

Ambientes Virtuales Interactivos a través de Sonido 3D para Niños No Videntes

Por

Jaime Sánchez y Mauricio Lumbreras
Departamento de Ciencias de la Computación
Universidad de Chile

La discapacidad visual en la niñez

El mundo de los niños con impedimentos visuales es muy diferente a nuestro mundo. El sonido y el tacto adquieren un nuevo significado cuando la vista es muy limitada o inexistente. La niñez es un período excitante para el descubrimiento del mundo circundante y debido a que la visión juega un rol primario en el aprendizaje, los niños con impedimentos visuales enfrentan el riesgo de retardar su desarrollo cognitivo. Más aún, este tipo de discapacidad impone una fuerte barrera para el acceso a medios de variada naturaleza, principalmente los interactivos. Para entender cómo impacta en los niños la discapacidad visual, revisaremos las diversas facetas cognitivas en el desarrollo temprano y posterior del niño ciego, contextualizados en el plano de los ambientes interactivos.

La comunidad médica actualmente posee un entendimiento más acabado de las causas que generan la ceguera y posee un conjunto más amplio de tratamientos para una gran cantidad de condiciones visuales anómalas. Desgraciadamente, junto con este progreso se ha conjugado también un incremento en la incidencia de multi-impedimentos. Esto es principalmente debido a la alta tasa de sobrevivencia de bebés prematuros como resultado de avances médicos en el área y, en un grado menor, debido a una disminución en el acceso a los medios de cuidado pre-natal en los grandes centros urbanos. Aunque parezca raro, la ceguera y los impedimentos visuales se han incrementado durante las últimas décadas.

Los niños con impedimentos usualmente experimentan retardos en alcanzar ciertos ítems fundamentales durante los primeros cinco años de vida. Debido a que la visión juega un rol preponderante en las primeras etapas del desarrollo, los niños muy pequeños que poseen un impedimento visual están sometidos al riesgo de un retardo en su desarrollo [HG82]. Es por ello que las metodologías de estimulación temprana y el uso de estrategias adecuadas hacen que esta brecha sea minimizada [Eri50, Pia52]. Según Piaget, la fase comprendida entre los 2 y 11 años de edad, es la más importante en la formación del niño, pues en ella ocurre el juego simbólico o imaginativo. Inclusive desde los 6 a 10 años de edad, es donde el juego simbólico e imaginativo alcanza realce como para poder utilizar un sistema computacional. Varios autores sostienen que este proceso también ocurre en niños no videntes, pero las fases usualmente se ven retrasadas en 2 años con respecto a niños videntes.

Piaget habla acerca de la construcción de la realidad a través de la categorización y guía de la acción, por medio de la cual se generan los esquemas de permanencia de objetos, espacio, tiempo y causalidad. La idea de permanencia implica que el niño aprende que el objeto no cesa de existir cuando este desaparece perceptivamente, pero niños ciegos este concepto aparece notoriamente disminuido [Fra77]. Por lo tanto, un sistema altamente interactivo que explote esta teoría como

mecanismo para estimular estas destrezas es de gran valor cognitivo. El concepto de tiempo también se va desarrollando a lo largo del crecimiento y un sistema interactivo debe ser capaz de disponer los eventos de modo de estimular las nociones de serialización temporal, inclusión de intervalos y la idea de métrica temporal

Si bien muchas de estas habilidades están ligadas a una exploración y conocimiento visual del ambiente circundante y sus objetos, también ocurre pero en otra modalidad, en los niños no videntes. El problema que surge es que no existe la variedad y calidad de juegos, entretenimientos y aplicaciones educativas que puedan ser usadas por niños no videntes. Un sistema que estimule las habilidades anteriores y sirva de apresto bajo el esquema de juego interactivo, no es un sistema para despreciar, teniendo en cuenta que muchas de estas actividades son volcadas a los niños no videntes a través de elementos pasivos tradicionales y con un nivel de entretenimiento acotado al medio pasivo de soporte (cartón, plastilina, bloques LEGO, punzón y papel, collage, etc.).

Los niños que están impedidos visualmente parecen pasivos y desinteresados por el mundo externo, debido a la falta de estimulación visual la cual motiva interacción, pues no pueden ver la conexión entre lo que ellos hacen y lo que sucede. Si un ambiente altamente interactivo obliga a la reacción rápida y atención, seguramente puede hacer que esta característica de desinterés sea minimizada. Así, la conceptualización de un sistema interactivo basado en audio debe tener en cuenta este esquema causa-efecto y obtención de feedback adecuado. A estos pequeños se les debe animar a explorar el ambiente activamente de manera de practicar habilidades motoras y conceptos relacionados con el espacio. La idea de encadenar eventos con un fin determinado y la movilidad dentro del ambiente para lograr un objetivo debe estar incluido en un ambiente virtual acústico, donde el niño se deba desenvolver para obtener o lograr un cierto propósito.

