

# Capítulo 2

## Introducción a la Realidad Virtual

Empezaremos por un apartado donde se explican los tres ingredientes elementales que tiene que tener un sistema de realidad virtual para ser considerado como tal: simulación interactiva, interacción implícita e inmersión sensorial.

### 2.1. Definición y conceptos básicos

A pesar de que en la literatura se pueden encontrar muchas definiciones de realidad virtual (en inglés, *virtual reality VR*), seguramente una de las más completas es la que propuso A. Rowell:

*“La Realidad Virtual es una simulación interactiva por computador desde el punto de vista del participante, en la cual se sustituye o se aumenta la información sensorial que recibe”.*

En esta definición aparecen los elementos básicos que tienen que estar presentes en todo sistema de realidad virtual y que se discuten a continuación:

- Simulación interactiva
- Interacción implícita
- Inmersión sensorial

#### 2.1.1. Simulación interactiva

Una aplicación de realidad virtual es una simulación en el sentido de que se recrea un mundo virtual que sólo existe como una representación digital

en la memoria de un ordenador. El hecho de que la simulación sea interactiva es lo que distingue la realidad virtual de una animación. En una animación, al igual que en el cine, los espectadores son individuos pasivos en el sentido que no pueden alterar el contenido de las imágenes que ven: éstas han sido grabadas previamente desde un determinado punto de vista o siguiendo una determinada trayectoria. En cambio, en un sistema de realidad virtual, el usuario puede escoger libremente su movimiento por la escena y, por tanto, sus acciones afectan de forma directa a las imágenes que verá. Además, el usuario puede improvisar su movimiento sobre la marcha, sin necesidad de tener que establecer un guión previo, pues el sistema de realidad virtual responderá en tiempo real (es decir, con un tiempo de respuesta despreciable) a sus acciones.

Dado que la trayectoria que seguirá el participante es imprevisible, y que hay infinitos puntos de vista diferentes, los sistemas de realidad virtual requieren tener una representación geométrica 3D del entorno virtual, de forma que puedan calcular la imagen correspondiente a un punto de vista arbitrario. Además de este modelo geométrico, los sistemas de realidad virtual también requieren utilizar algoritmos de visualización realista (algoritmos de síntesis de imágenes) para poder generar las imágenes a partir de estas representaciones digitales de la escena.

### **2.1.2. Interacción implícita**

La realidad virtual utiliza la interacción implícita en contraposición a la interacción explícita o interacción clásica.

En la interacción clásica, cuando un usuario quiere llevar a cabo una determinada acción (pongamos que quiere ver la parte del mundo virtual que tiene detrás) tiene que comunicar de forma explícita su voluntad al computador. Para esta tarea, el usuario utiliza el esquema de comunicación determinado por la interfaz de la aplicación, ya sea una interfaz basada en comandos o una interfaz gráfica de tipo W. I. M. P. (*window-icon-menu-pointing device*), que es como se denominan los sistemas basados en ventanas e íconos. En cualquiera de estos casos, el usuario tiene que realizar un pequeño esfuerzo, para recordar o buscar el comando o el elemento gráfico asociado a la acción que quiere realizar, y debe hacer llegar al ordenador esta orden a través de los dispositivos de interacción clásica, fundamentalmente el teclado y el ratón.

A pesar que de en los últimos años las interfaces gráficas en mejorado mucho la comunicación del hombre con el computador, la comunicación de

acciones repetidas (como la trayectoria que quiere seguir el usuario) no deja de suponer un esfuerzo, y normalmente requiere de un cierto tiempo de entrenamiento o formación.

En cambio, en la realidad virtual el sistema captura la voluntad del usuario implícita en sus movimientos naturales. El ejemplo más claro es el control de la cámara virtual: en una sistema de realidad virtual, la cámara se actualiza en función de los movimientos de la cabeza del usuario. Si el usuario quiere ver la parte de mundo virtual que tiene detrás, no tiene que utilizar ningún comando ni mover el ratón, sino que simplemente debe hacer el mismo gesto natural (girar la cabeza) que haría en el mundo real:

Otro ejemplo es la interacción con los objetos de la escena. Si el usuario quiere abrir una puerta o mover un objeto, lo único que tiene que hacer es coger la puerta o el objeto y llevar a cabo con la mano los mismos movimientos que haría con un objeto real.

Como podemos ver, tanto en interacción clásica como en interacción implícita se utilizan dispositivos periféricos para la entrada de datos, pero la diferencia fundamental es la percepción que tiene el usuario de estos dispositivos. La revolución de la realidad virtual en la interacción hombre-máquina radica en que el usuario deja de percibir los dispositivos (incluso deja de percibir el propio ordenador) para pasar a interactuar directamente con los objetos de la escena. La diferencia psicológica es que el usuario deja de mirar por una ventana y pasa a estar dentro del entorno virtual.

### **2.1.3. Inmersión sensorial**

El tercer elemento fundamental de todo sistema de realidad virtual es la inmersión sensorial.

La palabra inmersión tiene significados muy variados según el contexto en que se utilice, pero en realidad virtual es un concepto muy claro. Podemos definir la inmersión sensorial como la desconexión de los sentidos del mundo real y la conexión al mundo virtual. Como consecuencia, el usuario deja de percibir el entorno que le rodea y pasa a estar inmerso dentro del mundo virtual que recrea el computador en tiempo real. De los diferentes órganos de los sentidos, la vista es el que nos proporciona una mayor cantidad de información y a la vez es la que da una mayor sensación de presencia. Por este motivo, todo sistema de realidad virtual debe proporcionar estímulos adecuados como mínimo para el sentido de la vista (es decir, debe generar

las imágenes correspondientes al mundo virtual), a pesar de que a menudo los sistemas de realidad virtual proporcionan también inmersión acústica.

