



UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS
CENTRO DE INFORMÁTICA INDUSTRIAL

SISTEMA DE SONIDO 3D PARA MEJORAR LA SENSOPERCEPCIÓN EN EL ESTUDIO DE FUNCIONES VISUALES

TESIS PRESENTADA EN OPCIÓN AL TÍTULO DE MÁSTER EN
INFORMÁTICA APLICADA

Autor: Ing. Yoander Cabrera Díaz

Tutor: Dr. Santiago Martín González

Co-Tutor: MSc. Yanoski Rogelio Camacho Román

Ciudad de La Habana

Febrero de 2011

Declaración jurada de autoría

Yo, Yoander Cabrera Díaz, con carné de identidad 83081115889, declaro que soy el autor principal del resultado que expongo en la presente memoria titulada Sistema de Sonido 3D para Mejorar la Sensopercepción en el Estudio de Funciones Visuales, para optar por el título de Máster en Informática Aplicada.

Este trabajo fue desarrollado durante el período 2009-2010 en colaboración con mis colegas de equipo, quienes me reconocen la autoría principal del resultado expuesto en esta memoria.

A todos mis colegas del equipo de trabajo les estoy muy agradecido. En especial deseo agradecer a mis tutores Dr. Santiago Martín González y MSc. Yanoski Rogelio Camacho Román que contribuyeron a mi formación como máster. También a mis padres, esposa, y amistades quienes también apoyaron a mi crecimiento profesional y humano en general. A todos ellos, así como a otros colegas y amigos que no he mencionado por razones de espacio, les doy las más sinceras gracias.

Finalmente declaro que todo lo anteriormente expuesto se ajusta a la verdad, y asumo la responsabilidad moral y jurídica que se derive de este juramento profesional.

Y para que así conste, firmo la presente declaración jurada de autoría en la Ciudad de la Habana a los ___ días del mes de febrero del año 2011.

Ing. Yoander Cabrera Díaz

Resumen

En este trabajo se desarrolla un sistema de sonido 3D sobre una biblioteca de visión estereoscópica y realidad virtual. Esta biblioteca tiene distintos campos de aplicación, lo que hace necesario implementar distintos comportamientos para el módulo de sonido. En primer lugar y como característica singular y más importante, la biblioteca es utilizada en aplicaciones de evaluación y entrenamiento de funciones visuales como la agudeza y la visión binocular. En ellas el módulo de sonido contribuye a la mejora de la sensopercepción del usuario. El audio tridimensional es resultado de combinar la posición de distintas fuentes de sonido con la posición de un puntero con movimiento tridimensional. De esta forma, el sonido 3D se convierte en una señal de profundidad complementaria a la visual, que ayuda al entrenamiento del paciente con problemas visuales.

En segundo lugar, la biblioteca es utilizada para simular ambientes de realidad virtual con seguimiento posicional. Es requisito que el sistema de sonido también sea compatible con este tipo de aplicaciones. Por ello se desarrolla una funcionalidad complementaria donde el audio tridimensional es resultado de combinar la posición de distintas fuentes de sonido con la posición del usuario, determinada esta última mediante un dispositivo de seguimiento.

Finalmente, el sistema desarrollado también incluye la generación de sonido 2D para la reproducción de música de fondo y sonidos de eventos. En este caso la función del sonido es mejorar la motivación y disposición del usuario frente a las aplicaciones desarrolladas.

Palabras Clave: sensopercepción, sonido envolvente, sonido posicional 3D, funciones visuales

Índice de contenidos

1. Introducción	1
Estructura del documento	3
2. Desarrollo	5
2.1. Mecanismos de sensorpercepción	5
2.2. El sonido	5
2.2.1. Características	6
2.2.2. Percepción por el ser humano	6
2.3. Tecnologías de sonido	9
2.3.1. Sonido envolvente 5.1	10
2.3.2. Sonido posicional 3D	11
2.3.3. Comparación de los paradigmas de sonido 3D estudiados	15
2.3.4. Dispositivos de escucha	17
2.3.5. Bibliotecas para programación de sonido	17
2.3.6. Formatos	22
2.4. Funciones visuales y sonido posicional 3D	23
2.4.1. Funciones visuales	23
2.4.2. Utilización del sonido posicional 3D en las discapacidades visuales . .	24
2.5. Descripción del problema	29
2.5.1. Justificación de la incorporación del sistema de sonido posicional 3D a la GLSVe	29
2.5.2. Funcionalidades	29
2.6. Soluciones técnicas	31
2.6.1. Graphics Library for Stereoscopic Vision (GLSVe)	31
2.6.2. Funcionalidades incorporadas a la GLSVe:	32
2.6.3. Prototipos de actividades	37
3. Conclusiones finales	43
4. Recomendaciones	44

5. Publicaciones derivadas	45
Referencias	46
A. Glosario de términos	51
B. Acrónimos	53

Índice de figuras

1.	Sistema auditivo. [Kalat, 2008]	7
2.	Diferencia de intensidad inter-aural. [J. L. Fernandez, 2006]	8
3.	Diferencia de tiempo inter-aural. [J. L. Fernandez, 2006]	8
4.	Frecuencia proveniente de dos direcciones distintas. [García, 2005]	9
5.	Sistema envolvente 5.1.	11
6.	Sonido emitido desde un objeto. [M. Lumbreras, 2010]	12
7.	Combinación de dos ángulos y una distancia.	13
8.	Diferentes posiciones de la fuente de sonido. [M. Lumbreras, 2010]	14
9.	Efecto <i>Doppler</i> . [P. Turcan, 2004]	15
10.	Distintas situaciones del entorno virtual AudioChile. [J. Sánchez, 2005a] . . .	25
11.	Sistema de sonido 3D con seguimiento de la cabeza. [Ch. Frauenberger, 2003]	27
12.	Objeto emisor de sonido. [Sánchez, 2008]	27
13.	Diagrama de clases del sistema de sonido posicional 3D.	39
14.	Diagrama de clases del sistema de visualización de video estereoscópico. . . .	40
15.	Búsqueda de objetos parcialmente ocultos por paralaje.	41
16.	Trazado de contornos 3D.	41
17.	Búsqueda de objetos con penalización del ojo sano.	42

Índice de tablas

1.	Comparación de los paradigmas de sonido 3D estudiados	16
2.	Comparación de las bibliotecas de sonido estudiadas	21
3.	Aplicación de sonido posicional 3D en las discapacidades visuales	28

1. Introducción

Los ambientes de realidad virtual tienen el objetivo de sumergir al usuario en un mundo generado por ordenador, homologando su entorno a la realidad y logrando establecer operatividad e interacción en tiempo real entre la aplicación y el hombre. Para lograr esto en teoría los sistemas sensoriales como el tacto, la visión y la audición deben ser estimulados en su forma natural [T. Lentz, 2006] [D. Schroder, 2007]. En la práctica, los Sistemas de Realidad Virtual (SRV) más modernos muestran un establecido conjunto de métodos para incentivar al ser humano desde el punto de vista visual. A medida que la visión se convierte en una importante fuente de información a la percepción humana al igual que el tacto, la estimulación acústica es considerada como un aspecto elemental de percepción natural dentro de los SRV. Ello permite al usuario poder detectar correctamente objetos virtuales que se encuentren alejados de su posición actual mediante el sonido 3D [I. Assenmacher, 2004].

Para aumentar la inmersión del usuario en un escenario virtual, es necesario no perder de vista la posición de la fuente y el oyente, así como la dirección y orientación de este, para que llegue a cada oído la información sonora correcta en cada momento. Igualmente debe respetarse la sincronización del audio con los estímulos visuales [I. Assenmacher, 2004]. Aprovechando esta tecnología es posible montar sistemas que recreen sonido 3D útiles en diversas aplicaciones proporcionando vitalidad y credibilidad del entorno generado [T. Lentz, 2006].

Uno de los usos más recientes de las técnicas de visión estereoscópica (VE) y realidad virtual (RV) ha sido en el campo de la oftalmología. Se ha constatado su uso en la evaluación y entrenamiento de las funciones visuales: agudeza visual, dominancia ocular, convergencia y visión binocular. Estas funciones visuales pueden ser medidas con pruebas clínicas estandarizadas de la vista [Thomas T. Norton, 2002] y, a la vez, su funcionamiento puede ser mejorado después de un entrenamiento. Algunos estudios presentados en [Polat, 2009][A. L. M. Pambakian, 2004] así lo corroboran.

Otros estudios confirman que el uso de sonido 3D en aplicaciones vinculadas a las discapacidades visuales permite mejorar las habilidades de los usuarios en el reconocimiento de objetos en espacios virtuales [Ch. Frauenberger, 2003][Y. Eriksson, 2004][J. Sánchez, 2005a]

[M. Ohuchi, 2005][J. Sánchez, 2006a].

Por las potencialidades que ofrece el sonido en aplicaciones de RV que entrenan personas con discapacidades visuales y la carencia de este en la herramienta *Graphics Library for Stereoscopic Vision* (GLS_{Ve}), el proyecto *Herramientas de Desarrollo para Sistemas de Visión Estereoscópica* (HDS_{VE})¹ se ha propuesto entre sus líneas de investigación el estudio del uso del sonido en la evaluación y entrenamiento de funciones visuales.

GLS_{Ve} había sido creada anteriormente con propósitos más generales, en su concepción inicial tenía algunas deficiencias de diseño y todavía no tenía incorporada funcionalidades de sonido, ni de forma general ni específicamente para mejorar la sensopercepción ² en la evaluación y entrenamiento de funciones visuales de los usuarios.

Se plantea entonces el siguiente **problema de investigación** ¿Cómo incorporar sensopercepción acústica en la biblioteca GLS_{Ve} para el estudio de funciones visuales?

El **objeto de estudio** de la investigación para dar solución al problema planteado, se centra en la sensopercepción acústica y realidad virtual para el estudio de funciones visuales. El **campo de acción** se enfoca en la sensopercepción acústica y realidad virtual para el estudio de funciones visuales mediante el sonido 3D.

Con el propósito de brindar solución al problema, se plantea como **objetivo de la investigación**, desarrollar un sistema de sonido 3D en la biblioteca GLS_{Ve} para contribuir a la sensopercepción en el estudio de funciones visuales.

Se espera que con el desarrollo de un sistema de sonido 3D en la GLS_{Ve}, se contribuya a mejorar la sensopercepción en la evaluación y entrenamiento de funciones visuales, propiciand^o una herramienta de ayuda para la navegación e interacción del usuario en las aplicaciones.

¹El proyecto HDS_{VE} lo conformaban inicialmente un grupo de profesores de la Universidad de Oviedo (España). Como parte de un convenio de colaboración se ha comenzado a trabajar en conjunto con la Universidad de las Ciencias Informáticas (UCI), y para esta nueva línea de investigación colaboran médicos oftalmólogos del Hospital de Arriondas (España) y el Hospital Ramón Pando Ferrer (Cuba).

²Conocimiento sensorial de una realidad, basado directamente en la información que el individuo recibe de sus sentidos (audición, tacto, visión, gusto y olfato) [Calvo, 1986]. Es algo muy complejo que en realidad está compuesto por dos procesos que se encuentran muy relacionados, que son la sensación y la percepción de los estímulos [C. A. Ortíz, 2004].

No obstante, la biblioteca es utilizada además para simular ambientes de realidad virtual con seguimiento posicional. Por ello es requisito que el sistema de sonido también sea compatible con este tipo de aplicaciones. Finalmente, se incluye un sistema de sonido 2D cuya función es mejorar la motivación y disposición del usuario.

Para dar cumplimiento al objetivo, se plantearon las siguientes **tareas de investigación**:

- Elaboración del marco teórico a partir del estado del arte sobre el tema existente en la actualidad, para que la solución que se proponga sea actualizada.
- Diseño de una arquitectura lógica e implementación de una solución en la GLS_{Ve}, para la generación de sonido 2D y 3D en aras de contribuir a la sensopercepción en el estudio de funciones visuales.
- Validación de la solución propuesta en la GLS_{Ve}, partiendo de la construcción de prototipos de actividades para el estudio de funciones visuales.

Estructura del documento

El presente documento se encuentra estructurado en **Introducción, Desarrollo, Conclusiones y Recomendaciones**. Los cuatro primeros epígrafes del desarrollo constituyen el Marco Teórico de la investigación y los restantes la solución propuesta:

En **Mecanismos de sensopercepción** se realiza un breve resumen de los sistemas sensoriales que están especializados en responder a un tipo diferente de estímulo en los seres humanos.

En **El sonido** se realiza una breve descripción del sonido, para comprender en que consiste. Se describen además algunas de sus características de importancia para el trabajo y cómo es percibido biológicamente por los seres humanos.

En **Tecnologías de sonido** se estudian los paradigmas de implementación de sonido utilizados en video-juegos y aplicaciones interactivas, continuando con los dispositivos de escucha, bibliotecas de sonido y formatos de audio. El apartado presenta varias comparativas con el fin de elegir las tecnologías más adecuadas para el desarrollo del trabajo.

En **Las funciones visuales y el sonido posicional 3D** se exponen muy brevemente las funciones visuales (Visión binocular y Agudeza visual), luego se realiza una revisión bibliográfica actual referida a la utilización y beneficios del sonido posicional 3D en las discapacidades visuales.

En **Descripción del problema** se justifica la necesidad de un sistema de sonido 3D en la biblioteca GLSVe, luego se explican las funcionalidades necesarias que se deben brindar y se concluye con la mención de un aporte adicional al trabajo.

