espaciales presentadas al oyente, ya que las señales que llegan a los oídos pueden ser controladas independientemente y porque no llega sonido indirecto al oyente. Sin embargo, los auriculares pueden no ser prácticos para aquellas aplicaciones en al que los oyentes no desean utilizar hardware adicional sobre su cuerpo.

4.1. PROPIEDADES DE LOS DISPOSITIVOS DE AUDIO

Existen muchas propiedades de los dispositivos de audio que afectan a cualquier experiencia de RV. Estas propiedades deben considerarse a la hora de decidir si utilizar auriculares o altavoces.

En general, el sonido es creado y procesado con mucho menos poder de cómputo que los gráficos 3D. Sin embargo, es muy importante que el sonido sea computado rápidamente y entregado exactamente en sincronización con la información visual correspondiente.

Son muchos los aspectos que deben considerarse a la hora de elegir un dispositivo de audio para una aplicación de RV particular. Los aspectos más importantes están descritos a continuación.

4.1.1. CONTAMINACIÓN ACÚSTICA O RUIDO

En RV, los ruidos provenientes del mundo real pueden contaminar la experiencia. Por esta razón, es muy importante considerar el ambiente en el que la experiencia será presentada. Por ejemplo, si el sistema de RV utiliza altavoces, las voces de otras personas o sonidos del ambiente agregarían ruido a la experiencia, reduciendo considerablemente la inmersión.

Los sistemas que utilizan altavoces requieren un ambiente que sea razonablemente silencioso y sin eco. Al contrario, los auriculares permiten que los participantes escuchen sin oír el mundo exterior y sin que los demás escuchen el sonido.

4.1.2. MOVILIDAD DE LOS USUARIOS

La movilidad de los usuarios es algo a considerar a la hora de diseñar una experiencia de RV. Hoy en día existen varios tipos de auriculares inalámbricos que utilizan tecnologías como Bluetooth. Sin embargo, suelen ser más costosos que los normales por lo que no siempre son accesibles. Los auriculares con cable se conectarán al sistema y restringirían los movimientos del usuario de forma considerable, siempre y cuando el sistema sea estacionario.

Los altavoces, por otro lado, no restringen los movimientos de los participantes, pero el sonido se degrada a medida que el participante se aleja de ellos. Así, el participante debe estar dentro del rango audible de los altavoces. Sin embargo, existen técnicas que automáticamente ajustan la amplificación de los altavoces según la posición del participante.

4.2. SIMULACIÓN UTILIZANDO AURICULARES

Los auriculares se mueven junto con la cabeza del participante, son para uso personal y proveen un ambiente aislado. Se puede obstruir el mundo real al utilizar auriculares cerrados o permitir pasar el sonido del mundo real mediante la utilización de auriculares abiertos. Dado que los auriculares generalmente son *dual-channel* y se ubican cerca de cada oído, se pueden utilizar para reproducir el sonido 3D espacial de forma más fácil que utilizando altavoces.

Para poder simular cualquier fuente de sonido en algún lugar del espacio a través de auriculares, se debe reproducir una señal estéreo que recree en los tímpanos la forma de onda acústica exacta que llegaría desde una fuente en la posición deseada. Como los auriculares presentan sonido con respecto a la cabeza, cuando un sonido en un mundo virtual 3D debe parecer provenir de una determinada posición, es importante registrar la posición y orientación de la cabeza.

Este proceso no es simple, pero existen diversas técnicas para realizarlo. Algunas de éstas son simples de implementar pero no generan una representación espacial completa; otras son muy complejas y computacionalmente costosas, pero pueden generar una representación espacial cercana a la real.

4.2.1. DISPLAYS DIÓTICOS

La forma más simple de reproducir sonido a través de auriculares es presentando señales idénticas a ambos oídos. En estos casos, todas las fuentes de sonido son percibidas como si provinieran del centro de la cabeza. Mientras que de esta forma no se requiere ningún procesamiento auditivo especial, tampoco se provee información espacial al oyente.

Estos dispositivos serán útiles si la posición de la fuente de sonido no es conocida o no es importante. Así, los displays dióticos son los displays menos realistas. Cuando los oyentes requieren monitorear múltiples fuentes de sonido al mismo tiempo, los displays auditivos espacializados son claramente superiores a los displays dióticos ([103]).

4.2.2. DISPLAYS DICÓTICOS

En tanto la señal interaural normal varía con la frecuencia de forma compleja, las ITDs e ILDs independientes afectan la posición lateral percibida de una fuente de sonido ([151]). Las señales estéreo que solo contienen una ITD y/o ILD independientes se las conoce como señales dicóticas. Este término se utiliza comúnmente en los casos donde el sonido del oído izquierdo y derecho es diferente.

Las señales dicóticas resultan en fuentes de sonido que aparentan estar localizadas en una línea imaginaria dentro de la cabeza, contactando ambos oídos. Esta línea se llama el eje interaural (Figura 4.2). Por esta razón es que a veces a las fuentes de sonido ubicadas en esta línea se dice que están lateralizadas en lugar de localizadas.

A través de los auriculares se pueden generar las ITDs o ILDs constantes de manera muy simple. Solo se requiere que la señal sea retardada o escalada en cada oído. Así, la variación de ITD o la ILD causan que la posición lateral de la fuente de sonido percibida se mueva hacia el oído que reciba la señal más fuerte y/o más temprana.

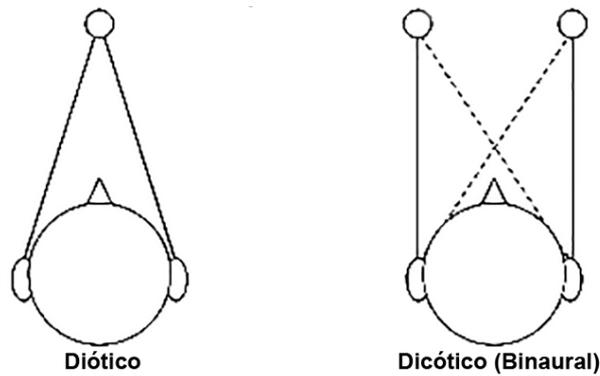


Figura 4.5: Representación de los modos de audio dióticos (la misma fuente de sonido a ambos oídos) y dicóticos (sonido binaural).

Estos displays dicóticos son fáciles de implementar pero sólo son útiles para indicar si la fuente de sonido se posiciona a la izquierda o a la derecha del oyente. La Figura 4.5 muestra la representación de señales dióticas y dicóticas llegando a los oídos del espectador.

4.2.3. FUNCIONES DE TRANSFERENCIAS EN RELACIÓN A LA CABEZA

Los seres humanos estiman la posición de las fuentes de sonido mediante la comparación de señales, como diferencias de tiempo o intensidad, recibidas en ambos oídos. El sonido original es modificado por muchos factores antes de entrar en el canal auditivo para ser procesado por el sistema auditivo. Esta modificación puede ser capturada por una respuesta de impulso en relación a la posición de la fuente de sonido y a la posición del oído. Esta respuesta de impulso es llamada **HRIR** (Head Related Impulse Response)³ ya que depende de las posiciones del sonido y el oído. La **HRTF** (Head Related Transfer Function)⁴ es la transformada de Fourier de la HRIR y describe cómo un sonido arribará al oído desde un punto particular en el espacio.

³Respuesta de Impulso en Relación a la Cabeza

⁴Función de Transferencia en Relación a la Cabeza

Para una persona particular, un sonido es reproducido desde muchos ángulos diferentes a su alrededor. Así, la HRTF es medida para cada posición de la fuente de sonido y describe cómo un sonido emitido desde un punto particular es escuchado por esa persona particular. Por esto, cualquier sonido proveniente de cualquier dirección puede ser simulado mediante el uso de la HRTF correspondiente. En los casos en que la HRTF no haya sido medida para un ángulo particular, se interpolan las HRTFs más próximas.

Como se ve en la Figura 4.6, la medición puede realizarse mediante el uso de pequeños altavoces ubicados en la sala y micrófonos miniatura ubicados dentro del canal auditivo del oyente. El altavoz es ubicado en un ángulo horizontal y vertical específico relativo a la cabeza y las funciones de transferencia o de impulso de los parlantes y micrófonos son medidas; a partir de los cuales se calculan las HRTFs y HRIRs. La medición es luego repetida para muchas otras posiciones alrededor de la cabeza y almacenada en un catálogo de HRTFs que será usado para espacializar el sonido en la posición espacial deseada. Finalmente, cuando un sonido es filtrado o convolucionado con la HRTF o HRIR y el resultado es reproducido a través de auriculares, parecerá como si el sonido proviniera del ángulo en que el parlante medido fue colocado.

Las HRTFs son diferentes para cada ángulo de incidencia y son específicas para cada persona individual ([152]). Debido a que dependen de las características anatómicas como el tamaño y forma de la cabeza y los oídos, las HRTFs varían considerablemente de persona a persona. Debido a este hecho es que emergieron dos enfoques distintos: HRTFs individualizadas, hechas para cada oyente a partir de sus características individuales, y HRTFs genéricas (o no individualizadas).

Dada la variabilidad en HRTFs entre distintos sujetos, se puede discutir el hecho de que todas las simulaciones auditivas deberían utilizar HRTFs individualizadas. Sin embargo, esto es extremadamente difícil de obtener en la práctica. Las grabaciones de HRTFs requieren mucho esfuerzo y son costosas, requieren salas sin eco, conjuntos de parlantes posicionados adecuadamente, micrófonos miniatura, software especializado y técnicos que manipulen el sistema.



Figura 4.6: Arreglo de altavoces al rededor del usuario.



Figura 4.7: Grabación de HRTFs mediante el uso de un maniquí.

La Figura 4.7 muestra una alternativa. Conjuntos genéricos pueden ser medidos en maniqués o sistemas de cabeza y torso equipados con oídos artificiales diseñados para aproximar lo mejor posible un humano “promedio”. Hoy en día, existen varias bases de datos de HRTFs de cabezas ficticias. Además, se están desarrollando métodos prometedoras para medir rápidamente las HRTFs individuales ([153]).

4.3. SIMULACIÓN UTILIZANDO ALTAVOCES

La naturaleza estacionaria de los altavoces hace que el sonido que producen sea con referencia al mundo, lo que es ventajoso, ya que el sonido con referencia al mundo es generalmente preferido en sistemas de RV. Sin embargo, la creación de sonido espacializado mediante el uso de altavoces puede ser más difícil que con auriculares.

La señal acústica total que llega a cada oído es simplemente la suma de las señales que llegan al oído desde cada fuente en el ambiente. Mediante esta propiedad, es posible variar las señales auditivas espaciales (como por ejemplo ITD y ILD) mediante el control de señales reproducidas a través de múltiples altavoces dispuestos alrededor del oyente.

A diferencia de las simulaciones con auriculares, la señal en ambos oídos no puede ser manipulada independientemente. Es decir, al cambiar la señal de cualquiera de los altavoces, cambiará la señal que llega a ambos oídos. Como resultado, es difícil controlar precisamente las señales espectrales y las diferencias interaurales de la señal que llegan al oyente. Esto es necesario para poder replicar las señales de una fuente de audio del mundo real.

Para reducir la variabilidad de las señales de audio que llegan a los oídos, se debe prestar especial atención a la colocación de los altavoces y a la acústica de la habitación. Los altavoces incorrectamente colocados pueden reducir dramáticamente la sensación de inmersión ([154]), especialmente al tratar con habitaciones pequeñas. Más aún, la colocación y la cantidad de altavoces variará también dependiendo de las dimensiones y forma de la habitación. Si el sistema de RV es portátil, el sistema de sonido deberá ser reajustado para cada nueva locación.

4.3.1. DISPLAY NO ESPACIAL

Muchos sistemas utilizan altavoces en donde cada altavoz presenta una señal idéntica. Dichos sistemas son análogos a los sistemas de auriculares dióticos; aunque son fáciles de implementar, estos displays no proveen información espacial al oyente. Estos sistemas pueden ser utilizados cuando la información espacial no importa o cuando la superposición de múltiples señales auditivas no es crítica. Por ejemplo, si el único sonido de interés se encuentra dentro del campo visual y la interferencia entre sonidos no es relevante, este tipo de display puede ser adecuado.

4.3.2. DISPLAY ESTÉREO

En este caso, las señales desde dos altavoces son presentadas simultáneamente para poder controlar la lateralidad percibida de una fuente de sonido virtual. La variación de las señales causa que la ITD y la ILD varíen sistemáticamente, produciendo que la fuente de sonido virtual varíe su posición sistemáticamente con respecto al nivel de los dos altavoces ([155]).

Cuando solo el altavoz izquierdo se reproduce, el oyente escucha la fuente de sonido en la dirección izquierda. Análogamente, cuando solo el altavoz derecho se reproduce, el oyente escucha la fuente de sonido en dirección a la derecha. De esta forma, con solo variar el nivel de las señales idénticas reproducidas en un par de altavoces, se puede alterar la lateralidad percibida de la fuente de sonido virtual.

Esta técnica es simple y produce una percepción robusta de una fuente de sonido en diferentes posiciones laterales. Sin embargo, es casi imposible controlar precisamente la posición exacta de la fuente de sonido. Además, cuando el oyente se mueve fuera del área restringida, la simulación se degrada dramáticamente. El eco también puede distorsionar las señales interaurales.

No obstante, estos sistemas proveen información sobre la lateralidad de la fuente de sonido y pueden ser muy efectivos cuando se desea simular sonidos en posiciones dentro del área comprendida por los altavoces. La mayoría de los altavoces estéreo comerciales están basados en variaciones de esta técnica.

4.3.3. ALTAVOCES MULTICANAL

Los sistemas de audio con altavoces multicanal pueden generar una sensación espacial natural sobre una determinada área, llamada "sweet spot". El tamaño de esta área depende del número de canales de audio utilizados. Actualmente, un sistema con 5 o 7 canales se refiere como audio envolvente (surround). Estos sistemas han sido parte de varias tecnologías, por ejemplo, en los cines. También han sido utilizadas satisfactoriamente en aplicaciones de realidad aumentada, como en el caso de sistemas de sensibilización del tráfico de vehículos ([155]), y son generalmente una parte integral de aplicaciones de música o videojuegos.

Existen otras tecnologías que proveen un mejor sonido espacial, como en 10.2-channel ([156]), Vector Based Amplitude Panning (VBAP) ([157]), Ambisonics ([158]) o Wave Field Synthesis (WFS) ([159, 160]). WFS puede crear una sensación espacial correcta sobre una gran área al utilizar grandes parlantes rodeando dicha área (típicamente más de 100 canales). Actualmente WFS se ha combinado con principios basados en objetos, en donde el campo de sonido deseado es sintetizado por señales de entrada separadas que representan el sonido de los objetos e información sobre la acústica de la sala ([161]).

La técnica VBAP anteriormente mencionada puede ser usada para colocar una fuente de sonido virtual en un espacio 3D con altavoces rodeando al oyente ([157]). Una terna de altavoces pueden usarse con amplitudes particulares para poder crear la imagen de un sonido dentro del triángulo formado por ellos. Grandes áreas alrededor del oyente pueden ser cubiertas con varias de estas ternas.

Ambisonics es la presentación de sonido 3D espacializado mediante la utilización de múltiples altavoces estacionarios. Las investigaciones en Ambisonics continúan avanzando y podrían resultar en un sistema usable ([158]), pero el hecho de que ambos oídos puedan escuchar el sonido de cada altavoz hace que esto sea difícil de lograr.

5. REVISIÓN DE LA LITERATURA

Como se mencionó anteriormente, maximizar la sensación de presencia e inmersión es una característica deseada en todo tipo de sistemas de RV. En este sentido, el audio juega un rol fundamental ya que compone una gran parte de la percepción humana de la realidad.

Los sonidos no pre-grabados generados en tiempo real han sido utilizados para mejorar todo tipo de ambientes virtuales e incluso videojuegos 3D. Por ejemplo, se ha estudiado el sonido de los pasos que acompañan al movimiento del usuario, así como también a otros humanos virtuales. Nordahl et al. ([162]) propusieron un sistema que permite la síntesis en tiempo real del sonido de pasos sobre diferentes materiales. Testearon la habilidad de los participantes de reconocer la superficie a la que estaban expuestos. También trataron de mejorar el sistema con sonido ambiente. Este experimento indicó la habilidad de los participantes de reconocer las superficies virtualmente simuladas, al mismo tiempo de la importancia de agregar sonido ambiente para mejorar el reconocimiento y la presencia.

Como el sonido es fundamental para generar una sensación de presencia e inmersión en cualquier sistema de RV, la percepción del sonido en el espacio ha sido un tema de investigación desde los inicios de la RV. El audio puede combinarse con otros sentidos y existen trabajos que aprovechan esta propiedad para mejorar la sensación de presencia e inmersión en sus aplicaciones.

La RV ha sido utilizada en diversos campos de la medicina mediante la integración de múltiples sentidos al mismo tiempo. Suied et al., por ejemplo, utilizaron RV como una herramienta de diagnóstico y terapia para cinofobia ([163]). En su aplicación, la información auditiva no es usada solo para complementar la información visual. Debido a que pueden surgir fuertes reacciones emocionales mediante la audición ([164]), ellos aprovecharon el potencial del audio 3D para incrementar el realismo y la riqueza del ambiente inmersivo, teniendo en cuenta también cómo diferente información sensorial debería ser combinada para activar reacciones efectivas en ambientes virtuales.

No solo se ha combinado la percepción del audio con la percepción visual. En [165] la realimentación auditiva con háptica fue combinada para mejorar tareas de navegación. En este trabajo no se utilizaron señales visuales de ningún tipo.

Taffour et al. ([166]) realizaron una evaluación con participantes con miedo a las multitudes para explorar la influencia de ambientes virtuales auditivos y visuales en las reacciones afectivas. Utilizaron simulaciones visuales de alta fidelidad y sonido 3D interactivo. Su estudio sugiere que la manipulación de la presentación sensorial junto a la posición espacial de la estimulación puede proveer una forma de modular las reacciones afectivas.

Bronkhorst y Houtgast investigaron el efecto del eco en la percepción del sonido en ambientes con eco usando acústica virtual ([118]). A los oyentes se les presentó un sonido y se les pidió que calificaran la distancia aparente en una escala cuantificada. Su experimento demostró que la distancia percibida es determinada por la distancia a la fuente de sonido virtual, la cantidad de reflejos y el nivel relativo de los mismos. Además, presentan un modelo matemático que expresa la distancia auditiva percibida a una fuente de sonido como una función de la relación entre las energías directas e indirectas.

En [167] se examinó la fidelidad visual de la percepción y la eficiencia de las tareas en la presencia de varias señales de sonido espacial. Se incluyó eco, oclusión y efectos de difracción, mientras los participantes completaban una simple tarea dentro de un ambiente virtual.

Bonneel et al. ([168]) examinaron la influencia del nivel de detalle en estímulos auditivos y visuales en la percepción de la calidad del material auditivo-visual. Observaron que la percepción del nivel de detalle visual se incrementa cuando el nivel de detalle auditivo se incrementa.

En [169] mostraron que los efectos de sonido hacen que las animaciones lentas se vean más suaves que las rápidas. Además mostraron que el agregado de sonido de pasos a las animaciones de caminar incrementan la suavidad de la percepción de la animación.

Rojas et al. ([170]) examinaron la percepción de la fidelidad visual definida con respecto a la cantidad de polígonos y la resolución de la textura en ambientes 3D estereoscópicos y no estereoscópicos bajo varias condiciones ambientales auditivas.

Muchos trabajos han investigado sobre la simulación de audio espacial y las distintas formas en que los oyentes pueden construir la imagen mental de los sonidos simulados. Existen estudios que han investigado diversos métodos para entrenar personas con disminución visual en campos de sonido simulados.

Shimizu et al. ([171]) desarrollaron un sistema de representación 2D mediante patrones, utilizando ILD y la frecuencia del sonido. En su sistema, la posición del sonido horizontal era controlada por el ILD y la posición vertical por la frecuencia.

En [172] se desarrolla un sistema de navegación en RV usando audio espacial. Reportaron que los sujetos con disminución visual podrían usar su sistema para localizar sonidos satisfactoriamente y recordar la escena 3D en pocas semanas.

González et al. ([173]) desarrollaron un sistema de percepción espacial utilizando un sistema de RV acústico que consistía de dos estaciones para capturar un objeto con cámaras estereoscópicas para simulaciones de HRTF. Su sistema puede representar el objeto capturado por las cámaras como un sonido 3D. Reportaron que la persona con disminución visual podía moverse dentro de la sala experimental utilizando su sistema sin depender del tacto.

Krueger et al. ([174]) desarrollaron un mapa de RV acústico, llamado “KnowWare” para que usuarios con disminución visual puedan interactuar con la computadora de manera sencilla. Sánchez et al. ([175]) también desarrollaron sistemas de RV acústicos, llamados “AudioVida” y “AudioChile”, para la educación de niños con disminución visual. En [176] desarrollaron un juego de RV auditivo llamado “MentalMapper” que ayuda a generar una mejor representación mental de mapas cognitivos a personas con disminución visual.

Seki et al. ([177]) desarrollaron un sistema de entrenamiento de orientación y movilidad para personas con disminución visual utilizando RV acústica. Sus resultados sugieren que el entrenamiento virtual puede reducir el efecto de desviación más que el entrenamiento real y también reduce el estrés tanto como el entrenamiento real.

La utilización de auriculares ha sido la forma más popular de representar sonido espacial en aplicaciones de RV. La técnica de sonido diótico es comúnmente utilizada para simular sonidos ambiente ya que se suele asumir que el sonido ambiente no proviene de ninguna posición particular. Nordahl et al. ([162]) utilizan sonido ambiental para mejorar la sensación de presencia en tareas de navegación. En [178] investigaron la forma de aliviar los síntomas del estrés mediante el uso de sonido de naturaleza en un bosque virtual.

En cuanto a los altavoces, las técnicas de implementación mencionadas sufren del hecho de que se necesitan muchos altavoces para procesar una gran cantidad de señales. Para instalaciones de RV con displays envolventes, un problema común es el conflicto entre la disposición de los altavoces y las pantallas de video. No se debe bloquear la vista a la pantalla de video con los altavoces, ni tampoco la pantalla de video debe bloquear el sonido de los altavoces. Existen pantallas de video acústicamente transparentes que solucionan este problema, pero la resolución de la imagen es significativamente menor.

Las HRTFs han sido un tópico de interés en muchas investigaciones relacionadas al audio para RV. Schissler et al. ([179]) presentaron una técnica de renderizado de audio espacial para manipular fuentes de sonido que pueden ser representadas por áreas o volúmenes en ambientes virtuales. Ellos utilizaron HRTFs individualizadas y obtuvieron una técnica eficiente cuya complejidad computacional y requerimientos de memoria son independientes de la complejidad de la fuente de sonido.

Existen muchos esfuerzos adicionales para minimizar las diferencias entre HRTFs genéricas e individualizadas ([180–182]) y se han propuesto algunos modelos mixtos ([183]). Sin embargo, aún se debaten los méritos y *trade-offs* entre HRTFs genéricas e individualizadas.

Existen trabajos que sugieren que la percepción espacial del sonido con HRTFs no individualizadas podría cambiar con respecto al tiempo. Begault y Wenzel ([120]) observaron varias diferencias individuales, lo que sugiere que algunos oyentes pudieron adaptarse más fácilmente que otros a las señales espectrales de las HRTFs no individualizadas.

Mendoncca et al. ([184]) observaron que la mera exposición a sonidos virtuales procesados con HRTFs genéricas no mejoró el desempeño de los participantes en la localización del sonido, pero los entrenamientos cortos generaron aprendizaje activo y la realimentación llevó a resultados significativamente mejores. Ellos concluyeron que la simple exposición no es suficiente para una evolución significativa en cortos períodos de tiempo. Sin ningún entrenamiento, es muy difícil localizar muchas fuentes de sonido. Por esto, ellos argumentan que testear sonidos virtuales procesados a través de HRTFs no individualizadas debería considerar siempre un período de aprendizaje y de adaptación.

Existen indicaciones de que los efectos de HRTFs individualizadas difieren de las genéricas de forma notoria ([185, 186]). Un incremento significativo de la sensación de presencia se produce cuando los sonidos virtuales son procesados con filtros individualizados, en lugar de genéricos ([187]). En un estudio que compararon la audición en la vida real con grabaciones realizadas en sujetos reales y en cabezas artificiales, se encontró que la precisión en la localización con grabaciones es peor que en la vida real y que las de cabezas artificiales son peores que las de los sujetos reales ([188]).

Cuando se usan las HRTFs genéricas, se produce un efecto en donde parece que el sonido no proviene de afuera, sino de adentro de la cabeza. Otro problema es el alto grado de percepción inversa en las posiciones espaciales de las fuentes de sonido virtual, donde las señales binaurales son ambiguas (en el cono de confusión) ([189]). Se cree que estos problemas se reducen cuando se utilizan HRTFs individualizadas y tracking de la cabeza ([180]).

Actualmente se ha vuelto popular el utilizar información antropométrica (medidas del oído y movimientos de la cabeza) para elegir alguna HRTF “personalizada” de un catálogo de muchos individuos ([190]). Sin embargo, como el sistema auditivo expresa gran plasticidad en el dominio de localización espacial, una persona puede adaptarse a la localización con alguna HRTF genérica. Este proceso se puede pensar como re-aprender a escuchar con un oído distinto, como se muestra en [132]. Esta habilidad natural para adaptarse a nuevas HRTFs especialmente modificadas se puede implementar para mejorar el desempeño de la localización y por ejemplo, reducir problemas de confusiones ([191]).

6. DISCUSIÓN Y CONCLUSIONES

6.1. PERCEPCIÓN Y ESTADO PREVIO

En este Capítulo se analizó la importancia del audio en nuestra vida diaria y la forma de conseguir una representación auditiva confiable en RV. Las personas utilizan las señales auditivas para la localización espacial, incrementando así la percepción de los alrededores, mejorando la atención visual y reuniendo una variedad de información compleja sin sobrecargar el sistema visual. Cuando se trabaja con interacciones multimodales, el audio puede mejorar la percepción visual, mejorando así la sensación de presencia en todo el sistema. Por esto, se requiere una buena correspondencia entre el sistema visual y el auditivo para lograr una representación realista en RV.

La percepción espacial auditiva está también determinada por el estado del oyente. Por ejemplo, un oyente familiar con el tenis tendrá un mejor conocimiento sobre el sonido de la raqueta golpeando la bola. Si el sonido no está apropiadamente representado, la sensación de presencia será reducida. Por otro lado, un oyente que nunca ha jugado al tenis podría no detectar esta diferencia. Aunque el sistema auditivo computa la posición de la fuente de sonido mediante la combinación de toda la información acústica disponible, el conocimiento previo y la información de otros canales sensoriales tendrán un efecto directo en la percepción espacial.

6.2. DISEÑO Y SOFTWARE

Como se mencionó anteriormente, una implementación de audio apropiadamente diseñada incrementa la sensación promedio de presencia en una simulación o RV. Si tenemos la posibilidad de crear nuestros propios sonidos, las señales espectrales del sonido, como la amplitud, frecuencia, forma, dirección y tiempo de llegada, deberían ser consideradas. La inclusión de eco puede incrementar significativamente el realismo percibido en RV, mejorando así la percepción de la distancia.

Para aquellas aplicaciones en las que la dirección o el realismo de la fuente de sonido es más importante que la distancia, no se recomienda implementar el eco. Por el otro lado, es importante notar que un sonido muy realista suele ser percibido como emocionalmente plano. Los efectos de sonido suelen ser diseñados para exagerar la versión de la realidad para transmitir emoción o satisfacer las expectativas de los usuarios. El diseño del sonido para RV necesita entonces balancear una reproducción precisa con la necesidad de involucrar emocionalmente al usuario.

Cuando la información auditiva espacial no es importante, los displays dióticos deberían ser utilizados. En estos casos, el mismo sonido es presentado a ambos oídos. Esta técnica se utiliza generalmente para sonido ambiental, como el sonido de los ambientes naturales en un bosque o en la playa, o para sonidos ambientales ruidosos como los de una ciudad.

Los displays dichóticos pueden usarse cuando una espacialización del audio simple es requerida. Mediante la alteración de la ILD, la ITD y las señales de intensidad, el oyente puede percibir un movimiento lateral y un cambio de la distancia del sonido. Aunque esta representación es muy limitada, la combinación con señales visuales puede mejorar bastante la percepción espacial del sonido. Además, esta técnica es muy fácil de implementar comparada con las HRTFs.

Las HRTFs pueden ser utilizadas para mostrar sonido binaural que parezca provenir de diferentes puntos en el espacio. La desventaja principal es que las HRTFs son específicas para cada persona individual y el proceso de medición lleva mucho esfuerzo y es costoso. Por esta razón, las HRTFs no individualizadas están siendo utilizadas con más frecuencia. Sin embargo, el uso de las HRTFs no individualizadas presentan efectos de localización negativos, como la confusión del sonido que está delante con el de atrás (y viceversa) o errores en la elevación del sonido.

Como se mencionó anteriormente, el sistema auditivo expresa una gran plasticidad en el dominio de localización espacial. Los oyentes pueden aprender a adaptarse a señales espectrales que no son naturales cuando se exponen a suficiente entrenamiento a largo plazo. Para aplicaciones en las que los oyentes deben ser entrenados, puede que no sea necesaria una simulación perfecta de las señales espaciales. Incluso si la percepción del sonido espacial con las HRTF no individualizadas es inferior a las individualizadas, esto puede cambiar con el paso del tiempo y cualquier persona podría adaptarse a la localización usando algún catálogo de HRTFs genéricas.

6.3. HARDWARE

Cuando se elije un display auditivo para utilizar en un sistema de RV, varios aspectos deben ser considerados. El ambiente real en donde la experiencia se realiza es importante. Como el usuario debe estar inmerso en el mundo virtual, el sonido del mundo real no debería ser escuchado. La movilidad del usuario también es otro elemento importante a considerar. De hecho, la mayoría de las aplicaciones de RV, la posibilidad de que el usuario pueda moverse es un requerimiento.

Como los auriculares son generalmente *dual-channel* y se ubican cerca de los oídos, es más fácil mostrar sonido espacializado a través de ellos. Los auriculares se mueven con la cabeza del participante, son para una sola persona y proveen un ambiente aislado.

La habilidad de generar una simulación espacial precisa utilizando altavoces se incrementa dramáticamente con el número de altavoces utilizados. Con un número infinito de altavoces alrededor del oyente, simplemente se podría reproducir la señal deseada desde el altavoz de la posición deseada para lograr una reproducción perfecta. Sin embargo, es importante notar que al agregar altavoces adicionales se pueden generar efectos negativos en la sensación de inmersión, especialmente en habitaciones pequeñas. Por esta razón, se debe tener especial cuidado con la acústica de la habitación, el diseño y ubicación de los altavoces y la posición del oyente dentro de la sala.

En general, los auriculares están asociados con displays visuales para la cabeza, mientras que los altavoces lo están con los displays de proyección. Los auriculares pueden combinarse fácilmente con los displays visuales de la cabeza de tal forma que los sentidos visuales y auditivos sean controlados por un solo dispositivo. Además, ambos HMDs y auriculares están asociados con la vista y el audio privado, mientras que los displays de proyección y los altavoces son adecuados para las presentaciones grupales. En general, siempre que se requiere audio espacializado, los auriculares suelen ser la mejor opción.

A veces, la experiencia de RV está diseñada para ser mostrada a muchas personas. En estos casos, cada persona requeriría un par de auriculares. Si hay más personas que auriculares disponibles, entonces éstos deberían ser compartidos entre todos los oyentes. Sin embargo, si se utilizaran altavoces, todos podrían escuchar de manera más fácil. Así, para grandes multitudes, los altavoces son la mejor opción.

En cuanto al costo, un par de auriculares suele ser menos caro que unos altavoces de una calidad equivalente. Esto se debe a que los altavoces suelen requerir amplificadores más poderosos y costosos. En los casos donde cada participante requiere un par de auriculares, es importante considerar la cantidad de participantes y el costo de la respectiva cantidad de auriculares, contra el costo de los altavoces.

6.4. RELACIÓN CON LOS OTROS SENTIDOS

La mayoría de los sistemas de RV tienen fuentes de sonido dinámicas que se mueven a través del ambiente virtual. Cualquier cambio en las señales espaciales de la fuente de sonido o del oyente también influye en la percepción de la posición de la fuente y ayuda a resolver ambigüedades. Ya sea por movimientos del oyente o del sonido, las señales de audio deben ser actualizadas de manera acorde para que el usuario perciba una representación auditiva adecuada. Para lograr esto, la rotación de la cabeza del usuario debe ser trackeada. De esta forma, al utilizar esta información, las señales de audio pueden ser actualizadas correctamente.

6.5. CONCLUSIÓN

En conclusión, es posible generar señales estéreo que contengan la mayoría de las señales espaciales normales disponibles en el mundo real. Es importante recalcar que para lograr la mejor representación virtual, las señales auditivas deben combinarse adecuadamente con las señales de los otros sentidos, maximizando así la sensación de presencia e inmersión. Utilizando técnicas de audio especializadas, una variedad de fuentes de sonido puede ser representada simultáneamente en diferentes direcciones y distancias. La mayoría de los sistemas de RV actualmente equipados con auriculares de suficiente calidad pueden reproducir audio espacializado de forma relativamente fácil. Si se combina esto con un dispositivo de head-tracking, el audio espacializado puede proveer una interfaz auditiva virtual verdadera. Finalmente, desde que las tecnologías de hardware y software han mejorando en los últimos años, los costos han bajado y el nivel de procesamiento ha aumentado, por lo que ahora es posible incluir audio espacializado en la mayoría de los sistemas de RV.

5

Olfato y Gusto en Realidad Virtual

1. INTRODUCCIÓN

La percepción humana del ambiente está influenciada significativamente por todas nuestras entradas sensoriales ([192, 193]). Entre estas entradas, el olfato es una de las más intrigantes, debido a su compleja naturaleza y el hecho de que su percepción por los humanos todavía no está completamente comprendida ([194]).

En los últimos años, Buck y Axel ([195]) lograron un mejor entendimiento de los procesos de reconocimiento y percepción de diferentes aromas. Sin embargo, un sistema de clasificación de aromas aún no existe. Como mencionó Kaye ([196, 197]), no existen aromas primarios como en el caso de los colores. Un gran número de sistemas de clasificación fueron propuestos por Ghinea y Ademoye ([198]), pero éstos se basan en bases industriales, como perfumes, comida o vino. Además, es sabido que los aromas son percibidos de forma distinta según la edad de la persona ([199]), su nacionalidad o cultura ([200]), su género ([201]), su estado de ánimo y sus experiencias en la vida ([198]).

Los usos potenciales y los beneficios de los sistemas olfativos están creciendo en cuanto a distintas aplicaciones en diversas industrias como entretenimiento ([202]), salud ([203]), educación ([204]) y turismo ([205]). Desde el punto de vista de la salud, la incorporación de aromas a un ambiente ha demostrado que reduce el estrés ([206]) y puede ser usado en diagnóstico y tratamiento de enfermedades. En cuanto a la educación, el olfato ha demostrado mejorar la memoria ([207, 208]) y mejorar el rendimiento en tareas de resolución de problemas como, por ejemplo, reducir el tiempo de respuesta ([209]). También se ha demostrado que el estímulo olfativo ayuda a mejorar la productividad, la alerta y rendimiento físico ([210]).

El poder reproducir aromas con la calidad e intensidad requeridas ha sido particularmente desafiante. Aunque existen varios dispositivos que emiten aromas, las mayores dificultades siguen siendo el control de la dirección del olor, la cantidad del aroma emitido, el tiempo que se tarda en emitir un aroma, el tiempo que el aroma permanece en el ambiente y la cantidad de diferentes aromas que pueden ser almacenados en el dispositivo emisor al mismo tiempo.

El gusto es el sentido más estrechamente relacionado con el olfato. Al igual que el olfato, éste es uno de los sentidos que puede activar zonas del cerebro que los otros sentidos no pueden, haciéndonos recordar y experimentar sensaciones de distinto tipo.

Como este sentido es uno de los sentidos denominados “de corto alcance”, existe un problema fundamental a la hora de simularlo en sistemas de IHC y RV. Para poder hacer esto, el usuario debería poder colocar algo en su boca, ya sea un dispositivo, un electrodo, alimentos, etc. Este tipo de intrusión suele ser extremadamente incómodo para los usuarios y además, reduce considerablemente la inmersión del sistema.

Igualmente, como se verá en la sección 4, existen trabajos de investigación que analizan el sentido del gusto, pero generalmente lo hacen mediante interacciones multimodales que involucran a los otros sentidos. Igualmente, con los avances de la tecnología, no se descarta que en los próximos años desarrollen sistemas cómodos y usables.

En la Sección 2 del presente Capítulo se describe brevemente el funcionamiento del sistema olfativo en los humanos. Luego, se presenta una clasificación de los tipos de displays olfativos existentes en la Sección 3. La Sección 4 presenta una revisión del estado del arte de los dispositivos olfativos existentes, haciendo hincapié en aquellos utilizados o que pueden utilizarse en sistemas de RV. Finalmente, la Sección 5 finaliza con una discusión y conclusión al respecto.

2. OLFATO HUMANO

El sentido del olfato es uno de los sentidos más subestimados del ser humano. Sin embargo, nos permite percibir el mundo de forma muy diferente a los otros sentidos, activando zonas de nuestro cerebro que nos recuerdan sensaciones que estaban ocultas. Esta Sección se enfoca en el sistema psicofísico olfativo y en cómo los humanos perciben los aromas.

2.1. SISTEMA PSICOFÍSICO OLFATIVO

El estudio psicofísico olfativo se refiere a cómo los humanos perciben los aromas y la sensación que los humanos identifican cuando las células receptoras olfativas son estimuladas. Esta sensación es la reacción psicofísica que resulta de la estimulación sensorial. Para que dicha sensación se produzca, deben presentarse cuatro condiciones. Primero, debe existir un estímulo en el ambiente. Segundo, un receptor sensorial debe convertir el estímulo en impulsos nerviosos. Tercero, estos impulsos nerviosos deben ser llevados desde los receptores hasta el cerebro. Finalmente, el cerebro debe producir la sensación según la recepción e integración de los impulsos nerviosos. La sensación de olor es el resultado de la estimulación de los receptores de olor, o quimio-receptores, por un estímulo físico-químico.

2.2. PERCEPCIÓN DEL OLOR

La percepción subjetiva del olor es un proceso complejo que usualmente se describe utilizando varias dimensiones. Agrado, familiaridad o intensidad, por ejemplo, son tres dimensiones usuales que suelen utilizarse para describir las cualidades subjetivas de un aroma ([211]). Algunas investigaciones han mostrado que cuando se presenta una gran variedad de aromas con una intensidad similar, el agrado es la dimensión más utilizada para la percepción ([212]). Por esto, el agrado es usualmente la propiedad más prominente en la experiencia olfativa y tiene la mayor influencia en el humor, donde los aromas placenteros inducen humores positivos y los aromas no placenteros inducen humores negativos ([213]).

3. TIPOS DE DISPLAY

Barfield y Danas definieron un display olfativo virtual como “una colección de hardware, software y químicos que pueden ser usados para presentar información olfativa al participante” ([214]). Los displays olfativos requieren un proceso en dos etapas. La primera es la creación del aroma, que implica la grabación o síntesis del aroma; y la segunda es el renderizado de un aroma listo para utilizar. Crear un aroma es un gran desafío y los grandes laboratorios de perfumes son los mejores en hacer esto. Sin embargo, debido a la limitación de la extracción de los aromas, la mayoría de ellos contiene muy poco porcentaje del elemento original. Existen varias tecnologías para almacenar aromas para displays olfativos, incluyendo líquidos, gel y ceras sólidas. Para liberar un aroma, estos métodos de almacenamiento requieren que el compuesto se vaporice usando métodos electrostáticos o calor.

La mayoría de los trabajos académicos se enfocan en la segunda etapa, es decir, en renderizar un aroma listo para utilizar. Este renderizado es la contrapartida de los HMDs para la vista, por ejemplo. Sin embargo, al contrario de los displays visuales, no se han desarrollado tantos displays olfativos hasta el momento.

Los displays olfativos (DOs) son dispositivos controlados por computadora que liberan aromas en el aire. Los desafíos principales son presentar el aroma con una calidad aceptable en términos de intensidad y realismo, y con una duración correcta. Murray et al. ([194]) presenta una clasificación de DOs según su ubicación y el modo de presentar aromas al usuario. A continuación daremos una breve descripción de cada una de estas clasificaciones.

3.1. DISPLAYS OLFATIVOS UBICADOS EN EL AMBIENTE

La mayor ventaja de un display ubicado en el ambiente es la no intrusión, es decir, que el usuario no se da cuenta de la existencia del display. La mayor desventaja es la susceptibilidad al movimiento del aire en el ambiente.

En los sistemas de Emisión Natural, los DOs emiten aromas al usuario mediante dispersión natural, directamente en el ambiente. El problema principal de este tipo de displays es la naturaleza lenta del movimiento del aroma a través del aire. Otro problema es que el aroma dependerá directamente de la dirección y movimiento del aire que exista en el ambiente. Esta clasificación es más adecuada para presentar aromas en un ambiente completo y no para aplicaciones que requieren intercambiar aromas en tiempos determinados. Por ejemplo, un sistema de este estilo sería útil para la representación de una cocina virtual, en donde se quiere que exista determinado aroma en toda la cocina.

Los ODs de Flujo de Aire entregan aromas al usuario mediante ventiladores, bombas de aire, aerosoles o sopladores. Estos sistemas tratan de mover el aire para poder llegar al campo olfativo del usuario, pero sin apuntar a él directamente. Su mayor desventaja es el poder mantener el aroma el tiempo requerido y en el lugar requerido.