A menudo se dice que un sistema de realidad virtual crea el efecto de que los objetos del entorno virtual existen con independencia del dispositivo de visualización. Cuando miramos un modelo en una pantalla de un computador (sin realidad virtual), claramente distinguimos que los objetos están proyectados en la superficie de la pantalla, y no tenemos ninguna sensación que estos objetos tengan una existencia material, por más realista que sean las imágenes.

En cambio, la utilización de los dispositivos de visualización de realidad virtual crea en el usuario el efecto que los objetos no están proyectados en ninguna superficie, sino que se encuentran a diferentes distancias “flotando” en el espacio que envuelve al observador, como si tuviesen existencia material propia. La clave de este proceso es la visión estereoscópica. La visión estereoscópica se basa en proporcionar dos imágenes ligeramente diferentes del mundo virtual, una por cada ojo, de forma que el sistema visual humano deduce la profundidad de los objetos a partir de las diferencias en las imágenes. La visión estereoscópica es un elemento imprescindible en todo sistema de realidad virtual, pues es el único medio para hacer que los objetos tengan una fuerte presencia espacial. La visión estereoscópica también es una herramienta muy útil para aplicaciones que no se pueden considerar de realidad virtual, ya que permite ubicar espacialmente los objetos de forma mucho más precisa, gracias a que la escena se percibe en relieve.

## 2.2. Arquitectura de un sistema de realidad virtual

En este apartado veremos cuales son los diferentes elementos hardware y software que componen un sistema de realidad virtual, cual es la interrelación entre estos elementos, y como estos elementos afectan a los niveles de simulación, interacción e inmersión del sistema.

### 2.2.1. Componentes de un sistema de realidad virtual

En un sistema de realidad virtual se pueden distinguir elementos hardware y elementos software.

Los componentes hardware más importantes son el computador, los periféricos de entrada y los periféricos de salida.

Los componentes software más importantes son el modelo geométrico 3D y los programas de simulación sensorial (simulación visual, auditiva, táctil, . . .), simulación física (movimiento de la cámara virtual, detección de colisiones, cálculo de deformaciones, . . .), y recogida de datos. La siguiente figura ilustra los componentes de un sistema típico de realidad virtual:

A continuación describiremos brevemente cada uno de estos componentes:

#### **Periféricos de entrada (sensores)**

Los periféricos de entrada se encargan de capturar las acciones del participante y enviar esta información al computador. Los periféricos de entrada más frecuentes en realidad virtual son los posicionadores (que permiten al sistema conocer en tiempo real la posición y la orientación de la cabeza, de la mano, o de todo el cuerpo del usuario), los guantes (que permiten detectar movimiento de los dedos de la mano) y los micrófonos (que graban la voz del participante).

#### **Periféricos de salida (efectores)**

Los periféricos de salida se encargan de traducir las señales de audio, video, etc. generados por el computador en estímulos para los órganos de los sentidos (sonido, imágenes, . . .). Los efectores se clasifican según el sentido al que va dirigido: existen efectores visuales (cascos estereoscópicos, pantallas de proyección, . . .), y de audio (sistemas de sonido, altavoces, . . .) de fuerza y tacto (dispositivos táctiles), y del sentido del equilibrio (plataformas móviles).

## **Computador**

El computador se encarga de llevar a cabo la simulación de forma interactiva, basándose en el modelo geométrico 3D y en el software de recogida de datos, simulación física y simulación sensorial. Debido a que el proceso más crítico en realidad virtual es la simulación visual (síntesis de imágenes a partir de modelos 3D), los computadores que se utilizan para realidad virtual son estaciones de trabajo con prestaciones gráficas avanzadas, donde la mayor parte de las etapas del proceso de visualización están implementadas por hardware.

## **Modelo geométrico 3D**

Dado que un sistema de realidad virtual tiene que permitir explorar la escena de forma interactiva y ver el mundo virtual desde cualquier punto de vista, es necesario disponer de una representación geométrica 3D de este mundo, que permita hacer los cálculos de imágenes, generación de sonido espacial, cálculo de colisiones, etc. a los módulos que describiremos más adelante.

## **Software de tratamiento de datos de entrada**

Los módulos de recogida y tratamiento de datos se encargan de leer y procesar la información que proporcionan los sensores. Esto incluye los controladores de los dispositivos físicos, así como los módulos para el primer tratamiento de los datos suministrados. Por ejemplo, los datos de posición y orientación de la cabeza del usuario normalmente se tienen que transformar para expresarlas en un sistema de coordenadas de la aplicación y se deben filtrar para evitar saltos repentinos como consecuencia de lecturas erróneas de los valores de posición. Los sistemas que permiten la comunicación con el ordenador mediante órdenes orales requieren un sistema de reconocimiento de voz. Otros sistemas utilizan un esquema de comunicación basado en gestos de la mano (una especie de lenguaje de sordo-mudos pero más sencillo) y que requiere el reconocimiento de gestos a partir de una secuencia de movimientos.

## **Software de simulación física**

Los módulos de simulación física se encargan de llevar a cabo las modificaciones pertinentes en la representación digital de la escena, a partir de las acciones del usuario y de la evolución interna del sistema. Por ejemplo, si el módulo de recogida de datos indica que el usuario tiene que hacer el gesto

correspondiente a abrir una puerta, el sistema debe aplicar la transformación geométrica correspondiente al objeto del modelo 3D que representa esta puerta. Estos módulos varían mucho dependiendo de la aplicación concreta. La función más básica consiste en calcular en tiempo real los parámetros de la cámara virtual de acuerdo con los movimientos del usuario, aunque también puede encargarse del cálculo de colisiones, deformaciones, comportamiento y otras actualizaciones que afecten a la evolución en el tiempo del entorno virtual representado.