En **Soluciones técnicas** se presenta inicialmente la biblioteca GLSVe. A continuación se describen las funcionalidades incorporadas en GLSVe y por último se presentan las actividades realizadas que permiten validar el sistema de sonido posicional 3D incorporado.

2. Desarrollo

A continuación se presenta el desarrollo de cada uno de los epígrafes realizados en la investigación. Se recogen los principales conceptos, trabajos y soluciones relacionadas con el problema de investigación.

2.1. Mecanismos de sensopercepción

Se realiza un breve resumen de los sistemas sensoriales que están especializados en responder a un tipo diferente de estímulo en los seres humanos.

Cada uno de los sistemas sensoriales está especializado en responder a un tipo diferente de estímulo. La visión responde a la luz [C. A. Ortiz, 2004], se plantea que es el sentido más importante y ocupa mayor espacio en la corteza cerebral [Bayona, 2007]. La audición responde a variaciones de la presión del aire o de algún otro medio. El olfato y el gusto responden ambos a sustancias químicas [C. A. Ortiz, 2004]. El tacto es otro de los mecanismos de sensopercepción, a través de este, el cuerpo percibe el contacto con las distintas sustancias y objetos en el mundo [Bayona, 2007]. El trabajo se enmarca solamente en la sensopercepción relacionada con la audición.

2.2. El sonido

En este epígrafe se realiza una breve descripción del sonido, para comprender en que consiste. Se describen además algunas de sus características de importancia para el trabajo y cómo es percibido biológicamente por los seres humanos.

El sonido consiste en variaciones de la presión del aire o de algún otro medio. Estas variaciones de presión se originan en algún punto del espacio desde el cual irradian como una serie de ondas. Los sonidos se caracterizan por su volumen y tono (es decir lo graves o agudos que

parecen) [G. Pocock, 2005]. Cuando la presión ejercida por el paquete de ondas choca con las aurículas del oído se logra escuchar el sonido [Bartlett, 2009].

2.2.1. Características

- Amplitud: La altura de la onda corresponde a su amplitud. Los sonidos fuertes tienen altas amplitudes (implicando grandes cambios de presión), mientras que los sonidos suaves tienen una baja amplitud (implicando pequeños cambios de presión). La unidad de medida está proporcionada en decibelios (Db) [Bartlett, 2009][Kalat, 2009][Kalat, 2008][Weiten, 2010][Boer, 2002].
- Frecuencia: Es el número de vibraciones u oscilaciones por unidad de tiempo. Dependiendo de la frecuencia el sonido tendrá un tono grave (frecuencia baja) o un tono agudo (frecuencia alta). Se mide el *hertz*(Hz), correspondiendo un Hz a una vibración por segundo [Isbert, 1998][E. S. B., 2003]. Las variaciones de frecuencia de una fuente de sonido en movimiento respecto de un observador explican el efecto *Doppler*. Ver más en 2.3.2.

2.2.2. Percepción por el ser humano

El sonido es una parte importante de la experiencia sensorial cotidiana del ser humano, ya que la audición está comprendida dentro de los cinco sentidos de la persona [Henderson, 2010].

El ser humano convierte las ondas sonoras en desplazamientos mecánicos que un conjunto de células del receptor (oído) pueden detectar. El oído es capaz de percibir altas frecuencias como altos tonos y bajas frecuencias como bajos tonos, para que sean audibles deben estar entre los 20Hz y los 20 000Hz. Cuando las ondas de sonido llegan al tímpano hacen que este vibre, véase la figura 1. El tímpano está conectado a tres pequeños huesos, el martillo, el yunque y el estribo, que convierten las ondas sonoras en una serie de fuertes vibraciones en la cóclea, véase la figura 1, estas vibraciones se desplazan a las células de pelo a lo largo de la membrana basilar en la cóclea. Cuyas células de pelo que actúan como receptores del tacto

sobre la piel, conectan las neuronas formándose el nervio auditivo. Y finalmente este nervio envía impulsos a las áreas del cerebro responsables de la audición [Kalat, 2009][Kalat, 2008].

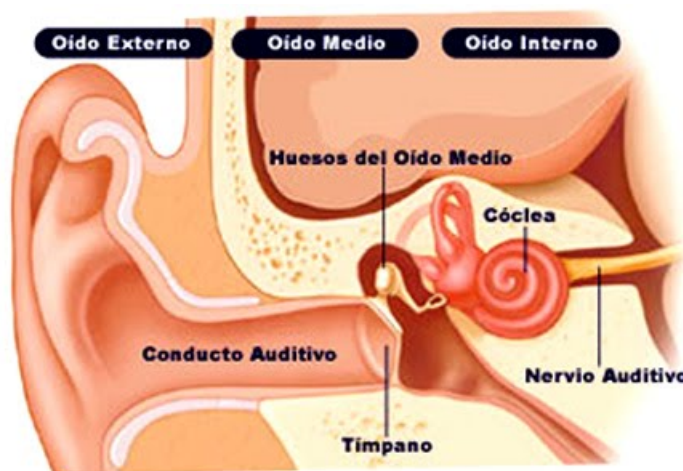


Figura 1: Sistema auditivo. [Kalat, 2008]

Los científicos que estudian la percepción del sonido han descubierto que el cerebro se basa en la siguientes claves para ubicar una fuente de sonido espacialmente.

Volumen: Diversos elementos producen sonido a diferentes volúmenes naturales, mientras más alto se encuentra un sonido más cerca el cerebro piensa que este está [McCuskey, 2003].

Diferencia de intensidad inter-aural (DII): Un sonido que provenga directamente desde la izquierda de la cabeza se escuchará con mayor fortaleza por el oído izquierdo y en menor medida por el oído derecho, indicación realizada por el cerebro para interpretar el sonido como procedente de la izquierda, véase la figura 2 [McCuskey, 2003].

Diferencia de tiempo inter-aural (DIT): Este es similar al concepto anterior, los sonidos procedentes de la izquierda llegarán al oído derecho con un milisegundo o menos más tarde que su llegada al oído izquierdo (en función de la distancia inter-aural), véase la figura 3 [McCuskey, 2003].

Muffling: Término del inglés, que define la forma en que la oreja está adecuada para conocer

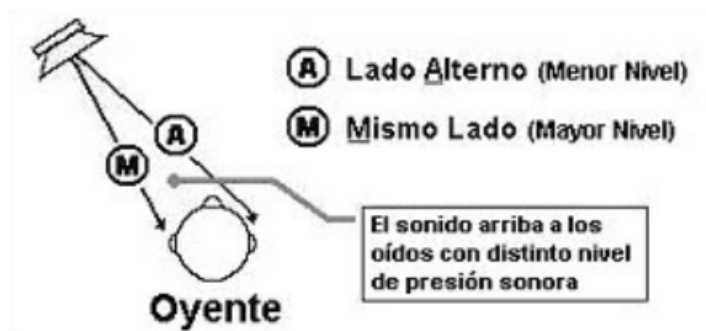


Figura 2: Diferencia de intensidad inter-aural. [J. L. Fernandez, 2006]

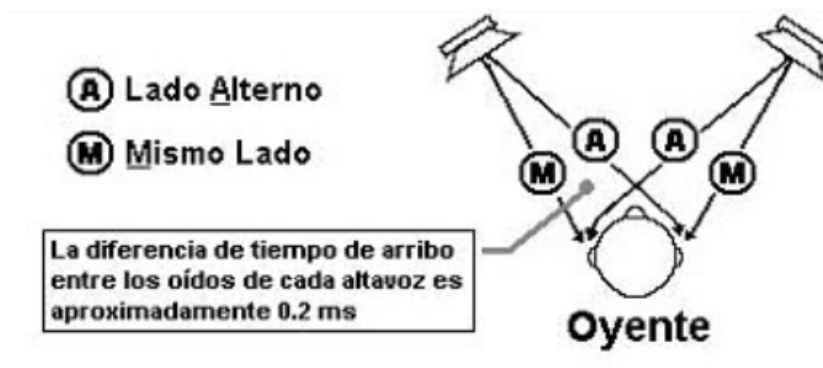


Figura 3: Diferencia de tiempo inter-aural. [J. L. Fernandez, 2006]

los sonidos que le llegan de frente. Por ejemplo un sonido escuchado por detrás sonará ahogado [McCuskey, 2003].

El pabellón auditivo actúa como una cavidad de resonancia, donde amplifica algunas frecuencias, y su geometría genera ciertos efectos de interferencia que atenúan otras frecuencias. Dicho de otra forma, su respuesta en frecuencia depende de la dirección. La figura 4 mide las respuestas en frecuencia de dos direcciones distintas. En ambos casos se ven los caminos que siguen las ondas [García, 2005].

Si se presenta un sonido de banda limitada con frecuencias centrales de 300 Hz o 3000 Hz la imagen sonora siempre se formará delante del sujeto. Si la frecuencia central es de 8000 Hz la imagen estará siempre arriba. Y si la frecuencia central es de 1000 o 10 000 Hz la imagen

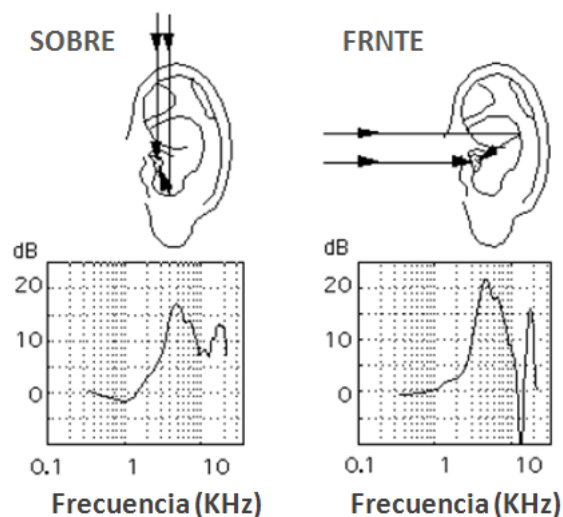


Figura 4: Frecuencia proveniente de dos direcciones distintas. [García, 2005]

se formará detrás [García, 2005].

2.3. Tecnologías de sonido

En el epígrafe se estudian los paradigmas de implementación de sonido utilizados en videojuegos y aplicaciones interactivas, continuando con los dispositivos de escucha, bibliotecas de sonido y formatos de audio. El apartado presenta varias comparativas con el fin de elegir las tecnologías más adecuadas para el desarrollo del trabajo.

El sonido 3D puede contribuir a aumentar la sensación de inmersión en un entorno 3D, posibilitando que los usuarios interactúen de forma eficaz, siempre que las acciones estén acompañadas de sonidos apropiados emitidos desde una posición. El sonido 3D proporciona un importante canal de retroalimentación que puede contribuir a redundancia visual o proporcionar información para las acciones y situaciones que están fuera del campo de vista del oyente. Tiene una ventaja sobre la visión y es que las múltiples fuentes de sonido virtual se pueden sintetizar al ocurrir en cualquier lugar en un espacio de 360° alrededor de un oyente. En una pantalla, el foco de atención entre las fuentes de sonido virtual puede ser cambia-

do a voluntad, la visión, por el contrario, requiere del movimiento de los ojos o la cabeza [[Begault, 2000](#)].

Elementos que forman parte del sonido en aplicaciones interactivas:

1. Los sonidos propios de objetos y/o personajes en escena, estos por lo general son sonidos monofónicos y que son posicionados dependiendo de la ubicación relativa de sus respectivas fuentes en el espacio tridimensional que rodea al oyente.
2. La banda sonora o música de fondo, que es generalmente estéreo.
3. Los sonidos ambientales principales, también en estéreo por lo general [[Fernandez, 2006](#)].

2.3.1. Sonido envolvente 5.1

Este paradigma de implementación de sonido desde el punto de vista técnico exige el uso de cinco parlantes y un canal de baja frecuencia, denominado subwoofer.

En su configuración el sonido envolvente se logra con:

- Tres canales frontales (izquierdo, central y derecho) de banda ancha de frecuencias.
- Dos canales envolventes traseros (izquierdo y derecho) que crean la sensación envolvente, también de banda ancha.
- Un canal para efectos de baja frecuencia. Este último canal sólo necesita aproximadamente una décima parte del ancho de banda que utilizan los demás [[Cid, 2006](#)], véase la figura 5.

El uso del subwoofer busca separar el manejo de las frecuencias bajas, permite la reducción en tamaño y en coste de los demás parlantes, además de reducir la distorsión resultante del sistema. La ubicación de los parlantes debe decidirse teniendo en consideración la geometría,

instalación eléctrica y acústica del recinto. Idealmente se espera que tengan paredes paralelas y techo, una altura mínima de 3 (m) y que el volumen del recinto tenga unos 300 (m^3) con un área de 30 (m^2) [Fernandez, 2006].

