

Aspectos matemáticos del proyecto multidisciplinar de I+D “Espacio Acústico Virtual”

SCTM03



Haresh M. Chulani

Ingeniero Electrónico del Area de Instrumentación del Instituto de Astrofísica de Canarias

Antonio F. Rodríguez Hernández

Licenciado en Medicina y Cirugía e Investigador de Proyecto del Departamento de Fisiología de la Universidad de La Laguna

1. Introducción

Es proverbial la capacidad del invidente para moverse, orientarse y reconocer e interactuar con el entorno apoyándose en los otros sentidos. Braille, con el desarrollo de su sistema de lectura, nos muestra que es posible realizar de una manera efectiva y útil tareas que habitualmente consideramos propias, si no exclusivas, de un sentido a través de otra modalidad sensorial. El presente proyecto incide en esta concepción y aborda una línea de investigación y desarrollo en la percepción del entorno inmediato del invidente a partir de sonidos.

1.1. “Sonidos que se extienden”. La hipótesis se fundamenta en el siguiente razonamiento. La imagen visual que experimentamos se genera a partir de las imágenes del entorno impresionadas en nuestras retinas. Cada punto de una imagen retiniana dada se corresponde con un punto concreto del espacio abarcado por la mirada. La información (en forma de onda luminosa) procedente de cada uno de los puntos del campo perceptivo visual es la que finalmente utiliza el cerebro para generar la imagen visual que experimentamos. Si esa misma información espacial procedente del objeto viaja en forma de onda sonora, se plantea la hipótesis de si es posible que se genere también en el receptor una imagen de la ocupación espacial por parte del objeto u objetos presentes en el campo perceptivo, útil para el reconocimiento y la interacción con los mismos.

Este razonamiento, junto a toda una serie de indicios científicos que parecen apoyarlo, procedentes tanto del campo de la neurofisiología como de la psicología y la psicoacústica (por ejemplo: Andersen, 1997; Konishi, 1986; Takahashi, 1994; Lakatos, 1993; Bach-Y-Rita, 1969; Rice, 1965) hace gestar pues la idea de que:

“A través del estímulo sonoro adecuado, un sujeto puede percibir de una forma global, de forma similar a como ocurre con la visión, una imagen de los objetos del entorno, con la opción de percatarse de sus características espaciales como su posición, las dimensiones en los tres ejes, sus límites, etc.”

Se aborda así una línea de investigación dirigida a validar tal hipótesis y a establecer las propiedades que debe cumplir dicho estímulo. Por la aplicación última de la misma y por la indicación de los propios argumentos neurofisiológicos y psicológicos referidos, se plantea desde el inicio explorar la viabilidad de la hipótesis en el grupo de la población invidente, sin cuya valiosa colaboración ninguno de los avances habría sido posible.

Lo que se pretende es, pues, generar una ***percepción global y mantenida en el tiempo del conjunto del entorno captado***, pero para ser experimentada ***como una imagen*** de la

ocupación espacial por parte de los objetos de dicho entorno, lo cual se acerca más a la experiencia de la visión que a la propia de la audición.

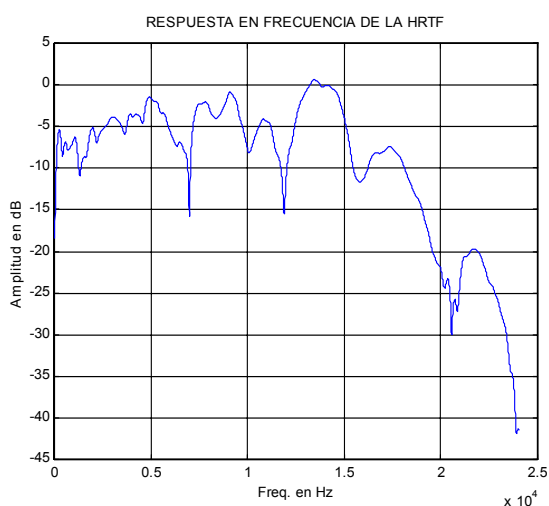
Existen múltiples líneas de investigación sobre percepción del entorno mediante sonidos. En algunas se ha utilizado el sonido únicamente como indicador de presencia o no de un objeto y su posible distancia, en el caso de los denominados detectores de obstáculos [por ejemplo: Doods, 1981; Benjamin, 1974; Loomis JM, Golledge RG, Klatzky RL, 1998]. Las líneas más cercanas a nuestro planteamiento pretenden generar un conocimiento de la escena en sí de una manera global, pero tras aprender un código y basando el reconocimiento de los objetos en el aprendizaje de un patrón sonoro concreto o específico [Kay, 1984; Capelle et al, 1998].

1.2. Las bases de la localización humana de fuentes sonoras. La información espacial es transmitida en nuestro caso utilizando el código natural, es decir: **codificamos cada punto del espacio ocupado por un objeto mediante un sonido que es percibido por el sujeto como si procediera de dicha posición.** El estímulo sonoro final correspondiente a uno o varios objetos suficientemente grandes de una escena, equivale pues, en líneas generales, al efecto de cubrir con pequeñas fuentes sonoras la superficie de los objetos a percibir. Para entender cómo es posible generar en un sujeto por medio del sonido que se le envía a través de unos auriculares, la ilusión perceptiva de que tal sonido procede de múltiples posiciones que componen la superficie de un objeto, es conveniente repasar las bases de la capacidad humana de localizar fuentes sonoras aisladas.

Para localizar una fuente sonora externa, el sistema auditivo humano se apoya en una serie de pistas contenidas en el sonido que está recibiendo. Tales son:

- La diferencia interaural de tiempo de llegada de un mismo sonido a un oído y al otro (ITD).
- La diferencia interaural de intensidad con que llega un mismo sonido a un oído frente al otro (IID).
- Las modificaciones que sufre el sonido en cada oreja por efecto de reflexiones, difracciones, etc. en sus repliegues (HRTF).

Dada una fuente sonora a nuestra derecha, el sonido emitido llega antes al oído derecho que al izquierdo. La diferencia del tiempo de llegada depende de cuánto a la derecha esté la fuente respecto a nosotros. Igualmente ocurre con la intensidad: la presencia de los relieves faciales y craneales ejercen un efecto de sombra sobre determinadas frecuencias procedentes de una fuente sonora. Se genera así una diferencia de intensidad percibida en cada oído que el cerebro reconoce como correspondiente a una determinada posición en el espacio (concretamente en el eje horizontal). Además, cada oreja, junto al resto de la cabeza y del cuerpo en menor medida, da lugar a una serie de modificaciones en la señal acústica que llega hasta ella y la sobrepasa para llegar al tímpano. Estas modificaciones son dependientes y específicas de la dirección relativa de



aproximación de la onda acústica a la oreja, y el cerebro es capaz de extraerlas de la señal que finalmente le llega al tímpano. Estas modificaciones constituyen una pista más de localización, en particular para determinar la altura de la que procede el sonido, si se encuentra delante o detrás de nosotros, y el hecho de que proceda de fuera de nosotros y no sea un sonido de producción corporal. Con la técnica apropiada (que se abordará más adelante), dichas modificaciones pueden ser medidas y representadas en forma de una función matemática. Esta función corresponde a las modificaciones espectrales que ejerce la oreja y el resto del cuerpo sobre la onda sonora incidente, y consisten en una serie de picos y de depresiones de la función de la intensidad sonora respecto de la frecuencia. Es conocida por las siglas HRTF (de Head Related Transfer Function).

1.3. “Sonidos que se ven”. Varios de los colaboradores invidentes han referido que perciben en el punto donde acústicamente localizaban la fuente sonora, unos destellos luminosos que describen como “pequeñas lucecitas o estrellitas”, simultáneos a la percepción acústica e inducidos por esta. Este fenómeno se ha repetido a lo largo del tiempo, siempre que no se han variado las condiciones del estímulo. Con la utilización del prototipo, se ha reproducido en sólo una ocasión en el caso de la única colaboradora de aquellas dos que lo ha utilizado, y refiere diferencias cualitativas con respecto a las demás ocasiones en que no lo ha percibido, en el sentido de una mayor facilidad para la percepción de dimensiones y formas, aparte de producirle una impresión muy agradable y que le recordaba la experiencia de la visión anterior. Se ha iniciado un estudio de la naturaleza de este fenómeno así como sobre las condiciones en que se puede dar.

2. Aspectos metodológicos

Dada la hipótesis, se plantea una experimentación en la que se explora la percepción y la capacidad de reconocimiento de patrones espaciales de fuentes sonoras reales, i.e., figuras construidas mediante altavoces y que se hacen sonar de una determinada manera. Posteriormente se desarrolla un prototipo de laboratorio (descrito más adelante) que hace percibir a un sujeto la situación virtual de que los objetos presentes en el campo perceptivo frontal del mismo, sin estarlo, parezcan estar recubiertos de fuentes sonoras emitiendo el estímulo. Las recientes técnicas de **auralización o tridimensionalización** de sonidos [Wightman & Kistler, 1989; para una introducción general al tema: Begault, 1994] y el desarrollo de los sistemas de **captación y reconocimiento del entorno** de la visión artificial [Faugheras, 1993], ofrecen la posibilidad de una generación de dicho efecto virtual. De esta manera, se explora si esta codificación particular del entorno puede llegar a ser de utilidad práctica para el invidente.

