



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

DEPARTAMENTO DE INGENIERÍA GRÁFICA

Tesis Doctoral

Diseño, integración y  
experimentación de un  
sistema 3D de navegación  
para personas invidentes

Valencia, abril de 2012

Autor: D. Victor M. Santiago Praderas

Director: Dr. D. Guillermo Peris Fajarnés  
Codirectora: Dra. Dña. Larisa Dunai



# INDICE GENERAL

1. INTRODUCCIÓN.....	21
1.1 Generalidades .....	23
1.2 Motivación.....	30
1.3 Estado del arte .....	33
1.3.1 Sistemas Tecnológicos Desarrollados .....	33
1.3.2 Sistemas Patentados (Dispositivos electrónicos de movilidad) .....	62
1.3.3 Publicaciones científicas y patentes .....	74
1.4 Objetivos.....	78
1.5 Conclusiones.....	80
2. CASBLiP, DESCRIPCIÓN DEL PROYECTO .....	83
2.1 Resumen .....	83
2.2 Objetivos del proyecto CASBLiP .....	84
2.3 Diseño y características del sistema .....	89
2.3.1 Prototipo acústico: M1 .....	90
2.3.1.1 Método de medida de distancias con 3D CMOS.....	92
2.3.1.2 Sensor 3D CMOS .....	94
2.3.2 Prototipo de visión: M2 .....	98
2.3.2.1 Módulo de visión artificial .....	99
2.3.2.2 Estimación de profundidad y segmentación de imágenes .....	101
2.3.2.3 Detección de obstáculos y estimación de movimiento...	102
2.3.2.4 Detección de objetos usando mapas de profundidad.....	103
2.3.2.5 Sensor inercial .....	104
2.3.2.6 Sistema GPS .....	104

2.3.3 Interface de audio.....	106
2.3.3.1 Medida de HRTF .....	108
2.3.3.2 Señal de sonido utilizada .....	112
2.3.3.3 Estrategia de sonificación .....	114
2.4 Conclusiones .....	118
3. INTERFAZ DE COMUNICACIÓN E INTEGRACIÓN.....	121
3.1 Resumen.....	121
3.2 Integración de los módulos .....	121
3.3 Conclusiones .....	121
4. VALIDACIÓN .....	129
4.1 Resumen.....	129
4.2 Test de navegación.....	130
4.2.1 Diseño del experimento .....	130
4.2.2 Usuarios .....	143
4.2.3 Resultados .....	144
4.3 Conclusiones .....	164
5. CONCLUSINES GENERALES.....	167
6. FUTUROS PLANES DE TRABAJO.....	171
7. REFERENCIAS.....	173

## INDICE DE FIGURAS

Figura1.1 Informe sobre la contratación a personas con discapacidad en España durante el primer trimestre de 2011 [OED Inf] .....	27
Figura1.2 Población invidente total y por edades en España [OED] ...	28
Figura1.3 Población invidente total por deficiencias causadas en España [OED].....	29
Figura 1.4 Lindsay Pathsounder .....	45
Figura 1.5 Mowat Sonar Sensor .....	46
Figura 1.6 Nottigham Obstacle Detector.....	47
Figura 1.7 Principio de funcionamiento del Laser cane .....	48
Figura 1.8 Sonic Torch .....	49
Figura 1.9 Sonicguide MkII .....	50
Figura 1.10 Sonicguide MkII .....	51
Figura 1.11 Funcionamiento del Talking Signs System.....	52
Figura 1.12 Funcionamiento básico del GPS .....	53
Figura 1.13 Principio de funcionamiento del PGS [Loomis, 1985].....	55
Figura 1.14 Última versión del PGS.....	55
Figura 1.15 Imagen SIG del municipio 900 parcela 6 según SIGPAC	57
Figura 1.16 Sistema de identificación de parcelas agrícolas mediante SIGPAC.....	57

Figura 1.17 Dispositivo NAVI.....	58
Figura 1.18 Dispositivo NAVI - KINECT .....	58
Figura 1.19 Principio funcionamiento del sistema SWAN.....	60
Figura 1.20 Componentes del sistema The vOICe formado por cámaras web 3D modelo Vizix Wrap 920AR HMD .....	61
Figura 2.3.1 Prototipo M1 .....	89
Figura 2.3.2 Prototipo M2.....	89
Figura 2.3.1.1 Gafas del M1 .....	91
Figura 2.3.1.2 FPGA M1 .....	91
Figura 2.3.1.1.1 Principio de medida 3D CMOS.....	93
Figura 2.6 Método de proyección de la luz a través de las lentes ópticas [Bernhard 2008] .....	94
Figura 2.7 Sistema Sensor 3D CMOS .....	95
Figura 2.8 Principio de cálculo de las propiedades del pixel: x=n×l×m×h .....	95
Figura 2.9 Circuito del pixel CDS readout stage [Elkhalili et al] .....	96
Figura 2.10 Componentes del sensor 3D CMOS y su integración en gafas .....	97
Figura 2.11 Mapa de Profundidad .....	98
Figura 2.12 Cámara Fireware Flea 2 color .....	100
Figura 2.13 Fusión de mapas de profundidad y segmentación de imágenes .....	102

Figura 2.14 Detección de obstáculos basados y mapas de profundidad...	103
Figura 2.15 Dead Reckoning Compass .....	105
Figura 2.16 Versión experimental del sistema GPS .....	106
Figura 2.17 Sistema de obtención de HRTFs .....	109
Figura 2.18 Sonido utilizado en el estrategia .....	115
Figura 2.19 Representación acústica de vías libres .....	116
Figura 3.1 Integración física del prototipo definitivo.....	122
Figura 3.2 Unión de M1 y M2.....	123
Figura 3.3 Flujo de comunicación entre módulos .....	125
Figura 4.1 Ejercicio con una columna estática.....	131
Figura 4.2 Ejercicio con dos columnas estáticas .....	132
Figura 4.3 Ejercicio con dos columnas juntas .....	133
Figura 4.4 Ejercicio con superficie estática.....	134
Figura 4.5 Ejercicio con hueco .....	135
Figura 4.6 Ejercicio con una superficie detrás de una columna .....	136
Figura 4.7 Ejercicio en movimiento .....	137
Figura 4.8 Laberinto experimental .....	139
Figura 4.9 Test de movilidad.....	140
Figura 4.10 Resultados BPL de tres pruebas consecutivas .....	146

Figura 4.11 Resultados totales por ejercicio .....	147
Figura 4.12 Comparativa entre tiempos RWT y PWT .....	149
Figura 4.13 Colisiones no deseadas .....	151
Figura 4.14 Número de correcciones .....	152
Figura 4.15 PWT y RWT para escenario MOB-A .....	154
Figura 4.16 Colisiones no deseadas en la prueba MOB-A .....	156
Figura 4.17 Correcciones en el test MOB-A .....	157
Figura 4.18 PWT y RWT para escenario MOB-B.....	158
Figura 4.19 Colisiones no deseadas en la prueba MOB-B .....	159
Figura 4.20 Comparativa del índice NUH_Ind entre MOB-A y MOB-B .....	160
Figura 4.21 Correcciones en el test MOB-B.....	162
Figura 4.22 Índices de corrección entre MOB-A y MOB-B .....	163

# INDICE DE TABLAS

Tabla 1.1 Características que afectan a la movilidad .....	32
Tabla 1.2 Clasificación de dispositivos .....	36
Tabla 2.1 Características de los sonidos .....	113
Tabla 2.2 Correspondencia de puntos .....	114
Tabla 4.1 Perfil de los participantes .....	144



# Dedicatorias

Esta tesis y todo el esfuerzo realizado para llevar a cabo este trabajo, se lo dedico de forma muy especial a mi hija Ana.

Valencia, abril de 2012



# Agradecimientos

En primer lugar quiero agradecer a mi director de tesis y amigo personal, Guillermo Peris Fajarnés, el entusiasmo y la insistencia que ha puesto para que se realizara esta tesis. Juntos hemos conseguido en estos últimos años muy buenos resultados, y ha sido un placer compartirlos con él, tanto desde el punto de vista profesional como humano.

Por otra parte, a mi codirectora Larisa Dunai, compañera en nuestras investigaciones, agradecer la ayuda que me ha prestado para poder encauzar este trabajo desde el principio apoyándome e interesándose en los resultados que iba alcanzando día a día.

Además, agradecer especialmente por su interés, tanto en lo referente a la calidad de la tesis, como en la exigencia y profesionalidad mostrada, a mi amigo Eduardo Lluna, persona clave para la realización de la tesis.

Por último, quiero mencionar también a mi compañero del día a día Ismael Lengua, por sus observaciones realizadas respecto a cuestiones puntuales, para mejorar el aspecto y seriedad de la tesis.

En general, doy las gracias a todos los que se han interesado de alguna forma por mí durante la realización de este trabajo.

Valencia, abril de 2012



# Resumen

La presente tesis refleja el trabajo de investigación y de integración llevado a cabo durante los últimos años con el propósito de aplicar los avances tecnológicos al mundo de las personas invidentes. A principio de la década del 2000, se llevaron a cabo numerosas reuniones y encuentros entre diferentes científicos para compartir opiniones acerca de las ayudas existentes para invidentes. Las que había, estaban muy enfocadas hacia personas de baja visión, pero no había casi nada para personas con ceguera total. Por tanto este fue el punto de partida para comenzar una larga trayectoria hasta llegar a conseguir los resultados reflejados en esta tesis.

A lo largo de ella, se detalla el estado del arte y los módulos y desarrollos empleados para la fabricación y validación de un prototipo compuesto por tecnología láser, estereovisión y audio, con el fin de ofrecer una herramienta más y complementaria a las ayudas que actualmente tienen estas personas: el bastón y el perro guía.

Además, tras la investigación y desarrollo, se presentan una serie de experimentos con el fin de comprobar la aceptación y las mejoras que se han conseguido en ayudas para la navegación a personas ciegas.

Por tanto, este trabajo muestra todo el proceso de investigación, científico, tecnológico, de integración y de ingeniería para, además de abrir nuevos caminos a futuras investigaciones, poder concluir diciendo que uniendo tecnología de visión y de audio se puede llegar a conseguir que las personas invidentes puedan detectar obstáculos, sentir distancias y caminar por áreas libres y mejorar así su percepción espacial, además de conseguir una mayor integración en la sociedad.

**Palabras clave:** ciego, invidente, laser, visión, estereovisión, audio, casb lip, mapas de profundidad, integración, navegación.



# Resum

La present tesis reflexa el treball d'investigacio i d'integracio portat a veta durant els ultims anys en el proposit d'aplicar els avanços tecnològics al mon de les persones cegues. A principi de la decada del 2000, se portaren a veta numerosa reunions i trobaments entre diferents científics per a compartir opinions acosta de les ajudes existents per a cegues. Les que havia, estaven molt enfocades cap a persones de baixa visio, pero no havia casi res per a persones en orbera total. Per tant este fon el punt de partida per a escomençar una llarga trayectoria fins aplegar a conseguir els resultats reflectits en esta tesis.

A lo llarc d'ella, se detalla l'estat de l'art i els moduls i desenrolls empleats per a la fabricacio i validacio d'un prototip compost per tecnologia laser, visió artificial, d'àudio en el fi d'oferir una ferramenta mes i complementaria a les ajudes que actualment tenen estes persones: el bastó i el gos guia.

Ademes, despres de l'investigacio i desenroll, se presenten una serie d'experiments en el fi de comprovar l'acceptacio i les millors que s'han conseguit en ajudes per a la navegacio a persones cegues.

Per tant, este treball mostra tot el proces d'investigacio, científic, tecnològic, d'integracio i d'ingenieria per a, ademes d'obrir nous camins a futures investigacions, poder concloure dient que unint tecnologia de visio i d'àudio se pot aplegar a conseguir que les persones puguen detectar obstacles, sentir distancies i caminar per arees lliures i millorar aixina eguesla seua percepcio espacial, ademes de conseguir una major integracio en la societat.

**Paraules clau:** cego, invident, làser, visió, audio, casblip, mapes de fondaria, integració, navegació.



# Abstract

This doctoral thesis writes about the researching and integration carried out over the recent years in order to apply technological advances to the world of the blind people. In the early 2000s, different meetings were held between different scientists to share views on the existing aid for the blind people. Many of them were very focused on people with low vision and partially blind people, but there was almost nothing for people with total blindness. So it was the starting point to begin a long way to be able to get the results reflected in this thesis.

Through this thesis, I detail the state of the art and all developments and modules used in the manufacture and validation of a prototype consisting of lasers, stereovision and audio to provide a complementary tool these kinds of people: cane and guide dog.

Moreover, after research and development, I present a series of experiments to test the acceptance and all improvements we have been made to help blind people to move and navigate around the space.

Therefore, this work shows the whole process of research, scientific, technological, integration and engineering, as well as opening new hopes for future research to conclude that, combining vision technology and audio it is possible achieve the blind people can detect obstacles, and walking distances, feel free areas and improve their spatial awareness, while they achieving greater integration in society.

**Keywords:** blind, sightless, blind people, laser, stereovision, audio, casblip, depth map, integration, navigation.



# 1. INTRODUCCIÓN

Como inicio a esta tesis, este capítulo presenta aspectos generales relativos a las personas invidentes y su contexto. En primer lugar, se habla de la necesidad de investigar y dedicar esfuerzo para mejorar su calidad de vida, incluyendo cifras y datos sobre la cantidad de personas invidentes existentes y como puede la tecnología ayudar a mejorar la calidad de vida de estas personas.

A continuación, se abordan las motivaciones personales para iniciar el proyecto y cuáles son las características que pueden mejorar la calidad de vida de las personas ciegas y las que pueden influir en su movilidad.

Vistas las generalidades y la motivación de este estudio, se analiza el estado de la técnica, incluyendo cuales son los sistemas tecnológicos más importantes ya desarrollados que han intentado mejorar la calidad de vida de las personas ciegas. Se añaden datos referentes a las patentes que han surgido como consecuencia de estas investigaciones y los artículos de investigación que se han publicado con las experiencias adquiridas y dar la correspondiente difusión científica.

Se describe de forma detallada el estado de arte de distintos dispositivos de navegación, comerciales o en fase de prototipo desde los años 50. Se presta especial atención en la adaptabilidad de los dispositivos en la comunidad de personas con discapacidad visual, demostrando los avances más significativos en el desarrollo y diseminación de los dispositivos electrónicos, mediante una evaluación, tanto objetiva como subjetiva, con medidas directas e indirectas en los programas de entrenamiento.

Respecto a las necesidades y requerimientos de las personas invidentes, se ha analizado qué información del entorno tiene más importancia para ser tratada, procesada e interpretada a la hora de transferirla al usuario final. Al mismo tiempo, se comentan algunas de las técnicas más pioneras y avanzadas desde el punto de vista tecnológico, para sustituir a la visión humana e interpretar el espacio de la misma forma, sencilla y clara, de cómo lo hacen los ojos de personas videntes, con el único fin de integrar a las personas con discapacidad en nuestro entorno social.

Los dispositivos de ayuda a la movilidad se han clasificado en tres grupos básicos: detectores de obstáculos, sensores del entorno y sistemas de navegación. Las diferencias entre estas categorías y los dispositivos más relevantes de cada grupo también son comentados en el capítulo.

Las características de los distintos dispositivos se han obtenido de las diferentes publicaciones de congresos y revistas indexadas incluidas en la sección de referencia.

Además, se enumera un listado de las patentes más importantes que tienen que ver con sistemas de ayuda a la movilidad y navegación de personas con discapacidad visual, dando una breve descripción de en qué consiste cada una de ellas con la información extraída de la propia base de datos de la patente y con las palabras exactas de su definición por parte de los inventores.

Por último, se presentan los objetivos de esta tesis, haciendo una introducción al proyecto CASBLiP, sus componentes, su desarrollo y la integración final. En este apartado, se indican cuáles son las hipótesis de partida para elaborar esta tesis, es decir, cual es el punto de partida y QUE se puede conseguir en el contexto y entorno de las personas invidentes. A continuación, se describe COMO se ha resuelto el problema y como se ha llegado a un sistema capaz de satisfacer las hipótesis de partida. Como en cualquier trabajo de investigación, una vez visto el “qué” y el “cómo”, la tesis dará paso a la fase experimental, en al cual se describe la metodología seguida para demostrar que el “COMO” funciona sobre los usuarios finales, es decir, se describirán las pruebas a efectuar con personas invidentes para llegar a la conclusión de que esta metodología y estos experimentos, pueden ser aplicados a cualquier sistema tecnológico futuro, sea cual sea sus entradas y sus salidas, además de presentar y demostrar que las pruebas efectuadas y los resultados obtenidos demuestran que lo llevado a cabo en este trabajo funciona, cubriendo las expectativas de los experimentos.