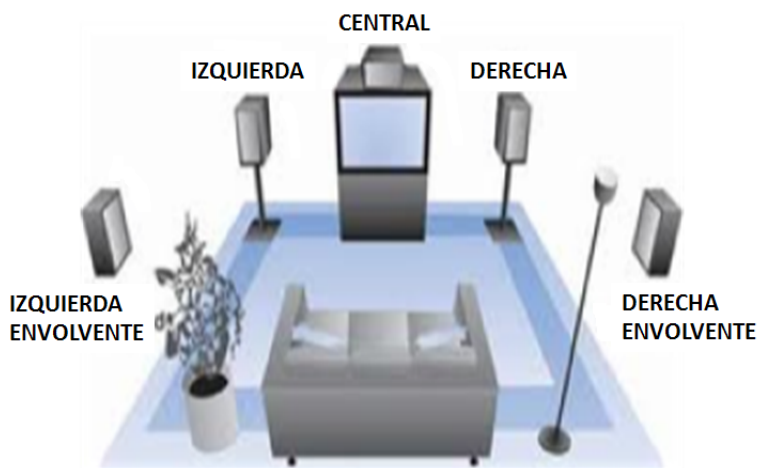


Figura 5: Sistema envolvente 5.1.

Este paradigma de sonido envolvente está concebido para su escucha desde una posición fija en el centro de la sala. La sensación tridimensional que se logra con el sistema no es totalmente realista. Si bien el sonido se recrea en un campo de 360°, consiguiendo sensaciones muy precisas en cuanto a la anchura y profundidad, no se logra la obtención de imágenes precisas en la dimensión altura [Cid, 2006].

2.3.2. Sonido posicional 3D

El paradigma de audio o sonido posicional 3D desde el punto de vista técnico se puede implementar bien mediante dos parlantes o bien mediante auriculares. El uso de auriculares permitiría el movimiento del observador en un entorno dotado de seguimiento posicional.

En este paradigma se definen fuentes generadoras de sonido (objetos), localizadas en un entorno o espacio 3D en relación al oyente. En la figura 6 se muestra un usuario (oyente) que ve un cubo. El sonido posicional 3D está ubicado en este caso en el centro del cubo. Cuando

el objeto es dinámico y/o el usuario navega en el espacio virtual, este tiene que escuchar que el sonido es emitido continuamente desde el objeto [M. Lumbreras, 2010].

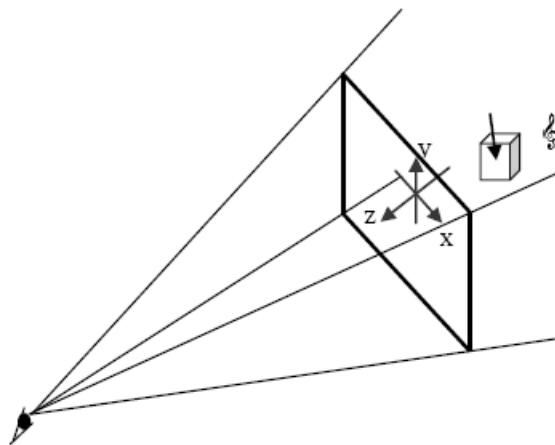


Figura 6: Sonido emitido desde un objeto. [M. Lumbreras, 2010]

El sonido posicional 3D utiliza la posición de cada fuente de sonido en relación con la posición del oyente para crear una imagen sonora realista. Utilizando un sistema de coordenadas esféricas es posible ubicar tridimensionalmente, mediante la combinación de dos ángulos y una distancia, las distintas fuentes sonoras respecto del usuario, véase la figura 7 [García, 2005].

En la figura 8 se indican diferentes posiciones del sonido respecto al oyente. En la (1) la fuente de sonido está cerca del oyente, en la (2) se encuentra lejos del oyente, en la (3) a la izquierda del oyente y en la (4) a la derecha del oyente.

Por lo tanto si un objeto con una fuente de sonido es ubicado a la izquierda del oyente, el sonido debe escucharse por el lado izquierdo de este. Lo cual quiere decir que en el oído izquierdo se escuchará más alto que en el oído derecho. Los anteriores planteamientos se sustentan por la DII, véase 2.2.2. Complementariamente, el sonido llegará antes al oído izquierdo que al derecho, lo cual puede ser simulado por el sistema de sonido posicional 3D para recrear el efecto de la DIT, véase 2.2.2.

Cuando el sonido está posicionado exactamente en frente o detrás del oyente ambos oídos deberán escuchar el sonido con el mismo volumen, pero con frecuencias diferentes, dado por

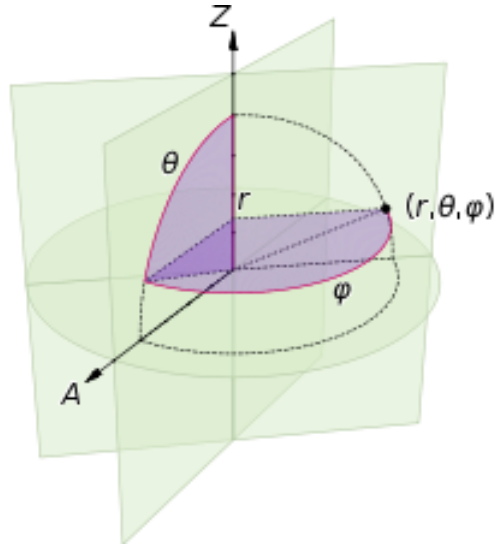


Figura 7: Combinación de dos ángulos y una distancia.

la forma con que está conformado el pabellón auditivo, véase 2.2.2.

El volumen con que se escuche el sonido depende de la distancia entre el objeto emisor y el oyente. Cuando la fuente de sonido se aleja el observador espera que su volumen disminuya. Distintas ecuaciones pueden ser utilizadas para modelizar la rapidez con que el sonido se desvanece [M. Lumbreras, 2010].

El sonido posicional 3D también permite simular el efecto *Doppler*, denominado así por el matemático y físico austriaco *Christian Doppler* [McCuskey, 2003]. Si una fuente sonora se desplaza hacia un observador, comprimirá las ondas sonoras emitidas, por lo que el sonido se apreciará con mayor frecuencia de la que realmente tiene. Al alejarse, la frecuencia decae, apreciándose una ampliación de las ondas sonoras que el objeto está emitiendo. En este caso, el sonido del objeto parece tener una frecuencia más baja de la que realmente tiene. Si el objeto no se mueve rápidamente, el efecto *Doppler* es prácticamente imperceptible [P. Turcan, 2004]. Un ejemplo habitual se pone de manifiesto al escuchar un tren pasando por una estación [McCuskey, 2003], véase la figura 9.

Para el cálculo de la frecuencia como forma de percibir el efecto *Doppler* se plantea la fórmula

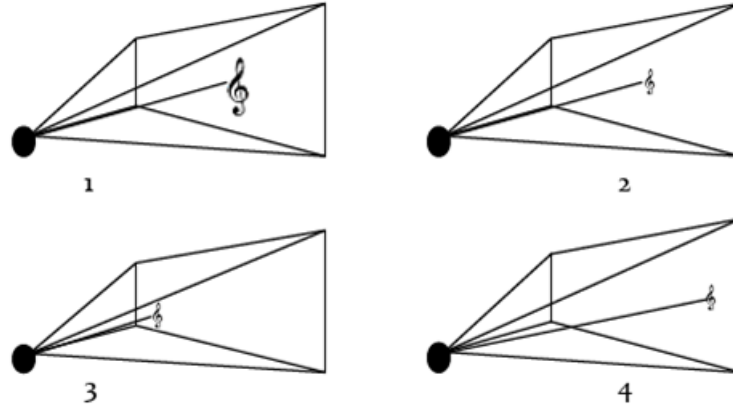


Figura 8: Diferentes posiciones de la fuente de sonido. [M. Lumbreras, 2010]

1.

$$f' = DF * f * \left[\frac{v - v_1}{v} \right] \quad (1)$$

Donde:

f : Tono original de la fuente

f' : Tono adquirido por el efecto *Doppler*

v_1 : Velocidad de la fuente respecto del oyente

v : Velocidad del sonido (igual a 343.3 m/s en el aire seco a 20° C)

DF : factor *Doppler* [0,1]

Ventajas del sonido posicional 3D como herramienta inmersiva en un mundo virtual

- Provee información extra para usuarios en la búsqueda de objetos cuando este navega en

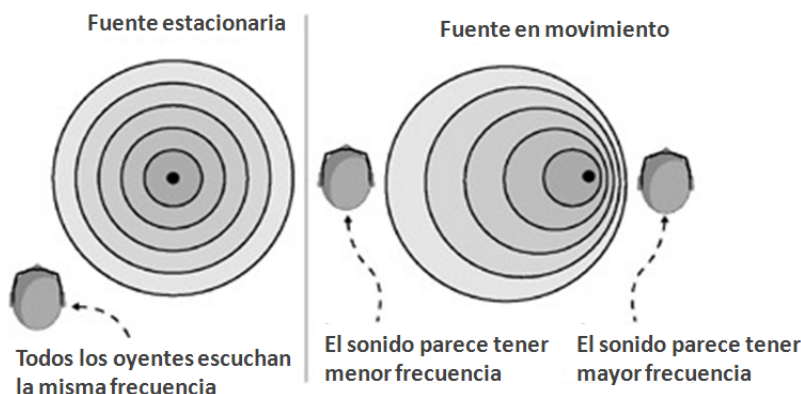


Figura 9: Efecto *Doppler*. [P. Turcan, 2004]

un espacio virtual 3D, ya que el sistema auditivo es capaz de determinar la localización de la fuente de sonido.

- Produce altos niveles de inmersión en un ambiente virtual.
- Ayuda a interpretar distancias entre objetos.
- Facilita interacción natural porque es similar al sonido en el mundo real.
- Provee información adicional porque ayuda a los usuarios a entender más sin esfuerzo adicional [M. Lumbreras, 2010].

2.3.3. Comparación de los paradigmas de sonido 3D estudiados

En la tabla 1 se realiza una comparación de las tecnologías de sonido 3D estudiadas, atendiendo a varios parámetros.

La implementación de sonido en aplicaciones de RV se realiza usualmente hoy en día bajo dos tecnologías diferentes como: el sonido envolvente 5.1 o superior y el audio posicional 3D. Si bien la calidad de sonido de ambos sistemas puede ser comparada debido a las características propias de cada una de ellas, no existe una opinión única de que sistema es más eficiente o

Paradigmas de sonido 3D	Cantidad de canales	Dispositivos de escucha	Calidad de escucha	Costo de puesta en práctica
Envolvente 5.1	6	Parlantes con sub-woofer	Buena (360º en la horizontal, no percibe sonidos en la vertical)	Alto
Posicional 3D	2	Parlante o Auriculares	Buena (3D)	Muy bajo

Tabla 1: Comparación de los paradigmas de sonido 3D estudiados

de cual presenta mejores resultados en cuanto a calidad del sonido reproducido. No existe una metodología para poder evaluarlos objetivamente entre sí, pero existen recomendaciones específicas para la evaluación de cada sistema por separado.

El sistema envolvente 5.1 necesita seis canales para la transmisión del audio, hardware específico como es el caso del subwoofer con los respectivos parlantes y configuración específica del recinto donde se vaya a instalar[Fernandez, 2006]. Aunque posee buena calidad de escucha en el plano horizontal en 360º grados alrededor del oyente todavía no logra simular los sonidos en la vertical [Cid, 2006], incluyendo a esto el alto costo de puesta en práctica dado a los requisitos de configuración y la adquisición del sistema. Mientras que con el sonido posicional 3D se logra recrear un ambiente tridimensional con solo dos canales para transmisión de sonido y un par de parlantes o auriculares, además su costo de puesta en práctica es mínimo ya que no contempla requisitos geométricos para la sala y la adquisición del dispositivo de escucha no es costosa. Finalmente, el sonido posicional 3D, en el caso de utilizar auriculares, es el único compatible con un sistema de seguimiento posicional. Si bien para el estudio de funciones visuales es posible diseñar aplicaciones donde el usuario permanezca en una posición fija, no hay que olvidar que la biblioteca GLSve es utilizada también para el modelizado de entornos de realidad virtual.

En esta comparación se determinó el empleo de la tecnología de audio posicional 3D para la

implementación del sistema de sonido 3D en la biblioteca GLS_{Ve}.

2.3.4. Dispositivos de escucha

El presente epígrafe se centra en la elección del dispositivo de escucha empleado para la solución del trabajo, pues como se ha indicado el sonido posicional 3D puede obtenerse tanto mediante parlantes como mediante auriculares.

Los auriculares simplifican el problema de enviar distintas señales a cada oído pero sin embargo presentan algunos inconvenientes. A muchas personas no les gusta, son incómodos en dependencia de su diseño, incluso si son inalámbricos. Atenuan los sonidos externos y aíslan socialmente al usuario.

Los parlantes resuelven la mayor parte de estos problemas, pero poseen deficiencias graves en la generación de sonido posicional 3D, dado que el oído izquierdo percibe información proveniente del parlante derecho y viceversa, que con los auriculares no sucede.

Finalmente se decidió el uso de auriculares porque la solución que se propone en el trabajo sobre la utilización del sonido posicional 3D para actividades que evalúen y entrenen funciones visuales está pensada para emplearse tanto en las clínicas como en las casas de los pacientes. Esta es una solución barata y no lleva limitantes de configuración como lo sería un sistema de dos o más parlantes. Además si más de un paciente se está tratando en el mismo local el uso de parlantes sería un problema por las interferencias ocasionadas.

2.3.5. Bibliotecas para programación de sonido

En el epígrafe se describen algunas de las bibliotecas utilizadas para el tratamiento de audio posicional 3D y sonido 2D en aplicaciones informáticas actuales. Finalmente se muestra una tabla comparativa para decidir cuales serán empleadas en la solución del trabajo.