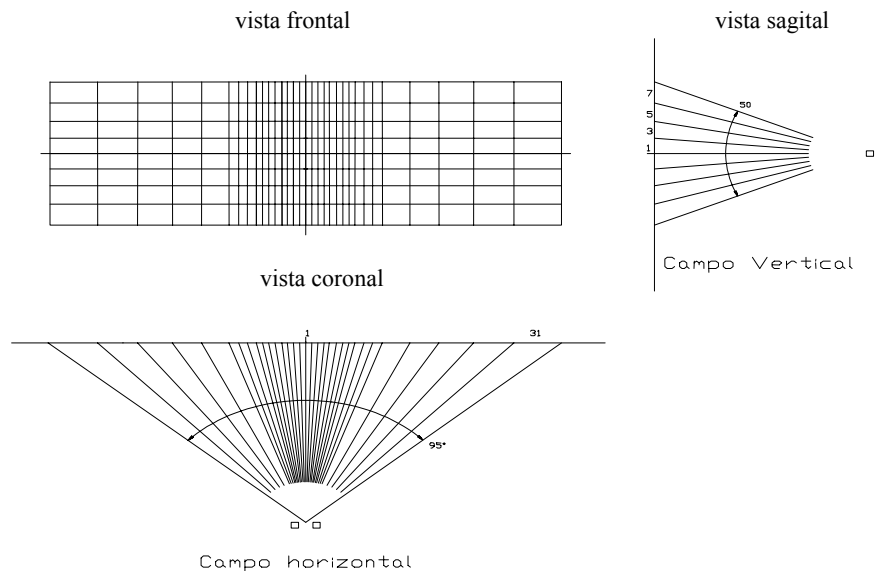
La técnica de auralización empleada, como se explicará más adelante, requiere la realización de sesiones de grabación para cada sujeto de los sonidos correspondientes a las posiciones que se quieren presentar de forma virtual. Estas sesiones pueden prolongarse excesivamente en el tiempo. Entre distintos abordajes posibles para solucionar esta cuestión (uno de ellos, la interpolación de unas funciones a partir de un número menor de funciones medidas), se han realizado estudios para evaluar la posibilidad de sustituir a cada sujeto por una reproducción anatómica precisa del mismo. En colaboración con el departamento de Escultura y Pintura de la facultad de Bellas Artes de la Universidad de la Laguna, se ha elaborado en silicona un maniquí individualizado, y se han realizado las primeras pruebas de comparación del comportamiento acústico del mismo en relación al sujeto original. Los resultados iniciales se

muestran favorables a la hipótesis. Actualmente trabajamos en la cuantificación matemática de la similitud gráfica observada.

3. Descripción del prototipo realizado

Se ha construido un prototipo que obtiene información de las distancias de los objetos situados delante del usuario y reproduce un conjunto de sonidos virtualmente espacializados de manera que se perciban como si estuviesen situados donde dichos objetos. Dos microcámaras situadas en un soporte adecuado que el usuario porta en su cabeza recogen la información del entorno y la envían a los ordenadores, los cuales tras las oportunas operaciones enviarán a los auriculares que el usuario tiene puestos el conjunto de sonidos correspondiente. El usuario tiene así la sensación de que los objetos emiten sonidos, y generan todos ellos un mapa de sonidos virtuales que le permiten extraer información de la forma y volumen de su entorno.

El espacio que rodea al sujeto debe ser discretizado, es decir, transformado de un todo continuo con infinitas distancias a considerar, a un entorno discontinuo con un número finito de distancias a reproducir, con el objeto de que sea manejable electrónica/informáticamente. A tal efecto se trabaja con un sistema de coordenadas esféricas, cuyo centro es la cabeza del sujeto, y del que únicamente reviste interés el sector esférico que se encuentra inmediatamente delante del individuo. Dicho sector se descompone en "estereopíxeles", en cada uno de los cuales estará situada una porción del entorno, con un valor de profundidad (distancia respecto al usuario) determinado.

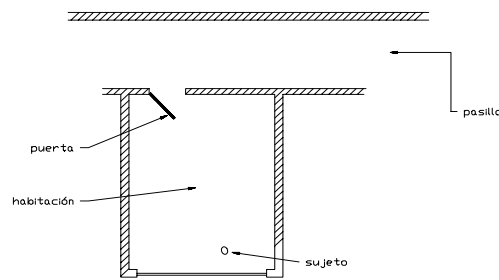


Descomposición en estereopíxeles de un campo rectangular.

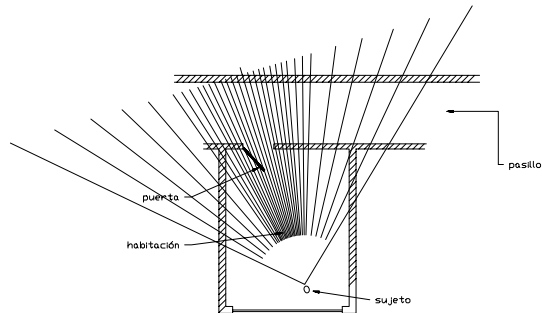
La figura anterior representa esquemáticamente un ejemplo de distribución de estereopíxeles en un campo rectangular, de 32×8 elementos, los cuales tienen ángulos diferentes desde el punto de vista del sujeto, de tal forma que la resolución en el centro del campo aparece reforzada frente a los laterales.

A modo de ejemplo, y al objeto de visualizar el efecto de la división del entorno en estereopíxeles sobre la percepción del mismo, seguidamente se representa un caso en el que el sujeto se encuentra en una habitación normal, dirigiendo su campo de visión a una puerta entreabierta.

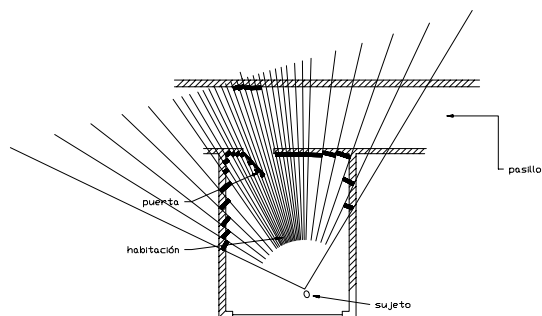
a) Entorno ejemplo:



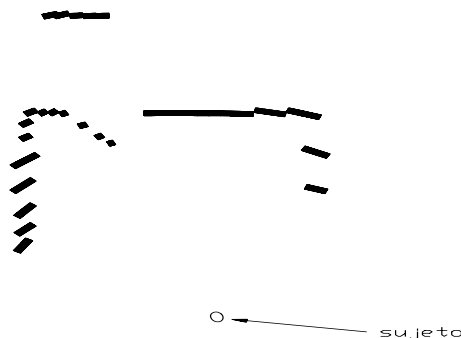
b) Trazado de los estereopíxeles:



c) Determinación de valores de los estereopíxeles:



d) Percepción resultante:



Puede observarse la manera en la que sobre el entorno inicial (a), tras el trazado de los estereopíxeles (b), se pueden obtener los valores de las distancias de las porciones del entorno encerradas por los mismos (c), y con ello la percepción a la que el conjunto daría lugar (d).

El número de estereopíxeles finalmente escogido ha sido de 17 en azimut (horizontal) × 9 en elevación, con 8 profundidades posibles por estereopíxel.

La solución técnica conceptual para el prototipo consiste en dos subsistemas independientes, Subsistema de Visión y Subsistema de Sonidos, los cuales se comunican mediante un enlace de comunicaciones para la transmisión de las informaciones pertinentes entre ellos. Esto se muestra en la figura siguiente. El subsistema de captación o sistema de visión artificial registra las formas del entorno del sujeto mediante dos microcámaras de video situadas en la cabeza a la altura de los ojos, colocadas adecuadamente sobre un soporte diseñado al efecto. Una vez el subsistema de visión ha calculado el mapa de profundidades del entorno, envía esta información al subsistema de sonidos. Éste selecciona y suma los sonidos adecuados a reproducir, y envía esta señal acústica a unos auriculares que también forman parte del soporte que el usuario porta. Toda esta operación se realiza unas 10 veces por segundo.

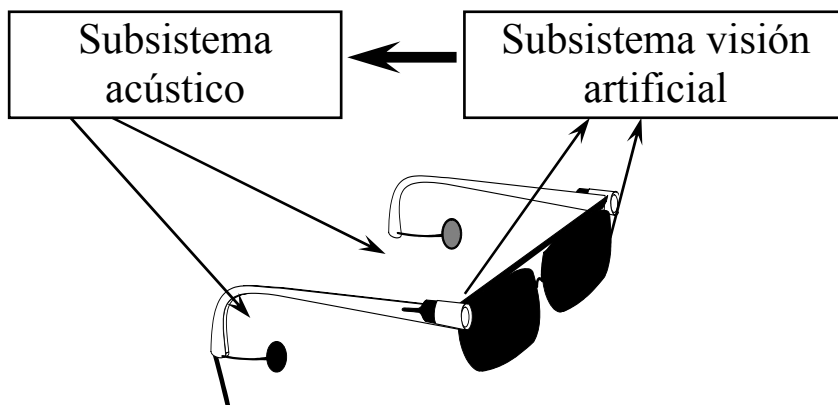


Diagrama conceptual del prototipo desarrollado.