Por tanto, la base de esta tesis doctoral va a ser el integrar en un sistema para la navegación de personas invidentes las últimas tecnologías, y elaborar unas metodologías y experimentos para ser aplicados en futuros sistemas, con el fin de estudiar su viabilidad comprobar si el resultado mejora o no en algo las prestaciones de otros sistemas predecesores, y la calidad de las personas con discapacidad visual.

## **1.1 Generalidades**

Las cifras de personas totalmente ciegas son muy dispares, y las distintas fuentes y organismos públicos que informan acerca de la cantidad de personas con discapacidad visual en el mundo, no disponen de una base de datos exacta y precisa. Uno de los datos más fiables que conocemos es el de Alemania, en la que se estima un total de 150.000 ciegos totales y 500.000 personas con discapacidad visual [DBSV]. En base a estas cifras, sobre un total de 82.000.000 millones de personas se corresponden con un 0,18% y 0,60% de su población, respectivamente.

Según la Organización Mundial de la Salud (OMS), en el informe número 282 publicado en Abril 2011 [OMS Nota descriptiva], en el mundo hay aproximadamente 284 millones de personas con discapacidad visual, de las cuales 39 millones son ciegas, 245 millones presentan baja visión y 2 millones de ellas viven entre Europa y EEUU. Además:

- Aproximadamente un 90% de la carga mundial de discapacidad visual se concentra en los países en desarrollo.
- En términos mundiales, los errores de refracción no corregidos constituyen la causa más importante de discapacidad visual, pero en los países de ingresos medios y bajos las cataratas siguen siendo la principal causa de ceguera.
- El número de personas con discapacidad visual atribuible a enfermedades infecciosas ha disminuido considerablemente en los últimos 20 años.
- El 80% del total mundial de casos de discapacidad visual se pueden evitar o curar.
- El número de personas ciegas se duplicará para el año 2020, entre otros factores motivados por la mayor longevidad de la población.
- Solamente en EEUU el crecimiento de la población mayor de 65 años en los últimos años ha pasado de 13% a 19%, cuando en China fue de 7% a 13%.

Respecto al país más grande del mundo, China, también existen datos oficiales. La legislación China tiene registrado que un 5% de la población representada por 60 millones de personas, tienen alguna discapacidad o problema de independencia [ChinaDis]. De todas ellas, hay un total de 5 millones de ciegos, incluidos de baja visión y, según la administración “China

Association of the Blind” cada año va aumentando en 450.000 personas. Si continua así hasta 2020, se llegará fácilmente a los 20 millones de ciegos en el país.

Con arreglo a la Clasificación Internacional de Enfermedades [CIE-10], en su actualización y revisión de 2006, la función visual se subdivide en cuatro niveles:

- Visión normal
- Discapacidad visual moderada
- Discapacidad visual grave
- Ceguera

La discapacidad visual moderada y la discapacidad visual grave se reagrupan comúnmente bajo el término “baja visión”; la baja visión y la ceguera representan conjuntamente el total de casos de discapacidad visual.

La distribución mundial de las principales causas de discapacidad visual es como sigue:

- Errores de refracción (miopía, hipermetropía o astigmatismo) no corregidos: 43%
- Cataratas: 33%
- Glaucoma: 2%

#### Personas de 50 años o mayores

Alrededor de un 65% de las personas con discapacidad visual son mayores de 50 años, si bien este grupo de edad apenas representa un 20% de la población mundial. Con el creciente envejecimiento de la población en muchos países, irá aumentando también el número de personas que corren el riesgo de padecer discapacidades visuales asociadas a la edad.

#### Niños menores de 15 años

Se estima que el número de niños con discapacidad visual asciende a 19 millones, de los cuales 12 millones la padecen debido a errores de refracción, fácilmente diagnosticables y corregibles. Unos 1,4 millones de menores de 15 años sufren ceguera irreversible.

### Evolución en los últimos 20 años

En términos generales, las tasas mundiales de discapacidad visual han disminuido desde comienzos de los años noventa, y ello pese al envejecimiento de la población en el mundo entero. Esa disminución es principalmente el reflejo de una actuación concertada de salud pública destinada a reducir las discapacidades visuales causadas por enfermedades infecciosas.

Otra fuente de datos sobre la cantidad de personas ciegas en el mundo es el de la World Blind Union (WBU 2009), según la cual existen 314 millones de personas en el mundo con discapacidad visual, y 45 millones de ellos son invidentes [WBU 2009].

Las principales causas de la ceguera son las cataratas con un 39% que aparece con la edad, el glaucoma con un 2%, el errores de refracción así como miopía, hipermetropía o astigmatismo que ocupa un 43% es decir unos 153 millones de personas, también influye la degeneración macular relacionada con la edad, el diabetes no controlado, las afecciones oculares infantiles, y la carencia de la vitamina A. Un factor muy importante es el tracoma que afecta 40 millones de personas, que proviene de enfermedades infecciosas. En lo que se trata de los niños, las causas de ceguera son las cataratas, la retinopatía del prematuro y la avitaminosis A.

Según los diversos estudios realizados por los organismos, organizaciones o fundaciones para las personas ciegas o con discapacidad visual, los objetivos más importantes en la vida cotidiana son:

- Tener un vida independiente y productiva
- Comunicar
- Crear una familia
- Tener una vida social
- Movilidad
- Estudiar
- Etc.

Para conseguir estos objetivos necesitan mucho esfuerzo y ayuda tanto psicológica como física. El problema más importante es la movilidad, que limita a la comunidad de personas invidentes a conseguir los demás objetivos y necesidades. Ellos tienen dificultades en desplazarse de un lugar al otro porque no tienen una orientación en el espacio, y tampoco tienen información sobre la posición de los objetos estáticos o móviles del entorno. Aunque tengan el bastón o un perro lazarillo, la información percibida se encuentra a una distancia de metro y medio como mucho.

Además de la ceguera, existen muchos tipos de problemas visuales. Las más destacadas son: la visión borrosa, puntos ciegos, halos, moscas volantes y otros síntomas. Los citados síntomas pueden llevar a la ceguera.

La visión borrosa representa la pérdida de la agudeza visual y la incapacidad para visualizar pequeños detalles.

Los puntos ciegos se caracterizan por los agujeros oscuros en el campo de visión, que impide ver algo en este campo.

Las moscas volantes representan pequeñas estrellitas que flotan a través del ojo. Este síntoma puede llevar al desprendimiento de la retina que crea unos destellos de luz o una sensación de sombra en una de las partes del campo visual.

Cualquier sea la causa de pérdida de visión, lleva a una serie de incomodidades en la vida social de las personas. En estos casos, por la falta de la información visual, aparecen problemas de lectura, escritura, reconocimiento de personas o lugares y movilidad.

Algunos datos de interés, sólo en España, según el Observatorio Estatal de la Discapacidad [OED] son:

*“La evolución de los contratos a personas con discapacidad en el primer trimestre de 2011 es positiva respecto al mismo periodo del año 2010 y en estos tres meses se han registrado mejores cifras que este mismo periodo de 2010, según el último informe realizado por el Observatorio Estatal de la Discapacidad sobre la materia. Durante el primer trimestre de 2011 se han realizado 15.166 contratos a personas con discapacidad en nuestro país frente a los 13.971 contratos de este mismo periodo de 2010”.*

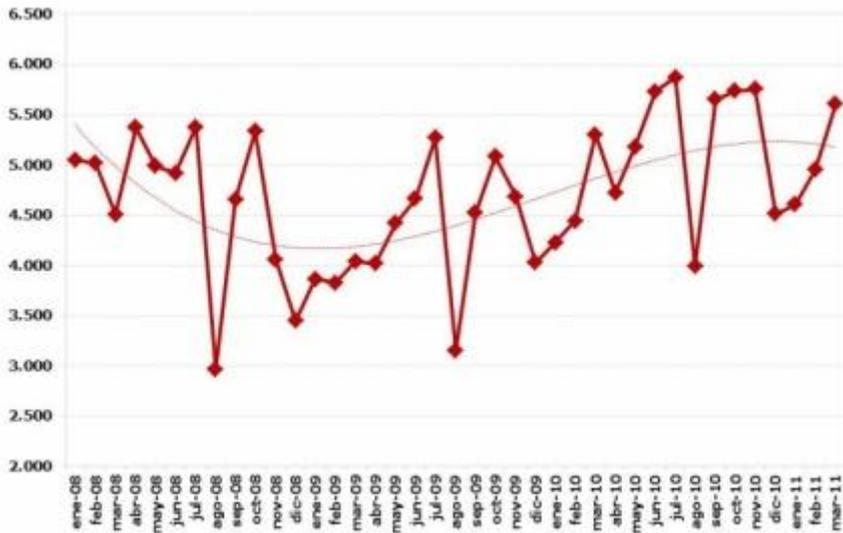


Figura1.1 Informe sobre la contratación a personas con discapacidad en España durante el primer trimestre de 2011 [OED Inf]

<b>POBLACIÓN TOTAL, ESTIMACIÓN DEL NÚMERO DE PERSONAS CON DISCAPACIDAD Y DE LA PREVALENCIA DE LAS SITUACIONES DE DISCAPACIDAD POR GRANDES GRUPOS DE EDAD. ESPAÑA, 1986, 1999 y 2008.</b> (Datos referidos a personas residentes en domicilios particulares)			
<i>Intervalos de edad</i>	<i>Población total</i>	<i>Población con discapacidad</i>	
		<i>Número</i>	<i>Porcentaje sobre población total</i>
<b>EDDM 1986</b>			
<b>TOTAL</b>	<b>38.341.126</b>	<b>5.743.291</b>	<b>14,98</b>
Menores de 6 años	2.338.048	43.047	1,84
De 6 a 64 años	30.856.653	2.804.837	9,09
65 años y más	5.146.425	2.895.407	56,26
<b>EDDES 1999</b>			
<b>TOTAL</b>	<b>39.247.019</b>	<b>3.528.221</b>	<b>8,99</b>
Menores de 6 años	2.209.504	49.577	2,24
De 6 a 64 años	30.602.991	1.405.992	4,59
65 años y más	6.434.524	2.072.652	32,21
<b>EDAD 2008</b>			
<b>TOTAL</b>	<b>45.031.743</b>	<b>3.847.854</b>	<b>8,54</b>
Menores de 6 años	2.809.038	60.408	2,15
De 6 a 64 años	34.863.979	1.560.361	4,48
65 años y más	7.358.726	2.227.085	30,26

Figura1.2 Población invidente total y por edades en España [OED]

POBLACIÓN CON DISCAPACIDADES SEGÚN LAS DEFICIENCIAS QUE LAS HAN CAUSADO, POR GRANDES GRUPOS DE EDAD Y SEXO (Datos referidos a personas de 6 y más años residentes en domicilios particulares). ESPAÑA, 2008						
	De 6 a 64 años		De 65 años y más		Total	
	Número de personas	Tasa por 1.000 hab.	Número de personas	Tasa por 1.000 hab.	Número de personas	Tasa por 1.000 hab.
<b>TOTAL</b>						
Deficiencias mentales	324.161	9,3	400.622	54,4	724.783	16,1
Deficiencias visuales	245.738	7,0	551.870	75,0	797.608	17,7
Deficiencias del oído	266.069	7,6	641.721	87,2	907.790	20,2
Deficiencias del lenguaje, habla y voz	35.202	1,0	51.234	7,0	86.436	1,9
Deficiencias osteoarticulares	515.935	14,8	970.554	131,9	1.486.489	33,0
Deficiencias del sistema nervioso	235.872	6,8	256.257	34,8	492.129	10,9
Deficiencias viscerales	166.783	4,8	409.777	55,7	576.560	12,8
Otras deficiencias	32.853	0,9	289.732	39,4	322.585	7,2
No consta	104.343	3,0	144.585	19,6	248.928	5,5
<b>Total personas con discapacidad (*)</b>	<b>1.560.361</b>	<b>44,8</b>	<b>2.227.086</b>	<b>302,6</b>	<b>3.787.447</b>	<b>84,1</b>
<b>VARONES</b>						
Deficiencias mentales	193.228	11,0	114.908	36,5	310.136	13,9
Deficiencias visuales	115.147	6,5	184.999	58,8	300.146	13,5
Deficiencias del oído	136.376	7,7	248.945	79,2	385.322	17,3
Deficiencias del lenguaje, habla y voz	25.699	1,5	24.538	7,8	50.237	2,3
Deficiencias osteoarticulares	191.514	10,8	225.204	71,6	416.718	18,7
Deficiencias del sistema nervioso	102.835	5,8	103.221	32,8	206.056	9,3
Deficiencias viscerales	80.730	4,6	151.414	48,2	232.144	10,4
Otras deficiencias	11.451	0,6	78.284	24,9	89.735	4,0
No consta	59.691	3,4	56.896	18,1	116.587	5,2
<b>Total varones con discapacidad (*)</b>	<b>754.498</b>	<b>42,7</b>	<b>756.757</b>	<b>240,7</b>	<b>1.511.255</b>	<b>67,9</b>
<b>MUJERES</b>						
Deficiencias mentales	128.932	7,5	285.714	67,8	414.647	18,2
Deficiencias visuales	130.592	7,6	366.871	87,0	497.462	21,8
Deficiencias del oído	129.692	7,5	392.776	93,2	522.468	22,9
Deficiencias del lenguaje, habla y voz	9.503	0,6	26.696	6,3	36.199	1,6
Deficiencias osteoarticulares	324.421	18,9	745.349	176,9	1.069.770	47,0
Deficiencias del sistema nervioso	133.038	7,7	153.036	36,3	286.074	12,6
Deficiencias viscerales	86.053	5,0	258.363	61,3	344.416	15,1
Otras deficiencias	21.402	1,2	211.448	50,2	232.850	10,2
No consta	44.652	2,6	87.689	20,8	132.341	5,8
<b>Total mujeres con discapacidad (*)</b>	<b>805.863</b>	<b>46,9</b>	<b>1.470.329</b>	<b>348,9</b>	<b>2.276.192</b>	<b>100,0</b>

Figura1.3 Población invidente total por deficiencias causadas en España [OED]

## **1.2 Motivación**

Tras algunos años trabajando en el desarrollo de proyectos de ingeniería en distintos sectores (automoción, aviación, siderúrgico, etc.), es decir, trabajando con máquinas, se me brindó la posibilidad de introducirme en el ámbito social y poder aplicar la tecnología al servicio de las personas para mejorar su bien estar e integración social. Por ello arranca la ilusión de utilizar todo el conocimiento adquirido en el desarrollo e investigación de herramientas para las personas invidentes.

Técnicamente también existen numerosos aspectos por los que me he motivado a seguir adelante con este trabajo. Principalmente, pienso que la tecnología ha de servir como una herramienta para conseguir diferentes objetivos. Estos objetivos pueden ser, desde cuestiones relativas a sectores industriales, hasta diferentes fines sociales relacionados con la integración y calidad de vida de las personas discapacitadas. Por ello pienso que, a pesar de todos los avances de los que hoy en día disponemos, aún existe muchísimo camino por delante para explotar y utilizar todos los recursos que los avances tecnológicos nos ofrecen.

La ceguera consiste en la pérdida de la percepción visual debido a factores neurológicos o psicológicos. La pérdida de visión viene a menudo acompañada de la pérdida de independencia. Las personas con discapacidad visual, y en particular las personas invidentes, son incapaces de aprovechar las ventajas de distintos servicios, así como interacción social, comunicación y movilidad.