Finalmente, los DOs de Cañones de Aire, aunque se basen en ventiladores, están diseñados para entregar un aroma de forma precisa. El uso de cañones de aire es otra idea para introducir aromas hacia el campo olfativo del usuario. Esto incluye el lanzamiento de anillos de esencia que mantienen su forma hasta que alcanzan la posición deseada. Los cañones de aire intentan resolver el problema de la posición del aroma, que resuelven siempre y cuando el usuario se encuentre a una distancia corta del cañón.

3.2. DISPLAYS OLFATIVOS USABLES

Los DOs usables tienen un gran potencial ya que se encuentran cerca del campo olfativo del usuario. Esto incluye tanto a aquellos sistemas que se encuentran en dispositivos que utiliza el usuario, como a los que el usuario utiliza como accesorio.

Como se verá luego en la Sección 4, existen distintos tipos de dispositivos que funcionan como accesorios. Por ejemplo, algunos sistemas están diseñados para ser utilizados como si fuesen joyería o prendas de vestir. Estos sistemas están pensados para emitir distintas fragancias en distintas ocasiones.

Los sistemas montados en la cabeza usualmente se utilizan en combinación con los HMDs para sistemas de RV. La gran ventaja de estos dispositivos es que están muy cerca del campo olfativo del usuario. El principal inconveniente es el grado de intrusión que pueden generar en el usuario.

4. REVISIÓN DE LA LITERATURA

Las primeras interfaces olfativas desarrolladas en el campo de la Interacción Humano Computadora utilizaban químicos para generar aromas ([215–219]). Estos sistemas son complejos, costosos y necesitan mucho mantenimiento para poder operar correctamente. Además, existen otras limitaciones como la dificultad de cambiar rápidamente entre aromas, la de controlar el área de estimulación a lo largo del tiempo y la de mantener una concentración uniforme. De esta forma, el poder generar sensaciones de aromas sin utilizar químicos se ha convertido en una necesidad de nuestro mundo digital. De este modo, las personas podrán experimentar sensaciones de aromas digitalmente a través de internet, como lo hacemos actualmente con sensaciones visuales y auditivas.

Desde la aparición del Sensorama ([220]), se ha realizado muy poco trabajo en cuanto a displays sensoriales que no sean visuales, auditivos o hápticos. En el Sensorama se experimentaba con aromas, exponiendo a los participantes a distintos aromas, incluyendo por ejemplo aromas a comida.

Los ambientes virtuales actuales están mayormente diseñados para estimular la modalidades visuales, auditivas y táctiles del humano. La información olfativa como input al participante generalmente es ignorada, aún cuando los receptores olfativos provean una fuente importante de información al humano. En términos de aplicaciones específicas, Cater ([221]) postula que para crear una gran sensación de presencia en ambientes virtuales realistas, se deberían agregar otras modalidades sensoriales más allá de la visión, la audición y el tacto. Lo que es más, como reportan en Robinett ([222]), los displays virtuales olfativos pueden ser útiles para incrementar el nivel de presencia en sitios remotos.

Existen sistemas basados en la emisión natural de los aromas. El sistema de emisión usado por Brkic et al. ([223]) consiste de un atomizador que expulsa una pequeña porción de esencia automáticamente a intervalos de tiempo definidos. Reportaron que la sensibilidad de los usuarios a la calidad de un video presentado se reducía con la incorporación de aromas.

Kim et al. ([224]) reportaron un DO que utilizaba calor para generar y emitir aromas. Pequeñas tarjetas de gel con aroma eran calentadas para difundir aromas. Reportaron que los usuarios tardaban aproximadamente siete segundos en detectar el aroma. Tan et al. ([225]) describieron un DO electromecánico como atomizador de botellas de perfume. Este tipo de sistemas es adecuado para aquellas aplicaciones en las que no se requiere controlar el flujo del aroma. Ariyakul et al. ([226]) desarrollaron un display olfativo compuesto de bombas electrosmóticas miniatura y un dispositivo de ondas acústicas para presentar aromas de forma rápida.

Para controlar el flujo del aroma, existen diversos sistemas de flujo de aire. Haque et al. ([227, 228]) desarrollaron un sistema interactivo que crea zonas olfativas dinámicas. Los aromas son emitidos según los movimientos del usuario, con un flujo de aire generado a través de un arreglo de ventiladores. Un sistema similar fue presentado por Emsenhuber y Ferscha ([229]).

Nambu et al. ([230]) presentaron un sistema visual-olfativo en donde la generación de aromas se basaba en bombas de aire. Estas bombas dirigían el aire con aroma hacia el campo olfativo del usuario. Discutieron la relación entre el contenido visual y olfativo y concluyeron que no es necesario disponer de muchos tipos de esencias a la vez, ya que los participantes asociaban la esencia directamente con el contenido visual.

Tanikawa et al. ([231]) presentaron dos displays olfativos-visuales: el *nioni cafe* y el *flavor of color*. Cada uno utilizaba un DO con tres componentes principales: un controlador, un mecanismo de generación de aroma y un mecanismo de presentación. A los participantes se les presentaban esencias en conflicto con lo que veían. Los resultados resaltaron la influencia de los estímulos visuales sobre los olfativos. Los aromas eran generados al hacer soplar aire a través de cartuchos de esencias.

Okada y su equipo de trabajo han realizado muchas investigaciones al respecto ([232–240]) y construyeron un emisor de aromas basado en tecnología de inyección. Aceites aromáticos son almacenados en pequeñas cámaras y son emitidos a través de pequeños ventiladores. Además se han investigado técnicas para la presentación de aromas para evitar la adaptación olfativa ([235]), expresar dos aromas al mismo tiempo ([234]), cambiar instantáneamente entre dos aromas ([232]) y medir las características temporales del olfato en una inhalación ([233]).

Como parte de un *Teatro de Realidad Virtual*, Park et al. ([241]) utilizaron cinco contenedores con esencias para generación de aromas. La vaporización del líquido fue realizada al utilizar platillos calentadores y la esencia se esparce mediante sistemas de ventilación.

Tominaga et al. ([242]) agregaron aromas a un sistema de RV de proyección utilizando un dispositivo de generación de aromas de Fukurara Co. Ltd. Este dispositivo se basaba en la compresión de aire y el almacenado líquido de esencias. El aire es bombeado a través del líquido hacia los usuarios.

GameSkunk de Sensory Acumen ([243]) es un dispositivo que está disponible comercialmente para el mercado médico. Este dispositivo emite aromas al soplar aire alrededor de substratos aromáticos sólidos. El dispositivo puede almacenar hasta 12 cartuchos, cada uno conteniendo una esencia.

Matsukura et. al. ([244]) desarrollaron un display que puede presentar aromas en posiciones arbitrarias de una pantalla. El dispositivo da la sensación de que un aroma emana de un punto particular de la pantalla. Se disponen ventiladores en las cuatro esquinas de la pantalla, de tal forma de poder manipular la corriente de aire hasta el punto de la pantalla deseado. Al introducir aromas en la corriente de aire, el aroma se esparce desde el punto virtual de la pantalla hacia el usuario. El usuario puede mover su cabeza libremente para sentir varios aromas en varias posiciones.

Actualmente, se han propuesto sistemas que disparan una porción de aire con aroma a la nariz del usuario, dirigida por un sistema de cámaras. Los *ART Media Information Science Laboratories* han creado un display olfativo controlable realmente cómodo ([245]). Este sistema es un sistema de proyección que utiliza un “cañón de aire” para presentar un cúmulo de aire con aroma en una posición específica cerca de la nariz del usuario. El sistema se compone de un tracker de nariz, una cañón de aire y un generador de aromas.

Este tipo de sistemas brinda a los usuarios la posibilidad de disfrutar aromas sin tener que usar dispositivos incómodos. El problema es que no son muy prácticos para grandes ambientes, o exteriores, debido al rango limitado y a la susceptibilidad hacia factores ambientales, como el viento.

Yanagida et al. ([245]) también han utilizado cañones de aire para “disparar” aromas hacia el usuario. Yanagida et al. ([246]) se enfocaron en entregar aromas a la nariz del usuario mediante cañones de aire. El cañón de aire crea vórtices de aire con olor, que mantienen su forma hasta llegar al campo olfativo del usuario. Reportaron una precisión del 72 % en cuanto a la detección de olores.

Tomono et al. ([247–249]) reportaron una propuesta para la presentación de aromas a través de una pantalla en dirección al usuario. Sakari et al. ([250]) presentaron un sistema de proyección basado en tracking de la nariz, centrado en el control espacio-temporal de aromas. Yanagida et al. ([245]) describen otro dispositivo basado en aros de vórtice que generan aromas a través de un cañón de aire. El sistema genera aromas mediante la vaporización de aceites esenciales.

Por otro lado, en los últimos años, la miniaturización de la tecnología ha permitido la inclusión de sistemas que pueden utilizarse sobre el cuerpo. Yamada et al. ([251]) presentaron un display con una unidad de generación de aromas compuesta de bombas de aire y filtros de esencias (algodón infiltrado de perfume). Un tubo proveniente del sistema (que se usa en la espalda) se conecta a una interfaz situada en la cara del usuario.



Figura 5.1: Usuario utilizando el sistema de generación de aromas. ([251])

Ferraro y Ugur ([252]) discuten la inclusión de DOs en diferentes materiales para facilitar experiencias multi sensoriales. Jenkins et al. ([253]) reportaron prototipos inspirados en insectos. Entre ellos incluyen joyas, ropa y bolsos, todos ellos utilizando Microfluidics, una tecnología que facilita la creación de dispositivos que soporten el manejo de pequeñas cantidades de fluidos. Además, se suelen utilizar también biosensores para proveer realimentación.

El sistema desarrollado por Yamada et al. ([251]) inyecta pequeñas gotas de fragancia directamente en la nariz del usuario o cerca de ella. El material aromático es almacenado en estado líquido, haciendo que el dispositivo sea pequeño. Además, el sistema dispone de una unidad que detecta respiración para poder reducir la cantidad de fragancia transmitida cuando el usuario exhala. Esto ahorra el consumo del material aromático, permitiendo que el display pueda ser usado por más tiempo. La Figura 5.1 muestra a un usuario utilizando este sistema, que se detalla en la Figura 5.2.

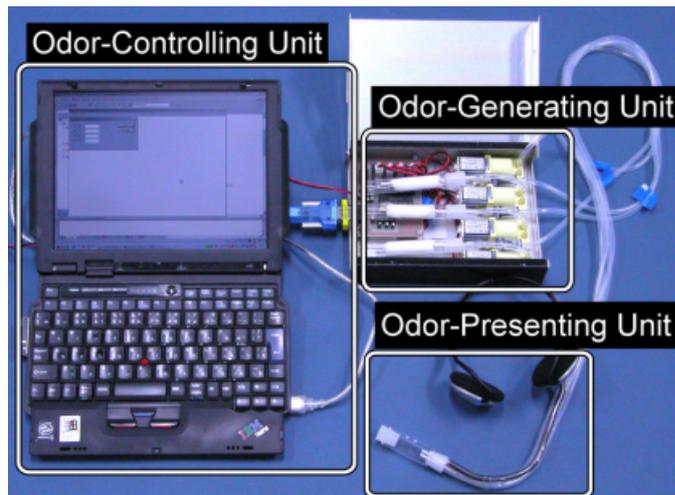


Figura 5.2: Sistema de generación de aromas. ([251])

Matsukura et al. ([254–257]) reportaron una técnica interesante para presentar aromas que considera la dinámica de fluidos del ambiente. La concentración de olor se generaba al utilizar un mezclador de olores que podía cambiar la concentración al añadir aire sin aroma. El vapor con olor era luego entregado a la nariz del usuario mediante una interfaz especial. En sus experimentos, el 65 % de los participantes pudieron encontrar la posición del aroma.

Narumi et al. ([258]) presentaron un dispositivo multisensorial que integraba visión, olfato y gestos. El objetivo de este trabajo era determinar la influencia de diferentes modalidades en el sabor. El DO podía presentar hasta seis aromas, a través de bombas de aire, cerca de la nariz del usuario.

En [259] presentaban aromas a través de inyección directa de pequeñas gotas de líquido dentro de las fosas nasales. Experimentaron con el conflicto entre el sistema visual y el olfativo de una galleta, donde los participantes comparaban comer la galleta con estimulación olfativa y sin ella. El 79 % de los participantes reportaron cambios en el sabor al aplicar estímulos olfativos.

Nakamoto et al. ([260]) desarrollaron un display olfativo capaz de mezclar aromas en tiempo real. El sistema mezcla varios componentes en una gran variedad de “recetas”, ofreciendo una gran variedad de aromas. Sus estudios revelaron que una escena con los aromas apropiados atrae e incrementa la atención de los usuarios. En este trabajo, realizaron una experiencia de cocina virtual, en donde el participante debía preparar curry.

Algunos trabajos han demostrado la fidelidad entre las señales visuales y las olfativas ([230]). En estos trabajos se demuestra que la inclusión de imágenes visuales incrementa la capacidad de los participantes de percibir distintos tipos de aromas. Nishizawa et al. ([261]) propusieron un sistema de realidad aumentada con proyección para modificar la apariencia de comida en tiempo real. Realizaron un experimento psicofísico para investigar si la evaluación del gusto puede ser modificada por el color de la comida en una porción de torta. Sus resultados mostraron que la dulzura de la torta se incrementa según la cantidad de saturación del color y que el sabor cambia según el matiz.

Narumi et al. ([262]) desarrollaron “MetaCookie+”, un sistema que le permite al usuario experimentar varios sabores sin agregar químicos a las comidas, mediante la influencia de las distintas modalidades. El sistema puede verse en la Figura 5.3 y cambia el sabor percibido de una galleta al superponer la información visual y olfativa en una galleta real con un marcador especial. El usuario puede elegir entre distintos tipos de galletas, como se ve en la Figura 5.4. Luego, la apariencia y aroma de la galleta seleccionada se superpone a la galleta real. La aumentación realista visual y olfativa crea efectos inter-modales entre la visión, el olfato y el gusto, haciendo que los usuarios experimenten el sabor seleccionado en su galleta simple. En su experimento, el 72.6 % de los participantes reportaron sentir el sabor de la galleta seleccionada. Los resultados del experimento sugieren que el sistema puede alterar el sabor percibido.



Figura 5.3: Usuario utilizando el sistema MetaCookie+. ([262])



Figura 5.4: Usuario seleccionando el tipo de galleta virtual. ([262])

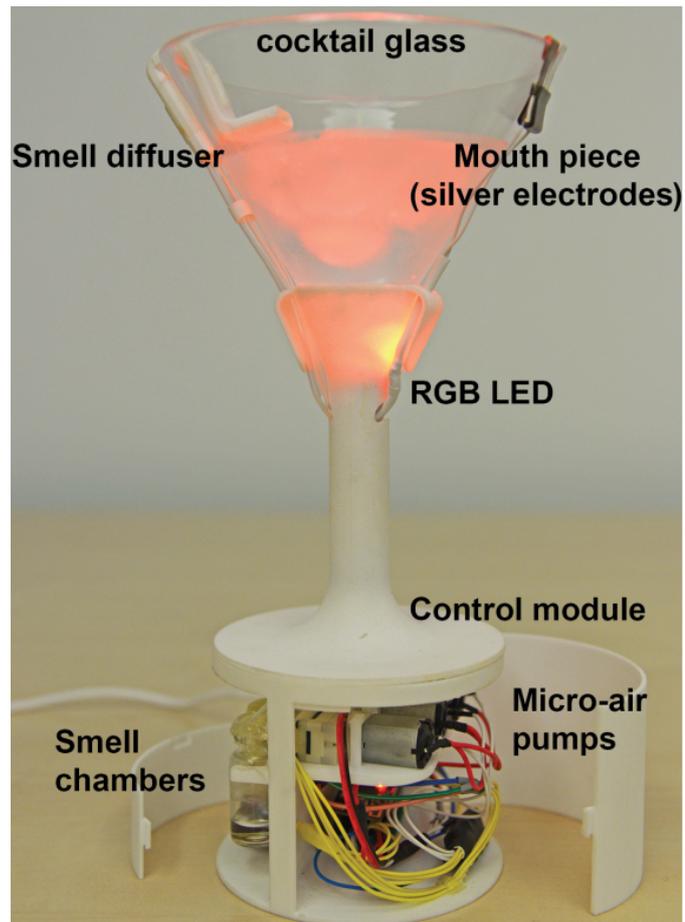


Figura 5.5: Diferentes componentes del sistema Vocktail. Cuando el usuario bebe, impulsos eléctricos controlados son aplicados a través de electrodos para simular sabores. Los olores también son expulsados cerca de la nariz del usuario, mientras una luz colorida se proyecta en la bebida. ([263])

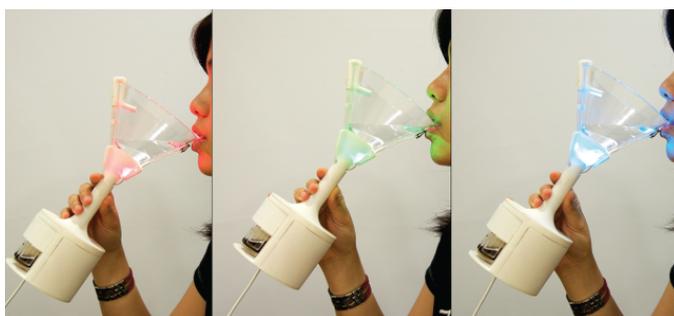


Figura 5.6: Un participante interactuando con el sistema Vocktail. ([263])

En [263] presentan “Vocktail”, un utensilio de bebidas interactivo que digitalmente simula experiencias de sabor multisensorial. Este sistema puede verse en la Figura 5.5. El sistema utiliza tres modalidades, sabor, olor y color, para crear sabores virtuales en la bebida. La Figura 5.6 muestra un usuario interactuando con el sistema. Cuando el usuario bebe del sistema, una luz de determinado color se proyecta sobre la bebida, una estimulación eléctrica se presenta sobre la punta de la lengua del usuario y un estímulo olfativo se presenta al usuario. La combinación de estas tres cosas crean la sensación virtual de sabor.

El “Project Nourished” ([264]) es un sistema similar que combina un HMD, un difusor aromático, auriculares de conducción por hueso, utensilios virtuales y comida impresa en 3D, para simular cualquier tipo de alimento. El sistema “Ohroma” ([265]) también promete hacer sentir distintos aromas mientras las personas están comunicadas a través de telepresencia. Estos sistemas están comercialmente disponibles, pero no existen resultados formales sobre su funcionamiento.

Hariri et al. ([266]) proponen una interfaz para activar sensaciones aromáticas mediante la estimulación de los receptores nasales utilizando pulsos eléctricos. Su trabajo todavía es un prototipo y aún no han realizado experimentos representativos.

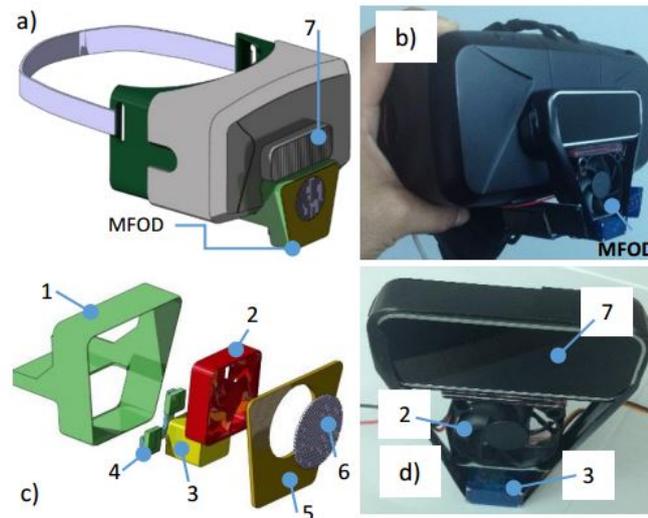


Figura 5.7: (a) y (b) Oculus Rift con el MFOD y Leap Motion integrados. (c) Componentes del MFOD. (d) Detalles del prototipo real. ([267])

En Covarrubias et al. ([267]), se muestra el desarrollo de un sistema de RV para rehabilitación de las extremidades superiores del cuerpo, que está complementado con la estimulación de aromas. El sistema utiliza un *Multi-Fragrance Olfactory Display (MFOD)* el cual entrega aromas a través de la modulación del aire que pasa a través de tabletas de fragancias sólidas. Esto reduce la complejidad comparado con los sistemas de fragancias líquidas o vaporización. El sistema puede verse en la Figura 5.7, unido a un Oculus Rift. En la Figura 5.8 puede verse la variación de la fragancia según la rotación del servomotor integrado al sistema. El servomotor hace rotar las tabletas aromáticas de tal forma de seleccionar aquella cuyo aroma pasará por el pequeño ventilador, llegando a la nariz del usuario. Reportaron resultados muy positivos en cuanto a la sensación de presencia y disfrute que generó la incorporación de aromas al sistema.

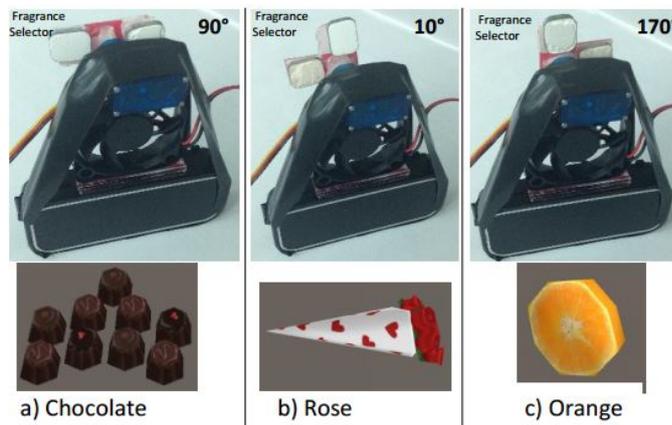


Figura 5.8: Variaciones de las fragancias según la rotación de un servomotor. ([267])

En [268] discuten las ventajas de utilizar este tipo de tecnologías para el tratamiento de desórdenes de estrés post-traumático. En [269] realizaron un experimento para analizar si la incorporación de aromas placenteros y no placenteros mejoran la sensación de presencia en sistemas de RV. Su experimento sugiere que la inclusión de aromas no placenteros incrementa considerablemente el nivel de presencia.

Hashimoto et al. ([270] desarrollaron un display olfativo portátil utilizando un dispositivo de ondas acústicas de superficie y micro-bombas. En su experimento, confirmaron que el display olfativo puede presentar olores con la intensidad deseada.

Otras aplicaciones potenciales incluyen el entrenamiento para el reconocimiento de materiales peligrosos, la identificación de habitaciones por el aroma y la demarcación de caminos de salida de emergencia ([221]). Los displays olfativos integrados con los ambientes virtuales pueden usarse también para la visualización de procesos dinámicos (por ejemplo, reacciones químicas), asistir en el diagnóstico de enfermedades y para recreación, por ejemplo, al incrementar las experiencias sensoriales de una película o videojuego ([271]). Además, la información olfativa puede utilizarse para proveer señales de dirección y distancia de un objeto.

5. DISCUSIÓN Y CONCLUSIONES

Desde el surgimiento del “Sensorama” que fue el primer sistema en proveer estímulos olfativos al usuario, se han desarrollado varios displays olfativos en las últimas décadas ([221]). Hoy en día, existen distintos tipos de interfaces olfativas pensadas para RV, como el *head-mounted olfactory display* ([272]) o el *arm-worn olfactory display* ([273]). Sin embargo, todavía existen algunos problemas, como mejorar la generación de aromas y la detección precisa de aromas en el ambiente. Estos problemas deben solucionarse para poder crear un sistema cómodo y fácil de utilizar.

Aún no se ha desarrollado un sistema comercial que pueda recrear aromas de forma precisa en cuanto a intensidad, calidad, cantidad y presentación. Existe una gran variedad de técnicas a la hora de crear un DO, con una gran variedad de mecanismos de generación de aromas, técnicas de presentación y posicionamiento. El poder lograr un dispositivo de forma no intrusiva es todo un desafío.

En cuanto a los tipos de display, cada una de las clasificaciones tiene sus ventajas y desventajas. Los DOs ubicados en el ambiente tienen la ventaja de no ser intrusivos. El problema es que dependen de las condiciones del ambiente y el flujo natural del aire. Esto no es problema para los dispositivos usables, ya que los aromas son presentados muy cerca del campo olfativo del usuario. Sin embargo, se debe tener en cuenta la intrusión que estos dispositivos pueden generar en el usuario.

La realización de un DO requiere una aproximación multidisciplinaria. Se requieren contribuciones de diversos campos como la química y la biotecnología. Esta claro que el hecho de no tener un esquema de aromas, como lo tienen los colores, es un impedimento clave. Hasta que esto se logre, el almacenaje de esencias será limitado. Por otro lado, por su precio elevado, los displays olfativos no son económicamente accesibles para el usuario común.

Antes de que los displays olfativos puedan integrarse a la RV, se deben realizar varias mejoras en la tecnología. Específicamente, los aromas deben cumplir requerimientos en cuanto a intensidad y duración. Además, los aromas deben poder intercambiarse rápidamente para generar una sensación realista. Más allá de la tecnología, las capacidades humanas también deben tenerse en cuenta. Nuestro vocabulario para describir aromas es pobre y conocemos muy poco sobre el continuo físico en el que tienen lugar los estímulos olfativos. La falta de conocimiento nos prohíbe desarrollar las bases para los displays olfativos, condiciones necesarias si queremos ofrecer a distintos individuos la misma experiencia olfativa virtual. Aún más, se deberían desarrollar estrategias para tratar con las diferencias individuales que existen en el reconocimiento de aromas. También debemos considerar que distintos aromas producen distintas emociones y reacciones físicas en personas distintas.

Para una sinopsis detallada sobre las interfaces olfativas desde el punto de vista del sensado y la presentación, puede verse Nakamoto ([274]). Ghinea y Ademoye ([198]) proveen una revisión sobre la percepción olfativa considerando la edad, el género, el estado de ánimo y la cultura. Ellos también discuten el uso de dispositivos olfativos en distintas industrias. Yanagida et al. ([259]) revisaron y categorizaron los diferentes tipos de tecnologías utilizadas en el desarrollo de displays olfativos modernos en término de generación y transporte de aromas.

En resumen, los displays olfativos virtuales prometen mucho a la hora de mejorar una experiencia de RV, aunque aún se debe realizar mucho trabajo de investigación en cómo utilizar el potencial de los displays olfativos.

6

Tacto en Realidad Virtual

1. INTRODUCCIÓN

El sentido del tacto se relaciona con una de las zonas más grandes e importantes del cuerpo humano. Sin embargo, es uno de los menos estudiados en sistemas de RV. Los sistemas existentes que incluyen el sentido del tacto se componen de grandes exoesqueletos en donde el usuario se sitúa para poder desplazarse en entornos virtuales y poder sentir determinadas fuerzas sobre su cuerpo. Este tipo de sistemas, además de grandes, complejos y costosos, son incómodos para la mayoría de los usuarios, lo que es contraproducente cuando pensamos en la inmersión del sistema.

Otro tipo de dispositivos utiliza vibraciones para simular distintas fuerzas sobre el cuerpo del usuario. Estos dispositivos están muy limitados en cuanto al nivel de realismo que pueden alcanzar mediante solo la utilización de vibraciones.

Existen muchos otros dispositivos y otras tecnologías que se han utilizado para generar interfaces físicas que actúen sobre el cuerpo del usuario. Muy pocos de estos dispositivos han llegado al consumo comercial y la mayoría de ellos solo existen en laboratorios especializados.

El objetivo de este Capítulo es presentar las características del sistema táctil y háptico del humano, para poder comprender los sistemas interactivos que intentan estimularlo. La Sección 2 presenta cuestiones relacionadas con el sistema del tacto humano. Para poder comprender la complejidad de estos dispositivos, se debe conocer sus distintas clasificaciones. Una clasificación sobre los diferentes tipos de displays hápticos existentes se presenta en la Sección 3. La Sección 4 presenta una revisión de la literatura de aquellos dispositivos que existen tanto en el ámbito académico como en el ámbito comercial. Finalmente la Sección 5 discute los aspectos relevantes del presente Capítulo, finalizando con una conclusión correspondiente.

2. SISTEMA TÁCTIL HUMANO

El sentido del tacto representa uno de los sentidos más importantes del cuerpo humano. El ser humano percibe su entorno cercano a través de este sentido, que repercute a lo largo de toda la piel, el órgano más grande del cuerpo.

El término háptico deriva de las palabras griegas *haptikos* y *haptesthai*, y significa la pertenencia al sentido del tacto ([275]). La realimentación háptica es provista por el estímulo somatosensorial a través de interfaces físicas que proveen información táctil o cinestésica al usuario. En los últimos años, el campo de los dispositivos hápticos ha contribuido significativamente en el entendimiento de cómo la sensación táctil se utiliza para explorar y entender el ambiente y el rol de la sensación en el movimiento funcional.

Se han desarrollado dispositivos hápticos capaces de generar sensaciones táctiles de objetos virtuales ([276]). Algunos ejemplos más comunes son los vibradores ubicados en los teléfonos móviles y joysticks de videojuegos que generan fuerza.

Las personas que desarrollan sistemas de RV se están dando cuenta que mejorar los parámetros visuales de una simulación, como por ejemplo la resolución de la pantalla, las dimensiones, la frecuencia de refresco o el renderizado 3D, hoy en día solo mejoran modestamente la sensación de presencia del usuario. Y algunas veces, ni siquiera genera un cambio notorio en la performance de los usuarios en un ambiente dado ([94, 277–279]).

Hasta el momento, ninguno de los dispositivos existentes es capaz de reproducir una interacción táctil o háptica real y completa con los objetos virtuales ([280]). Cada vez que agarramos un objeto, obtenemos información sobre su estructura (en términos de señales propioceptivas) pero también recibimos información sobre su superficie (si está mojado, suave, rugoso, etc.). Los diferentes receptores de la piel contribuyen para deducir todos estos diferentes aspectos de un estímulo táctil. Un guante háptico puede simular el contacto con un objeto, su forma e incluso su solidez, pero no todas las características de su superficie, o incluso su peso. Así, la naturaleza multisensorial del tacto aún está lejos de ser realista al utilizar dichos dispositivos.

La información táctil somatosensorial es vital para las actividades normales de agarrar objetos. La información sobre la textura o forma de un objeto, por ejemplo, es vital para asegurar las correctas presiones de un agarre ([281]). La información táctil también es necesaria para tareas que requieren movimientos precisos y rápidos de los dedos, como tocar el piano o el clarinete ([282, 283]). La ausencia de una realimentación táctil normal resulta en la realización de las actividades diarias de manera torpe ([284]).

Existen dos sistemas somatosensoriales: el sistema cinestésico (llamado también propioceptivo), que provee información sobre la posición y el movimiento del cuerpo y las extremidades, y el sistema táctil (llamado también cutáneo), que provee realimentación del mundo externo. Las señales cinestésicas son la combinación de entradas nerviosas que sensan las articulaciones, la longitud de los músculos, la tensión y resistencia al esfuerzo muscular.

El sistema cinestésico provee realimentación sobre la posición de los dedos y las manos permitiendo percibir el tamaño y la forma de los objetos. Además, estas señales son usadas por el cerebro para determinar la información sobre el mundo, como la firmeza y forma de los objetos, las fuerzas como el viento o la gravedad que actúan sobre el cuerpo, etc. Existen muchas articulaciones en todo el cuerpo capaces de recibir señales cinestésicas, haciendo muy difícil que un solo dispositivo pueda controlar cada posible punto de fuerza en el cuerpo del usuario.

El sistema somatosensorial también se relaciona con el sentido del tacto, que transmite información sobre varios parámetros físicos, mayormente aquéllos relacionados con las propiedades de la superficie de los objetos ([285]). Esto permite que las personas puedan detectar cuándo un objeto está caliente, frío, filoso, suave, rígido, etc.

El sistema táctil responde a las propiedades termo-conductivas de los objetos (por ejemplo para detectar si el objeto es de metal o plástico), a la textura (el objeto es suave o rugoso) y a la presión que se ejerce durante la manipulación (el objeto es redondo o tiene esquinas). El sistema táctil también responde a los parámetros físicos que son capaces de generar daño en la piel al generar la percepción del dolor. Los receptores del sistema táctil se encuentran en la piel a lo largo de todo el cuerpo pero son más prominentes en las zonas de las manos, labios y genitales ([285]).

Las señales táctiles son aquéllas que usan los receptores de la piel para transmitir la información sobre el mundo externo. Las terminaciones se asocian a distintos tipos de receptores táctiles en la piel, que pueden dividirse en mecano-receptores, termo-receptores, electro-receptores y nocio-receptores. Los mecano-receptores están asociados a la percepción de la presión, posición y desplazamiento de la piel ([286]). La percepción cinestésica o propioceptiva se obtiene de una variedad de mecano-receptores que proveen información acerca de la posición y el movimiento de las uniones y la longitud y tensión de los músculos. Los mecanoreceptores en la piel también contribuyen a sentir cuando la piel se estira ([287]).

Los termo-receptores se dividen en dos grupos, receptores calientes y fríos, que son más numerosos en las capas de la piel epidermis y dermis ([285]). Los nocio-receptores indican el potencial de daño del tejido y sensan el dolor y daño de la piel. Entre ellos, los nocio-receptores mecánicos responden al exceso de presión, y deformación mecánica de la punción de la piel. Los termo-receptores sensan el ritmo de transferencia de calor entre la piel y los objetos. Finalmente los electro-receptores sensan el flujo de corriente eléctrica a través de la piel.

La habilidad del cerebro de discriminar estímulos táctiles en espacios cercanos varía según la región del cuerpo. La distancia más corta entre dos puntos detectables de la piel humana también ha sido utilizado para medir la sensibilidad táctil y tiene claras implicaciones en el diseño de dispositivos hápticos ([288, 289]). Se ha reportado la mayor sensibilidad en las superficies del pulgar y dedo índice ([290]). Aunque existe algo de variabilidad entre individuos, las observaciones han demostrado que la distancia más pequeña detectable entre dos puntos, a los que se les aplica un estímulo simultáneamente, es de aproximadamente 1,88mm ([291]), aunque este valor se incrementa con la edad ([292]). En la espalda, los estímulos que se encuentran a menos de 70 mm de distancia son sentidos como si ocurrieran en el mismo lugar; en el antebrazo, esta distancia baja a 30 cm; en la punta de los dedos es de solo 2 mm. De esta forma, la simulación de una superficie plana puede ser lograda al utilizar puntos de contacto separados por distancias menores a este umbral. Darghai et al. ([293]) demostraron que la aplicación de un arreglo de actuadores de 10x15 mm puede ser utilizado para transmitir la percepción de una superficie suave.

Como los dispositivos hápticos se enfocan en la habilidad de la piel para interpretar los estímulos, esta información nos dice que se necesitaría por ejemplo un dispositivo de alta resolución para mostrar textura a los dedos, pero se requerirá menor para el antebrazo. Se deben tener en cuenta las propiedades del cuerpo humano y la parte del cuerpo donde se desea mostrar información háptica a la hora de diseñar un nuevo dispositivo o a la hora de elegir un dispositivo ya existente.

3. TIPOS DE DISPLAYS

Existe un gran número de dispositivos que simulan los sentidos somáticos y cinestésicos. En las últimas décadas la investigación sobre dispositivos hápticos se ha vuelto muy popular. En los 90s, esta investigación estaba dominada por dispositivos complejos muy similares a los de robótica (como exoesqueletos). La miniaturización de la mecánica ha resultado en dispositivos de escala más pequeña y más variados.

En general, las interfaces hápticas tienen una función básica: mostrar fuerzas de contacto y posiciones al usuario. Entre estas posiciones y fuerzas de contacto, las variables sensoriales dependen del diseño de hardware y de software, así como también de las tareas que se lleven a cabo.

Los dispositivos hápticos se diferencian según los componentes del display háptico que generan la fuerza o la sensación táctil. También difieren según la fuerza que ejercen y de dónde se referencian para poder ejercer dicha fuerza. Estos dispositivos también difieren según el tipo de señales que proveen, su respuesta, cómo se conectan al cuerpo y su logística. Esta sección explica estas características y cómo éstas son importantes para diferentes aplicaciones.

3.1. CLASIFICACIÓN SEGÚN LA REFERENCIA FÍSICA

Los dispositivos hápticos generan sensaciones sobre la piel del usuario. La mayoría de estos dispositivos aplican una fuerza, y para poder hacerlo, deben sostenerse contra algo. La forma más común de clasificar a los dispositivos hápticos es según su referencia, es decir, si están unidos a algún objeto del mundo para poder sostenerse, o si se sostienen directamente sobre el cuerpo del usuario.

3.1.1. DISPOSITIVOS HÁPTICOS REFERENCIADOS AL MUNDO

Los dispositivos hápticos referenciados al mundo generan un vínculo físico entre el usuario y el mundo. Estos dispositivos pueden anclarse a un escritorio, pared, techo o piso, para poder soportar su peso. Entre éstos se incluyen los joysticks que ejercen fuerza, dispositivos con forma de lápiz, dispositivos con cuerdas, plataformas y grandes brazos robóticos articulados.

Este tipo de dispositivos provee una forma de hacer que los usuarios sientan objetos virtuales. Dependiendo del mecanismo usado, estos dispositivos pueden proveer una gran cantidad de presión resistiva, lo que crea la sensación de contacto físico con los objetos. Los dispositivos referenciados al mundo típicamente utilizan tecnología de actuadores eléctricos, neumáticos o hidráulicos.

Generalmente, un display referenciado al mundo es un dispositivo mecánico que provee una fuerza en las extremidades del usuario. Estos dispositivos pueden simular resistencia y requerir que el usuario se esfuerce para lograr una tarea, alcanzando hasta seis grados de libertad.

Aunque más común en los dispositivos referenciados al cuerpo, los dispositivos referenciados al mundo también pueden proveer estimulaciones táctiles simples. Los displays táctiles como actuadores de temperatura o de vibración pueden montarse en el extremo que hace contacto con el usuario. De esta forma se puede lograr, por ejemplo, la simulación de textura de los objetos.

3.1.2. DISPOSITIVOS HÁPTICOS REFERENCIADOS AL CUERPO

En los dispositivos hápticos referenciados al cuerpo, el dispositivo está unido al usuario de alguna forma. Estos dispositivos le proveen al usuario mucha más libertad de movimiento que los dispositivos referenciados al mundo ya que el usuario no está unido directamente al mundo físico. Sin embargo, el usuario debe soportar el peso del dispositivo. Además, estos dispositivos tienen limitaciones en cuanto a tamaño y peso, haciéndolos más difíciles de diseñar. Este tipo de dispositivo generalmente usa tecnología de actuadores eléctricos o neumáticos.

Algunos ejemplos incluyen brazos exoesqueléticos unidos a la espalda del usuario o guantes que ejercen fuerza, unidos al antebrazo del usuario. Estos dispositivos típicamente utilizan cables o poleas para transmitir fuerzas a la mano y a los dedos, con actuadores ubicados remotamente. Caso contrario los actuadores se colocan en la mano del usuario, reduciendo la complejidad del dispositivo. Sin importar el tipo de referencia del dispositivo, como son utilizados por el usuario, éstos requieren un tiempo para colocarlos y para calibrarlos.

3.2. CLASIFICACIÓN SEGÚN EL ESTÍMULO

Los dispositivos hápticos pueden clasificarse según si ejercen activamente una fuerza sobre el usuario, mediante algún tipo de sistema o tecnología especial, o si es un objeto puramente pasivo. Este último suele ser un objeto que no intenta generar fuerzas sobre el usuario, sino que está pensado para que mediante sólo su forma física pueda generar una sensación háptica particular sobre el usuario. A continuación se discuten las tecnologías y las técnicas utilizadas para los dispositivos hápticos activos y pasivos.

3.2.1. DISPOSITIVOS HÁPTICOS CON FEEDBACK ACTIVO

Los dispositivos hápticos utilizan distintos tipos de actuadores para estimular el sentido táctil del usuario. Existen distintos tipos de dispositivos y muchas modalidades para proveer una estimulación háptica táctil, como por ejemplo cámaras de aire inflable, vibradores, arreglos de pines, corrientes eléctricas, y reguladores de temperatura. Los dispositivos táctiles pueden operar a través de un gran número de mecanismos, incluyendo la aplicación de un campo magnético o eléctrico, mediante la incorporación de cristales piezoeléctricos, sistemas neumáticos o el uso de fluidos que cambian de viscosidad en diferentes ambientes.

Estos dispositivos principalmente operan mediante la aplicación de presión, vibración, campos eléctricos o cambios en el flujo térmico. Como resultado, los displays táctiles se describen como capaces de replicar de cerca los parámetros táctiles de un objeto incluyendo temperatura, textura, forma y rugosidad ([294]).

Los actuadores de cámara de aire son bolsillos que pueden ser expandidos o contraídos mediante el control del flujo de aire (neumáticos) o líquido (hidráulicos) que entra o sale de ellos. El posicionamiento estratégico de los bolsillos crean la sensación de presión en diferentes áreas del cuerpo. Esta tecnología presenta varias dificultades. Primero, como muchos otros dispositivos, es difícil diseñar un dispositivo que funcione bien para muchos usuarios. Luego, el hardware suele ser incómodo para los usuarios, difícil de mantener y bastante delicado. El tiempo que demora en llenar y vaciar los bolsillos puede ser lento, especialmente en los sistemas neumáticos.