Para que los sonidos se escuchen por los auriculares pero se perciban provenientes del entorno exterior, es necesario haberlos procesado previamente con las denominadas HRTFs (Head Related Transfer Functions), que deben haber sido medidas para cada usuario del prototipo, para cada una de las posiciones de interés.

3.1. Subsistema de visión. Para el sistema de captación del entorno en los comienzos del proyecto se estudiaron varias posibilidades, considerándose en un principio como requerimientos la utilización de sistemas pasivos de percepción del entorno que ofreciesen un rango de medidas amplio, posibilidades de mejora considerable en requerimientos de resolución en cualquiera de las tres dimensiones espaciales, y captación de características del entorno que aportasen información adicional a la espacial.

La solución escogida fue la utilización de sistemas pasivos de visión artificial basados en la adquisición de imágenes con una o varias cámaras CCD y su posterior procesamiento mediante computadores digitales, para la determinación del mapa de profundidades y otras propiedades de los objetos como color, textura, etc. El trabajo inicial se ha centrado únicamente en la obtención del mapa de profundidades, para el envío mediante sonidos de información simplemente espacial, con idea de validar la idea de percepción de superficies y formas mediante sonidos.

Se han estudiado dos técnicas para la obtención del mapa de profundidades de la escena:

- **Estereovisión o visión estereoscópica.** Se basa en la diferencia existente entre las imágenes de un objeto visto simultáneamente desde dos posiciones diferentes.
- **Profundidad a partir del desenfoque (DFD depth from defocus).** Estos métodos utilizan dos imágenes simultáneas de un objeto tomadas desde la misma posición y con distintos parámetros ópticos (focal de las lentes, diámetro del diafragma de apertura, distancia entre lente y sensor), lo cual da lugar a que el desenfoque en cada punto sea distinto en ambas imágenes y dependiente de la distancia entre la cámara y el objeto.

Actualmente el prototipo implementa únicamente el método de estereovisión. En un futuro se pretende fusionar ambos métodos para aprovechar las ventajas de cada uno de ellos.

Estereovisión. Suponemos dos cámaras con objetivos de distancia focal f separadas una cierta distancia b . Un punto (x, y, z) del espacio se proyectará en los puntos (x_1, y_1) y (x_2, y_2) de los planos imagen de cada cámara. Suponiendo que las cámaras tienen los ejes ópticos paralelos entre sí y perpendiculares a los planos imagen, la distancia z se puede obtener a partir de la diferencia en las coordenadas x : $z = bf/d$, donde $d = x_1 - x_2$ se conoce como disparidad. Para conocer la distancia, por tanto, tenemos que encontrar en una imagen el punto que corresponde a un punto dado de la otra imagen: es el llamado problema de la correspondencia.

En la práctica es necesario un paso previo al proceso de correspondencia, que es el proceso de rectificación y consiste en una transformación de las imágenes obtenidas de manera que se consigan las imágenes que corresponderían al caso ideal (cámaras paralelas, lentes sin distorsión). En las imágenes rectificadas, los pares de puntos correspondientes tienen la misma coordenada y , lo cual permite restringir la búsqueda a los píxeles que están en la misma fila que el pixel original.

Para poder aplicar la rectificación y para poder calcular la profundidad z a partir de la disparidad d necesitamos conocer los valores de los parámetros ópticos del sistema (coeficientes de distorsión y distancia focal de las lentes, ángulos de rotación de una cámara respecto a la otra, etc.). Estos valores se calculan en un procedimiento previo de calibración del sistema, basado en tomar 5 pares de imágenes de un patrón de calibración (consistente en cuadros blancos y negros de un determinado tamaño previamente conocido) y buscar las coordenadas imagen de sus esquinas.

El problema fundamental de la obtención de profundidad mediante estereovisión es, pues, el problema de la correspondencia; es decir, dado un pixel de una imagen, buscar (en la

misma fila) el pixel correspondiente en la otra imagen y obtener de esta manera un mapa de disparidades a partir del cual se puede obtener el mapa de profundidades. El algoritmo de correspondencia usado por nosotros consta de los siguientes pasos:

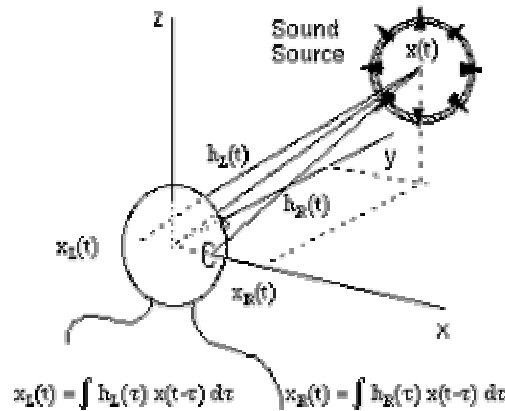
- Rectificación de las imágenes tomadas por las cámaras.
- Extracción de características: convolución con el laplaciano de la gaussiana. El par de imágenes así obtenido contiene información sobre los contornos de las imágenes originales.
- Correlación mediante la suma del valor absoluto de las diferencias de ambas imágenes a lo largo de una pequeña ventana. Para cada pixel de la imagen izquierda se busca el máximo de la correlación en la imagen derecha, hasta un valor máximo de disparidad. Se interpola la posición del máximo de correlación hasta una precisión de $\frac{1}{4}$ pixel.
- Filtrado: se eliminan correspondencias poco fiables. En este paso se emplea una medida de confianza basada en la energía del contorno y también se hace una prueba de consistencia derecha/izquierda (es decir, se busca el pixel en la imagen izquierda que corresponde a un determinado pixel de la imagen derecha y se comprueba que la correspondencia sea la misma en ambos sentidos).

3.2. Subsistema de sonidos. Como repuesta a la información proporcionada por el sistema de captación del entorno, el subsistema acústico se encarga de reproducir sonidos espacializados, es decir, escuchados por el usuario a través de los auriculares pero percibidos como si estuviesen situados *en* los objetos del entorno. Los sonidos han sido procesados previamente fuera de tiempo real de manera que se perciban situados en el entorno exterior al sujeto. Posteriormente se han almacenado en la memoria del subsistema de sonidos. El subsistema selecciona en cada momento los sonidos que correspondan, los suma y reproduce el resultado por los auriculares.

Las tareas que realiza el subsistema de sonidos son:

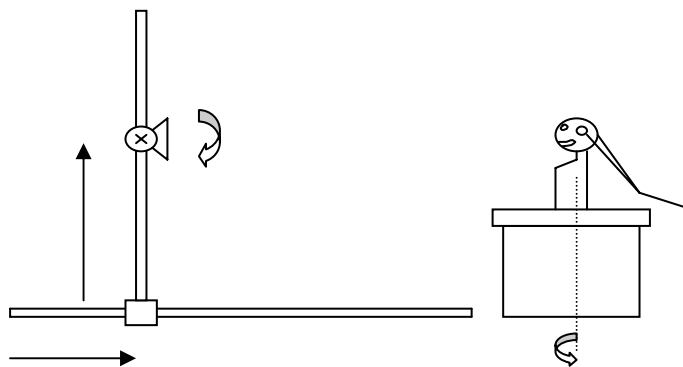
- Recepción de los mapas de profundidades enviados por el subsistema de visión
- Cálculo del orden de presentación de sonidos correspondientes al conjunto de estereopíxeles recibidos, aleatorio y diferente para cada imagen. Esta aproximación evita patrones sonoros repetitivos que distraen la atención del usuario respecto de la información espacial contenida en los sonidos.
- Reproducción de los sonidos correspondientes a cada estereopíxel, uno a uno, manteniendo entre ellos una separación temporal programable. En este momento sólo se utiliza un tipo de sonido, corto e impulsivo, sin cualidades tonales.

Para la correcta percepción del sonido en tres dimensiones, es necesario procesarlo con las denominadas HRTF (Head Related Transfer Functions). La transformación que sufre una señal acústica desde que se emite en una determinada posición espacial relativa al oyente, rebota en todas las superficies del entorno, incluidas las del cuerpo del oyente, hasta que excita el tímpano de este último, se considera una transformación lineal, y por lo tanto, caracterizable por medio de una función de transferencia. Para cada posición espacial relativa al oyente tendremos un par de funciones de transferencia, una para cada oído, y las llamamos HRTFs.



Head Related Transfer Functions.

Se ha desarrollado un sistema de medidas de HRTFs que permite medir HRTFs individuales para cada usuario del prototipo. Un robot manipulador sitúa un altavoz respecto del usuario en la posición espacial relativa de interés. Se reproduce un sonido especial de medida, denominado MLBS (Maximum Length Binary Sequence) a través del altavoz y se graban las respuestas de dos micrófonos miniatura situados en los canales auditivos del usuario (o bien de un maniquí réplica del busto del mismo). El proceso se repite para todas las posiciones espaciales de interés, y además para el caso en que el sonido se reproduzca desde los auriculares seleccionados para el prototipo. Esto es así porque los sonidos que finalmente se reproducen han de estar ecualizados para compensar la reproducción a través de esos auriculares en concreto. El sistema descrito se sitúa dentro de una sala acondicionada para tener el suficiente aislamiento acústico y un pequeño grado de reverberación.



Grados de libertad de movimiento del robot manipulador de altavoz y voluntario.