Existen importantes necesidades que las personas invidentes requieren:

- a. **Vida independiente y productiva**  
Para todo ser humano, la independencia es una necesidad vital para la calidad de vida. Las personas invidentes no son menos, por lo que ayudar a que cualquier persona o ser humano a que pueda valerse por sí misma es un avance importante.
- b. **Comunicación**  
La comunicación es un factor importante en la vida humana, así como leer, escribir y hablar hace que la gente sea más comunicativa y amistosa. La comunicación también ayuda a intercambiar conocimientos, explicar situaciones y sentimientos de felicidad y tristeza. Por otra parte, con la

comunicación ayuda a las personas a enriquecerse intelectual y espiritualmente.

Varias técnicas de lectura y escritura han sido desarrolladas para la comunidad de personas invidentes, así como Braille, libros que transmiten voz, lectores que convierten el texto impreso en voz, etc. También se han aplicado tecnologías informáticas tales como software, hardware, móviles escáneres y aplicaciones en general de reconocimiento óptico y lectores de pantalla.

c. Crear una familia

La gente necesita tener descendencia, convivir con personas queridas y poder sacarla adelante. Por ello, una persona invidente que pueda apoyarse en la tecnología para llevar a cabo esta tarea le va a aportar felicidad y bienestar.

d. Tener vida social

Esta facultad puede ir unida al resto de las anteriores, es decir, la comunicación, la independencia y la relación con el resto de personas forman un entorno social importante y, sobre todo, ayuda al ser humano a desarrollarse y convivir con respeto y aceptación de los demás.

e. Movilidad

Una de las necesidades más importantes de las personas invidentes es la falta de movilidad, la cual se convierte en una limitación importante para su estilo de vida: no pueden viajar solos debido a que no pueden determinar su posición y la localización de los objetos en el entorno, ya que se necesita bastante información.

f. Poder ejercer estudios superiores para igualarse al resto de personas

g. Ocio y hobbies

La falta de comunicación y movilidad impide a las personas invidentes disfrutar de los deportes y vivir momentos de diversión y agradables.

Por tanto, las personas invidentes siempre necesitan tener a alguien a su lado para acompañarles en todas sus necesidades.

La World Blind Union tiene como objetivo ayudar a las personas invidentes o con discapacidad visual a aprender a utilizar varias técnicas y métodos de

lectura, escritura y navegación. A su vez, les enseña cómo mejorar algunas habilidades físicas que les ayudará a percibir y orientarse en el entorno. De este modo, aprovechando los sonidos, el ruido, la percepción de sensaciones y la temperatura les integrarán de mejor forma en el contexto social.

La mayoría de las personas ciegas, aprenden a compensar su falta de visión con el desarrollo del aparato auditivo. La información del entorno permite a los humanos y animales aprender sobre las fuentes sonoras situadas en diferentes posiciones y, en particular, aquellas que están fuera del radio de acción visual [Fay R. R. and Popper A. N., 2005]. La localización de sonidos es importante para la supervivencia y seguridad en la navegación. Las personas invidentes utilizan mucho el sonido, no solamente para detectar la presencia de los obstáculos, si no para determinar su posición y su nivel de peligrosidad.

La tabla 1.1 representa las características más importantes que afectan a la movilidad.

Características Sicológicas	Expectativas individuales
	Forma de vida
	Motivación
Características Sensoriales y Motoras	Capacidad visual
	Capacidad auditiva
	Capacidad motora
Características perceptuales cognitivas	Interpretación del eco
	Conciencia espacial
	Modo de proceso y Capacidad cognitiva
Características del entorno	Urbanos o rurales
	Clima
	Tiempo de luz al día
	Entrenamiento y capacidad de aprendizaje
	Entorno familiar

**Tabla 1.1 Características que afectan a la movilidad**

Dado el conocimiento de las limitaciones que existen en el entorno vital de las personas invidentes, como se explica anteriormente y entendiendo la problemática desde el punto de vista tecnológico y científico, debido a mis conocimientos y trayectoria profesional, se ha pensado unir todos los esfuerzos para el desarrollo de un equipo electrónico que pueda mejorar la calidad de vida de este colectivo.

Por ello, la presente tesis presenta el diseño, desarrollo, implementación y las posteriores pruebas experimentales de un sistema electrónico de detección de obstáculos y navegación utilizando una mezcla de tecnologías de visión y audio.

## **1.3 Estado del arte**

### **1.3.1 Sistemas Tecnológicos Desarrollados**

Foulke [Foulke E., 1971] definió la movilidad como “la habilidad de viajar con seguridad, confortablemente y de forma independiente por el entorno”. Existen tres formas de conseguir movilidad. La primera de ellas es utilizando un guía humano. La segunda y más extendida es con la ayuda del bastón, y la tercera con ayuda de perros lazarillos.

Los perros lazarillos son capaces de detectar y analizar las situaciones complejas como intersecciones y cruces, escalones y peligros potenciales.

Durante varios siglos, el bastón ha sido el más importante y popular sistema de movilidad en personas invidentes. Sin embargo su importancia en la comunidad de las personas invidentes antes de 1964, cuando Russell C. Williams publicó las “Especificaciones del bastón (Typhlocane)”, que ayudaron a establecer el modelo estándar del bastón [Farmer W., 1978]. Las personas usaban cualquier palo o bastón para moverse por el espacio. Mediante el bastón, los usuarios perciben el entorno más cercano, alrededor de 1 metro. El bastón representa el sentido táctil del usuario, como una extensión de la mano [WBU, 2009]. En 1971, durante la conferencia National Research Council en Washington, se aprobaron las características físicas y funcionales del bastón [NRC, 1972]. Estas características, se refieren a la barra, la empuñadura y la punta:

1. El eje vertical de la barra ha de ser recto y sin oscilaciones
2. La barra debe estar dividida en segmentos desde la empuñadura hasta la punta.
3. Debe estar disponible en diferentes longitudes para ajustarse a la altura del usuario.
4. Debe de tener suficiente longitud para proporcionar al usuario la información importante, con suficiente antelación para poder reaccionar, pero no darle al usuario libertad física excesiva.
5. Su peso ha de ser ligero sin afectar al equilibrio del resto de requerimientos.
6. Debe tener una pequeña resistencia al viento.
7. Debe ser lo suficientemente rígido para permitir al usuario establecer la distancia y posición de los objetos detectados, de manera exacta y sin modificar su forma original.
8. No debe permitir la conductividad térmica y eléctrica.
9. Debe ser capaz de transmitir vibraciones desde la punta hasta la empuñadura para proporcionar estímulos táctiles y aurales (auditivos).
10. El bastón ha de ser rígido y visible para peatones y conductores.
11. Debe ser capaz de producir un mínimo nivel de ruido cuando se utiliza sin dispositivos amplificadores.
12. Deber ser resistente.
13. Debe ser remplazado fácilmente sin herramientas especiales.
14. Debe tener un buen equilibrio.
15. Debe tener buen aspecto.

A pesar de que el bastón está considerado como el dispositivo de movilidad por excelencia, tiene sus limitaciones. Entre ellas, el bastón no proporciona suficiente información cuando el objeto cercano es peligroso o antes de su alcance. La longitud del bastón limita el rango de percepción del entorno, perdiendo una cantidad sustancial de información.

Después de la segunda guerra mundial, con el desarrollo del sensor, se emplearon muchos esfuerzos para diseñar y desarrollar sistemas electrónicos capaces de percibir el entorno circundante. En 1897, Starkiewicz y Kuliszewsky desarrollaron el primer Dispositivo Electrónico de Ayuda a la Movilidad (DEAM), denominado “Neiszewski’s Electroftalm”, y en 1912 Nye y Bliss desarrollaron el “D’Albe’s Exploring Optophone”. En 1970, Nye y Bliss publican el mejor estado de arte sobre los DEAM [Nye P. and BlissJ, 1970]. Farmer [Farmer L. W., 1975] menciona que los DEAM deben, por ellos mismos, con la ayuda del bastón o del perro lazarillo, informar al usuario

invidente sobre los objetos del camino desde el suelo hasta la posición de la cabeza, así como advertir de cualquier superficie discontinua que pueda provocar una amenaza para la seguridad de la persona.

El Electroftalm utiliza una sola celda de selenio colocada en la parte frontal con el objetivo de controlar la intensidad del sonido de salida [Starkiewicz W. and Kuliszewsky T., 1963]. Con esta tecnología, el usuario invidente era capaz de distinguir entre luz y oscuridad. El Exploring Optophone, desarrollado por D'Albe's, tiene un procedimiento de funcionamiento similar al Electroftalm, que convierte la luz en sonidos [Capp y Picton, 2000], [Meijer P., 1992], en el que se utiliza un detector de selenio y un reproductor de sonidos para emitir un sonido cuando disminuye la intensidad de luz. Este dispositivo fue presentado en 1912 en la Convención Óptica en el Museo de Ciencia de South Kensington. DEAM permite a los usuarios invidentes evitar los obstáculos en el área de visión del dispositivo, e informan de la presencia de los obstáculos lejanos. Los DEAM determinan la dirección, distancia, altura y el ancho de los objetos, además de otros objetos dentro de la escena y otras particularidades de los objetos. DEAM representa una gran esperanza para la comunidad invidente permitiéndoles moverse con más independencia y seguridad. Viajando con más independencia y seguridad ayuda a los usuarios invidentes a moverse con menos estrés y tensión. Algunos autores [Farmer L. W., 1975] dicen que los DEAM son dispositivos complementarios al bastón y perro lazarillo que ayudan a las personas invidentes a moverse por el entorno.

Se han propuesto diferentes teoremas sobre DEAM: Benham [Benham T.A., 1954] y Benjamin [Benjamin, 1968] sugieren que los DEAM deben detectar obstáculos e indicar su localización y distancia aproximada, tanto obstáculos grandes como pequeños, ligeros o pesados, escalones o pasillos libres. Además, el dispositivo como información de entrada adicional, debe dar al usuario ciego información sobre orientación y navegación adicional. Esta información no debe interferir con los ruidos del entorno y no transmitir datos erróneos [Dupress J. K., 1963] y debe ser sincronizada con otros parámetros. La información proporcionada ha de ser clara y simple, además de rápidamente interpretada sin un entrenamiento excesivo. Para Farmer [Farmer L. W., 1975], un DEAM debe ser un dispositivo que contribuya a la independencia del usuario para su movilidad y no agresivo.

Además de los requerimientos y especificaciones mencionadas arriba, un DEAM debe estar diseñado para tener el mínimo número posible de accesorios (cajas, electrónica, cascos, conexiones, cables, etc.) con el objetivo de no molestar al usuario. Es muy importante la técnica usada para representar la información del entorno.

Desde el punto de vista técnico, los dispositivos deben ser diseñados y desarrollados, no solo para ser ligeros y pequeños sino para ser fiables, duraderos y con una estética agradable.

Técnicamente, un DEAM se basa en tres interfaces: interfaz de entrada, interfaz de proceso e interfaz de salida. La interfaz de entrada es la encargada de adquirir la información del entorno, que puede ser clasificada en: ultrasonidos, láser, visión artificial y GPS. El módulo de proceso contiene las técnicas hardware y software para procesar toda la información adquirida y transformarla en datos necesarios para la interfaz de salida (sonidos, táctil, habla, vibraciones, etc.). El interfaz de salida representa la forma de transmitir la información hacia el usuario. Esta información ha de ser lo más concisa y precisa posible, para no confundir y molestar al usuario final. La tabla 1.3 resume los DEAM más importantes así como las especificaciones básicas. Las características de estos dispositivos se comentan a continuación.

### CLASIFICACIÓN DE DISPOSITIVOS SEGÚN SISTEMA DE ENTRADA

Nº	DEAM	Especificaciones
<b><u>ULTRASONIDOS</u></b>		
1.	Nottingham Detector	<ul style="list-style-type: none"> <li>➤ Dispositivo detector de obstáculos</li> <li>➤ Dispositivo de mano</li> <li>➤ Emite pulsos sonidos de alta frecuencia (40 kHz) similar a la escala musical mayor</li> <li>➤ 8 salidas</li> <li>➤ Rango de percepción máximo de 7 pies.</li> </ul>

2.	Sonic Torch	<ul style="list-style-type: none"> <li>➤ Dispositivo de mano con haz luminoso estrecho</li> <li>➤ Señal de salida audio</li> <li>➤ La amplitud de la señal corresponde a la distancia</li> <li>➤ El timbre corresponde a la variación de la textura del objeto, con señales de ultrasonidos moduladas a frecuencia de banda ancha (40-80 kHz)</li> <li>➤ Ángulo de visión es cónico de 30 ° a cada lado de la línea media de visión</li> </ul>
3.	Lindsay Russell Pathsounder	<ul style="list-style-type: none"> <li>➤ Dispositivo detector de obstáculos</li> <li>➤ Rango de detección 6 pies</li> <li>➤ Salida táctil y auditiva</li> <li>➤ Es una caja colgada al cuello</li> </ul>
4.	Mowat Sensor	<ul style="list-style-type: none"> <li>➤ Dispositivo de mano</li> <li>➤ Salida táctil y sonar</li> <li>➤ Emite señales al espacio en forma cónica y elíptica con 15 ° de ancho y 30 ° de alto</li> <li>➤ Vibra con frecuencia inversa a la distancia</li> <li>➤ Rango de trabajo en distancia: de 1 m a 4 m</li> </ul>
5.	Sonic Pathfinder	<ul style="list-style-type: none"> <li>➤ Dispositivo montado en la cabeza</li> <li>➤ Salida con señales de audio</li> </ul>
6.	Sonic Guide	<ul style="list-style-type: none"> <li>➤ System binaural (stereo)</li> <li>➤ Dispositivo montado en la cabeza</li> <li>➤ Rango de trabajo 6 m</li> <li>➤ La salida es acústica, táctil y con vibraciones</li> </ul>
7.	Walkmate	<ul style="list-style-type: none"> <li>➤ Dispositivo detector de objetos</li> <li>➤ Salida mediante tonos acústica</li> </ul>

8.	Polaron	<ul style="list-style-type: none"> <li>➤ Dispositivo detector de obstáculos o indicador de pasillos libres</li> <li>➤ Salida vibro táctil y acústica</li> <li>➤ Rango de trabajo 16 pies</li> </ul>
9.	Sensory 6	<ul style="list-style-type: none"> <li>➤ Dispositivo detector de obstáculos</li> <li>➤ Salida tonal</li> </ul>
10.	Ultracane (Batcane)	<ul style="list-style-type: none"> <li>➤ Dispositivo detector de obstáculos</li> <li>➤ Sensor de entorno</li> <li>➤ Detecta textura de objetos y modelos</li> </ul>
11.	Light Probes	<ul style="list-style-type: none"> <li>➤ Fococélulas conectadas a circuito electrónico</li> <li>➤ Detecta fuentes de luz</li> <li>➤ Salida acústica, táctil y vibro táctil</li> </ul>
12.	Palmsonar	<ul style="list-style-type: none"> <li>➤ Fococélulas conectadas a circuito electrónico</li> <li>➤ Detecta fuentes de luz</li> <li>➤ Salida acústica, táctil y vibro táctil</li> </ul>

<b><u>LASER</u></b>		
13.	Laser cane	<ul style="list-style-type: none"> <li>➤ Laser de galio y Arsenio</li> <li>➤ Salida señales acústicas</li> <li>➤ Rango de detección 4 m</li> </ul>
14.	Pilot Light	<ul style="list-style-type: none"> <li>➤ Mini-radar</li> <li>➤ Sistema infrarrojos</li> <li>➤ Dispositivo manual</li> <li>➤ Rango de detección 3,5m</li> <li>➤ Salida mediante habla</li> </ul>
15.	Talking light	<ul style="list-style-type: none"> <li>➤ Comunicación wireless e infrarrojos</li> <li>➤ Dispositivo manual</li> <li>➤ Sensor de entorno</li> <li>➤ Salida mediante voz</li> </ul>
16.	Talking signs	<ul style="list-style-type: none"> <li>➤ Comunicación wireless e infrarrojo</li> <li>➤ Sistema RIAS</li> <li>➤ Dispositivo de mano</li> <li>➤ Luz infrarroja</li> <li>➤ Sensores externos infrarrojos</li> </ul>
17.	SONA [Kelly G. W., 1981]	<ul style="list-style-type: none"> <li>➤ Sistema inteligente de entorno</li> <li>➤ Señal acústica mediante altavoz</li> </ul>
18.	MARCO	<ul style="list-style-type: none"> <li>➤ Transmisión y recepción de infrarrojos</li> <li>➤ Dispositivo manual</li> <li>➤ Rango de detección 12,5m</li> <li>➤ Salida mediante voz humana</li> </ul>