### **Software de simulación sensorial**

Estos módulos se encargan de calcular la representación digital de las imágenes, sonidos, etc. que el hardware se encargará de traducir a señales y finalmente a estímulos para los sentidos. Entre los módulos de simulación sensorial, lo más importante es el de simulación visual, que se basa en algoritmos de visualización en tiempo real del modelo geométrico. Los algoritmos de visualización que se utilizan en realidad virtual son parecidos a los que se han descrito en los capítulos anteriores pero, dado que el rendimiento es crítico, se utilizan técnicas de aceleración de imagen con el propósito de reducir al mínimo posible el tiempo de generación de cada fotograma. Respecto a la simulación auditiva, es preciso comentar que la generación de sonido realista requiere tener en cuenta las propiedades acústicas de los objetos y que los algoritmos son tan complicados como los algoritmos de visualización. Respecto a la simulación táctil, es necesario distinguir entre los dispositivos que proporcionan sensación de tacto (a menudo limitado a la mano), sensación de contacto (también limitado a la mano) y realimentación de fuerza (impiden u ofrecen resistencia a hacer movimientos con la mano cuando ésta choca virtualmente con un objeto virtual). En cualquier caso, es imprescindible que el sistema sea capaz de detectar en tiempo real las colisiones que se puedan producir entre la mano del usuario y los objetos de la escena, ya que es este evento el que activa los dispositivos hardware apropiados.

### **2.2.2. Funcionamiento de un sistema de realidad virtual**

A continuación se describe el funcionamiento de un sistema típico de realidad virtual, en función de los elementos hardware y software descritos anteriormente.

Las acciones del usuario son registradas en tiempo real por los periféricos de entrada o sensores. Un requerimiento básico es capturar la posición de la cabeza del usuario (según el sistema de visualización bastará conocer la

posición de la cabeza o necesitaremos también su orientación en el espacio), aunque también es frecuente capturar movimientos de la mano y de los dedos de la mano. Esta información se envía al computador, donde es procesada por los módulos de recogida y tratamiento de datos de entrada. Una vez las acciones del usuario están en un formato adecuado, se utilizan en la parte de simulación, tanto para introducir posibles cambios en el mundo virtual (por ejemplo, el usuario ha movido una silla y este cambio se tiene que reflejar en el modelo geométrico) como para generar imágenes, sonidos y otros datos para completar la simulación sensorial. En el caso de la simulación visual, los algoritmos de síntesis de imágenes calculan la imagen en el *frame buffer* con los parámetros de cámara virtual correspondientes a la posición y orientación de la cabeza del usuario. La información digital del *frame buffer* se traduce en una señal de video por la tarjeta gráfica y esta señal se traduce en una imagen física en el dispositivo de visualización. Lo mismo pasa con los estímulos para el resto de sentidos, que son calculados por los algoritmos de simulación, traducidos en señales digitales para el hardware y finalmente convertidos en estímulos físicos para los dispositivos de salida.

Las etapas descritas anteriormente completan un círculo que se debe repetir continuamente lo más rápido posible como para que el usuario sea incapaz de percibir el retardo. En este sentido, es necesario distinguir entre las frecuencias de actualización (frecuencia con la que los sensores envían los datos de posición, frecuencia con la que se repintan las imágenes . . .) y el tiempo de latencia (que es el tiempo que pasa desde que se produce un evento, como por ejemplo un movimiento de la cabeza, hasta que tiene respuesta por parte del sistema y se visualiza la imagen con el nuevo punto de vista). Aunque están muy relacionadas, es preciso tener en cuenta la diferencia entre ambas magnitudes, como pone de manifiesto la siguiente comparación. Un sistema con una frecuencia de actualización muy elevada, como son 60 Hz, pero un tiempo de latencia también muy grande, por ejemplo 4 segundos, será capaz de detectar movimientos rápidos del usuario (pues se toman muestras de su posición cada 1/60 de segundo), pero desde que el usuario hace un movimiento de la cabeza hasta que recibe la imagen correspondiente pasa bastante tiempo como para que se pueda hablar de interactividad. En cambio, un sistema con una frecuencia de actualización de 10 Hz pero una latencia de 0.2 segundos, será más interactivo, aunque las imágenes se percibirán a saltos y no como una secuencia continua.

### 2.2.3. Niveles de simulación, interacción e inmersión

En este apartado veremos cual es la incidencia de los diferentes elementos de la arquitectura de un sistema de realidad virtual en los tres componentes básicos que tienen que estar presentes en un sistema para que los podamos considerar de realidad virtual (simulación interactiva, interacción implícita e inmersión sensorial), y que condicionan el grado de realidad virtual de un sistema.

- Factores que intervienen en la simulación interactiva
- Factores que intervienen en la interacción implícita
- Factores que intervienen en la inmersión sensorial

#### Factores que intervienen en la simulación interactiva

El nivel de simulación depende básicamente del abanico de leyes físicas que se simulan en el mundo virtual. Aunque normalmente los sistemas de realidad virtual se limitan a utilizar únicamente modelos de iluminación empíricos para poder simular leyes básicas de la óptica y así proporcionar una imagen más o menos realista de la escena, algunos sistemas pueden ir más allá y simular los siguientes fenómenos:

- Cinética y dinámica  
Algunos sistemas permiten simular las fuerzas de gravedad y las leyes básicas de movimiento de los objetos, como la inercia. Por ejemplo, si cogemos un objeto virtual y lo soltamos, el objeto cae por efecto de la gravedad o permanece inmóvil en la misma posición dependiendo del grado de simulación del sistema. Lo mismo pasa si empujamos un objeto virtual, que puede continuar en movimiento por efecto de la inercia o permanecer inmóvil.
- Deformaciones  
La mayoría de sistemas consideran que los objetos son sólidos rígidos y por tanto indeformables. Pero algunos sistemas, calculan la deformación de los objetos según las fuerzas que se le aplican, prestación muy frecuente en los sistemas CAE (*Computer-Aided Engineering*).
- Detección de colisiones  
Algunos sistemas permiten detectar colisiones entre objetos y evitar así que haya dos objetos virtuales en la misma posición del espacio. La detección de colisiones también se utiliza en realidad virtual para interactuar directamente con los objetos de la escena.