Los preescolares que son impedidos visualmente necesitan oportunidades para actuar a través de exploración táctil, usando las nociones de discriminación y comparación. En particular, ellos necesitan desarrollar un entendimiento de la relación todo-parte, discriminación auditiva, localización y memoria. El concepto de permanencia o conservación de objetos, es entre otras cosas, provisto por la localización acústica. Esta propiedad será explotada en un sistema tal como es el propuesto en este proyecto, por medio de la creación de objetos acústicos en el espacio virtual circundante, en el cual es posible discriminar la posición de la fuente sonora. Este esquema está basado subyacentemente en la memoria, la cual le permite reconocer los atributos de cada sonido y su significado.

Es deseable proveer al niño con oportunidades para desarrollar habilidades sensoriales. Esto incluye discriminación táctil con características relevantes de los objetos, discriminación auditiva de las voces de las personas, sonidos relacionados a las actividades cotidianas o contextualizadas, sonidos de animales, instrumentos musicales y otros sonidos del ambiente. Además, se deben proveer oportunidades para desarrollar la fuerza en la mano y su destreza asociada. Así, dispositivos de entrada o *input* que estimulen la actividad háptica con una modalidad cruzada acústica sirven como fuerte refuerzo para el tipo de actividades anteriormente citadas. Además, la riqueza de sonidos y timbres acústicos le darán a un sistema interactivo la variedad suficiente como para aumentar la habilidad de discriminación acústica, fundamental en niños no videntes. En general, con respecto al feedback acústico, tal como el visual, se debe seguir un esquema consistente y sistemático. Un nivel persistente de feedback, aún cuando provee valiosa información, frecuentemente confunde y cansa a los usuarios [MK96].

A través del proceso del juego, las actividades cotidianas y las actividades deliberadas de aprendizaje, los niños aprenden su mundo de manera segura. Un concepto importante de definir es el de orientación, el cual denomina la noción de conocerse uno mismo como una entidad que está en un

lugar determinado de un ambiente, que conoce cómo moverse en el espacio, y qué sabe que hacer para llegar a un sitio determinado. La movilidad involucra el proceso físico de moverse a través del espacio de una manera segura.

Por ejemplo, si bien los niños de 5 o 6 años van a la escuela todos los días recorriendo calles y puntos de referencia, usualmente no pueden representar tridimensionalmente las relaciones de edificios, distancias y relaciones topográficas. Así se puede ver que una memoria motora no necesariamente representa una reconstrucción simultánea unificada. El primer obstáculo para este tipo de operaciones es el problema de representar mentalmente lo que han experimentado al nivel de acción.

Contexto cognitivo del niño ciego y su implicancia en el diseño de sistemas interactivos

Los empiricistas asumen que el conocimiento surge de una copia de la realidad, agregando que la inteligencia es un mecanismo realizado por la percepción. Esta aproximación plantea que un sistema interactivo tal como el que se propone en este proyecto, puede incrementar el conocimiento de relaciones espaciales a través de un ambiente virtual que revivan los atributos espaciales, y que la inteligencia puede ser estimulada generando un esquema rico en experiencias perceptuales.

La inclusión de objetos virtuales acústicos con vida sonora finita y dinamismo en su representación, establecen en el niño el concepto de que el objeto no cesa de existir cuando este desaparece perceptivamente. El realce de este mecanismo de percepción tiene impacto directo en los servicios que involucran la actividad sensitiva motora.

El proceso de descentralización y autonomía no sólo se aplica al universo físico sino también a lo intrapersonal o universo social. Social no sólo es educación, cultura o moral, sino también cognición y componente afectivo. La afectividad es el componente energético de los patrones de comportamiento. Por ello, el control y autonomía del niño en el ambiente virtual es una característica accesoria extra, la cual potencia las posibilidades en el contexto personal.

En resumen, un ambiente acústico virtual interactivo y navegable posee el sustento teórico como para ser considerado una aplicación de valor educativo y de estimulación de destrezas de valor significativo para ser utilizado por niños no videntes.

Sistemas de información y aplicaciones interactivas para no videntes. Estado del arte de la tecnología.

La tecnología a lo largo de los años ha provisto diferentes alternativas para el acceso a la información por no videntes. Posteriormente, revisaremos cómo nuevas aplicaciones de cómputo y tecnologías de interacción son utilizadas sin pistas visuales y así dan un legado que contextualiza al presente trabajo en el área.

Desde el sistema Braille a las GUIs

Una de las técnicas más tradicionales para la transferencia de información y almacenamiento para no videntes surge de la creación de caracteres explorados táctilmente. Louis Braille, creó un sistema basado en puntos dispuestos en dos columnas de tres, determinando una celda que representa un carácter. Las hojas de papel o plástico grabadas con estos caracteres constituyen para las personas ciegas elementos permanentes de lectura, tales como los libros tradicionales funcionan para los videntes. Actualmente existen lo que se conocen como *células braille*, que es un conjunto de

elementos eléctricamente móviles los cuales dispuestos en línea constituye una línea braille. Esta línea usada como terminal de la computadora, es capaz de reproducir en braille, mediante software y la interfaz adecuada, una línea de texto convencional. El usuario lee esa línea pasando el dedo sobre ella como si se tratara de una línea impresa. Una vez leída, un nuevo conjunto de caracteres ocupa el lugar de los anteriores, y de esta manera se prosigue hasta completar un texto dado. El uso de mecanismo bidimensionales usando líneas braille y dispositivos hápticos se muestra también como una alternativa viable a través de una mejora del confort [Ram96].