La programación de audio en video-juegos u otras aplicaciones es un aspecto generalmente subestimado. En grupos de desarrollo reducidos raramente se tiene en cuenta la contratación de algún especialista relacionado al tema y dentro del esquema general de las bibliotecas, el audio se encuentra relegado. Sin embargo, el sonido y la música en un video-juego o aplicación interactiva es algo muy importante, basta con que juegue su video-juego favorito fijando el volumen al mínimo para notar la diferencia, sin sonido existe gran pérdida de inmersión [Boer, 2002].

Se estudiarán las bibliotecas DirectSound, OpenAL y SDL Mixer para elegir entre ellas las que se ajusten a las necesidades del sistema de sonido 3D a implementar. Se seleccionaron estas y no otras para el estudio dado que existe un conocimiento previo acerca de sus características y campos de aplicación.

DirectSound

DirectSound: Proporciona una interfaz directa entre las aplicaciones y los controladores de la tarjeta de sonido, permitiendo a las aplicaciones producir sonidos y música. Además de proporcionar el servicio esencial de pasar datos de audio a la tarjeta de sonido, proporciona muchas capacidades necesarias tales como grabación y mezclado de sonido; adición de efectos al sonido como reverberación³, eco⁴, *flanger*⁵; usando hardware controlado por *buffers*⁶ para velocidad extra; posicionar sonidos en espacio 3D. Proporciona acceso de bajo nivel para el hardware del dispositivo, es usada por los programadores para la reproducción básica de recursos de audio de un video-juego o aplicación que requiera sonido 3D [T. M. Fay, 2004]. Soporta solamente el formato de audio (wav) y está únicamente disponible en aplicaciones de Windows.

En teoría si se desea expresar al máximo las posibilidades de sonido de la PC (y se tienen

³Fenómeno derivado de la reflexión del sonido consistente en una ligera permanencia del mismo una vez que se ha extinguido el original.

⁴Fenómeno acústico producido cuando la onda sonora se refleja y regresa con retardo suficiente para superar la persistencia acústica.

⁵Término del inglés que define efecto de audio que produce un sonido metalizado oscilante, sobre todo en frecuencias medias y altas.

⁶Espacio de memoria, en el que se almacenan datos para evitar que el programa o recurso que los requiere, ya sea hardware o software, se quede en algún momento sin información.

los recursos necesarios para brindarle) DirectSound sería una opción, pero lo cierto es que los tiempos de desarrollo no son eternos y es posible que se desee inclinarse por opciones con una relación esfuerzo/resultado mas favorable [Boer, 2002]. Esta biblioteca realiza todos los cálculos necesarios como la atenuación debida a la distancia, la diferencia de intensidad inter-aural, la diferencia de tiempo inter-aural, la simulación del *muffling* y la frecuencia para percibir el efecto *Doppler*.

OpenAL

OpenAL es una interfaz de programación multiplataforma y libre para audio multicanal 3D apropiada en el uso de aplicaciones relativas al tratamiento de sonido 3D.

La biblioteca permite modelizar una colección de fuentes de audio moviéndose en un espacio tridimensional que son oídas por un único oyente en algún lugar de ese espacio. Los objetos básicos en OpenAL son un oyente (*Listener*), una fuente de audio (*Source*) y una zona de memoria (*Buffer*) que contiene la información de audio. Cada *buffer* puede ser asignado a una o más fuentes que representan posiciones (definidas por coordenadas en un espacio tridimensional) de donde brota el audio. Siempre hay un oyente, que representa el punto donde se escuchan los sonidos que generan las fuentes. OpenAL tiene la posibilidad de construir extensiones y usar la Extensión de Audio Ambiental (del inglés Environmental Audio Extensions) (EAX), tecnología propia de (Creative y SoundBlaster), que proporciona alto nivel de simulación acústica [Hiebert, 2007]. Soporta nativamente el formato de audio (wav).

La funcionalidad de OpenAL se estructura en base a estos objetos:

1. Una fuente (*source*) contiene un puntero a una zona de memoria (*buffer*), la velocidad, posición y dirección e intensidad del sonido [Hiebert, 2007].
2. El oyente (*listener*) representa la velocidad, posición y dirección del mismo, así como la ganancia asociada a todos los sonidos. Aunque se pueden definir varios oyentes solo uno puede estar activo [Hiebert, 2007].
3. Los *buffers* contienen audio en formato (PCM) modulación por impulsos codificados,

(muestreado cada cierto tiempo y codificado en una serie de bits [Jordi D. Gavalda, 2008]), en muestras de 8 o 16 bits, tanto en monofónico como en formato estéreo.

El motor de OpenAL realiza todos los cálculos necesarios como la atenuación debida a la distancia, la diferencia de intensidad inter-aural, la diferencia de tiempo inter-aural, la simulación del *muffling* y la frecuencia para percibir el efecto *Doppler*. El resultado para el usuario final es que las aplicaciones realizadas con OpenAL recrean un escenario aural cercano mientras el oyente se mueve en un espacio tridimensional [Ope, 2005].

Un típico programa que utiliza OpenAL empieza con instrucciones para abrir un dispositivo de sonido del sistema, que procesará el sonido y lo reproducirá mediante el hardware asociado (parlantes o auriculares). Después se crea un contexto y se asocia con un dispositivo, dentro del cual también se definirá un oyente. A partir de ese momento, sobre el contexto activo, se disponen fuentes de sonido, en las que se pueden modificar las propiedades relativas a sus coordenadas espaciales y a su orientación; junto a operaciones de procesamiento del audio que afectan cómo se modifica la señal de audio original en la escena (por ejemplo la atenuación). Las fuentes se inicializan con sonidos que se obtienen a partir de los *buffers*, que almacenan la descripción de un sonido desde un fichero o a partir de una función de síntesis de una señal básica. El *buffer* se asignará a una fuente y desde ese momento puede activarse, empezará a reproducir el sonido en cuestión con las modificaciones que la ejecución del resto del programa imponga. Terminada la aplicación hay que liberar los recursos creados antes de salir de la misma [Ope, 2005].

SDL Mixer

La biblioteca *Simple DirectMedia Layer* (SDL) está diseñada para que sea fácil de usar en video-juegos y aplicaciones interactivas que se ejecutan en *Linux*, *MacOS*, *Win32* y *BeOS*, usa diferentes medios de comunicación nativos de alto rendimiento con interfaces para video, audio, entre otras. SDL es una Interfaz de Programación de Aplicaciones (del inglés *Application Program Interface*) (API) libre, de nivel bajo, pero con ella, se pueden desarrollar aplicaciones completamente portátiles con gran flexibilidad y eficiencia [Lantinga, 2001].

SDL Mixer es un complemento que mejora el subsistema de audio de SDL. Está preparada

Bibliotecas	Archivos de sonido nativo	Atenuación	Efecto <i>Doppler</i>	<i>Muffling</i>	DII-DIT	Presente en Tao Framework
DirectSound	wav	si	si	si	si	no
OpenAL	wav	si	si	si	si	si
SDL Mixer	wav, mp3, midi, Ogg Vorbis, MOD, IT, S3M y VOC	si	no	no	no	si

Tabla 2: Comparación de las bibliotecas de sonido estudiadas

para manejar múltiples sonidos al mismo tiempo además de la música. Es más, si se es capaz, se puede especificar la manera de mezclar la música y aplicar varios efectos en su aplicación manejando punteros a funciones que realicen esta tarea.

SDL Mixer se encarga de realizar la mezcla de canales de audio de forma automática lo que ahorra tiempo en el desarrollo de un sistema de mezclado. Para reproducir sonido en la aplicación, SDL diferencia el audio referente a la música del juego, a la que coloca en un canal independiente, de los sonidos de efectos propios del juego. Los formatos con los que trabaja esta biblioteca son (wav, mp3, midi, Ogg Vorbis, MOD, IT, S3M y VOC) [[Visznecki, 2009](#)].

Comparación de las bibliotecas de sonido 3D estudiadas

En la tabla 2 se realiza una comparación de las bibliotecas de sonido estudiadas, atendiendo a varios parámetros.

Como se observa en la tabla 2 la biblioteca OpenAL y DirectSound coinciden con el cumplimiento de los cinco primeros aspectos. Independientemente de la similitud entre ellas se decidió el uso de OpenAL para el desarrollo del sistema de sonido posicional 3D. La razón de la elección se basa fundamentalmente en tres aspectos, primero, está contenida en el recubrimiento Tao Framework (Interfaz de programación de C# para acceso a bibliotecas escritas en C++) utilizado en el desarrollo de la biblioteca GLSvE. Segundo, proporciona

flexibilidad para obtener la atenuación mediante `alDistanceMode()`, activada por defecto permitiendo usar hasta siete constantes diferentes, algunas representan funciones lineales y exponenciales. Y tercero, es posible obtener la frecuencia con que se percibe el efecto *Doppler* mediante `alDopplerFactor()`, activada por defecto o también puede calcularse mediante la fórmula 1. Se eligió para la creación de audio 2D que se encuentra dentro del sistema de sonido 3D, a la biblioteca SDL Mixer aunque no proporcione efecto *Doppler*, DII ni DIT (no son necesarios para el sonido 2D), soporta varios formatos de sonido como es el caso de *ogg* de utilidad para archivos grandes, forma parte del recubrimiento Tao Framework,

2.3.6. Formatos

En el epígrafe se exponen algunos de los formatos de audio empleados para la reproducción de sonidos en aplicaciones informáticas.

- **Wav:** Archivo de Microsoft (wav) es un tipo de RIFF (Formato de Intercambio de Recursos de Archivo), que se basó originalmente en el IFF (Formato de Intercambio de Archivo), diseñado para trabajar más eficientemente con procesadores Intel [Boer, 2002]. Permite guardar una pequeña cabecera, donde se informa de la frecuencia que se ha utilizado para la digitalización, el tamaño en bits de la muestra y el número de canales. Después de la cabecera, se almacena la información digital que representa el sonido sin ningún tipo de compresión. Este tipo de fichero ocupa mucho espacio de memoria, pero reproducirlo es muy sencillo [Jordi D. Gavaldà, 2008].
- **Ogg:** Este formato permite aprovechar mejor el espacio en disco, pero añade un coste en tiempo de cálculo para poder descomprimir el sonido [Jordi D. Gavaldà, 2008]. Es totalmente libre de patentes, la codificación de audio profesional y la tecnología de *streaming*⁷ con todos los beneficios es de código abierto [Xiph.Org, 2008]. Esto significa que el algoritmo de compresión se puede utilizar libremente por todos los productores de software, y es un recurso para el desarrollo y el lanzamiento de numerosas herramientas y librerías libres de derecho [Kio, 2008].

⁷Se refiere a un proceso continuo.

Se determinó el uso del formato (wav) para el sonido 3D por su fácil reproducción sin coste de descompresión, además de informar en su cabecera la frecuencia y tamaño de la señal, lo cual es de utilidad para llenar el buffer. Para el sonido 2D se determinó el formato (ogg) dado que ocupa poco espacio en disco, ventaja de utilidad para bandas sonoras ya que estas suelen ser grandes.

2.4. Funciones visuales y sonido posicional 3D

En el epígrafe se exponen muy brevemente las funciones visuales (Visión binocular y Agudeza visual), luego se realiza una revisión bibliográfica actual referida a la utilización y beneficios del sonido posicional 3D en las discapacidades visuales.

2.4.1. Funciones visuales

El Sistema Visual Humano (del inglés Human Visual System) (HVS) agrupa todos los elementos necesarios para lograr el sentido de la vista. La percepción visual es la capacidad de interpretar la información y el entorno, de los efectos de la luz visible que llega al ojo [V. Bruce, 2003]. Estas capacidades del HVS se les denominan funciones visuales [M. Cay Holbrook, 2000] [A. L. Corn, 1996]. La clasificación de las funciones visuales es diversa, en general sobresalen agudeza visual, campo visual, discriminación de color, adaptación a la oscuridad, la sensibilidad al contraste, visión monocular y binocular, convergencia y dominancia ocular [Dif, 2010][Psy, 2010].

- **Visión binocular:** La visión binocular permite percibir el mundo desde diferentes puntos de vista, debido a que se tiene una separación de alrededor de seis centímetros entre los dos ojos, lo que se conoce como disparidad retinal [Ian P. Howard, 1995]. Esta diferencia se interpreta como una medida de la profundidad y es la fuente elemental de percepción de profundidad [Coon, 2005].

- **Agudeza visual:** Es la capacidad para discriminar detalles finos de un objeto en el campo visual, es la inversa del ángulo desde el cual los objetos son contemplados [D. Cline, 1997]. La prueba estándar en el mundo para medir la agudeza visual es denominada ETDRS (del inglés Early Treatment Diabetic Retinopathy Study), que es una extensión del test original de *Snellen*, que se aplicaba desde 1862. La enfermedad visual en que se pone de manifiesto la agudeza visual es la Ambliopía. La misma consiste en la reducción de la agudeza visual en uno o ambos ojos debido a experiencias visuales anormales en la primera parte de la infancia, dando lugar a cambios funcionales de los centros visuales del cerebro [J. M. Holmes, 2006].

El sonido posicional 3D contribuirá al estudio de la agudeza visual y la visión binocular, se proponen estas y no otras dado que se pueden evaluar y entrenar mediante la localización de objetos en un espacio tridimensional.