- Simulación del funcionamiento de vehículos  
En algunas aplicaciones de realidad virtual, la simulación del funcionamiento de un vehículo es el elemento más importante. Esto pasa en los simuladores de vuelo, simuladores de conducción, etc.
- Simulación del comportamiento de un sistema físico  
Otras aplicaciones permiten simular el comportamiento de un sistema para estudiar fenómenos como por ejemplo dinámica de partículas o el funcionamiento de una cadena de montaje. Dado que la complejidad de los cálculos que se pueden llevar a cabo en tiempo real es muy limitada, a menudo estos cálculos se hacen en una etapa de preproceso, y el sistema de realidad virtual se limita a visualizar estos resultados para facilitar su interpretación.

Además de estos y otros fenómenos, los sistemas de realidad virtual pueden simular en mayor o menor medida las leyes ópticas y acústicas. Dado que en este caso la simulación tiene una incidencia directa en la percepción, estos factores los estudiaremos en el apartado de inmersión.

El nivel de interactividad depende básicamente del abanico de acciones que el usuario puede llevar a cabo en el mundo virtual y que tienen respuesta por parte del sistema. El nivel de interacción básico de un sistema de realidad virtual tiene que permitir al usuario moverse libremente por la escena y observar los objetos desde el punto de vista que elija. Algunos sistemas ofrecen un grado de interactividad más elevado y permiten que el usuario pueda interactuar con los objetos de la escena: el usuario puede abrir puertas, coger y mover objetos, accionar dispositivos virtuales, y hasta llegar a pilotar un avión o conducir un coche.

Es necesario considerar que si el tiempo de respuesta del sistema a las acciones del usuario es lo suficientemente grande como para que el usuario pueda percibir este retardo, el nivel de interactividad se ve afectado negativamente.

### **Factores que intervienen en la interacción implícita**

Como hemos dicho anteriormente, la interacción implícita se basa en capturar la voluntad del usuario que está implícita en sus movimientos naturales. Con el uso de dispositivos de realidad virtual adecuados, como sensores y guantes, podemos capturar tanto el movimiento del usuario con sus acciones sobre los objetos de la escena. Asimismo, todo sistema de realidad virtual

tiene que proporcionar soporte a una serie de acciones que no tienen un equivalente en forma de gesto natural. Por ejemplo, abrir el cajón de una mesa virtual es una acción que se puede capturar a partir de un movimiento de la mano del usuario, pero cargar otro modelo 3D o seleccionar otros parámetros de visualización son acciones que no tienen una traducción directa en un gesto natural y por tanto requieren un esfuerzo adicional por parte del usuario.

Asimismo, hasta para llevar a cabo estas últimas acciones podemos encontrar diferentes niveles de interacción. Algunos sistemas utilizan reconocimiento de voz como medio para comunicar órdenes al computador, en sustitución del ratón y del teclado. Otros sistemas utilizan dispositivos de entrada 3D para seleccionar opciones en un menú virtual que forma parte de la escena como el resto de objetos. Otros sistemas usan un lenguaje de signos con la mano para indicar las acciones más frecuentes. Por ejemplo, señalar un punto con el dedo índice de la mano puede significar “avanzar en la dirección señalada”.

En cualquier caso, es importante remarcar que aunque las posibilidades de interacción de un sistema de realidad virtual se limiten a que el usuario pueda elegir libremente su trayectoria, esta prestación ya es bastante importante para justificar el uso de estas tecnologías, pues en muchas aplicaciones como la revisión de diseños de CAD (*Computer Aided Design*), una parte considerable del tiempo que invierte el usuario esta dedicada a la navegación por la escena. En este sentido, algunas aplicaciones de realidad virtual utilizan la expresión tiempo de aprendizaje cero para indicar que la mayor parte de la interacción se basa en movimientos naturales que no requieren un período de instrucción previo.

### **Factores que intervienen en la inmersión sensorial**

El grado de inmersión sensorial depende, en primera instancia, de cuales son los órganos de los sentidos para los cuales el sistema proporciona estímulos adecuados y, en segunda instancia, del alcance, calidad, velocidad y coherencia de estos estímulos.

Los humanos tenemos cinco sentidos: la vista, el oído, el tacto, el olfato y el gusto. A esta lista podemos añadir el sentido de equilibrio, que permite conocer la orientación del cuerpo respecto a la fuerza de la gravedad, y la cinestesia, que es la sensación por la cual uno percibe el movimiento muscular, el peso y la posición de los miembros del cuerpo.

De estos sentidos, los más utilizados en realidad virtual son la vista y el oído, y en un segundo nivel, el tacto, el equilibrio y la cinestesia. Los dos últimos sentidos son especialmente importantes en los simuladores de conducción de vehículos, para poder percibir las fuerzas de aceleración. Existen algunos prototipos experimentales que sintetizan olores, pero su utilidad práctica no está clara y todavía no existen versiones comerciales.

Como acabamos de ver, cuanto más grande sea el número de sentidos estimulados por el sistema de realidad virtual, más alto será el grado de inmersión. La inmersión también será mayor cuando el sistema nos desconecta totalmente (es decir, nos aísla) del mundo real, que cuando lo hace sólo parcialmente. Por ejemplo, respecto al sentido de la vista, hay sistemas inmersivos (basados en cascos estereoscópicos con pantallas situadas muy cerca de los ojos) que impiden la visión del entorno que nos rodea y nos presentan únicamente las imágenes del mundo virtual, y hay también sistemas semi-inmersivos (basados en proyección en una o más pantallas) en los cuales sólo percibimos el entorno real que hay entre el ojo del observador y la pantalla.

Por otro lado, aspectos como la calidad del estímulo (calidad de las imágenes, calidad del sonido), la velocidad de actualización y tiempo de latencia (tiempo que pasa desde que el participante lleva a cabo una acción, por ejemplo girar la cabeza, hasta que percibe el resultado) y la coherencia (si por ejemplo se cierra una puerta en el mundo virtual, tenemos que ver como se cierra en el mismo instante que oímos el sonido del golpe al cerrar) también determinan el grado de inmersión sensorial.