Hay que tener en cuenta que la salida táctil no es la única manera para transmitir a la persona ciega la información de textos codificados en memoria del computador. Los sintetizadores de voz permiten a la información escrita hacerla inteligible a través del oído. El hardware suele constar de tarjetas que se adicionan a la computadora, diseñadas usando chips dedicados o cierto software que utiliza alguna placa de sonido estándar. El sintetizador de voz puede leer una pantalla en modo texto, pero no es suficiente para manejar programas y formatos complicados de pantalla. Para conseguirlo es necesario utilizar además un software lector de pantalla (*screen reader*). Este software permite recorrer la pantalla y así acceder a cualquier software basado en modo texto. De esta manera, un no-vidente puede aprovechar gran parte de las posibilidades de una computadora. Un elemento crítico en este tipo de interfaces es el diseño adecuado del diálogo establecido entre el usuario y la computadora pues sino la usabilidad cae a un nivel inoperable [PE96]. Inclusive este es uno de los primeros trabajos que comienza a hablar de usabilidad para usuarios no videntes.

El advenimiento y proliferación de GUIs durante la última década y la reciente introducción de conceptos multimediales en la interacción humano-computador trae consigo nuevas posibilidades, pero también muchos problemas para usuarios ciegos con referencia al acceso a sistemas de cómputo e información digitalmente soportada. Hasta hace poco, las GUIs fueron totalmente inaccesibles para usuarios ciegos. Estas interfaces han sido diseñadas para explotar las capacidades visuales de videntes por medio de complicadas representaciones pictóricas, haciendo para los *screen readers* la tarea sumamente difícil. Dado el estado del arte de los lectores de pantalla, dos proyectos intentan generar un acceso genérico a aplicaciones basadas en GUI. Ellos son el proyecto GUIB [WK93], auspiciado por el consorcio TIDE de la CEE (Comunidad Económica Europea); y el otro, el proyecto denominado Mercator, desarrollado en el Instituto de Tecnológico de Georgia, USA [ME92].

El proyecto GUIB (*GUI for the Blind*) se basa fundamentalmente en el uso de un dispositivo especial que integra líneas braille, sonido estéreo, un panel sensitivo a la presión y teclas para manipulación del cursor [WK93]. La idea es remapear los eventos y simbología de cada artefacto gráfico de Windows a este dispositivo. Por medio de las líneas braille, el usuario es capaz de explorar la pantalla. Teclas especiales permiten activar un sintetizador de voz que describe el estado de la interfaz y lee caracteres ingresados. La idea del proyecto Mercator [ME92] es diferente a la de GUIB, pues no utiliza dispositivos especiales y además impone un cambio de metáfora en la representación de la interfaz. La plataforma de uso de Mercator es cualquiera que corra el protocolo X (X-Window). La información referente a la interfaz gráfica es modelada como una estructura de árbol, la cual representa los objetos gráficos en la interfaz (botones, menús, áreas de texto, etc.) y la relación jerárquica entre estos objetos. De esta manera, un ciego interactúa con el sistema independientemente de su representación gráfica. Es decir, hay un cambio de metáfora. El contenido de la interfaz es expresado a través de voz sintetizada y *nonspeech audio* (audio que no representa voz).

Acceso a WWW para no videntes

Si bien una solución que permita acceso genérico a una GUI permitiría acceder a cualquier aplicación y por ende a un *browser* o navegador de la World Wide Web (WWW), existen aproximaciones que intentan ofrecer un browser especial para acceder a la WWW evitando representaciones visuales y con

presentación de información totalmente acústica. El problema fundamental proviene del hecho que mucha de la información en la WWW es de carácter gráfico y que además todos los browsers corrientes se basan en la metáfora *point-and-click* para acceder a un documento. Sin embargo, gran cantidad de información es textual y se podría pensar en un esquema que filtre la información y así sea accesible a un no vidente. El caso de V-Lynx [KC96] es uno de ellos. Por medio de síntesis de voz y una adaptación al browser de modo texto llamado Lynx, los autores implementaron una adaptación de él, en el cual el usuario navega secuencialmente el documento a través de síntesis de voz y puede obtener meta-información para determinar la estructura del documento. Si bien esta iniciativa es interesante, [PMM97] explora cómo debe ser diseñado un documento hipermedial para un no vidente. Una esquema interesante es el que existe operacional a través de un proxy en el ETH [ETH98]. Si uno configura su *browser* o navegador para que el proxy sea *ea.ethz.ch:8080*, entonces obtendrá todas las páginas filtradas y estructuradas en modo texto. De tal modo, si usamos un navegador de solo texto y un sintetizador de voz, podremos interpretar y conocer una vasta cantidad de páginas web. Mas recientemente [Roth99] establece como browsers multimodales intentan proveer acceso a la WWW para no videntes.