La HRTF pura no la podemos medir, puesto que las funciones que medimos siempre estarán contaminadas de la respuesta de altavoz, micrófono y demás elementos de la cadena de audio. Sin embargo, el método es válido, y para entenderlo, quizás convenga formular matemáticamente el problema.

Sea $x_1(n)$ el registro digital del sonido a reproducir por un altavoz, y sea $y_1(n)$ el registro resultante de grabar con el micrófono situado en el canal auditivo del voluntario la respuesta a $x_1(n)$. Similarmente, sea $x_2(n)$ el sonido a reproducir por un auricular e $y_2(n)$ la respuesta grabada del mismo. Nuestro objetivo es entonces, para un determinado $x_1(n)$, generar un $x_2(n)$ tal que la resultante $y_2(n)$ sea igual a $y_1(n)$, ya que así habremos conseguido que un estímulo sonoro generado desde un altavoz y otro generado desde un auricular produzcan la misma repuesta en el canal auditivo del individuo, y por tanto, la misma impresión acústica y espacial. Esto lo hacemos diseñando un filtro digital que transforme $x_1(n)$ en $x_2(n)$.

La descripción del diseño del filtro es más sencilla en el dominio de la frecuencia. Llamemos X_1 a la representación de $x_1(n)$ en el dominio transformado de la frecuencia, Y_1 a la de $y_1(n)$, etc. La respuesta registrada como resultado de reproducir $x_1(n)$ se puede escribir como:

$$Y_1 = X_1 LFM \quad (1)$$

donde L , por simplicidad, es la función de transferencia conjunta del altavoz más toda la cadena de audio de reproducción, F es la función de transferencia del espacio libre entre altavoz y canal auditivo, que hemos llamado anteriormente *HRTF*, y M es la función de transferencia conjunta de micrófono más toda la cadena de audio de grabación. La respuesta registrada por el micrófono al reproducir $x_2(n)$ a través del auricular se puede expresar como:

$$Y_2 = X_2 HM \quad (2)$$

donde H es la función de transferencia del auricular (y toda la cadena de reproducción) al canal auditivo. Haciendo $Y_1 = Y_2$ y despejando X_2 obtenemos:

$$X_2 = X_1 LF / H \quad (3)$$

lo cual muestra que el filtro digital que buscamos viene dado por:

$$T = LF / H \quad (4)$$

Por tanto, si la señal $x_1(n)$ la filtramos con este filtro y la señal resultante $x_2(n)$ la reproducimos por el auricular, la señal registrada por el micrófono situado en el canal auditivo cabe esperarse sea $y_1(n)$, la misma que se registraría al reproducir $x_1(n)$ por el altavoz.

El filtro descrito por (4) se refiere a un único altavoz en una determinada posición espacial y un sólo oído. Necesitaremos dos filtros, uno para cada oído, para sintetizar el estímulo que simule cada fuente sonora en una determinada posición espacial.

Medimos las funciones de transferencia del tipo Y_1 / X_1 , a las que denominamos ‘funciones de transferencia altavoz-micrófono’ G_{LM} , para las diferentes posiciones espaciales de interés y para los dos oídos simultáneamente:

$$G_{LM} = \frac{Y_1}{X_1} = L \cdot F \cdot M \quad (5)$$

Posteriormente medimos para los dos oídos, simultáneamente, las dos funciones de transferencia Y_2 / X_2 , a las que denominamos ‘funciones de transferencia auricular-micrófono’ G_{HM} :

$$G_{HM} = \frac{Y_2}{X_2} = H \cdot M \quad (6)$$

Los filtros digitales necesarios para la simulación de un sonido proveniente de las posiciones espaciales de interés (a los que denominamos ‘filtros totales’) se obtienen multiplicando las funciones altavoz-micrófono G_{LM} de cada oído por la inversa de la función auricular-micrófono G_{HM} del mismo oído (ver ecuación (4)), así para ambos oídos:

$$T = \frac{G_{LM}}{G_{HM}} = \frac{L \cdot F \cdot M}{H \cdot M} = \frac{L \cdot F}{H} \quad (7)$$

Tanto para la medida de la función de transferencia altavoz-micrófono como para la del auricular-micrófono se ha utilizado la técnica de medida de respuestas impulsionales por medio de secuencias de máxima longitud (MLBS, de ‘Maximum Length Binary Sequences’), con la posterior correlación cruzada entre la respuesta del sistema y la MLBS de entrada.

La respuesta impulsional del sistema a identificar se obtiene por medio de una correlación cruzada circular entre la MLBS de entrada al sistema y la respuesta a la salida. Es decir, si excitamos al sistema con una MLBS que denominamos $s(n)$, y medimos la a la salida la señal $y(n)$ durante el intervalo de tiempo que dura la MLBS, la respuesta impulsional $h(n)$ viene dada por:

$$h(n) = \Omega_{sy}(n) = s(n) \Phi y(n) = \frac{1}{L+1} \sum_{k=0}^{L-1} s(k) \cdot y(n+k) \quad (8)$$

donde Φ representa la operación correlación cruzada circular o periódica; por tanto, $h(n)$ es una respuesta impulsional periódica, contaminada de aliasing temporal, y no una respuesta impulsional pura. Siempre y cuando se escoja una secuencia lo suficientemente larga, el aliasing resultante se podrá despreciar.

Puesto que la implementación directa de la ecuación para secuencias largas es inviable por el tiempo de cálculo que ello conllevaría, se ha utilizado la equivalencia entre correlación y convolución cruzadas periódicas y se ha pasado al dominio frecuencial, donde la operación de convolución se traduce en una multiplicación de vectores. La mencionada equivalencia es:

$$a(n) \Phi b(n) = \frac{1}{L+1} a(-n) * b(n) \quad (9)$$

donde la inversión de la primera secuencia es circular, al igual que la convolución. Aún así el tiempo de cómputo es excesivo, puesto que las FFTs utilizadas son de longitud 2^k-1 , y no de longitud potencia de dos, que es cuando realmente son eficientes.

Finalmente, se ha conseguido reducir el tiempo de cómputo en dos órdenes de magnitud, empleando la FHT (Fast Hadamard Transform). La ecuación que rige el cálculo de la respuesta impulsional $h(n)$ es:

$$h(n) = \frac{1}{(L+1)s[0]} P_2 \langle S_2 \{ H_{L+1} [S_1 (P_1 y(n))] \} \rangle \quad (10)$$

donde se ha excitado el sistema con una MLBS $s(n)$ de longitud L , y se ha registrado la respuesta $y(n)$. Las matrices P son matrices de permutación, las matrices S son matrices de redimensionamiento, y H_{L+1} es la matriz de Hadamard de orden $L+1$.



Subsistema de generación/procesado de sonidos. A la izquierda, rack donde está ubicado el sistema Huron, con 40 salidas analógicas, 8 entradas analógicas y 8 DSPs 56002. Se ha utilizado para experimentos como el “multifuente”, para el sistema de medida de HRTFs y para el subsistema de sonidos del prototipo. A la derecha, ordenador en el que se calculan las HRTFs y los sonidos espacializados off-line.

4. Resultados con el prototipo actual

Los primeros resultados se basan en las respuestas de una muestra de diez colaboradores invidentes, heterogénea en cuanto a sexo, edad, grado, duración y causa de la ceguera, etc., si bien en su mayoría son personas caracterizadas por tener desarrollado un fino sentido de la audición y de la orientación en entornos conocidos. Tal respuesta se ha medido a partir de la descripción verbal y del registro informatizado de gestos o señalamientos realizados por el voluntario para describir las distintas figuras y patrones espaciales presentados, así como en la realización de tareas de orientación y movilidad concretas, tanto ante fuentes reales como virtuales.

Presentados de una manera general, y considerando que los objetos a percibir varían desde formas puntuales hasta, mono, bi y tridimensionales (lo cual incluye, líneas horizontales o verticales, superficies planas y curvas cóncavas o convexas, aisladas o agrupadas componiendo figuras, como por ejemplo, cuadrados, o columnas o filas paralelas, etc.), podríamos destacar los siguientes aspectos observados:

- Un objeto presente en el campo de percepción del usuario, a partir de la información sonora recibida, puede ser descrito correctamente, en aspectos espaciales significativos como son: su posición, su distancia y las dimensiones en los ejes horizontal y vertical e incluso en el eje z de profundidad.
- Dos objetos separados por un determinado espacio, todo dentro del campo perceptivo captado por el sistema, pueden ser percibidos igualmente en sus posiciones, distancias - que pueden ser distintas, más próximas o más lejanas -, y con sus límites externos e internos.
- Los objetos, transcurrido un breve periodo de exploración, habitualmente inmediato, son percibidos en su disposición espacial de una manera global, i.e., presentes como un todo en el seno del campo perceptivo (por ejemplo, los lados de un cuadrado presentados en las distintas combinaciones posibles: una o dos líneas horizontales; una o dos verticales, etc., hasta el cuadrado total).

El invidente puede reconocer la presencia de un panel o superficie rectangular en su posición, a su distancia correcta, y con sus dimensiones en anchura y en altura. Se percibe además la condición de *superficie*, de continuidad espacial, en un grado suficiente –si bien esperamos mejorable- como para, como suelen referir los sujetos, poder decir que “por ahí no pasaría”. Dos paneles a modo de paredes formando un pasillo, son percibidos como objetos a ambos lados, con sus dimensiones verticales y de profundidad, y que dejan en medio un espacio de silencio por el que se puede pasar y de hecho pasa el invidente.