2.4.2. Utilización del sonido posicional 3D en las discapacidades visuales

En la literatura científica consultada se ha encontrado una gran variedad de aplicaciones que proponen la incorporación de sonido posicional 3D, en aras de mejorar la interacción en la localización de objetos dentro de escenarios virtuales por parte de usuarios con discapacidades visuales:

1. En *AudioChile* [J. Sánchez, 2005a][J. Sánchez, 2006b] se centran en usar el sonido posicional 3D como interfaz para mejorar las habilidades de los usuarios en el reconocimiento de espacios virtuales a través del sonido. Los niños con discapacidades visuales tienen una percepción distinta del mundo que les rodea, frente a los niños que ven con normalidad. De modo que suelen desarrollar otros sentidos, como la audición, que les ayuda a orientarse y por tanto, a moverse por el espacio. Así, necesitan aprender a interpretar esos sonidos. En *AudioChile*, se proporcionan una serie de escenarios virtuales que recrean las diferentes regiones de Chile. Los niños tienen que navegar por estos entornos interactuando a través del sonido posicional 3D, teniendo que realizar determinadas actividades, como tomar ciertas decisiones dentro de tales mundos, véase la figura 10.



Figura 10: Distintas situaciones del entorno virtual AudioChile. [J. Sánchez, 2005a]

2. El estudio [Jaime H. Sánchez, 2001] indica que es posible estimular el desarrollo y uso de estructuras espaciales de aprendices discapacitados de la vista a través del uso de la interacción con mundos virtuales provistos de interfaces de audio posicional 3D. Este trabajo de investigación presenta los resultados de un estudio que involucra el diseño, desarrollo y evaluación de la usabilidad de un editor de software para aprendices débiles visuales, *VirtualAurea*, y una aplicación que se ejecuta sobre la base de este editor, el *Castillo Musical*. De esta forma, se intenta determinar el real impacto en el desarrollo y uso de estructuras mentales de tipo espacio-temporal, de intervenciones con el apoyo de un editor basado en sonido tridimensional. Los resultados cuali-cuantitativos indican que cuando se utiliza software basado en editores con interfaces de sonido espacial, en conjunto con una cuidadosa metodología de aplicación en niños discapacitados con tareas cognitivas, pruebas de representación de espacios y tiempos navegados inicialmente en mundos virtuales de audio, el aprendiz puede desarrollar estructuras mentales temporo-espaciales, diversificar y profundizar las experiencias de su uso y transferirlas a tareas cotidianas.

Como se ha señalado, *VirtualAurea* provee una plataforma genérica sobre la cual pueden construirse aplicaciones específicas. Para este estudio se construyó una aplicación llamada

Castillo Musical.

Para lograr la ubicación espacial, el juego emite sonidos provenientes de diversas posiciones del espacio. La idea es que en la medida que el infante recoja los trozos de una canción en orden correcto, obtendrá un mayor puntaje y logrará escuchar la canción completa. Para el trabajo de las nociones de espacialidad, cada objeto en el mapa emite un sonido que se ubica en una determinada posición y dirección dentro del espectro auditivo. Así, el niño puede ejercitar conceptos, buscando objetos que se encuentran a la izquierda, derecha, adelante o atrás, según sea la ubicación del sonido emitido.

Los autores concluyeron que es posible utilizar sonido posicional 3D para estimular en niños con discapacidades visuales la construcción mental de espacios navegados en un mundo virtual interactivo. [Jaime H. Sánchez, 2001].

3. El trabajo [Ch. Frauenberger, 2003] hace frente a modos alternativos para la interacción visual de personas discapacitadas a usar las computadoras. El objetivo del enfoque propuesto es aprovechar las capacidades de la audición humana para establecer interacción de usuarios con problemas visuales, mediante computadoras. Este artículo presenta un sistema de realidad virtual enfocado en sonido posicional 3D, que permite a los usuarios de computadoras explorar un entorno virtual sólo por su sentido de la audición. Utilizando representación de audio binaural a través de auriculares para proporcionar una simulación auténtica.

La localización de fuentes de sonido se puede mejorar de manera significativa, si se consideran movimientos de la cabeza. Por esta razón, el sistema presentado es utilizando un dispositivo de seguimiento de la cabeza que se monta en los auriculares, véase la figura 11.

4. AudioVida [Sánchez, 2008] hace hincapié en la aplicación de diferentes vías para el desplazamiento en un entorno virtual complejo, basado en estimulación por audio posicional 3D facilitando así llegar a un destino específico y localizar un objeto en particular. Para lograr este objetivo, el usuario debe analizar e interpretar el espacio virtual mediante la aplicación de nociones de la espacialidad y la temporalidad. Esto favorece la capacidad del niño para reconocer las diferentes posibilidades de desplazamiento, haciendo una representación mental del espacio virtual en movimiento, y elaborar estrategias que utilizan para navegar por el

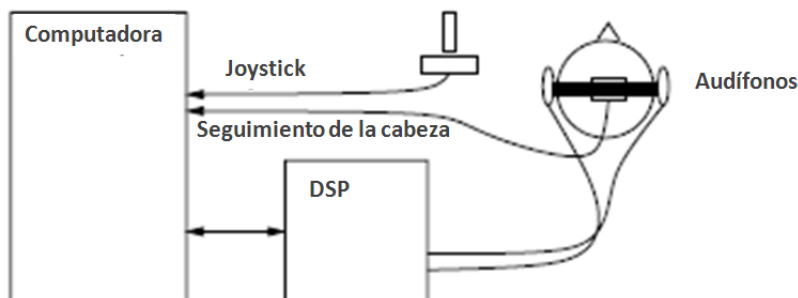


Figura 11: Sistema de sonido 3D con seguimiento de la cabeza. [Ch. Frauenberger, 2003]

medio ambiente, véase la figura 12. El usuario navega por el laberinto con la asistencia de efectos de sonido espacial que proporcionan referencias de los objetos, como paredes, puertas y elementos con los que se puedan interactuar. Esto los motiva a caminar a través del laberinto virtual como lo harían físicamente.



Figura 12: Objeto emisor de sonido. [Sánchez, 2008]

Este software fue desarrollado con lenguaje de programación C++, usando las bibliotecas OpenGL y DirectSound para gráficos y sonidos respectivamente [J. Sánchez, 2005b].

En toda la bibliografía consultada en el presente epígrafe se pone de manifiesto la premisa sobre la efectividad del sonido posicional 3D para mejorar la interacción de usuarios de poca

Autor	Trabajo	Objetivo	Tecnología de sonido 3D
[J. Sánchez, 2005a]	3D Sound interactive environments for problem solving	Reconocimiento de objetos	Posicional 3D
[Jaime H. Sánchez, 2001]	Cognición de ciegos con ambientes virtuales basados en sonido	Reconocimiento de objetos	Posicional 3D
[Ch. Frauenberger, 2003]	3d audio interfaces for the blind	Aprovechar las capacidades de la audición humana para establecer interacción con objetos virtuales	posicional 3D
[Sánchez, 2008]	User-centered technologies for blind children	Localización de objetos en el espacio	posicional 3D

Tabla 3: Aplicación de sonido posicional 3D en las discapacidades visuales

o ninguna visión con aplicaciones de realidad virtual que apoyan el uso de terapias en este sentido.

Se evidencia en las investigaciones presentadas que las incursiones en este campo son prometedoras. En la tabla 3 se relacionan a manera de resumen trabajos sobresalientes en el área, el objetivo que persiguen y la tecnología de sonido 3D. Por lo que puede aportarse nuevos resultados a este campo de aplicación, con la combinación de visión estereoscópica y sonido posicional 3D, como contribución para mejorar la sensopercepción visual de objetos en la escena para su uso en el estudio de funciones visuales.

2.5. Descripción del problema

En el epígrafe se justifica la necesidad de un sistema de sonido 3D en la biblioteca GLS_{Ve}, luego se explican las funcionalidades necesarias que se deben brindar y se concluye con la mención de un aporte adicional al trabajo.

2.5.1. Justificación de la incorporación del sistema de sonido posicional 3D a la GLS_{Ve}

La biblioteca GLS_{Ve} todavía no tenía incorporada funcionalidades de sonido, ni de forma general ni específicamente para mejorar la sensopercepción en la evaluación y entrenamiento de funciones visuales de los usuarios, así como tampoco para ambientes de realidad virtual con seguimiento posicional. Después de realizar una revisión bibliográfica en 2.4.2 quedaron reflejadas las potencialidades del uso del sonido 3D para mejorar en usuarios con discapacidades visuales la navegación e interacción en las aplicaciones con fines de entrenamiento.

2.5.2. Funcionalidades

Una vez realizada la investigación de cómo pueden influir las tecnologías de sonido 3D en usuarios con discapacidades visuales, se han definido un grupo de funcionalidades necesarias para la GLS_{Ve}, que debe brindar el sistema de sonido 3D para que pueda utilizarse en el desarrollo de actividades que permita la evaluación y entrenamiento de funciones visuales, así como también en ambientes de realidad virtual con seguimiento posicional. Finalmente se realizó una funcionalidad adicional que aporta un valor agregado a la GLS_{Ve} para las aplicaciones que estudian las funciones visuales.

Las funcionalidades necesarias que se deben brindar son las siguientes:

- Permitir la reproducción de sonido 2D en las aplicaciones virtuales que evalúan y entrenan las funciones visuales, así como también en ambientes de realidad virtual con

seguimiento posicional. Dentro de esta funcionalidad se debe brindar opciones de (*pause*, *stop*, *resume*, *ajuste de volumen*, efectos de: *encendido* y *apagado* del sonidos por instantes de tiempos determinados).

- Simular sonido 3D a partir del puntero de GLS_{Ve} en las aplicaciones que evalúan y entrenan las funciones visuales. A partir de los movimientos realizados por el usuario en la escena y que se tenga en cuenta la posición en tiempo real de las fuentes de sonido respecto al puntero.
- Generar sonido 3D en ambientes de realidad virtual con seguimiento posicional. Esta funcionalidad deberá permitir que los usuarios que se muevan libremente por el sistema logren percibir el sonido correctamente por el oído que le corresponda en la medida de su posición, orientación y punto de mira respecto a las fuentes de sonido que puedan estar activas.

Funcionalidad adicional:

- Permitir la visualización de video estereoscópico, útil como alternativa de acomodación visual diseñada para las actividades que estudian las funciones visuales.

Contar con una herramienta que cumpla estas funcionalidades, permitirá contribuir a mejorar la sensopercepción en las aplicaciones que evalúan y entrenan funciones visuales como la agudeza y vision binocular. El proyecto HDS_{VE} cuenta con la biblioteca GLS_{Ve} que se pretende usar para estos fines. La biblioteca ha sido desarrollada en conjunto por los investigadores del proyecto, en su concepción inicial no tenía incorporada funcionalidades de sonido, ni de forma general ni específicamente para mejorar la sensopercepción en el estudio de funciones visuales de los usuarios, así como tampoco para ambientes de realidad virtual con seguimiento posicional. Por lo que fue necesario incorporarlas para asumir estos nuevos campos de aplicación.

2.6. Soluciones técnicas

Se han definido en el epígrafe 2.5 las funcionalidades necesarias para contribuir a mejorar la sensopercepción en el estudio de funciones visuales, también para la correcta simulación del sonido 3D en ambientes de realidad virtual con seguimiento posicional y una funcionalidad adicional para mejorar la acomodación visual del usuario en los prototipos de actividades diseñados para el estudio de funciones visuales. Se presenta inicialmente en este epígrafe la biblioteca GLS_{Ve}. A continuación se describen las funcionalidades incorporadas en GLS_{Ve} y por último se presentan las actividades realizadas que permiten validar el sistema de sonido posicional 3D incorporado.

2.6.1. Graphics Library for Stereoscopic Vision (GLS_{Ve})

GLS_{Ve} es una biblioteca estructurada en clases que permite representar objetos a distintas profundidades, que se observen por delante del monitor, en este, o detrás, según los paralajes correspondientes, brinda una interfaz fácil e intuitiva para su uso y tiene las siguientes características:

- La visualización de la escena puede ser en modo monoscópico.
- La visualización de la escena puede ser en modo estereoscópico según las técnicas: visión paralela, visión cruzada, anaglifo, polarización y obturación (en los formatos de representación entrelazado horizontal, *alternated images* y *quad buffering*).
- Brinda un puntero 3D, útil para la selección y manipulación de los objetos de la escena a distintas profundidades.
- Brinda sonido 2D para motivar al usuario con música de fondo y sonidos de eventos en la realización de actividades.
- Permite sonido 3D para ayudar al usuario moverse en la escena a través del puntero 3D logrando mayor nivel de inmersión y comprensión del ambiente.

- Permite sonido 3D para ambientes de realidad virtual con seguimiento posicional.
- Una vez compilada la biblioteca es posible añadir nuevas primitivas gráficas.
- Permite realizar transformaciones de *pan*, *zoom* y *parallax*.
- El usuario puede gestionar las funciones implementadas para hacer su propio código.

La biblioteca ha sido desarrollada en el lenguaje C#, utilizando el recubrimiento Tao Framework para emplear las herramientas OpenGL, OpenAL y SDL Mixer, las cuales brindan facilidades para implementar las técnicas estereoscópicas y el sistema de sonido 3D. Para más detalles de las funcionalidades, arquitectura y características de diseño de GLSve consultar la Memoria Colectiva en su Capítulo 3.