Acceso a información sin pistas visuales

Si un sistema interactivo intenta transferir un contenido a un niño no vidente deberemos conocer como se puede expresar información estructurada más allá de un simple cassette de audio. Uno de los trabajos más interesantes surge con HyperPhone [MFC92], un prototipo experimental para comunicaciones de voz soportadas por computadora desarrollado en Bellcore, en el cual se provee un ambiente para interactuar con entidades denominadas *voice documents* (documentos acústicos) y aplicaciones operadas por la voz o para potenciales usuarios que no pueden acceder a displays o teclados. La idea de proveer interfaces de audio, y adicionarle inteligencia haciéndolas fáciles de operar inclusive vía telefónica por reconocimiento de voz y minimizando las interacciones innecesarias es tratado en [YLM95]. En este trabajo se rescata la importancia de un modelo conversacional como modo de interacción con el sistema y el error que ocurre al tratar de trasladar las modalidades gráficas de interacción a una modalidad enteramente auditiva. Con respecto a la organización de información acústica automática [Aro93] exhibe técnicas para segmentar automáticamente grabaciones monolíticas en pequeños *chunks* o porciones atómicas de información en un sistema denominado Speech Skimer de modo de incluir automáticamente tramos de conversaciones en sistemas de información acústica. Asociada a la idea de la organización automática de documentos de voz [SAS93] presenta con Voice Notes, un esquema para controlar un pequeño dispositivo que se puede llevar en forma autónoma y portátil, que permite capturar, manejar y obtener lo que denominan *voice notes* o clips de audio. Plantea un simple modelo jerárquico basado en categorías y notas dependientes, como el más favorable para este dominio. Interfaces modales vs. no modales, señales no vocales como feedback de interacción, conservación fundamental del recurso tiempo y estructuración de la información en forma dinámica, son los puntos interesantes a tener en cuenta en un sistema interactivo para niños no videntes.

Si la información no va a ser organizada de manera jerárquica sino a modo de hipertexto, [Aro91] describe en Hyperspeech, como se puede navegar en una hiper base de información acústica. Los resultados más relevantes indican la dificultad en la autoría de documentos de esta naturaleza, la posibilidad de efectos de audio para indicar algún atributo de la información que se va recibir, la posibilidad de crear conversaciones virtuales entre locutores que nunca se reunieron en la realidad, la fuerte tendencia a la desorientación en un ambiente sin pistas visuales, etc. De aquí se extraen ideas para desarrollar un soporte hipertextual a un sistema interactivo.

Interfaces y esquemas de interacción sin pistas visuales

Si bien [Aro91] con su Hyperspeech no utiliza una metáfora espacial, deja abierto el interrogante de cómo realizar algo tal como “point & click” en un medio acústico. Por ello, la utilización de una metáfora espacial asociada a un modelo funcional conversacional a través de sonido 3D y manipulación kinestésica, es un trabajo desarrollado y propuesto por [LR95, Lum96, LSB96] en el cual existe una manera de hacer point & click sobre un medio altamente temporal como es el sonido. Este sistema se provee el acceso a información hipermedial a través de un esquema de interacción con locutores virtuales previamente digitalizados, los cuales son involucrados en una conversación en la cual el usuario puede regular el flujo de la información, seleccionando qué locutor es el que puede hablar. Uno de los legados interesantes de este trabajo es la inclusión de unas entidades acústicas denominadas *audicons*, las cuales juegan un papel similar al icono en su contraparte gráfica. Otro resultado interesante que surge es que la interacción con dispositivos físicos que impliquen manipulación háptica son más adecuados para aplicaciones de entretenimientos que para uso y acceso a sistemas de información, siendo este un resultado relevante para aprovechar en este proyecto. [RV95] presenta un excelente resumen del de interfaces de audio, tratando de extender la idea básica de menús, listas y formularios. Como resultado más importante se destaca la economía de espacio que se debe hacer en una interfaz, en este caso el ahorro pasa por el tiempo, debido a que el display es acústico y no gráfico.

Uno de los resultados más categóricos al respecto de sistemas interactivos basados en sonido 3D aparece en [LS98, LS99] en el cual a través de un test experimental con un sistema denominado AudioDoom los autores presentan evidencias sustanciales que niños no videntes pueden percibir un ambiente acústico 3D, navegarlo y crear representaciones espaciales usando bloques LEGO. Los resultados de estos trabajos sirven como prueba del concepto y señalan importantes hipótesis para la continuación de esta labor, las cuales se plasman en este proyecto y tienen concordancia con lo presentado en [Ken97, Sat96]. Más aún, existen excelentes medios para recrear y estudiar como funciona la navegación en no videntes y videntes usando los mecanismos acústicos, extendiendo el estudio de la navegación de [LKG93]. La evaluación realizada en [LS99] siguió un esquema de experimento, pero es de recalcar que ya hay delineado un esquema para evaluar este tipo de tecnologías asistivas[SE96].