La inmersión total no siempre es una cualidad deseable. Por ejemplo, los usuarios de sistemas inmersivos (cascos estereoscópicos) pierden la visión de su propio cuerpo y esto provoca muy a menudo un cierto grado de desorientación. Otro ejemplo donde los sistemas inmersivos no son adecuados son las aplicaciones que requieren trabajar en grupo (por ejemplo, la revisión de un modelo de CAD por parte de un equipo de especialistas), ya que los sistemas inmersivos limitan mucho la interacción con los otros miembros del grupo. Para estos casos sistemas semi-inmersivos como la CAVE.

En algunos casos, la finalidad del sistema no es sustituir la información sensorial que recibe el participante, sino sólo aumentarla, siendo necesario mantener la percepción del mundo real. Un ejemplo es el sistema que llevan los pilotos de algunos aviones de combate basado en gafas especiales que les permiten ver información crítica superpuesta con las imágenes del entorno. Otro ejemplo es los sistemas que se utilizan en medicina para indi-

car al médico la posición exacta de un punto concreto de la anatomía de un paciente al que previamente se ha aplicado una exploración para reconstruir sus estructuras internas. Esta técnica, cuyo objetivo es solamente ampliar la percepción del mundo real con información crítica generada por computador, se denomina *realidad aumentada*, y tiene muchas aplicaciones en el campo de entrenamiento y la medicina.

## 2.3. Periféricos de realidad virtual

En este apartado estudiaremos los componentes más específicos de un sistema de realidad virtual: los dispositivos periféricos.

Los periféricos de un sistema de realidad virtual se clasifican según el sentido de la información entre el participante y la máquina.

Los dispositivos de entrada, también denominados **sensores**, capturan las acciones del participante (por ejemplo movimientos de la cabeza) y envían esta información al computador encargado de llevar a cabo la simulación.

Los dispositivos de salida, también denominados **efectores**, generan los estímulos necesarios para los sentidos del participante, traduciendo en imágenes, sonidos, etc. las señales de video, audio, etc. que reciben del computador.

- Dispositivos de entrada (sensores)
  - Posicionadores
    - magnéticos
    - ópticos
    - acústicos
    - mecánicos
    - de inercia
  - Guantes de datos
  - Registro de voz
  - Dispositivos de entrada 3D
- Dispositivos de salida (efectores)
  - Efectores visuales (casos estereoscópicos, sistemas basados en proyección)
    - Tecnologías de visualización
    - Factores que determinan la calidad de la visualización
    - Dispositivos de visualización
  - Efectores auditivos (altavoces)
  - Efectores táctiles (guantes táctiles, dispositivos de realimentación de fuerza) y efectores de equilibrio (plataformas móviles)

### 2.3.1. Dispositivos de entrada (sensores)

Los dispositivos de entrada capturan las acciones del participante (por ejemplo movimientos de la cabeza) y envían esta información al computador encargado de llevar a cabo la simulación.

Los dispositivos que veremos a continuación se consideran de realidad virtual bien porque utilizan el paradigma de la interacción implícita (se captura la voluntad del usuario implícita en sus movimientos y acciones naturales), o bien porque proporcionan entrada 3D al sistema, para hacer más sencilla la exploración del mundo virtual.

- Posicionadores
- Guantes de datos (guantes con una función de entrada de datos)
- Registro de voz
- Dispositivos de entrada 3D

#### Posicionadores

Los posicionadores son sensores que tienen como misión capturar la posición o/y orientación de un objeto real y enviar esta información al computador.

#### Prestaciones

Hay posicionadores que solamente registran la posición de un objeto, y por tanto se dice que miden tres grados de libertad (en inglés, *degrees of freedom*, *DOF*). Estos tres grados son las tres coordenadas X, Y, Z del objeto respecto un sistema de coordenadas conocido que proporciona el posicionador. Otros posicionadores sólo registran la orientación de un objeto. En este caso también tenemos tres grados de libertad, que son los ángulos de rotación sobre los ejes X, Y, Z. Estos ángulos reciben nombres particulares en realidad virtual, especialmente cuando medimos la orientación de la cabeza:

- **Yaw (azimuth)**. Ángulo de rotación respecto al eje vertical. Este ángulo varía por ejemplo cuando miramos de derecha a izquierda.
- **Pitch (elevation)**. Ángulo de rotación respecto un eje horizontal que une las dos orejas. Este ángulo mide la elevación respecto al horizonte, y varía por ejemplo cuando miramos de arriba a abajo.

- **Roll.** Ángulo de rotación respecto al eje determinado por la dirección de visión. Este ángulo varía por ejemplo cuando acercamos la cabeza a la espalda.

Finalmente, la tercera clase de posicionadores proporcionan posición y orientación del objeto rastreado, por lo que se denominan posicionadores con seis grados de libertad (*6 DOF trackers*).

## Aplicaciones

El uso habitual de los posicionadores en realidad virtual es para capturar los movimientos de la cabeza del participante, lo que nos permite generar las imágenes de acuerdo con su punto de vista. También es frecuente registrar los movimientos de la mano para permitir al usuario interactuar con los objetos de la escena de forma natural. Acciones como coger o mover objetos de la escena se pueden realizar de forma mucho más intuitiva mediante un reconocimiento de los movimientos del participante.

Algunos sistemas inmersivos utilizan una representación virtual del cuerpo del participante (denominada avatar) que se obtiene mediante un conjunto de sensores situados en determinadas articulaciones clave del participante. En este caso, normalmente se escoge un conjunto reducido de articulaciones y puntos clave (como por ejemplo pies, tobillos, rodillas, codos, hombros, etc.) a partir de los cuales se reconstruye el cuerpo completo. Estos sensores también se utilizan para obtener los movimientos del usuario en aplicaciones que no son de realidad virtual. Es el caso del análisis de los movimientos de los deportistas de alta competición que llevan a cabo algunas aplicaciones usadas por los centros de alto rendimiento, así como las secuencias de movimientos que se capturan a partir de un actor real y que se aplican a un carácter virtual. De esta forma se puede conseguir que el actor virtual haga unos movimientos mucho más complicados y realistas de los que se podrían conseguir con otros métodos. Esta técnica se utiliza actualmente en muchos juegos y también en dibujos animados.