Los actuadores con vibradores pueden ser integrados a un dispositivo tipo guante por ejemplo. Típicamente, solo se colocan unos pocos vibradores en un mismo display. Estos tipos de actuadores son más robustos y fáciles de controlar que las cámaras de aire, por lo que suelen utilizarse comúnmente para simular la sensación de presión. La presión ejercida por una pelota de tenis virtual, por ejemplo, puede ser representada mediante vibraciones en la punta de los dedos. Así, al apretarse la pelota, la cantidad de vibración se puede incrementar, indicando así un incremento en la presión en los dedos de la mano.

La utilización de altavoces de baja frecuencia (*subwoofers*) presentan otra forma de generar vibraciones. Estos altavoces pueden ser colocados bajo el suelo, añadidos a una silla o escondidos en las prendas del usuario.

La tecnología de actuadores de *pinnes* aún está siendo investigada y no se ha utilizado en muchas aplicaciones. Un sistema basado en pinnes puede usarse para mostrar una superficie con textura. Muchos pinnes se disponen en arreglos que pueden moverse hacia adentro o hacia afuera, haciendo contacto con la piel. Sistemas con pequeños arreglos de pinnes en cada dedo están siendo investigados, incluso sistemas con pinnes en un cilindro que rota.

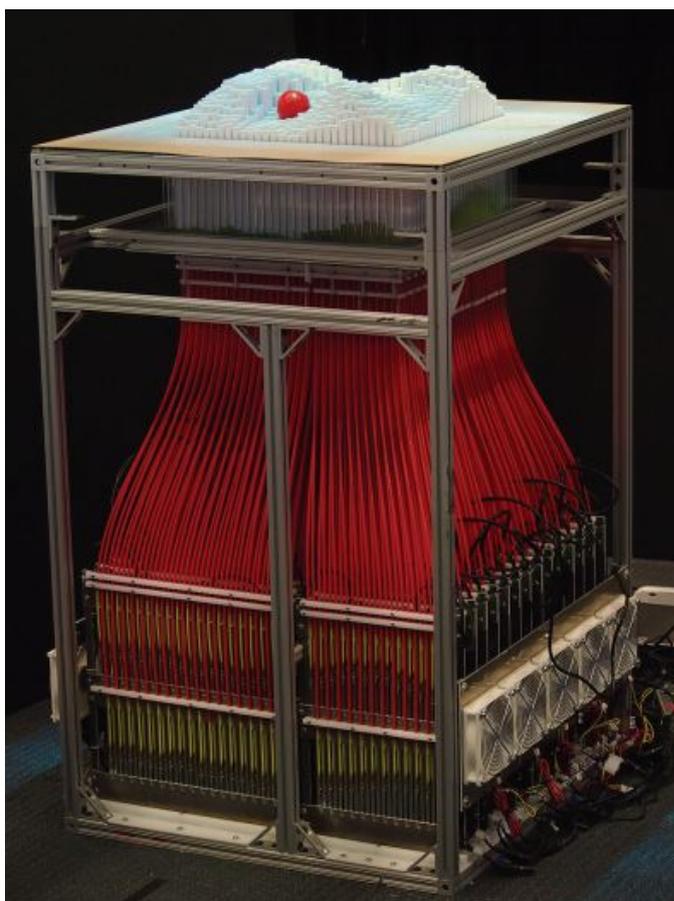


Figura 6.1: Sistema con arreglo de pines ([295]).

Leithinger et al. ([295]) han desarrollado un sistema que consiste de 120 pines para representar formas en una superficie. El sistema puede verse en la Figura 6.1 y su complejidad puede notarse a simple vista. La dificultad mecánica para poder controlar cada uno de los pines, el costo y el tamaño, son las mayores desventajas de este dispositivo. Sin embargo, si se llegara a disminuir su tamaño, podría utilizarse en muchas aplicaciones de RV para generar una sensación háptica al usuario sobre todo tipo de formas distintas.

El diseño de dispositivos para proveer estimulación háptica térmica necesita considerar el rango término que puede ser aplicado sin llegar a causar dolor o daño térmico en la piel. También es necesario considerar la habilidad de detectar cambios en la temperatura. Si la temperatura se mantiene en el rango de 30-36°C, los individuos podrían no detectar los cambios si la tasa de cambio es menor a 0.5°C/min ([296]). La sensibilidad térmica también es dependiente de la edad de la persona y del área de la piel testeada. La piel también tiene una resolución espacial pobre en cuanto a la detección de los cambios de temperatura ([285]).

La información sobre las propiedades térmicas de un objeto pueden facilitar la identificación del objeto cuando la información visual o auditiva no está disponible ([296]). Merrett et al. ([297]) demostraron que la modulación temporal provista por sistemas que cambian de forma crean una percepción de sostener objetos cálidos que puede ser útil para simular objetos como una taza de té.

Los actuadores de temperatura disponibles pueden variar rápidamente su temperatura y típicamente se aplican a la punta de los dedos. Estos dispositivos pueden ser lo suficientemente calientes o fríos como para dañar la piel. Por esto, es importante pensar en la seguridad. Una solución es presentar solo un cambio relativo en la temperatura, como para indicar que un objeto esta frío o caliente, pero sin llegar a la temperatura real. Aunque estos dispositivos pueden integrarse fácilmente a aplicaciones donde la temperatura de los objetos es importante, todavía se están utilizando poco en sistemas de RV.

La estimulación electrotáctil involucra pasar una corriente eléctrica a través de la piel, para generar sensaciones táctiles, en lugar de activar los receptores musculares ([298]). La locación del estímulo aplicado afecta el umbral absoluto, es decir, la corriente eléctrica más pequeña que puede ser detectada debido al grosor variable de las capas de la piel a lo largo del cuerpo. El umbral absoluto también se ve afectado por la tasa de repetición, la onda y el tipo de configuración del electrodo, incluyendo su material y el tipo de pasta eléctrica. El voltaje aplicado también genera distintas sensaciones ([299]). Además de la posibilidad de confiar en la realimentación del usuario para controlar la intensidad de la estimulación ([300]), el medir la impedancia eléctrica de la piel también sirve como un método de control ([301]). De esta forma, se pueden incorporar circuitos de realimentación en tiempo real en el diseño de dispositivos. Así, se pueden determinar los parámetros de intensidad de cualquier estimulación aplicada basándose en las propiedades de la piel.

3.3. DISPOSITIVOS HÁPTICOS PASIVOS

Aunque no haya ninguna fuerza activa, igualmente se puede sentir un objeto, lo que constituye de cierta forma una realimentación háptica. Un volante sin realimentación de fuerza igualmente se siente realista y es más fácil de manipular que uno virtual.

Los dispositivos hápticos pasivos o interfaces, generalmente conocidos como *props*, son dispositivos de la vida real embebidos en la RV, haciéndolos más fáciles de utilizar por el usuario. La mayoría de las interfaces de usuario tangibles se usan para acciones de manipulación y son poderosas, ya que muchas acciones en un mundo virtual pueden ser representadas por un objeto conocido de la vida real.

Como se verá más en detalle en el Capítulo 7, un *prop* es un objeto físico que se utiliza para representar algún objeto dentro del mundo virtual. En una experiencia de RV, por ejemplo, un círculo de cartón puede representar un volante o un pequeño cilindro puede representar un sable láser. Algunos *props* sólo son formas simples (esferas, conos, platos, etc) que pueden aproximarse a una gran cantidad de objetos.

Como los *props* son objetos físicos, éstos proveen información háptica a los usuarios. Por ejemplo, si se utilizara una raqueta de tenis como *prop*, esto proveería la textura del *grip*, el peso correcto y el centro de gravedad que daría un momento realista necesario para manipular una raqueta virtual. Estas características no se pueden obtener con guantes especiales, teclados, *mice* o *joysticks*. Los *props* generalmente se construyen para un propósito específico en una aplicación.

Los *props* permiten una interacción más flexible e intuitiva con los mundos virtuales. El principal objetivo del uso de *props* es el crear una interfaz que los usuarios puedan manipular de la forma más natural posible. Otro beneficio es que el realismo que aporta la incorporación de dispositivos hápticos a los objetos virtuales hace que el resto del mundo virtual parezca incluso más real.

La forma del dispositivo es la forma de la unidad física con la que el participante interactúa. La forma puede ser un *prop* que representa una forma particular, como un palo, pelota o raqueta, un arma o una forma que cambia según las necesidades del display, como un guante.

Nos referimos al uso de *props* inactivos por sus propiedades hápticas como háptica pasiva. Lo único que siente el usuario es la forma del objeto, que no genera ninguna fuerza activa.

Un tipo de display que puede considerarse háptico, aunque no sea particularmente interactivo, se conoce como estereolitografía. Esto es una forma de crear modelos físicos basados en modelos de computadora. El modelo provee así una sensación háptica y visual. Actualmente, con las impresiones 3D, se pueden lograr dispositivos hápticos pasivos con características muy realistas y de bajo costo.

4. REVISIÓN DE LA LITERATURA

Los dispositivos hápticos táctiles han sido desarrollados para simular, entre otros, el sentido de presión sobre la piel. Esto incluye dispositivos que usan motores para contraer una banda alrededor de los dedos ([297, 302]), dispositivos que utilizan fluidos eléctricos o magnéticos para cambiar su viscosidad ([303]), y sistemas que cambian de forma cuando una corriente eléctrica pasa a través de ellos generando presión ([288, 297, 304]).

Existen varios mecanismos para proveer una respuesta de fuerza, incluyendo palancas mecánicas o actuadores neumáticos conectados a la mano, que forman una interfaz entre los dedos del usuario y la computadora. Los movimientos de la mano son trasladados en acciones en la pantalla y el feedback es provisto a los dedos a través de palancas mecánicas ([297]). Esto genera la sensación de las propiedades físicas del objeto virtual, de tal forma que el usuario pueda sentir el tamaño, peso o textura del objeto virtual.

La información cinestésica de la fuerza puede ser experimentada mediante el uso de un dispositivo robótico, un lápiz especializado o un guante especial con un exoesqueleto ([305]). Los exoesqueletos de mano o muñeca, como por ejemplo HWARD ([306]) y PERCO ([307]) pueden ser utilizados para monitorear, e imponer movimiento en cada articulación. Sin embargo, estos dispositivos suelen ser complejos y costosos.

Los dispositivos con fuerza de respuesta (o Force Feedback) son claramente eficientes en término de simular la interacción entre la mano y un objeto. Sin embargo, suelen ser más complejos desde la perspectiva ingenieril y generalmente son físicamente más grandes que un simulador vibrotáctil ([308]). Estos últimos están generalmente limitados por la mano ya que son incómodos si se aplican en varias partes del cuerpo ([309, 310]).

La estimulación táctil a través de toda la superficie del cuerpo requiere la implementación de grandes exoesqueletos extremadamente incómodos. Estos pueden proveer una respuesta de fuerza a cualquier parte de nuestros cuerpos, así como también podría simular cada clase de receptores táctiles de la piel.

Por el momento, la tecnología actual está lejos de lograr una interfaz de cuerpo completo. La presencia de un exoesqueleto no siempre puede ser compatible con los sistemas necesarios para llevar estimulación táctil a la superficie del cuerpo (debido a los puntos de contacto entre el exoesqueleto y la piel) y además presentará una gran complejidad para poder simular movimientos con el grado de libertad requerido en un espacio muy pequeño. Uno de los mayores problemas relacionados al desarrollo de un exoesqueleto para el cuerpo completo es el hecho de que el funcionamiento de cualquier dispositivo mecánico siempre tendrá asociada la presencia de fricción, algo que no siempre está presente en situaciones de la vida real ([311]).

Takei et al. ([312]) desarrollaron WholeGrip, un sistema de agarre con un display electro-táctil en toda su superficie. El usuario agarra WholeGrip con su mano completa e interactúa con objetos virtuales utilizando su mano.

En [313] presentan ThirdHand, un brazo robótico que provee realimentación de fuerza con 5 grados de libertad, para enriquecer las experiencias de juegos en dispositivos móviles. El sistema se encuentra referenciado a los brazos del usuario para poder ejercer fuerza sobre el dispositivo móvil. Se demostró su uso con un juego de carreras. En este caso, cuando el usuario acelera, la pantalla se aleja de su cara; cuando el usuario frena, la pantalla es empujada hacia su cara; cuando el usuario choca con algo, el dispositivo se inclina hacia el lugar de la colisión.

Lyu et al. ([314]) desarrollaron Elastylus, un sistema háptico para poder pintar con los dedos. Un pequeño resorte se comprime cuando el dedo del usuario se acerca a una pantalla. Así el usuario siente una mayor fuerza sobre su dedo cuanto más cerca de la pantalla está. Esto permite que el usuario pueda pintar con distintos niveles de presión.



Figura 6.2: *Botomless Joystick* ([317]).

En [315] presentan un dispositivo capaz de generar una sensación de superficies dispares en las manos del usuario. Este sistema consiste de dos partes. Por un lado una superficie cubierta de imanes. Por el otro lado, un objeto portátil con imanes ubicados en posiciones estratégicas. Así, al cambiar las posiciones de estos imanes, se pueden generar sensaciones de textura distintas.

Kato et al. ([316]) diseñaron un dispositivo con forma de palillos chinos. Este dispositivo permite que el usuario pueda abrir y cerrar los palillos normalmente, pero también puede inmovilizarse mecánicamente. De esta forma, brinda la sensación háptica de que se está tomando un objeto de determinado tamaño entre los palillos.

En [317] crearon un dispositivo llamado el Botomless Joystick para interactuar en sistemas de RV. Este sistema puede verse en la Figura 6.2 y dispone de un contrapeso que puede moverse en cualquier dirección, de tal forma de generar una fuerza en esa dirección. Aunque dispone de un peso fijo, este dispositivo puede utilizarse para simular la sensación de estar agarrando objetos virtuales de distintas formas y pesos, según a donde se direcciona el contrapeso.



Figura 6.3: Usuario interactuando con el sistema educativo virtual de geografía ([318]).

Varios sistemas aprovechan los beneficios de los sistemas hápticos pasivos para incrementar la percepción de objetos dentro de los sistemas de RV, especialmente dentro del ámbito educativo. En [318] utilizan el Leap Motion para detectar las manos del usuario, mientras éste las ve dentro de la RV. Además, como vemos en la Figura 6.3, el usuario deberá interactuar con distintas piezas que representan distintos países. Dichas piezas son trackeadas mediante una cámara y marcadores. De esta forma, el usuario puede no sólo ver cómo manipula estas piezas, sino que también puede sentirlas con sus manos.

Shapira et al. ([319]) crearon un entorno en donde los niños puedan interactuar con distintos objetos físicos mientras experimentan distintos mundos virtuales. Estos elementos eran en su mayoría simples cubos pasivos, pero los resultados mostraron que su utilización fue más motivante comparado con no utilizar ningún objeto físico.



Figura 6.4: HMD con matriz de emisores ultrasónicos que aplican una fuerza háptica sobre la mano del usuario ([321]).

En [320] plantean una técnica interesante que combina sensaciones hápticas con estímulos visuales. En este experimento, el usuario pasa su dedo sobre una superficie cilíndrica, pero la salida visual muestra que está tocando otra cosa en lugar de un cilindro. Esto engaña al cerebro y lo hace creer que en realidad se está tocando la forma que se está viendo, en lugar de la que en realidad se está tocando. Esto prueba que no es necesario disponer de muchos objetos distintos para representarlos en RV, sino que se puede utilizar alguna forma básica y jugar luego con los efectos visuales correspondientes.

Sand et al. ([321]) presentan un sistema que genera una realimentación háptica en el aire, que se coloca en cualquier tipo de HMD. Este sistema dispone de una matriz de emisores ultrasónicos para poder generar la sensación háptica en la mano del usuario frente al HMD. El sistema puede verse en la Figura 6.4.

En entrenamiento médico, es importante que los cirujanos perciban la rigidez de los tejidos. En [322] desarrollaron un dispositivo háptico de levitación magnética basado en principios electromagnéticos para mejorar la percepción de la rigidez de los tejidos en ambientes virtuales. El usuario puede detectar directamente los tejidos virtuales al mover un lápiz magnético en el campo magnético generado por el sistema. Los experimentos han demostrado que el sistema provee una percepción precisa de la rigidez de los tejidos.

Existen trabajos que demuestran que los efectos visuales ayudan a mejorar la percepción de las sensaciones hápticas. Hashiguchi et al. ([323]) realizaron un experimento con generadores de temperatura ubicados en el antebrazo, que se enfriaban o calentaban según lo que el usuario estaba viendo en RV. Demostraron que la percepción de temperatura está altamente afectada por la estimulación visual.

Varios dispositivos se centran en realimentación cinestésica con sensaciones táctiles limitadas que tienen la desventaja de estar referenciados al piso, son pesados o difíciles de integrar con los sistemas de RV existentes. Esto hace que sean difíciles de usar por un largo tiempo y restringen el movimiento de los usuarios.

En Hummel et al. ([324]) presentan un dispositivo liviano y pequeño que provee información táctil en RV mediante estimulaciones electrotáctiles en la punta de los dedos. Esto involucra corriente que pasa a través de la piel, que genera una sensación al excitar directamente los nervios correspondientes.

En [325] diseñaron un dispositivo háptico para simular la condición en que los usuarios agarran objetos con los dedos pulgar, índice y mayor. El sistema consiste de tres pequeños actuadores con esferas en la punta que hacen contacto con cada dedo. Cuando los usuarios mueven su mando mientras agarran un objeto virtual, los pequeños actuadores se mueven para representar la posición de la superficie del objeto.



Figura 6.5: Brazo robótico que simula la interacción con objetos virtuales ([326]).

En Vonach et al. ([326]) presentaron un sistema de RV que puede proveer feedback háptico en ambientes virtuales. Un brazo robótico mueve objetos reales para asimilar la posición y rotación de objetos virtuales. El sistema no requiere de estructuras externas como exoesqueletos, y en combinación de alguna plataforma para que el usuario pueda caminar, puede lograr una interacción háptica sin restricciones. El sistema puede verse en la Figura 6.5. La mayor desventaja es el tiempo de respuesta del brazo robótico y el espacio físico, ya que el brazo robótico se encuentra restringido a una superficie fija en el entorno.

La estimulación vibrotáctil se utiliza ampliamente en videojuegos, RV y telecomunicaciones. Los controles de videojuegos de la Sony Playstation o el Nintendo Wii incorporan feedback vibrotáctil para mejorar la experiencia de juego ([327, 328]). Los CyberTouch, desarrollados por CyberGlove Systems ([329]), han sido utilizados en neuro-rehabilitación para proveer estimulación vibrotáctil a la palma de la mano o al dorso de los dedos, mientras el usuario interactúa con el mundo virtual ([330]).

Maeda et al. ([331]) presentan HapticAid, un sistema que amplifica las sensaciones hápticas. Se compone de un sensor que se coloca en el dedo índice, un sistema de filtrado y un actuador háptico de vibración que se coloca en la muñeca. Al pasar el dedo por cualquier superficie, el sensor traduce la textura de dicha superficie en una señal determinada, que posteriormente es filtrada y amplificada en la muñeca del usuario mediante respectivas vibraciones. De esta forma se amplifican las sensaciones hápticas de todo tipo de objetos, incluso aquéllos con texturas suaves.

Oliveira et al. ([332]) desarrollaron una cinta compuesta de vibradores que se coloca sobre la cabeza del participante, con el objetivo de ayudar a la localización de objetos dentro del entorno virtual. Los vibradores vibrarán para poder dar indicaciones al usuario. Estas indicaciones le sugerirán la dirección en la que debe girar para alcanzar el objetivo indicado.

Como se mencionó, también se pueden usar altavoces para generar sensaciones hápticas mediante su vibración. En [333] presentaron un teclado virtual con realimentación háptica realista. El sistema usa dos guantes especiales para trackear las manos del usuario y micro-altavoces para generar vibración en los dedos. Cuando los usuarios presionan una tecla virtual del teclado, el sistema provee un *click* háptico realista, lo que mejora la performance de la actividad.

Existen exoesqueletos más pequeños como el Cyberglove ([329]), pero igualmente siguen siendo dispositivos costosos para el público general. Otros dispositivos, como el MIT-Manus ([334]), el Haptic Knob ([335]) y el Phantom ([336]) son más simples y más baratos de implementar.

El sistema Phantom permite la simulación del contacto de la punta de los dedos con objetos virtuales. Un dispositivo en forma de lápiz trackea los ángulos de rotación y la posición en coordenadas cartesianas de la mano. Sus actuadores comunican fuerzas a los dedos del usuario cuando detectan colisiones con objetos virtuales, simulando el sentido del tacto ([337]). El Phantom se ha utilizado como un dispositivo de entrenamiento para simular tareas como cirugías ([338]) que requieren gran precisión en el mundo real.

Song et al. ([339]) utilizan el Phantom Omni ([340]) para generar una sensación háptica de profundidad de la cámara de entornos 3D. El usuario utiliza este dispositivo (como si fuese un mouse) para mover la cámara, pero además la distancia hacia la superficie que se encuentra en el centro de la cámara es reflejada en el Phantom Omni. De esta forma el usuario percibe una sensación háptica de la “textura” de las superficies 3D. Por ejemplo, al pasar la cámara por una escalera, el sistema provee una sensación háptica de profundidad cuando se pasa por los distintos escalones.

En [341] desarrollaron un simulador de cirugías muy completo. Este sistema utiliza el Phantom Omni ([340]) para generar una respuesta háptica muy rápida y precisa.

Otros trabajos, como por ejemplo en [342], han investigado sobre la estimulación háptica producida por el viento y cómo esto mejora la sensación de presencia dentro de la RV. Dicho trabajo mostró que la inclusión de viento genera una sensación háptica que mejora la presencia, que es difícil detectar la posición del viento desde atrás de la cabeza y que no hay diferencia en la detección entre los distintos géneros.

Tecnologías como guantes hápticos están siendo usadas en un gran número de campos, especialmente donde se requiere una reproducción visual precisa de los movimientos de la mano en simulaciones virtuales. Por ejemplo, varios estudios están comenzando a investigar el uso de guantes hápticos para la rehabilitación de pacientes que han perdido la habilidad de mover sus brazos o la sensibilidad en éstos, como consecuencia de un derrame cerebral ([343-349]).

En los últimos años, varios estudios han mostrando los beneficios potenciales de los dispositivos hápticos cinestéticos para la neurorrehabilitación de la mano. Almari et al. ([275]) desarrollaron una interfaz háptica que utilizaba el guante CyberGrasp para realizar actividades en RV. Capturaron datos de personas saludables que interactuaban con el sistema y mostraron su potencial para capturar métricas de performance para evaluar la performance de las tareas de forma cuantitativa.

Admovich et al. ([330]) demostraron mejoras en el alcance, velocidad y fuerza en el movimiento de los dedos, en varios participantes entrenados con el CyberGlove y el Rutgers Master II-ND Haptic Glove en un ambiente virtual. Merians et al. ([347]) investigaron el uso de entrenamiento de RV hápticamente computarizado en la mano hemiparética de ocho participantes después del accidente cerebrovascular. Qiu et al. ([350]) utilizaron entrenamiento háptico en RV para entender los mecanismos de la recuperación de la mano y el brazo. Evidencia de personas saludables también mostró los beneficios de la RV háptica comparados con la RV tradicional ([291, 351, 352]).

Un área en donde los guantes hápticos son muy útiles es en los videojuegos. De hecho, el agregar contenido táctil a los videojuegos genera interacciones más realistas y entretenidas, mejorando el placer del usuario.

Sistemas como el Wiimote ([353]) son herramientas usadas para controlar un videojuego, por lo que requieren ser intuitivas. Usar un Wiimote se asemeja al uso de una raqueta de tenis o de un palo de golf, cosas que realmente no requieren una manipulación háptica. Esto es, la interfaz se usa para controlar el videojuego y la sensación de estar sosteniendo el objeto en la mano contribuye a la sensación de realismo del juego. En el caso de los Power Gloves ([329]) los movimientos de la mano se usaron para controlar el juego, algo que no se asemeja a la interacción natural con el mundo.

Varios trabajos focalizados en la realimentación táctil se han basado en generar una realimentación en la punta de los dedos. Por ejemplo, Bouzit et al. ([354]) desarrollaron Rutgers Master II-ND, un guante diseñado para soportar interacciones en RV usando fuerzas en las puntas de los dedos y el pulgar. Kron et al. ([355]) combinaron un dispositivo cinestésico en la mano (DeKiFeD₃) con otro en los dedos (HapticGlove) para la exploración de objetos rígidos. Prattichizzo et al. ([356]) propusieron el prototipo de un *joystick* en que la realimentación cinestésica es substituida por realimentación táctil aplicando fuerzas en las puntas de los dedos.

Existen dispositivos que permiten la simulación de la sensación proveniente de la mano y sus movimientos, como por ejemplo los Haptic Gloves ([357]). Hoy en día existen varias versiones de estos dispositivos en una gran variedad de precios. Algunos de ellos reproducen los movimientos precisos de la mano en el ambiente virtual, otros también proveen información (como vibración) cuando hay contacto entre la mano del usuario y objetos virtuales. Otros, solo proveen señales propioceptivas sobre alguna interacción con los objetos. En este último caso, la información sobre la solidez o estructura interna de los objetos, se provee mediante la simulación de fuerzas mecánicas sobre la manipulación de los objetos (es decir, una fuerza de respuesta).

Últimamente están surgiendo dispositivos comerciales que aseguran que se puede alcanzar una sensación realista mediante la incorporación de sistemas hápticos novedosos. Taclim es un conjunto de zapatos y guantes que generan sensación háptica al usuario ([358]). El sistema puede verse en la Figura 6.6 y asegura generar la sensación de estar parado sobre distintos tipos de suelo (desierto, césped, agua). El problema principal de este tipo de dispositivos es que resulta imposible generar la sensación de estar parado o caminar por una superficie particular mediante solo la utilización de vibración.

Pareciera que los guantes hápticos, para lograr ser una interfaz efectiva, deberán transmitir información intuitiva al usuario. El hecho de no entender para qué se usa el tacto y la propiocepción, y cómo interpreta nuestro cerebro esas señales en el mundo real, generalmente resulta en el desarrollo de interfaces hápticas de RV insatisfactorias.



Figura 6.6: Sistema Taclim que asegura simular el contacto con distintas superficies ([358]).

5. DISCUSIÓN Y CONCLUSIONES

Los displays hápticos ofrecen un estímulo táctil y de fuerza al usuario, generado por el contacto con los dispositivos que emulan los objetos en el mundo virtual. La mayoría de los displays hápticos comercialmente disponibles proveen un estímulo táctil o un estímulo háptico, pero no ambos. Dos sistemas separados pueden combinarse en un solo sistema de RV. Los requerimientos de una aplicación dada deberían dictar el tipo de display háptico que se debería usar.

Como existen muchos aspectos diferentes de realimentación háptica (tacto, propiocepción, termocepción, electrocepción, etc.) y muchos lugares diferentes donde el display puede ser puesto, los displays hápticos tienen una gran variedad de cualidades logísticas.

5.1. CUERPO HUMANO Y DESAFÍOS DE LA TECNOLOGÍA

La mayoría de los displays táctiles se enfocan en presentar estímulos a las manos, particularmente a los dedos. Esto se debe a que generalmente utilizamos nuestras manos y dedos para manipular el mundo real. Además, la mayoría de nuestros sensores táctiles se encuentran en la punta de los dedos.

Es importante notar que fuera de las restricciones mencionadas, el objetivo de lograr tacto virtual de cuerpo completo puede ser menos desafiante que lo esperado ([359]). Esto se debe a que investigaciones han mostrado que los campos táctiles receptivos son más grandes en ciertas áreas del cuerpo que en otras ([360]). Consecuentemente, un número más pequeño de estimuladores puede usarse en esas partes del cuerpo que tengan menor sensibilidad para simular la presencia de un dado objeto. Además, las personas parecen no darse cuenta de la cantidad de estímulos cuando más de tres estímulos se presentan en su cuerpo al mismo tiempo ([361, 362]).

Esto sugiere que el estimular solo un número limitado de posiciones del cuerpo podría resultar en la percepción de patrones de percepción más complejos, incluso sin la necesidad de proveer una estimulación completa de alta resolución a lo largo de la superficie de todo el cuerpo ([363]). Este factor es incluso más relevante cuando cualquier falta de estimulación táctil se compensa por la presencia de información consistente de otras modalidades sensoriales. Lo que es más, estas observaciones también sugieren que, dado el número limitado de estímulos que una persona puede percibir al mismo tiempo, la RV debe proveer solo aquella información estrictamente relevante para la tarea que los usuarios estén realizando. Esto es, no siempre sentimos la presencia de la ropa sobre nuestra piel o la silla en donde nos estamos sentando. Por esto, proveer esta información en un escenario virtual complejo sería de poca utilidad.

Reproducir la textura de una superficie es más complicado. Sensar la textura de la superficie se basa en la alta densidad de los sensores de presión en la punta de los dedos y en los movimientos de los dedos sobre la superficie. Generar la textura virtual requiere un sensado rápido y preciso de los movimientos de los dedos y una realimentación rápida de los elementos de sensado de presión.

Un método para mostrar la textura consiste en mover un dedo sustituto o falso sobre una superficie virtual y sentir los golpes. Este método es más fácil de implementar ya que solo hay un punto de contacto entre la superficie y el dedo. Además, la sensación de textura proviene de la realimentación cinestésica.

La detección cutánea de vibración es independiente de la frecuencia de la vibración entregada ([364]). Los mecanorreceptores son extremadamente sensibles a la estimulación de alta frecuencia debido a su estructura. Esto asegura que solo las estimulaciones de alta frecuencia alcancen el nervio en cuestión ([293]). No hay ventaja al proveer múltiples puntos de estimulación a la punta de los dedos; un actuador por dedo es considerado suficiente ([365]), con una sensibilidad óptima a aproximadamente 250 Hz ([366]).

5.2. PROPIEDADES FÍSICAS

Comparado con los dispositivos visuales y de audio, los dispositivos hápticos no están tan avanzados tecnológicamente y suelen estar más unidos a aplicaciones particulares. El mercado para estos dispositivos es pequeño, y se ha realizado poca investigación al respecto.

Retardos más largos, o latencia, entre las acciones y la respuesta del sistema puede degradar la ilusión de la solidez del mundo. Como los displays hápticos requieren contacto físico con el participante, los movimientos del participante pueden ser trackeados mecánicamente. Este tipo de *tracking* tiene latencia más pequeña y precisión mayor que muchos otros sistemas de *tracking*.

El tamaño del dispositivo juega un rol importante según el tipo de interacción que puede ser simulado. Un display grande generalmente permite un mayor grado de movimiento, permitiendo la realización de más tareas. Sin embargo, un sistema grande puede tener más problemas de seguridad. Los sistemas pequeños son buenos para aplicaciones que simulan tareas en pequeñas áreas de trabajo limitadas.

Los sistemas grandes pueden ser más difíciles de transportar si estos requieren un gran equipamiento para generar distintas fuerzas. Los sistemas que están montados físicamente al piso o al techo están más limitados en su portabilidad. Además, los sistemas grandes como exoesqueletos generalmente incomodan mucho a los usuarios. Por esto, este tipo de sistemas grandes y complejos solo se utiliza en laboratorios especializados o en grandes eventos especiales.

Por otro lado, los sistemas pequeños como guantes suelen ser más cómodos y pueden ser transportados sin mucha dificultad. Esto hace que su costo baje y puedan llegar al público general.

Los displays de fuerza grandes generalmente requieren una habitación especial donde operar. Esta habitación puede estar equipada con bombas de presión hidráulicas o neumáticas. Los displays de fuerza más pequeños pueden caber en un pequeño espacio en cualquier habitación.

La mayoría de los dispositivos hápticos son costosos debido a que no son comercializados o utilizados masivamente. Además, generalmente los displays hápticos suelen ser más complejos de construir que los otros elementos de RV, debido a que contienen muchas partes móviles y suelen estar diseñados para aplicaciones particulares.

5.3. SEGURIDAD

Existen varios *trade-offs* involucrados en determinar la fidelidad de los displays hápticos. El primero es la seguridad. Los sistemas de alta fidelidad requieren altos niveles de energía, lo que puede ser peligroso. El tamaño del display también es importante, ya que la inercia del mismo dispositivo complica el control preciso de las fuerzas. En el caso de un dispositivo de temperatura, una medida de fidelidad se traduce en qué tan rápido puede cambiar de una temperatura a otra y qué rangos de temperatura maneja. Se debe tener cuidado de no quemar al usuario por ejemplo.

La seguridad es una preocupación importante cuando se trabaja con displays que ejercen grandes fuerzas. Los robots grandes pueden generar golpes fatales a los humanos y los exoesqueletos pueden atrapar al cuerpo. Para incrementar el nivel de seguridad al trabajar con estos sistemas, se suele colocar un botón especial en el dispositivo. Tan pronto como el botón deje de estar presionado, se corta el poder del sistema. Incorporar un sistema similar a los displays más pequeños suele ser algo bueno en caso de mal funcionamiento ya que, aunque las fuerzas no sean suficientemente fuertes para herir al usuario, el dispositivo puede dañarse a sí mismo.

5.4. LOGÍSTICA

La tarea a realizar es algo que se debe tener en cuenta a la hora de elegir una interfaz. Para poder agarrar algún objeto virtual, se requiere que al menos dos partes del cuerpo del usuario tengan interfaces al dispositivo háptico. Dos o más dedos podrían usar dispositivos distintos, o el movimiento entre los dedos puede ser limitado con un exoesqueleto. Un sistema de exoesqueleto puede ser usado para agarrar o empujar objetos del mundo real.

Los dispositivos hápticos referenciados al mundo requieren que el usuario permanezca cerca del dispositivo, lo que limita la movilidad del usuario. Para aquellas aplicaciones donde el usuario está quieto, esto no influye mucho. Los sistemas auto-referenciados pueden ser usados por el usuario, lo que los hace móviles. Sin embargo, los cables del dispositivo pueden generar incomodidad.

Los dispositivos vibratorios pueden colocarse en muchas interfaces distintas. Algunos estudios han demostrado que los movimientos del usuario decrementan la sensibilidad de detectar estímulos táctiles presentados en su cuerpo ([362, 367–372]). Por esto, para que un dispositivo vibratorio sea efectivo, se deben implementar sistemas adicionales para variar la intensidad de la vibración en función de los movimientos del usuario. Estudios sobre supresión sensorial han demostrado que la habilidad de un individuo de detectar estímulos táctiles se decrementa en función de la distancia entre la parte del cuerpo movida y la posición donde el estímulo táctil es presentada.

Generalmente el diseño de una herramienta para RV se hace mediante el análisis de los movimientos de la mano al realizar una dada tarea y prestando poca atención a los aspectos más cognitivos, emocionales y perceptuales, relacionados con el uso de la herramienta ([280, 373]). Estos aspectos sin duda se vuelven incluso más importantes cuando los prototipos virtuales necesitan ser testeados con los usuarios potenciales finales. Considerando que, para que el prototipado háptico sea efectivo, debería estar basado en nuestro conocimiento acerca de cómo la información táctil es procesada por nuestro sistema neurocognitivo y cómo esta información provoca una variedad de respuestas emocionales en el usuario. En particular, la exploración háptica de una superficie puede ser afectada por la presentación simultánea de información visual y auditiva ([374–376]). Esto es, la adopción de un display multisensorial compensaría la gran complejidad requerida para renderizar correctamente la información háptica y táctil en el prototipado virtual ([377]).

La mayoría de las interfaces para RV limitan la presentación del estímulo solo a la superficie de las manos. Sin embargo, un sistema de RV completamente inmersivo debería presentar información sobre todo el cuerpo del usuario. Se debe notar que, además de manipular objetos 3D durante nuestras actividades diarias, las sensaciones táctiles están constantemente activas desde todas las partes de nuestro cuerpo.

Como veremos en el Capítulo 7, se pueden incorporar sensores mecánicos directamente en el sistema. Esto es una característica beneficiosa ya que el *tracking* mecánico generalmente es más rápido y preciso, requerimientos para un buen display háptico.

5.5. CONCLUSIONES

El sentido del tacto es uno de los sentidos menos estudiados a la hora de conectar al usuario con los sistemas de RV. Esto se debe en gran parte a que este sentido se encuentra dominado por la piel, el órgano más grande del cuerpo humano.

Una interfaz que presente estímulos hápticos a toda la piel con la tecnología actual, sería una interfaz grande, compleja, costosa e incómoda. Esto último es una de las mayores desventajas a la hora de trabajar con sistemas de RV, ya que uno de los objetivos fundamentales es que el usuario se sienta lo más natural y cómodo posible. Sin embargo, existen estudios que demuestran que no es necesario abarcar todo el cuerpo del usuario, sino áreas estratégicas en donde el usuario esté consciente. Así, la complejidad de las interfaces hápticas se podría reducir considerablemente.

Es también por este motivo que la mayoría de los sistemas comerciales se enfocan solamente en presentar información háptica a las manos del usuario. Los sistemas con vibradores son los más populares debido a su bajo costo y complejidad. Se colocan pequeños vibradores en todo tipo de interfaces para generar un estímulo háptico en el usuario. El principal problema de esto es que las vibraciones no logran generar sensaciones hápticas muy realistas. Si bien éstas le proveen al usuario una realimentación de que algo está ocurriendo en sus manos, por ejemplo, la sensación de que se está agarrando un objeto no se puede lograr únicamente mediante vibraciones.

Como se vio en este Capítulo, existe diversas técnicas para generar sensaciones hápticas sobre el cuerpo. Los dispositivos hápticos desarrollados en laboratorios suelen estar pensados para una aplicación particular. En un futuro, los avances tecnológicos harán que estos dispositivos sean más pequeños, cómodos y baratos, para que puedan comercializarse y llegar al público general. En ese momento podremos no solo ver y escuchar, sino también sentir distintos estímulos físicos sobre nuestro cuerpo, experimentando otra forma de RV inexistente hasta el momento.

7

Tracking en Realidad Virtual

1. INTRODUCCIÓN

En un sistema de RV ideal, el usuario vería su cuerpo dentro del mundo virtual y no podría distinguir su cuerpo virtual del real. Cada movimiento que el usuario realizara en el mundo real, se vería mapeado al mundo virtual de manera exactamente igual, sin que el usuario note alguna diferencia. Actualmente, este nivel de inmersión no es tecnológicamente posible, aunque el *tracking* del usuario permitirá acercarnos lo más posible a este caso ideal.

Desde los orígenes de la RV, se ha buscado que el usuario no solo sienta que está viendo una película de 360°, sino que sienta que es parte del mundo virtual que está viendo. Para lograr esto, los movimientos del usuario deben ser trackeados de alguna forma para luego trasladar esta representación al mundo virtual. Así, por ejemplo, cuando el usuario mueva su brazo, debería ver que el brazo de su avatar virtual se mueve de la misma forma.

Existen diversas técnicas para trackear la posición en 3D de un objeto en el espacio. Éstas serán discutidas en la Sección 2. Cada técnica tiene ventajas y desventajas, y existen numerosos estudios que discuten cada una de ellas. Todas estas técnicas sirven entonces para poder realizar el *tracking* del cuerpo del usuario. Sin embargo, las partes del cuerpo que se deseen trackear dependerán de la aplicación particular. En la Sección 3 se verán todas las partes del cuerpo que son interesantes o útiles de trackear para lograr una sensación más natural.

Un gran número de dispositivos especiales han sido creados para poder comunicar al humano con la computadora. El hecho de detectar que un usuario está presionando los botones de un control es una forma de trackear la mano de dicho usuario. Otros dispositivos envuelven al usuario para poder trackear, por ejemplo, que éste se encuentra caminando o no, de forma de trasladar estos movimientos al mundo virtual. La Sección 4 discute todos los tipos de dispositivos especiales existentes.

La literatura presenta numerosos trabajos que crean, analizan y discuten los diferentes tipos de *tracking*, las partes del cuerpo que conviene trackear y por qué, y los distintos dispositivos que se requieren para poder hacerlo. La Sección 5 presenta una revisión exhaustiva de la literatura que abarca estos temas. Finalmente, la Sección 6 presenta una discusión sobre los conceptos vistos en este Capítulo, finalizando con una conclusión al respecto.

2. TIPOS DE TRACKING

Desde los inicios de la RV, el *tracking* del usuario ha sido algo fundamental. Diversas técnicas de *tracking* han surgido a lo largo de los años, algunas de ellas provenientes de áreas como la robótica. Esta sección presenta y analiza dichas técnicas, dando ejemplos actuales y significativos para ayudar en la comprensión de las mismas.

2.1. TRACKING MAGNÉTICO

El *tracking* magnético o electromagnético es un método que utiliza un transmisor para generar un campo magnético de baja intensidad a partir de tres bobinas ortogonales. Así, un receptor puede determinar su posición y orientación relativa con respecto a esta fuente magnética. Para esto, la unidad transmisora debe estar fija en una posición y orientación conocida. De esta forma se puede calcular la posición absoluta de la unidad receptora.

Generalmente los sistemas de *tracking* magnéticos son muy precisos, alrededor de 0.1 pulgadas en la posición y 0.1 grados en la orientación. Sin embargo, cualquier objeto conductor ferromagnético (es decir cualquier metal) puede distorsionar el campo magnético, reduciendo así la precisión. Además, la precisión también disminuye notoriamente cuando se incrementa la distancia entre el transmisor y el receptor. Este cambio en la precisión a veces puede ser muy severo, volviendo inutilizable al sistema.

También existen otras limitaciones. Generalmente, el corto alcance del campo magnético generado suele ser un problema. Por esto, esta técnica no es adecuada para grandes displays o hasta HMDs en donde los usuarios deban moverse por el espacio. Si se desea trackear múltiples objetos a la vez, se necesitarán múltiples unidades receptoras. Sin embargo, estas unidades pueden interferir entre ellas y esto puede ser un gran problema.



Figura 7.1: Razer Hydra, sensor de trackeo y controles para cada mano ([378]).