Los resultados son pues muy alentadores en las pruebas realizadas con personas ciegas. Se ha comprobado que es posible a partir del estímulo acústico recibido notar la presencia, en su sitio y con sus dimensiones, de los objetos captados y que no eran en un principio sonoros en sí mismos.

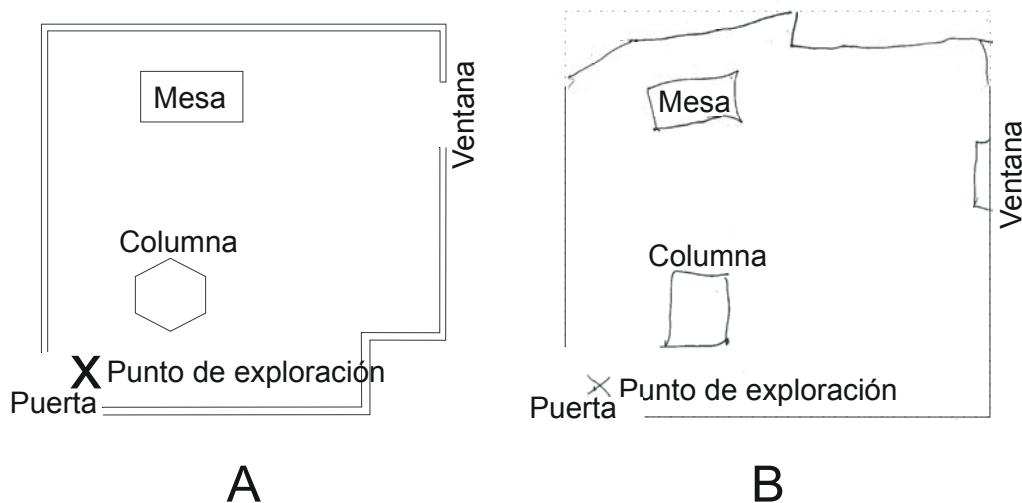
No se trata de que se localizan muchos sonidos en sus posiciones más o menos alejadas cada uno, como los sonidos que se pueden apreciar en una reunión multitudinaria, sino que **se percibe un solo sonido que se extiende, forma una superficie continua y termina en unos bordes netos que puedes delimitar, y que además tiene volumen.**



Voluntario señalando los límites de una apertura a modo de ventana en un panel del laboratorio.

De esta forma, la persona es capaz por medio del sonido de reconocer una **columna** o la presencia y dimensiones del hueco de una **ventana** en el seno de una **pared**, pasar a través del hueco de una **puerta** o caminar dejando al lado la pared de un **pasillo**.

Simulando las condiciones de la vida cotidiana en una habitación experimental relativamente simple, ha sido posible para la persona ciega moverse y extraer información suficiente para hacer una descripción verbal y gráfica de la escena global .



A. Croquis de una habitación experimental con una disposición de objetos desconocida para una usuaria invidente. B. Dibujo de la habitación con la distribución de los objetos reconocidos por la sujeto usuaria del dispositivo tras una exploración del entorno sin apoyarse en el tacto.

Por otro lado, el estímulo utilizado permite la captación y comprensión de otros sonidos procedentes del ambiente como ruidos o una conversación.

Otro resultado obtenido en las pruebas ha sido la aparición de los denominados **fosfenos** evocados por sonidos. En algunas de las personas ciegas, determinados sonidos con los que hemos estado trabajando provocan una sensación de luz (denominada fosfeno). Son como luciérnagas o estrellitas que la persona percibe simultáneamente en el mismo punto donde oye la fuente sonora, o el escenario acústico.

Existen referencias a este fenómeno pero con muy pocos casos y además la mayoría transitorios y sin esa relación espacial tan marcada entre posición oída y posición de las luces [Lessel & Cohen, 1979; Page et al, 1982].

Este hecho ha dado lugar a la apertura de una línea de investigación que pensamos podría tener importantes implicaciones en nuestro objetivo global de ayuda a la orientación y la movilidad, así como en otros aspectos clínicos y de rehabilitación en la ceguera.

Los primeros estudios con la técnica de **Resonancia Magnética funcional**, que muestra cortes del cerebro donde se aprecian las áreas cerebrales que se activan durante la realización de una determinada tarea mental (perceptiva, imaginativa, memorística, de control de movimientos, etc.) muestran una activación de la corteza visual cuando el sujeto, mediante el estímulo sonoro, percibe fosfenos.

Como en toda investigación, especialmente en las ciencias de la vida, la estadística juega un papel muy importante tanto en el diseño experimental de las pruebas planteadas como en el análisis de los resultados obtenidos.

5. Líneas de investigación y desarrollo abiertas

En la actualidad se está culminando el desarrollo de un **dispositivo portátil** para su utilización cotidiana en **ambientes habituales**, es decir, ya fuera de los límites del laboratorio, por personas ciegas. Esto permitirá la consiguiente evaluación de validez y utilidad del prototipo.

Además, se continúa la necesaria línea de investigación para la mejora del estímulo sonoro de cara a conseguir las percepciones de una manera más fina, así como de los distintos elementos implicados en el fenómeno perceptivo bajo estudio (papel de los factores individuales, papel del aprendizaje, etc.).

Paralelamente se mantiene la línea de investigación neurofisiológica destinada a valorar el potencial de adaptación del cerebro humano ante este tipo de estímulo particular y ante la posibilidad de generar sensaciones luminosas de forma controlada mediante estímulos auditivos.

Las matemáticas pueden ejercer un papel determinante en la consecución de toda una serie de mejoras del dispositivo actual. La obtención de las HRTFs individuales, puede beneficiarse del desarrollo de los oportunos métodos de obtención de funciones a partir de un número limitado de las mismas, normalmente referidos en la literatura científica como métodos de interpolación. La obtención de una información de profundidad (mapa de profundidad) estable y precisa, al menos con respecto a las técnicas actuales de estereovisión, se puede beneficiar del desarrollo de métodos adecuados para la eliminación de ruidos por inestabilidades periódicas de las imágenes de partida, etc. etc.

Ha surgido también la idea de utilizar este estímulo acústico para generar una interfaz entre la persona invidente y el ordenador, concretamente en forma de un juego para los niños. Éste podría tener valor, además del meramente lúdico, en la por otro lado dificultosa introducción de los niños al ferviente mundo actual de la informática..

Creemos que nuestro prototipo será una herramienta de ayuda más a la movilidad de personas ciegas o con deficiencia visual severa, como lo es por ejemplo también el bastón, con el que se podría utilizar de forma conjunta y complementaria. Además de **seguridad** en los

desplazamientos aporta **nuevas sensaciones**, como son la de **volumen, globalidad**, etcétera. **Es un dispositivo con el que puedes seleccionar el objeto o la escena que quieres reconocer con sólo girar la cara hacia donde están situados.**

También consideramos que puede ser de gran interés la aplicación de un sistema de este tipo como **herramienta de rehabilitación** en el **desarrollo del niño** ciego congénito o con déficit visual severo temprano. La ausencia de información del mundo distante, con todas las repercusiones que ello conlleva para el desarrollo tanto psicomotor como cognitivo e incluso emocional, podría compensarse con un estímulo acústico que se comporta como lo haría uno visual.

Los resultados, muy alentadores, no nos ocultan sin embargo que estamos a las puertas de un área de investigación muy reciente y en la que queda mucho por desvelar.

Reconocimientos

Este es un proyecto desarrollado por la Universidad de la Laguna y el Instituto Astrofísico de Canarias en colaboración con el Instituto Tecnológico de Canarias, la O.N.C.E., el Hospital Universitario de Canarias, el Instituto Canario de Investigación y Desarrollo, Imásde y pixel 486, con financiación a lo largo de sus diferentes etapas por fondos de la Consejería de Economía y Hacienda del Gobierno de las Islas Canarias, IMSERSO (Ministerio Español de Trabajo y Asuntos Sociales), fondos FEDER y fondos del Ministerio Español de Ciencia y Tecnología (TIC2001-3916 y FIT-070000-2002-549). **Nuestro mayor agradecimiento a las personas colaboradoras en la investigación, sin las cuales no hubiera sido posible su realización.**

Bibliografía

- Andersen, R. A. MULTIMODAL REPRESENTATION OF SPACE IN THE POSTERIOR PARIETAL CORTEX AND ITS USE IN PLANNING MOVEMENTS. *Annual Review Neuroscience*, **20**, 303-330 (1997).
- Bach-Y-Rita, P. et al. VISUAL SUBSTITUTION BY TACTILE IMAGE PROJECTION. *Nature*, **1221**, 963-964 (1969).
- Begault, D. R. 3-D SOUND FOR VIRTUAL REALITY AND MULTIMEDIA, 1994.
- Benjamin, J. M. THE LASER CANE. *Bulletin of Prosthetics Research, Veterans Administration, Washington*, 1974 .
- Bregman, A. S. AUDITORY SCENE ANALYSIS: THE PERCEPTUAL ORGANIZATION OF SOUND. The MIT Press, 1990.
- Doods, A. G. et al. THE NOTTINGHAM OBSTACLE DETECTOR: DEVELOPMENT AND EVALUATION. *Ibid.*, **75**, 203-209 (1981).
- Faugeras, O. THREE-DIMENSIONAL COMPUTER VISION: A GEOMETRIC VIEWPOINT. The MIT Press, 1993.
- González Mora, J.L., Rodríguez-Hernández, A., Rodríguez-Ramos, L.F., Díaz-Saco, L., Sosa, N. DEVELOPMENT OF A NEW SPACE PERCEPTION SYSTEM FOR BLIND PEOPLE, BASED ON THE CREATION OF A VIRTUAL ACOUSTIC SPACE. *Proceedings of the International Work Conference on Artificial and Natural Networks*, Vol. 2, pp 321-330. Springer, 1999.