## Tecnologías

- Magnéticos

Utilizan bobinas para obtener la posición y orientación basándose en las variaciones de tensión eléctrica inducidas por una fuente de campo magnético, que debe estar siempre próxima a los sensores. Requieren cables entre éstos y la unidad central, son sensibles a interferencias magnéticas y tienen un radio de acción de pocos metros, en el que pierde precisión conforme se va alejando del centro. Acostumbran a ser pequeños, con una alta libertad de movimientos y no precisan de una “línea de visión” entre emisor y receptor.



Figura 2.1: Posicionador magnético *Polhemus Liberty*

- Ópticos

Se basan en el procesamiento de imágenes captadas por más de una cámara de óptica y posición conocidas. Este análisis de imagen introduce un retardo, y tiene el problema de que siempre se ha de mantener una “línea de visión” entre emisor y receptor.

- Acústicos

Utilizan sonido ultrasónico y micrófonos. Soportan distancias mayores que los magnéticos, y su latencia acostumbra a ser pequeña (proporcional a la distancia), pero son poco utilizados porque la señal sonora se distorsiona con facilidad, a parte de que presentan el mismo problema de pérdida de “línea de visión”.

- Mecánicos  
Utilizan potenciómetros montados sobre una estructura articulada para medir los ángulos. Son los más precisos y poseen latencias muy bajas. No tienen el problema de pérdida de “línea de visión” ni padecen interferencias externas, pero limitan la libertad de movimientos y acostumbran a ser voluminosos y pesados, y son más caros.
- De inercia  
Se basan en pequeños dispositivos que permiten medir la aceleración a la que están sometidos. Tienen un radio de acción prácticamente ilimitado pero el principal inconveniente es que el error en la medida es acumulativo debido a que cada posición se calcula a partir de la última posición calculada. Son adecuados para detectar movimientos pero no una posición absoluta.



Figura 2.2: *Tracker inercial InterSense InterTrax2*

## Guantes de datos

Los guantes de datos que se utilizan en realidad virtual permiten detectar la posición de los dedos de la mano, normalmente expresada como los ángulos de flexión de cada dedo. Los guantes de datos se utilizan para interactuar con los objetos de la escena mediante gestos naturales, así como para dar órdenes al sistema mediante un lenguaje basado en signos.



Figura 2.3: Guantes de datos *5DT Ultra Wireless Kit*

### Registro de voz

En realidad virtual el reconocimiento de voz se utiliza como una forma natural de entrada de órdenes. Actualmente los sistemas de reconocimiento de voz están poco desarrollados y normalmente requieren de entrenamiento para tener un funcionamiento aceptable.

### Dispositivos de entrada 3D

Estos dispositivos permiten interactuar de forma más cómoda con un modelo 3D. Normalmente se clasifican según el número de grados de libertad (*DOF*).

- *Space Ball*  
Este dispositivo es parecido, superficialmente, a un *trackball*. Está diseñado para que el operador con la mano, y lo empuje o lo intente girar. El dispositivo emite información sobre la fuerza o el momento angular que se le aplica. De esta forma se pueden obtener datos con seis grados de libertad disponibles en el espacio (tres de posición y tres de giro).
- *Joystick 3D*  
Este dispositivo es una variante de la clásica palanca de los videojuegos, donde se ha incorporado un posicionador dentro del *joystick*. De esta forma es posible utilizarlo dentro de una CAVE.
- *Stylus*  
El *Stylus* es un puntero virtual, representado en la escena por un rayo con el que podemos apuntar a objetos de la escena virtual, seleccionarlos

y manipularlos mediante botones del propio *Stylus*. Con este dispositivo el usuario puede seleccionar y mover objetos, o también navegar a través de menús tridimensionales.



Figura 2.4: *IBM Spaceball 3D*

### 2.3.2. Dispositivos de salida (efectores)

Los dispositivos de salida generan los estímulos necesarios para los sentidos del participante, traduciendo en imágenes, sonidos, etc. las señales de video, audio, etc. que reciben del computador.

Los dispositivos de salida se clasifican en:

- Efectores visuales (casos estereoscópicos, sistemas basados en proyección)
- Efectores auditivos (altavoces)
- Efectores táctiles (guantes táctiles, dispositivos de realimentación de fuerza) y efectores del equilibrio (plataformas móviles)

#### **Efectores visuales**

Los efectores visuales tienen como objetivo proporcionar estímulos para el sentido de la vista (imágenes). De este tipo de efectores, veremos los siguientes aspectos:

- Tecnologías de visualización
- Factores que determinan la calidad de la visualización

- Dispositivos de visualización

### **Tecnologías de visualización**

Independientemente del paradigma de visualización, hay tres tecnologías básicas de generación de imágenes: CRT, LCD y DLP.

La tecnología CRT (*Catode Ray Tube*, tubo de rayos catódicos) es la tecnología que utilizan la mayor parte de los aparatos de TV. Se basa en utilizar una pantalla con la cara interna recubierta con tres tipos de fósforo, que emiten luz roja (R), verde (G) o azul (B) cuando incide un haz de electrones. Esta tecnología permite diseñar dispositivos con mucha resolución, pero son pesados y requieren un voltaje considerable.

La tecnología LCD (*Liquid Cristal Display*, cristal líquido) es la tecnología que utilizan la mayor parte de ordenadores portátiles. Se basa en una matriz de pequeños componentes que varían su transparencia cuando se le aplica un voltaje. Estos elementos no emiten luz y requieren de una fuente de luz externa detrás de la matriz LCD. La resolución de las pantallas LCD está muy limitada por sus dimensiones físicas, de forma que es muy difícil construir pantallas pequeñas con mucha resolución. Normalmente son más o menos sensibles al ángulo de visión. Dentro de esta tecnología hay muchas variantes: TFT, LCD-Polysilicon . . .