19.	Verbal Landmarks	<ul style="list-style-type: none"><li>➤ Dispositivo de orientación y navegación</li><li>➤ Señales de radio de bucle inductivo</li><li>➤ Rango de detección 2m</li><li>➤ Omnidireccional</li><li>➤ Salida mediante voz humana</li></ul>
20.	Easy Walker	<ul style="list-style-type: none"><li>➤ Sistema de orientación y navegación</li><li>➤ Entrada láser infrarrojos</li></ul>
21.	CASBlIP M1	<ul style="list-style-type: none"><li>➤ Dispositivo de detección de objetos</li><li>➤ Sensor 3D CMOS</li><li>➤ Rango de detección 5m</li><li>➤ Salida acústica</li></ul>

<b><u>GPS</u></b>		
22.	MoBIC	<ul style="list-style-type: none"> <li>➤ Sistema de orientación y navegación</li> <li>➤ Entrada mediante teclado</li> <li>➤ Salida mediante señales acústicas</li> </ul>
23.	Makino	<ul style="list-style-type: none"> <li>➤ Sistema de orientación y navegación</li> <li>➤ Teléfono móvil digital</li> <li>➤ Conexión con Unidad Central</li> <li>➤ Salida mediante voz sintetizada</li> </ul>
24.	Electronic Guide Dog	<ul style="list-style-type: none"> <li>➤ Sistema de orientación y navegación</li> <li>➤ Teléfono móvil y unidad central de datos</li> <li>➤ Guía mediante el habla directa de persona</li> <li>➤ Dispositivo manual</li> </ul>
25.	PNG	<ul style="list-style-type: none"> <li>➤ Sistema de orientación y navegación</li> <li>➤ Mapas digitales</li> <li>➤ Conexión SIG</li> <li>➤ Salida vibro táctil y voz sintetizada</li> </ul>
26.	GPS Braille Note	<ul style="list-style-type: none"> <li>➤ GPS, mapas y puntos de interés</li> <li>➤ Salida mediante voz sintetizada y braille</li> </ul>
27.	Trekker [Lagace Y., 2005]	<ul style="list-style-type: none"> <li>➤ GPS y mapas digitales</li> <li>➤ Salida vocal</li> </ul>
28.	TORMES [CIDAT]	<ul style="list-style-type: none"> <li>➤ Dispositivo de mano</li> <li>➤ Tecnología EGNOS, GPS y Braille</li> <li>➤ Salida mediante voz sintetizada</li> <li>➤ Rango de detección 2m</li> </ul>

29.	Easy Walk	<ul style="list-style-type: none"> <li>➤ GPS y teléfono móvil con Symbian</li> <li>➤ Receptor GPS y Bluetooth</li> <li>➤ Conexión con call center 24 horas</li> <li>➤ Contacto mediante SMS o llamada</li> <li>➤ Salida mediante voz sintetizada</li> </ul>
30.	Geotact	<ul style="list-style-type: none"> <li>➤ Sistema de orientación y navegación</li> <li>➤ GPS y sensor inercial (posición del usuario)</li> <li>➤ Salida mediante voz sintetizada</li> </ul>
31.	Talk & Wayfinder	<ul style="list-style-type: none"> <li>➤ Sistema de orientación y navegación</li> <li>➤ GPS</li> <li>➤ Salida mediante voz</li> </ul>
32.	Loadstone GPS [Loadstone, 2008]	<ul style="list-style-type: none"> <li>➤ Teléfono móvil con sistema operativo Symbian, Bluetooth GPS, USB y auriculares</li> <li>➤ Salida mediante voz</li> <li>➤ Informa sobre puntos de interés</li> </ul>
<b><u>VISIÓN ARTIFICIAL</u></b>		
33.	vOICe	<ul style="list-style-type: none"> <li>➤ Sistema de orientación y navegación</li> <li>➤ Cámaras de visión para entrada al sistema</li> <li>➤ Salida mediante señales acústicas y táctiles</li> </ul>
34.	CASBlIP M2	<ul style="list-style-type: none"> <li>➤ Sistema de orientación y navegación</li> <li>➤ Entrada con dos cámaras, sensor inercial</li> <li>➤ Detección de obstáculos estáticos y en movimiento y pasillos libres</li> <li>➤ Salida mediante señales acústicas</li> </ul>

35.	EPEL	<ul style="list-style-type: none"> <li>➤ Sistema sonar estereoscópico</li> <li>➤ Entrada con sensores sonar</li> <li>➤ Salida vibro táctil</li> </ul>
<b><u>DISPOSITIVOS MIXTOS</u></b>		
36.	CASBlIP	<ul style="list-style-type: none"> <li>➤ Sistema de detección de objetos, orientación y navegación</li> <li>➤ Sensor 3D CMOS, cámaras estéreo, sensor inercial y PC</li> <li>➤ Detección de objetos estáticos y en Movimiento y pasillos libres</li> <li>➤ Rango de detección 15 metros</li> <li>➤ Salida acústica</li> </ul>
37.	SWAN	<ul style="list-style-type: none"> <li>➤ Sistema de navegación</li> <li>➤ GPS, 4 cámaras de video, PC táctil</li> <li>➤ Salida de audio</li> </ul>
38.	TANIA	<ul style="list-style-type: none"> <li>➤ Sistema de navegación</li> <li>➤ Dispositivo de mano</li> <li>➤ Entrada mediante mapas, sensores de entorno, PC</li> <li>➤ Informa verbalmente y de forma táctil sobre distancias y objetos</li> </ul>
39.	Tyflos	<ul style="list-style-type: none"> <li>➤ Dispositivo de navegación orientación y lectura</li> <li>➤ Entrada con 2 cámaras, sensor de rango, GPS y lector de radiofrecuencia</li> <li>➤ Salida mediante vibraciones y voz sintetizada</li> </ul>
40.	Eye PlusPlus (FSRS)	<ul style="list-style-type: none"> <li>➤ Sistema de navegación</li> <li>➤ Entrada con videocámaras</li> <li>➤ Salida con cinturón vibro táctil</li> </ul>

**Tabla 1.2 Clasificación de dispositivos**

La mayoría de los dispositivos de movilidad existentes, están basados en tecnología de ultrasonidos y óptica. Actualmente, existen más de quince dispositivos de ultrasonidos diferentes.

Desde el comienzo de los estudios, los dispositivos de movilidad electrónicos se han clasificado en tres grupos:

- 1. Detectores de obstáculos:** Dispositivos que solo detectan objetos muy cercanos (1-2 m)
- 2. Dispositivos de sensores de entorno:** Dispositivos que detectan objetos a distancias mayores de 4 m y la textura de la superficie.
- 3. Dispositivos de orientación y navegación:** Sistemas compuestos por electrónica de usuario y dispositivos que se colocan en distintos puntos del entorno o GPS.

- 1. Detectores de obstáculos:** informan solamente de la presencia de los obstáculos muy cercanos. Estos dispositivos se consideran como dispositivos go-no-go y proporcionan información, pero limitada.

Dispositivos detectores de obstáculos ultrasónicos:

El dispositivo más conocido de esta categoría es Lindsay Russell Pathsounder [Russell, 1965], el cual es considerado el primer dispositivo de movilidad basado en ultrasonidos [Mann R.W., 1970]. Este dispositivo envía impulsos ultrasónicos de 30° a una distancia máxima de seis pies. El dispositivo informa de la presencia de los objetos mediante interfaces de audio y táctil. La distancia es dividida en dos zonas: la zona de para protección y la zona de no protección. Para personas con discapacidad auditiva se les implementó la salida vibratoria [Russell, 1965]. Durante los últimos años, el dispositivo ha estado sujeto a múltiples modificaciones, desde el modelo E hasta el dispositivo actual [Farmer W., 1978].

El modelo E o Pathsounder Special E (Figura 1.4), está formado por dos componentes: la parte de percepción del entorno y la caja de control y procesamiento. Las desventajas del dispositivo son:

- a) No es resistente al agua, es decir, no funciona en caso de lluvia.
- b) Las baterías se agotan rápidamente a bajas temperaturas.

Las ventajas serían:

- a) Se ha probado con niños y está considerado como el dispositivo electrónico de movilidad mejor valorado para el entrenamiento en orientación y movilidad.
- b) Garantiza seguridad en su manejo, sobre todo en niños pequeños.

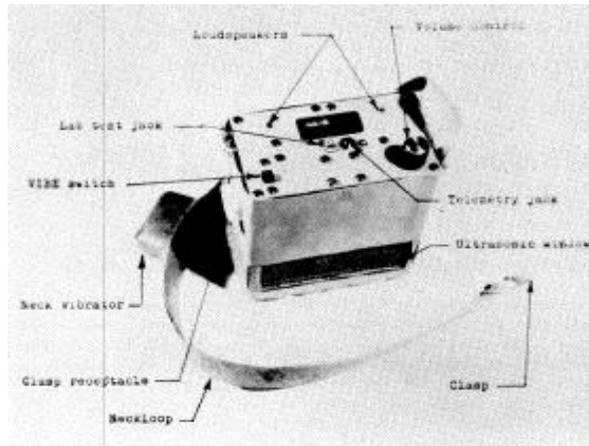


Figura 1.4 Lindsay Russell Pathsounder

Otro dispositivo de movilidad electrónica es el *Mowat Sonar Sensor* (Figura 1.5), que ha sido desarrollado por G.C. Mowat en Nueva Zelanda. El dispositivo es manual con salida sonora y táctil. Fue considerado uno de los más simples y más reconocidos durante los años 70 [Morrissette D. L. et al, 1971]. El dispositivo emite un ultrasonido de forma cónica y elíptica de 15° de ancho y 30° de altura, cubriendo un área equivalente a la forma humana.

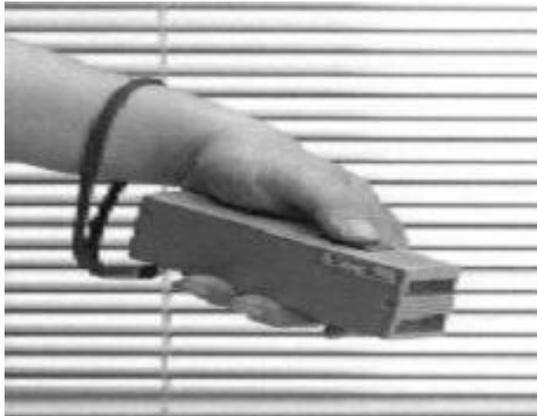


Figura 1.5 Mowat Sonar Sensor

El Mowat Sonar Sensor detecta solamente los obstáculos más cercanos. El rango de distancia es de 1 a 4 metros. Al detectar un obstáculo, el sensor vibra con una frecuencia inversa a la distancia desde el objeto al dispositivo. Tiene un conmutador en la parte superior con tres posiciones. Estas posiciones permiten al usuario elegir uno de los dos niveles, distancias cortas o distancias largas. La distancia más larga de 4 m puede ser seleccionada desplazando el conmutador hacia delante desde la posición central. La distancia corta de 1 m, se seleccionaba desplazando el conmutador hacia atrás.

Al igual que Lindsay Russell Pathsounder, el Mowat Sonar Sensor fue diseñado para ser utilizado como dispositivo adicional complementando el bastón o los métodos tradicionales, tanto para entornos cerrados como abiertos, interpretar señales en estaciones de autobús, detectar objetos en movimiento, etc. Debido a su pequeño tamaño, el sensor se podía llevar en el bolsillo sin dificultad [Freiberger H. A. M., 1974], [Book, 1986].

Otro dispositivo es el *Nottigham Obstacle Detector (NOD)* (Figura 1.6). Se trata de un dispositivo de ultrasonidos manual, parecido al Mowat Sonar Sensor con la excepción de que el dispositivo emite pulsos sonoros de alta frecuencia de 40 kHz [Farmer W., 1978]. Los pulsos sonoros representan las 8 notas de la escala musical [Freiberger H. A. M., 1974].

Este dispositivo tiene 8 salidas que representan los datos de entrada. La distancia total está dividida en 7 distancias menores, donde cada nota representa una distancia. Cuando no hay objetos en el área de visión del

sensor, el dispositivo no emitirá ninguna señal sonora. Si se detecta más de un objeto en el área de visión, el dispositivo sonificará el objeto más cercano. La distancia máxima de detección es de 2,1 m dividida en pequeñas distancias de 30 cm.



Figura 1.6 Nottingham Obstacle Detector

El NOD ha sido el dispositivo más utilizado en localización de obstáculos en situaciones difíciles. El dispositivo tiene un conmutador de on/off y un pequeño altavoz para representar la señal. Este dispositivo fue probado con salida táctil [Amstrong, 1976].

Dispositivos detectores de obstáculos con entrada láser:

***Laser Cane*** (Light Amplification by Stimulated emission of Radiation) (Figura 1.7), representa la combinación del bastón con la tecnología láser [Farmer W., 1978], [Ando 2003]. En 1963 se desarrolló el primer Laser Cane llamado C-3 el cual fue evaluado y validado por 50 personas invidentes. El dispositivo fue miniaturizado modificando el tamaño de sus circuitos integrados, los láseres de galio arsénico y utilizando pequeñas baterías. Tras los resultados obtenidos con las pruebas de las personas invidentes en la versión C-3, en 1971 se desarrolla la versión C-4. Esta versión fue probada por 8 ciegos veteranos de guerra que tenían al menos un año de experiencia en movilidad. El

entrenamiento duró cinco semanas, y tras dicho entrenamiento cada uno de las personas involucradas en las pruebas se llevó un ejemplar a sus casas para seguir testeando y validando el sistema. Las conclusiones obtenidas tras estas pruebas desembocaron en el desarrollo de la siguiente versión C-5, más ligera y que detectaba obstáculos en un rango de 4 metros. El sistema escaneaba el entorno informando al usuario sobre los obstáculos más cercanos mediante señales acústicas, como se muestra en la figura 1.4.

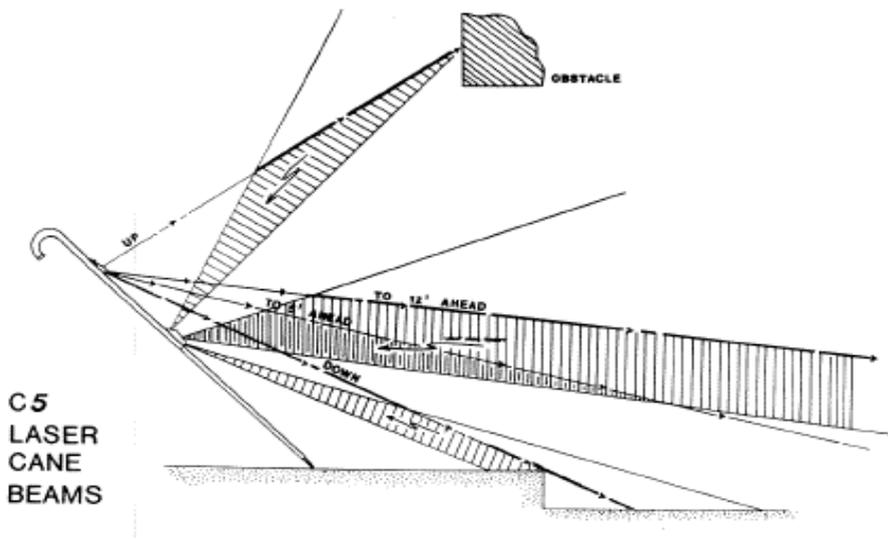


Figura 1.7 Principio de funcionamiento del Laser cane

Este sistema fue desarrollado por J. Malvern Benjamín y sus colegas de Bionic Instrument Inc. La versión C.5 está formado por tres láseres de inyección de GaAs (Arsénico de Galio) que emiten pulsos de 9050 Angstrom de  $0,2 \mu s$ , entre 40 y 80 veces por segundo, y tres receptores fotosensibles, permitiendo detectar los obstáculos con una precisión alta en altura.

Su funcionamiento se basa en el principio Granberg, que consiste en triangularización óptica aplicada a la detección de obstáculos. El láser emite un impulso de luz, cuando rebota con un obstáculo y se refleja, siendo detectado por las lentes de los foto receptores.