Medios subyacentes para la creación de un ambiente acústico virtual: sonido espacializado 3D

La representación de un ambiente acústico que reviva los atributos espaciales de una fuente sonora es de fundamental importancia si intentamos representar un ambiente basado en una metáfora espacial. Debido a que el usuario destino del sistema no contará con pistas visuales, es deseable utilizar al máximo los sentidos disponibles, siendo el del oído el que mayor ancho de banda y poder de expresividad ofrece. Para ello el usos de un sistema de sonido 3D es de vital importancia. Un sistema de sonido 3D es uno tal que reemplaza o complementa los atributos espaciales de una cierta fuente de sonido en su reproducción acústica [Wen92]. La utilización de sonido 3D es sumamente valiosa pues mejora la respuesta para cualquier usuario, y para usuarios no videntes es importantísimo pues pueden alcanzar precisión y versatilidad en la manipulación de la interfaz [MK96]. Si la idea es utilizar un sistema de sonido denominado transaural, la representación de sonido será finalmente hecha con parlantes desacoplados de la cabeza, eliminando el uso de auriculares para la espacialización del sonido [Int96, Gar95]. La teoría básica que describe la localización a través de estas pistas se denomina “teoría duplex” y proviene desde principio de siglo [Ray07]. El problema de esta teoría es que no describe como el ser humano puede distinguir un sonido desde el frente o el fondo, pues desde esas posiciones las pistas de retardo de tiempo e intensidad son similares. Surge más recientemente la

teoría de las HRTFs, las cuales permiten describir que una fuente sonora se localiza por el efecto morfológico del pabellón de la oreja [Bla83]. Este tipo de modelo es el que actualmente se utiliza en las placas de sonido de PC que generan sonido 3D [Dia96].

Ambientes acústicos virtuales: sonificación y pantallas virtuales acústicas

El estudio de los sonidos desde el punto de vista físico o sicoacústico sólo refleja una parte en el diseño de un display acústico o como medio subyacente para expresar una aplicación interactiva para niños no videntes. La idea fundamental de lo que se denomina un display acústico, es la de poseer un medio preciso para transferir información a un ser humano usando modalidades auditivas, las cuales combinan características direccionales y semánticas de objetos dinámicos, los cuales representan entidades de un ambiente simulado [Wen92]. Esta información puede referirse a orientación en un cierto ambiente, el estado de un sistema o inclusive la representación acústica de un conjunto de datos (sonificación). La manipulación de las variables que definen un sonido, permite mapear cada una de ellas a diferentes parámetros del sistema en función de la aplicación del display acústico. Una extensión a este concepto surge en la idea de pantalla virtual acústica 3D. Si la codificación sonora expresada por el sistema se presenta a través de voz, entonces el DA está explotando el repertorio del lenguaje y su significado cognitivo. Si la codificación está expresada a través de sonidos denominados *nonspeech*, es decir señales acústicas que no son voz humana, el DA explotará el conocimiento evolucionario adquirido a través del uso del sistema y las pistas cognitivas que generan la naturaleza de las señales percibidas. Un DA permite manipular un ambiente con 'ojos libres', da una rápida detección y alerta de condiciones anómalas, permite orientación, percepción paralela, *backgrounding*, resolución temporal de eventos, y formación de *gestalts* auditivos [Kra94]

En ambientes virtuales y en el caso de no videntes donde no hay feedback visual de la actividad desarrollada, las experiencias de interacción no son fácilmente expresables. Se denomina *sinestesia* a la sustitución de una modalidad sensorial por otra, y esta es una técnica utilizada para solucionar el problema anterior. Un ejemplo común en aplicaciones de RV (Realidad Virtual) es el siguiente: para proveer una sensación que reemplace el feedback táctil generado al tocar un objeto virtual, el sistema genera un sonido especial para dar el feedback indicando que el usuario ha tomado el objeto en cuestión. Este tipo de representación es de vital importancia pues permite al niño ciego obtener conciencia de su actividad en el ambiente y del estado de interacción inmediata con un objeto dado. Hay que tener en cuenta que un DA también posee limitaciones tal como baja resolución de muchas variables auditivas, limitada precisión espacial, falta de valores absolutos y ortogonalidad, interferencia con comunicador oral, ausencia de persistencia y limitaciones acústicas propias del usuario.

Un tema de singular importancia en un ambiente interactivo para niños no videntes es el de la sonificación. Sonificación es el *mapping* de relaciones representadas numéricamente en algún dominio de estudio, a relaciones en el dominio acústico con el propósito de interpretar, entender o comunicar relaciones en el dominio estudiado [Sca92]. Si un ambiente debe expresar un estado dado, una condición de las entidades del ambiente o similar este concepto debe estar incluido. Un ejemplo claro es la expresión de las distancias relativas desde la posición virtual del niño con respecto a una entidad, a modo de bastón virtual. Este tipo de modalidad es totalmente olvidado en los pocos sistemas interactivos para niños no videntes [BCC97].