2.6.2. Funcionalidades incorporadas a la GLSve:

La biblioteca GLSve no tenía incorporada funcionalidades de sonido, ni de forma general ni específicamente para mejorar la sensopercepción en la evaluación y entrenamiento de funciones visuales de los usuarios, así como tampoco para la simulación en ambientes de realidad virtual con seguimiento posicional, ni funcionalidad de visualización de video estereoscópico. Se describen a continuación los aportes realizados.

Incorporación de sonido 2D

Para crear sonido 2D en las aplicaciones que se desarrollen con la GLSve, inicialmente se parte de la lectura de formatos de audio usados para música de fondo o sonidos de eventos, para esta solución se dio soporte a (wav) y a (ogg). Se crea una instancia de **Sound2D** para cada uno de los sonidos que se deseen almacenar en **SceneSound2D** ubicada en la controladora de sonidos **SoundManager**, en el constructor de **Sound2D** se importa el fichero de audio mediante **LoadFile()**. Luego se gestionan los canales con **AllocateChannels()** a los cuales se le adjuntan las fuentes de sonido disponibles, la tarea de esta función es contabilizar la creación y asignación de canales por instancias de sonidos en **Sound2D**, aportando la ventaja que

se aprovecha eficientemente los recursos de la máquina, dado que la generación de canales está en dependencia del número de archivos de sonido que se importan. Una vez que estén las fuentes en sus respectivos canales es posible reproducir el que se desee mediante `Play()` accediendo a `SceneSound2D`, internamente para la reproducción se contabiliza el canal, la fuente y si se desea una reproducción cíclica o no.

Las fuentes de sonido solo podrán ser pausadas, paradas o continuadas a partir del punto donde se hallan quedado solo si estén previamente en reproducción. Igualmente para poner en ejecución alguna de las funcionalidades antes comentadas se deberá acceder a `SceneSound2D` para indicar la fuente. A los sonidos se les puede modificar su volumen en cualquier momento mediante `AjustVolume()` con un valor entre 0 y 128. Se ofrecen efectos de encendido y apagado del sonido con `FadeIn()` y `FadeOut()` a partir de un valor de tiempo determinado, estos efectos suelen ser de utilidad en las aplicaciones para activar o desactivar un sonido después de transcurrido un tiempo que se halla indicado como aviso de un determinado evento.

La biblioteca SDL Mixer usada para el cumplimiento de este aporte tiene una limitante en la gestión de canales de música. Solo admite un solo canal activo, lo cual presenta una desventaja desde el punto de vista de rendimiento dado que si se desea reproducir más de una música en la aplicación se deberá destruir la actual e importar el nuevo fichero. Esta deficiencia fue resuelta con la funcionalidad antes comentada `AllocateChannels()`.

Simulación de sonido 3D para prototipos de actividades que estudian funciones visuales

Para crear sonido 3D en las aplicaciones que se desarrollen con la GLSve para el estudio de funciones visuales, inicialmente se parte de la lectura de un formato de audio mono, para esta solución se dio soporte a (wav) porque el mismo no es necesario descomprimirlo, además informa en su cabecera sobre la frecuencia y tamaño del fichero, información de utilidad para el posterior llenado del *buffer*. Se crea una instancia de `Sound3D` para cada una de las fuentes que se deseen almacenar en `SceneSound3D` ubicada en la controladora de sonidos `SoundManager`, en el constructor de `Sound3D` se importa la fuente de audio mediante `LoadFile()`, también se manda a generar el *buffer* y la fuente con `GenBuffers()` y `GenSources()` respectivamente

y por último se llena el *buffer* con `FillBuffer()`. Las fuentes importadas se les deberá indicar como propiedad su posición en el espacio tridimensional además de su velocidad. Este aporte considera que pueden haber una o más fuentes en la escena con posición, la cual se actualizará con `SetSource` en tiempo real si está en constante movimiento, la actualización es necesaria para tener una constancia de la fuente en relación al oyente. El oyente en este caso es el puntero 3D de la GLS_{Ve}, objeto mediante el cual el usuario interactuará con la aplicación. Para obtener sonido 3D se toma como posición del oyente la ubicación del puntero 3D. Para conocer donde está el oyente y su orientación respecto a la fuente se utiliza la función `SetListener()` ubicada en la controladora `SoundManager`, la cual realiza cálculos vectoriales entre el vector horizontal del observador, en este caso el puntero 3D y un vector perpendicular a la pantalla para obtener otro vector ortogonal a ellos que se denomina orientación del puntero, con este vector resultante y el vector perpendicular a la pantalla se obtiene otro vector ortogonal denominado punto de mira del puntero 3D. De esta forma el usuario es capaz de sentir los sonidos provenientes de distintas posiciones y ser escuchados en la medida correcta por el oído que le corresponda. Logrando con este aporte contribuir a la sensopercepción en las aplicaciones vinculadas al estudio de las funciones visuales.

Las fuentes de sonido también podrán ser pausadas, paradas o continuadas a partir del punto donde se hallan quedado solo si están previamente en reproducción. Igualmente para poner en ejecución alguna de estas funcionalidades se deberá acceder a `SceneSound3D` para indicar la fuente. A los sonidos se les puede modificar su volumen en cualquier momento mediante `AjustVolume()` con un valor entre 0 y 1. Para obtener efectos de cambios de tonalidad se puede usar la función `AjustPitch(int pitch)`.

Generación sonido 3D en ambientes de realidad virtual con seguimiento posicional

Para crear sonido 3D en ambientes de realidad virtual con seguimiento posicional desarrollados con la GLS_{Ve}, inicialmente se parte de la lectura de un formato de audio mono, para esta solución se dio soporte a (wav) porque el mismo no es necesario descomprimirlo, además informa en su cabecera sobre la frecuencia y tamaño del fichero, información de utilidad para el posterior llenado del *buffer*. Se crea una instancia de `Sound3D` para cada una de las fuentes que se deseen almacenar en `SceneSound3D` ubicada en la controladora de sonidos `SoundManager`, en el constructor de `Sound3D` se importa la fuente de audio mediante `LoadFile()`, también

se manda a generar el *buffer* y la fuente con `GenBuffers()` y `GenSources()` respectivamente y por último se llena el *buffer* con `FillBuffer()`. Las fuentes importadas se les deberá indicar como propiedad su posición en el espacio tridimensional además de su velocidad. Este aporte considera que pueden haber una o más fuentes en la escena con posición, la cual se actualizará con `SetSource` en tiempo real si está en constante movimiento, la actualización es necesaria para tener una constancia de la fuente en relación al oyente. El oyente en este caso es un usuario con un sistema de seguimiento posicional, pudiendo interactuar en este caso personalmente con la aplicación. Para obtener sonido 3D se toma como posición del oyente la ubicación obtenida desde el sistema de seguimiento. Para conocer donde está el oyente y su orientación respecto a la fuente se utiliza la función `SetListener()` ubicada en la controladora `SoundManager`, la cual realiza cálculos vectoriales entre el vector horizontal del observador y un vector perpendicular a la pantalla para obtener otro vector ortogonal a ellos que se denomina orientación del oyente en el ambiente virtual, con este vector resultante y el vector perpendicular a la pantalla se obtiene otro vector ortogonal denominado punto de mira del oyente. De esta forma el usuario es capaz de sentir dentro del ambiente virtual los sonidos provenientes de distintas posiciones y ser escuchados en la medida correcta por el oído que le corresponda.

Las fuentes de sonido también podrán ser pausadas, paradas o continuadas a partir del punto donde se hallan quedado solo si están previamente en reproducción. Igualmente para poner en ejecución alguna de estas funcionalidades se deberá acceder a `SceneSound3D` para indicar la fuente. A los sonidos se les puede modificar su volumen en cualquier momento mediante `AjustVolume()` con un valor entre 0 y 1. Para obtener efectos de cambios de tonalidad se puede usar la función `AjustPitch(int pitch)`.

Para un mayor entendimiento del sistema de sonido posicional 3D, véase la figura 13.

Visualización de video estereoscópico

En aras de ofrecer mayores prestaciones a la GLSVe se desarrolló también un componente de visualización de video estereoscópico. Visualiza de forma diferenciada para cada proyección el video que le corresponde. Este aspecto que no existía en la biblioteca aporta un valor agregado posibilitando mejora en la acomodación visual de los niños para adaptarse al modo

de visualización presentado en los prototipos diseñados para el estudio de funciones visuales, también de esta forma se brinda información sobre las actividades que realizará.

Para visualizar los videos se parte inicialmente de importar el par de videos en formato *avi*. **VideoPlayer** carga el video izquierdo y derecho, **VideoFrame** crea una instancia de **VideoReader**, esta permite abrir el archivo a través de **Open()**, extraerle los fotogramas mediante **GetFrame()**, cuya textura es guardada en una dirección de memoria, luego a través de **ExportStream()** se separa el audio del video y se guarda este primero en memoria física. Finalmente se cierra el fichero de video con **Close()**. **VideoFrame** se hace responsable de cargar los fotogramas con **LoadGLTexture()** del video izquierdo y derecho que fueron extraídos anteriormente en la clase **VideoReader**. Mediante **DrawHead()** se proyecta cada fotograma, esta función es heredada de la clase **Textura** de la GLSve, en **VideoFrame** también se implementa la reproducción del audio **PlaySound()** del componente de sonido 3D incorporado a la GLSve como aporte principal del trabajo. Finalmente **VideoPlayer** visualiza para cada proyección la secuencia de fotogramas que fueron cargados y proyectados por **VideoFrame**. En la misma se sincroniza la reproducción de fotogramas con el sonido mediante **CurrentPosition()**, con **DrawLeftEye()** se visualiza el video para la proyección izquierda y con **DrawRightEye()** para la proyección derecha, ambas funciones son heredadas de **Model**.

Resumiendo: Para obtener la solución dado que el control del renderizado lo debe tener el núcleo de la GLSve, se descomprimió cada video en una cadena de fotogramas para ser proyectados posteriormente sobre el plano, esto lógicamente deja muda la proyección, por tanto aquí es donde entra el componente de sonido 3D realizado, su función es tomar el audio extraído de los videos y reproducirlo, pero para lograr una correcta sincronización entre audio y fotogramas visualizados, se realizó una funcionalidad adicional que controla este aspecto según el procesamiento de la PC, como suele suceder en los reproductores profesionales cuando existe algún desfasaje.

Para un mayor entendimiento del sistema de visualización de video estereoscópico, véase la figura 14.

2.6.3. Prototipos de actividades

Con el cumplimiento en GLSVe de las funcionalidades necesarias, se logró proveer en los prototipos de actividades diseñados para el estudio de funciones visuales la sensopercepción acústica, se planteó el uso del componente de video estereoscópico diseñado para iniciar las actividades con el objetivo de preparar la visión del usuario antes de comenzar.

- **Búsqueda de objetos parcialmente ocultos por paralaje:** Se presenta una escena compleja donde se debe localizar varios objetos ayudado por el sonido posicional 3D que emiten los mismos. Estos objetos están solapados por otros elementos más cercanos, de forma que, por efecto del paralaje, se ve con suficiente detalle con un ojo y sin embargo permanecen prácticamente ocultos con el otro. Se utiliza el ratón en un espacio 3D para localizarlos. La actividad permite ejercitar la agudeza visual, la visión binocular y la convergencia. No se han encontrado herramientas similares que intenten el uso de sonido posicional 3D en este tipo de ejercicios, véase la figura 15.
- **Trazado de contornos 3D:** Se debe trazar un contorno tridimensional seleccionando con el ratón puntos a distintas profundidades, ayudado por el sonido posicional 3D que emiten los mismos, hasta descubrir una figura oculta, posibilitando que el sujeto se sienta inmerso en el escenario virtual mejorando la sensopercepción del mismo. Se utiliza por tanto el ratón en un espacio 3D. Esta actividad permite ejercitar la agudeza visual, la visión binocular y la convergencia, pero además de ello incorpora la posibilidad de medir la estereopsis independiente para cada ojo, véase la figura 16.
- **Búsqueda de objetos con penalización del ojo sano:** Se presenta una escena compleja donde el paciente tiene que seleccionar distintos elementos que emiten sonido posicional 3D percibidos nítidamente para el ojo vago y no visualizados correctamente con el ojo sano. Para ello, a estos objetos se les aplica un filtro de transparencia en una de las proyecciones. El niño utiliza el ratón en un espacio 3D. La actividad permite ejercitar la agudeza visual, la visión binocular y la convergencia, véase la figura 17.

Con el cumplimiento en GLSVe de las funcionalidades necesarias puestas en práctica a través de los prototipos de actividades antes descritos, se logró conseguir el objetivo del trabajo

presentado. Las actividades diseñadas cuentan con sonido 2D para motivar los pacientes en la realización de las mismas con todas las funcionalidades descritas en 2.6.2, así como también aportan sensación realista referente al sonido 3D, dado que este ayuda al usuario en la comprensión de la escena. El paciente interactúa mediante el puntero 3D localizando objetos en diferentes posiciones del espacio. En la medida que el paciente se aleja o se acerca de los objetos emisores, el sonido 3D le indica aproximadamente cual es su posición respecto al objeto emisor. Finalmente quedó validado el sistema de sonido posicional 3D, con el uso de auriculares en los prototipos de actividades diseñados para la evaluación y entrenamiento de funciones visuales.