A pesar de las desventajas de este tipo de sistema de *tracking*, también tiene algunas ventajas notorias. La mayor ventaja es que el sistema electromagnético no requiere una línea de visión directa entre el emisor y el receptor. Esto permite que los usuarios puedan moverse libremente en ambientes con obstáculos. Además, la tecnología inalámbrica siempre es más cómoda para los usuarios.

El Razer Hydra, que puede verse en la Figura 7.1, es uno de los primeros sistemas de *tracking* magnéticos producidos masivamente para el mercado de videojuegos [378]. Nos brinda la posición y orientación de dos controles con 3DOF, cada uno con un joystick y 8 botones. Este dispositivo fue el primero en utilizarse en conjunto con el Oculus Rift para interactuar en ambientes virtuales de forma más inmersiva que un mouse, teclado o joystick convencional.



Figura 7.2: Phantom Omni, sistema de trackeo y respuesta háptica ([340]).

2.2. TRACKING MECÁNICO

Los sistemas de *tracking* mecánico poseen una estructura rígida con varias uniones mecánicas interconectadas. Esto se suele combinar con transductores electromecánicos, como potenciómetros o *shift encoders*¹. Un extremo se encuentra fijo mientras que el otro está unido al objeto que será trackeado, por ejemplo, la mano o la cabeza del usuario. De esta forma, cuando el objeto trackeado se mueve, las uniones se mueven también. A partir de los transductores se obtiene la información sobre la posición y la orientación del objeto trackeado.

Estas técnicas de *tracking* mecánico son muy precisas y pueden transmitir información con muy poco retardo. Sin embargo, suelen ser grandes e incómodas, limitando la movilidad del usuario y dificultando las técnicas de navegación físicas.

Uno de los primeros dispositivos y uno de los más utilizados inclusive hoy en día es el Phantom Omni ([340]). Este dispositivo puede verse en la Figura 7.2 y permite simular con gran precisión que el usuario está utilizando alguna herramienta en su mano. Además de trackear los movimientos de la mano del usuario, este dispositivo provee también una sensación háptica, es decir, genera fuerzas sobre el usuario de tal forma de simular distintas colisiones o texturas de los objetos virtuales.

¹Registros de Desplazamiento

2.3. TRACKING ÓPTICO

Los sistemas de *tracking* óptico miden la luz reflejada o emitida para poder trackear usuarios u objetos físicos. Este tipo de *tracking* utiliza técnicas de Visión por Computadora y sensores ópticos como cámaras, emisores infrarrojos o diodos de efecto lateral, los cuales generan señales proporcionales a la posición de la luz incidente sobre un eje particular.

Se puede utilizar una gran variedad de cámaras para trackear objetos o personas. Estas cámaras generalmente deben estar en una posición fija para que los algoritmos de Visión por Computadora puedan determinar la posición del objeto. Cómo las cámaras normales solo proveen imágenes 2D, no hay información sobre la profundidad. Mediante la utilización de más de una cámara al mismo tiempo, el sistema puede triangular la posición y orientación del objeto trackeado, generando así información sobre la posición 3D. La calidad de los resultados dependerá mayormente del número y configuración de las cámaras.

Una de las mayores ventajas de esta técnica es que el usuario está completamente desconectado de la computadora. Sin embargo, la mayor limitación es que la línea visual entre la cámara y el objeto a trackear debe estar siempre despejada. Aunque se pueden agregar más cámaras para reducir los problemas de oclusión, esto incrementará la complejidad del sistema y del algoritmo de *tracking*. Esto es debido al que el sistema deberá considerar la cantidad de cámaras, su ubicación, el ambiente, entre otras cosas.



Figura 7.3: Vista de los LEDs infrarrojos del Oculus Rift CV1 con los cuales el sistema trackea su posición y rotación ([380]).

Hoy en día, los sistemas de RV más populares utilizan *tracking* óptico para obtener la posición y orientación del usuario, así como también de controles especiales. La última versión del Oculus Rift, llamada CV1, está cubierta completamente con LEDs infrarrojos en posiciones estratégicas. Esto puede verse en la Figura 7.3. Los controles del Oculus, llamados Oculus Touch y mostrados en la Figura 7.5, también disponen de un arreglo interno de LEDs infrarrojos para ser trackeados. En la Figura 7.4 podemos ver el sensor de Oculus, que básicamente es una cámara especializada para detectar los LEDs infrarrojos del visor y de los controles. De esta forma el sistema de Oculus detecta la posición y orientación tanto del HMD como de los dos controles. Sin embargo, es importante mencionar que el rango de detección estará acotado por el área de visión de las cámaras. Pueden utilizarse entre una y tres cámaras para tener un mayor rango de detección.



Figura 7.4: Sensor óptico del Oculus Rift que detecta los LEDs infrarrojos del HMD y de los controles ([380]).



Figura 7.5: Controles del HCT Vive (izquierda) y del Oculus Rift (derecha) ([380]).

El HTC Vive ([379]) es otro de los dispositivos de RV más utilizados hoy en día y su sistema de *tracking* es muy similar al del Oculus Rift. La Figura 7.6 muestra el visor, así como también los controladores de los que dispone y los sensores correspondientes. Estos sensores, de la misma forma que los del Oculus Rift, detectan LEDs infrarrojos posicionados estratégicamente en el HMD y en los controles.



Figura 7.6: Visor HTC Vive junto a sus sensores ópticos y controles ([379]).

2.4. TRACKING VIDEOMÉTRICO

La Odometría, en robótica, trata de estimar los cambios en la posición y orientación de los robots ([381]). Los robots utilizan la información de sus sensores para determinar su posición según un punto inicial. La Odometría Visual o *tracking* Videométrico es una técnica utilizada para localizar un robot utilizando solamente un conjunto de imágenes adquiridas por una o varias cámaras unidas al robot ([382]). Para un *review* sobre Odometría Visual ver [383]. Las imágenes contienen suficiente información significativa (color, textura, formas, etc.) para estimar el movimiento de la cámara en un ambiente estático ([384]).

La navegación de robots móviles es uno de los principales objetivos de la Visión por Computadora y de la robótica ([385]). Se busca posicionar efectivamente los robots, particularmente en aplicaciones de exterior ([386]). El *tracking* Videométrico provee una estimación incremental de la posición del objeto mediante el análisis de la secuencia de imágenes capturadas por una cámara ([385, 387]).

El *tracking* Videométrico es una técnica más precisa que las técnicas convencionales, como es el GPS y la Odometría sobre las ruedas ([388]). Tiene un buen *trade-off* entre costo, confiabilidad y complejidad de implementación ([389]). Puede estimar la posición de un objeto de forma económica mediante el uso de cámaras de consumo general, en lugar de sensores o sistemas costosos, como GPS especializado u otros ([389, 390]). Además, no requiere de ningún otro sensor particular ([391]).

En comparación al uso de otros sensores, el uso de cámaras para la localización de objetos tiene la ventaja de la reducción de costos y permite una integración simple con otros algoritmos de visión, como por ejemplo la detección de obstáculos, sin necesidad de calibrar varios sensores a la vez ([392]). Las cámaras son pequeñas, baratas, livianas, de bajo consumo y versátiles. Por esto, éstas pueden ser montadas sobre casi cualquier objeto.

Estas técnicas de *tracking* son muy utilizadas en robótica y se han ido expandiendo a la RV para ser utilizadas, en este caso, para hacer *tracking* del usuario u otros objetos. En lugar de tener una cámara fija, en el *tracking* Videométrico la cámara está unida al objeto que será trackeado y estará mirando los alrededores. El sistema trackea el ambiente para localizar marcadores y derivar la posición de la cámara a partir de ellos. La posición de esos marcadores debería ser conocida para poder determinar la posición absoluta del dispositivo que está sensando.

Sin embargo, el análisis de imagen suele ser computacionalmente costoso. En localización visual, la computación involucra varias etapas, entre ellas la adquisición de las imágenes de la cámara, la extracción de las características de la imagen (bordes, esquinas, líneas, etc.), la comparación de las imágenes y el cálculo de la posición mediante el cálculo del desplazamiento de los píxeles entre cada *frame*. Además, los algoritmos de visión son altamente sensibles a las condiciones ambientales, como la iluminación, las texturas o las sombras, por ejemplo ([393]). Por este motivo, estos algoritmos suelen trabajar bien bajo determinadas condiciones, pero se vuelven poco confiables en diferentes condiciones ambientales ([394]). Los principales desafíos del *tracking* Videométrico están relacionados con el costo computacional y las condiciones de las imágenes ([386, 395–397]) Si los marcadores presentan distintas formas o colores, podrían ser registrados más fácilmente. Generalmente, se suele ubicar los marcadores en aquellos lugares que son susceptibles a caer dentro de la visión de la cámara.

Uno de los problemas al utilizar una sola cámara es que no se puede escalar el mundo correctamente ya que no se dispone de ninguna referencia. Algunos trabajos, como [398], proponen una alternativa con dispositivos móviles. En este caso, la cámara trasera del dispositivo se utilizará para la Videometría, mientras que la cámara frontal estará midando hacia la cara del usuario. De esta forma se puede tener una estimación de la cámara trasera mediante la relación que existe con la cámara frontal y las proporciones de la cara del usuario.

Uno de los primeros dispositivos en utilizar *tracking* óptico fue el Nintendo Wiimote ([353]). Este sistema se compone de un control especial que contiene una cámara infrarroja. Por otro lado, el sistema dispone de una tira especial con LEDs infrarrojos que se encontrará fija en la habitación. El usuario tendrá este control en su mano y la cámara detectará los LEDs infrarrojos, de tal forma de estimar la posición y orientación del control. Además, este control dispone de sensores inerciales (acelerómetro y giroscopio) para complementar la detección de la cámara infrarroja y poder interpretar distintos tipos de gestos. El dispositivo Wiimote puede verse en la Figura 7.7.



Figura 7.7: Sistema Wiimote ([353]).

La mayoría de los sistemas de Realidad Aumentada utilizan *tracking* Videométrico. Esto se debe a que utilizan una cámara que siempre se encuentra buscando y detectando marcadores especiales para poder inferir su posición en el mundo.

2.5. TRACKING ULTRASÓNICO

Los sensores ultrasónicos funcionan con un principio similar a los radares o sonares que evalúan la información de un objetivo mediante la interpretación de ecos provenientes de una onda de radio o de sonido, respectivamente. Los sensores ultrasónicos generan ondas de sonido de alta frecuencia y evalúan los ecos que son recibidos por el sensor. Los sensores calculan el intervalo de tiempo entre el envío de la señal y el eco recibido para determinar la distancia al objeto. De esta forma, los sensores ultrasónicos pueden utilizarse para detectar el movimiento de objetos. La posición y la orientación se estima mediante el uso de múltiples receptores y técnicas de triangulación.

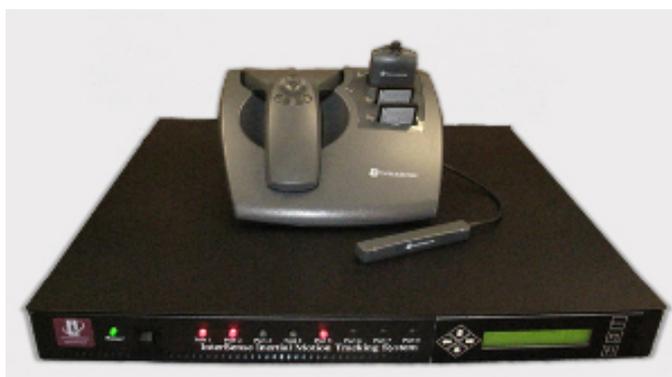


Figura 7.8: InterSense 900 ([399]).

Una de las mayores ventajas de estos sistemas es su bajo costo. Además, como los sensores ultrasónicos utilizan sonido en lugar de luz, pueden trabajar en aplicaciones donde los sensores de luz no funcionarían. Sin embargo, suelen trabajar en cortos rangos de distancia y poseen baja frecuencia de muestreo. Además, la precisión disminuye si hay superficies acústicamente reflectantes en la sala. Los sonidos externos del ambiente también pueden interferir con la señal.

El InterSense 900 es un dispositivo de *tracking* ultrasónico que se ha utilizado para todo tipo de aplicaciones, incluso sistemas de RV para hacer *tracking* del usuario y de sus manos ([399]). El sistema, que puede verse en la Figura 7.8, realiza *tracking* de posición y orientación con baja latencia.

2.6. TRACKING INERCIAL

Los sistemas de *tracking* inercial utilizan una variedad de dispositivos de medición inerciales, como giroscopios y acelerómetros. Estos dispositivos proveen medidas derivativas con alta frecuencia de muestreo, por lo que pueden integrarse para obtener la información sobre la posición y la orientación.

Aunque pareciera que los cambios en la posición pueden ser medidos confiando solo en giroscopios (para orientación) y acelerómetros (para medir la distancia desde un punto inicial), se deben considerar algunos problemas técnicos. Debido a que los acelerómetros proveen medidas relativas, presentan errores de acumulación, ruido y *drift*, a lo largo del tiempo, generando información imprecisa. Por esta razón, la mayoría de los sistemas de *tracking* puramente inerciales solo trackean la orientación. Aunque los giroscopios también sufran de errores de acumulación, esto es menos severo y existen métodos para solucionar este problema.

Los sistemas que solo trackean la orientación pueden ser utilizados en RV para trackear la cabeza en aquellas ocasiones en las que el usuario esté posicionado en un punto fijo y pueda rotar su cabeza. Esto puede ser combinado posteriormente con otros métodos de *tracking* de posición, para lograr un *tracking* más completo.

Como los sistemas de *tracking* que solo utilizan Visión por Computadora generalmente manejan velocidades de movimiento limitadas, varios trabajos han propuesto la fusión con sistemas inerciales ([400–404]). Los sensores inerciales son buenos para capturar movimientos rápidos, mientras que los sensores de visión proveen referencias absolutas para resetear el error, el que se acumula rápidamente.

Particularmente, la integración de la información del acelerómetro introduce inestabilidades cuando se utiliza para estimar la posición. Esto se debe no solo a que los acelerómetros miden la aceleración, sino que también miden la gravedad. Los acelerómetros entonces suelen usarse para estabilizar solo la inclinación de la cámara, en lugar de la posición. Esto es algo común de los trackers de orientación comerciales, como el Oculus Rift ([16]), que internamente fusionan las medidas de los acelerómetros, magnetómetros y giroscopios, para brindar la orientación absoluta. Finalmente, debido a estos problemas de los acelerómetros, algunos trabajos utilizan solo giroscopios para apoyar el *tracking* de visión.



Figura 7.9: Persona utilizando el traje Rokoko y un HMD de RV ([405]).

Los sistemas inerciales son muy útiles cuando se conoce de la distancia entre ellos. Con esta información se puede generar una representación espacial de los mismos. De esta forma, sistemas como Rokoko ([405]) se componen de trajes especiales basados en sensores inerciales. Por ejemplo, al tener un giroscopio en cada articulación del cuerpo y conociendo la distancia entre dichas articulaciones, se podría estimar una buena aproximación en 3D de la posición de cada articulación trackeada, es decir, de todo el cuerpo. En la Figura 7.9 puede verse a un usuario utilizando el traje, que está diseñado para ser cómodo.

2.7. TRACKING NEURAL

El *tracking* neurológico o muscular es un método de sensando del movimiento de partes del cuerpo individuales, en relación a otras partes del cuerpo. Se colocan pequeños sensores en los dedos, en las manos o en las extremidades. Estos sensores miden los cambios en las señales nerviosas o contracciones musculares y reportan la postura de la parte trackeada al sistema de RV.

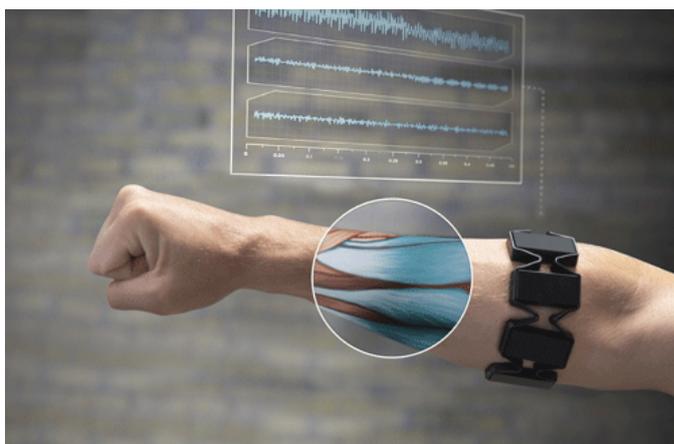


Figura 7.10: Sistema Myo que trackea distintos gestos del brazo y de la mano mediante la detección de los músculos del antebrazo.

Al monitorear los impulsos eléctricos en áreas específicas de la piel, es posible determinar la activación de los músculos que controlan la flexión de los dedos u otros movimientos similares. Esta tecnología se ha testeado en dispositivos protésicos para controlar los movimientos de las prótesis al monitorear el estímulo nervioso a través de la extremidad.

Un sistema que utiliza esta tecnología, aunque no específicamente para RV, es el sistema Myo ([406]). Este sistema interpreta los impulsos eléctricos de los músculos del antebrazo permitiendo controlar una computadora mediante gestos. En la Figura 7.10 puede verse qué tan pequeño y discreto es este sistema, ya que está diseñado para ser utilizado todo el día. En un futuro, los sistemas de este estilo podrían usarse para reemplazar otros sistemas de *tracking* específicos para RV.

2.8. TRACKING HÍBRIDO

Como se mencionó anteriormente y como se verá en la Sección 5, el *tracking* híbrido utiliza más de una tecnología de *tracking* para incrementar así la precisión, reducir la latencia y proveer una mejor experiencia de interacción 3D. En general, se suelen combinar distintas tecnologías individuales de *tracking* para compensar por las debilidades de las demás.

La mayor dificultad con el *tracking* híbrido es que se incrementa tanto la cantidad de componentes a utilizar como la complejidad del sistema. Sin embargo, según cada sistema particular, si se logra incrementar la precisión del sistema la complejidad está justificada.

3. TRACKING DEL CUERPO

La mayoría de los sistemas de RV necesitan trackear el cuerpo del usuario para poder crear una representación adecuada dentro del mundo virtual. El *tracking* del cuerpo se refiere al sensado de la posición y las acciones de los participantes. Los componentes particulares del movimiento que son trackeados dependen de la parte del cuerpo y de cómo está implementado el sistema. Cualquier parte del cuerpo podría ser trackeada con uno o más grados de libertad, siempre y cuando exista un mecanismo de *tracking* con un tamaño y peso apropiado.

3.1. POSTURA CORPORAL Y GESTOS

Se conoce como postura o pose a la posición estática de una parte del cuerpo o grupo de partes, como un dedo extendido o un puño cerrado. Un movimiento de usuario específico que ocurre a lo largo del tiempo se conoce como gesto. Los gestos se utilizan como interfaces intuitivas; por ejemplo, cerrar el puño puede indicar el deseo de agarrar algún objeto.

Las posturas y los gestos proveen un repertorio expandido del cual se pueden derivar comandos de entrada al sistema. Sin embargo, como no todas las personas tienen la misma intuición, los usuarios podrán requerir algún entrenamiento para poder utilizar un sistema de reconocimiento de gestos.

Uno de los sistemas más utilizados en estudios de Interacción Humano Computadora para detectar poses y gestos de los usuarios, es el Microsoft Kinect ([407]). Este dispositivo puede verse en la Figura 7.11 y dispone de un emisor y de un sensor de profundidad infrarrojo que permite detectar así a las personas que se encuentran frente al mismo. De esta forma, el sistema puede detectar las poses y los gestos de las personas.



Figura 7.11: Microsoft Kinect, sistema que puede trackear la posición estimada de una persona completa.

El sistema Kinect es de utilidad para detectar los movimientos de un usuario de un sistema de RV y trasladar dichos movimientos a un avatar virtual que el usuario podrá ver, incrementando de esta forma la inmersión del sistema. El inconveniente principal del Kinect es su baja precisión y su retardo en la adquisición de datos, lo que puede disminuir la inmersión del sistema. Sin embargo, la nueva versión del Kinect mejora notoriamente estas cuestiones. Además, el Kinect requiere de una PC para funcionar, por lo que no podría utilizarse en sistemas de RV completamente móviles de forma cómoda.

3.2. TRACKING DE LA CABEZA

En casi todos los sistemas de RV se trackea la cabeza del usuario ya que la posición y orientación de la cámara virtual debería cambiar respectivamente según la posición y orientación de la cabeza del usuario.

Los HBDs requieren al menos el *tracking* de la orientación. Mientras los usuarios rotan sus cabezas, la escena debe adaptarse y renderizarse apropiadamente según la dirección de la visión, sino los usuarios no estarán físicamente inmersos. La mayoría de los HBDs utilizan giroscopios para trackear la orientación de la cabeza de los usuarios. Como el mover la cabeza es un movimiento muy natural, trackear además la posición de la cabeza incrementa considerablemente la inmersión.

Los sistemas visuales estacionarios, como los de escritorio, necesitan la posición de la cabeza para actualizar la representación visual de la escena virtual. En este caso, como la pantalla es generalmente pequeña, el sistema puede asumir que el usuario estará mirando al centro de la misma; de esta manera, no sería necesario trackear la rotación de la cabeza del usuario. Cuando se utilizan pantallas más grandes, como en los sistemas de proyección, sí sería necesario trackear la rotación de la cabeza del usuario, porque no se puede asumir que el usuario esté mirando siempre al centro de la pantalla.

3.3. TRACKING DEL TORSO

El *tracking* del torso es algo muy poco común en sistemas de RV. Esto se debe a que no existen muchas aplicaciones que requieran este tipo de *tracking*. Además, generalmente se asume que el torso apunta al mismo lugar que apunta la cabeza del usuario, por lo que trackear la cabeza generalmente es suficiente.

La posición y la rotación del torso puede ser requerida cuando se muestra el avatar del usuario pero generalmente se deduce de las posiciones de la cabeza y de las manos. Sin embargo, el torso es un mejor indicador de la dirección a la que apunta el cuerpo que la cabeza o las manos.

Algunas aplicaciones pueden requerir que los usuarios caminen en una dirección particular mientras están mirando hacia otro lado. En ese caso, la cabeza y el torso deben trackearse por separado. Además, existen aplicaciones en donde el torso juega un rol crucial a la hora de interactuar con el mundo virtual. Éste es el caso de ciertas aplicaciones de deportes, que requieren realizar movimientos del cuerpo especiales.

El sistema Kinect ([407]), aunque está pensado para trackear todo el cuerpo, puede utilizarse para trackear solo el torso del usuario, independientemente de la cabeza y las manos. Otros sistemas que utilicen sensores inerciales como giroscopios también pueden utilizarse para este propósito.

3.4. MANOS Y DEDOS

Trackear la mano y los dedos generalmente se hace para que los usuarios puedan interactuar con el mundo virtual de manera más inmersiva. Una mano puede ser trackeada mediante una unidad de *tracking* inercial en la muñeca, por ejemplo. Sin embargo, si se necesita información sobre la forma o movimiento de la mano, se podría utilizar un dispositivo de entrada de tipo guante para trackear la posición de los dedos del usuario y otras flexiones de la mano. En este caso, el *tracking* de posición de la mano se suele montar directamente en un guante.

Los guantes con sensores que detectan dobleces (bend-sensing) son dispositivos puramente pasivos que se usan para detectar las posturas de la mano, de los dedos y algunos gestos. Los guantes típicamente tienen entre 5 y 22 sensores. Por ejemplo, un guante con 5 sensores generalmente mide una articulación en cada dedo, mientras que un guante con 18 sensores puede medir al menos dos articulaciones en cada dedo y la posición y orientación de la muñeca, entre otros.

Los dispositivos tipo guantes presentan algunas desventajas. Primero, son difíciles de poner y sacar. Luego, estos dispositivos suelen ser difíciles de calibrar y, peor aún, difíciles de mantener calibrados para que el sistema mida de forma precisa.



Figura 7.12: Sistema Leap Motion para trackeo preciso de la mano y los dedos del usuario.

La mayoría de estos dispositivos no trackean la información completa sobre la forma de la mano. Generalmente reportan información sobre las articulaciones de la misma.

El Leap Motion ([408]) es un dispositivo que surgió hace pocos años y está específicamente diseñado para detectar las manos del usuario de manera muy precisa y veloz. De forma similar al Kinect, utiliza sensores infrarrojos para detectar las manos del usuario y algoritmos especializados para trackear la posición y la orientación de la mano y de los dedos. Actualmente está siendo utilizado además para sistemas de RV y soporta la integración con el Oculus Rift, por ejemplo. En la Figura 7.12 puede verse el Leap Motion detectando las manos del usuario y en la Figura 7.13 puede verse la integración con el Oculus Rift en una aplicación de RV.



Figura 7.13: Sistema Leap Motion siendo utilizado junto con el Oculus Rift. El usuario puede ver e interactuar con sus manos en RV.

3.5. TRACKING DE OJOS

Actualmente, los sistemas gráficos deben renderizar la imagen en toda la pantalla con el mejor detalle posible ya que el usuario podría estar mirando cualquier parte de la pantalla. Si se supiese con precisión dónde está mirando el usuario, se podría renderizar con lujo de detalle esa parte y disminuir el detalle del resto. De esta forma se reduciría notoriamente la carga computacional del sistema. Por éste y otros motivos, se busca la forma de poder realizar el *tracking* de los ojos del usuario en los sistemas de RV.

Para realizar el *tracking* de los ojos del usuario se utilizan dispositivos que determinan hacia dónde está mirando. Esta tecnología se basa principalmente en técnicas de Visión por Computadora. Los dispositivos ópticos trackean las pupilas del usuario utilizando reflexiones en la córnea que son detectados por una cámara. Estos dispositivos pueden ser colocados sobre el cuerpo o embebidos en la pantalla, logrando una interfaz mucho menos molesta.

Otra técnica es la electro-oculografía, que mide la diferencia de potencial eléctrico en la piel utilizando electrodos alrededor del ojo y poniendo objetos mecánicos u ópticos en lentes de contacto que son colocados directamente en el ojo.

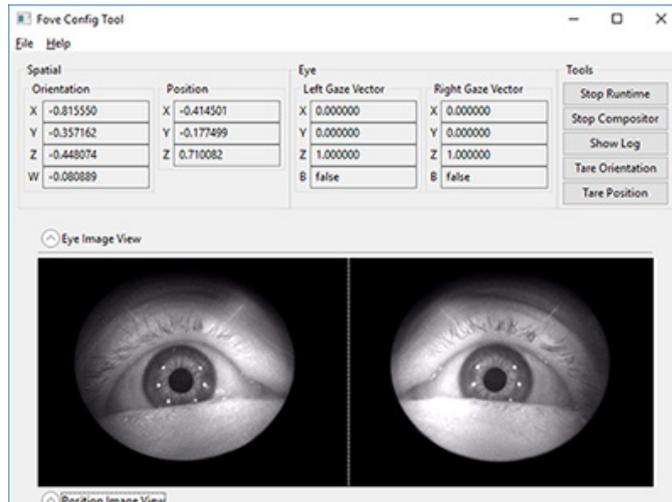


Figura 7.14: Detección de ojos del sistema Fove.

Los sistemas de *tracking* se suelen utilizar tanto como herramientas de evaluación como para interactuar con una aplicación. Estos dispositivos se utilizan para recolectar información acerca de los movimientos del ojo, para obtener patrones de uso que ayuden a mejorar una interfaz o para entrenamiento en tareas de inspección visual.

El Fove ([409]) es uno de los primeros HMDs en incluir un sistema de *Eye tracking*. Este dispositivo tiene las mismas dimensiones y comodidades que un HMD corriente pero incluye el hardware y software necesarios para trackear los ojos del usuario. En la Figura 7.14 se puede apreciar la aplicación de configuración del sistema Fove y cómo detecta los ojos del usuario.

3.6. RECONOCIMIENTO DE VOZ

Las interfaces más comunes que utilizan la boca son definitivamente los sistemas de reconocimiento de la voz y del habla. Los sistemas de reconocimiento de voz proveen una oportunidad de comunicación más natural con los sistemas de computación. Esto es de especial importancia para las aplicaciones de RV, donde uno de sus objetivos es lograr una interfaz lo más natural posible.

En general, los sistemas de reconocimiento de voz funcionan mejor cuando son entrenados para un usuario particular. Si la aplicación está hecha para ser utilizada por varios usuarios que no tendrán tiempo de entrenar el sistema con su voz, será necesario que el sistema no dependa de este entrenamiento. Si la aplicación va a ser utilizada por unos pocos especialistas que utilizarán un vocabulario complejo, entonces un sistema entrenado funcionará mejor en este caso.

Una de las consideraciones a tener en cuenta es cuándo el sistema de reconocimiento de voz debe encontrarse escuchando y cuándo no. La solución más simple es dejar que el sistema siempre esté escuchando al usuario. Sin embargo, esto puede causar problemas cuando el usuario quiere hablar con otras personas. En este caso el sistema estará constantemente tratando de entender la conversación del usuario, siendo esto más complejo e ineficiente. Una opción más inteligente sería permitir una atención selectiva a la voz del usuario.

El método “Presionar y Hablar” utiliza algún botón para activar el sistema de reconocimiento de voz. De esta forma, el usuario decide cuándo quiere que el sistema reconozca sus comandos de voz. Otro método se conoce como “Nombrar y Hablar” y es cuando el usuario dice alguna palabra de activación seguida por una instrucción. Este es el caso de la aplicación Google Now de Google, en la que el usuario primero debe decir “Ok Google Now” y luego dar una instrucción, como por ejemplo “qué hora es en España?”. Otro método más relacionado con RV es el “Mirar y Hablar”. De esta forma el usuario deberá mirar algún objeto virtual significativo que detectará cuándo lo están observando y procederá a reconocer los comandos de voz del usuario. Un ejemplo de esto sería un sistema de chat virtual, en donde el usuario debe mirar a la persona (es decir al avatar) con quien desea hablar.

La ventaja de la comunicación por voz es que es una forma de comunicación natural y cómoda. En situaciones donde se deben utilizar las manos para otras tareas, el reconocimiento de voz es ventajoso porque el usuario puede disparar comandos mientras utiliza sus manos. Sin embargo, debido a la naturaleza de la voz, hay varias tareas y situaciones en donde los sistemas de reconocimiento de voz no son la mejor solución. Aquellas tareas que requieran entradas rápidas, funcionan mejor con controles físicos. Uno de los inconvenientes principales es que, como las personas están acostumbradas a conversar con seres inteligentes, van a asumir lo mismo al comunicarse con entidades artificiales en mundos virtuales. Al no producirse un diálogo natural, esto puede romper la inmersión del sistema.

3.7. OTROS

Otras partes y aspectos del cuerpo pueden ser trackeados para controlar varios aspectos del mundo virtual. Entre ellos podemos encontrar la temperatura corporal, la transpiración, el ritmo cardíaco, el ritmo de respiración, el estado emocional y ondas cerebrales. Estos factores pueden medirse monitoreando la condición del participante mientras experimenta el mundo virtual, o pueden utilizarse también para interactuar con el mundo.

Los avances tecnológicos actuales permiten la creación de sistemas de *tracking* más económicos, pequeños y livianos. Esto permite que se pueda realizar el *tracking* de partes del cuerpo que hasta hace unos años ni se hubiese pensado. Por ejemplo, el sistema BinaryVR ([410]) provee una solución de *tracking* en tiempo real de las expresiones faciales del usuario. Una ventaja de este sistema es que es compatible con cualquier HMD como podemos ver en la Figura 7.15. Esta información puede luego trasladarse al avatar virtual, como se ve en la Figura 7.16, logrando así sensaciones más inmersivas.



Figura 7.15: Dispositivo BinaryVR ([410]) colocado en diversos visores comerciales, demostrando su facilidad de adaptación.



Figura 7.16: Ejemplo de detección facial del BinaryVR ([410]). Cuando el usuario sonr e, el personaje virtual tambi n lo hace acordemente.

4. DISPOSITIVOS DE TRACKING

Los dispositivos físicos son otro elemento de la interfaz entre el usuario y el mundo virtual. Estos dispositivos generalmente se diseñan para aplicaciones particulares, aunque algunos se diseñan para ser genéricos para distintas aplicaciones. En esta sección veremos algunos de estos dispositivos, prestando especial atención a cómo son utilizados para trackear al usuario. Cuestiones como sensaciones tangibles o hápticas referentes a estos dispositivos fueron presentadas en el Capítulo 6.

4.1. CONTROLES Y PROPS

Los controles físicos son aquéllos que tienen botones o interruptores individuales para que los usuarios puedan proveer entradas directamente a los sistemas de RV. Por ejemplo, los controles Touch del Oculus Rift ([380]), además de ser trackeados mediante luces infrarrojas, también disponen de botones que serán detectados solo cuando el usuario los presione. Estos dispositivos pueden ser diseñados genéricamente, permitiendo que puedan ser utilizados en diversas aplicaciones. No se debe olvidar que, por ejemplo, el simple hecho de detectar que el usuario ha pulsado un botón también se considera *tracking*. El sistema recibirá esta información y realizará el procesamiento correspondiente.

Un *prop* es un objeto físico que se utiliza para representar algún objeto dentro del mundo virtual. Por ejemplo, en una experiencia de RV, un círculo de cartón puede representar un volante o un pequeño cilindro puede representar un sable láser. Algunos *props* sólo son formas simples (esferas, conos, platos, etc) que pueden aproximarse a cualquier número de objetos.



Figura 7.17: Props del Nintendo Wii con forma de raqueta, volante y palo de golf.



Figura 7.18: Props del PS Move con distintas formas para videojuego de deportes.

Los *props* permiten una interacción más flexible e intuitiva con los mundos virtuales. El principal objetivo del uso de *props* es el crear una interfaz que los usuarios puedan manipular de la forma más natural posible. Otro beneficio es que el realismo que aporta la incorporación de dispositivos hápticos a los objetos virtuales hace que el resto del mundo virtual parezca incluso más real. Distintos *props* pueden verse en la Figura 7.17 y en la Figura 7.18.

4.2. PLATAFORMAS

Las plataformas son grandes estructuras físicas que se utilizan para conectar al usuario con el mundo virtual. Son dispositivos donde se sitúa el usuario y son diseñadas para replicar dispositivos del mundo real encontrados en el mundo virtual, o simplemente proveer un lugar donde sentarse o pararse.

En nuestra vida cotidiana, la forma más intuitiva de navegación a través de nuestro entorno es a pie. Sin embargo, la mayoría de los sistemas modernos de RV no ofrecen una forma de caminar libremente a través de los mundos virtuales, y aquéllos que lo hacen, solo lo hacen de forma muy restrictiva. Por otro lado, en la mayoría de los casos, los usuarios simplemente navegan mediante la utilización de teclados, mouse, joysticks, o algún otro dispositivo de entrada similar. Esto genera un conflicto sensorial, ya que el usuario visualmente percibe que se está moviendo, pero físicamente no lo está. Estudios de comportamiento sugieren que este conflicto sensorial impide la formación de una representación espacial precisa, perjudicando así el desempeño en la navegación. Además, este conflicto incrementa el riesgo de generar *simulator sickness* en los usuarios.

Sherman et al. ([411]) presenta distintas categorías para clasificar los distintos tipos de plataformas. Una Plataforma Anillo es una plataforma donde el usuario está sujeto con un anillo alrededor de su cintura para restringir su posición.

Las Plataformas Ambulatorias son todas aquéllas que requieren que el usuario mueva su cuerpo para poder moverse por el espacio virtual. Esto incluye dispositivos que requieren que el usuario camine, ande en bicicleta o utilice una silla de ruedas, por ejemplo. Las caminadoras permiten que el usuario pueda caminar sobre ellas y generan la sensación de estar caminando en el mundo virtual de manera natural. La plataforma puede tener sensores en su base u otro tipo de sensores sobre el usuario para poder trackear los movimientos del usuario y mapear dichos movimientos al mundo virtual.

Las Plataformas Vehículos son aquéllas en donde el usuario puede sentarse o estar parado para controlar un vehículo virtual. Es muy común que estas plataformas provean controles muy realistas, como es el caso de los simuladores de vuelo. Algunos simuladores de conducción ponen al usuario dentro de un auto real en donde se controla un auto virtual. Algunos de estos sistemas disponen de actuadores que mueven e inclinan el asiento para generar más inmersión en el usuario.

El uso de una plataforma para RV no se limita al display visual. Hay ejemplos de plataformas que utilizan tanto HMDs como displays estacionarios o displays de proyección. Sin embargo, algunas plataformas pueden estar diseñadas para un tipo de display específico.

5. REVISIÓN DE LA LITERATURA

5.1. TIPOS DE TRACKING

En cuanto a los diferentes tipos de *tracking*, el *tracking* magnético ha sido muy utilizado en los sistemas de RV. Huang et al. ([412]) propusieron un sistema de *tracking* de movimiento que utiliza marcadores pequeños, livianos e inalámbricos. Mediante un principio de inducción electromagnética, excitan pequeñas bobinas que son trackeadas por un arreglo de otras bobinas. Su trabajo sugiere un *tracking* con 6-DOF y con una velocidad razonable. Estos pequeños marcadores son utilizados para trackear pequeños objetos o incluso los dedos de la mano. En la Figura 7.19 se puede ver como se trackean pequeñas esferas de colores. Su principal inconveniente es que el sistema requiere de mucha calibración.

En [413] presentan un sistema de *tracking* personal basado en campos magnéticos AC. Una antena transmisora de bajo poder actúa como referencia para tres módulos de *tracking*. Un módulo se une al HMD y trackea la posición y orientación de la cabeza del usuario, mientras que los otros trackean las manos del mismo. Su sistema logra una precisión de 1 mm en la posición y 0.1 grados en la orientación a una distancia menor a 2 m de la antena. La precisión va disminuyendo hasta llegar 20 cm de posición en una distancia de 2 m.

El *tracking* óptico ha sido la técnica de *tracking* de posición más utilizada para sistemas de RV debido a su alta confiabilidad. Aunque el usuario deba estar siempre frente a la cámara, esto es algo común en la mayoría de los sistemas.

En [414] utilizan una cámara fija en el techo apuntando hacia abajo para localizar unos marcadores especiales. De esta forma, al tener un marcador en la cabeza, el sistema trackea la posición de la cabeza del usuario. Esto lo usan para complementar en *tracking* inercial ya existente en los HMDs comerciales. El principal inconveniente es que agregan demasiada complejidad para generar resultados poco significativos.



Figura 7.19: Tracking de pelotas de colores en una bolsa de tela ([412]).

En [415] utilizan el sistema OptiTrack para realizar el tracking de la cabeza del usuario dentro del sistema CAVE. Sistemas de este estilo son generalmente una buena elección para entornos acotados, como es el caso del CAVE. Esto puede verse en la Figura 7.20.

Park et al. ([416]) desarrollaron un sistema y un algoritmo para reconocer las posiciones de la mano a la distancia, e incluso en la oscuridad. Utilizan luces infrarrojas y el mapa de profundidad provisto por una Kinect, junto con una cámara infrarroja de alta resolución. Obtuvieron una precisión de alrededor de un 92 %.

En [417], proponen un método para posicionar un conjunto de cámaras en un espacio de trabajo dado, para maximizar la visibilidad de los puntos con un error de reconstrucción mínimo, incluso en presencia de oclusión dinámica.

Con el avance tecnológico, las cámaras están siendo cada vez mejores, con más resolución, más pequeñas y livianas. Esto hace que el *tracking* videométrico se haya convertido en una de las técnicas de *tracking* más utilizadas.

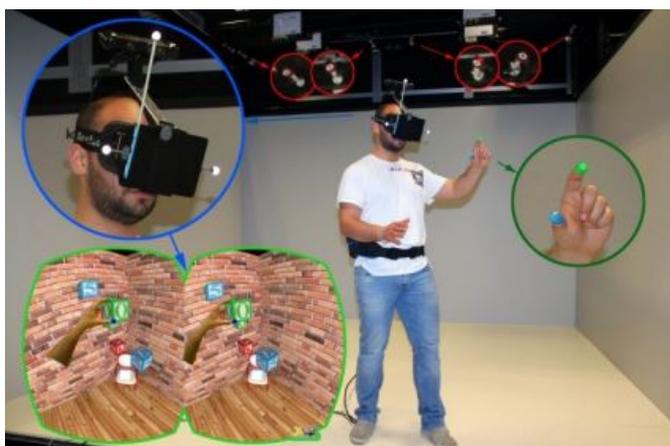


Figura 7.20: Tracking de la cabeza del usuario mediante el sistema óptico OptiTrack ([415]).

García et al. ([418]) presentan una técnica muy interesante. Proponen un sistema de Realidad Aumentada con ambientes virtuales interactivos para rehabilitación de la enfermedad de Parkinson. Al cubrir por completo una habitación con marcadores especiales como se ve en la Figura 7.21, el sistema de RV puede trackear la posición de la cámara respecto de los mismos. Entonces, si el usuario dispone de un HMD con una cámara, podrá detectar su posición en la habitación y actualizar la representación virtual de manera acorde.

Esta misma técnica puede utilizarse tanto en Realidad Aumentada como en RV ya que, al tener una posición 3D del usuario dentro del espacio, se lo puede ubicar dentro de cualquier mundo virtual. Sin embargo, se debe tener en cuenta que la aplicación debe conocer con precisión la ubicación de los marcadores con respecto a la habitación. Además, este tipo de aplicaciones son costosas computacionalmente ya que el sistema se encontrará buscando los marcadores en todo momento, por lo que debe considerarse el dispositivo que se utilizará.

Carozza et al. ([419]) utilizan Videometría con sensores inerciales para trackear una persona en una habitación determinada. Utilizan distintas técnicas de recolección y procesamiento de las imágenes.