Diseño de un display virtual acústico 3D y falta de esquemas establecidos

Son bien conocidos las pautas que permiten diseñar cierta GUI o diseñar una aplicación que corra sobre una GUI. Esta modalidad de interfaz gráfica, se ha tornado en un exitoso elemento en la interacción humano-computador, permitiendo que múltiples aplicaciones usen simultáneamente un display visual y que el usuario interactúe con ellas aprovechando el ancho de banda que provee el sentido de la vista. Estas capacidades provienen realmente de la combinación de 2 características: El usuario puede controlar la organización espacial de múltiples objetos visuales (ventanas) por medio de cierta interfaz física y un manejador de ventanas, y además el usuario puede cambiar su atención visual entre varios objetos representados.

Si bien desde [CL91] en donde se expresan las primeras ideas acerca de la extensión de un sistema de ventanas potenciado con sonido 3D para manipular más fácilmente el ambiente y con [Gav89] en donde se extiende el desktop de McIntosh para producir un feedback adecuado en función de la actividad del usuario, han surgido pocos resultados concisos y categóricos en el diseño de displays acústicos que exploten características espaciales. Recientemente en [KS97] surgen algunos items interesantes en el desarrollo de representación espacial acústica de información.

Hoy por hoy existen pocas pautas precisas y estandarizadas para llevar a cabo la tarea de esta tesis y por ello se debe recolectar la experiencia de los trabajos previos, los cuales plantean algunos puntos a tener en cuenta para partir en el diseño.

La metodología para diseñar un DAV (Display Acústico Virtual) en el contexto de una aplicación de educativa tradicional o de edutainment para niños no videntes, no sólo involucra el dominio de la aplicación, sino que también incluye aspectos culturales, físicos y de implementación. La literatura es escasa con respecto al diseño de display acústicos sin pistas visuales. Por esta razón, parte del trabajo original de este proyecto es determinar qué modalidades son relevantes de incluir en un DAV.

El diseño de un DAV cubre básicamente dos aspectos: el aspecto técnico de la manipulación de sonidos, y el aspecto estético y conceptual de la organización espacial de ellos. La creación de un ambiente genérico y reusable basado en la idea de un display acústico en sonido 3D, apuntamiento 3D, reconocimiento de gestos y comandos activados por reconocimiento de voz con el fin de permitir una navegación jerárquica es explorado en [SSK96]. Si bien esta aproximación es interesante, todavía no está madura aún una especificación abierta y extensa que pueda ser utilizada por cualquier diseñador que intente explotar esta aproximación. Un concepto que apareció en [LR95] y posteriormente en [Kri96] es la utilización de un anillo virtual de objetos acústicos que rodean al usuario para poder activar ellos en función de una cierta acción de interfaz. A esta organización del display acústico el autor lo denomina topología de anillo.

Referencias

- [ALK92] Appino P., Lewis J., Koved L. et al. An Architecture for Virtual Worlds. *Presence*, Volume 1, Number 1, Winter 1992, MIT Journals, 1992.
- [Aro91] Arons B., Hyperspeech: Navigating in speech-only hypermedia, in *Proceedings of Hypertext '91*, ACM, 1991, pp. 133-146
- [Aro93] Arons B., SpeechSkimmer: Interactively Skimming Recorded Speech, In *Proceedings of the ACM*