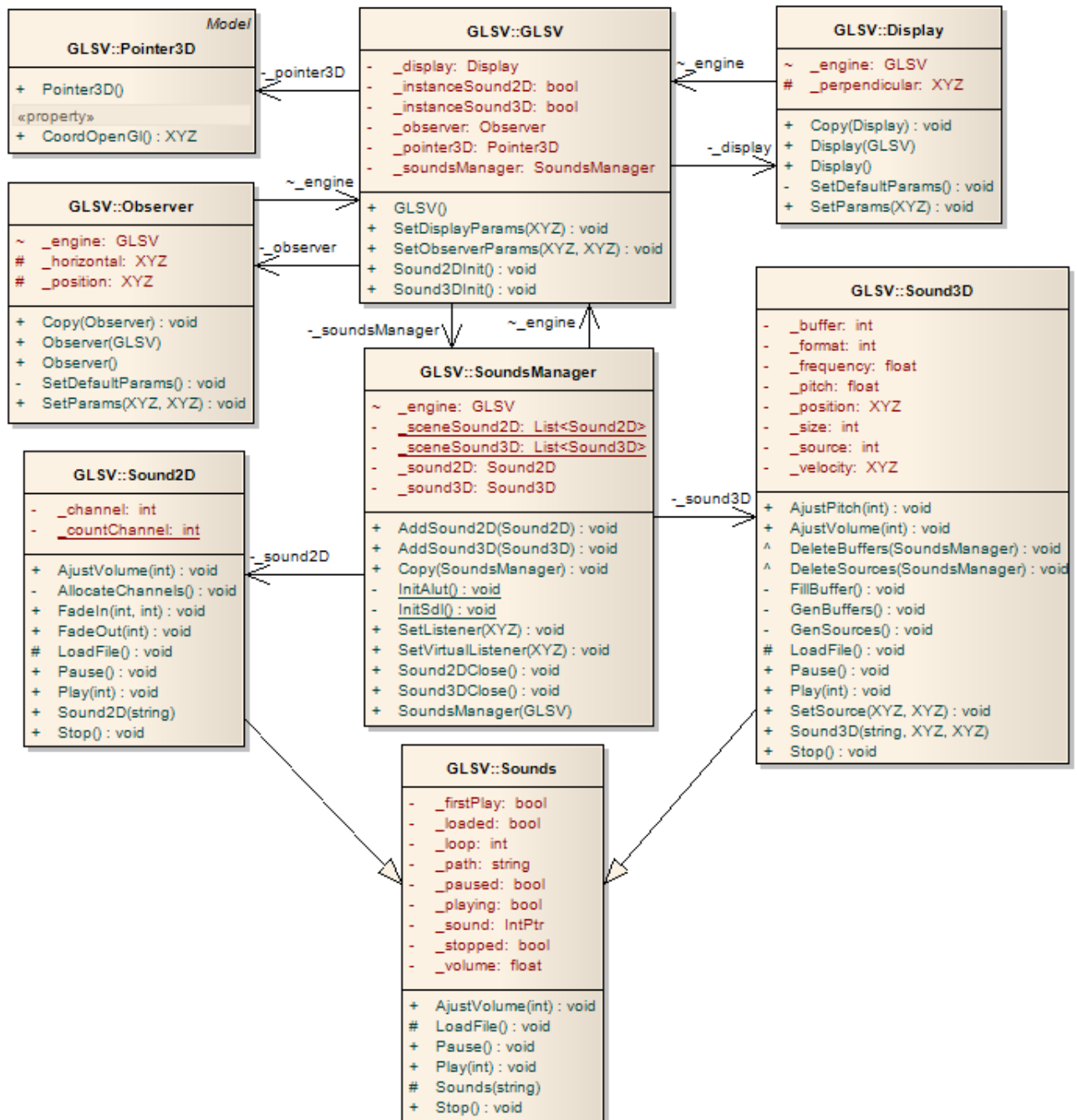


Figura 13: Diagrama de clases del sistema de sonido posicional 3D.

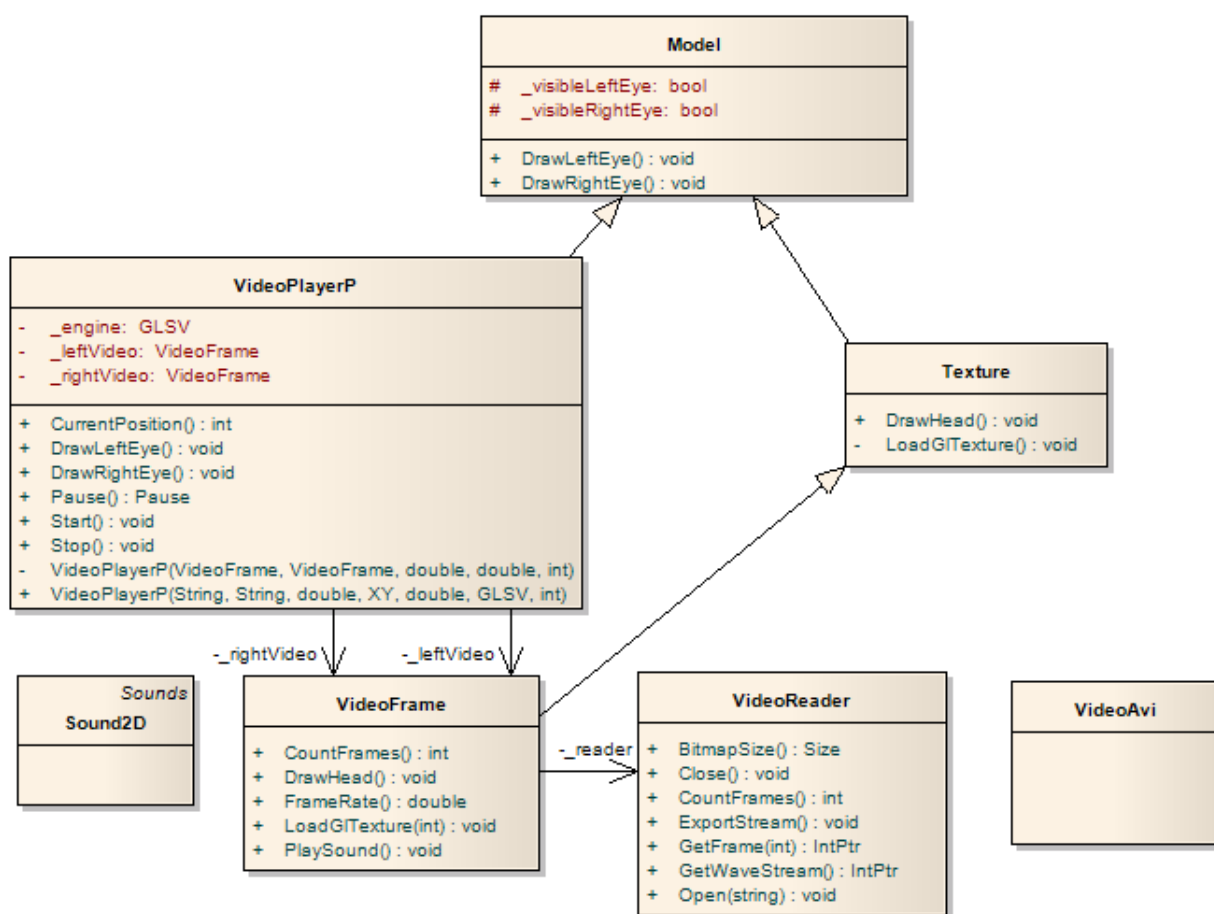


Figura 14: Diagrama de clases del sistema de visualización de video estereoscópico.

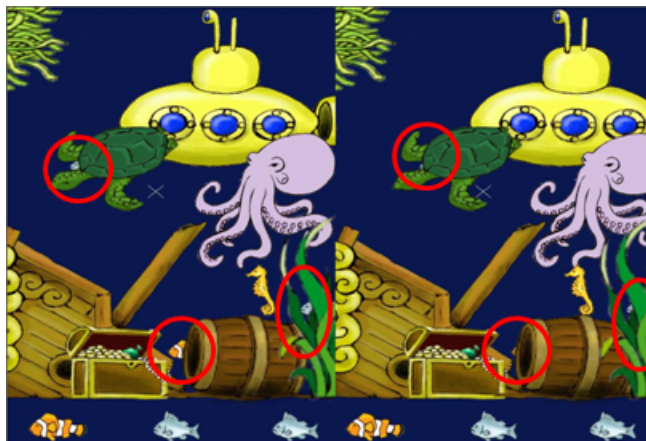


Figura 15: Búsqueda de objetos parcialmente ocultos por paralaje.

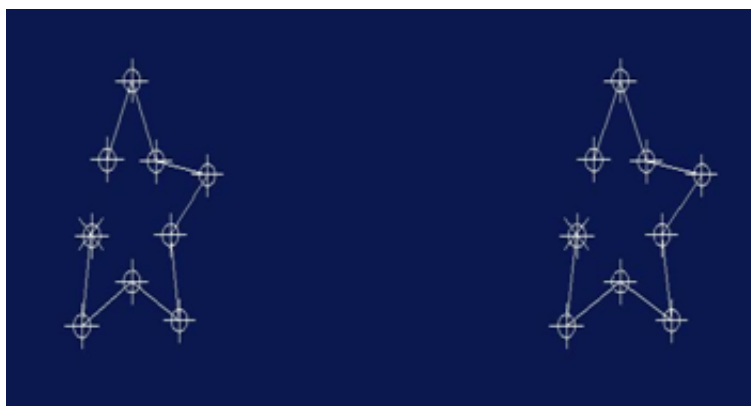


Figura 16: Trazado de contornos 3D.



Figura 17: Búsqueda de objetos con penalización del ojo sano.

3. Conclusiones finales

En el transcurso de la investigación realizada se han llegado a las siguientes conclusiones:

- La biblioteca GLS_{Ve} con del sistema de sonido posicional 3D se puede usar en aplicaciones oftalmológicas para el estudio de funciones visuales.
- El sistema de sonido posicional 3D con auriculares es una solución adecuada y económica para la percepción espacial.
- Con el cumplimiento en GLS_{Ve} de las funcionalidades necesarias, se logró proveer en las actividades la sensopercepción acústica, además de su usabilidad en ambientes de realidad virtual con seguimiento posicional.
- Se logró la visualización de video estereoscópico con sonido ambiental 2D para los prototipos de actividades diseñados.

4. Recomendaciones

- Comenzar en conjunto con los médicos un proceso de ensayo con pacientes de acuerdo a los procedimientos establecidos, en aras de realizar un análisis estadístico sobre el uso de visualización estereoscópica combinada con sonido posicional 3D para el estudio de funciones visuales.
- Desarrollar nuevas actividades con incorporación de sonido posicional 3D ya sea para la evaluación y entrenamiento de funciones visuales u otras.
- Investigar sobre posibles aplicaciones del sonido posicional 3D en el estudio de otros elementos de percepción visual.

5. Publicaciones derivadas

Durante la investigación presentada se realizaron las siguientes publicaciones y presentaciones:

- Informática 2011⁸

AUTORES: L. Pupo, Y. Cabrera, S. Martín, R. Rubio

TÍTULO: Análisis de GLSve como biblioteca para estudios sobre el Sistema Visual Humano

CONGRESO: VIII Congreso Internacional de Informática en Salud en la XIV Convención y Feria Internacional Informática

PUBLICACIÓN: Actas del congreso

LUGAR DE CELEBRACIÓN: La Habana

AÑO: 2011

- UCIENCIA 2010:

AUTORES: L. Pupo, S. Martín, Y. Cabrera, R. Rubio

TÍTULO: Biblioteca de clases para la visión estereoscópica

PUBLICACIÓN: Actas del congreso

LUGAR DE CELEBRACIÓN: UCI, Cuba

AÑO: 2010

- XI Congreso Internacional de Interacción Persona-Ordenador:

AUTORES: S. Martín, L. Pupo, Y. Cabrera, R. Rubio, P. Aduriz

TÍTULO: Aplicación de las técnicas de realidad virtual en el tratamiento de la ambliopía

PUBLICACIÓN: Actas del congreso

LUGAR DE CELEBRACIÓN: Valencia

AÑO: 2010

⁸Aprobado para presentar en ponencia presencial en fecha de impresión de este documento

Referencias

- [Ope, 2005] (2005). *OpenAL 1.1 Specification and Reference*. Creative Labs, Inc, 1.1 edition. Consultado: Noviembre, 2010.
- [Kio, 2008] (2008). Kioskea.net. <http://es.kioskea.net/contents/audio/ogg-vorbis.php3>. Consultado: Enero, 2011.
- [Dif, 2010] (2010). Different visual functions. <http://www.leatest.fi/en/eyes/visfunct.html>. Consultado: Diciembre, 2010.
- [Psy, 2010] (2010). Psychology dictionary. <http://www.tuition.com.hk/psychology/v.htm>. Consultado: Diciembre, 2010.
- [A. L. Corn, 1996] A. L. Corn, A. J. K. (1996). *Foundations of low vision: clinical and functional perspectives*. AFB Press, American Foundation for the Blind. Consultado: Diciembre, 2010.
- [A. L. M. Pambakian, 2004] A. L. M. Pambakian, S. K. Mannan, T. L. H. C. K. (2004). Saccadic visual search training: a treatment for patients with homonymous hemianopia. *J Neurol Neurosurg Psychiatry*, 75(10):1443–1448. Consultado: Noviembre, 2010.
- [Bartlett, 2009] Bartlett, B. (2009). *Practical Recording Techniques*. Elsevier's Science and Technology Rights, 5 edition. Consultado: Octubre, 2010.
- [Bayona, 2007] Bayona, G. H. (2007). *Psicopatología básica*. Pontificia Universidad Javeriana, 4 edition. Consultado: Noviembre, 2010.
- [Begault, 2000] Begault, D. R. (2000). *3-D Sound for Virtual Reality and Multimedia*. Ames Research Center, Moffett Field, California. Consultado: Noviembre, 2010.
- [Boer, 2002] Boer, J. (2002). *Game Audio Programing*, volume working paper 9004 of *Charles River Media Game Development*. Charles River Media, Hingham, Mass. Consultado: Noviembre, 2010.
- [C. A. Ortiz, 2004] C. A. Ortiz, J. d. C. P. (2004). *Manual de prácticas de psicología*. 1 edition. Consultado: Noviembre, 2010.