- Heikkila, Silven. A FOUR-STEP CAMERA CALIBRATION PROCEDURE WITH IMPLICIT IMAGE CORRECTION.
- Horn. ROBOT VISION. The MIT Press, 1986.
- Kanade, Kano, Kimura, Yoshida, Oda. DEVELOPMENT OF A VIDEO-RATE STEREO MACHINE. Proceedings of International Robotics and Systems Conference (IROS 95), August 1995.
- Kanade, Okutomi. A STEREO MATCHING ALGORITHM WITH AN ADAPTIVE WINDOW: THEORY AND EXPERIMENT. PAMI, **16**, no. 9, September 1994.
- Kay, L. ELECTRONICS AIDS FOR BLIND PERSONS: AN INTERDISCIPLINARY SUBJECT. IEEE Review. IEEE Proceedings, **131**, pt. A, no. 7, September 1984.
- Kistler, Doris J. and Wightman, Frederic L. A MODEL OF HEAD-RELATED TRANSFER FUNCTIONS BASED ON PRINCIPAL COMPONENT ANALYSIS AND MINIMUM-PHASE RECONSTRUCTION. J. Acoust. Soc. Am., **91**, 1637-1647(1992).
- Konishi, M. CENTRALLY SYNTHESIZED MAPS OF SENSORY SPACE. TINS, 163-168, April 1986.
- Krotov. ACTIVE COMPUTER VISION BY COOPERATIVE FOCUS AND STEREO. Springer-Verlag, 1989.
- Lai, Chang. A GENERALIZED DEPTH ESTIMATION ALGORITHM WITH A SINGLE IMAGE. PAMI, **14**, no. 4, April 1992.
- Lakatos, S. RECOGNITION OF COMPLEX AUDITORY SPATIAL PATTERNS. Perception. **22**, 363-374 (1993).
- Lessell, S. and Cohen, M.M. PHOSPHENES INDUCED BY SOUND. Neurology, **29**, 1524-1527 (1979).
- Loomis, J.M., Golledge, R.G., Klatzky, R.L. NAVIGATING SYSTEM FOR THE BLIND: AUDITORY DISPLAY MODES AND GUIDANCE. Presence, **7**(2), 193-203 (1998).
- Marr, D. VISION: A COMPUTATIONAL INVESTIGATION INTO THE HUMAN REPRESENTATION AND PROCESSING OF VISUAL INFORMATION. W. H. Freeman and Company, New York, 1982.
- Martens, W. L. PRINCIPAL COMPONENTS ANALYSIS AND RESYNTHESIS OF SPECTRAL CUES TO PERCEIVED DIRECTION. En J. Beauchamp (editor): Proceedings of the International Computer Music Conference, pp. 274-281. International Music Association, San Francisco, CA (1987).
- Nishihara, Poggio. STEREOVISION FOR ROBOTICS. En R. Paul and M. Bradu (editores): Robotics Research, The First International Symposium, pp. 489-505. MIT Press, 1984.
- Page N. G. R. et al. AUDITORY EVOKED PHOSPHENES IN OPTIC NERVE DISEASE. Journal of Neurology, Neurosurgery and Psychiatry, **45**, 7-12 (1982).
- Pentland, Scherrock, Darrell, Girod. SIMPLE RANGE CAMERAS BASED ON FOCAL ERROR. Journal Optical Society of America A, **11**, no. 11.
- Rice, Ch. E. et al. ECHO-DETECTION ABILITY OF THE BLIND: SIZE AND DISTANCE FACTORS. Journal of Experimental Psychology, **70**, 3, 246-251 (1965).
- Rodríguez Hernández, A; Sosa, N , Rodríguez Ramos, L.F. , Chulani, H. Díaz Saco, L, y González-Mora, J.L. PERCEPCIÓN DEL ENTORNO EN PERSONAS CIEGAS A TRAVÉS DE UN ESTÍMULO SONORO ESPACIAL VIRTUAL GENERADO POR COMPUTADOR. Actas del Congreso Iberoamericano 3º de CAA, 1º de Tecnologías de Apoyo para la Discapacidad, pp. 81-84, Octubre 2000.
- Rodríguez Hernández, A.F, González Mora, J.L., Pujol, J., López García, J.A., Muñoz Montes, J.R., Rodríguez Ramos, L.F., Chulani, H.M. and Burunat, E. FUNCTIONAL MAGNETIC RESONANCE IMAGING OF SPATIAL AUDITORY PERCEPTION IN

- BLIND AND SIGHTED HUMAN SUBJECTS. Aceptado para presentación en el XXXII Congreso de la Sociedad Española de Ciencias Fisiológicas, 2003.
- Rodríguez-Ramos, L.F., Chulani, H.M., Díaz-Saco, L., Sosa, N., Rodríguez-Hernández, A., González Mora, J.L. IMAGE AND SOUND PROCESSING FOR THE CREATION OF A VIRTUAL ACOUSTIC SPACE FOR THE BLIND PEOPLE. *Signal Processing and Communications*, 472-475 (1997).
- Sadato, N., Pascual-Leone, A., Grafman, J., Ibáñez, V., Daiber, M.P., Dold, G., Hallett, M. (1996). ACTIVATION OF PRIMARY VISUAL CORTEX BY BRAILLE READING IN BLIND PEOPLE. *Nature*, **380**, 526-527 (1996).
- Schlag, Sanderson, Neumann, Wimberly. IMPLEMENTATION OF AUTOMATIC FOCUSING ALGORITHMS FOR A COMPUTER VISION SYSTEM WITH CAMERA CONTROL. Technical Report CMU-RI-TR **314**, Carnegie-Mellon University, August 1983.
- Shih, Hung, Lin. ACCURATE LINEAR TECHNIQUE FOR CAMERA CALIBRATION CONSIDERING LENS DISTORTION BY SOLVING AN EIGENVALUE PROBLEM. *Optical Engineering* **32**(1), 138-149 (1993).
- Shirai. THREE-DIMENSIONAL COMPUTER VISION. Springer-Verlag, 1987.
- Takahashi, T. T. et al. REPRESENTATION OF MULTIPLE SOUND SOURCES IN THE OWL'S AUDITORY SPACE MAP. *Journal of Neuroscience*, **14**, no. 8, 4780-4793 (1994).
- Thurlow, W. R. et al. HEAD MOVEMENTS DURING SOUND LOCALIZATION. *Journal of the Acoustic Society of America*, **42**, 489-493 (1967).
- Tsai. A VERSATILE CAMERA CALIBRATION TECHNIQUE FOR HIGH-ACCURACY 3D MACHINE VISION METROLOGY USING OFF-THE-SHELF TV CAMERAS AND LENSES. *IEEE Journal of Robotics and Automation* RA -3 (4): 323-344.
- Tsai. AN EFFICIENT AND ACCURATE CAMERA CALIBRATION TECHNIQUE FOR 3-D MACHINE VISION. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 364-374 (1986).
- Wightman, F., Kistler, D.J. HEADPHONE SIMULATION OF FREE FIELD LISTENING I: STIMULUS SYNTHESIS. *Journal of the Acoustical Society of America*, **85**, 858-867 (1989).
- Wei, De, Ma. IMPLICIT AND EXPLICIT CAMERA CALIBRATION: THEORY AND EXPERIMENTS. *IEEE Transactions on PAMI*, **16**, no. 5 (1994).
- Wilson. FOURIER SERIES AND OPTICAL TRANSFORM TECHNIQUES IN CONTEMPORARY OPTICS. Wiley-Interscience. 1995
- Xu, Zang. EPIPOLAR GEOMETRY IN STEREO, MOTION AND OBJECT RECOGNITION: A UNIFIED APPROACH. Kluwer Academic Publishers, 1996.
- Zölzer, Udo. DIGITAL AUDIO SIGNAL PROCESSING. John Wiley & Sons, 1997.

En Internet

<http://www.iac.es/eav>

Espacio Acústico Virtual

Sitio web del Proyecto.

La idea

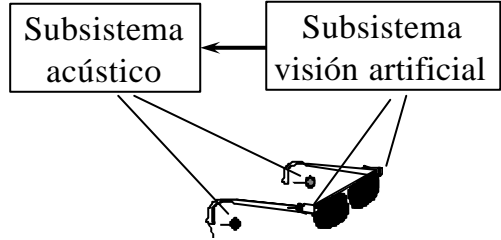
- Si el cerebro recibiese la información espacial que normalmente le proporciona la vista, pero a través del oído, ¿podría componer una imagen 3D del entorno?