Figura 7.21: Espacio físico con marcadores en las paredes y en el piso ([418]). De esta forma el sistema puede trackear la posición del usuario en cualquier punto de la habitación.



Figura 7.22: Escenario virtual de cruzar la calle ([418]). Mediante los marcadores situados en la habitación de la Figura 7.21, el participante puede caminar por este mundo virtual.

Maesen et al. ([420]) realizaron un sistema que determina la posición de un objeto con 6 DOF mediante el uso de tiras de LEDs montadas en el techo de una habitación. Una cámara montada en la cabeza del usuario observa estos patrones de LEDs. Así, la cámara determina su propia posición de manera independiente, por lo que no hay restricción en el número de objetos trackeados. El sistema tiene un error de unos pocos milímetros en posición y menos de un grado en orientación. El principal problema de este sistema es la portabilidad, ya que deben instalarse LEDs en disposiciones especiales en el techo de la habitación donde se desee usar el sistema.

Tecchia et al. ([415]) utilizan una cámara RGBD montada sobre el HMD para detectar y trackear las manos del usuario. Este tipo de cámaras detectan profundidad, por lo que puede distinguir fácilmente las manos del resto del mundo.

El *tracking* puramente inercial no ha sido muy utilizado en sistemas de RV. Si bien se ha utilizado en otras áreas, como por ejemplo la robótica, diversos estudios realizados sugieren que los sensores inerciales no son útiles para realizar el *tracking* de posición de un usuario ya que acarrean mucho error a lo largo del tiempo. Sin embargo, el giroscopio es utilizado en casi todos los sistemas de RV actuales para poder medir la orientación de los HMDs.

En [421], además de rastrear el teléfono mediante el uso de GPS, utilizan el acelerómetro para estimar la actividad que está realizando el usuario. En [419] y en [422] combinan los sensores en el display (acelerómetro y giroscopio) con técnicas de Videometría para hacer *tracking* del usuario.



Figura 7.23: Dispositivo Mask unido a un Oculus Rift ([423]).

El *tracking* neural es aquél que puede sentir señales eléctricas del cuerpo, ya sea musculares o cerebrales por ejemplo, y actualmente están surgiendo alternativas muy atractivas en el campo de la RV. Sistemas como Mask ([423]) están surgiendo para lograr una unión más emocional entre los usuarios y el mundo virtual. Este sistema utiliza técnicas de *Machine Learning* y procesamiento de bio-señales para decodificar expresiones faciales incluso antes de que aparezcan en la cara del usuario, replicando instantáneamente esa expresión en el avatar virtual. Mediante el análisis de los impulsos eléctricos de la cara, el sistema crea un patrón neural de las expresiones del usuario sin necesidad de ningún tipo de entrenamiento o calibración. En la Figura 7.24 se muestra un ejemplo de esto con dos usuarios diferentes. Otra ventaja del sistema es que es liviano y puede ser acoplado a cualquier HMD, desde aquéllos que utilizan dispositivos móviles hasta los que no. En la Figura 7.23 se pueden apreciar los sensores del sistema Mask acoplados al Oculus Rift.



Figura 7.24: Representación de dos avatars con distintas expresiones faciales ([423]).

El sistema Emotiv Epoc ([424]) es otra tecnología, especializada en el área denominada *Brain-Computer Interface*². Este sistema es un Electroencefalograma inalámbrico de alta resolución que puede utilizarse para interpretar las señales cerebrales del usuario. Este tipo de sistemas se ha utilizado principalmente para analizar las señales neurológicas de los usuarios en determinadas situaciones especiales ([425–427]), pero muy pocas veces se ha utilizado como un dispositivo de interacción con los mundos virtuales. En la Figura 7.25 puede verse un usuario utilizando este dispositivo para jugar un videojuego. Zhao et al. ([428]), por ejemplo, utilizaron un sistema de este tipo para intentar conducir un auto en RV. Sugieren que es posible lograr este tipo de tareas con estos sistemas.

Como los sistemas de GPS no funcionan en ambientes de interior, muchos estudios se han enfocado en el *tracking* de interior utilizando diversas técnicas como Wifi ([429]), Radiofrecuencia ([430, 431]), Bluetooth ([432]), visión infrarroja ([433]), magnética ([434]), *tracking* óptico ([435, 436]) y métodos acústicos, como ultrasonido ([437–440]). Cada método tiene sus ventajas y desventajas en término de precisión, área de cobertura, costo, infraestructura, disponibilidad y privacidad ([441–443]).

²Interfaz Cerebro-Computadora

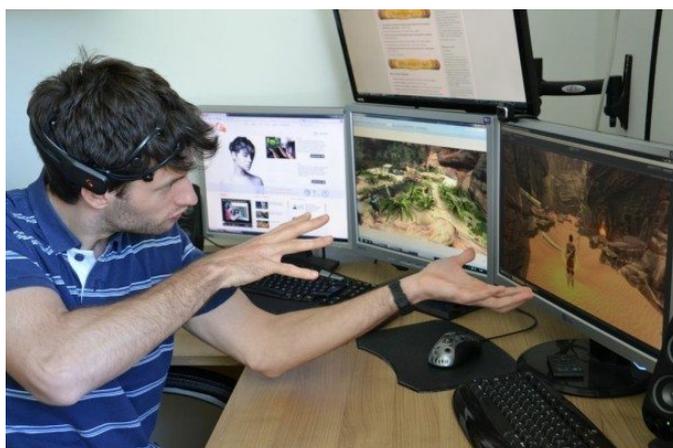


Figura 7.25: Emotiv Epoc siendo utilizado para controlar un videojuego.

Hodgson et al. ([444]) desarrollaron un HMD para RV con capacidad de *tracking* de posición. Su sistema utiliza GPS, por lo que solo puede ser utilizado en exteriores. Además tiene baja precisión con un error de no menos de un metro.

Shin et al. ([445]) utilizan también ultrasonido para trackear a un usuario dentro de un ambiente determinado. Se enfocan en maximizar la precisión al estimar la posición del usuario mediante el uso de varios sensores ultrasónicos omnidireccionales que cubren un gran espacio.

Existen sistemas comerciales como el Pozyx ([446]), por ejemplo, que utilizan ultrasonido de alta banda para hacer el *tracking* espacial de los objetos. Sin embargo, sistemas como éste pueden ser muy costosos (más de 1000 euros).

En los últimos años se han implementado diferentes métodos para representar posicionamiento en interiores utilizando las redes Wifi existentes ([447-451]). RADAR ([430]), desarrollado por Microsoft, fue uno de los primeros sistemas en utilizar este método. La principal desventaja de esta técnica es que presenta un error de algunos metros, lo que la vuelve muy imprecisa a la hora de trackear al usuario para representarlo en un mundo virtual.

Hemos mencionado que el uso de dispositivos móviles para sistemas de RV se ha popularizado en los últimos años. Fudickar et al. ([452]) discuten sobre el uso de las redes Wifi para trackear dispositivos móviles en interiores. Su propuesta utiliza señales pre-grabadas que luego son comparadas en tiempo real con la señal actual detectada por el dispositivo móvil, mediante la utilización de distintos algoritmos. El principal problema es que logran localizar el dispositivo con un error de no menos de 5 metros, lo que lo vuelve inviable para sistemas de RV donde los pequeños movimientos son importantes.

En [453], utilizan esta técnica para localizar un teléfono móvil dentro de un campus y lo combinan con Realidad Aumentada para mostrar información a los usuarios. También obtuvieron un gran error en metros.

Gao et al. ([454]) también utilizaron esta técnica de Wifi, pero utilizando técnicas de Machine Learning para estimar la posición del dispositivo móvil. Asimismo, lograron un *tracking* con un error de aproximadamente 2 metros. En [455] realizan algo similar para localizar dispositivos móviles mediante la combinación de distintas tecnologías como Wifi, GSM y Bluetooth.

Algunos trabajos como Bargh et al. ([456]) o Figueiredo et al. ([457]) solo utilizan la tecnología Bluetooth para localizar un dispositivo móvil en un espacio determinado. En [456], por ejemplo, sugieren un 98 % de precisión pero el dispositivo debe mantenerse estático por al menos 3 minutos, volviéndolo inviable.

En [457] combinaron señales de Wifi con radio frecuencia para localizar dispositivos dentro de una sala particular. De la misma forma que otros estudios, obtuvieron un error no menor a 3 metros.

Otros casos buscan determinar la posición de un dispositivo bajo tierra, más específicamente en el subterráneo ([458]). En este trabajo, no quieren depender de sensores como el GPS ya que éste no funciona bien bajo tierra. Para esto utilizan el barómetro del teléfono celular. Sin embargo, solo pueden obtener la locación del dispositivo entre estaciones, es decir con decenas de metros o hasta kilómetros de distancia.

5.2. TRACKING DEL CUERPO

Como se mencionó anteriormente, la mayoría de los sistemas de RV tratan en una primera instancia de trackear la cabeza del usuario para poder representar el mundo virtual acordemente.

En [418] y [419] utilizan una cámara para realizar *tracking* Videométrico. Por este motivo, la posición de la cámara será trackeada y se representará directamente en la posición de la cámara virtual. En este caso, se estaría trackeando la posición de la cabeza del usuario en el espacio.

Maesen et al. ([420]) utilizan una cámara en la cabeza del usuario para trackear luces LED en el techo de la habitación. En [414] trackean la cabeza del usuario mediante una cámara fija en el techo y marcadores especiales en la cabeza.

Trackear las manos y hasta los dedos del usuario ha sido también un tema muy estudiado. Desde el surgimiento de la Interacción Humano Computadora se ha buscado que el usuario pueda interactuar con la computadora mediante sus manos, logrando una comunicación más natural. Esta tendencia fue adoptada en los sistemas de RV y está siendo estudiada actualmente.

Como se puede apreciar en la Figura 7.26, el sistema presentado por Huang et al. ([412]) utiliza pequeños sensores electromagnéticos para trackear los dedos de la mano. Este sistema solo funciona en un rango espacial muy acotado y requiere de una compleja calibración.

En [418] no realizan *tracking* de las manos del usuario ya que utilizan un filtro de imagen especial para que el usuario pueda verlas, cosa que mejora notablemente la inmersión.



Figura 7.26: Sistema que utiliza sensores electromagnéticos para trackear los dedos de la mano ([412]).

En [415] el usuario puede caminar por un ambiente virtual e interactuar utilizando sus manos. Para simplificar el *tracking* de los dedos, el usuario utiliza dos dedales de colores, uno azul en el pulgar y uno verde en el dedo índice. El *tracking* de los dedos se realiza mediante una cámara RGBD que detecta los colores de los dedales y utiliza la información de profundidad para “recortar” todo aquello que esté más lejos que las manos del usuario. Así, se realizan determinados gestos según la distancia entre ambos dedales.

Park et al. ([416]) utilizan una Kinect para generar un mapa de profundidad y luces infrarrojas que son detectadas por una cámara infrarroja. De esta forma, logran detectar los distintos gestos de la mano con gran precisión, incluso cuando las manos están alejadas y cuando hay poca luz en el ambiente.

Sagayam et al. ([459]) presenta un Survey sobre posturas de manos y técnicas de reconocimiento de gestos para aplicaciones de RV.

En [460] presentan un método para reconocer posturas de manos en tiempo real, incluso cuando la luz ambiental no puede configurarse apropiadamente. Disponen de un guante especial monocromático con marcadores especiales en la palma y en cada dedo. Su sistema puede detectar al menos 96 posiciones de la mano mediante las características de los dedos.

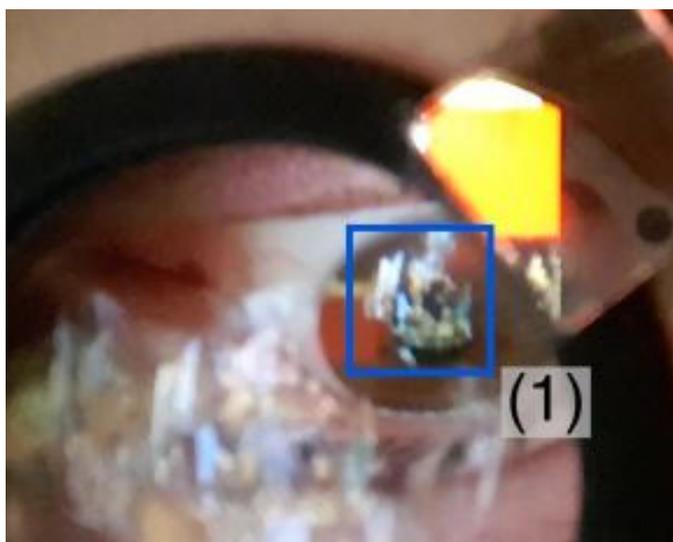


Figura 7.27: Imagen capturada por la cámara frontal del dispositivo móvil mientras detecta la pupila del usuario a través de los lentes del HMD ([461]).

Hemos visto que el *tracking* de los ojos en sistemas de RV puede resultar de gran utilidad, ya sea como un método de interacción, o de estudio del comportamiento humano. El principal inconveniente es la incorporación de un sistema de *Eye Tracking*³ a los HMDs existentes o la creación de un sistema que no resulte incómodo para el usuario. En [461] desarrollaron un sistema *Open Source* de *Eye Tracking* para HMDs basados en dispositivos móviles. Como puede verse en la Figura 7.27, este sistema utiliza la cámara frontal del dispositivo para trackear la pupila del usuario a través de las lentes del HMD. Aunque lograron interacciones limitadas, su sistema puede ser de utilidad para distintos tipos de aplicaciones.

Actualmente, están surgiendo visores de RV comerciales que disponen de sistemas de *Eye tracking* integrados. Un ejemplo es el visor Fove ([409]). La incorporación de *Eye Tracking* en RV no solo serviría para el análisis visual del usuario sino también para mejorar la performance de las aplicaciones de RV.

³Seguimiento de Ojos

5.3. DISPOSITIVOS

Un gran número de dispositivos ha sido desarrollado para lograr una comunicación más natural entre el humano y la computadora. Sin embargo, la mayoría de estos dispositivos fueron creados para interactuar con una computadora de escritorio y muy pocos han sido desarrollados puramente para un uso en RV.

Las plataformas sí han sido pensadas, en su mayoría, para ser utilizadas directamente en entornos de RV. Desde los comienzos de la RV, los investigadores han notado la importancia de poder caminar o trasladarse dentro de los mundos virtuales y gran variedad de plataformas han sido creadas para este propósito.

El Virtual Perambulator ([462]) presenta una plataforma en donde los usuarios pueden caminar de forma omnidireccional. Los usuarios utilizan dispositivos de deslizamiento en sus pies, lo que genera la sensación de caminar mientras mantienen su posición en el mundo físico. Un aro es situado sobre su cintura de tal forma de limitar su posición. En una primera instancia se intentó utilizar un arnés para sostener al usuario en lugar de un aro, que se optó por quitar posteriormente esto debido a que no permitía que el usuario pueda saltar o agacharse, además de su incomodidad evidente. Además, en un principio, en lugar de utilizar dispositivos deslizantes en los pies, se utilizaban patines. Su altura y peso ocasionaban problemas. El nuevo dispositivo de desplazamiento consiste entonces en sandalias de goma con láminas de baja fricción puestas en la suela. El material de la plataforma está seleccionado especialmente para conseguir la menor fricción con la suela. El sensado del movimiento se realiza con sensores magnéticos utilizando sistemas de terceros y con sensores de presión en la suela.

El Virtual Perambulator es uno de las primeras plataformas diseñada exclusivamente para RV, que ha inspirado el diseño de plataformas posteriores. El Virtuix Omni ([463]), otra plataforma que ha surgido recientemente y que se encuentra comercialmente disponibles, se basa en gran medida en este modelo. Puede verse su similitud en la Figura 7.28 y la Figura 7.29.



Figura 7.28: Boceto del Virtual Perambulator, una de las primeras plataformas para caminar en RV ([462]).



Figura 7.29: Virtuix Omni, una de las plataformas más modernas para caminar en RV ([463]).



Figura 7.30: Plataforma desarrollada en el VyGLab⁺ como caso de estudio del sistema Anaru ([464]).

Selzer et al. ([464]) desarrollaron un framework que conecta un HMD móvil con un microcontrolador Arduino. Como caso de estudio se desarrolló una plataforma en donde el usuario puede caminar al deslizarse sobre la misma. Esta plataforma puede verse en la Figura 7.30.

El Virtual Motion Controller ([465]) consiste en un cinturón suspendido con tiras elásticas, conectado a un aro circular que puede girar sobre sí mismo. Cuando el usuario se mueve del centro, las tiras elásticas aplican una fuerza que lo llevan nuevamente hacia el centro.

The Magic Carpet ([466]) es una plataforma construida con una matriz de cables piezoeléctricos que detectan presión, logrando así detectar la posición del usuario sobre la misma. Un par de radares doppler (sensores de movimiento de microondas) se han utilizado para sensar movimientos del usuario.

El Omni-Directional Treadmill (ODT) ([467]) es un dispositivo de locomoción que habilita un movimiento bipedal en cualquier dirección. El principio básico consiste en dos caminadoras perpendiculares, una dentro de la otra. La caminadora superior, que se compone de un arreglo de rodillos que se mueven libremente, se encuentra sobre otra cinta orientada perpendicularmente a la primera, también compuesta por un arreglo de rodillos. Cada una de estas cintas tiene aproximadamente 3400 rodillos. A diferencia de las cintas caminadoras pasivas, en donde el usuario debe hacer fuerza para moverlas, el ODT responde activamente a los movimientos del usuario, utilizando servomotores. Un brazo mecánico de seguimiento se extiende desde el techo hasta la cintura del usuario. Este brazo se utiliza para localizar la posición y orientación del usuario relativa a la plataforma, de tal forma que los servomotores puedan actuar para tratar de situar al usuario nuevamente en el centro de la cinta, cuando éste se desplace.

Además de ser un dispositivo extremadamente caro de fabricar, es bastante ruidoso. Otro inconveniente es que se mueve a una velocidad constante ajena al usuario. Entonces, distintos usuarios con velocidades distintas pueden sufrir inconvenientes como caídas, por ejemplo, en casos en los que la cinta se mueva más rápido de lo esperado. Además, si la plataforma no responde lo suficientemente rápido, el usuario podría caerse de la cinta. Por otro lado, la tecnología de seguimiento de la época no permite un rápido y adecuado seguimiento del usuario.

El ATLAS (ATR Locomotion Interface for Active Self Motion) ([468]) es una cinta caminadora bidireccional que permite que el usuario camine en forma natural sin ningún equipamiento extra. Para permitir que el usuario pueda girar en todo momento, el sistema permite que toda la caminadora pueda girar libremente. De esta forma, puede contrarrestar los movimientos del usuario cuando éste trata de girar, manteniéndolo en su lugar. El sistema estima la velocidad a la que debe correr para mantener al usuario en el centro de la cinta. Esto lo hace mediante sensores infrarrojos ubicados en el usuario, estimando su velocidad y estableciendo una velocidad opuesta en la cinta.

El Torus Treadmill ([469]) es un dispositivo que utiliza una superficie en forma de toroide como una interfaz de locomoción. Este sistema utiliza 12 conjuntos de caminadoras conectadas lado a lado operadas en dirección perpendicular. De esta forma generan una superficie infinita. Los movimientos del usuario son medidos mediante la utilización de sensores magnéticos. El piso se mueve en dirección opuesta al usuario, cancelando el movimiento de cada paso. La posición del usuario se mantiene centrada en el mundo con la ayuda del movimiento del piso controlado por computadora. El usuario puede cambiar de dirección de movimiento cuando lo desee, sin ningún tipo de problemas. Una limitación del sistema es el área donde es posible caminar. Los usuarios deben tener mucho cuidado de no llegar al borde de la cinta y no pueden realizar pasos muy largos. Para permitirles caminar más rápido, la plataforma debería ser más grande. Actualmente el tamaño de la misma es de 2m por 1.8m.

El Ground Surface Simulator (GSS) ([470]) es una caminadora especial que aprovecha las facilidades de ATLAS ([468]) para brindar a los usuarios la sensación de caminar, incluso en terrenos desparejos. Además, dispone de una cinta movable y un control de velocidad activo de la misma. Se concentra en dos métodos: cancelar el movimiento del usuario para mantener su posición y simular superficies naturales de terreno. Mediante la implementación del ATLAS, el usuario puede caminar en cualquier dirección en una superficie plana. Luego, el sistema permite el movimiento de pequeños paneles en el piso para simular distintas superficies. Para esto, la cinta se divide en distintas etapas o partes, cada una de ellas representada por un actuador situado bajo la cinta. Cuando este actuador sube o baja, esa parte de la cinta se moverá acordeamente. La altura máxima que puede subir cada una de estas etapas es de 6cm, la cual es suficiente para simular cierto nivel de superficie. La mayor desventaja del sistema es que es muy costoso y complejo.

El Gait Master ([471]) es una interfaz de locomoción que genera la sensación de estar caminando sobre superficies irregulares. El dispositivo utiliza un plato giratorio en donde se montan dos plataformas de movimiento. Éstas rastrean los pies para llevarlos a su posición inicial. De esta forma, el usuario puede caminar en un mundo virtual mientras que su posición es mantenida. Las plataformas también se pueden mover verticalmente, simulando de esta forma superficies irregulares. El plato giratorio gira de tal forma que el usuario pueda cambiar su orientación. La eficiencia del rastreo de los pies no es suficiente. El sensor de posición utiliza cuerdas y tiene un retraso de 0.3s, que genera un *offset* considerable entre el pedal y el pie.

El Omni-directional Ball-bearing Disc Platform (OBDP) ([472]) provee una forma natural para que los usuarios puedan caminar en entornos virtuales. En lugar de utilizar trackers 3d para la detección del usuario, este sistema utiliza un disco con un arreglo de sensores de presión con forma de pequeñas esferas. No se necesita ningún otro sensor, salvo un tracker para detectar los movimientos de la cabeza. Además, la forma del plato, con la ayuda de la gravedad, hacen que el usuario resbale hasta el centro del mismo. Este dispositivo es pasivo, ya que no requiere de ningún motor para impulsar al usuario. El arreglo de sensores permite detectar los movimientos del usuario y su forma permite que éste pueda caminar en forma omnidireccional. Cada sensor se compone de pequeñas bolas para reducir el rozamiento. Hay un total de 975 sensores en toda la superficie de la plataforma. Para sostener al usuario y evitar que caiga, un aro rígido se sitúa en su cintura. El arreglo de sensores devuelve información sobre qué sensores de toda la matriz se están presionando en todo momento. A partir de estos datos, un complejo algoritmo determina si el usuario se encuentra caminando o no y en qué dirección lo hace. Una desventaja es que el aro fijo restringe los movimientos del usuario; por ejemplo, el usuario no podrá saltar o agacharse. Por otro lado, los sensores contruidos para este sistema fueron hechos a mano, lo que dificulta su instalación y mantenimiento. Por último, el algoritmo de sensado no puede detectar pasos hacia los lados.



Figura 7.31: Sistema Cybersphere ([473]). El usuario camina dentro de una esfera que detecta su movimiento. La esfera también proyecta el mundo virtual sobre su superficie.

El Cybersphere ([473]) permite a los usuarios caminar, correr, saltar o gatear en cualquier dirección, mientras pueden observar un complejo ambiente virtual. Este sistema se compone de una gran esfera hueca, de unos 3.5 metros de diámetro. Los movimientos del usuario hacen que la esfera rote, permitiendo que el usuario navegue y explore el mundo virtual de forma natural. Estos movimientos rotacionales son transferidos a una esfera secundaria más pequeña, cuyos movimientos son efectivamente sensados. Estos datos son luego utilizados para actualizar las imágenes proyectadas en la esfera, dándole al observador la ilusión de caminar libremente a través de un ambiente virtual generado computacionalmente. Las imágenes son proyectadas en segmentos de la esfera exterior mediante proyectores de gran potencia. Un proyector se encuentra en el techo, mientras que otros cuatro en las paredes. La desventaja principal es que tiene un costo de construcción elevado y, como puede verse en la Figura 7.31, posee un tamaño muy grande.



Figura 7.32: Sistema Circulafloor ([475]). Las baldosas robóticas se mueven para posicionarse justo debajo de los pies del usuario.

El sistema presentado en [474] consiste en una plataforma con dos brazos que controlan dos pedales, en donde el usuario se parará. Cada brazo tiene 4 grados de movimiento. El sistema utiliza sistemas de terceros para rastrear al usuario. Las desventajas del sistema es que es muy grande y caro. Además, requiere de mantenimiento constante.

El Circulafloor ([475]) consiste de un conjunto de baldosas robóticas inteligentes que se mueven para posicionarse justo en la posición donde el usuario se encuentra caminando. Además, realizan un movimiento de tal forma que el usuario siempre mantiene su posición espacial en todo momento. Las baldosas y la posición de la persona son rastreadas mediante sensores ultrasónicos de terceros. La principal desventaja de este sistema es su costo y la baja velocidad de movimiento de las baldosas. Éstas logran posicionarse con una velocidad muy baja en comparación a la velocidad promedio en que camina una persona, por lo que ésta debe realizar pasos muy cortos y lentos para contrarrestar esto. El sistema puede verse en la Figura 7.32.



Figura 7.33: Sistema Powered Shoes ([476]). Estos patines se mueven para mantener al usuario siempre en la misma posición.

Los Powered Shoes ([476]) son una interfaz de locomoción que utiliza patines actuados por dos motores con ejes flexibles. El dispositivo es liviano y cómodo de usar. Habilita al usuario a caminar en ambientes virtuales mientras mantiene su posición en el lugar. El usuario puede cambiar su dirección si así lo desea. El movimiento de los pies es sensado por sensores ópticos. Los patines se mueven en sentido opuesto a la dirección del usuario, para que de esta forma el movimiento de los pies sea cancelado. De esta forma la posición del usuario en el mundo real se mantiene fija. La mayor desventaja es que permite al usuario caminar a un máximo de 600mm/s, velocidad mucho más lenta que el promedio. El sistema puede verse en la Figura 7.33.

El String Walker ([477]) es una interfaz de locomoción que utiliza 8 cuerdas impulsadas por motores montadas sobre un plato giratorio. De esta forma permite una forma de caminar omnidireccional, mientras la posición del usuario se mantiene. Cuatro cuerdas se conectan a cada zapato y son manipuladas por motores. Cada uno de ellos está equipado con *encoders* rotatorios y mecanismos para poder medir así la posición y orientación de los zapatos. Las cuerdas jalan con dirección opuesta a la de caminar, cancelando así el paso.

El Cyberwalk ([478]) es una cinta motorizada que se encarga de manipular la velocidad con la que gira una matriz de pequeñas bolitas de acero. Éstas se encuentran sujetas dentro de una superficie con agujeros de tal forma que permitan su libre rotación en cualquier dirección. Como puede apreciarse en la Figura 7.34, un usuario puede caminar libremente sobre una superficie de bolitas a una velocidad normal. La ventaja de este dispositivo es que el usuario no debe utilizar zapatos especiales, ni estar sujeto mediante algún arnés o aro especial. La desventaja es que se debe ser muy cuidadoso a la hora de seleccionar los materiales de las bolitas y de las suelas de los zapatos, ya que a velocidades medianamente altas, estas bolitas logran salirse de su lugar, y además, las suelas de los zapatos pueden deteriorarse rápidamente.

El Cybercarpet ([479]) es una plataforma que dispone de dos grados de movimiento: uno rotacional provisto por un plato giratorio y uno lineal generado por una cinta. Este sistema presenta una grilla de esferas de un diámetro de 800 mm y contiene 4332 bolas de 8 mm de diámetro. Los movimientos de la cinta afectan a las pequeñas esferas de la superficie para poder mantener al usuario en el centro de la misma todo el tiempo. Este plato giratorio es activado mediante un servomotor. De esta forma se alcanzan velocidades de hasta 200 cm/s. Un problema particular de este tipo de plataformas es que las pequeñas esferas pueden salirse de su lugar. Por todo esto, la combinación de los materiales de la cinta, las bolas, la alfombra y las suelas de los zapatos, es muy importante.

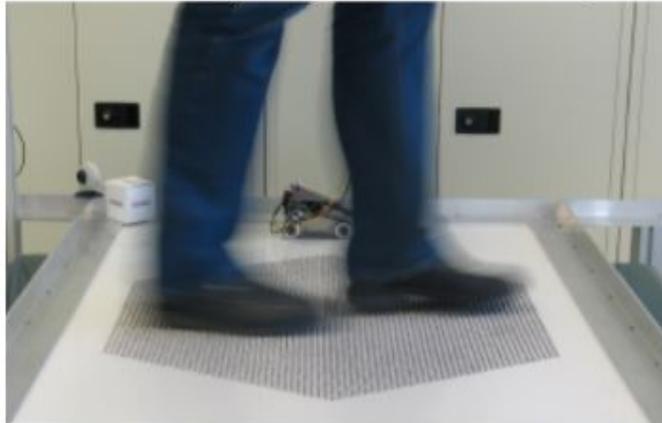


Figura 7.34: Cyberwalk ([478]). Esta plataforma está cubierta de pequeñas bolas metálicas permitiendo que el usuario deslice al caminar sobre ellas.

La Virtosphere ([480]) consiste en una esfera hueca situada sobre una plataforma especial que permite que la esfera rote en cualquier dirección en la que el usuario se mueva. Sus desventajas principales son su precio y su tamaño. Por otro lado, los usuarios tuvieron problemas al caminar rápido y al doblar. La aceleración persistente de la esfera a la hora de acelerar o frenar difiere de las del usuario, haciendo que éste pierda su estabilidad. Además, estos usuarios reportan que en todo momento, aún con una visualización en primera o tercera persona, siguen sintiendo que se encuentran dentro de una esfera gigante. Esto última reduce considerablemente la inmersión.

El Wizzdish ([481]) es una superficie con forma de plato cóncavo, en donde una persona puede caminar. Para poder hacerlo, la persona debe utilizar zapatos especiales con poca fricción. Por su forma, cuando la persona aleja su pie del centro, se deslizará y volverá naturalmente hasta el centro del plato. Para realizar el *tracking* de la posición y rotación del usuario se utiliza un sistema terciarizado. La locomoción es medida solamente con el movimiento hacia adelante del pie. Cuando un pie se deslice hacia adelante, el otro lo hará hacia atrás.

Intelligent Floor Surface ([482]) es un sistema que está diseñado para permitir a los usuarios interactuar naturalmente de pie en ambientes virtuales inmersivos. La interfaz consiste en un arreglo de baldosas rígidas, interconectadas, distribuidas sobre un área de varios metros cuadrados. Cada baldosa es capaz de sensar fuerzas aplicadas por los pies del usuario en 3 grados de libertad. Las fuerzas estimadas son usadas para rastrear las posiciones de contacto, permitiendo así una forma de que los usuarios puedan caminar en entornos virtuales. El sistema se asemeja a una gran pantalla táctil ubicada en el piso.

El Omniwalker ([483]) está basado en el OBDP y utiliza una plataforma de pequeñas bolas para que un usuario pueda caminar sobre ellas de forma omnidireccional. Esta plataforma tiene forma de cono con un centro aplanado; tiene una inclinación de 15 grados para facilitar al desplazamiento de los pies. Un soporte de goma se sitúa debajo del plato para generar una mejor amortiguación. El *tracking* es realizado usando tres cámaras web. El usuario dispone de 10 marcadores LED situados sobre sus articulaciones que son trackeados por las cámaras web y los datos son posteriormente procesados para obtener los movimientos del usuario. Uno de los inconvenientes reportados es que algunas veces las cámaras pierden rastro de los marcadores cuando el usuario rota.

Shake Your Head ([484]) no es una plataforma en sí, sino una técnica para que el usuario pueda interactuar con el sistema mediante movimientos de su cabeza, estando sentado o de pie. Estos movimientos pueden ser capturados mediante cualquier tipo de sistema de *tracking*; en este caso se utilizan webcams convencionales. De esta forma, el usuario puede percibir un desplazamiento en el mundo virtual mediante los movimientos de la cámara virtual en todos los ejes de movimiento, mejorando así la sensación de caminar.

El Joyman ([485]) es un joystick a escala humana, que permite a los usuarios indicar sus intenciones de movimiento en la navegación virtual mediante respectivas inclinaciones de su cuerpo. Joyman apunta a preservar el equilibrio-percepción para mejorar la sensación de inmersión durante tareas de movimiento virtual. El dispositivo mapea su estado en un vector de velocidad virtual que posteriormente moverá la cámara o personaje virtual. No es un dispositivo que permita que una persona camine en ambientes virtuales; sin embargo, es un buen sistema de navegación que mejora la sensación de inmersión de un joystick convencional.

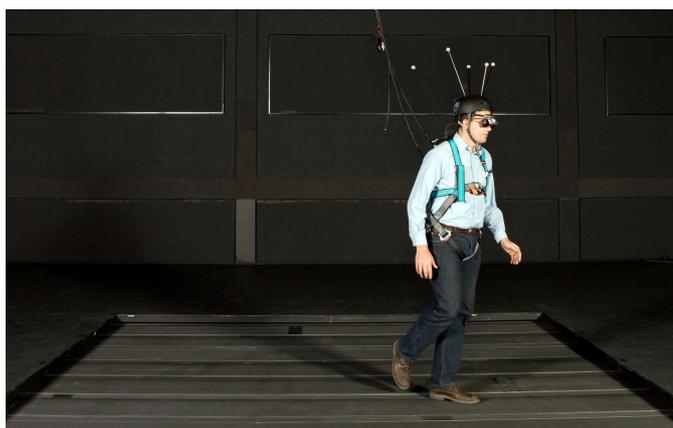


Figura 7.35: Caminadora Cyberwalk ([486]).

El Cyberwalk ([486]) es un sistema de caminadora omnidireccional, que habilita a los usuarios el poder caminar infinitamente en cualquier dirección, al mismo tiempo que se mantienen en los confines de la superficie de caminata. La caminadora consiste en 25 cintas segmentadas de 5m de longitud y medio metro de ancho, formando dos cintas entrelazadas perpendicularmente. Estas cintas llegan a velocidades de 2m/s y 3m/s, aunque se limitan a una velocidad de 1.4 m/s por cuestiones de seguridad. Todo el sistema pesa aproximadamente 12000 kg. El sistema se encuentra embebido en un piso elevado, lo que solo deja una superficie para caminar de 4x4m. Los usuarios deben utilizar un arnés de seguridad conectado al techo para prevenir caídas. La posición del usuario con respecto a la cinta es trackeada y la velocidad de la cinta es regulada de acuerdo a esta posición y a la forma de caminar del usuario. De esta forma, se requiere una superficie lo suficientemente grande para que el usuario pueda caminar sin caerse de la cinta. Como puede verse en la Figura 7.35, la mayor desventaja es que es un dispositivo grande y pesado.



Figura 7.36: Cyberith Virtualizer ([487]), una de las plataformas más modernas para caminar en RV.

El Cyberith Virtualizer ([487]) puede verse en la Figura 7.36 y es un dispositivo que permite que los usuarios puedan caminar a través de cualquier ambiente virtual. Combina un principio de baja fricción y sensores de alta precisión. La elección de los materiales de las suelas de los zapatos y de la plataforma habilita un buen deslizamiento de los pies, sin ser demasiado resbaladizo. El usuario se encuentra fijo en un sistema de cinturones, que está directamente conectado a un aro externo, que puede rotar hasta 360 grados. Esto permite que el usuario pueda caminar y correr sin cambiar su posición real. Los movimientos del usuario son detectados por una combinación de diferentes sensores, que se sitúan en el piso, en el aro y en los pilares. La información provista por estos sensores se envía a un microcontrolador integrado. Esta información es la orientación del usuario, la altura de su cintura y la velocidad y dirección de caminata. El software permite emular el teclado y la entrada de controladores, por lo que es posible su utilización en aplicaciones que no fueron pensadas para ser utilizadas por este dispositivo.

Por otro lado, existen varias técnicas para poder caminar o trasladarse en mundos virtuales que no requieren de complejas plataformas. La técnica más común se llama **walking in place**⁵ (WIP) y requiere que el usuario esté en una posición fija y mueva sus pies como si estuviese caminando, pero manteniéndose en el mismo lugar.

⁵Caminando en el Lugar

La Pressure Mat ([488]) es una plataforma que retorna la cantidad de presión aplicada en varios puntos fijos de la misma. Consiste en una matriz de resistencias sensibles a la presión conectadas a una computadora a través de una tarjeta de conversión analógica/digital. Un sistema de reconocimiento de patrones determina si la persona está detenida o caminando y la dirección en la que camina. Sin embargo, los patrones de presión fueron tan simétricos para determinar la dirección del usuario que se tuvo que adicionar un sensor extra a la cintura del mismo. Una de sus ventajas fundamentales es el bajo costo de su construcción.

Low-Latency continuous-motion Walking-in-Place ([489]) mide la posición y velocidad de los talones del usuario para luego utilizar un algoritmo que indica la velocidad de movimiento. El sistema presenta los problemas más comunes de los sistemas de WIP, como por ejemplo que los usuarios tienen problemas a la hora de comenzar y terminar de caminar. Esto genera inconvenientes cuando se experimenta con escenarios repletos de obstáculos que el usuario debe tratar de esquivar.

En el Walking-Pad ([490]), las acciones del usuario son realizadas sobre una plataforma que dispone una grilla de pulsadores que detectan la presión de los pies. A partir de la información provista por estos pulsadores, el sistema calcula el valor de diferentes variables que representan el comportamiento del usuario, como por ejemplo la dirección, la velocidad, o si está o no saltando. Una de las desventajas reportadas en el prototipo es el pequeño tamaño de la plataforma, lo que hace que los usuarios se salgan de ella.

6. DISCUSIÓN Y CONCLUSIONES

6.1. TIPOS DE *TRACKING*

En cuanto a los distintos tipos de *tracking*, los sistemas de *tracking* magnéticos, aunque muy precisos, solo funcionan adecuadamente en rangos cortos y sin la presencia de metales cerca. Sin embargo presenta una ventaja notoria ya que no requiere una línea de visión libre entre el transmisor y el receptor y el sistema inalámbrico suele ser más cómodo para los usuarios.

Aunque sea una de las técnicas más precisas, el *tracking* mecánico no es muy utilizado a la hora de trackear el cuerpo del usuario ya que involucra sistemas complejos e incómodos. Para trackear todo el cuerpo del usuario se necesitaría algo similar a un exoesqueleto mecánico. Sin embargo, el *tracking* mecánico es de utilidad cuando el sistema no se encuentra unido al usuario, como es el caso del sistema Phantom Omni ([340]) para la simulación de pequeñas herramientas.

Los sistemas de *tracking* óptico son de los primeros sistemas de *tracking* y se siguen utilizando hoy en día. Esto es gracias a que los avances tecnológicos han facilitado mucho la implementación de esta técnica. Al utilizar cámaras para el seguimiento del usuario, éste se encuentra completamente desconectado de la computadora. Sin embargo, la mayor limitación es que la línea visual entre la cámara y el objeto a trackear debe estar siempre despejada. Como una sola cámara solo nos provee información en 2D, se debe usar más de una para poder estimar la posición 3D mediante triangulaciones. De esta forma se logra una buena estimación de la posición del usuario, cosa que es muy difícil de lograr con otros sistemas de *tracking*. Siempre se debe tener en cuenta que al agregar más cámaras se estará incrementando la complejidad del sistema y del algoritmo de *tracking*.

A diferencia del *tracking* óptico, en el *tracking* Videométrico las cámaras son las que están en movimiento. Hoy en día, estas cámaras son pequeñas, baratas, livianas, de bajo consumo y versátiles, por lo que pueden ser montadas sobre casi cualquier objeto. Dado que en el *tracking* Videométrico la cámara está unida al objeto que será trackeado, esto agrega la complejidad de tener que poner marcadores especiales en la habitación, en posiciones conocidas. Además, las condiciones ambientales también pueden interferir en la detección de la imagen, reduciendo así la precisión del sistema. Sin embargo, aunque el análisis de la imagen en tiempo real suele ser un recurso muy costoso, los avances tecnológicos actuales lo convierten en algo completamente viable, incluso sobre dispositivos móviles.

El *tracking* ultrasónico suele ser barato de implementar. Sin embargo, presenta serias desventajas. Suelen trabajar en cortos rangos de distancia y poseen una baja frecuencia de muestreo. Además, la precisión disminuye si hay superficies acústicamente reflectantes en la sala o cualquier otro obstáculo. Los sonidos externos del ambiente también pueden interferir con la señal. Por este motivo, no suelen utilizarse para el *tracking* del usuario en sistemas de RV.

Los sistemas de *tracking* inercial utilizan sensores inerciales, como acelerómetros y giroscopios. Debido a que los acelerómetros proveen medidas relativas, estos presentan errores de acumulación, ruido y drift, a lo largo del tiempo, generando información imprecisa. Por esta razón, la mayoría de los sistemas de *tracking* puramente inerciales solo trackean la orientación. Aunque los giroscopios también sufran de errores de acumulación, esto es menos severo y existen métodos para solucionar este problema.

La mayoría de los HMDs comerciales utilizan sensores inerciales para determinar la orientación del dispositivo. Así se puede ajustar la orientación de la cámara virtual considerando hacia dónde se encuentre mirando el usuario. Estos sistemas suelen combinarse con otros sistemas de *tracking* para determinar las traslaciones en la posición del usuario, en caso de que la aplicación lo requiera.

Los sistemas de *tracking* neural son relativamente novedosos en el ámbito de la RV. Pueden trackear partes del cuerpo que ninguno de los otros sistemas puede hacer, como por ejemplo las contracciones de los músculos o distintas ondas cerebrales. La mayor desventaja de estos sistemas es que requieren el uso de varios electrodos colocados en la zona a trackear, lo que vuelve incómodo al sistema. Sin embargo, en un futuro cercano, esta tecnología promete nuevas formas de interactuar con los sistemas de manera más natural e intuitiva.