2. **Dispositivos de sensores de entorno:** se trata de dispositivos detectores de obstáculos a largas distancias. Estos dispositivos, además de detectar obstáculos muestran información sobre la textura de los objetos y sus patrones [Book, 1986], [Ando, 2007]. El dispositivo más conocido de este tipo es el *Kay Sonic Torch* desarrollado en 1959 (Figura 1.8).



Figura 1.8 Sonic Torch

El dispositivo Sonic Torch es un dispositivo manual basado en ultrasonidos de banda estrecha que proporciona señales auditivas a través de auriculares, donde el tono corresponde al rango de detección y el timbre a la variación de la textura de la superficie de los objetos. Las señales de ultrasonido son de banda ancha en frecuencia modulada entre 40 y 80 kHz.

El funcionamiento de este dispositivo se basa en escanear el entorno mediante un cono de 30° a cada lado de la línea central. Sin embargo, tras sus pruebas el sistema no fue lanzado al mercado por ser un dispositivo que requería el ser transportado manualmente, además del bastón.

Con el fin de mejorar el dispositivo Sonic Torch, y evitar que se tratara de un dispositivo manual, en 1966, se desarrolló el *Binaural Sonar Electronic Travel Aid*, conocido como *Sonicguide*. Este dispositivo está implementado en unas gafas, y se le añadió un segundo canal acústico, creando la sensación de espacio a través de señales binaurales. El sistema fue diseñado para proporcionar más información al usuario sobre el entorno a una distancia de cuatro metros mediante sonidos estéreo, tacto y vibro tacto [Kuc R., 2002].

Además, el usuario es capaz de reconocer distancias según la variación del tono [Farmer L. and Smith D., 1997].

El Sonicguide informa al usuario de la distancia, de la dirección horizontal (azimut) y la identificación del objeto mediante características del tono. El dispositivo emite una señal sónica de forma cónica de 45° en referencia a la izquierda y la derecha de la línea central de visión del usuario. La desventaja es que el sistema no proporciona información sobre objetos bajos y escalones.

La segunda versión del Sonicguide fue el *Sonicguide MkII* (Figura 1.9 y 1.10), el cual tenía una precisión en distancia de 6 m y, dependiendo de la superficie y textura de los objetos, podía variar de 3,6 a 4,5 m para superficies difusas.



Figura 1.9 Sonicguide MkII

**Dispositivos de orientación y navegación:** Estos dispositivos tienen dos características que los diferencian de los anteriores: familiarización con el entorno y navegación. La familiarización se consigue mediante sensores instalados en diferentes puntos: centros comerciales, estaciones, pasos de peatones, etc., por donde las personas invidentes van a navegar [Brabyn J., 1982] y la navegación mediante GPS [Book, 1986].

Como ejemplo de dispositivos de familiarización con el entorno, podemos nombrar a *Talking Signs System* (Figura 1.11), que representa un dispositivo de comunicación inalámbrico mediante infrarrojos basado en la tecnología Remote Infrared Audible Signage (RIAS), que proporciona mensajes de voz humana a distancia informando al usuario acerca de su posición en el entorno [Brabyn J., 1982], [Brabyn J., 1997]. Fue una tecnología pionera desarrollada por Smith-Kettlewell Eye Research Institute, Rehabilitation Engineering Research Center de San Francisco, California en 1979. Cada transmisor instalado en el entorno envía constantemente señales codificadas sobre la su identidad. Al mismo tiempo, el usuario cuando orienta su receptor hacia el transmisor se establece una comunicación vocal describiendo las características de la entidad representada por el transmisor, guiando así al usuario hacia el lugar donde desea ir.

El rango de distancia es, para outdoor (aire libre) 30 m y poco menos para indoor (interiores).



Figura 1.10 Sonicguide MkII

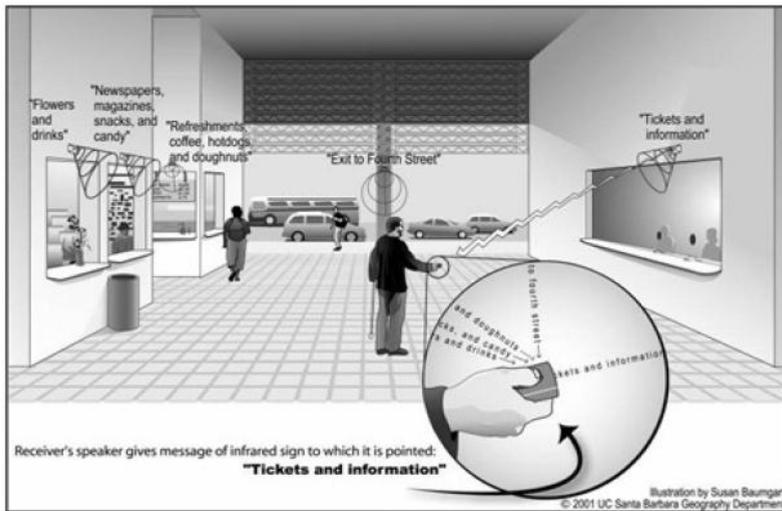


Figura 1.11 Funcionamiento del Talking Signs System

Los dispositivos de navegación son dispositivos basados en GPS, donde los satélites, o entidades tipo centro de datos (Call Center), envía señales al receptor del usuario [Farmer L. and Smith D., 1997], [Collins C.C., 1985], [Loomis, 1985].

Los sistemas de posicionamiento global (GPS) son herramientas para identificar cada uno de los puntos de la superficie de la tierra utilizando coordenadas de latitud y longitud. Esta información puede ser proporcionada en tiempo real dando al usuario su posición exacta e indicándole parámetros acerca de su contexto, tales como calles, números, edificios, servicios, velocidad, intersecciones y direcciones. El dispositivo informa al usuario de forma verbal asegurando la exactitud en un rango de 20 m.

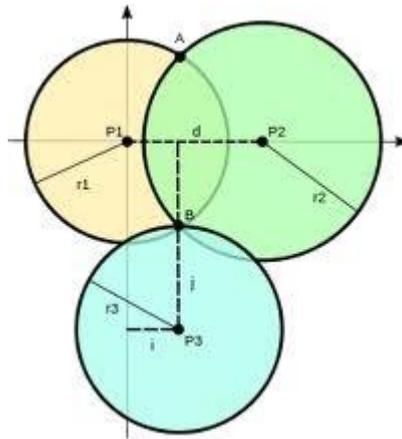


Figura 1.12 Funcionamiento básico del GPS

Estos sistemas se basan en el principio matemático de la triangularización (ver Figura 1.12), a través del cual podemos saber el punto donde nos encontramos situados. Cuando el receptor de GPS detecta una señal de radiofrecuencia procedente de un satélite, se genera una esfera virtual que envuelve al satélite, y es el propio satélite el que actúa como centro de esfera, y su superficie se extenderá hasta el punto que posee la antena receptora, por tanto, el radio de la esfera será la distancia que separa el satélite del receptor. Mientras tanto, el receptor hará el mismo proceso, es decir, calculará las distancias que les separa con dos satélites más como mínimo (de ahí el proceso de triangularización), calculando el tiempo de demora de cada señal para posteriores cálculos matemáticos.

Las ondas de radio por parte de los satélites se transmiten a la velocidad de la luz (300 mil Km/s) medida en el vacío, por lo que podemos calcular la distancia entre un transmisor y un receptor si se conoce el tiempo que tarda la señal en viajar de un punto a otro. Para ello, los relojes del emisor y transmisor han de estar perfectamente sincronizados. Esto se consigue porque el satélite emite cada cierto tiempo señales de control a los receptores, la cual llega al receptor de GPS con más retraso que la señal propia de radiofrecuencia, y el retraso entre ambas señales indican el tiempo que tarda la señal de radio en ir desde el satélite al receptor GPS.

Las distancias entre los satélites y los receptores son calculadas por los propios receptores multiplicando el tiempo de retraso de la señal de control

por la velocidad de la luz. Si la señal viaja en línea recta, el resultado será la distancia entre el receptor y el satélite.

Entre los dispositivos de navegación, podemos encontrar el **MoBIC** [Petrie H. et al, 1997], [Strothotte T. et al, 1997], **Makino** [Makino H. et al., 1996], **Electronic Guide Dog** [ELD, 1995], **GPS Braille Note** [Loomis, 1985], [HumanWare], etc.

**MoBIC** está formado por dos sistemas: el MoBIC Pre-Journey System (MoPS) el cual se encarga de planificar actividades, y el MoBIC Outdoor System (MoODS) que tiene como objetivo interpretar planos [Strothotte T. et al, 1997], [Douglas G. et al., 1997]. Su funcionamiento consiste en que el usuario mediante teclado introduce una pregunta al sistema y recibe la respuesta mediante sonidos [Douglas G. et al., 1997].

**Personal Guidance System (PGS)** se basa en el sistema de navegación GPS [Loomis J. M. et al, 2001]. Uno de los objetivos del dispositivo fue conseguir que fuera portable y que ayudara a personas invidentes a navegar tanto en entornos conocidos como en entornos nuevos y desconocidos sin ayuda de un guía acompañante [Loomis J.M. et al, 2006]. El dispositivo se basa en Fluxgate Compass (electromagnético), mapas digitales [Loomis, 1985], un GPS y un ordenador [Loomis J.M. et al, 2006], [Loomis J. M. et al, 1998], según figuras 1.13 y 1.14.

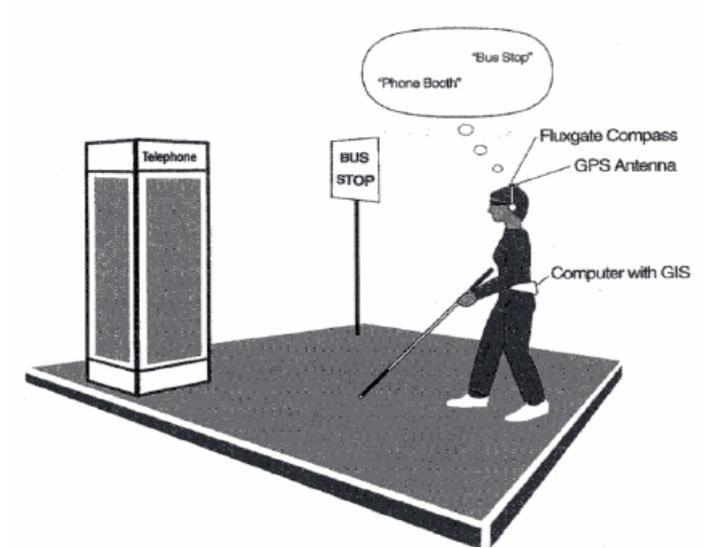


Figura 1.13 Principio de funcionamiento del PGS [Loomis, 1985]

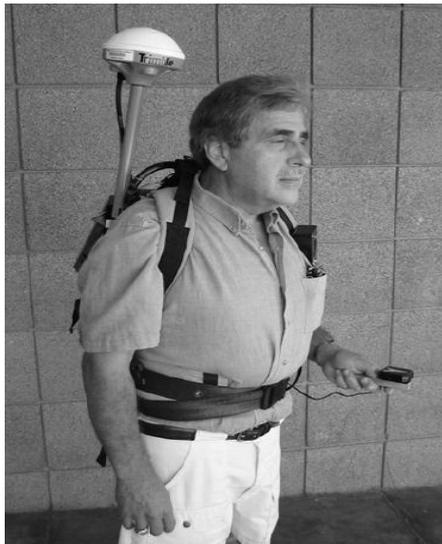


Figura 1.14 Última versión del PGS

El sistema GPS determina la posición del usuario en el entorno, y se comunica con el Sistema de Información Geográfica (SIG), un ordenador y unos auriculares para informar al usuario con voz sintética [Loomis J.M. et al, 2005], o mediante sensaciones vibro táctiles [Klatzky R. L. et al, 2006] da información sobre su posición.

La idea original del PGS era que el usuario, mientras navegaba por la calle pudiera oír nombres de calles y números de edificios con la sensación de que le llegaran del propio objeto y de forma espacializada. La figura 1.13, representa el principio de funcionamiento de la primera versión del PGS. Fue diseñado para trabajar en distancias largas teniendo en cuenta que las distancias cortas pudieran ser detectadas con el bastón. Como interfaces de entrada el dispositivo aceptaba teclados Qwerty tradicionales y teclados Braille.

El rango de trabajo del dispositivo es de 20 m y como importante aportación el usuario puede aprender y recabar información sobre puntos de interés tales como estaciones, restaurantes, etc. contribuyendo a la construcción de una representación espacial cognitiva sobre el entorno.

Los sistemas de información geográfica (SIG) o Geographic Information System (GIS) están formados por un conjunto de elementos hardware, software y bases de datos, interactuando entre ellos, para poder interpretar información geográfica del terreno y analizarla, manipularla y almacenarla, con el fin de poder resolver problemas de gestión en general, es decir, reflejan modelos de la realidad en base a coordenadas terrestres.

El objetivo más importante de un sistema de información geográfica es el mostrar al usuario (figura 1.15 y Figura 1.16), mediante orto fotografías o técnicas similares, información geográfica del terreno. Una vez el usuario tiene en su poder toda esta información, puede interactuar con ella y poder realizar consultas con el fin de explotarla con muchos y muy diferentes objetivos: riegos, tratamientos, características del terreno, parámetros del suelo y presentar resultados geográficos en general.



Figura 1.15 Imagen SIG del municipio 900 parcela 6 según SIGPAC

GOBIERNO DE ESPAÑA MINISTERIO DE MEDIO AMBIENTE Y MEDIO RURAL Y MARINO SUBSECRETARÍA FONDO ESPAÑOL DE GARANTÍA AGRARIA (FEGA) SIGPAC sistema de identificación de parcelas agrícolas

Fecha de vuelo: 06/2008 Año de renovación catastral: 1994

Provincia	Municipio	Agregado	Zona	Polígono	Parcela	Superficie (Ha)
46 - VALENCIA	900 - VALENCIA	0	0	6	6	0,2235

Recinto	Superficie (Ha)	Pendientes (%)	Uso	Coef. Pastoreo	Coef. Regadío	Incidencias
1	0,2013	5,2	TA		100	12
2	0,0221	5,2	CA			

Figura 1.16 Sistema de identificación de parcelas agrícolas mediante SIGPAC

**Navigation Assistance for Visually Impaired (NAVI)** Se desarrolló en la Universidad de Malasia bajo la dirección del profesor Sainarayanan y sus colegas [Sainarayanan G. et al, 2007]. El dispositivo está compuesto por una cámara de video, unos auriculares estéreos, un SBPS (Single Board Processing System) y una chaqueta con baterías recargables (figura 1.17).

El sistema detecta los obstáculos en el área de visión con una resolución de 32x32 píxeles de resolución y en escala de grises. El SBPS utiliza la red

neuronal Fuzzy LVQ para clasificar los píxeles y transformar la escala de grises en sonidos estéreo. Con el dispositivo, el usuario es capaz de detectar objetos estáticos y objetos con movimientos lentos. La desventaja del dispositivo es que no detecta la distancia de los objetos.

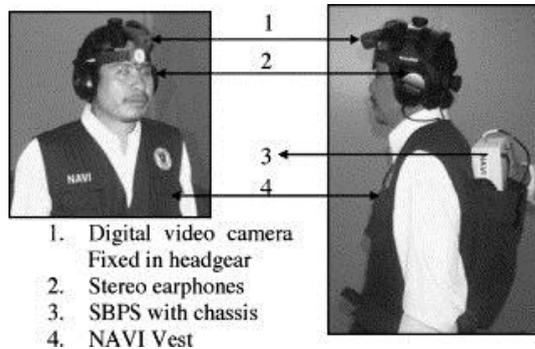


Figura 1.17 Dispositivo NAVI

Una versión más avanzada del NAVI es el KINECT, desarrollado por estudiantes de la universidad alemana de Kostanz, donde pretendieron utilizar una cámara Kinect para detectar el entorno de una persona con deficiencias visuales y navegar de forma segura en ambientes cerrados.

La cámara va ensamblada en un casco (Figura 1.18), y envía los datos capturados del mundo de alrededor del usuario a un cinturón, provocando sensaciones vibro táctiles basándose en un diseño del contexto predefinido.