- [Calvo, 1986] Calvo, C. H. (1986). *Teoría y Práctica de la Propaganda Contemporánea*. Andrés Bello, 1 edition. Consultado: Noviembre, 2010.
- [Ch. Frauenberger, 2003] Ch. Frauenberger, M.Ñ. (2003). 3d audio interfaces for the blind. *International Conference on Auditory Display*. Consultado: Diciembre, 2010.
- [Cid, 2006] Cid, M. S. (2006). *Capacidad comunicativa del sonido envolvente 5.1 en la producción publicitaria radiofónica en España*. Tesis doctoral, Universidad Rey Juan Carlos, España. Consultado: Enero, 2011.
- [Coon, 2005] Coon, D. (2005). *Fundamentos de Psicología*. International Thomson, tenth edition. Consultado: Diciembre, 2010.
- [D. Cline, 1997] D. Cline, H. Hofstetter, J. d. G. (1997). *Diccionario de la ciencia visual*. Butterworth-Heinemann, 4 edition. Consultado: Diciembre, 2010.
- [D. Schroder, 2007] D. Schroder, T. Lentz, M. V. I. A. (2007). Virtual reality system with integrated sound field simulation and reproduction. *EURASIP Journal on Advances in Signal Processing*. Consultado: Septiembre, 2010.
- [E. S. B., 2003] E. S. B., M. C. G. (2003). *Física general*. Casa Editorial Mares - Editorial Tébar, 32a edition. Consultado: Febrero, 2011.
- [Fernandez, 2006] Fernandez, J. L. (2006). *Introducción a Sonido en video Juegos (1ª parte)*. codepixel.
- [G. Pocock, 2005] G. Pocock, C. D. R. (2005). *Fisiología Humana: La base de la Medicina*. Masson, S.A, Travessera de Gràcia, 17-21 - Barcelona (España), 2 edition. Consultado: Enero, 2011.
- [García, 2005] García, A. (2005). Sistema de sonido 3d. Trabajo de electroacústica, Escuela Superior de Ingenieros de Bilbao. Consultado: Enero, 2011.
- [Henderson, 2010] Henderson, T. (2010). *The Physics Classroom, The Nature of a Sound Wave*. University of Illinois, thenderson@glenbrook.k12.il.us. Consultado: Octubre, 2010.
- [Hiebert, 2007] Hiebert, G. (2007). *OpenAL Programmer's Guide*. Creative Technology Limited, version 1.0 and 1.1 edition. Consultado: Noviembre, 2010.

- [I. Assenmacher, 2004] I. Assenmacher, T. Kuhlen, T. L. M. V. (2004). Integrating real-time binaural acoustics into vr applications. In *In Eurographics Symposium on Virtual Environments*, pages 129–136. Consultado: Septiembre, 2010.
- [Ian P. Howard, 1995] Ian P. Howard, B. J. R. (1995). *Binocular vision and stereopsis*. 29. Oxford University Press, first edition. Consultado: Diciembre, 2010.
- [Isbert, 1998] Isbert, A. C. (1998). *Diseño acústico de espacios arquitectónicos*. Politecnos Series. UPC, 1 edition. Consultado: Enero, 2011.
- [J. L. Fernandez, 2006] J. L. Fernandez, M. L. (2006). *Introducción a Sonido en video Juegos (2ª parte)*. codepixel.
- [J. M. Holmes, 2006] J. M. Holmes, M. P. C. (2006). Amblyopia. Technical Report 367, Department of Ophthalmology, Mayo Clinic College of Medicine, Rochester, USA. Consultado: Diciembre, 2010.
- [J. Sánchez, 2006a] J. Sánchez, M. E. (2006a). Aprendizaje de ciencias a través de audio en niños ciegos. 2:11–21. Consultado: Diciembre, 2010.
- [J. Sánchez, 2005a] J. Sánchez, M. S. (2005a). 3d sound interactive environments for blind children problem solving. *Department of Computer Science, University of Chile, Chile*. Consultado: Noviembre, 2010.
- [J. Sánchez, 2006b] J. Sánchez, M. S. (2006b). Three-dimensional virtual environments for blind children. *CyberPsychology and Behavior*, 9(2):200–206. Consultado: Enero, 2011.
- [J. Sánchez, 2005b] J. Sánchez, N. B. (2005b). Modeling audio-based virtual environment for children visual disabilities. *World Conference on Educational*. Consultado: Enero, 2011.
- [Jaime H. Sánchez, 2001] Jaime H. Sánchez, María L. Jorquera, E. L. M. E. E. V. (2001). Cognición de ciegos con ambientes virtuales basados en sonido. Consultado: Diciembre, 2010.
- [Jordi D. Gavalda, 2008] Jordi D. Gavalda, H. T.Ñ. (2008). *Sonido, Interacción y Redes*. UOC, www.uoc.edu. Consultado: Octubre, 2010.

- [Kalat, 2008] Kalat, J. W. (2008). *Introduccion to Psychology*. 8 edition. Consultado: Octubre, 2010.
- [Kalat, 2009] Kalat, J. W. (2009). *Biological Psychology*. 10 edition. Consultado: Octubre, 2010.
- [Lantinga, 2001] Lantinga, S. (2001). *SDL Library Documentation*. SDL Documentation Project, 1.2 edition. Consultado: Noviembre, 2010.
- [M. Cay Holbrook, 2000] M. Cay Holbrook, A. J. K. (2000). *Foundations of Education: Instructional strategies for teaching children and youths with visual impairments*, volume 2. AFB Press, American Foundation for the Blind, 2 edition. Consultado: Diciembre, 2010.
- [M. Lumbreras, 2010] M. Lumbreras, M. A. (2010). *Design of Virtual Environments Using Multi-screen, 3D Sound and Stereoscopic Systems*. Thesis doctoral, Ciencias de la Computación. Departamento de Computación, Electrónica y Mecatrónica, Escuela de Ingeniería, Universidad de las Américas Puebla, Santa Catarina Mártir, San Andrés Cholula, Puebla. CP 72820. México. Consultado: Noviembre, 2010.
- [M. Ohuchi, 2005] M. Ohuchi, Y. Iwaya, Y. S. T. M. (2005). Training effect of a virtual auditory game on sound localization ability of the visually impaired. *05-Eleventh Meeting of the International Conference on Auditory Display*. Consultado: Diciembre, 2010.
- [McCuskey, 2003] McCuskey, M. (2003). *Beginning Game Audio Programming*. Stacy L. Hiquet, Premier Press, a division of Course Technology 25 Thomson Place Boston, MA 02210. Consultado: Noviembre, 2010.
- [P. Turcan, 2004] P. Turcan, M. W. (2004). *Fundamentals of Audio and Video Programming for Games*. Microsoft Press A Division of Microsoft Corporation One Microsoft Way Redmond, Washington 98052-6399. Consultado: Noviembre, 2010.
- [Polat, 2009] Polat, U. (2009). Making perceptual learning practical to improve visual functions. *Vision Research*, 49(21):2566–2573. Perceptual Learning.
- [Sánchez, 2008] Sánchez, J. (2008). User-centered technologies for blind children. *An Interdisciplinary Journal on Humans in ICT Environments*, 4(2):96–122. Consultado: Diciembre, 2010.

- [T. Lentz, 2006] T. Lentz, I. Assenmachery, M. V. T. K. (2006). Precise near-to-head acoustics with binaural synthesis. *Journal of Virtual Reality and Broadcasting*, 3(2). Consultado: Septiembre, 2010.
- [T. M. Fay, 2004] T. M. Fay, S. Selfon, T. J. F. (2004). *DirectX 9 Audio Exposed: Interactive Audio Development*. Wordware Publishing, Inc, 2320 Los Rios Boulevard Plano, Texas 75074. Consultado: Noviembre, 2010.
- [Thomas T. Norton, 2002] Thomas T. Norton, David A. Corliss, J. E. B. J. E. B. (2002). *The Psychophysical Measurement of Visual Function*. Butterworth-Heinemann, 2 edition. Consultado: Noviembre, 2010.
- [V. Bruce, 2003] V. Bruce, Patrick R. Green, M. A. G. (2003). *Visual perception: physiology, psychology, and ecology*. Psychology Press, 4 edition. Consultado: Diciembre, 2010.
- [Visznecki, 2009] Vizsnecki, D. (2009). Sdl mixer. http://sdl.beuc.net/sdl.wiki/SDL_mixer. Consultado: Noviembre, 2010.
- [Weiten, 2010] Weiten, W. (2010). *Psychology: Themes and Variations*. L. Schreiber, 8 edition. Consultado: Octubre, 2010.
- [Xiph.Org, 2008] Xiph.Org (2008). Vorbis.com. <http://www.vorbis.com/>. Consultado: Noviembre, 2010.
- [Y. Eriksson, 2004] Y. Eriksson, D. G. (2004). Computer games for children with visual impairments. *5th Intl Conf. Disability, Virtual Reality Assoc. Tech.* Consultado: Noviembre, 2010.

A. Glosario de términos

Términos en español:

Agudeza visual: Es la capacidad para discriminar detalles finos de un objeto en el campo visual, es la inversa del ángulo desde el cual los objetos son contemplados.

Ambliopía (u ojo vago o perezoso): Consiste en la reducción de la agudeza visual en uno o ambos ojos debido a experiencias visuales anormales en la primera parte de la infancia, dando lugar a cambios funcionales de los centros visuales del cerebro.

Biblioteca: Desde el punto de vista informático es una colección o conjunto de subprogramas usados para desarrollar un software.

Funciones visuales: Capacidades del HVS de interpretar la información y el entorno, de los efectos de la luz visible que llega al ojo.

Realidad virtual: Representación de escenas u objetos producidos por un sistema informático, dando la sensación de su existencia real.

Visión binocular: Es la posibilidad de percibir el mundo desde diferentes puntos de vista debido a que se tiene una separación de alrededor de seis centímetros entre los dos ojos, esta diferencia se interpreta como una medida de la profundidad.

Visión estereoscópica: Visión binocular.

Visualización estereoscópica: Es la simulación de la visión estereoscópica por computadora.

Ergonomía: Estudio de datos biológicos y tecnológicos aplicados a problemas de mutua adaptación entre el hombre y la máquina.

Sensopercepción: Conocimiento sensorial de una realidad, basado directamente en la información que el individuo recibe de sus sentidos (audición, tacto, visión, gusto y olfato). Es algo muy complejo que en realidad está compuesto por dos procesos que se encuentran muy relacionados, que son la sensación y la percepción de los estímulos.

Renderizado: Proceso de generar una imagen desde un modelo. Este término técnico es utilizado por los animadores o productores audiovisuales y en programas de diseño en 3D.

Tao Framework : Interfaz de programación de C# para acceso a bibliotecas escritas en C++.

Reverberación: Fenómeno derivado de la reflexión del sonido consistente en una ligera permanencia del mismo una vez que se ha extinguido el original.

Eco: Fenómeno acústico producido cuando la onda sonora se refleja y regresa con retardo suficiente para superar la persistencia acústica.

Términos en inglés:

Muffling: Término del inglés, que define la forma en que la oreja está adecuada para conocer los sonidos que le llegan de frente.

Buffer: Término del inglés, que define una ubicación de la memoria en una computadora o en un instrumento digital reservada para el almacenamiento temporal de información digital, mientras que está esperando ser procesada.

Streaming: Término del inglés, que se refiere a un proceso continuo.

Flanger: Término del inglés, que define el efecto de audio que produce un sonido metalizado oscilante, sobre todo en frecuencias medias y altas.

B. Acrónimos

AV: Agudeza visual.

API: Interfaz de Programación de Aplicaciones (del inglés Application Program Interface).

CAVE: Cueva de Realidad Virtual (del inglés Computer-Animated Virtual Environment).

GLSVe: Graphics Library for Stereoscopic Vision engine.

HDSVE: Herramientas de Desarrollo para Sistemas de Visión Estereoscópica.

HVS: Sistema Visual Humano (del inglés Human Visual System).

RV: Realidad Virtual.

SRV: Sistema de Realidad Virtual.

UCI: Universidad de las Ciencias Informáticas.

VE: Visión estereoscópica.

3D: Tres dimensiones.

UO: Universidad de Oviedo.

DII: Diferencia de intensidad inter-aural.

DIT: Diferencia de tiempo inter-aural.

RIFF: Formato de intercambio de recursos de archivo.

IFF: Formato de intercambio de archivo.

PCM: Modulación por impulsos codificados.

Fecha de impresión: 15 de febrero de 2011

Generado con L^AT_EX