6.2. TRACKING DEL CUERPO

Como se mencionó anteriormente, es fundamental que un sistema de RV pueda trackear al menos la orientación de la cabeza del usuario para que éste esté físicamente inmerso. Si además de esto se puede trackear la posición y traslación del usuario, esto mejoraría aún más la inmersión. No es muy común realizar el *tracking* de otras partes del cuerpo que no sean la cabeza o las manos, pero hemos visto distintos estudios en los que se detalla la manera de hacerlo. En el caso ideal, el usuario podría ver todo su cuerpo en el mundo virtual y además, ver cómo este cuerpo virtual se mueve acorde con su cuerpo real. Sistemas como el Microsoft Kinect ([407]) podrían ayudar a realizar esto, pero se debe tener en cuenta la complejidad que esto agregaría al sistema. Es la aplicación particular la que determinará el *trade-off* entre precisión y complejidad.

Los gestos de las manos son una de las formas más efectivas para realizar interacciones entre los humanos y las computadoras. Entre los sistemas de *tracking* de manos más populares se encuentran los que tienen forma de guantes. La ventaja de estos dispositivos es que la mayoría de los sensores pueden estar incluidos en los mismos guantes. La desventaja es que los usuarios deben ponerse y sacarse los guantes, y generalmente no son muy cómodos.

Actualmente existen dispositivos de *tracking* de manos que no necesitan colocar sensores sobre las mismas. El Leap Motion ([408]) es uno de los que, mediante el *tracking* óptico, los usuarios pueden utilizar libremente sin añadir nada a sus propias manos. El inconveniente principal de este dispositivo es que el usuario pierde la sensación de estar tocando los objetos virtuales.

Por este motivo, algunos de los nuevos controles ofrecen técnicas intermedias. Por ejemplo, los controles Touch de Oculus ([380]) son controles físicos que el usuario toma con sus manos, pero están hechos de una forma especialmente ergonómica que logra que el usuario se olvide rápidamente que los está sosteniendo. Además, al proveer una realimentación háptica, el usuario siente que está tocando o sosteniendo los objetos virtuales, lo que incrementa considerablemente la inmersión.

El *tracking* de ojos presenta muchas ventajas al ser implementado en sistemas de RV. Una de las mayores ventajas es que ayudaría a optimizar el renderizado de las escenas virtuales ya que se podría renderizar con lujo de detalles solamente la región a la que el usuario está mirando y disminuir el detalle en las otras regiones de la pantalla. El inconveniente principal surge en la incorporación de sistemas de *tracking* de ojos en los HMDs existentes.

Por otro lado, la RV utilizando dispositivos móviles está creciendo rápidamente. La habilidad de trackear los ojos en estos dispositivos sería una gran oportunidad como técnica de interacción y como estudio del comportamiento visual humano.

En general, los sistemas de reconocimiento de voz juegan un rol importante en lograr una experiencia de RV más inmersiva y natural. Sin embargo, hasta que los sistemas no sean capaces de reconocer la voz perfectamente, de forma continua y natural, la elección del sistema a utilizar dependerá de la aplicación particular. Además, sin importar que tan bueno sea el sistema de reconocimiento, existirán tareas en las que la entrada por voz no será apropiada.

Por otro lado, la mayoría de los estudios solo utilizan comandos de voz para activar eventos particulares. Para lograr una sensación más inmersiva, es necesario que el reconocimiento de voz se combine con técnicas de inteligencia artificial para que el sistema actúe según lo que el usuario intente decir y no solo para activar eventos predeterminados.

6.3. DISPOSITIVOS

Existe una diversa variedad de dispositivos creados para lograr una comunicación más natural entre el usuario y la computadora. Muchos de estos dispositivos solo están pensados para ser utilizados con una PC de escritorio, mientras que muchos otros fueron creados exclusivamente para ser utilizados en RV.

Sin mencionar la técnica de *tracking* que se utilice, los controles son aquellos dispositivos con los que el usuario puede interactuar con el sistema. Generalmente el usuario interactuará mediante sus manos, pero éste no es siempre el caso. Hemos mencionado dispositivos para que el usuario interactúe con sus pies, con su voz, con sus ojos, con gestos o con cualquier parte del cuerpo.

Desde los orígenes de la RV se ha buscado que los usuarios puedan desplazarse o caminar por el mundo virtual. Esto dio origen a las plataformas especializadas para RV. Diversas estrategias han sido utilizadas para tratar de lograr una plataforma cómoda que permita a los usuarios sentir que están caminando en el mundo virtual. Sin embargo, incluso hoy en día, las plataformas existentes siguen resultando incómodas y poco naturales para la mayoría de los usuarios.

6.4. CONCLUSIÓN

En cuanto a los diferentes tipos de *tracking*, los sistemas magnéticos y ultrasónicos han demostrado depender tanto de las características del ambiente que son muy poco prácticos a la hora de lograr una localización suficientemente precisa. El *tracking* óptico ha sido uno de los primeros sistemas de *tracking* y sigue utilizándose incluso en los sistemas de RV comerciales más utilizados, como es el Oculus Rift ([16]) o el HTC Vive ([17]). Esto se debe a la alta precisión y comodidad que proveen al usuario.

La utilización de dispositivos móviles para RV está en constante crecimiento. Los sistemas de *tracking* ópticos que utilizan cámaras fijas dejan de ser prácticos en sistemas de RV móviles donde los usuarios se mueven por grandes espacios. Gracias a los avances tecnológicos, las cámaras son más pequeñas, más portátiles y más eficientes que nunca antes. Por este motivo, el *tracking* Videométrico se está haciendo más popular.

Es importante mencionar que los sistemas de *tracking* inerciales, aunque imprecisos a la hora de trackear la posición, están siendo utilizados en la mayoría de los sistemas de RV actuales para determinar la orientación de la cabeza del usuario, o de otras partes del cuerpo. Los dispositivos móviles disponen de estos sensores, facilitando su utilización para aplicaciones de RV.

En cuanto al *tracking* del cuerpo del usuario, existen muchos dispositivos y técnicas para trackear casi cualquier parte del cuerpo. El problema fundamental radica en que cuanto más dispositivos se le agregan al usuario, mayor será el nivel de incomodidad y menor será la inmersión. Entonces, esto dependerá de la aplicación particular que se desee utilizar.

Sin duda será fundamental trackear al menos la orientación de la cabeza del usuario para poder actualizar la cámara virtual. De lo contrario, el usuario sentirá que está viendo una película en lugar de sentirse inmerso en el mundo virtual. Realizar el *tracking* de la posición de la cabeza mejorará aún más la inmersión ya que permitirá que el usuario pueda desplazarse y vea los cambios reflejados en el mundo virtual.

Una de las primeras cosas que hace cualquier usuario cuando utiliza un sistema de RV es querer ver su propio cuerpo, o al menos sus manos, frente a su rostro. Esta es una de las razones por las que el siguiente paso en los sistemas de RV es poder hacer el *tracking* de las manos del usuario. Como vimos, existen actualmente muchas técnicas y dispositivos para realizar esto. Nuevamente, algunos serán más cómodos o precisos que otros, y la elección dependerá de la aplicación particular. Incluso hoy en día, no existe un dispositivo que pueda trackear las manos de forma precisa y natural exclusivamente para dispositivos móviles.

Como se mencionó en el Capítulo 2, la comunicación entre el usuario y la computadora dependerá exclusivamente de las interfaces que los comuniquen. Todos los sistemas e interfaces que se encuentren en el medio agregarán ruido y el usuario percibirá ese ruido como algo que no es natural. El objetivo fundamental de todas los dispositivos e interfaces vistos es disminuir este ruido lo más posible para así lograr realizar un *tracking* del usuario de la forma más cómoda y natural posible.

8

Conclusiones

La RV es una tecnología que busca sumergir a los usuarios en mundos virtuales de tal forma de hacerlos sentir como si estuviesen en el mundo real. A lo largo de esta tesis hemos visto cómo el ser humano percibe la realidad a través de cada uno de sus sentidos y cómo éstos influyen en el diseño y el desarrollo de nuevas interfaces y dispositivos para poder conectar al usuario con el mundo virtual.

En los Capítulos anteriores se han discutido los aspectos relevantes, ventajas y desventajas, de aquellos dispositivos asociados a cada uno de los sentidos en particular; se ha presentado un análisis sobre el funcionamiento de cada sentido, generando una mejor comprensión sobre cómo el ser humano percibe la realidad y la RV a través de los mismos; y se ha realizado un relevamiento exhaustivo del estado del arte sobre los sistemas y técnicas de RV existentes, puntualizando las ventajas y desventajas de cada uno.

En este Capítulo se presenta una síntesis de los conceptos más relevantes extraídos de los Capítulos anteriores, organizando esta información en 4 categorías que destacan los aspectos esenciales a considerar en todo sistema de RV. Estas categorías incluyen los conceptos sobre hardware, logística, presencia e inmersión, y sistemas multi-modales.

Además, durante el desarrollo del presente Magíster se han publicado diversos trabajos de investigación. Las contribuciones correspondientes a cada una de ellas se desarrollan en la Sección 5. Finalmente, el Capítulo concluye con el trabajo que está planificado a ser realizado en el futuro, que constituye con el plan de trabajo del Doctorado en Ciencias e Ingeniería de la Computación.

1. CARACTERÍSTICAS DE HARDWARE

Distintos tipos de displays visuales fueron presentados en el Capítulo 3. Los más populares actualmente son los HMDs debido al grado de inmersión completo que producen. Esto se debe en parte a los avances tecnológicos que han permitido el correcto funcionamiento de los mismos, cosa que no era así hasta hace pocos años. Además, el costo ha disminuido considerablemente, siendo más asequibles que nunca antes. A esto debe agregarse que, como los dispositivos móviles pueden “convertirse” en HMDs, un gran número de personas tiene la posibilidad de utilizar un HMD, en comparación a otros tipos de display.

Se debe tener en cuenta que un HMD solo funciona para una persona a la vez, por lo que no es adecuado en todo momento, particularmente en aquellas ocasiones en las que existen varios espectadores al mismo tiempo. La comodidad de los HMDs también ha mejorado; sin embargo, muchos usuarios suelen sentirlos incómodos luego de un tiempo de uso, cosa que no ocurre con los otros tipos de display, como displays de escritorio o de proyección.

Con respecto a los sistemas auditivos, debe resaltarse que la mayoría de las personas disponen del hardware necesario para obtener una buena estimulación auditiva. Un par de simples auriculares bastarían para lograr una experiencia de audio espacial bastante realista. La gran ventaja de estos dispositivos es su bajo costo. Además, existen varios tipos de auriculares, de distintas formas y tamaños, por lo que la posibilidad de encontrar unos que resulten cómodos es muy alta.

Para aquellas situaciones donde no se puedan utilizar auriculares, siempre se pueden utilizar altavoces. En este caso, como se vio en el Capítulo 4, se deben tener en cuenta los cambios logísticos que hay que realizar. La gran ventaja de los auriculares en comparación a los altavoces es que aíslan al usuario del mundo exterior, incrementando así el nivel de inmersión.

Los sentidos del gusto y el olfato son los menos estudiados, especialmente en sistemas de RV. Esto se debe principalmente a que aún no se ha encontrado una forma de reproducir distintos aromas de manera práctica, económica, y principalmente, no intrusiva. Esto es aún peor en los sistemas que intentan simular el gusto. Por esta razón, existen muy pocos sistemas de este estilo que se vendan comercialmente y su potencial no ha sido demostrado formalmente. Por el otro lado, aquellos sistemas que utilizan atomizadores para presentar un aroma en todo un ambiente, solamente son útiles para representar un aroma ambiental en la RV.

Como la piel es el órgano más grande del cuerpo, se requiere de grandes dispositivos para poder estimularla. Afortunadamente, los estudios sugieren que no se debe estimular todo el cuerpo, basta con estimular solo algunas zonas estratégicas de las que el cerebro esté consciente. La mayoría de los dispositivos hápticos son complejos y costosos. Más importante, como estos dispositivos deben estar en contacto con la piel, la mayoría de ellos son incómodos para el usuario. Por estas razones, no existen muchos dispositivos de este estilo en el mercado.

El tracking del usuario presenta la mayor variedad de dispositivos y tecnologías. Los dispositivos comerciales más populares actualmente, como el Oculus Rift ([16]) o el HTC Vive ([17]), utilizan sensores inerciales en el HMD para trackear la inclinación de la cabeza del usuario. Esto se debe a que dichos sensores son económicos, pequeños, y precisos. Es gracias a estos sensores que podemos utilizar los dispositivos móviles como HMDs. Por el otro lado, el Oculus Rift y el HTC Vive utilizan sensores ópticos, compuestos de cámaras y LEDs infrarrojos, para poder trackear la posición del HMD y de controles especiales. Aunque este tipo de sensores posee sus desventajas, como oclusión o generación de calor, funcionan muy bien y son extremadamente precisos en ambientes acotados.

2. LOGÍSTICA

Actualmente se utilizan muchos dispositivos que no son los más precisos, o los más económicos, o los más cómodos. Sin embargo, muchos de ellos siguen siendo muy utilizados.

Uno de los elementos más importantes a la hora de elegir un dispositivo para RV es la aplicación particular para la que éste se va a utilizar. Aunque existen dispositivos comerciales que son económicos, cómodos y eficientes, éstos suelen ser buenos para aplicaciones o sistemas “generales”. Sin embargo, existen sistemas particulares que requieren de dispositivos especiales.

Debe considerarse también que no se utilizan los mismos dispositivos en eventos al aire libre que en eventos bajo techo, en un sistemas dedicados a un único usuario que en sistemas multi-usuario, cuando el usuario debe estar sentado que cuando debe estar parado, etc. Todas estas consideraciones generalmente son subestimadas pero deben tenerse en cuenta.

Para poder elegir los dispositivos más adecuados, es importante primero considerar factores como la aplicación particular, el sistema, el ambiente, o la cantidad de usuarios. Incluso todo esto debe ser considerado al momento de diseñar y desarrollar un nuevo dispositivo.

3. PRESENCIA E INMERSIÓN

Aunque la presencia y la inmersión son factores cuyo estudio escapa de los objetivos de esta tesis, hemos visto que son factores que deben ser considerados a la hora de desarrollar, diseñar, o elegir una interfaz de usuario. Los sistemas de RV buscan maximizar la presencia y la inmersión y hemos visto algunas técnicas o factores que las impactan negativamente.

Con respecto a la visión, existen muchas señales visuales que pueden cambiar radicalmente la forma en que el usuario percibe el mundo virtual. Como se vio en el Capítulo 3, no siempre es necesario implementar cada una de estas señales visuales. Es muy importante considerar el tipo de aplicación que se está desarrollando y solo implementar las señales relevantes, ahorrando así poder computacional.

Uno de los efectos más estudiados es el *simulator sickness* y se encuentra más asociado a los displays visuales. Este efecto se produce generalmente cuando los sentidos del usuario experimentan situaciones conflictivas. El caso más común es cuando el usuario se encuentra físicamente sentado en una silla, pero a través del display visual está viendo que se encuentra caminando. El cerebro no sabe cómo interpretar este conflicto entre la percepción visual y física, por lo que genera sensaciones de mareo y malestar. El Capítulo 3 presenta diversos problemas que generan este efecto y diversas formas de solucionarlo.

El audio es uno de los sentidos fundamentales que nos permite incrementar considerablemente el nivel de inmersión, de forma relativamente fácil. Aunque las técnicas más simples permiten simular el sonido ambiente de forma precisa, la representación espacial del sonido hace que el usuario se sienta verdaderamente inmerso. Evidentemente, esto debe estar acompañado de un buen sistema de tracking ya que, por ejemplo, si el usuario tiene una fuente de sonido virtual en frente suyo, al girar la cabeza hacia la izquierda deberá percibir la fuente de sonido a su derecha.

Hemos visto que es posible generar señales estéreo que contengan la mayoría de las señales espaciales del mundo real. Utilizando técnicas de audio especializadas, una variedad de fuentes de sonido puede ser representada simultáneamente en diferentes direcciones y distancias.

Otros de los factores más comunes que reducen directamente la inmersión es la intrusión, y se presenta en cada uno de los sentidos. Los dispositivos intrusivos generan incomodidad en el usuario e inmediatamente lo “desconecta” del mundo virtual. Algunos HMDs muy pesados, displays olfativos incómodos, o grandes dispositivos hápticos mecánicos, suelen ser interfaces tan intrusivas e incómodas que rompen completamente la ilusión de pertenecer a un mundo virtual.

Por esto, a la hora de diseñar una nueva interfaz, no importa si se trata del HMD con la mejor resolución de todos, o un exoesqueleto sumamente flexible. Si estos dispositivos son incómodos para el usuario, no serán prácticos ni eficientes a la hora de utilizarse en sistemas de RV.

4. MULTI-MODALIDAD

Como se vio a lo largo de toda esta tesis, la experimentación aislada de cada uno de los sentidos del cuerpo humano es beneficiosa para testear algunos dispositivos en condiciones aisladas o particulares. Sin embargo, los sentidos funcionan mucho mejor en conjunto. Esto quiere decir que por ejemplo, el rendimiento de una interfaz háptica o auditiva se ve muy influenciado por las señales visuales que el usuario esté percibiendo. Esto no es algo menor ya que nos brinda estrategias para el diseño y desarrollo de nuevas interfaces y dispositivos más simples y económicos que puedan aprovechar la percepción multi-modal del cuerpo humano.

5. PUBLICACIONES

Durante el transcurso de la presente tesis de Magíster, se han publicado los siguientes trabajos:

- “AnArU, a virtual reality framework for physical human interactions” ([491]). El trabajo presenta un Framework que interconecta el sistema de desarrollo Unity3D, el sistema operativo de dispositivos móviles Android, y el microcontrolador Arduino, para sistemas de RV. Además, presenta el desarrollo de una plataforma en donde el usuario pueda caminar dentro de un mundo virtual.
- “Interacción humano computadora mediante interfaces hápticas en RV” ([492]). El trabajo presenta nuestra línea de investigación, que busca diseñar y desarrollar interfaces hápticas usables en el contexto de RV y obtener una tecnología que reduzca los costos involucrados, logrando una mayor inserción de las interfaces hápticas.
- “Braille Messages in a Haptic Wearable Device for Visually Impaired People” ([493]). El trabajo presenta el desarrollo de un dispositivo háptico que permite a los usuarios no videntes recibir mensajes de texto de forma háptica en su muñeca, en formato braille.
- “Real-Time Estimation of Illumination Direction for Augmented Reality with Low-Cost Sensors” ([494]). El trabajo presenta el diseño y desarrollo de un dispositivo que permite sensor la orientación de la luz incidente para así poder proyectarla dentro de una sistema de realidad aumentada de forma correcta.
- “Modelos de interacción y aplicaciones en realidad virtual mediante dispositivos móviles” ([495]). Este trabajo presenta nuestra línea de investigación, que estudia el uso de dispositivos móviles como núcleo computacional y displays de RV, analizando el impacto de los dispositivos móviles en los modelos de interacción de RV, y desarrollando nuevas aplicaciones de esta tecnología.

- “Estimating Illumination Direction for Augmented Reality in Real-Time by using Low-Cost Sensors” ([496]). El trabajo presenta la continuación del trabajo desarrollado en [494], agregando la complementación del dispositivo con aplicaciones móviles de realidad aumentada, en donde se presenta la sombra virtual del objeto aumentado según la orientación de la luz sensada.

Los siguientes trabajos se encuentran pendientes de aprobación:

- “Analysis of the relation between Presence and Learning for Low-Cost Virtual Reality”. El trabajo realiza un análisis sobre la relación existente entre la presencia y el aprendizaje adquirido en diversos entornos de realidad virtual. Los resultados sugieren que los sistemas de realidad virtual de bajo costo, aunque son menos inmersivos que los sistemas de punta, generan un grado de aprendizaje similar.

6. TRABAJO A FUTURO

Hemos visto que la presencia y la inmersión son factores fundamentales a tener en cuenta a la hora de diseñar una experiencia de RV. La diferencia fundamental entre la presencia y la inmersión es que la primera depende principalmente de factores psicológicos del usuario. Por ejemplo, en un simulador de Tenis virtual, no sentirá el mismo nivel de presencia aquel usuario que nunca ha jugado al tenis, comparado con aquel usuario que juega al tenis una vez por semana. La inmersión, por el otro lado, depende principalmente de factores objetivos de los dispositivos utilizados.

La tecnología ha avanzado mucho en las últimas décadas y hemos visto la gran cantidad de técnicas, sistemas, y dispositivos que existen para conectar a los usuarios con el mundo virtual. Sin embargo, no existe una métrica para poder cuantificar el grado de inmersión de un sistema de RV dado. Es decir, dados por ejemplo, un HMD particular, un dispositivo háptico particular, y una determinada técnica de tracking de alguna parte del cuerpo particular, no existe una métrica que nos brinde información del grado de inmersión que dicho conjunto de dispositivos le provee al usuario.

La investigación y desarrollo de dicha métrica será el objetivo del Doctorado en Ciencias de la Computación que será realizado al finalizar con el presente Magíster. Dicha métrica, además de cuantificar la inmersión de los dispositivos existentes, servirá como guía de diseño a la hora de desarrollar nuevos sistemas inmersivos para RV.

Bibliografía

- [1] K Brad Wray. *Knowledge and Inquiry: Readings in Epistemology*. Broadview Press, May 2002.
- [2] Alison Reiheld and Rory Kraft. Brain in the vat. *Questions: Philosophy for Young People*, 8:4–4, July 2008.
- [3] Michael Huemer. Direct realism and the Brain-in-a-Vat argument. *Philos. Phenomenol. Res.*, 61(2):397–413, 2000.
- [4] Mordekhai Velger. *Helmet-mounted displays and sights*. Norwood, MA: Artech House Publishers, 1998., 1998.
- [5] Jan-Philipp Stauffert, Florian Niebling, and Marc Erich Latoschik. Reducing application-stage latencies of interprocess communication techniques for real-time interactive systems. In *Virtual Reality (VR), 2016 IEEE*, pages 287–288. IEEE, 2016.
- [6] Kielan Yarrow, Ingvild Sverdrup-Stueland, Warrick Roseboom, and Derek H Arnold. Sensorimotor temporal recalibration within and across limbs. *Journal of Experimental Psychology: Human Perception and Performance*, 39(6):1678, 2013.
- [7] J Guo, D Weng, H Been-Lirn Duh, Y Liu, and Y Wang. Effects of using HMDs on visual fatigue in virtual environments. In *2017 IEEE Virtual Reality (VR)*, pages 249–250, March 2017.
- [8] Lisa Rebenitsch and Charles Owen. Review on cybersickness in applications and visual displays. *Virtual Real.*, 20(2):101–125, 1 June 2016.
- [9] Martin Mauve, Jürgen Vogel, Volker Hilt, and Wolfgang Effelsberg. Local-lag and timewarp: providing consistency for replicated continuous applications. *IEEE transactions on Multimedia*, 6(1):47–57, 2004.

- [10] Sophie Jörg, Aline Normoyle, and Alla Safonova. How responsiveness affects players' perception in digital games. In *Proceedings of the ACM Symposium on Applied Perception*, pages 33–38. ACM, 2012.
- [11] Kjetil Raaen, Ragnhild Eg, and Carsten Griwodz. Can gamers detect cloud delay? In *Proceedings of the 13th Annual Workshop on Network and Systems Support for Games*, page 6. IEEE Press, 2014.
- [12] Katerina Mania, Bernard D Adelstein, Stephen R Ellis, and Michael I Hill. Perceptual sensitivity to head tracking latency in virtual environments with varying degrees of scene complexity. In *Proceedings of the 1st Symposium on Applied perception in graphics and visualization*, pages 39–47. ACM, 2004.
- [13] Richard Held and Nathaniel Durlach. Telepresence, time delay and adaptation. *Pictorial communication in virtual and real environments*, pages 232–246, 1991.
- [14] Stephen R Ellis, Katerina Mania, Bernard D Adelstein, and Michael I Hill. Generalizeability of latency detection in a variety of virtual environments. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 48, pages 2632–2636. SAGE Publications Sage CA: Los Angeles, CA, 2004.
- [15] Carolina Cruz-Neira, Daniel J Sandin, Thomas A DeFanti, Robert V Kenyon, and John C Hart. The cave: audio visual experience automatic virtual environment. *Communications of the ACM*, 35(6):64–73, 1992.
- [16] Oculus rift. <https://www.oculus.com/rift/>. accedido en abril 2017.
- [17] Htc vive. <https://www.vive.com/us/>. accedido en abril 2017.
- [18] Google glass. <https://developers.google.com/glass/>. accedido en abril 2017.
- [19] Microsoft hololens. <https://www.microsoft.com/en-us/hololens/developers>. accedido en abril 2017.
- [20] H Okumura, T Sasaki, A Hotta, and K Shinohara. Monocular hyperrealistic virtual and augmented reality display. In *2014 IEEE Fourth International Conference on Consumer Electronics Berlin (ICCE-Berlin)*, pages 19–23, 2014.
- [21] E Viirre, H Pryor, S Nagata, and T A Furness, 3rd. The virtual retinal display: a new technology for virtual reality and augmented vision in medicine. *Stud. Health Technol. Inform.*, 50:252–257, 1998.

- [22] Avegant. <https://www.avegant.com/video-headset>. accedido en abril 2017.
- [23] C Krumbholz, J Leigh, A Johnson, and others. Lambda table: high resolution tiled display table for interacting with large visualizations. *Proceedings of 5th*, 2005.
- [24] Timothy A Sandstrom, Chris Henze, and Creon Levit. The hyperwall. In *Proceedings of the conference on Coordinated and Multiple Views In Exploratory Visualization*, page 124. IEEE Computer Society, 15 July 2003.
- [25] Bo-Wen Shen, Wei-Kuo Tao, and Bryan Green. Coupling advanced modeling and visualization to improve High-Impact tropical weather prediction. *Computing in Science and Engineering*, 13(5):56–67, 1 September 2011.
- [26] Nasa. 2014. visualization system: Hyperwall-2. http://www.nas.nasa.gov/hecc/resources/viz_systems.html. accedido en abril 2017.
- [27] Maurice Ten Koppel, Gilles Bailly, Jörg Müller, and Robert Walter. Chained displays: configurations of public displays can be used to influence actor-, audience-, and passer-by behavior. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 317–326. ACM, 5 May 2012.
- [28] Multitaction. 2014. ultra thin bezel. <http://www.multitaction.com/>. accedido en abril 2017.
- [29] Antonietta Grasso, Alain Karsenty, and Dave Snowdon. A bench for all moods. In *CHI '00 Extended Abstracts on Human Factors in Computing Systems*, pages 197–198. ACM, 1 April 2000.
- [30] Peter Peltonen, Esko Kurvinen, Antti Salovaara, Giulio Jacucci, Tommi Ilmonen, John Evans, Antti Oulasvirta, and Petri Saarikko. It's mine, don't touch!: interactions at a large multi-touch display in a city centre. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1285–1294. ACM, 6 April 2008.
- [31] Simo Hosio, Marko Jurmu, Hannu Kukka, Jukka Riekkö, and Timo Ojala. Supporting distributed private and public user interfaces in urban environments. In *Proceedings of the Eleventh Workshop on Mobile Computing Systems & Applications*, pages 25–30. ACM, 22 February 2010.

- [32] Timo Ojala, Hannu Kukka, Tomas Lindén, Tommi Heikkinen, Marko Jurmu, Simo Hosio, and Fabio Kruger. UBI-Hotspot 1.0: Large-Scale Long-Term deployment of interactive public displays in a city center. In *Proceedings of the 2010 Fifth International Conference on Internet and Web Applications and Services*, pages 285–294. IEEE Computer Society, 9 May 2010.
- [33] Ville Valkama and Timo Ojala. Stakeholder value propositions on open community testbed of interactive public displays. In *Proceedings of the International Conference on Management of Emergent Digital EcoSystems*, pages 107–113. ACM, 21 November 2011.
- [34] Tom Peterka, Daniel J Sandin, Jinghua Ge, Javier Girado, Robert Kooima, Jason Leigh, Andrew Johnson, Marcus Thiebaut, and Thomas A DeFanti. Personal varrier: autostereoscopic virtual reality display for distributed scientific visualization. *Future Generation Computer Systems*, 22(8):976–983, 2006.
- [35] Daniel J Sandin, Todd Margolis, Jinghua Ge, Javier Girado, Tom Peterka, and Thomas A DeFanti. The varrier tm autostereoscopic virtual reality display. In *ACM Transactions on Graphics (TOG)*, volume 24, pages 894–903. ACM, 2005.
- [36] Juan-J Arino, M-Carmen Juan, Jose-Antonio Gil-Gómez, and Ramón Mollá. A comparative study using an autostereoscopic display with augmented and virtual reality. *Behaviour & Information Technology*, 33(6):646–655, 2014.
- [37] Wolfgang Krueger and Bernd Froehlich. Responsive workbench. In *Virtual Reality'94*, pages 73–80. Springer, 1994.
- [38] Marek Czernuszenko, Dave Pape, Daniel Sandin, Tom DeFanti, Gregory L Dawe, and Maxine D Brown. The immersadesk and infinity wall projection-based virtual reality displays. *ACM SIGGRAPH Computer Graphics*, 31(2):46–49, 1997.
- [39] Andreas Simon and Martin Gobel. The i-cone/spl trade/-a panoramic display system for virtual environments. In *Computer Graphics and Applications, 2002. Proceedings. 10th Pacific Conference on*, pages 3–7. IEEE, 2002.
- [40] Dongsik Jo, Hyun Kang, Gun A Lee, and Wookho Son. Xphere: A pc cluster based hemispherical display system. In *Virtual Reality Conference, 2006*, pages 312–312. IEEE, 2006.

- [41] Tobias Höllerer, JoAnn Kuchera-Morin, and Xavier Amatriain. The allosphere: a large-scale immersive surround-view instrument. In *Proceedings of the 2007 workshop on Emerging displays technologies: images and beyond: the future of displays and interacton*, page 3. ACM, 2007.
- [42] C Canada, Tim Harrington, Robert Kares, Dave Modl, Laura Monroe, and Steve Stringer. La cueva grande: a 43-megapixel immersive system. In *Virtual Reality Conference, 2006*, pages 316–316. IEEE, 2006.
- [43] Thomas A DeFanti, Gregory Dawe, Daniel J Sandin, Jurgen P Schulze, Peter Otto, Javier Girado, Falko Kuester, Larry Smarr, and Ramesh Rao. The starcave, a third-generation cave and virtual reality optiportal. *Future Generation Computer Systems*, 25(2):169–178, 2009.
- [44] DM Traill, JD Bowshell, and PJ Lawrence. Interactive collaborative media environments. *BT Technology Journal*, 15(4):130–140, 1997.
- [45] Carolina Cruz-Neira, Daniel J Sandin, and Thomas A DeFanti. Surround-screen projection-based virtual reality: the design and implementation of the cave. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 135–142. ACM, 1993.
- [46] Chao Li, Hai Lin, and Jiaoying Shi. A survey of multi-projector tiled display wall construction. In *Image and Graphics (ICIG'04), Third International Conference on*, pages 452–455. IEEE, 2004.
- [47] Ed Lantz. Future directions in visual display systems. *ACM SIGGRAPH Computer Graphics*, 31(2):38–42, 1997.
- [48] Ed Lantz. Large-scale immersive displays in entertainment and education. In *2nd Annual Immersive Projection Technology Workshop*, pages 11–12, 1998.
- [49] Robert A Fisher and HM Tong. A full-field-of-view dome visual display for tactical combat training. In *Proc. Image Conference IV*, pages 23–26, 1987.
- [50] E Lantz, M Hutton, S Savage, and C Ward. Graphics design and production for hemispheric projection. In *ACM SIGGRAPH*, 1995.
- [51] Ed Lantz and US Route. The digital planetarium. In *Proc. of 2002 International Planetarium Society Conference*. Citeseer, 2002.
- [52] Kevin Scott and Richard McColman. Report of the ips technical committee: Full-dome video systems. *The Planetarian*, 28(1):25–33, 1999.

- [53] Martin Ratcliffe. The planetarium revolution. *Astronomy*, 29(10):78–83, 2001.
- [54] Ryan Wyatt. Planetarium paradigm shift. *Planetarian*, 2005.
- [55] Ka Chun Yu. Digital full-domes: The future of virtual astronomy education. *Planetarian*, 2005.
- [56] Don Davis. Some thoughts from an artist on fulldome theaters. *Journal of the International Planetarium Society*, 34(3):20–22, 2005.
- [57] C Emmart. The powers of ten with a steering wheel on the global visualization superhighway. *The Planetarian*, 34:19–26, 2005.
- [58] Ka Chun Yu, Kenji Williams, Daniel Neafus, Leslie Gaston, and Greg Downing. Gaia journeys: a museum-based immersive performance exploration of the earth. *International Journal of Digital Earth*, 2(1):44–58, 2009.
- [59] Phillippe Chiwy. Vr in immersive and interactive theatres: The bridge between movie and video-game experiences. In *9th TiLE Conference Proceedings*, pages 145–148, 2000.
- [60] Ed Lantz. A survey of large-scale immersive displays. In *Proceedings of the 2007 workshop on Emerging displays technologies: images and beyond: the future of displays and interacton*, page 1. ACM, 2007.
- [61] Daniel R Schikore, Richard A Fischer, Randall Frank, Ross Gaunt, John Hobson, and Brad Whitlock. High-Resolution multiprojector display walls. *IEEE Comput. Graph. Appl.*, 20(4):38–44, 1 July 2000.
- [62] Timo Ojala, Vassilis Kostakos, Hannu Kukka, Tommi Heikkinen, Tomas Linden, Marko Jurmu, Simo Hosio, Fabio Kruger, and Daniele Zanni. Multipurpose interactive public displays in the wild: Three years later. *Computer*, 45(5):42–49, 1 May 2012.
- [63] Akio Shinohara, Junji Tomita, Tamio Kihara, Shinya Nakajima, and Katsuhiko Ogawa. A huge screen interactive public media system: mirai-tube. In *Proceedings of the 12th international conference on Human-computer interaction: interaction platforms and techniques*, pages 936–945. Springer-Verlag, 1 July 2007.
- [64] An affordable Surround-Screen virtual reality display. 18(836-843), 1 October 2010.

- [65] Carolina Cruz-Neira, Dirk Reiners, and Jan P Springer. An affordable surround-screen virtual reality display. *Journal of the Society for Information Display*, 18(10):836–843, 2010.
- [66] Gaku Narita, Yoshihiro Watanabe, and Masatoshi Ishikawa. Dynamic projection mapping onto a deformable object with occlusion based on high-speed tracking of dot marker array. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, pages 149–152. ACM, 13 November 2015.
- [67] Johnny C Lee, Scott E Hudson, and Edward Tse. Foldable interactive displays. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 287–290. ACM, 19 October 2008.
- [68] Claudio Pinhanez. Using a steerable projector and a camera to transform surfaces into interactive displays. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, pages 369–370. ACM, 31 March 2001.
- [69] Johnny C Lee, Scott E Hudson, Jay W Summet, and Paul H Dietz. Moveable interactive projected displays using projector based tracking. In *Proceedings of the 18th annual ACM symposium on User interface software and technology*, pages 63–72. ACM, 23 October 2005.
- [70] John Bolton, Kibum Kim, and Roel Vertegaal. A comparison of competitive and cooperative task performance using spherical and flat displays. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, pages 529–538. ACM, 11 February 2012.
- [71] Gilbert Beyer, Florian Alt, Jörg Müller, Albrecht Schmidt, Karsten Isakovic, Stefan Klose, Manuel Schiewe, and Ivo Haulsen. Audience behavior around large interactive cylindrical screens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1021–1030. ACM, 7 May 2011.
- [72] Jin-Yao Lin, Yen-Yu Chen, Ju-Chun Ko, Huishan Kao, Wei-Han Chen, Tsun-Hung Tsai, Su-Chu Hsu, and Yi-Ping Hung. i-m-tube: an interactive multi-resolution tubular display. In *Proceedings of the 17th ACM international conference on Multimedia*, pages 253–260. ACM, 23 October 2009.
- [73] OrbeVR: a handheld concave spherical virtual reality display. In *ACM SIGGRAPH 2017 Emerging Technologies*, 30 July 2017.

- [74] Hrvoje Benko, Andrew D Wilson, and Ravin Balakrishnan. Sphere: multi-touch interactions on a spherical display. In *Proceedings of the 21st annual ACM symposium on User interface software and technology*, pages 77–86. ACM, 19 October 2008.
- [75] A Steed. Supporting multiple immersive configurations using a shape-changing display. In *2016 IEEE Virtual Reality (VR)*, pages 289–290, March 2016.
- [76] Ismo K Rakkolainen and Artur K Lugmayr. Immaterial display for interactive advertisements. In *Proceedings of the international conference on Advances in computer entertainment technology*, pages 95–98. ACM, 13 June 2007.
- [77] Measurements and experiments of the immaterial virtual reality display. In *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2008*, 28 June 2008.
- [78] Playstation vr. <https://www.playstation.com/en-au/explore/playstation-vr/>. accedido en abril 2017.
- [79] Gear vr. www.samsung.com/latin/promotions/galaxynote4/feature/gearvr/. accedido en abril 2017.
- [80] Google cardboard. <https://vr.google.com/cardboard/>. accedido en abril 2017.
- [81] Kaan Akşit, Jan Kautz, and David Luebke. Slim near-eye display using pinhole aperture arrays. *Appl. Opt.*, 54(11):3422–3427, 10 April 2015.
- [82] Dewen Cheng, Yongtian Wang, Hong Hua, and Jose Sasian. Design of a wide-angle, lightweight head-mounted display using free-form optics tiling. *Opt. Lett.*, 36(11):2098–2100, 1 June 2011.
- [83] Dewen Cheng and Yongtian Wang. Highly immersive head-mounted displays based on aspherical and freeform optics. In *2015 Conference on Lasers and Electro-Optics Pacific Rim*, page 26B3_1. Optical Society of America, 24 August 2015.
- [84] Yunqi Wang, Weiqi Liu, Xiangxiang Meng, Hanyi Fu, Daliang Zhang, Yusi Kang, Rui Feng, Zhonglun Wei, Xiuqing Zhu, and Guohua Jiang. Development of an immersive virtual reality head-mounted display with high performance. *Appl. Opt.*, 55(25):6969–6977, 1 September 2016.

- [85] Fangwang Gou, Haiwei Chen, Ming-Chun Li, Seok-Lyul Lee, and Shin-Tson Wu. Submillisecond-response liquid crystal for high-resolution virtual reality displays. *Opt. Express*, 25(7):7984–7997, 3 April 2017.
- [86] Ismo Rakkolainen, Matthew Turk, and Tobias Höllerer. A compact, wide-FOV optical design for head-mounted displays. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, pages 293–294. ACM, 2 November 2016.
- [87] A MacQuarrie and A Steed. Cinematic virtual reality: Evaluating the effect of display type on the viewing experience for panoramic video. In *2017 IEEE Virtual Reality (VR)*, pages 45–54, March 2017.
- [88] Fabio Buttussi and Luca Chittaro. Effects of different types of virtual reality display on presence and learning in a safety training scenario. *IEEE Trans. Vis. Comput. Graph.*, 16 January 2017.
- [89] Maxime Cordeil, Tim Dwyer, Karsten Klein, Bireswar Laha, Kim Marriott, and Bruce H Thomas. Immersive collaborative analysis of network connectivity: CAVE-style or Head-Mounted display? *IEEE Trans. Vis. Comput. Graph.*, 23(1):441–450, January 2017.
- [90] Marc Lambooij, Marten Fortuin, Ingrid Heynderickx, and Wijnand IJsselstein. Visual discomfort and visual fatigue of stereoscopic displays: A review. *Journal of Imaging Science and Technology*, 53(3):30201–1, 2009.
- [91] Takahiro Kawabe, Warrick Roseboom, and Shin’ya Nishida. The sense of agency is action–effect causality perception based on cross-modal grouping. In *Proc. R. Soc. B*, volume 280, page 20130991. The Royal Society, 2013.
- [92] DB Chertoff, SL Schatz, KS Hale, and KM Stanney. Beyond presence: how holistic experience drives training and education. *Handbook of Virtual Environments: Design, Implementation, and Applications*, pages 857–872, 2014.
- [93] Thomas B Sheridan. Further musings on the psychophysics of presence. *Presence: Teleoperators & Virtual Environments*, 5(2):241–246, 1996.
- [94] Claudia Hendrix and Woodrow Barfield. The sense of presence within auditory virtual environments. *Presence: Teleoperators & Virtual Environments*, 5(3):290–301, 1996.

- [95] Russell L Storms et al. Auditory-visual cross-modal perception. In *In ICAD*. Citeseer, 2000.
- [96] Barbara Shinn-Cunningham. Learning reverberation: Considerations for spatial auditory displays. 2000.
- [97] Barbara G Shinn-Cunningham. Object-based auditory and visual attention. *Trends in cognitive sciences*, 12(5):182–186, 2008.
- [98] William A Yost. *Fundamentals of hearing: An introduction*. Academic Press, 1994.
- [99] E Colin Cherry. Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, 25(5):975–979, 1953.
- [100] Durand R Begault. Virtual acoustic displays for teleconferencing: Intelligibility advantage for 'telephone-grade' audio. *Journal of the Audio Engineering Society*, 47(10):824–828, 1999.
- [101] Durand R Begault. Head-up auditory displays for traffic collision avoidance system advisories: A preliminary investigation. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 35(4):707–717, 1993.
- [102] Durand R Begault and Elizabeth M Wenzel. Techniques and applications for binaural sound manipulation. *The International Journal of Aviation Psychology*, 2(1):1–22, 1992.
- [103] Ellen C Haas, Charles Gainer, Dennis Wightman, Michael Couch, and Russell Shilling. Enhancing system safety with 3-d audio displays. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 41, pages 868–872. SAGE Publications, 1997.
- [104] Robert P Carlyon. How the brain separates sounds. *Trends in cognitive sciences*, 8(10):465–471, 2004.
- [105] Stephen A Brewster, Peter C Wright, and Alistair DN Edwards. An evaluation of earcons for use in auditory human-computer interfaces. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 222–227. ACM, 1993.
- [106] Song Hui Chon and Stephen McAdams. Investigation of timbre saliency, the attention-capturing quality of timbre. *The Journal of the Acoustical Society of America*, 131(4):3433–3433, 2012.