Figura 1.18 Dispositivo NAVI - KINECT

***SWAN (System for Wearable Audio Navigation)*** Se trata de un dispositivo que consiste en un interface de entrada táctil y salida audio que ayuda a la detección de objetos y a la navegación (Figura 1.19).

El sistema utiliza el SIG con el objetivo de espacializar la información [Wilson J. et al, 2007], incluyendo un GPS, un sensor inercial, un podómetro, tarjetas RFID, una brújula y sensores de detección.

Finalmente, para guiar a la persona invidente a través del trayecto a seguir e indicar la localización de los objetos más importantes del entorno, es sistema emite sonidos acústicos mediante auriculares estéreo. En resumen, los objetivos del sistema son:

- Emitir sonidos de navegación para guiar a los usuarios por el camino predeterminado desde el punto de partida a través de distintos puntos de interés, hasta llegar al punto destino.
- Los sonidos de los objetos indican la localización y tipo de objeto que conforman el contexto del usuario, así como mobiliario, puertas, etc.
- Sonidos para el cambio de superficie de navegación así como el cambio de acera a césped, desde una alfombra a un suelo de azulejos, de pasillo a escaleras interiores, bordillos, etc.
- Sonidos de localización para identificar oficinas, aulas, tiendas, edificios, estaciones, etc.
- Sonidos aclarativos formados por mensajes vocales muy cortos grabados por el usuario que proporciona detalles adicionales sobre el entorno. Ej.: ¡Aquí hay un charco profundo cuando llueve!

## SWAN System Overview

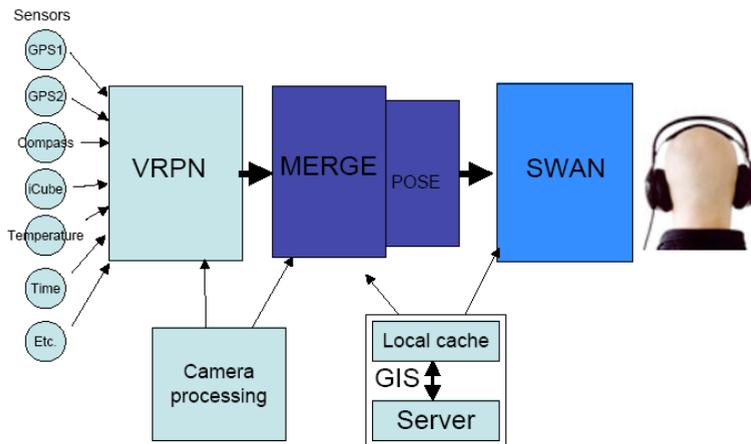


Figura 1.19 Principio funcionamiento del sistema SWAN

Los sonidos son sonidos 3D y están diseñados de manera que parece de provengan desde el entorno. El dispositivo utiliza la convolución entre el sonido y las head-related transfer function dinámicos (HRTF), funciones de transferencia las cuales proporcionan coordenadas espaciales del sonido (altura, anchura y profundidad)

*The vOICE (system Seeing with sounds)* [Meijer P., 2005] (Figura 1.20). Es un dispositivo para complementar los sistemas de navegación existentes, proporcionando información sensorial adicional mediante simulaciones acústicas. Está compuesto por una cámara, auriculares y un ordenador. La cámara escanea el campo visual, y el ordenador convierte las imágenes adquiridas por la cámara en sonidos o mapas táctiles [Meijer P., 1992]. Las escenas de enfrente de los usuarios con convertidas en imágenes estéreo, lo cual significa que los objetos serán oídos por el lado de donde procedan (izquierda o derecha). La luminosidad de la imagen capturada en blanco y negro se representa por intensidad de volumen, y la elevación por el tono del sonido.



Figura 1.20 Componentes del sistema The vOICe formado por cámaras web 3D modelo Vizix Wrap 920AR HMD

### 1.3.2 Sistemas Patentados (Dispositivos electrónicos de movilidad)

La siguiente lista resume las patentes más relevantes de los últimos años, reflejando el interés científico que ha habido para intentar conseguir sistemas comerciales que finalmente puedan llegar a las personas invidentes como usuarios finales. Cito textualmente desde la referencia del autor el texto original que describe la patente.

1)

**Alonzi, L. W., Smith, D. C., Burlak, G. J., Mirowski, M. (1992). Radio frequency message apparatus for aiding ambulatory travel of visually impaired persons. U.S. Patent #[5,144,294](#) issued Sept. 1, 1992.**

*“A radio frequency message apparatus for aiding ambulatory travel by handicapped persons such as blind individuals. The apparatus generally comprises a portable, radio frequency transceiver and a stationary, radio frequency base transceiver unit. The portable radio frequency transceiver is carried on the person of the handicapped individual and transmits a message request signal in response to manual activation of a transmit button thereon by the handicapped individual. The message request signal is received by the base transceiver, which causes the base transceiver to transmit a pre-recorded message signal back to the portable transceiver unit in radio frequency form. The message signal contains location identifying information such as the streets of an intersection at which the base unit is located. The portable transceiver has a limited transmission range of preferably about 20-50 feet to enable it to interrogate a single base transceiver unit located at an intersection within a metropolitan area or at a display/exhibit within a recreational facility such as a zoological park without accidentally interrogating base transceivers in the near vicinity of the desired base transceiver. In a preferred embodiment, an electronic compass is included within the portable transceiver to further aid a visually handicapped individual in orientating himself/herself with respect to North, South, East and Westerly directions.”*

2)

**Balbuena, A. U., Cantabrana, A. L.(1995) System to assist the guiding of the non-sighted U.S. Patent# [5,409,380](#) issued April 25, 1995**

*“The invention consists of a modular circuit installed at a physical site where it is desired to signal the current situation to a visually impaired person and which is activated by means of radiations emitted by the person carrying the circuit. It includes a fixed modular circuit, a radiation sensor, an analog/digital interface triggering device connected with a microprocessor with an input for a microphone and outputs for an amplifier connected with a loudspeaker that generates the sound messages emitted. The triggering device includes outputs connected with a control to adjust the sensitivity of the device and the amplifier includes a volume control. Also included are pulsators for the recording and reproduction of the signals introduced by the microphone/tape, with the ability to choose between two possible messages and capable of being deactivated. This system is used for the emission of an oral message upon activation of a circuit by the user.”*

3)

**Borenstein, J.(1997). User-driven active guidance system. U. S. Patent #[5,687,136](#) issued Nov. 11, 1997**

*“A user-driven, active guidance system for guiding visually impaired users through obstacle filled routes of travel. The system includes an obstacle detection system, having an array of ultrasonic sensors, which detects the distance to and location of obstacles and a controller for receiving obstacle data and determining an optimal path around the obstacle so as to return the user back to the original path of travel without losing orientation or direction. The system provides active guidance by exerting physical force upon the user to intuitively direct the user around the obstacle. The system is driven by the user's motion and comprises a cane, as well as the my of ultrasonic sensors and controller supported on a pair of guide wheels.”*

4)

**Chi-Sheng, H. (1992). Electronic talking stick for the blind. U.S. Patent #[5,097,856](#) issued Mar. 24, 1992**

*“This invention relates to an electronic talking stick for the blind and more particularly to a stick which talks to instruct a blind man to walk and go upstairs and downstairs, and to warn a blind man of dangerous depression in road, and which calls for help when a blind man who uses the stick falls. It is generally comprised of a supporting rod, a control box, a handle, a free steering caster, and a horizontal scanning device, in which the free steering caster helps the whole assembly to slide; two detectors in the control box and a range finder and moving object detector in the handle are arranged for trouble detection to let detected signals be converted into voice, by means of the processing through a control circuit, to instruct the user through an earphone connected to the handle.”*

5)

**Crandall, William (1997). Accessible automatic teller machines for sight-impaired persons and print-disabled persons U.S. Patent# [5,616,901](#) issued April 1, 1997**

*“Systems which comprise (a) an automatic teller machine which includes a plurality of customer interfaces such as a bank card reader, a banking record dispenser, a cash dispenser, and a receptacle for receiving bank deposits; (b) infrared remote communication emitters and (c) individual short range infrared communication emitters located in the teller machine. The emitters (b) are adapted to provide repeating, directionally sensitive frequency modulated message signals identifying the direction to and location of the teller machine. Thus a person having a portable receiver for such signals is led to the machine and is enabled to position himself/herself in front of the machine in order to operate it. The respective emitters of (c) provide a separate repeating, directionally sensitive frequency modulated message signal which at least identifies the location of the respective customer interfaces on the teller machine so that by movement of the portable receiver in front of the machine, the location on the teller machine of the respective customer interfaces can be determined. Feedback concerning the transactions can also be provided from the system to the customer through the portable receiver.”*

6)

**Crandall, Jr. William F. (1998). Signal transmitter with automatic output control and systems utilizing the same U.S. Patent# [5,757,530](#) issued May 26, 1998**

*“Described are apparatus for transmitting light wave signals, the apparatus being capable of adjusting their signal strength to automatically adapt to changes in the ambient light conditions of their surrounding environment. In one embodiment, the apparatus for transmitting lightwave signals comprises (a) an emitter (I) for sending a frequency modulated lightwave signal of variable intensity, (b) a light energy receiver (CDS) for detecting a change in the level of surrounding ambient light energy and for producing a response to such change in proportion to the change detected, and (c) a controller for adjusting, directly or indirectly, the frequency modulated lightwave signal in response to the light energy receiver response to the change detected. Audio transmission systems employing apparatus of this invention are also described.”*

7)

**Friedman, M. B. (1991). Computerized vocational task guidance system U.S. Patent# [5,032,083](#) issued July 16, 1991.**

*“A system for training or aiding an impaired individual to perform a series of tasks at spaced work sites comprises a mobile device for being worn or carried by the individual. The mobile device comprises circuits and devices for detecting the position of the mobile device relative to the spaced work sites. It further comprises a programmed computer having a memory for storing a data structure associating each task with the ordinal number of the task, a task description, the site at which the task is to be performed, a flag or pointer indicative of whether a task has been performed and messages containing cues and/or instructions for each task. The programmed computer also comprises a memory and processor for storing and executing a program which based upon the task flag or pointer and the location of the mobile module for cueing and/or instructing the individual to perform a task at a site.”*

8)

**Fruchterman, J. R., Schwegler, W. C., Merritt, B. W., & LaPierre, C. (1995). System and method for tracking a pedestrian. U.S. Patent #[5,470,233](#) issued Nov. 28, 1995**

*The present invention is a global positioning system that helps a blind pedestrian navigate through a city. This system uses the Department of Defense Global Positioning System (GPS) and a Differential GPS receiver to determine a pedestrian's longitude and latitude. Once these coordinates have been determined, they are correlated with a computerized map database. The map database holds the names and coordinates of specific locations, such as intersections, in a particular region. The system of the present invention retrieves the names of locations from the map database that are near the pedestrian. These names are then output to a voice synthesizer.*

9)

**Gesink, J., Guth, D., Fehr, B. (1998) Learning and assessment aid for a severely visually impaired individual U. S. Patent# [5,803,740](#) issued September 8, 1998**

*“A learning and assessment aid (10) for helping visually impaired individuals develop desirable ambulatory motion habits. The learning aid includes a housing (14) designed to be attached to an article of clothing worn by the individual. Internal to the housing is an inertial transducer (16) that monitors the rotation of the housing and the individual. The individual enters commands to the learning aid through a keypad (20) in response to audible instructions generated over a speaker (22). To foster desirable turning motions, the learning aid informs of the extent to which the individual turns. To foster straight-line walking, the learning aid generates a message over the speaker whenever the individual rotates more than a pre-set individual-entered maximum veer.”*

10)

**Hancock; Michael B. (1998). Electronic autorouting navigation system for visually impaired persons [5,806,017](#) issued September 8, 1998**

*“The present invention is an autorouting navigation system for directing one or more visually impaired persons to a physical location. The present invention includes a portable autorouting navigation unit that includes an input device for selecting a target location corresponding to a physical*

*location of a plurality of location beacons, a receiver configured to receive signals from the location beacons, a memory storing information correlating the locations of the location beacons, and an output device for communicating with the user. A microcontroller or microprocessor under program control processes the signals received from the location beacons, analyzes those signals and correlation data, and determines a routing path from the user's present position to the target location. The microcontroller may also provide an indication, such as an audible sound, to the user of whether or not the user is following the routing path to the target location. Significantly, the present invention relies upon beacons that send out location and direction information without being polled by the portable unit determines a routing path that allows for autorouting. Beacons may also be organized in groups of beacons that can correlate to geographical landmarks such as hallways in a building to reduce the size of the correlating data stored in the memory of the portable unit and thereby reduce processing time."*

**11)**

**Hashimoto, J. (1995). Identification Device for Visually Impaired Person. Japanese Patent # [JP7334076](#) issued 22 December 1995**

*"PURPOSE:To provide an identification device for a visually impaired person which is easily usable by a visually impaired person, enables the visually impaired person to easily and exactly recognize the changing points and dangerous points of a course and enables the visually impaired person to recognize places, positions and names of objects, etc. CONSTITUTION:This identification device has an information output device 1 for the visually impaired person which is installed in or on the prescribed place or object and transmits an information signal relating to the position or object based on the place or the object and a device 6 which receives the information signal transmitted from the information output device and converts the information signal to identification information to enable the visually impaired person to recognize the information signal."*

**12)**

**Jorgensen, A. A. (1990). Echo Location System for Vision-Impaired Persons. U.S. Patent #[4,907,136](#) issued Mar. 6, 1990**

*"Echo locating apparatus for a vision-impaired person which includes: a sound emitter for emitting a stream of sound bursts of ultrahigh frequency; at least on receive channel having a microphone for receiving echoes of the*

*sound bursts and generating echo signals; an echo profile detector for generating an echo profile signal of each echo signal; a delay circuit for adding a delay to the echo profile signal, consisting of the sum of a fixed delay and a variable delay that increases with the distance to the object reflecting the echo. The sound burst emitter is preferably arranged to emit a beam of sound bursts having a given beam angle that can be pointed in any direction.”*

13)

**Kelk, G. F. (1996). Mobility Aid for Blind Persons. U.S. Patent #[5,487,669](#) issued Jan. 30, 1996**

*“The present invention provides a mobility aid, for example a mobility aid handled in the manner of a flashlight, which allows a totally blind person to hand-scan, such as back and forth or up and down, the surrounding scene in order to hear the unbroken tonal phrases of varying frequencies which continuously indicate to the user the distance to, and, from the position of the hand, the position of objects and surfaces which make up the user's surroundings. A narrow laser beam is used, resulting in much higher angular resolution and distance detail, due to the much shorter wavelength, than that obtainable from practical ultrasonic devices, including those of the said prior art.”*

14)

**Massimino, M. J. & Sheridan, T. B. (1995) Apparatus for providing sensory substitution of force feedback. U.S. Patent #[5,451,924](#) issued Sept. 19, 1995**

*“A feedback apparatus for an operator to control an effector that is remote from the operator to interact with a remote environment has a local input device to be manipulated by the operator. Sensors in the effector's environment are capable of sensing the amplitude of forces arising between the effector and its environment, the direction of application of such forces, or both amplitude and direction. A feedback signal corresponding to such a component of the force, is generated and transmitted to the environment of the operator. The signal is transduced into an auditory sensory substitution signal to which the operator is sensitive. Sound production apparatus present the auditory signal to the operator. The full range of the force amplitude may be represented by a single, audio speaker. Auditory display elements may be stereo headphones or free standing audio speakers, numbering from one to*

*many more than two. The location of the application of the force may also be specified by the location of audio speakers that generate signals corresponding to specific forces. Alternatively, the location may be specified by the frequency of an audio signal, or by the apparent location of an audio signal, as simulated by a combination of signals originating at different locations.”*

**15)**

**Madey, Julius M. J.(1997). Discriminating infrared signal detector and systems utilizing the same. U.S. Patent#[5,623,358](#) issued April 22, 1997**

*“Described is a portable directional receiver. It comprises a non-visual communicator, a self-contained source of electrical energy, a detector for receiving a sensed continuous frequency modulated infrared signal, and electronics converting the sensed signal into intelligible non-visual communication emanating from the communicator both in the presence and in the absence of an ambient background of light energy, such as sunlight. The device is adapted for use by the visually impaired in finding desired locations from which suitable modulated infrared signals are continuously emitted.”*