El desarrollo tecnológico

- Un aparato de ayuda a la movilidad y percepción 3D para los ciegos

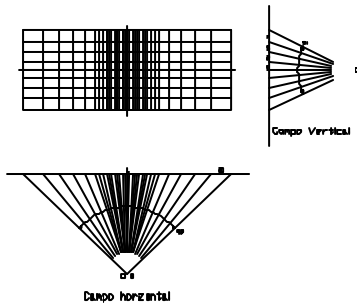
1

El desarrollo tecnológico. El concepto.



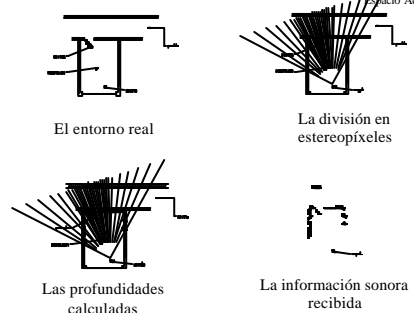
2

El estereopíxel



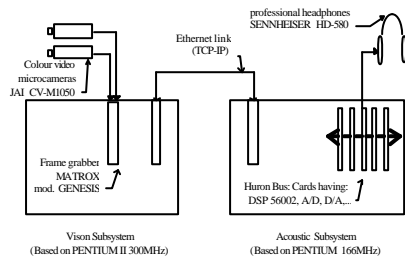
3

Ejemplo 2D



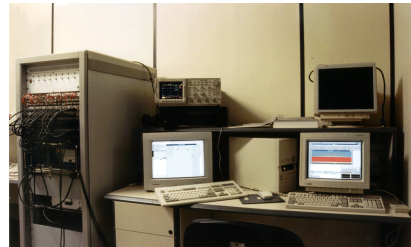
4

El prototipo de validación en laboratorio



5

Sala de control



Subsistema de generación/procesado de sonidos. A la izquierda, rack donde está ubicado el sistema Huron, con 40 salidas analógicas, 8 entradas analógicas y 8 DSPs 56002. Se ha utilizado para experimentos como el "maquillaje", para el sistema de medida de RTDs y para el subsistema de sonidos del prototipo. A la derecha, ordenador en el que se calculan las IRDTs y los sonidos especializados off-line.

6



Sala de experimentación, detalle



El prototipo portátil



Maqueta de las gafas



Maqueta de las gafas



Subsistema de visión

Captación del entorno y generación de un mapa de profundidades



Métodos abordados

- **Visión estereoscópica o estereovisión**
Dos imágenes de la misma escena tomadas desde dos posiciones diferentes
- **Profundidad a partir del desenfoque (DFD Depth From Defocus)**
Dos imágenes de la misma escena con diferente desenfoque

Visión estereoscópica. Operaciones.



Espacio Acústico Virtual

- **Calibración:** Obtención de los parámetros ópticos
- **Rectificación:** Transformación al caso ideal (cámaras paralelas, lentes sin distorsión)
- **Correspondencia:** Obtención de mapa de disparidades
 - Extracción de características
 - Correlación e interpolación
 - Filtrado
- **Cálculo de la profundidad:**

$$z = bf/d, \text{ donde } d = x_1 - x_2$$

13

Depth from Defocus. Resultados.



Espacio Acústico Virtual



Entorno real. Dos imágenes con distinto desenfoque.

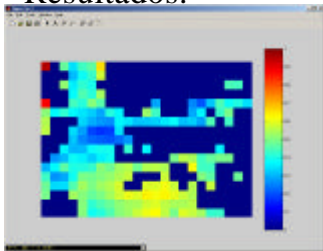
(Centre de Visió per Computador)

14

Depth from Defocus. Resultados.



Espacio Acústico Virtual



Distancias lejanas en azul. Medias en verde. Cercanas en rojo.

(Centre de Visió per Computador)

15

Subsistema acústico

Generación de un estímulo sonoro
formado por un mapa de sonidos 3D



Espacio Acústico Virtual

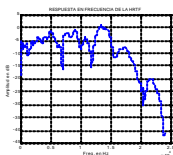
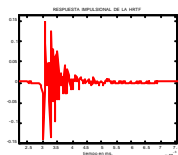
16

Respuesta impulsional y función de transferencia



Espacio Acústico Virtual

Respuesta impulsional: dominio del tiempo Función de transferencia: dominio de la frecuencia



17

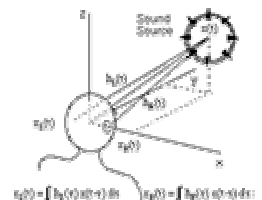
H.R.T.F. Definición.



Espacio Acústico Virtual

HRTF: Head Related
Transfer Function
(Función de
Transferencia Relativa
a la Cabeza)

Depende de la posición
relativa entre fuente
sonora y oyente

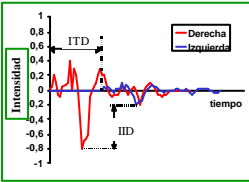


18



Información de azimut

RESPUESTAS IMPULSIONALES



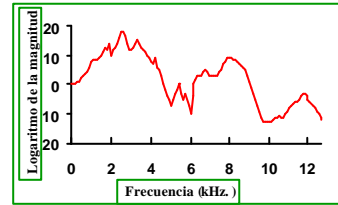
IID: Diferencia interaural de intensidad

ITD: Diferencia interaural de tiempo



Información de elevación

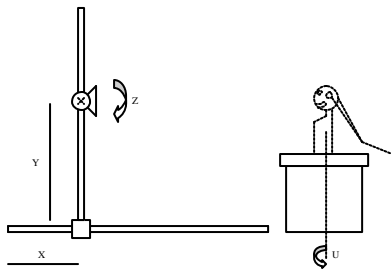
MAGNITUD DE LA HRTF



Crestas y valles en el dominio frecuencial



Stma automatizado de medidas de HRTFs



M.L.B.S. (Maximum Length Binary Sequences)

Entrada (M.L.B.S. de longitud L): $s(n)$ Salida: $y(n)$

Respuesta impulsional:

$$h(n) = \Omega_{sy}(n) = s(n) \Phi y(n) = \frac{1}{L+1} \sum_{k=0}^{L-1} s(k) \cdot y(n+k)$$

Propiedad:

$$\Omega_{ss}(n) = s(n) \Phi s(n) \equiv d(n)$$



Procesado de las medidas

L : Altavoz (más la cadena de reproducción)

F : HRTF

M : Micrófono (más la cadena de grabación)

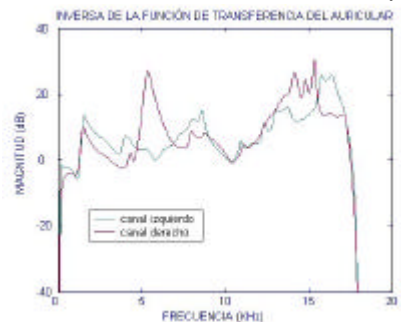
H : Auriculares (más la cadena de reproducción)

Filtro buscado:

$$T = \frac{G_{LM}}{G_{HM}} = \frac{L \cdot F \cdot M}{H \cdot M} = \frac{L \cdot F}{H}$$



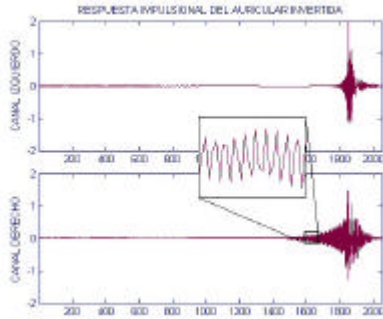
Ejemplo problemático



Ejemplo problemático



Espacio Acústico Virtual

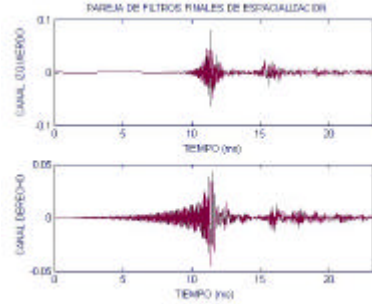


25

Ejemplo problemático



Espacio Acústico Virtual



26

Fast Hadamard Transform



Espacio Acústico Virtual

Problema con la MLBS: Su longitud es $L = 2^k - 1$
 \rightarrow FFT's de longitud $2^k - 1$

Reformulación del cálculo de $h(n)$ con la F.H.T.