- [107] Ross K Maddox and Barbara G Shinn-Cunningham. Influence of task-relevant and task-irrelevant feature continuity on selective auditory attention. *Journal of the Association for Research in Otolaryngology*, 13(1):119–129, 2012.
- [108] David R Badcock, Stephen Palmisano, and James G May. Vision and virtual environments. 2015.
- [109] Nuray Dindar, A Murat Tekalp, and Cagatay Basdogan. Dynamic haptic interaction with video.
- [110] Ben D Lawson and Bernhard E Riecke. The perception of body motion. *Handbook of virtual environments: Design, implementation, and applications*, 2:163–196, 2014.
- [111] ROBERT B WELCH and DAVID H WARREN. Intersensory interactions. *Handbook of perception and human performance.*, 1:25–1, 1986.
- [112] Pavel Zahorik, Douglas S Brungart, and Adelbert W Bronkhorst. Auditory distance perception in humans: A summary of past and present research. *Acta Acustica united with Acustica*, 91(3):409–420, 2005.
- [113] Douglas S Brungart and William M Rabinowitz. Auditory localization of nearby sources. head-related transfer functions. *The Journal of the Acoustical Society of America*, 106(3):1465–1479, 1999.
- [114] Richard O Duda and William L Martens. Range dependence of the response of a spherical head model. *The Journal of the Acoustical Society of America*, 104(5):3048–3058, 1998.
- [115] Barbara G Shinn-Cunningham, Scott Santarelli, and Norbert Kopco. Tori of confusion: Binaural localization cues for sources within reach of a listener. *The Journal of the Acoustical Society of America*, 107(3):1627–1636, 2000.
- [116] V Ralph Algazi, Carlos Avendano, and Richard O Duda. Elevation localization and head-related transfer function analysis at low frequencies. *The Journal of the Acoustical Society of America*, 109(3):1110–1122, 2001.
- [117] Soren H Nielsen. Depth perception-finding a design goal for sound reproduction systems. In *Audio Engineering Society Convention 90*. Audio Engineering Society, 1991.
- [118] Adelbert W Bronkhorst and Tammo Houtgast. Auditory distance perception in rooms. *Nature*, 397(6719):517–520, 1999.

- [119] Donald H Mershon. Phenomenal geometry and the measurement of perceived auditory distance. *Binaural and spatial hearing in real and virtual environments*, pages 257–274, 1997.
- [120] Durand R Begault and Elizabeth M Wenzel. Headphone localization of speech. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 35(2):361–376, 1993.
- [121] William M Hartmann. Localization of sound in rooms. *The Journal of the Acoustical Society of America*, 74(5):1380–1391, 1983.
- [122] Donald H Mershon, William L Ballenger, Alex D Little, Patrick L McMurtry, and Judith L Buchanan. Effects of room reflectance and background noise on perceived auditory distance. *Perception*, 18(3):403–416, 1989.
- [123] BG Shinn-Cunningham, N Kopco, and S Santarelli. Computation of acoustic source position in near-field listening. In *3rd International Conference on Cognitive and Neural Systems, Boston, MA*, 1999.
- [124] Durand R Begault, Elizabeth M Wenzel, and Mark R Anderson. Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source. *Journal of the Audio Engineering Society*, 49(10):904–916, 2001.
- [125] Douglas S Brungart and William D'Angelo. Effects of reverberation cues on distance identification in virtual audio displays. *The Journal of the Acoustical Society of America*, 97(5):3279–3279, 1995.
- [126] Hans Wallach. The role of head movements and vestibular and visual cues in sound localization. *Journal of Experimental Psychology*, 27(4):339, 1940.
- [127] Frederic L Wightman and Doris J Kistler. Headphone simulation of free-field listening. i: Stimulus synthesis. *The Journal of the Acoustical Society of America*, 85(2):858–867, 1989.
- [128] D Wesley Grantham. Auditory motion perception: Snapshots revisited. *Binaural and spatial hearing in real and virtual environments*, pages 295–313, 1997.
- [129] Birger Kollmeier and Robert H Gilkey. Binaural forward and backward masking: evidence for sluggishness in binaural detection. *The Journal of the Acoustical Society of America*, 87(4):1709–1719, 1990.

- [130] Rachel K Clifton, Richard L Freyman, Ruth Y Litovsky, and Daniel McCall. Listeners' expectations about echoes can raise or lower echo threshold. *The Journal of the Acoustical Society of America*, 95(3):1525–1533, 1994.
- [131] Richard L Freyman, Rachel K Clifton, and Ruth Y Litovsky. Dynamic processes in the precedence effect. *The Journal of the Acoustical Society of America*, 90(2):874–884, 1991.
- [132] Paul M Hofman, Jos GA Van Riswick, and A John Van Opstal. Relearning sound localization with new ears. *Nature neuroscience*, 1(5):417–421, 1998.
- [133] Barbara G Shinn-Cunningham, Nathaniel I Durlach, and Richard M Held. Adapting to supernormal auditory localization cues. i. bias and resolution. *The Journal of the Acoustical Society of America*, 103(6):3656–3666, 1998.
- [134] Barbara Shinn-Cunningham. Adapting to remapped auditory localization cues: A decision-theory model. *Perception & psychophysics*, 62(1):33–47, 2000.
- [135] George V Popescu, Grigore C Burdea, and Helmuth Trefftz. *Multimodal interaction modeling*. Lawrence Erlbaum Associates, Mahwah, NJ, 2002.
- [136] Maarten A Frens, A John Van Opstal, and Robert F Van der Willigen. Spatial and temporal factors determine auditory-visual interactions in human saccadic eye movements. *Perception & Psychophysics*, 57(6):802–816, 1995.
- [137] David R Perrott, Kouros Saberi, Kathleen Brown, and Thomas Z Strybel. Auditory psychomotor coordination and visual search performance. *Perception & Psychophysics*, 48(3):214–226, 1990.
- [138] David R Perrott, Toktam Sadralodabai, Kouros Saberi, and Thomas Z Strybel. Aurally aided visual search in the central visual field: Effects of visual load and visual enhancement of the target. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 33(4):389–400, 1991.
- [139] Robyn Kim, Megan AK Peters, and Ladan Shams. $0+1>1$ how adding noninformative sound improves performance on a visual task. *Psychological science*, 23(1):6–12, 2012.
- [140] David H Warren, Robert B Welch, and Timothy J McCarthy. The role of visual-auditory “compellingness” in the ventriloquism effect: Implications for transitivity among the spatial senses. *Perception & Psychophysics*, 30(6):557–564, 1981.

- [141] Herbert L Pick, David H Warren, and John C Hay. Sensory conflict in judgments of spatial direction. *Perception & Psychophysics*, 6(4):203–205, 1969.
- [142] Robert B Welch and David H Warren. Immediate perceptual response to intersensory discrepancy. *Psychological bulletin*, 88(3):638, 1980.
- [143] Adelbert W Bronkhorst, JA Hans Veltman, and Leo Van Breda. Application of a three-dimensional auditory display in a flight task. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 38(1):23–33, 1996.
- [144] Richard D Shilling and Barbara Shinn-Cunningham. Virtual auditory displays. *Handbook of virtual environment technology*, pages 65–92, 2002.
- [145] Joshua M Knapp and Jack M Loomis. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators and Virtual Environments*, 13(5):572–577, 2004.
- [146] Albert A Rizzo, Todd Bowerly, J Galen Buckwalter, Dean Klimchuk, Roman Mitura, and Thomas D Parsons. A virtual reality scenario for all seasons: the virtual classroom. *Cns Spectrums*, 11(01):35–44, 2006.
- [147] R Iglesias, S Casado, T Gutierrez, JI Barbero, CA Avizzano, S Marcheschi, and M Bergamasco. Computer graphics access for blind people through a haptic and audio virtual environment. In *Haptic, Audio and Visual Environments and Their Applications, 2004. HAVE 2004. Proceedings. The 3rd IEEE International Workshop on*, pages 13–18. IEEE, 2004.
- [148] Jaime Sánchez, Mauricio Lumbreras, and Luca Cernuzzi. Interactive virtual acoustic environments for blind children: computing, usability, and cognition. In *CHI'01 Extended Abstracts on Human Factors in Computing Systems*, pages 65–66. ACM, 2001.
- [149] Lorenzo Picinali, Amandine Afonso, Michel Denis, and Brian FG Katz. Exploration of architectural spaces by blind people using auditory virtual reality for the construction of spatial knowledge. *International Journal of Human-Computer Studies*, 72(4):393–407, 2014.
- [150] Erin C Connors, Lindsay A Yazzolino, Jaime Sánchez, and Lotfi B Merabet. Development of an audio-based virtual gaming environment to assist with navigation skills in the blind. *JoVE (Journal of Visualized Experiments)*, (73):e50272–e50272, 2013.

- [151] Nathaniel I Durlach and H Steven Colburn. Binaural phenomena. *Handbook of perception*, 4:365–466, 1978.
- [152] Jens Blauert. *Spatial hearing: the psychophysics of human sound localization*. MIT press, 1997.
- [153] Pascal Dietrich, Bruno Masiero, and Michael Vorländer. On the optimization of the multiple exponential sweep method. *Journal of the Audio Engineering Society*, 61(3):113–124, 2013.
- [154] Tomlinson Holman. *Surround sound: up and running*. CRC Press, 2014.
- [155] Benjamin B Bauer. Phasor analysis of some stereophonic phenomena. *The Journal of the Acoustical Society of America*, 33(11):1536–1539, 1961.
- [156] Roger Zimmermann, Chris Kyriakakis, Cyrus Shahabi, Christos Papadopoulos, Alexander A Sawchuk, and Ulrich Neumann. The remote media immersion system. *IEEE MultiMedia*, 11(2):48–57, 2004.
- [157] Ville Pulkki. Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society*, 45(6):456–466, 1997.
- [158] Michael A Gerzon. Ambisonics in multichannel broadcasting and video. *Journal of the Audio Engineering Society*, 33(11):859–871, 1985.
- [159] Augustinus J Berkhout. A holographic approach to acoustic control. *Journal of the audio engineering society*, 36(12):977–995, 1988.
- [160] Ulrich Horbach, Etienne Corteel, Renato Pellegrini, and Edo Hulsebos. Real-time rendering of dynamic scenes using wave fields synthesis. In *ICME 2002*, 2002.
- [161] Ulrich Horbach and Marinus M Boone. Future transmission and rendering formats for multichannel sound. In *Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction*. Audio Engineering Society, 1999.
- [162] Rolf Nordahl, Luca Turchet, and Stefania Serafin. Sound synthesis and evaluation of interactive footsteps and environmental sounds rendering for virtual reality applications. *IEEE transactions on visualization and computer graphics*, 17(9):1234–1244, 2011.
- [163] Clara Suied, George Drettakis, Olivier Warusfel, and Isabelle Viaud-Delmon. Auditory-visual virtual reality as a diagnostic and therapeutic tool for cynophobia. *Cyberpsychology, Behavior, and Social Networking*, 16(2):145–152, 2013.

- [164] J Douglas Bremner, Lawrence H Staib, Danny Kaloupek, Steven M Southwick, Robert Soufer, and Dennis S Charney. Neural correlates of exposure to traumatic pictures and sound in vietnam combat veterans with and without posttraumatic stress disorder: a positron emission tomography study. *Biological psychiatry*, 45(7):806–816, 1999.
- [165] Charlotte Magnusson, Henrik Danielsson, and Kirsten Rasmus-Gröhn. Non visual haptic audio tools for virtual environments. In *International Workshop on Haptic and Audio Interaction Design*, pages 111–120. Springer, 2006.
- [166] Marine Taffou, Jan Ondřej, Carol O’Sullivan, Olivier Warusfel, Stéphanie Dubal, and Isabelle Viaud-Delmon. Auditory-visual virtual environment for the treatment of fear of crowds. In *Laval Virtual*, 2015.
- [167] Brent Cowan, David Rojas, Bill Kapralos, Karen Collins, and Adam Dubrowski. Spatial sound and its effect on visual quality perception and task performance within a virtual environment. In *Proceedings of Meetings on Acoustics*, volume 19, page 050126. Acoustical Society of America, 2013.
- [168] Nicolas Bonneel, Clara Suied, Isabelle Viaud-Delmon, and George Dretakis. Bimodal perception of audio-visual material properties for virtual environments. *ACM Transactions on Applied Perception (TAP)*, 7(1):1, 2010.
- [169] V Hulusic, Matt Aranha, and Alan Chalmers. The influence of cross-modal interaction on perceived rendering quality thresholds. 2008.
- [170] D Rojas, B Kapralos, S Crsitancho, K Collins, C Conati, and A Dubrowski. The effect of background sound on visual fidelity perception. *Proc. ACM Audio Mostly*, pages 32–39, 2011.
- [171] Michio Shimizu, Kazunori Itoh, and Tatsuo Nakazawa. Pattern representation system using movement sense of localized sound. In *Proceedings of the HCI International’99 (the 8th International Conference on Human-Computer Interaction) on Human-Computer Interaction: Communication, Cooperation, and Application Design-Volume 2-Volume 2*, pages 990–994. L. Erlbaum Associates Inc., 1999.
- [172] Mayer L Max and Jesse R Gonzalez. Blind persons navigate in virtual reality (vr); hearing and feeling communicates reality”. *Studies in health technology and informatics*, 39:54–59, 1996.

- [173] José L González-Mora, A Rodríguez-Hernandez, LF Rodríguez-Ramos, L Díaz-Saco, and N Sosa. Development of a new space perception system for blind people, based on the creation of a virtual acoustic space. In *International Work-Conference on Artificial Neural Networks*, pages 321–330. Springer, 1999.
- [174] Myron W Krueger and Deborah Gilden. Going places with “knowware”: Virtual reality maps for blind people. In *International Conference on Computers for Handicapped Persons*, pages 565–567. Springer, 2002.
- [175] Jaime Sánchez. User-centered technologies for blind children. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, 4(2):96–122, 2008.
- [176] Makoto Ohuchi, Yukio Iwaya, and Yôiti Suzuki. Cognitive-map forming of the blind in virtual sound environment. 2006.
- [177] Yoshikazu Seki and Tetsuji Sato. A training system of orientation and mobility for blind people using acoustic virtual reality. *IEEE Transactions on neural systems and rehabilitation engineering*, 19(1):95–104, 2011.
- [178] Matilda Annerstedt, Peter Jönsson, Mattias Wallergård, Gerd Johansson, Björn Karlson, Patrik Grahn, Åse Marie Hansen, and Peter Währborg. Inducing physiological stress recovery with sounds of nature in a virtual reality forest—results from a pilot study. *Physiology & behavior*, 118:240–250, 2013.
- [179] Carl Schissler, Aaron Nicholls, and Ravish Mehra. Efficient hrtf-based spatial audio for area and volumetric sources. *IEEE transactions on visualization and computer graphics*, 22(4):1356–1366, 2016.
- [180] Jens Blauert. *Communication acoustics*, volume 2. Springer, 2005.
- [181] Dorte Hammersho, Henrik Mo, et al. Sound transmission to and within the human ear canal. *The Journal of the Acoustical Society of America*, 100(1):408–427, 1996.
- [182] John C Middlebrooks. Individual differences in external-ear transfer functions reduced by scaling in frequency. *The Journal of the Acoustical Society of America*, 106(3):1480–1492, 1999.
- [183] II Faller, Kenneth John, Armando Barreto, and Malek Adjouadi. Augmented hankel total least-squares decomposition of head-related transfer functions. *Journal of the Audio Engineering Society*, 58(1/2):3–21, 2010.

- [184] Catarina Mendonça, Guilherme Campos, Paulo Dias, José Vieira, João P Ferreira, and Jorge A Santos. On the improvement of localization accuracy with non-individualized hrtf-based sounds. *Journal of the Audio Engineering Society*, 60(10):821–830, 2012.
- [185] Timos Papadopoulos and Philip A Nelson. Choice of inverse filter design parameters in virtual acoustic imaging systems. *Journal of the Audio Engineering Society*, 58(1/2):22–35, 2010.
- [186] Elena Blanco-Martin, Francisco Javier Casajús-Quirós, Juan José Gómez-Alfageme, and Luis Ignacio Ortiz-Berenguer. Objective measurement of sound event localization in horizontal and median planes. *Journal of the Audio Engineering Society*, 59(3):124–136, 2011.
- [187] Er Väljamäe, Pontus Larsson, Daniel Västfjäll, and Mendel Kleiner. Auditory presence, individualized head-related transfer functions, and illusory ego-motion in virtual environments. In *in Proc. of Seventh Annual Workshop Presence 2004*. Citeseer, 2004.
- [188] Pauli Minnaar, S Krarup Olesen, Flemming Christensen, and Henrik Møller. Localization with binaural recordings from artificial and human heads. *Journal of the Audio Engineering Society*, 49(5):323–336, 2001.
- [189] Claudia Arias and Oscar A Ramos. Audición espacial en ambientes reverberantes: aspectos teóricos relevantes. *Revista interamericana de psicología= Interamerican journal of psychology*, 37(2):373–382, 2003.
- [190] Dmitry N Zotkin, Ramani Duraiswami, and Larry S Davis. Rendering localized spatial audio in a virtual auditory space. *IEEE Transactions on Multimedia*, 6(4):553–564, 2004.
- [191] Navarun Gupta, Armando Barreto, and Carlos Ordonez. Spectral modification of head-related transfer functions for improved virtual sound spatialization. In *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, volume 2, pages II–1953. IEEE, 2002.
- [192] Gemma Calvert, Charles Spence, and Barry E Stein. *The handbook of multisensory processes*. MIT press, 2004.
- [193] Touradj Ebrahimi. Quality of multimedia experience: past, present and future. In *MM'09: Proceedings of the seventeen ACM international conference on Multimedia*, pages 3–4, 2009.

- [194] Niall Murray, Brian Lee, Yuansong Qiao, and Gabriel-Miro Muntean. Olfaction-Enhanced multimedia: A survey of application domains, displays, and research challenges. *ACM Computing Surveys (CSUR)*, 48(4):56, May 2016.
- [195] L Buck and R Axel. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell*, 65(1):175–187, April 1991.
- [196] Joseph “jofish” Kaye. Making scents: Aromatic output for HCI. *Interactions*, 11(1):48–61, January 2004.
- [197] J N Kaye. Symbolic olfactory display. 2001.
- [198] Gheorghita Ghinea and Oluwakemi A Ademoye. Olfaction-enhanced multimedia: perspectives and challenges. *Multimed. Tools Appl.*, 55(3):601–626, December 2011.
- [199] N Murray, Y Qiao, B Lee, G M Muntean, and A K Karunakar. Age and gender influence on perceived olfactory and visual media synchronization. In *2013 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6, July 2013.
- [200] N Murray, Y Qiao, B Lee, G M Muntean ACM Transactions on, and 2014. User-profile-based perceived olfactory and visual media synchronization. *dl.acm.org*, 2014.
- [201] Ray Shih and Pieter Blignaut. The influence of gender and internet experience on the acceptability of smell as interaction modality. In *Proceedings of the South African Institute of Computer Scientists and Information Technologists Conference on Knowledge, Innovation and Leadership in a Diverse, Multidisciplinary Environment, SAICSIT '11*, pages 205–214, New York, NY, USA, 2011. ACM.
- [202] Scott Puopolo, Carlos Cordero, William Gerhardt, Kate Griffin, Leszek Izdebski, and David Parsons. The future of television: Sweeping change at breakneck speed. *Point of View document, Cisco Internet Business Solutions Group (IBSG)*, 2011.
- [203] B S Spencer. Incorporating the sense of smell into patient and haptic surgical simulators. *IEEE Trans. Inf. Technol. Biomed.*, 10(1):168–173, January 2006.

- [204] Ladan Shams and Aaron R Seitz. Benefits of multisensory learning. *Trends Cogn. Sci.*, 12(11):411–417, November 2008.
- [205] Graham M S Dann and Jens Kristian Steen. Leading the tourist by the nose. *The tourist as a metaphor of the social world*, page 209, 2002.
- [206] Joel S Warm, William N Dember, and R Parasuraman. Effects of olfactory stimulation on performance and stress. *J. Soc. Cosmet. Chem.*, 42(3):199–210, 1991.
- [207] Rachel S Herz. Are odors the best cues to memory? a cross-modal comparison of associative memory stimuli. *Ann. N. Y. Acad. Sci.*, 855(1):670–674, 1998.
- [208] Maureen Morrin and S Ratneshwar. The impact of ambient scent on evaluation, attention, and memory for familiar and unfamiliar brands. *J. Bus. Res.*, 49(2):157–165, August 2000.
- [209] Stephen Brewster, David McGookin, and Christopher Miller. Olfoto: Designing a smell-based interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '06*, pages 653–662, New York, NY, USA, 2006. ACM.
- [210] D A Washburn, L M Jones, R Vijaya Satya, C A Bowers, and A Cortes. Olfactory use in virtual environment training. *Modeling & Simulation*, 2(3):19–25, 2003.
- [211] Sylvain Delplanque, Didier Grandjean, Christelle Chrea, Laurence Ay-mard, Isabelle Cayeux, Bénédicte Le Calvé, Maria Inés Velazco, Klaus R Scherer, and David Sander. Emotional processing of odors: evidence for a nonlinear relation between pleasantness and familiarity evaluations. *Chem. Senses*, 33(5):469–479, June 2008.
- [212] Masaaki Yoshida. Psychometric classification of odors. *Chem. Senses*, 1(4):443–464, 1975.
- [213] A N Rétiveau, E Chambers Iv, and G A Milliken. COMMON AND SPECIFIC EFFECTS OF FINE FRAGRANCES ON THE MOOD OF WOMEN. *J. Sens. Stud.*, 19(5):373–394, October 2004.
- [214] Woodrow Barfield and Eric Danas. Comments on the use of olfactory displays for virtual environments. *Presence: Teleoperators and Virtual Environments*, 5(1):109–121, January 1996.

- [215] Adam Bodnar, Richard Corbett, and Dmitry Nekrasovski. AROMA: Ambient awareness through olfaction in a messaging application. In *Proceedings of the 6th International Conference on Multimodal Interfaces, ICMI '04*, pages 183–190, New York, NY, USA, 2004. ACM.
- [216] Yongsoon Choi, Adrian David Cheok, Xavier Roman, The Anh Nguyen, Kenichi Sugimoto, and Veronica Halupka. Sound perfume: Designing a wearable sound and fragrance media for face-to-face interpersonal interaction. In *Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology, ACE '11*, pages 4:1–4:8, New York, NY, USA, 2011. ACM.
- [217] Sue Ann Seah, Diego Martinez Plasencia, Peter D Bennett, Abhijit Karnik, Vlad Stefan Otrocol, Jarrod Knibbe, Andy Cockburn, and Sriram Subramanian. SensaBubble: A chrono-sensory mid-air display of sight and smell. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14*, pages 2863–2872, New York, NY, USA, 2014. ACM.
- [218] Christian Friedrich Schönbein. On some secondary physiological effects produced by atmospheric electricity. *Med. Chir. Trans.*, (1):205–220, 1851.
- [219] Gonzalo Ramos, Matthew Boulos, and Ravin Balakrishnan. Pressure widgets. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '04*, pages 487–494, New York, NY, USA, 2004. ACM.
- [220] 870 U S Patent 3, 050 and 1962. Sensorama simulator, 1962.
- [221] Cater J.P. The nose have it! *Presence: Teleoperators and Virtual Environments*, 1(4):493–494, 1992.
- [222] Robinett W. Comments on “.^a nose gesture interface device: Extending virtual realities”. *Presence: Teleoperators and Virtual Environments*, 1(4):493, 1992.
- [223] Belma R Brkic, Alan Chalmers, Kevin Boulanger, Sumanta Pattanaik, and James Covington. Cross-modal affects of smell on the real-time rendering of grass. In *Proceedings of the 25th Spring Conference on Computer Graphics, SCCG '09*, pages 161–166, New York, NY, USA, 2009. ACM.
- [224] D W Kim, Y H Cho, K Nishimoto, Y Kawakami, S Kunifujji, and H Ando. Development of aroma-card based soundless olfactory display. In *2009 16th IEEE International Conference on Electronics, Circuits and Systems - (ICECS 2009)*, pages 703–706, December 2009.

- [225] E C Tan, A Wahab, G H Goh, and S H Wong. PC-controlled scent system. *IEEE Trans. Consum. Electron.*, 44(1):130–136, February 1998.
- [226] Y Ariyakul and T Nakamoto. Olfactory display using a miniaturized pump and a SAW atomizer for presenting low-volatile scents. In *2011 IEEE Virtual Reality Conference*, pages 193–194, March 2011.
- [227] Usman Haque. The choreography of sensations: Three case studies of responsive environment interfaces. In *VSM 2004: Proceedings of the tenth International Conference on Virtual Systems and Multimedia: Hybrid realities & digital partners—explorations in art, heritage, science & the human factor, 17-19 November 2004, Softopia Japan, Ogaki City, Japan, 2004*.
- [228] Usman Haque. Scents of space: An interactive smell system. In *ACM SIGGRAPH 2004 Sketches, SIGGRAPH '04*, pages 35–, New York, NY, USA, 2004. ACM.
- [229] Bernadette Emsenhuber and Alois Ferscha. Olfactory interaction zones. In *International Conference on Pervasive Computing, Pervasive*, pages 11–14, 2009.
- [230] A Nambu, T Narumi, K Nishimura, T Tanikawa, and M Hirose. Visual-olfactory display using olfactory sensory map. In *2010 IEEE Virtual Reality Conference (VR)*, pages 39–42, March 2010.
- [231] T Tanikawa and M Hirose. A study of multi-modal display system with visual feedback. In *2008 Second International Symposium on Universal Communication*, pages 285–292, December 2008.
- [232] Daisuke Noguchi, Kaori Ohtsu, Yuichi Bannai, and Ken-Ichi Okada. Scent presentation expressing two smells of different intensity simultaneously. In *EGVE/ICAT/EuroVR*, pages 53–60, 2009.
- [233] Daisuke Noguchi, Sayumi Sugimoto, Yuichi Bannai, and Ken-Ichi Okada. Time characteristics of olfaction in a single breath. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11*, pages 83–92, New York, NY, USA, 2011. ACM.
- [234] A Kadowaki, D Noguchi, S Sugimoto, Y Bannai, and K Okada. Development of a High-Performance olfactory display and measurement of olfactory characteristics for pulse ejections. In *2010 10th IEEE/IPSJ International Symposium on Applications and the Internet*, pages 1–6, July 2010.

- [235] A Kadowaki, J Sato, Y Bannai, and K i. Okada. Presentation technique of scent to avoid olfactory adaptation. In *17th International Conference on Artificial Reality and Telexistence (ICAT 2007)*, pages 97–104, November 2007.
- [236] J Sato, K Ohtsu, Y Bannai, and Ki. Okada. Effective presentation technique of scent using small ejection quantities of odor. In *2009 IEEE Virtual Reality Conference*, pages 151–158, March 2009.
- [237] Junta Sato, Kaori Ohtsu, Yuichi Bannai, and Ken-Ichi Okada. Pulse ejection technique of scent to create dynamic perspective. In *Proceedings of the 18th International Conference on Artificial Reality and Telexistence*, pages 167–174, 2008.
- [238] K Ohtsu, J Sato, Y Bannai, and K i. Okada. Scent presentation technique of pulse ejection synchronized with breathing. In *2009 Ninth Annual International Symposium on Applications and the Internet*, pages 125–128, July 2009.
- [239] Kaori Ohtsu, Junta Sato, Yuichi Bannai, and Kenichi Okada. Measurement of olfactory characteristics for two kinds of scent in a single breath. In *Human-Computer Interaction – INTERACT 2009*, Lecture Notes in Computer Science, pages 306–318. Springer, Berlin, Heidelberg, August 2009.
- [240] Sayumi Sugimoto, Ryo Segawa, Daisuke Noguchi, Yuichi Bannai, and Kenichi Okada. Presentation technique of scents using mobile olfactory display for digital signage. In *Human-Computer Interaction – INTERACT 2011*, Lecture Notes in Computer Science, pages 323–337. Springer, Berlin, Heidelberg, September 2011.
- [241] Changhoon Park, Heedong Ko, Ig-Jae Kim, Sang Chul Ahn, Yong-Moo Kwon, and Hyoung-Gon Kim. The making of kyongju VR theatre. In *Proceedings IEEE Virtual Reality 2002*, pages 269–270, 2002.
- [242] Kentaro Tominaga, Shinkuro Honda, Takaharu Ohsawa, Hiroshi Shigeno, Ken-Ichi Okada, and Yutaka Matsushita. “friend park”-expression of the wind and the scent on virtual space. In *Virtual Systems and Multimedia, 2001. Proceedings. Seventh International Conference on*, pages 507–515, 2001.
- [243] Sensory acumen inc. <http://www.sensoryacumen.com>. accedido en febrero 11, 2016.

- [244] H Matsukura, T Yoneda, and H Ishida. Smelling screen: Technique to present a virtual odor source at an arbitrary position on a screen. In *2012 IEEE Virtual Reality Workshops (VRW)*, pages 127–128, March 2012.
- [245] Y Yanagida, S Kawato, H Noma, A Tomono, and N Tesutani. Projection based olfactory display with nose tracking. In *IEEE Virtual Reality 2004*, pages 43–50, March 2004.
- [246] Yasuyuki Yanagida, Haruo Noma, Nobuji Tetsutani, and Akira Tomono. An unencumbering, localized olfactory display. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '03, pages 988–989, New York, NY, USA, 2003. ACM.
- [247] Akira Tomono and Syunya Otake. The eye catching property of Digital-Signage with scent and a Scent-Emitting video display system. *IEEJ Transactions on Electronics, Information and Systems*, 130:668–675, 2010.
- [248] K Tomono, H Katsuyama, and A Tomono. A scent-emitting video display system. In *2011 IEEE International Symposium on VR Innovation*, pages 335–336, March 2011.
- [249] Keisuke Tomono, Hiroki Wakatsuki, Shigeki Kumazawa, and Akira Tomono. Display system for advertising image with scent and psychological effect. *Engineering Psychology and Cognitive Ergonomics*, pages 110–119, 2011.
- [250] Minoru Sakairi, Ayako Nishimura, and Daisuke Suzuki. Olfaction presentation system using odor scanner and odor-emitting apparatus coupled with chemical capsules of alginic acid polymer. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 92(2):618–629, 2009.
- [251] T Yamada, S Yokoyama, T Tanikawa, K Hirota, and M Hirose. Wearable olfactory display: Using odor in outdoor environment. In *IEEE Virtual Reality Conference (VR 2006)*, pages 199–206, March 2006.
- [252] Venere Ferraro and Secil Ugur. Designing wearable technologies through a user centered approach. In *Proceedings of the 2011 Conference on Designing Pleasurable Products and Interfaces*, DPPI '11, pages 5:1–5:8, New York, NY, USA, 2011. ACM.
- [253] G Jenkins, A Manz, and J Tillotson. Scent whisper. pages 97–104, January 2006.

- [254] H Matsukura, H Yoshida, H Ishida, A Saitoh, and T Nakamoto. Odor presentation with a vivid sense of reality: Incorporating fluid dynamics simulation into olfactory display. In *2009 IEEE Virtual Reality Conference*, pages 295–296, March 2009.
- [255] H Matsukura, H Yoshida, H Ishida, and T Nakamoto. Interactive odor playback based on fluid dynamics simulation. In *2009 IEEE Virtual Reality Conference*, pages 255–256, March 2009.
- [256] Haruka Matsukura, Akira Ohno, and Hiroshi Ishida. Fluid dynamic considerations for realistic odor presentation using olfactory display. *Presence: Teleoperators and Virtual Environments*, 19(6):513–526, December 2010.
- [257] Haruka Matsukura, Hitoshi Yoshida, Takamichi Nakamoto, and Hiroshi Ishida. Synchronized presentation of odor with airflow using olfactory display. *J. Mech. Sci. Technol.*, 24(1):253–256, January 2010.
- [258] Takuji Narumi, Shinya Nishizaka, Takashi Kajinami, Tomohiro Tanikawa, and Michitaka Hirose. Augmented reality flavors: Gustatory display based on edible marker and cross-modal interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11*, pages 93–102, New York, NY, USA, 2011. ACM.
- [259] Y Yanagida. A survey of olfactory displays: Making and delivering scents. In *2012 IEEE Sensors*, pages 1–4, October 2012.
- [260] Takamichi Nakamoto, Shigeki Otaguro, Masashi Kinoshita, Masahiko Nagahama, Keita Ohinishi, and Taro Ishida. Cooking up an interactive olfactory game display. *IEEE Comput. Graph. Appl.*, 28(1):75–78, January 2008.
- [261] J Nishida and K Suzuki. biosync: Wearable haptic I/O device for synchronous kinesthetic interaction. In *2016 IEEE Virtual Reality (VR)*, pages 243–244, March 2016.
- [262] T Narumi, S Nishizaka, T Kajinami, T Tanikawa, and M Hirose. Meta-Cookie #xoo2b;. In *2011 IEEE Virtual Reality Conference*, pages 265–266, March 2011.
- [263] Nimesha Ranasinghe, Thi Ngoc Tram Nguyen, Yan Liangkun, Lien-Ya Lin, David Tolley, and Ellen Yi-Luen Do. Vocktail: A virtual cocktail for pairing digital taste, smell, and color sensations. In *Proceedings of the 2017 ACM on Multimedia Conference*, pages 1139–1147. ACM, October 2017.

- [264] Project nourished. <http://www.projectnourished.com/>. accedido en noviembre 11, 2017.
- [265] Ohroma. <https://www.camsoda.com/products/ohroma/>. accedido en noviembre 11, 2017.
- [266] Surina Hariri, Nur Ain Mustafa, Kasun Karunanayaka, and Adrian David Cheok. Electrical stimulation of olfactory receptors for digitizing smell. In *Proceedings of the 2016 workshop on Multimodal Virtual and Augmented Reality*, page 4. ACM, November 2016.
- [267] Mario Covarrubias, Monica Bordegoni, Mauro Rosini, Eleonora Guanzi-rolì, Umberto Cugini, and Franco Molteni. VR system for rehabilitation based on hand gestural and olfactory interaction. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, pages 117–120. ACM, November 2015.
- [268] Mary P Aiken and Mike J Berry. Posttraumatic stress disorder: possibilities for olfaction and virtual reality exposure therapy. *Virtual Real.*, 19(2):95–109, June 2015.
- [269] Oliver Baus and Stéphane Bouchard. Exposure to an unpleasant odour increases the sense of presence in virtual reality. *Virtual Real.*, 21(2):59–74, June 2017.
- [270] K Hashimoto and T Nakamoto. Olfactory display using surface acoustic wave device and micropumps for wearable applications. In *2016 IEEE Virtual Reality (VR)*, pages 179–180, March 2016.
- [271] Heiling M.L. The cinema of the future. *Presence: Teleoperators and Virtual Environments*, 1(3):279–294, 1992.
- [272] Michitaka Hirose, T Tanikawa, S Tanaka, and S Sakikawa. A study of olfactory display. In *Proc. of the Virtual Reality Society of Japan 5th Annual Conference*, pages 193–196, 2000.
- [273] Arito Mochizuki, Takashi Amada, Sayuri Sawa, Tadayuki Takeda, Shogo Motoyashiki, Kazuhiro Kohyama, Masataka Imura, and Kunihiro Chihara. Fragra: A visual-olfactory VR game. In *ACM SIGGRAPH 2004 Sketches, SIGGRAPH '04*, pages 123–, New York, NY, USA, 2004. ACM.
- [274] Takamichi Nakamoto. *Human olfactory displays and interfaces: Odor sensing and presentation*. IGI Global Snippet, 2013.

- [275] Atif Alamri, Mohamad Eid, Rosa Iglesias, Shervin Shirmohammadi, and Abdulmotaleb El Saddik. Haptic virtual rehabilitation exercises for post-stroke diagnosis. *IEEE transactions on instrumentation and measurement*, 57(9):1876–1884, 2008.
- [276] Quan-Zen Ang, Ben Horan, Zoran Najdovski, and Saeid Nahavandi. Grasping virtual objects with multi-point haptics. In *Virtual Reality Conference (VR), 2011 IEEE*, pages 189–190. IEEE, 2011.
- [277] Woodrow Barfield and Claudia Hendrix. The effect of update rate on the sense of presence within virtual environments. *Virtual Reality*, 1(1):3–15, 1995.
- [278] Huong Q Dinh, Neff Walker, Larry F Hodges, Chang Song, and Akira Kobayashi. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Virtual Reality, 1999. Proceedings., IEEE*, pages 222–228. IEEE, 1999.
- [279] Christopher D Wickens, Oscar Olmos, Andrew Chudy, and Clark Davenport. Aviation display support for situation awareness. Technical report, DTIC Document, 1997.
- [280] Monica Bordegoni, Umberto Cugini, and Francesco Ferrise. Analysis of designers’ manual skills for the development of a tool for aesthetic shapes evaluation. In *Proceedings of IDMME*, 2010.
- [281] RS Johansson. Sensory guidance and movement, 1998.
- [282] Werner Goebel and Caroline Palmer. Tactile feedback and timing accuracy in piano performance. *Experimental Brain Research*, 186(3):471–479, 2008.
- [283] Caroline Palmer, Erik Koopmans, Janeen D Loehr, and Christine Carter. Movement-related feedback and temporal accuracy in clarinet performance. *Music Perception: An Interdisciplinary Journal*, 26(5):439–449, 2009.
- [284] Sarah F Tyson, Marie Hanley, Jay Chillala, Andrea B Selley, and Raymond C Tallis. Sensory loss in hospital-admitted people with stroke: characteristics, associated factors, and relationship with function. *Neuro-rehabilitation and Neural Repair*, 22(2):166–172, 2008.
- [285] Lynette A Jones and Susan J Lederman. *Human hand function*. Oxford University Press, 2006.

- [286] A Puertas, P Pures, AM Echenique, et al. Braille line using electrical stimulation. In *Journal of Physics: Conference Series*, volume 90, page 012091. IOP Publishing, 2007.
- [287] David F Collins, Kathryn M Refshauge, Gabrielle Todd, and Simon C Gandevia. Cutaneous receptors contribute to kinesthesia at the index finger, elbow, and knee. *Journal of Neurophysiology*, 94(3):1699–1706, 2005.
- [288] PM Taylor, A Moser, and A Creed. A sixty-four element tactile display using shape memory alloy wires. *Displays*, 18(3):163–168, 1998.
- [289] T Matsunaga, K Totsu, M Esashi, and Y Haga. Tactile display for 2-d and 3-d shape expression using sma micro actuators. In *Microtechnology in Medicine and Biology, 2005. 3rd IEEE/EMBS Special Topic Conference on*, pages 88–91. IEEE, 2005.
- [290] Michael F Nolan. Two-point discrimination assessment in the upper limb in young adult men and women. *Phys Ther*, 62(7):965–9, 1982.
- [291] Wen Qi, Russell M Taylor II, Christopher G Healey, and Jean-Bernard Martens. A comparison of immersive hmd, fish tank vr and fish tank with haptics displays for volume visualization. In *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization*, pages 51–58. ACM, 2006.
- [292] Joseph C Stevens and Kenneth K Choo. Spatial acuity of the body surface over the life span. *Somatosensory & motor research*, 13(2):153–166, 1996.
- [293] J Dargahi and S Najarian. Human tactile perception as a standard for artificial tactile sensing—a review. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 1(1):23–35, 2004.
- [294] Vasilios G Chouvardas, Amalia N Miliou, and Miltiadis K Hatalis. Tactile displays: Overview and recent advances. *Displays*, 29(3):185–194, 2008.
- [295] Daniel Leithinger, Sean Follmer, Alex Olwal, and Hiroshi Ishii. Shape displays: Spatial interaction with dynamic physical form. *IEEE Comput. Graph. Appl.*, 35(5):5–11, September 2015.
- [296] Hsin-Ni Ho and Lynette A Jones. Contribution of thermal cues to material discrimination and localization. *Attention, Perception, & Psychophysics*, 68(1):118–128, 2006.