**16)**

**McKiel, Jr., F. A. (1994) Method and system for enabling a blind computer user to locate icons in a graphical user interface U.S. Patent#[5,287,102](#) issued February 15, 1994**

*“Disclosed is a computer audio interface adapted to enable blind or visually impaired users to locate icons positioned in rows in the background of a graphical user interface. Whenever the pointer of the interface is positioned on the background, the system produces a distinctive tone. Whenever the pointer is positioned in an icon row that is occupied by one or more icons, the system produces a distinctive chord. Whenever the pointer is positioned in the background and not in an occupied row, the system produces stereo effects that give the user information as to the relative left/right position of the pointer. Whenever the pointer is positioned in an occupied row, the amplitude of the left and right stereo channels is controlled such that whenever the pointer is to the left of the leftmost icon, substantially all of the volume is produced by the left speaker and whenever the pointer is to the right of the rightmost icon, substantially all of the the volume is produced by the right speaker. Whenever the pointer is located between the leftmost and rightmost*

*icons, the volume is balanced substantially equally between the two speakers.”*

17)

**Meijer, Peter B. L. (1992). Image-audio transformation system. U. S. Patent #[5,097,326](#) issued Mar. 17, 1992**

*“In a device for converting visual images into representative sound information especially for visually handicapped persons an image processing unit is provided with a pipelined architecture with a high level of parallelism. An image is scanned in sequential vertical scanlines and the acoustical representatives of the scanlines are produced in real time. Each scanline acoustical representation is formed by sinusoidal contributions from each pixel in the scanline, the frequency of the contribution being determined by the position of the pixel in the scanline and the amplitude of the contribution being determined by the brightness of the pixel.”*

18)

**Montane, Ioan (1998). Method and device of navigating in a graphical computer interface for the blind. E.P. Patent# [EP874345](#) issued 28 October 1998.**

*“The method involves having objects, words or graphics of the graphics environment are represented by a word. Each character of the word is associated with a dynamic function defined according to the position of the character in the word. The word is formed by the designation of an object in the graphics interface to which is linked a Braille device (5). One or several supplementary characters are indicated by direct reading of the characteristics of the object. The functions are activated by keys (8) each of which corresponds to Braille cells (7) in a Braille range, which displays the characters of the word. The activation of the keys by a blind person produces the same result as a mouse operated by a sighted person. The system has a navigation keyboard that works with a voice system, each key of the keyboard corresponds to an exclusive function when one of the words uttered by the voice system has been marked by means of the keys (4). The Braille device (3) has keys (4) for Braille character sensing and has a tactile strip (5) forming a range of Braille cells (7) which can be modified for reading screen displayed data. The device (5) has also a row (6) of interactive buttons (8) each corresponding to a Braille cell (7). each interactive keys (8) situated above the Braille cells (7) corresponds to a function linked to a position in the word*

*of the character displayed in the Braille cell (7). The corresponding interactive keys (8) change function when the word is displaced over the row of Braille cells (5) so as to permit the linking of functions to characters of the words not yet displayed. This is done so that the corresponding function is always at the same position of the character in the word displayed, whatever its length in the cell range (5).”*

19)

**Ono....etal (1991) Guiding device for visually handicapped person. U. S. Patent# [5,032,836](#) issued July 16, 1991.**

*“A guide device for visually handicapped person which has a portable unit and a position indicator unit. The portable unit is carried by the visually handicapped person and includes a radio frequency signal oscillator. The position indicator unit is disposed to a desired place and includes a radio frequency signal receiver and a loud speaker.”*

20)

**Reiter, L. (1991). Electronic-automatic orientation device for walkers and the blind. U.S. Patent #[4,991,126](#) issued Feb. 5, 1991**

*“The automatic orientation device according to the invention for walkers and the blind is a small light weight and easily man-carried instrument. It contains an electronic-digital-calculating for continually determining the vectorial sum of the steps made by converting these steps into electrical pulses proportional to the strides and by automatic determination of the direction by means of a direction emitter in order to calculate the distance and direction covered from the point of departure. Whether by walking, running or horse riding the user is thus provided with optical or acoustic indicating or with perceptible vibrations indicating the distance and direction from the point of departure. In this way any person, including the blind, can find his way to the point of departure without any means of assistance, or geographical map or town plan, and in particular when the visibility is extremely poor.”*

21)

**Rohen, J. E. (1993) Virtual graphics display capable of presenting icons and windows to the blind computer user and method U.S. Patent# [5,186,629](#) issued February 16, 1993**

*“The disclosure teaches a computer mouse having tactile feedback to be used with audio computer output to provide a virtual graphic display to blind computer users. The mouse and audio feedback allow blind computer users to visualize computer graphic images and multiple screen windows in much the same way as these forms of computer output are visualized by persons with sight.”*

22)

**Silverman, H. S. (1996). Identifier/locator device for visually impaired. U.S. Patent #[5,508,699](#) issued Apr. 16, 1996**

*“The present invention is a locator device for the visually impaired including a transmitter and a dual sensing system having a receiver with earphones which are attached to opposite sides of a unit worn on the user's head. The sensing system is capable of determining a direction and distance of the signal, and capable of producing first and second audible notifiers in each of the earphones. The notifier produced in each earphone is different and is associated with the direction and distance of the transmitted signal. When the notifier produced in each earphone is the same, the user is directly in front of the desired object to be located.”*

23)

**Takeyama, M. (1997) Man-machine interface in aerospace craft that produces a localized sound in response to the direction of a target relative to the facial direction of a crew U. S. Patent# [5,647,016](#) issued July 8, 1997**

*“The direction of a target is determined relative to the facial direction of a crew of an aerospace craft and a localized sound is produced in this direction. This localized sound is then output binaurally to the crew to provide an indication of the location of the target. This use of localized sound allows the crew to rapidly locate targets while reducing eye strain. Pseudo-targets can also be located in a flight simulator.”*

24)

**Vogel, Karl-Heinz (1997). Ultrasonic Spectacles for the Blind and Visually Disabled. E.P. Patent# [EP777455](#) issued 11 June 1997**

*“Ultrasonic spectacles help the blind and visually disabled to better orient themselves. The direction, size, distance and space distribution of obstacles*

*located in the field of vision up to 6 m distance are quasidigitally displayed by vibration generators mounted on the forehead or by optoelectronic signal generators. In order to accurately locate obstacles, the ultrasonic spectacles have an ultrasonic emitter in the centre of the spectacle frame and two ultrasonic receivers that work separately, have separate evaluating logics and are arranged at the left and right in the spectacle frame. The ultrasonic spectacle is designed in such a way that the sensors and vibration generators of the electronic ultrasonic system may be built into optical spectacles in order to assist vision.”*

25)

**Yamamoto Tsuneki (1998) Walking Aid for Blind Person Japanese Patent # [JP10043263](#) issued 17 February 1998**

*“To provide an implement which can quickly detect an obstacle for a blind person to take safe measures to walk. SOLUTION: An obstacle detection mechanism comprising an ultrasonic sensor 1 to transmit/receive ultrasonic waves, a microcomputer 2 to determine of an obstacle exists or not based on received reflection waves of the ultrasonic waves, an oscillator for an alarm signal to generate a buzzer sound and/or vibration based on an obstacle detection signal of it, and a battery 4 for driving each part and an operation switch 5 for on/off switching is assembled to be an implements which can be carried when a blind person walks, and transmission modes for buzzer sound and/or vibration are changed in accordance with a distance to the obstacle.”*

### 1.3.3 Publicaciones científicas y patentes

En este punto resumo las aportaciones científicas llevadas a cabo hasta la fecha, referentes a libros de investigación, artículos y congresos, seguidos de los registros de propiedad intelectual, en relación con sistemas y ayudas para la navegación de personas invidentes.

#### Libros o capítulos de libros de investigación

1. TÍTULO: Virtual Moving Sound Source Localization through Headphones  
AUTORES: Larisa Dunai, Guillermo Peris-Fajarnés, Teresa Magal-Royo, Beatriz Defez, **Victor Santiago Praderas**  
EDITORIAL: InTech  
REF. REVISTA/LIBRO (ISBN, ISSN, SUPV): Advances in Sound Localization ISBN 978-953-307-224-1  
AÑO: 2011

#### Artículos de revista JCR

1. TÍTULO: The influence of the inter-click interval on moving sound source localization for navigation systems  
AUTORES: Larisa Dunai; Guillermo Peris Fajarnes; Beatriz Defez García; **Victor Santiago Praderas**; Ion Dunai  
REVISTA/LIBRO (ISBN, ISSN, SUPV): ACOUSTICAL PHYSICS (ISSN 1063-7710)  
VOLUMEN: 56 PÁGINAS: 370 - 375 AÑO: 2010

#### Artículos de congresos nacionales e internacionales

1. TÍTULO: Real-Time Assistance Prototype - a new Navigation Aid for blind people  
AUTORES: Larisa Dunai; Guillermo Peris Fajarnés; **Victor Santiago Praderas**; Beatriz Defez García, Ismael Lengua Lengua  
EDITORIAL: IEEE e IES (Institute of Electrical and Electronics Engineers & IEEE Industrial Electronics Society)  
REF. REVISTA/LIBRO (ISBN, ISSN, SUPV): 36th Annual Conference of the IEEE Industrial Electronics Society (IECON 2010) (ISBN 978-1-4244-5226-2)  
AÑO: 2010

2. TÍTULO: COGNITIVE AID SYSTEM FOR BLIND PEOPLE  
AUTORES: **Victor M. Santiago Praderas**; Ortigosa Araque, Nuria; Dunai, Larisa; Peris-Fajarnes, Guillermo (CASBLiP) EDITORIAL: Departamento de Ingeniería Grafica REF. REVISTA/LIBRO (ISBN, ISSN, SUPV): XXI Congreso Internacional de Ingeniería Grafica (ISBN 978-84-96351-57-8)  
AÑO: 2009 [Santiago Praderas V. et al, 2009]
3. TÍTULO: A Multiple Sensors Prototype for Visually Impaired Subjects Mobility Assistance Using Sound Map Generation  
AUTORES: Nuria Ortigosa; Larisa Dunai; Samuel Morillas; **Victor Santiago**; Guillermo Peris-Fajarnés; Ion Dunai  
EDITORIAL: Universidad de Valencia  
REF. REVISTA/LIBRO (ISBN, ISSN, SUPV): European Conference LIVINGALL 2009 (ISBN 0000000000)  
AÑO: 2009
4. TÍTULO: Acoustical Navigation System for Visual Impaired People  
AUTORES: Larisa Dunai; Nuria Ortigosa; Samuel Morillas; **Victor Santiago Praderas**; Guillermo Peris-Fajarnés; Ion Dunai  
EDITORIAL: Universidad de Valencia  
REF. REVISTA/LIBRO (ISBN, ISSN, SUPV): European Conference LIVINGALL 2009 (ISBN 0000000000)  
AÑO: 2009
5. TÍTULO: Sound Map Generation for a Prototype Blind Mobility System Using Multiple Sensors [Ortigosa Araque N. et al, 2009]  
AUTORES: Nuria Ortigosa; Larisa Dunai; Samuel Morillas; **Victor Santiago Praderas**; Guillermo Peris-Fajarnés; Ion Dunai  
EDITORIAL: companioable  
REF. REVISTA/LIBRO (ISBN, ISSN, SUPV): International Congress on Accesible and Assistive Technologies for dependent people. ABLETECH08 (ISBN 0000000000000000)  
AÑO: 2008
6. TÍTULO: Utilización de mapas de disparidad para la detección de áreas libres  
AUTORES: Nuria Ortigosa; Samuel Morillas; **Victor M. Santiago**; Guillermo Peris-Fajarnés  
EDITORIAL: Ibergaceta Publicaciones et Al. (Eds)

REF. REVISTA/LIBRO (ISBN, ISSN, SUPV): II Workshop de Reconocimiento de formas y análisis de imágenes.  
AERFAI (ISBN 978-84-92812-66-0)  
AÑO: 2010

7. TÍTULO: CASBlIP - a new cognitive object detection and orientation system for impaired people  
AUTORES: Guillermo Peris Fajarnes; Larisa Dunai; **Victor Santiago Praderas**; Ion Dunai  
EDITORIAL: Universidad de Zürich  
REF. REVISTA/LIBRO (ISBN, ISSN, SUPV): Actas de Congreso  
AÑO: 2010
  
8. TÍTULO Virtual acoustic environment interface for sound source localization training  
AUTORES: L.DUNAI; B. DEFEZ; **V. SANTIAGO PRADERAS**;  
N. ORTIGOSA ARAQUE; G.PERIS-FAJARNÉS  
EDITORIAL: Ingegraf  
REF. REVISTA/LIBRO (ISBN, ISSN, SUPV): INGEGRAF 2011  
AÑO: 2011
  
9. TÍTULO: Image processing and tactile methodologies of a navigation system for visually impaired people.  
AUTORES: N. ORTIGOSA; **V. SANTIAGO PRADERAS**;  
L.DUNAI; G.PERIS-FAJARNÉS; S. MORILLAS  
EDITORIAL: Ingegraf  
REF. REVISTA/LIBRO (ISBN, ISSN, SUPV): INGEGRAF 2011  
VOLUMEN: PÁGINAS: - AÑO: 2011 CLAVE:

### **Patentes presentadas**

SISTEMA DE AYUDA A LA MOVILIDAD DE PERSONAS CIEGAS  
MEDIANTE EL USO DE CÁMARAS Y SONIDO ESPACIALIZADO  
Referencia: R15646-2011

AUTORES:

Víctor Santiago Praderas  
Guillermo Peris Fajarnés  
Larisa Dunai  
Beatriz Defez

### **Registro de propiedad Intelectual**

Autor del software “EYE21” para la ayuda a la movilidad de personas ciegas mediante el uso de cámaras y sonido espacializado, registrado en el CTT de la Universidad Politécnica de Valencia con registro número R-15882-2011

## **1.4 Objetivos**

Tras un estudio riguroso y exhaustivo sobre la situación actual de las personas invidentes en el mundo, así como de los sistemas tecnológicos investigados y las patentes desarrolladas, parto de lo siguiente:

1. La integración de diferentes tecnologías en plataformas y estructuras funcionales son válidas para ayudar a navegar por el mundo real a personas invidentes.
2. A partir de la hipótesis 1, ¿qué metodologías y experimentos se pueden utilizar para demostrar que la integración diseñada e implementada obtiene resultados satisfactorios?

Por lo tanto, a partir de bloques constituyentes existentes y de funcionalidad probada, se va a construir un sistema que:

1. Integre diferentes tecnologías (CMOS, estéreo visión y mapas acústicos) para ayudar a navegar a personas invidentes.
2. Garantice la autonomía de las personas invidentes, y sobre todo consiga que estas personas sientan, y no que escuchen sintéticamente el espacio. El resto de sistemas estudiados trabajan con sonidos descriptivos del entorno, pero no sensibilizan la parte del cerebro que interpreta y entiende el espacio, tal como hace el sistema expuesto en esta tesis.

Finalmente, el trabajo doctoral se complementará:

1. Definiendo las condiciones del contexto y el entorno en las que se van a realizar las pruebas de navegación con diferentes usuarios invidentes.
2. Aplicando distintos métodos para demostrar experimentalmente que las pruebas realizadas sobre diferentes personas invidentes con la tecnología integrada es útil, funcional, y cumplen con la función perseguida, es decir, mejorar la movilidad de este grupo de personas.
3. Documentando los métodos empleados para poder ser aplicados a futuras líneas de trabajo, e incluso con otros sistemas desarrollados con nuevas tecnologías.

Los resultados obtenidos al finalizar este trabajo reflejan su aplicabilidad y abren nuevos caminos y líneas de investigación para futuros estudios.