• H_{L+1} : Matriz de Hadamard de orden $L+1$, relacionada con la M.L.B.S. $s(n)$

• P_i : Matrices de permutación

• S_i : Matrices de redimensionamiento

$$h(n) = \frac{1}{(L+1)s[0]} P_2 \{ S_2 \{ H_{L+1} [S_1 \{ P_1 y(n) \}] \} \}$$

27

HRTFs en plano horizontal



Espacio Acústico Virtual

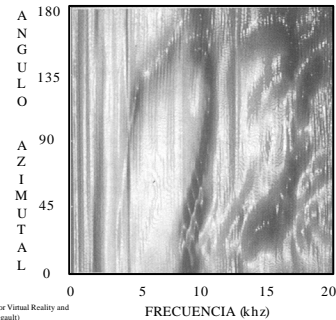


Foto tomada de "3D Sound for Virtual Reality and Multimedia" (Durand R. Begault)

28

HRTFs en plano medio



Espacio Acústico Virtual

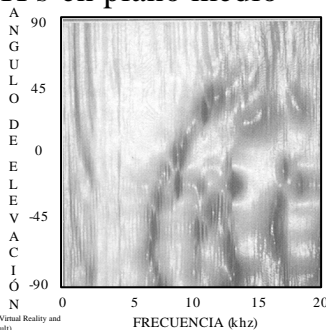


Foto tomada de "3D Sound for Virtual Reality and Multimedia" (Durand R. Begault)

29

Modelado de HRTF's



Espacio Acústico Virtual

- Objetivo: Obtener una representación funcional de las HRTFs dependiente de la frecuencia, posición espacial e individuo.
- Métodos recogidos en la bibliografía:
 - Representación de las HRTFs en función de formas espectrales básicas (funciones básicas)
 - Análisis de componentes principales sobre la magnitud de HRTFs de mínima fase (Wightman & Kistler, University of Wisconsin, Madison, 1991)
 - SFER (Spatial Features Extraction and Regularization) sobre magnitud y fase de las HRTFs (Chen, J. et al., University of Wisconsin, Madison, 1995).
 - Ajuste polinomial sobre los coeficientes (spline, lineal ...)

30



Modelado de HRTF's

- Resultados:
 - Utilización eficiente de la cantidad de memoria y potencia de cálculo disponibles.
 - HRTFs generales (no individuales)
 - Posibilidad de medir la similitud entre HRTFs de distintos individuos
 - distancia euclídeana entre coeficientes de las funciones base
 - análisis de escalado multidimensional
 - Interpolación espacial de HRTFs

PERCEPCIÓN GLOBAL DEL ENTORNO A TRAVÉS DE SONIDOS

PROYECTO ESPACIO ACÚSTICO VIRTUAL

LÍNEAS ACTUALES DE ACTUACIÓN ANTE LA CEGUERA

- PREVENCIÓN.
- RESTABLECIMIENTO DE LA VÍA visual NATURAL.
- RESTABLECIMIENTO DE LA VÍA visual POR PROTESIS SENSORIALES.
- APROVECHAMIENTO Y POTENCIACIÓN DE LOS **OTROS SENTIDOS**.

APROVECHAMIENTO DE LAS CAPACIDADES SENSORIALES INDEMNES

Es posible la realización de tareas propias de un sentido a través de otro u otros:
SUSTITUCIÓN SENSORIAL.

- Orientación y movilidad.
- Desarrollo del niño.
- Lectura.
- Reconocimiento de objetos.
- Interacción hombre - máquina.

PERCEPCIÓN DEL ENTORNO A TRAVÉS DE SONIDOS

OBJETOS SONOROS

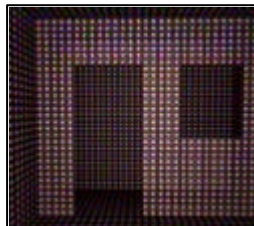
- Localización.
- Identificación.
- Percepción de objetos en su posición y forma.

OBJETOS NO SONOROS

- Dimensiones del entorno.
- Ecolocación.
- Bastón, cascabeles, etc.
- Dispositivos electrónicos:
 - Detectores de obstáculos.
 - Codificación de escenas.

PERCEPCIÓN DE SUPERFICIES SONORAS

- La hipótesis.
- Fundamentos:
 - **Procesamiento amodal**
 - **Reproducción de estrategia.**
 - **Condiciones adecuadas de la audición.**
 - **Precedentes.**



ABORDAJE EXPERIMENTAL DE LA HIPÓTESIS

- Estudio con fuentes sonoras reales
- Estudio con fuentes sonoras virtuales



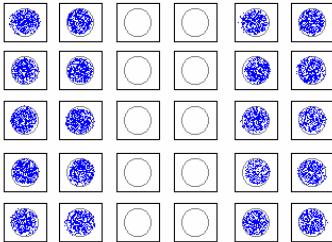


MATRIZ DE ALTAVOCES

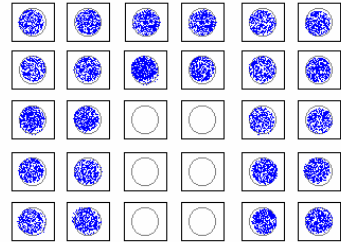
FUENTES SONORAS REALES: EXPERIMENTO MULTIFUENTE

Valoración de la percepción de figuras en el espacio a partir de superficies sonoras.

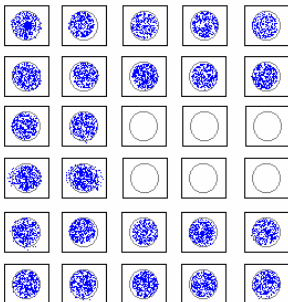
- Cada imagen definida por el conjunto de puntos de la superficie o escena a percibir.
- Imágenes cada 30, 90 y 300 ms.
- Sonido muy localizable.
- Secuencia espacial de presentación al azar.



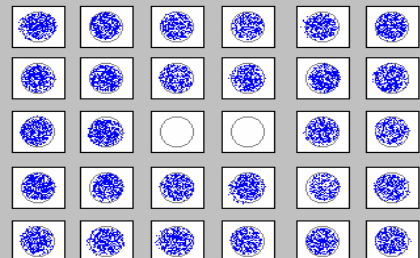
MATRIZ SONORA: BARRAS VERTICALES



MATRIZ SONORA: U INVERTIDA

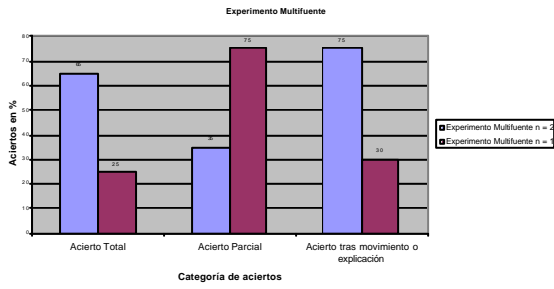


MATRIZ SONORA: C A LA DERECHA



MATRIZ SONORA: VENTANA ABIERTA

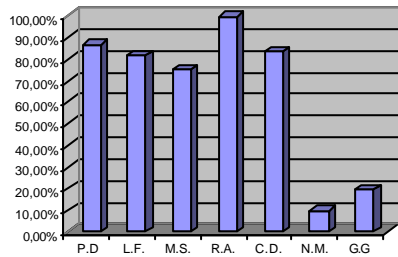
RESULTADOS PRELIMINARES



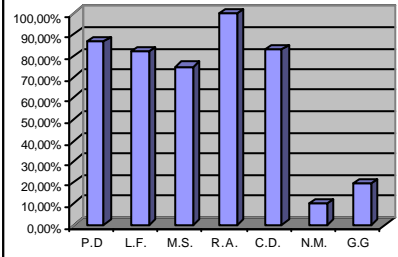
Resultados



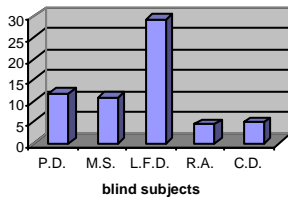
Percentage of correct responses about the shape of the figures



Percentage of correct responses about the shape and orientation of the figures



Mean error in the perception of the dimensions of the figures (cm)



Blind person pointing to two points horizontally located. Both them are perceived simultaneously present



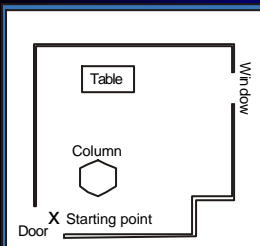
Resultados



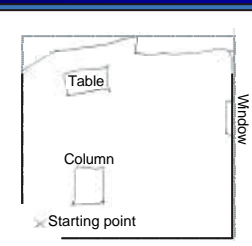
- Detección del entorno
- Reconocimiento de formas con sus tamaños
- Reconocimiento de la habitación con su contenido



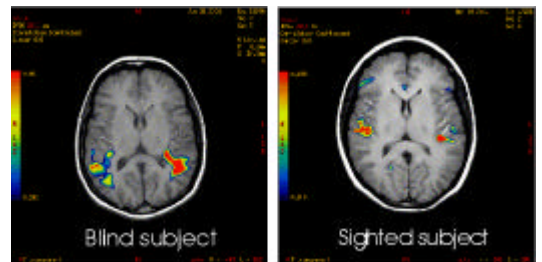
Resultados



Schematic representation of the experimental room, with a particular objects distributions.



Drawing made by a blind person after a very short exploration, using the prototype, without relying in touch



Imágenes de Resonancia Magnética Funcional mostrando las áreas cerebrales implicadas en la localización de una fuente sonora virtual aislada

El futuro



- Continuar el desarrollo del programa de investigación optimizando sonidos, vision artificial, y capacidad del sistema
- Poner a punto el prototipo portátil desarrollado, con voluntarios ciegos.
- Realizar un programa de transferencia tecnológica para el desarrollo de un prototipo comercial.

Reconocimientos

- Gobierno de las Islas Canarias
- Ministerio español de Asuntos Sociales
- Comunidad Europea
- Organización Nacional de Ciegos de España (ONCE)
- Ministerio Español de Ciencia y Tecnología
- Instituto Tecnológico de las Islas Canarias
- Hospital universitario de Canarias
- ImasDe Canarias SA
- Pixel SL



El equipo

Gracias por su atención.

www.iac.es/eav