La tecnología DLP (*Digital Light Processing*, procesamiento digital de la luz) es la que utilizan los proyectores más ligeros que se fabrican actualmente. Se basa en un microchip fabricado por *Texas Instruments* que incorpora una gran cantidad de pequeños espejos oscilantes. Cada espejo puede estar orientado en dos direcciones: en una refleja la luz de una lámpara hacia el objetivo y en la otra no. Cada microespejo controla la intensidad de un píxel según la fracción de tiempo que está orientado en "ON". Esta nueva tecnología incorpora algunas mejoras respecto a la tecnología LCD, pues permite diseñar proyectores más pequeños y ligeros.

Existen más tecnologías de visualización (pantallas de plasma, proyección de un láser directamente sobre la retina . . .) pero su uso en realidad virtual es poco extenso.

### **Factores que determinan la calidad de visualización**

En un sistema de visualización hay algunas características especialmente importantes que determinan la calidad de la imagen, como la resolución, la profundidad de color, la frecuencia de refresco y el campo visual.

- **Resolución**

La resolución es el número de píxeles de una imagen, cuyos colores se pueden controlar de forma individual.

Los sistemas basados en LCD y DLP tienen una resolución “nativa” que corresponde con el número de transistores y microespejos, respectivamente. En los sistemas basados en CRT, no se puede hablar de resolución nativa sino de resolución máxima.

Una resolución elevada, por ejemplo 1024x768, no es garantía de buena calidad de imagen, ya que entre otras cosas tenemos que tener en cuenta las dimensiones físicas de la pantalla, y la distancia del observador a esta pantalla. Por esta razón es común utilizar la resolución angular, definida como el número de píxeles para cada grado de visión de un observador situado a una determinada distancia de la pantalla. La resolución angular tiene una relación directa con la resolución de las imágenes que se proyectan en la retina del observador, y por tanto es una medida adecuada para comparar dispositivos como los cascos estereoscópicos. El ojo humano puede distinguir aproximadamente unos 30 ciclos por grado.

- **Profundidad de color**

La profundidad de color determina los niveles de intensidad con la que el dispositivo es capaz de visualizar un color, y se mide normalmente como el número de colores diferentes que se pueden visualizar. Normalmente, la profundidad de color está limitada no por el dispositivo de visualización sino por la placa gráfica que genera la señal de video correspondiente. Una excepción son los proyectores DLP, donde los niveles de intensidad de cada componente de color están limitados por la velocidad de oscilación de los microespejos.

- **Frecuencia de refresco**

La frecuencia de refresco es el número de imágenes que se visualizan por unidad de tiempo. Conviene distinguir la frecuencia de refresco del

dispositivo de visualización (*vertical frequency*), de la frecuencia de refresco de la aplicación gráfica determinada (*frame rate*).

La frecuencia de refresco del dispositivo de visualización determina el número de imágenes diferentes que es capaz de presentar en un segundo (se mide en Hz). El nombre proviene de las pantallas basadas en CRT, donde las imágenes se tiene que repintar continuamente haciendo un barrido de la pantalla. En los sistemas LCD y DLP, la frecuencia máxima de refresco depende del tiempo de respuesta de los componentes electrónicos y mecánicos utilizados.

La frecuencia de refresco de una aplicación gráfica es el número de imágenes diferentes que es capaz de calcular en un segundo. Normalmente se mide en fotogramas por segundo (fps).

La frecuencia de refresco del dispositivo depende de la señal de sincronismo vertical que genera la placa gráfica, mientras que la frecuencia de refresco de la aplicación depende de la velocidad en que la aplicación hace el redibujado del *frame buffer*. Si la frecuencia de refresco del dispositivo o de la aplicación es muy baja, las imágenes se perciben a saltos. Para que el usuario no perciba las imágenes con una secuencia continua, hay que proporcionar una frecuencia de refresco bastante elevada (alrededor de 30 imágenes por segundo).

- **Campo visual (*field-of-view, FOV*)**

El campo visual del ojo humano es aproximadamente de 180 grados de dirección horizontal, y un poco menos en dirección vertical. El campo visual de un sistema de visualización indica la parte del campo visual humano que cubre la imagen de la pantalla. En realidad virtual el campo visual es muy importante porque condiciona el grado de inmersión. Cuanto más se acerca el campo visual del dispositivo al campo visual humano, más sensación de inmersión tendremos.

## **Dispositivos de visualización**

En realidad virtual, existen dos paradigmas básicos de visualización: cascos estereoscópicos (anulan completamente la visión del entorno real) y sistemas basados en proyección (el usuario puede ver parte del entorno real).

## ■ Cascos estereoscópicos

Un casco estereoscópico es un dispositivo que se adapta a la cabeza del participante y le permite ver el mundo virtual en 3D. Los cascos estereoscópicos incorporan sensores que almacenan la posición y la orientación de la cabeza del participante para actualizar su punto de vista de la escena adecuadamente.

Los cascos estereoscópicos se pueden clasificar en dos categorías según si se adaptan y respaldan únicamente la cabeza del usuario, o por el contrario, si se deben sujetar con las manos como unos binoculares:

- Los **HMD** (*Head-Mounted Displays*) son cascos inmersivos que se adaptan a la cabeza del usuario. Los HMD incorporan dos pequeñas pantallas y los elementos ópticos necesarios para permitir el enfoque. Cada pantalla está alineada con un ojo y recibe una señal de video diferente. Las tecnologías más habituales son LCD y CRT. Los HMD con pantallas LCD tienen muy poca resolución, pero son más ligeros.
- Los **HCD** (*Head-Coupled Displays*) son similares a los HMD pero incorporan pantallas CRT muy pesadas, por lo que están montados sobre un soporte mecánico y no se apoyan directamente en la cabeza del usuario. Este soporte mecánico incorpora potenciómetros en las articulaciones con el propósito de capturar la posición y orientación de la cabeza del usuario.