- Symposium on User Interface Software and Technology*, UIST, ACM Press, Nov. 1993
- [Bla83] Blauert J., Spatial Hearing: The Psychophysics of Human Sound Localisation. Cambridge MA: MIT Press
- [Ble92] Blenkhorn P., Requirements for screen access software using synthetic speech, in Zagler W. (Ed.), *Computer for Handicapped Persons*, Oldenbourg: Wien; pp 31-37, 1996
- [BCC97] Blind Children Center. First Steps, a Handbook for teaching young children who are visually impaired. Los Angeles, California, USA, 1997.
- [BMC93] Burger D., Mazurier C., Cesarano S., Sagot J., The design of interactive auditory learning tools, Non-Visual Human-Computer Interactions, Eds. D. Burger, J.C. Sperandio. Colloque INSERM, John Libbey Eurotext, 1993, Vol. 228, pp. 97-114
- [Bre94] Bregman A., *Auditory Scene Analysis: The Perceptual Organization of Sound*, MIT Press, 1994
- [CL91] Cohen M., Ludwig L., Multidimensional Window Management, *International Journal of Man/Machine Systems*, 34: 319-336, 1991
- [Cry94] Alphasat User's Manual, Crystal River Engineering Inc., 490 California Ave, Suite 200, Palo Alto, CA 94306, 1995.
- [Den93] Denis M, Visual images and models of described environments, *Proceedings of Non-Visual Human-Computer Interactions Conference*, Eds. D. Burger, J.C. Sperandio. Colloque INSERM, John Libbey Eurotext, 1993, Vol. 228, pp. 3-12
- [Dia96] Diamond Multimedia, *Monster Sound Reference*, disponible en <http://www.diamondmm.com>
- [Die96] A. Dieberger. Browsing the WWW by interacting with a textual virtual environment - A framework for experimenting with navigational metaphors. *Proceedings of Hypertext '96*, pp.170-179, 1996.
- [Dur97] Durham S., Sounding Off with the RSX Library, *Dr.Dobb's Journal*, July 1997, pp.66-71
- [Eri50] Erikson, E. *Childhood and society*. New York:Norton, 1950
- [ETH97] World Wide Web Access for Blind People, Institut fuer Informations systeme, ETH Zurich, , 1997, accesible en www.inf.ethz.ch/departemente/IS/ea/publications
- [Fra77] Fraiberg, S. *Insights from the blind*. New York: Basic Books
- [Gar97] Gardner W., Transaural 3-D Audio, *Phd. Thesis*, MIT, 1997
- [Gav89] Gaver W., The Sonic Finder: An interface that uses auditory icons. *Human-Computer Interaction*, 4 (1), pp. 67-94, 1989
- [HG82] Herren H., Guillemet S., Estudio sobre la educación de los niños y adolescentes ciegos, ambliopes y sordo-ciegos, Editorial Médica y Técnica, Barcelona, España, 1982.
- [Hat93] Hatwell Y., Images and non-visual spatial representations in the blind, *Proceedings of Non-Visual Human-Computer Interactions Conference*, Eds. D. Burger, J.C. Sperandio. Colloque INSERM, John Libbey Eurotext, 1993, Vol. 228, pp. 13-35
- [Int97] *Realistic Sound Experience Developers Kit*, disponible via Internet en <http://developer.intel.com/ial/rsx>
- [Joi94] Joiner D., Real Interactivity in Interactive Entertainment. *Computer Graphics*, Vol 28, number 2: 97-99, 1994.
- [Kaw93] Kawalski R., *The Science of Virtual Reality and Virtual Environments*, Wokingham, England: Addison-Wesley
- [KC96] Krell M., Cubranic D., V-Lynx: Bringing the World Wide Web to Sight Impaired Users, in *proceedings of ASSETS '96*, *The second Annual ACM Conference on Assistive Technologies*, Vancouver, Canada, 1996. pp 23-26
- [Ken97] Kennedy J., How the Blind Draw, *Scientific American*, January 1997, pp. 60-65
- [Kra94] Kramer G., An introduction to Auditory Displays, en *ICAD 92, Auditory Display: Sonification, Audification and Auditory Interfaces*, pp.1-78, Addison-Wesley, 1994
- [Kri96] Krispien G., The GUIB spatial auditory display - Generation of an audio-based interface for blind computer users, *Proceedings of ICAD 96, International Conference on Auditory Displays*, USA, 1996
- [KS97] Kobayashi M., Schmandt C., Dynamic Soundscape : mapping time to space for audio browsing, *Proceedings of ACM CHI 97*, pp. 224-228, 1997
- [LR95] Lumberras M., Rossi G., A metaphor for the visually impaired: browsing information in a 3D auditory environment, in *Proceedings Companion CHI 95, ACM Conference on Human Interaction*, Denver, Colorado, Mayo 1995, pp.261-262