- [297] Geoff V Merrett, Cheryl D Metcalf, Deyi Zheng, Sarah Cunningham, Stuart Barrow, and Sara H Demain. Design and qualitative evaluation of tactile devices for stroke rehabilitation. In *Assisted Living 2011, IET Seminar on*, pages 1–6. IET, 2011.
- [298] Andrew YJ Szeto and Frank A Saunders. Electrocutaneous stimulation for sensory communication in rehabilitation engineering. *IEEE Transactions on Biomedical Engineering*, (4):300–308, 1982.
- [299] SH Peurala, K Pitkänen, J Sivenius, and IM Tarkka. Cutaneous electrical stimulation may enhance sensorimotor recovery in chronic stroke. *Clinical Rehabilitation*, 16(7):709–716, 2002.
- [300] Hiroyuki Kajimoto, Naoki Kawakami, Taro Maeda, and Susumu Tachi. Electro-tactile display with force feedback. In *Proc. World Multiconference on Systemics, Cybernetics and Informatics (SCI2001)*, volume 11, pages 95–99, 2001.
- [301] Hiroyuki Kajimoto. Electro-tactile display with real-time impedance feedback. *Haptics: Generating and Perceiving Tangible Sensations*, pages 285–291, 2010.
- [302] Kouta Minamizawa, Souichiro Fukamachi, Naoki Kawakami, and Susumu Tachi. Interactive representation of virtual object in hand-held box by finger-worn haptic display. In *Haptic interfaces for virtual environment and teleoperator systems, 2008. haptics 2008. symposium on*, pages 367–368. IEEE, 2008.
- [303] PM Taylor, A Hosseini-Sianaki, CJ Varley, and DM Pollet. Advances in an electrorheological fluid based tactile array. 1997.
- [304] Robert Scheibe, Mathias Moehring, and Bernd Froehlich. Tactile feedback at the finger tips for improved direct interaction in immersive environments. In *3D User Interfaces, 2007. 3DUI'07. IEEE Symposium on*. IEEE, 2007.
- [305] Margaret McLaughlin, Albert Rizzo, Younbo Jung, Wei Peng, S Yeh, and Weirong Zhu. Haptics-enhanced virtual environments for stroke rehabilitation. *Proc. IPSI*, 2005.
- [306] Craig D Takahashi, Lucy Der-Yeghiaian, VH Le, and Steven C Cramer. A robotic device for hand motor therapy after stroke. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pages 17–20. IEEE, 2005.

- [307] Antonio Frisoli, Fabrizio Rocchi, Simone Marcheschi, Andrea Dettori, Fabio Salsedo, and Massimo Bergamasco. A new force-feedback arm exoskeleton for haptic interaction in virtual environments. In *Eurohaptics Conference, 2005 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, 2005. World Haptics 2005. First Joint*, pages 195–201. IEEE, 2005.
- [308] Chester Gary Nelson, Michael Thomas Lee, and Michael G Schrenk. Apparatus and method for remote therapy and diagnosis in medical devices via interface systems, July 9 2002. US Patent 6,418,346.
- [309] RARC Gopura and Kazuo Kiguchi. Mechanical designs of active upper-limb exoskeleton robots: State-of-the-art and design difficulties. In *Rehabilitation Robotics, 2009. ICORR 2009. IEEE International Conference on*, pages 178–187. IEEE, 2009.
- [310] Ho Shing Lo and Sheng Quan Xie. Exoskeleton robots for upper-limb rehabilitation: State of the art and future prospects. *Medical engineering & physics*, 34(3):261–268, 2012.
- [311] J Edward Colgate, Michael C Stanley, and J Michael Brown. Issues in the haptic display of tool use. In *Intelligent Robots and Systems 95. 'Human Robot Interaction and Cooperative Robots', Proceedings. 1995 IEEE/RSJ International Conference on*, volume 3, pages 140–145. IEEE, 1995.
- [312] Seiya Takei, Ryo Watanabe, Ryuta Okazaki, Taku Hachisu, Michi Sato, and Hiroyuki Kajimoto. WholeGrip: grip type master hand with the whole hand tactile feedback. In *SIGGRAPH Asia 2015 Emerging Technologies*, page 27. ACM, November 2015.
- [313] Yi-Chi Liao, Shun-Yao Yang, Rong-Hao Liang, Liwei Chan, and Bing-Yu Chen. ThirdHand: wearing a robotic arm to experience rich force feedback. In *SIGGRAPH Asia 2015 Emerging Technologies*, page 24. ACM, November 2015.
- [314] Ruimin Lyu, Haojie Hao, Wei Chen, Yuan Liu, Feng Wang, and Allen C H Wu. Elastylus: flexible haptic painting stylus. In *SIGGRAPH Asia 2015 Emerging Technologies*, page 10. ACM, November 2015.
- [315] Kentaro Yasu and Yuichiro Katsumoto. Bump ahead: easy-to-design haptic surface using magnet array. In *SIGGRAPH Asia 2015 Emerging Technologies*, page 3. ACM, November 2015.

- [316] Ginga Kato, Yoshihiro Kuroda, Ilana Nisky, Kiyoshi Kiyokawa, and Haruo Takemura. HapSticks: tool-mediated interaction with grounding-free haptic interface. In *SIGGRAPH Asia 2015 Haptic Media And Contents Design*, page 7. ACM, November 2015.
- [317] Yuichiro Katsumoto. Bottomless joystick. In *SIGGRAPH ASIA 2016 Emerging Technologies*, page 6. ACM, November 2016.
- [318] Zhizhuo Yang and Dongdong Weng. Passive haptics based MR system for geography teaching. In *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry - Volume 1*, pages 23–29. ACM, December 2016.
- [319] L Shapira, J Amores, and X Benavides. TactileVR: Integrating physical toys into learn and play virtual reality experiences. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 100–106, 2016.
- [320] Yuki Ban, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. Displaying shapes with various types of surfaces using visuo-haptic interaction. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pages 191–196. ACM, November 2014.
- [321] Antti Sand, Ismo Rakkolainen, Poika Isokoski, Jari Kangas, Roope Raimo, and Karri Palovuori. Head-mounted display with mid-air tactile feedback. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, pages 51–58. ACM, November 2015.
- [322] Qianqian Tong, Zhiyong Yuan, Mianlun Zheng, Weixu Zhu, Guian Zhang, and Xiangyun Liao. A novel magnetic levitation haptic device for augmentation of tissue stiffness perception. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, pages 143–152. ACM, November 2016.
- [323] S Hashiguchi, F Shibata, and A Kimura. Psychophysical influence on temperature perception by mixed-reality visual stimulation. In *2016 IEEE Virtual Reality (VR)*, pages 177–178, March 2016.
- [324] J Hummel, J Dodiya, G A Center, L Eckardt, R Wolff, A Gerndt, and T W Kuhlen. A lightweight electrotactile feedback device for grasp improvement in immersive virtual environments. In *2016 IEEE Virtual Reality (VR)*, pages 39–48, March 2016.

- [325] T Handa, K Murase, M Azuma, T Shimizu, S Kondo, and H Shinoda. A haptic three-dimensional shape display with three fingers grasping. In *2017 IEEE Virtual Reality (VR)*, pages 325–326, March 2017.
- [326] E Vonach, C Gatterer, and H Kaufmann. VRRobot: Robot actuated props in an infinite virtual environment. In *2017 IEEE Virtual Reality (VR)*, pages 74–83, March 2017.
- [327] Dean Chang. Haptics: gaming’s new sensation. *Computer*, 35(8):84–86, 2002.
- [328] Johnny Chung Lee. Hacking the nintendo wii remote. *IEEE pervasive computing*, 7(3), 2008.
- [329] Available at: <http://www.cyberglovesystems.com>. accessed on: 24 november 2010.
- [330] Sergei V Adamovich, Alma S Merians, Rares Boian, Jeffrey A Lewis, Marilyn Tremaine, Grigore S Burdea, Michael Recce, and Howard Poizner. A virtual reality—based exercise system for hand rehabilitation post-stroke. *Presence: Teleoperators and Virtual Environments*, 14(2):161–174, 2005.
- [331] Tomosuke Maeda, Roshan Peiris, Nakatani Masashi, Yoshihiro Tanaka, and Kouta Minamizawa. HapticAid: wearable haptic augmentation system for enhanced, enchanted and empathised haptic experiences. In *SIGGRAPH ASIA 2016 Emerging Technologies*, page 4. ACM, November 2016.
- [332] V A de Jesus Oliveira, L Brayda, L Nedel, and A Maciel. Experiencing guidance in 3D spaces with a vibrotactile head-mounted display. In *2017 IEEE Virtual Reality (VR)*, pages 453–454, March 2017.
- [333] Chien-Min Wu, Chih-Wen Hsu, Tzu-Kuei Lee, and Shana Smith. A virtual reality keyboard with realistic haptic feedback in a fully immersive virtual environment. *Virtual Real.*, 21(1):19–29, March 2017.
- [334] Jerome J Palazzolo, Mark Ferraro, Hermano Igo Krebs, Daniel Lynch, Bruce T Volpe, and Neville Hogan. Stochastic estimation of arm mechanical impedance during robotic stroke rehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15(1):94–103, 2007.

- [335] Ludovic Dovat, Olivier Lambercy, Y Ruffieux, Dominique Chapuis, Roger Gassert, Hannes Bleuler, CL Teo, and Etienne Burdet. A haptic knob for rehabilitation of stroke patients. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 977–982. IEEE, 2006.
- [336] Available at: <http://www.sensable.com/products-hapticdevices.htm>. accessed on: 30 september 2011.
- [337] Kenneth Salisbury, David Brock, Thomas Massie, Nick Swarup, and Craig Zilles. Haptic rendering: Programming touch interaction with virtual objects. In *Proceedings of the 1995 symposium on Interactive 3D graphics*, pages 123–130. ACM, 1995.
- [338] Katsunori Miyahara and Yoshihiro Okada. A surgical simulation system supporting collada-based file format. In *Computer Graphics, Imaging and Visualisation, 2008. CGIV'08. Fifth International Conference on*, pages 126–131. IEEE, 2008.
- [339] Deok-Jae Song and Jinah Park. AnyHaptics: a haptic plug-in for existing interactive 3D graphics applications. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pages 27–30. ACM, November 2014.
- [340] Phantom omni. <http://www.geomagic.com/es/products/phantom-omni/overview>. accedido en junio 2017.
- [341] Qi Xing, Jim X Chen, Jihui Li, Ali Moshirfar, Mark M Theiss, and Qi Wei. A real time haptic simulator of spine surgeries. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, pages 121–124. ACM, November 2015.
- [342] T Nakano and Y Yanagida. Conditions influencing perception of wind direction by the head. In *2017 IEEE Virtual Reality (VR)*, pages 229–230, March 2017.
- [343] Amy Boos, Qinyin Qiu, Gerard G Fluet, and Sergei V Adamovich. Haptically facilitated bimanual training combined with augmented visual feedback in moderate to severe hemiplegia. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 3111–3114. IEEE, 2011.
- [344] Michele Confalonieri, Giovanni Guandalini, Mauro Da Lio, and Mariolino De Cecco. Force and touch make video games serious for dexterity rehabilitation. *pHealth*, pages 139–144, 2012.

- [345] Yong Mi Kim, Min Ho Chun, Gi Jeong Yun, Young Jin Song, and Han Eun Young. The effect of virtual reality training on unilateral spatial neglect in stroke patients. *Annals of rehabilitation medicine*, 35(3):309–315, 2011.
- [346] AS Merians, E Tunik, GG Fluet, Q Qiu, and SV Adamovich. Innovative approaches to the rehabilitation of upper extremity hemiparesis using virtual environments. *Eur J Phys Rehabil Med*, 45(1):123–133, 2009.
- [347] Alma S Merians, Howard Poizner, Rares Boian, Grigore Burdea, and Sergei Adamovich. Sensorimotor training in a virtual reality environment: does it improve functional recovery poststroke? *Neurorehabilitation and neural repair*, 20(2):252–267, 2006.
- [348] Giuseppe Placidi. A smart virtual glove for the hand telerehabilitation. *Computers in Biology and Medicine*, 37(8):1100–1107, 2007.
- [349] Jeonghun Ku, Richard Mraz, Nicole Baker, Konstantine K Zakzanis, Jang Han Lee, In Y Kim, Sun I Kim, and Simon J Graham. A data glove with tactile feedback for fmri of virtual reality experiments. *Cyberpsychology & behavior*, 6(5):497–508, 2003.
- [350] Qinyin Qiu, Gerard G Fluet, Soha Saleh, Ian Lafond, Alma S Merians, and Sergei V Adamovich. Integrated versus isolated training of the hemiparetic upper extremity in haptically rendered virtual environments. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 2255–2258. IEEE, 2010.
- [351] Patricia S Smith, Hubert R Dinse, Tobias Kalisch, Mark Johnson, and Delaina Walker-Batson. Effects of repetitive electrical stimulation to treat sensory loss in persons poststroke. *Archives of physical medicine and rehabilitation*, 90(12):2108–2111, 2009.
- [352] Dao M Vo, Judy M Vance, and Mervyn G Marasinghe. Assessment of haptics-based interaction for assembly tasks in virtual reality. In *EuroHaptics conference, 2009 and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics 2009. Third Joint*, pages 494–499. IEEE, 2009.
- [353] Nintendo wiimote. <https://es.wikipedia.org/wiki/wiimote>. accedido en abril 2017.
- [354] M Bouzit, G Popescu, G Burdea, and R Boian. The rutgers master II-ND force feedback glove. In *Proceedings 10th Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. HAPTICS 2002*, pages 145–152, 2002.

- [355] A Kron, M Buss, and G Schmidt. Exploration and manipulation of virtual environments using a combined hand and finger force feedback system. In *Proceedings. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000) (Cat. No.00CH37113)*, volume 2, pages 1328–1333 vol.2, 2000.
- [356] Domenico Prattichizzo, Claudio Pacchierotti, Stefano Cenci, Kouta Minamizawa, and Giulio Rosati. Using a fingertip tactile device to substitute kinesthetic feedback in haptic interaction. In *Haptics: Generating and Perceiving Tangible Sensations*, Lecture Notes in Computer Science, pages 125–130. Springer, Berlin, Heidelberg, July 2010.
- [357] Martin O Culjat, Ji Son, Richard E Fan, Christopher Wottawa, James W Bisley, Warren S Grundfest, and Erik P Dutson. Remote tactile sensing glove-based system. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*, pages 1550–1554. IEEE, 2010.
- [358] Taclim. <https://taclim.cerevo.com/en/>. accedido en junio 2017.
- [359] Hirro Iwata. History of haptic interface. *Human haptic perception: Basics and applications*, pages 355–361, 2008.
- [360] Sidney Weinstein. Intensive and extensive aspects of tactile sensitivity as a function of body part, sex and laterality. In *the First Int'l symp. on the Skin Senses, 1968*, 1968.
- [361] Alberto Gallace, Hong Z Tan, and Charles Spence. The failure to detect tactile change: A tactile analogue of visual change blindness. *Psychonomic bulletin & review*, 13(2):300–303, 2006.
- [362] A Gallace, E Imbornone, and G Vallar. When the whole is more than the sum of the parts: evidence from visuospatial neglect. *Journal of neuropsychology*, 2(2):387–413, 2008.
- [363] Daniel C Dennett. *Consciousness explained*. boston (little, brown and co) 1991. 1991.
- [364] Carl E Sherrick, Roger W Cholewiak, and Amy A Collins. The localization of low-and high-frequency vibrotactile stimuli. *The Journal of the Acoustical Society of America*, 88(1):169–179, 1990.
- [365] Yoseph Bar-Cohen. Haptic devices for virtual reality, telepresence, and human-assistive robotics. *Biol Inspired Intell Robots*, 122:73, 2003.

- [366] Ronald T Verrillo. Psychophysics of vibrotactile stimulation. *The Journal of the Acoustical Society of America*, 77(1):225–232, 1985.
- [367] C Elaine Chapman and Evelyne Beauchamp. Differential controls over tactile detection in humans by motor commands and peripheral reafference. *Journal of Neurophysiology*, 96(3):1664–1675, 2006.
- [368] Anita Cybulska-Klosowicz, El-Mehdi Meftah, Mélissa Raby, Marie-Line Lemieux, and C Elaine Chapman. A critical speed for gating of tactile detection during voluntary movement. *Experimental brain research*, 210(2):291–301, 2011.
- [369] Georgiana Juravle, Heiner Deubel, and Charles Spence. Attention and suppression affect tactile perception in reach-to-grasp movements. *Acta psychologica*, 138(2):302–310, 2011.
- [370] Georgiana Juravle, Heiner Deubel, Hong Z Tan, and Charles Spence. Changes in tactile sensitivity over the time-course of a goal-directed movement. *Behavioural brain research*, 208(2):391–401, 2010.
- [371] Georgiana Juravle and Charles Spence. Juggling reveals a decisional component to tactile suppression. *Experimental brain research*, 213(1):87–97, 2011.
- [372] Martin Voss, James N Ingram, Patrick Haggard, and Daniel M Wolpert. Sensorimotor attenuation by central motor command signals in the absence of movement. *Nature neuroscience*, 9(1):26–27, 2006.
- [373] M Bergamasco, A Frisoli, A Gucciardino, and S Scattareggia Marchese. Haptic interfaces for virtual prototyping. Technical report, DTIC Document, 2003.
- [374] Steve Guest, Caroline Catmur, Donna Lloyd, and Charles Spence. Audio-tactile interactions in roughness perception. *Experimental Brain Research*, 146(2):161–171, 2002.
- [375] Steve Guest and Charles Spence. Tactile dominance in speeded discrimination of textures. *Experimental Brain Research*, 150(2):201–207, 2003.
- [376] Steve Guest and Charles Spence. What role does multisensory integration play in the visuotactile perception of texture? *International Journal of Psychophysiology*, 50(1):63–80, 2003.

- [377] Ju-Hwan Lee, Ellen Poliakoff, and Charles Spence. The effect of multimodal feedback presented via a touch screen on the performance of older adults. In *International Conference on Haptic and Audio Interaction Design*, pages 128–135. Springer, 2009.
- [378] Razer hydra. <https://www2.razerzone.com/es-es/gaming-controllers/razer-hydra-portal-2-bundle>. accedido en junio 2017.
- [379] Htc vive accessories. <https://www.vive.com/us/accessory/>. accedido en junio 2017.
- [380] Oculus rift accessories. <https://www.oculus.com/accessories/>. accedido en junio 2017.
- [381] D Fernandez and A Price. Visual odometry for an outdoor mobile robot. In *IEEE Conference on Robotics, Automation and Mechatronics, 2004.*, volume 2, pages 816–821 vol.2. ieeexplore.ieee.org, December 2004.
- [382] Davide Scaramuzza and F Fraundorfer. Tutorial: visual odometry. *IEEE Robot. Autom. Mag.*, 18(4):80–92, 2011.
- [383] Mohammad O A Aqel, Mohammad H Marhaban, M Iqbal Saripan, and Napsiah Bt Ismail. Review of visual odometry: types, approaches, challenges, and applications. *Springerplus*, 5(1):1897, 28 October 2016.
- [384] William Rone and Pinhas Ben-Tzvi. Mapping, localization and motion planning in mobile multi-robotic systems. *Robotica*, 31(1):1–23, January 2013.
- [385] J Campbell, R Sukthankar, I Nourbakhsh, and A Pahwa. A robust visual odometry and precipice detection system using consumer-grade monocular vision. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, pages 3421–3427. ieeexplore.ieee.org, April 2005.
- [386] Keiji Nagatani, Ayako Ikeda, Genya Ishigami, Kazuya Yoshida, and Isaku Nagai. Development of a visual odometry system for a wheeled robot on loose soil using a telecentric camera. *Adv. Robot.*, 24(8-9):1149–1167, 1 January 2010.
- [387] Ramon Gonzalez, Francisco Rodriguez, Jose Luis Guzman, Cedric Pradaliere, and Roland Siegwart. Combined visual odometry and visual compass for off-road mobile robots localization. *Robotica*, 30(6):865–878, October 2012.

- [388] A Howard. Real-time stereo visual odometry for autonomous ground vehicles. In *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3946–3952. ieeexplore.ieee.org, 2008.
- [389] D Nister, O Naroditsky, and J Bergen. Visual odometry. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, volume 1, pages I–652–I–659 Vol.1. ieeexplore.ieee.org, June 2004.
- [390] N Nourani-Vatani, J Roberts, and M V Srinivasan. Practical visual odometry for car-like vehicles. In *2009 IEEE International Conference on Robotics and Automation*, pages 3551–3557. ieeexplore.ieee.org, May 2009.
- [391] K Ni and F Dellaert. Stereo tracking and Three-Point/One-Point algorithms - a robust approach in visual odometry. In *2006 International Conference on Image Processing*, pages 2777–2780. ieeexplore.ieee.org, October 2006.
- [392] Cailing Wang, Chunxia Zhao, and Jingyu Yang. Monocular odometry in country roads based on phase-derived optical flow and 4-DOF ego-motion model. *Industrial Robot: An International Journal*, 38(5):509–520, 2011.
- [393] Takeshi Takahashi. *2D localization of outdoor mobile robots using 3D laser range data*. PhD thesis, Doctoral dissertation, Carnegie Mellon University, 2007.
- [394] Aboelmagd Noureldin, Tashfeen B Karamat, and Jacques Georgy. *Fundamentals of Inertial Navigation, Satellite-based Positioning and their Integration*. Springer Science & Business Media, 28 October 2012.
- [395] Ramon Gonzalez, Francisco Rodriguez, Jose Luis Guzman, Cedric Pradalier, and Roland Siegwart. Control of off-road mobile robots using visual odometry and slip compensation. *Adv. Robot.*, 27(11):893–906, 1 August 2013.
- [396] Navid Nourani-Vatani and Paulo Vinicius Koerich Borges. Correlation-based visual odometry for ground vehicles. *Journal of Field Robotics*, 28(5):742–768, 2011.
- [397] Y Yu, C Pradalier, and G Zong. Appearance-based monocular visual odometry for ground vehicles. In *2011 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*, pages 862–867. ieeexplore.ieee.org, July 2011.

- [398] S B Knorr and D Kurz. Leveraging the user's face for absolute scale estimation in handheld monocular SLAM. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 11–17, 2016.
- [399] Inter sense 900. <http://www.intersense.com/pages/20/14>. accedido en junio 2017.
- [400] Y Yokokohji, Y Sugawara, and T Yoshikawa. Accurate image overlay on video see-through HMDs using vision and accelerometers. In *Proceedings IEEE Virtual Reality 2000 (Cat. No.00CB37048)*, pages 247–254. ieeexplore.ieee.org, 2000.
- [401] G Qian, R Chellappa, and Q Zheng. Robust structure from motion estimation using inertial data. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.*, 18(12):2982–2997, December 2001.
- [402] P Pinies, T Lupton, S Sukkariéh, and J D Tardos. Inertial aiding of inverse depth SLAM using a monocular camera. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 2797–2802. ieeexplore.ieee.org, April 2007.
- [403] G S W Klein and T W Drummond. Tightly integrated sensor fusion for robust visual tracking. *Image Vis. Comput.*, 22(10):769–776, 1 September 2004.
- [404] Michael Aron, Gilles Simon, and Marie-Odile Berger. Handling uncertain sensor data in Vision-Based camera tracking. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality, ISMAR '04*, pages 58–67, Washington, DC, USA, 2004. IEEE Computer Society.
- [405] Rokoko. <https://www.rokoko.com/en/>. accedido en junio 2017.
- [406] Myo. <https://www.myo.com/>. accedido en junio 2017.
- [407] Microsoft kinect. <https://developer.microsoft.com/en-us/windows/kinect>. accedido en abril 2017.
- [408] Leap motion. <https://leapmotion.com>. accedido en abril 2017.
- [409] Fove. <https://www.getfove.com/>. accedido en junio 2017.
- [410] Binary vr. <http://www.binaryvr.com/>. accedido en junio 2017.

- [411] William R Sherman and Alan B Craig. *Understanding virtual reality: Interface, application, and design*. Elsevier, 2002.
- [412] Jiawei Huang, Tsuyoshi Mori, Kazuki Takashima, Shuichiro Hashi, and Yoshifumi Kitamura. IM6D: magnetic tracking system with 6-DOF passive markers for dexterous 3D interaction and motion. *ACM Trans. Graph.*, 34(6):217, 4 November 2015.
- [413] M Singh and B Jung. High-definition wireless personal area tracking using AC magnetic field for virtual reality. In *2017 IEEE Virtual Reality (VR)*, pages 209–210, March 2017.
- [414] Nicholas T Swafford, Bastiaan J Boom, Kartic Subr, David Sinclair, Darren Cosker, and Kenny Mitchell. Dual sensor filtering for robust tracking of head-mounted displays. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pages 221–222. ACM, 11 November 2014.
- [415] Franco Tecchia, Giovanni Avveduto, Raffaello Brondi, Marcello Carozzino, Massimo Bergamasco, and Leila Alem. I’m in VR!: using your own hands in a fully immersive MR system. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pages 73–76. ACM, 11 November 2014.
- [416] Soonchan Park, Moonwook Ryu, Ju Yong Chang, and Jiyoung Park. A hand posture recognition system utilizing frequency difference of infrared light. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pages 65–68. ACM, 11 November 2014.
- [417] Pooya Rahimian and Joseph K Kearney. Optimal camera placement for motion capture systems in the presence of dynamic occlusion. In *Proceedings of the 21st ACM Symposium on Virtual Reality Software and Technology*, pages 129–138. ACM, 13 November 2015.
- [418] A Garcia, N Andre, D Bell Boucher, A Roberts-South, M Jog, and M Katchabaw. Immersive augmented reality for parkinson disease rehabilitation. In *Virtual, Augmented Reality and Serious Games for Healthcare 1*, Intelligent Systems Reference Library, pages 445–469. Springer, Berlin, Heidelberg, 2014.
- [419] L Carozza, F Bosché, and M Abdel-Wahab. [poster] visual-Inertial 6-DOF localization for a wearable immersive VR/AR system. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 257–258, 2014.

- [420] Steven Maesen, Patrik Goorts, and Philippe Bekaert. Scalable optical tracking for navigating large virtual environments using spatially encoded markers. In *Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology*, pages 101–110. ACM, 6 October 2013.
- [421] Andrew Ofstad, Emmett Nicholas, Rick Szcodronski, and Romit Roy Choudhury. AAMPL: accelerometer augmented mobile phone localization. In *Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*, pages 13–18. ACM, 19 September 2008.
- [422] L Carozza, F Bosché, and M Abdel-Wahab. Robust 6-DOF immersive navigation using commodity hardware. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology*, pages 19–22. ACM, 11 November 2014.
- [423] Mask. <https://www.mindmaze.com/mask/>. accedido en junio 2017.
- [424] Emotiv epoc. <https://www.emotiv.com/epoc/>. accedido en junio 2017.
- [425] A Lécuyer, F Lotte, R B Reilly, R Leeb, M Hirose, and M Slater. Brain-Computer interfaces, virtual reality, and videogames. *Computer*, 41(10):66–72, October 2008.
- [426] Johanne Tromp, David Peeters, Peter Hagoort, and Antje S Meyer. Combining EEG and virtual reality: The N400 in a virtual environment. In *the 4th edition of the Donders Discussions (DD, 2015)*, 2015.
- [427] Walter F Bischof and Pierre Boulanger. Spatial navigation in virtual reality environments: an EEG analysis. *Cyberpsychol. Behav.*, 6(5):487–495, October 2003.
- [428] Qibin Zhao, Liqing Zhang, and Andrzej Cichocki. EEG-based asynchronous BCI control of a car in 3D virtual reality environments. *Chin. Sci. Bull.*, 54(1):78–87, 1 January 2009.
- [429] Brian D Ferris, Dieter Fox, and Neil Lawrence. WiFi-SLAM using gaussian process latent variable models. 2007.
- [430] P Bahl and V N Padmanabhan. RADAR: an in-building RF-based user location and tracking system. In *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, volume 2, pages 775–784 vol.2. ieeexplore.ieee.org, 2000.

- [431] M Uddin and T Nadeem. SpyLoc: A light weight localization system for smartphones. In *2014 Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pages 72–80. ieeexplore.ieee.org, June 2014.
- [432] O Cruz, E Ramos, and M Ramírez. 3D indoor location and navigation system based on bluetooth. In *CONIELECOMP 2011, 21st International Conference on Electrical Communications and Computers*, pages 271–277. ieeexplore.ieee.org, February 2011.
- [433] S Satyavolu, G Bruder, P Willemsen, and F Steinicke. Analysis of IR-based virtual reality tracking using multiple kinects. In *2012 IEEE Virtual Reality Workshops (VRW)*, pages 149–150. ieeexplore.ieee.org, March 2012.
- [434] Janne Haverinen and Anssi Kemppainen. Global indoor self-localization based on the ambient magnetic field. *Rob. Auton. Syst.*, 57(10):1028–1035, 31 October 2009.
- [435] Q Yuan and I M Chen. 3-D localization of human based on an inertial capture system. *IEEE Trans. Rob.*, 29(3):806–812, June 2013.
- [436] Qilong Yuan, I-Ming Chen, and Shang Ping Lee. SLAC: 3D localization of human based on kinetic human movement capture. In *2011 IEEE International Conference on Robotics and Automation*, pages 848–853. ieeexplore.ieee.org, May 2011.
- [437] Yasuhiro Fukuju, Masateru Minami, Hiroyuki Morikawa, and Tomonori Aoyama. DOLPHIN: An autonomous indoor positioning system in ubiquitous computing environment. *WSTFES*, 3:53, 2003.
- [438] Mike Hazas and Andy Ward. A novel broadband ultrasonic location system. In *UbiComp 2002: Ubiquitous Computing*, Lecture Notes in Computer Science, pages 264–280. Springer, Berlin, Heidelberg, 29 September 2002.
- [439] Nissanka B Priyantha, Anit Chakraborty, and Hari Balakrishnan. The cricket location-support system. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, MobiCom '00, pages 32–43, New York, NY, USA, 2000. ACM.
- [440] A Ward, A Jones, and A Hopper. A new location technique for the active office. *IEEE Pers. Commun.*, 4(5):42–47, October 1997.
- [441] Rainer Mautz. *Indoor positioning technologies*. PhD thesis, 2012.

- [442] J Hightower and G Borriello. Location systems for ubiquitous computing. *Computer*, 34(8):57–66, August 2001.
- [443] Devesh Kumar Bhatnagar. Position trackers for head mounted display systems: A survey. *University of North Carolina, Chapel Hill TR93*, 10, 1993.
- [444] Eric Hodgson, Eric R Bachmann, David Vincent, Michael Zmuda, David Waller, and James Calusdian. WeaVR: a self-contained and wearable immersive virtual environment simulation system. *Behav. Res. Methods*, 47(1):296–307, March 2015.
- [445] Jongkyu Shin, Gwangseok An, Joon-Sang Park, Seung Jun Baek, and Kyogu Lee. Application of precise indoor position tracking to immersive virtual reality with translational movement support. *Multimed. Tools Appl.*, 75(20):12331–12350, 1 October 2016.
- [446] Prozyx. <https://www.pozyx.io/>. accedido en junio 2017.
- [447] C Pei, Y Cai, and Z Ma. An indoor positioning algorithm based on received signal strength of WLAN. In *2009 Pacific-Asia Conference on Circuits, Communications and Systems*, pages 516–519. ieeexplore.ieee.org, May 2009.
- [448] W Meng, W Xiao, W Ni, and L Xie. Secure and robust Wi-Fi fingerprinting indoor localization. In *2011 International Conference on Indoor Positioning and Indoor Navigation*, pages 1–7. ieeexplore.ieee.org, 2011.
- [449] N Marques, F Meneses, and A Moreira. Combining similarity functions and majority rules for multi-building, multi-floor, WiFi positioning. In *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–9. ieeexplore.ieee.org, November 2012.
- [450] M Gunawan, B Li, T Gallagher, A G Dempster, and G Retscher. A new method to generate and maintain a WiFi fingerprinting database automatically by using RFID. In *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–6. ieeexplore.ieee.org, November 2012.
- [451] D R Brown and D B Dunn. Classification schemes of positioning technologies for indoor navigation. In *2011 Proceedings of IEEE Southeastcon*, pages 125–130. ieeexplore.ieee.org, March 2011.

- [452] Sebastian Fudickar, Sebastian Amend, and Bettina Schnor. On the comparability of indoor localization systems' accuracy. In *Proceedings of the Fifth ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*, pages 21–28. ACM, 5 November 2013.
- [453] Ahmed Alnabhan and Brian Tomaszewski. INSAR: indoor navigation system using augmented reality. In *Proceedings of the Sixth ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*, pages 36–43. ACM, 4 November 2014.
- [454] Song Gao and Sathya Prasad. Employing spatial analysis in indoor positioning and tracking using wi-fi access points. In *Proceedings of the Eighth ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*, pages 27–34. ACM, 31 October 2016.
- [455] Philipp Bolliger. Redpin - adaptive, zero-configuration indoor localization through user collaboration. In *Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*, pages 55–60. ACM, 19 September 2008.
- [456] Mortaza S Bargh and Robert de Groot. Indoor localization based on response rate of bluetooth inquiries. In *Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*, pages 49–54. ACM, 19 September 2008.
- [457] Pedro Figueiredo e Silva and Elena Simona Lohan. Room-level indoor positioning with Wi-Fi and RFID fingerprints. In *Proceedings of the 5th International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments*, page 5. ACM, 3 November 2015.
- [458] Satoshi Hyuga, Masaki Ito, Masayuki Iwai, and Kaoru Sezaki. Estimate a user's location using smartphone's barometer on a subway. In *Proceedings of the 5th International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments*, page 2. ACM, 3 November 2015.
- [459] K Martin Sagayam and D Jude Hemanth. Hand posture and gesture recognition techniques for virtual reality applications: a survey. *Virtual Real.*, 21(2):91–107, 1 June 2017.
- [460] H Ishiyama and S Kurabayashi. Monochrome glove: A robust real-time hand gesture recognition method by using a fabric glove with design of structured markers. In *2016 IEEE Virtual Reality (VR)*, pages 187–188, March 2016.

- [461] Scott W Greenwald, Luke Loreti, Markus Funk, Ronen Zilberman, and Pattie Maes. Eye gaze tracking with google cardboard using purkinje images. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, pages 19–22. ACM, 2 November 2016.
- [462] H Iwata and T Fujii. Virtual perambulator: a novel interface device for locomotion in virtual environment. In *Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium*, pages 60–65, 265. ieeexplore.ieee.org, March 1996.
- [463] Virtuix omni. <http://www.virtuix.com/>. accedido en junio 2017.
- [464] Matías Nicolás Selzer and Martín Leonardo Larrea. AnArU, a virtual reality framework for physical human interactions. In *XXI Congreso Argentino de Ciencias de la Computación (Junín, 2015)*, 2015.
- [465] M Wells, B Peterson, and Jason Aten. The virtual motion controller: A sufficient-motion walking simulator. In *Proceedings of VRAIS*, volume 97, pages 1–8. pdfs.semanticscholar.org, 1996.
- [466] Joseph Paradiso, Craig Ablner, Kai-Yuh Hsiao, and Matthew Reynolds. The magic carpet: Physical sensing for immersive environments. In *CHI '97 Extended Abstracts on Human Factors in Computing Systems, CHI EA '97*, pages 277–278, New York, NY, USA, 1997. ACM.
- [467] Rudolph P Darken, William R Cockayne, and David Carmein. The omni-directional treadmill: A locomotion device for virtual worlds. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology, UIST '97*, pages 213–221, New York, NY, USA, 1997. ACM.
- [468] Haruo Noma. Design for locomotion interface in a large scale virtual environment. *ATLAS: ATR Locomotion Interface for Active Self Motion*, 64:111–118, 1998.
- [469] H Iwata. The torus treadmill: realizing locomotion in VEs. *IEEE Comput. Graph. Appl.*, 19(6):30–35, November 1999.
- [470] H Noma, T Sugihara, and T Miyasato. Development of ground surface simulator for Tel-E-Merge system. In *Proceedings IEEE Virtual Reality 2000 (Cat. No.00CB37048)*, pages 217–224. ieeexplore.ieee.org, 2000.
- [471] H Iwata, H Yano, and F Nakaizumi. Gait master: a versatile locomotion interface for uneven virtual terrain. In *Proceedings IEEE Virtual Reality 2001*, pages 131–137. ieeexplore.ieee.org, March 2001.

- [472] Jiung-Yao Huang. An omnidirectional stroll-based virtual reality interface and its application on overhead crane training. *IEEE Trans. Multimedia*, 5(1):39–51, March 2003.
- [473] Kiran J Fernandes, Vinesh Raja, and Julian Eyre. Cybersphere: The fully immersive spherical projection system. *Commun. ACM*, 46(9):141–146, September 2003.
- [474] Martin Schwaiger, Heinz Ulbrich, and Thomas Thümmel. A foot following locomotion device with force feedback capabilities. In *Proceedings of VIII Symposium on virtual Reality*, pages 309–321. lbd.dcc.ufmg.br, 2006.
- [475] H Iwata, H Yano, H Fukushima, and H Noma. CirculaFloor [locomotion interface]. *IEEE Comput. Graph. Appl.*, 25(1):64–67, January 2005.
- [476] Hiroo Iwata, Hiroaki Yano, and Hiroshi Tomioka. Powered shoes. In *ACM SIGGRAPH 2006 Emerging Technologies*, SIGGRAPH '06, New York, NY, USA, 2006. ACM.
- [477] Hiroo Iwata, Hiroaki Yano, and Masaki Tomiyoshi. String walker. In *ACM SIGGRAPH 2007 Emerging Technologies*, SIGGRAPH '07, New York, NY, USA, 2007. ACM.
- [478] Martin Schwaiger, Thomas Thümmel, and Heinz Ulbrich. Cyberwalk: Implementation of a ball bearing platform for humans. In *Human-Computer Interaction. Interaction Platforms and Techniques*, Lecture Notes in Computer Science, pages 926–935. Springer, Berlin, Heidelberg, 22 July 2007.
- [479] M C Schwaiger, T Thummel, and H Ulbrich. A 2D-Motion platform: The cybercarpet. In *Second Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems (WHC'07)*, pages 415–420. ieeexplore.ieee.org, March 2007.
- [480] Eliana Medina, Ruth Fruland, and Suzanne Weghorst. Virtosphere: Walking in a human size VR “hamster ball”. *Proc. Hum. Fact. Ergon. Soc. Annu. Meet.*, 52(27):2102–2106, 1 September 2008.
- [481] D Swapp, J Williams, and A Steed. The implementation of a novel walking interface within an immersive display. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 71–74. ieeexplore.ieee.org, March 2010.

- [482] Y Visell, A Law, J Ip, S Smith, and J R Cooperstock. Interaction capture in immersive virtual environments via an intelligent floor surface. In *2010 IEEE Virtual Reality Conference (VR)*, pages 313–314. ieeexplore.ieee.org, March 2010.
- [483] Minghadi Suryajaya, Tim Lambert, Chris Fowler, Phillip Stothard, David Laurence, and Chris Daly. OmniWalker: omnidirectional stroller-based walking platform. In *Proc. Virt. Real. Int. Conf., VRIC*, volume 10, pages 181–182. pdfs.semanticscholar.org, 2010.
- [484] Léo Terziman, Maud Marchal, Mathieu Emily, Franck Multon, Bruno Arnaldi, and Anatole Lécuyer. Shake-your-head: Revisiting walking-in-place for desktop virtual reality. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology, VRST '10*, pages 27–34, New York, NY, USA, 2010. ACM.
- [485] M Marchal, J Pettré, and A Lécuyer. Joyman: A human-scale joystick for navigating in virtual worlds. In *2011 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 19–26. ieeexplore.ieee.org, March 2011.
- [486] Jan L Souman, P Robuffo Giordano, Martin Schwaiger, Ilja Frissen, Thomas Thümmel, Heinz Ulbrich, A De Luca, Heinrich H Bühlhoff, and Marc O Ernst. CyberWalk: Enabling unconstrained omnidirectional walking through virtual environments. *ACM Trans. Appl. Percept.*, 8(4):25, 2011.
- [487] Tuncay Cakmak and Holger Hager. Cyberith virtualizer: a locomotion device for virtual reality. In *ACM SIGGRAPH 2014 Emerging Technologies*, page 6. dis.uniroma1.it, 2014.
- [488] Warren Couvillion. The pressure mat; a new device for traversing virtual environments using natural motion. In *Proc. Interservice/Industry Training Simulation and Education Conference, 2001*, pages 199–211. ci.nii.ac.jp, 2001.
- [489] J Feasel, M C Whitton, and J D Wendt. LLCM-WIP: Low-Latency, Continuous-Motion Walking-in-Place. In *2008 IEEE Symposium on 3D User Interfaces*, pages 97–104. ieeexplore.ieee.org, March 2008.
- [490] Laroussi Bouguila, Florian Evequoz, Michele Courant, and Beat Hirsbrunner. Walking-pad: A step-in-place locomotion interface for virtual environments. In *Proceedings of the 6th International Conference on Multimodal Interfaces, ICMI '04*, pages 77–81, New York, NY, USA, 2004. ACM.

- [491] Matías Nicolás Selzer and Martín Leonardo Larrea. AnArU, a virtual reality framework for physical human interactions. *J. Comput. Sci. Technol.*, 15, 2015.
- [492] Matías Nicolás Selzer. Interacción humano computadora mediante interfaces hápticas en RV. In *Congreso Internacional de Ciencia y Tecnología*, volume 3, 2016.
- [493] Matías Nicolás Selzer, Juan A. Biondi, Gamboa Florencia, and Escarza Sebastián. Braille messages in a haptic wearable device for visually impaired people. In *Congreso Argentino de Sistemas Embebidos, CASE2016*, volume 7, pages 84–87. Facultad de Ingeniería, Universidad de Buenos Aires, Buenos Aires, Argentina, 2016.
- [494] Kevin Soulier, Matías Nicolás Selzer, and Martín Leonardo Larrea. Real-Time estimation of illumination direction for augmented reality with Low-Cost sensors. In *XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016)*, 2016.
- [495] Matías Nicolás Selzer, Elisabet Arriata, Leonardo Segovia, Nicolás Fernando Gazcón, and Martín Leonardo Larrea. Modelos de interacción y aplicaciones en realidad virtual mediante dispositivos móviles. In *XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires)*, August 2017.
- [496] Kevin E Soulier, Matías Nicolás Selzer, and Martín Leonardo Larrea. Estimating illumination direction for augmented reality in Real-Time by using Low-Cost sensors. *J. Comput. Sci. Technol.*, 17, 2017.