## ■ Sistemas basados en proyección

En estos sistemas, las imágenes se proyectan en una o más pantallas (normalmente pantallas de proyección) que pueden adoptar diferentes configuraciones según el número, forma y disposición. Las configuraciones comerciales más utilizadas son: la habitación estereoscópica (*CAVE*) y la mesa estereoscópica (*WorkBench*).

- Habitación estereoscópica (*CAVE*). Una *CAVE* (*Cave Automatic Virtual Environment*) consiste en una habitación con unas dimensiones alrededor de 3x3x3 metros, en la cual de 2 a 4 paredes laterales, el suelo y/o el techo son pantallas donde se proyectan imágenes estereoscópicas. Lo más común es utilizar retroproyección

para las paredes (para evitar que el participante vea su sombra en la pantalla) y proyección directa para el suelo.

- Mesa estereoscópica (*WorkBench*). Una mesa estereoscópica consiste en una estructura donde las imágenes se proyectan en una pantalla horizontal en forma de mesa. Cuando las aplicaciones de realidad virtual requieren la manipulación de objetos en el espacio virtual, es muy importante el grado de coordinación entre la visión y los movimientos de la mano, coordinación que es muy débil en los sistemas de visualización basados en cascos estereoscópicos, que requieren utilizar una representación virtual de la mano del usuario. Las tablas estereoscópicas son especialmente útiles en aplicaciones de medicina, al no presentar problemas de coordinación entre la vista y la mano.

Otras configuraciones de los sistemas basados en proyección son el denominado *StereoWall* (con pantalla vertical), el *Holobench* (dos pantallas en forma de “L”), el *domus* (pantalla esférica) y la *reality room* (pantalla cilíndrica).

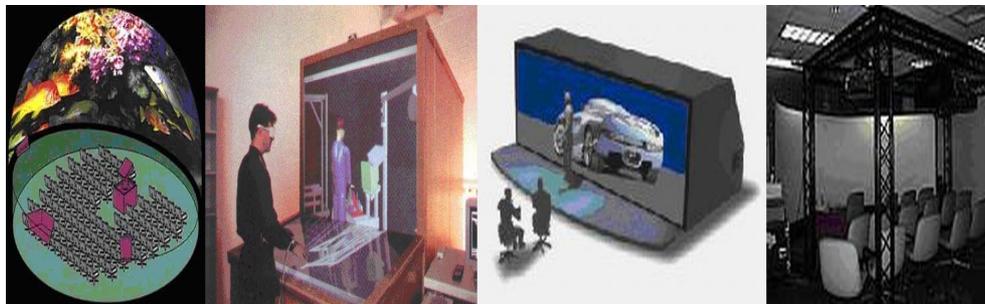


Figura 2.5: *Domus, HoloBench, StereoWall y reality room*

Los sistemas basados en proyección tienen algunas ventajas respecto a los cascos estereoscópicos:

- El usuario no tiene que transportar el dispositivo de visualización, que aísla del mundo real y limita en la interacción con los otros usuarios.
- La visión periférica es mucho mejor pues el campo visual del sistema se equipara al campo visual humano.
- El usuario puede ver su propio cuerpo, hecho que elimina el problema de la desorientación, mucho más frecuente en los HMD.

- La posición del usuario también tiene que ser monitorizada en tiempo real para proporcionar la perspectiva correcta, pero es muy poco sensible a los errores de precisión o al tiempo de latencia, pues solamente introduce pequeña corrección de la pirámide de visión. Si el usuario gira la cabeza en una CAVE, la imagen de las pantallas no se altera, mientras que en un HMD tiene que girar de acuerdo con la rotación.

### **Efectores auditivos**

Los efectores auditivos son simplemente altavoces que reproducen sonidos digitales.

En realidad virtual la generación de sonido realista requiere tener en cuenta las propiedades acústicas de los objetos y que los algoritmos son tan complicados como los algoritmos de visualización.

En realidad virtual frecuentemente se utiliza el sonido espacial, que consiste en generar el sonido de los objetos de forma que el participante es capaz de ubicar espacialmente la procedencia del sonido. En el mundo real, el sistema auditivo calcula la dirección de procedencia de un sonido a partir de las pequeñas diferencias en tiempo y amplitud con la que se registran las ondas auditivas en cada oreja, de forma que somos capaces de estimar la procedencia del sonido con los ojos cerrados. Un sistema de realidad virtual puede utilizar este hecho para sintetizar sonidos espaciales, utilizando múltiples altavoces.

También es muy frecuente encontrar aplicaciones de realidad virtual que utilizan el sonido como un sustituto del tacto, debido a que los altavoces son más económicos que los sistemas táctiles.

### **Efectores táctiles**

Se trata de dispositivos que proporcionan estímulos para el sentido del tacto.

Los guantes táctiles dan sensación de contacto con los objetos virtuales. Algunos guantes utilizan pequeñas cámaras neumáticas que se llenan/vacían de aire; otros utilizan dispositivos vibradores situados en los dedos, así como pequeñas descargas eléctricas.

Los dispositivos de realimentación de fuerza van más allá e impiden o ofrecen resistencia a hacer movimientos con la mano cuando ésta choca virtualmente con un objeto virtual o con un campo de fuerzas simulado. Frecuentemente sólo limitan el movimiento de los dedos, aunque los dispositivos tipo *Phantom* limitan el movimiento de uno de sus extremos.

### **Efectores del equilibrio**

Finalmente, algunos sistemas de realidad virtual, sobretodo en el campo de los simuladores y de los juegos por computador, incluyen módulos *software* para gestionar plataformas móviles que permiten recrear sensaciones de aceleración y desaceleración aprovechando la gravedad terrestre. El *software* de simulación se limita a calcular los grados de inclinación de la plataforma en función de parámetros del movimiento del vehículo virtual como son la aceleración y la fuerza centrípeta.