- [LKG93] Loomis J., Klatzky R., Golledge R., Nonvisual Navigation by Blind and Sighted: Assessment of Path Integration Ability, in *Journal of Experimental Psychology*, 1993, Vol. 122 No 1, pp 73-91
- [Lum96] Lumbreras M., Tesis de grado para obtener el grado de Licenciatura en Informática, Universidad Nacional de La Plata, Argentina, Febrero 1996
- [LS97] Lumbreras M., Sánchez J., Hyperstories: A Model to Specify and Design Interactive Educational Stories, *XII Conferencia Internacional de la Sociedad Chilena de Ciencias de la Computación*, Valparaíso, Noviembre de 1997. Publicado también como capítulo del libro de IEEE CS Press.
- [LS98] Lumbreras M., Sanchez J., 3D aural interactive hyperstories for blind children, in *Proceedings of The 2nd European Conference on Disability, Virtual REality and Associated Technologies*, Skovde, Suecia, Septiembre 1998, pp. 119-128. Also in *International Journal in Virtual Reality*, forthcoming.
- [LS99] Lumbreras M., Sanchez J., Interactive 3D Sound Hyperstories for Blind Children, in *Proceedings of ACM CHI 99 Conference*, pp.318-325, Pittsburgh, USA, May 1999
- [LSB96] M. Lumbreras, J. Sánchez, M. Barcia. A 3D sound hypermedial system for the blind. *Proceedings of the First European Conference on Disability, Virtual Reality and Associated Technologies*, pp. 187-191, Maidenhead, UK, 1996.
- [LSG84] Landau B., Spelke E., Gleitman H., Spatial Knowledge in a young blind children. *Cognition*, 16:225-260
- [ME92] Mynatt E., Edwards W., *The Mercator Environment. A Non Visual Interface to X-Windows and Unix Workstation*, GVU Tech Report GIT-GVU-92-05, Feb 1992
- [MFC92] Muller M., Farrel R., Cebulka K., Smith J., Issues in the Usability of Time-Varying Multimedia, *Multimedia Design*, ACM Press, 1992, pp. 7-38
- [MK96] Mereu S., Kazman R., Audio Enhanced 3D Interfaces for Visually Impaired Users, in *Proceedings of the ACM CHI 96*, Vancouver, Canada, 1996. pp. 72-78
- [PE96] Pitt I., Edwards A., Improving the Usability of Speech-Based Interfaces for Blind Users, in *Proceedings of ACM ASSETS '96*, Vancouver, British Columbia, Canada, pp. 124-133
- [Pia52] Piaget J., *The origins of intelligence in children* (M.Cook, Trans.) New York: International University, 1952
- [PMM97] Petrie H., Morley S., McNally P., O'Neill A., Majoe D., Initial Design and Evaluation of an Interface to Hypermedia Systems for Blind Users, in *proceedings of The Eight ACM Conference on Hypertext, Hypertext 97*, Southampton, UK. pp 48-56
- [Ram96] Ramstein C., Combining Haptic and Braille Technologies: Design Issues and Pilot Study, in *Proceedings of ACM ASSETS '96*, vancouver, British Columbia, Canada, pp. 37-44
- [Ray07] Rayleigh L., On our perception of sound direction, *Philosophy Magazine*, 13, pp. 214-232, 1907
- [Roth99] Roth P., Petrucci L., Pun T., Assimacopoulos A., Auditory browser for blind and visually impaired users, in *ACM CHI 99 Extended Abstracts*, pp. 218-219, Pittsburgh, USA.
- [RV95] Resnick P., Virzi R., Relief from the Audio Interface Blues: Expanding the Spectrum of Menu, List, and Form Styles, *ACM Transactions on Computer-Human Interaction*, Vol. 2, No. 2, june 1995, pp. 145-176
- [SAS93] Stiefelman L., Arons B., Schmandt C., Hulteen E., Voice Notes: A Speech Interface for a Hand-Held Voice Notetaker, in *Proceedings of INTERCHI '93*, ACM, April 1993, pp.179-186
- [Sat96] Satalich G., Navigation and Wayfinding in Virtual reality: Finding Proper Tools and Cues to Enhance Navigation Awareness, *Phd. Thesis*, Univeristy of Washington, USA, 1996.
- [Sca92] Scaletti C., Sound Synthesis Algorithms for Auditory Data Representations, en *ICAD 92, Auditory Display: Sonification, Audification and Auditory Interfaces*, pp.223-252, Addison-Wesley, 1994
- [SE96] Stevens R., Edwards A., An Approach to the Evaluation of Assistive Technology, in *Proceedings of ACM ASSETS '96*, Vancouver, British Columbia, Canada, pp. 64-71.
- [SL95] Sánchez, J., & M. Lumbreras. Interfaces for learning. In Anzai, Y., Ogawa, K. And Mori, H(Editors). *Advances in Human/Factors: Human-Computer Interaction.*, Symbiosis of Human and Artifact: Future Computing and Design for Human-Computer Interaction, 20A, pp. 865-870. New York: Elsevier Publishers, 1995.
- [SL97] Sánchez J. and Lumbreras M, HyperStories: Interactive Narrative in Virtual Words, *Fourth International Conference of Hypertexts And Hypermedia: Products, Tools, Methods*. Paris, Septiembre 25-26,1997. *Hypertextes et Hypermedias Revie*, Ed. Hermes.

- [SL99] Sánchez J. and Lumbreras M, Virtual environment interaction through 3D audio for blind children. Proceedings of Medicine Meets Virtual Reality, MMVR:7, San Francisco, USA, 1999.
- [SLB95] Sánchez J., Lumbreras M, Bibbo M, Hyperstories for Learning. Interantional Workshop on Hypermedia Design '95, pp. 239-248, Montpellier, Francia,1995. Publicado por Elsevier Publishing Lectures in Computer Science.
- [SMC94] J. Sánchez , A. Mallegas, L. Cernuzzi, G. Rossi, M. Lumbreras, A. Díaz, A conceptual framework for building hyperstories. *Ed-Media '94 Conference Proceedings*, p.761, 1994.
- [SSK96] Savidis A., Stepanidis C., Korte A., Crispian K., Felbaum K., A Generic Direct-Manipulation for Hierarchical Navigation in Non-visual Interaction, in *proceedings of ASSETS '96 , The second Annual ACM Conference on Assistive Technologies*, Vancouver, Canada, 1996.pp. 117-123
- [TFM97] M. Turine , M. Ferreira de Oliveira, P. Masiero. A Naviagtion-Oriented Hypertext Model Based on Statecharts. *Proceedings of Hypertext '97*, pp. 102-111, 1997.
- [WC93] Weber G., Cochaneck D. Et al, Access by blind people to interaction objects in MS Windows, en Proceedings ECART 2, European Conference on the Advancement of Rehabilitation Technology, Estocolmo, mayo 1993, pp. 2.2
- [Wen92] Wenzel E., Localization in Virtual Acoustic Displays, *Presence*, vol. 1 number 1, 1992, pp. 80-107, MIT Journals
- [YLM95] Yankelovich N., Levow G., Marx M, Designing SpeechActs: Issues in Speech User Interfaces, in *Proceedings of Human Factors in Computing Systems, CHI'95*, ACM, Denver, Colorado, Mayo 1995, pp. 369-376.