Por tanto, para poder llegar a resolver de forma experimental y científica los retos de esta tesis, los objetivos planteados son:

1. Diseñar y desarrollar un prototipo que integra diferentes tecnologías (CMOS, estéreo visión y mapas acústicos) para ayudar a navegar a personas invidentes.
2. Definir las condiciones del contexto y el entorno en las que se van a realizar las pruebas de navegación con diferentes usuarios invidentes.
3. Aplicar distintos métodos para demostrar experimentalmente que las pruebas realizadas sobre diferentes personas invidentes con la tecnología integrada, es útil, funcional y cumplen con la función perseguida, es decir, mejorar la movilidad de este grupo de personas.
4. Garantizar la autonomía de las personas invidentes, y sobre todo conseguir que estas personas sientan, y no que escuchen sintéticamente el espacio. Este objetivo no se ha cumplido en los sistemas estudiados, básicamente debido a las limitaciones tecnológicas.
5. Documentar los métodos empleados para poder ser aplicados a futuras líneas de trabajo, e incluso con otros sistemas desarrollados con nuevas tecnologías.

## **1.5 Conclusiones**

En el presente capítulo se han comentado y clasificado por su principio de funcionamiento y componentes, las características de 40 dispositivos de navegación. Se han clasificado en tres grupos: detectores de obstáculos, sensores de entorno y dispositivos de navegación. Este estudio ha demostrado que ninguno de los dispositivos desarrollados hasta el momento han conseguido llegar a cumplir los requerimientos de un dispositivo de movilidad para personas invidentes: seguridad, confort, estética e independencia.

Las críticas y comentarios más importantes de las personas invidentes sobre este tipo de dispositivos reflejan la falta de confianza en los sistemas existentes: sonidos estridentes y desagradables; confusión con los sonidos del entorno; las grabaciones vocales son largas y requieren tiempo; los dispositivos táctiles son grandes y han de trasportarse, lo que significa menos autonomía y movilidad; son dispositivos caros y en fase prototipo, y algunos de ellos son complejos y difíciles de usar.

Además, se han estudiado distintas patentes existentes de los últimos años y se puede concluir que muchas de ellas hablan de los componentes del sistema de manera funcional, pero no desde el punto de vista ergonómico ni cognitivo. Desgraciadamente, también se puede comprobar que el diseño definitivo de estas patentes en productos reales, licenciados y con explotación en la sociedad actual, no ha tenido mucho éxito, no habiendo conseguido comercializar un sistema con aceptación plena por parte de los usuarios y personas invidentes.

Por tanto, el campo de la ayuda a la navegación de personas invidentes y su inclusión plena y total en el sistema social y todos sus servicios, aún tiene enormes lagunas en lo que a la tecnología se refiere, es decir, tenemos que seguir haciendo grandes esfuerzos en seguir adelante investigando y analizando posibles nuevos sistemas. Es por ello por lo que este trabajo da un paso más a contribuir en el conocimiento de la ciencia para aportar nuevos métodos en el campo de la integración tecnológica, estableciendo las bases necesarias para ayudar a futuros sistemas e investigadores a seguir pautas metodológicas y asegurar que las tecnologías venideras cumplan con las expectativas de este grupo de personas.

En definitiva, esta tesis aporta un sistema compuesto por tecnología de estéreo visión, laser y bancos de sonidos, el cual demuestra que, tras una adaptación,

entrenamiento y siguiendo unos protocolos de aprendizaje, la secuenciación de los experimentos descritos en el punto 4 garantiza que las personas invidentes pueden detectar obstáculos y navegar, tanto en entornos abiertos como en entornos cerrados. También se cumplen con la mejoras de lo que las personas invidentes han demandado en los últimos años: confusión, calidad del sonido, autonomía, menos concentración en el uso del sistema, movilidad y ergonomía. Los experimentos llevados a cabo aseguran que los resultados obtenidos con los sistemas anteriores se han mejorado con la integración del prototipo presentado en este trabajo.



## 2. CASBLiP, DESCRIPCIÓN DEL PROYECTO

### 2.1 Resumen

En este capítulo se describe cómo nació la idea de desarrollar una herramienta para ayudar a las personas invidentes, así como cuales han sido los objetivos iniciales, tanto tecnológicos como científicos de la propuesta inicial. De la misma manera se describirán los socios tecnológicos del proyecto y las características técnicas, de funcionamiento y del diseño del sistema.

El dispositivo final está compuesto por dos partes de captura de datos del entorno. Una de ellas está formada por sensores 3D CMOS y la otra por una interfaz basada en estereovisión. Los datos capturados por estos dispositivos son enviados a un PC para ser procesados y, posteriormente ser transformados en señales acústicas, las cuales son extraídas de un banco de sonidos almacenado y pre procesado en un dispositivo FPGA (Field Programmable Gate Array). El primer sistema laser, trabaja en distancias comprendidas entre 0,5 y 5 metros y es capaz de detectar obstáculos, tanto estáticos como en movimiento, y en un azimut de 60°. Como complementario, el sistema de visión trabaja a partir de 5 metros y hasta 15 metros siendo capaz de detectar pasillos libres y objetos estáticos y en movimiento. Cuando la información es capturada por el sistema, los datos del láser son enviados al PC para ser procesados e interpretados, sacando resultados mediante algoritmos de medida y efectos rebote (Time of Flight) para obtener distancias. Por otra parte y en paralelo, los datos procedentes del sistema de visión son procesados a través de algoritmos de filtrado de imagen para obtener mapas de profundidad y a partir de ellos aplicar criterios de escala de grises, tracking y segmentación y conocer las distancias de los objetos y pasillos libres a sonificar.

El mapa de sonidos se basa en la convolución de un sonido con los HRTFs (Head Related Transfer Function) [Versenyi G., 2007] medidos en una cámara anecoica. Los HRTFs representan la Transformada Fourier de la Head Related Impulse Respons (HRIR), en otras palabras, representa los sonidos registrados en unas coordenadas especiales del espacio tanto para el oído izquierdo como para el oído derecho en el dominio temporal.

La convolución representa la unión de dos señales acústicas, es decir, componemos el HRTF de cada coordenada con el sonido seleccionado mediante el método de la convolución, manteniendo en todo momento las

características espaciales de los HRTFs pero con información complementaria para su interpretación.

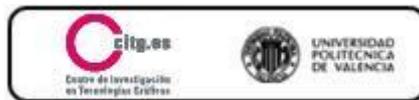
Veremos cómo se realizó una importante labor, no solo desde el punto de vista de la investigación, sino intentando integrar en un solo prototipo diferentes tecnologías y resultados procedentes de distintas fuentes de datos.

## **2.2 Objetivos del proyecto CASBLiP**

CASBLiP ha sido un proyecto financiado por la Unión Europea dentro del programa estratégico "Sistemas Cognitivos" del 6 ° PM. Se trata de un proyecto de investigación que se ha llevado a cabo entre el 1 de febrero de 2006 y el 31 de Enero 2009.

El proyecto CASBLiP nace con la intención de avanzar un paso más en los sistemas de ayuda hacia personas invidentes integrando diferentes tecnologías en una plataforma autónoma. Las tecnologías implicadas en el desarrollo del proyecto son de diferente índole, de ahí la complejidad encontrada a la hora de diseñar su integración e intercomunicación entre todos los datos que son capturados por dispositivos externos.

Siete socios han intervenido en la realización de este proyecto. Entre ellos están:



**Nombre:** Centro de investigación de Tecnologías Gráficas de la Universidad Politécnica de Valencia.

**País:** España.

**Misión:** Coordinador del Proyecto y colaboración en procesado y filtrado de imagen, así como abordar completamente la integración total del proyecto.



**Nombre:** Universidad de La Laguna. (Departamento de fisiología de la escuela de medicina).

**País:** España.

**Misión:** Equipo de médicos neurólogos, neurofisiólogos e ingenieros encargados de diseñar un banco de sonidos para excitar auditivamente el área del cerebro de personas invidentes y ayudarles a interpretar el espacio.



**Nombre:** Universidad Politécnica delle MARCHE (Departamento de Mecánica).

**País:** Italia.

**Misión:** Socio involucrado en tareas de localización utilizando sistemas de GPS y HPS.



**Nombre:** SIEMENS AG

**País:** Alemania.

**Misión:** Socio tecnológico responsable del diseño del sistema 3D CMOS.



**Nombre:** Servizio Tiflotecnico – CISAD (Instituto de ciegos Francesco Cavazza).

**País:** Italia.

**Misión:** Poner a personas invidentes en el proyecto para testear y comprobar los avances del mismo bajo condiciones reales y en diferentes contextos. Las personas invidentes eran ciegos de nacimiento, ciegos por accidente y de distintas edades y estatus social.



**Nombre:** Deutscher Blinden- und Sehbehindertenverband e.V. (Instituto de ciegos DBSV).

**País:** Alemania.

**Misión:** Poner a personas invidentes en el proyecto para testear y comprobar los avances del mismo bajo condiciones reales y en diferentes contextos. Las personas invidentes eran ciegos de nacimiento, ciegos por accidente y de distintas edades y estatus social.



**Nombre:** Departamento de informática de la Universidad de Bristol

**País:** Inglaterra.

**Misión:** Investigación y desarrollo de algoritmos para el procesamiento de imagen basados en sistemas de estereovisión.

La tecnología empleada en la adquisición de datos es estéreo visión, láser, y sonificación. Todo ello rodeado de diversos componentes electrónicos (FPGA), GPS, desarrollo software y comunicaciones.

Los objetivos del proyecto han sido diversos:

- Interpretar la información del entorno y transformarla en señales acústicas entendibles por personas ciegas para ayudarles en su movilidad.
- Crear un sistema acústico de representación del mundo 3D.
- Desarrollar algoritmos de procesamiento de imagen en tiempo real tanto en entornos cerrados como al aire libre.
- Crear un interface para el usuario final que integre toda esta tecnología y ofrezca seguridad antes riesgos y objetos, tanto estáticos como en movimiento.
- Analizar e implementar todos los protocolos necesarios para integrar todos los módulos intervinientes en el proyecto.
- Finalmente, generar e integrar en un prototipo final que pueda ser comercializado, el cual se comenta en esta tesis.

Estos objetivos se pueden dividir a su vez en objetivos tecnológicos y en objetivos científicos.

Más detalladamente, los objetivos tecnológicos son:

- Integrar en una plataforma común todos los datos adquiridos del mundo real y que proceden de fuentes diferentes: del sistema de visión artificial, del sistema 3D CMOS y del sistema de posicionamiento.
- Integrar en el prototipo un sistema de modelado en 3D y un sistema inteligente para la detección de obstáculos y riesgos.
- Desarrollar nuevas formas de interactuar con las personas: un interfaz visual (que producen imágenes en color y simplificado) para las personas parcialmente ciegas, una interfaz acústica (produciendo mapas acústicos pixelado) para la ceguera total.

Y de la misma manera, se pueden resumir los objetivos científicos de este proyecto en:

- Desarrollo de algoritmos de imagen en tiempo real para aplicarlos en entornos abiertos.
- Creación de un modelo 3D combinando información de diferentes fuentes con el fin de ayudar cognitivamente a personas invidentes.
- Aplicar criterios y filtrados para poder distinguir entre objetos estáticos y en movimiento, con o sin riesgos para las personas invidentes.
- Crear un interfaz final para el usuario con la característica de ser capaz de representar espacios reales usando información acústica.
- Investigar de forma general sobre las necesidades que tienen las personas ciegas para ser capaces de percibir el espacio circundante y definir la cantidad de información justa, necesaria y apropiada para su movilidad en diferentes situaciones.

Por tanto, el objetivo principal de CASBLIP fue desarrollar un sistema capaz de interpretar y gestionar la información del mundo real de diferentes fuentes para apoyar la movilidad a cualquier tipo de persona con discapacidad visual: ciegos, deficientes visuales y personas con pérdida progresiva de visión. La información tridimensional del mundo real se adquiere y se transforma, o bien en imágenes tratadas para las personas con deficiencias visuales, o en señales acústicas enviadas por los auriculares para personas invidentes totales.

## 2.3 Diseño y características del sistema

Para conseguir los objetivos mencionados en el punto anterior, así como para generar un producto tangible y real que pueda demostrar el resultado de todo el desarrollo del proyecto, se han desarrollado dos prototipos. Estos dos prototipos sirvieron también para la validación de los conceptos y las extensas pruebas se han realizado con usuarios con deficiencias visuales.

El primer prototipo acústico o M1, que trabaja con un sensor de tiempo de vuelo y una interfaz de audio para transformar los datos de la distancia en un mapa de sonido espacializado (Figura 2.3.1).

Prototipo M2 para la movilidad en tiempo real, el cual está equipado con cámaras de estereovisión y controlado por un mini PC para que los usuarios puedan navegar con seguridad en espacios interiores y exteriores. Uno de los objetivos, para complementar el prototipo M1, fue el cubrir más distancia con mayor precisión (Figura 2.3.2).

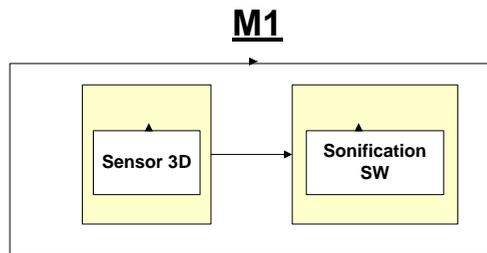


Figura 2.3.1 Prototipo M1

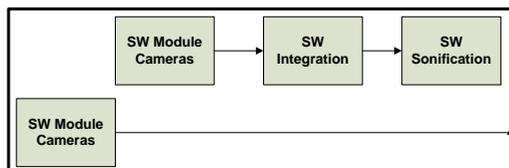


Figura 2.3.2 Prototipo M2

### **2.3.1 Prototipo acústico: M1**

Este primer prototipo, está formado por dos procesos: la captura de información procedente del mundo real y la transformación de dicha información en sonidos interpretables por los usuarios finales, personas ciegas en nuestro caso.

La información procedente del mundo real es capturada mediante un sensor de imagen CMOS (Complementary Metal Oxide Semiconductor), capaz de calcular información de la distancia de los objetos situados en un rango de 0 a 5 metros.

En comparación con los métodos tradicionales de sensores, el sensor CMOS 3D propuesto combina una alta percepción de sensibilidad, con bajos costes de producción y construcción y una integración hardware de pequeño tamaño.

Esta información, básicamente distancias de los objetos situados en el punto de mira del hardware CMOS, ha de ser transformada en mapas acústicos. Para ello es enviada a una FPGA (Field Programmable Gate Array) para finalmente, transformar en sonidos las distancias de los objetos y enviarla a las personas invidentes a través de unos auriculares.

El sistema final está compuesto por unas gafas (Figura 2.3.1.1), las cuales incluyen toda la electrónica láser necesaria para la captura de la información de las distancias, el ensamblado de los circuitos que componen la FPGA (Figura 2.3.1.2) encargada de sonificar la información [Dunai L. et al., 2009] procedente del láser y unos auriculares, además de la batería para alimentar la circuitería y demás cableado.



Figura 2.3.1.1 Gafas del M1



Figura 2.3.1.2 FPGA M1

El subsistema acústico es el encargado de reproducir sonidos espacializados, es decir, aquellos que serán 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, es decir, se han pre procesado previamente simulando puntos situados a distancias específicas, para que después se puedan percibir como procedentes de objetos reales. Posteriormente se almacenan 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.

### 2.3.1.1 Método de medida de distancias con 3D CMOS

Los métodos para calcular distancias son muy importantes, ya que el objetivo de ello es aplicar distintas técnicas a distintos contextos de la vida cotidiana: tráfico, robótica, aviación, etc. El método más antiguo, es el método clásico “Medida de Distancias Directas” que consiste en la comparación directa entre la distancia y un sistema de medida (por ejemplo una regla) calibrada. En los últimos tiempos se han desarrollado métodos de “Medida de Distancias Indirectas”, donde el cálculo de la distancia se obtiene de mediante la aplicación de reglas físicas y matemáticas.

El principio de funcionamiento del sistema 3D CMOS se basa en el método de medidas de distancias indirectas Time of Flight (TOF). Este método consiste en que el sensor emite pulsos de luz, que al ser reflejados con la superficie de un objeto, mide el tiempo utilizado para la ida y vuelta del pulso luminoso. De esta manera se calcula la distancia entre el CMOS y el objeto (1).

$$d = v * \frac{\tau_{TOF}}{2} \quad (1)$$

Donde  $d$  representa la distancia a medir,  $v$  es la velocidad de propagación del impulso luminoso y  $\tau_{TOF}$  es el tiempo medido entre la señal enviada al objeto y la recibida por parte del sensor.

En nuestro caso, TOF del impulso emitido no se mide directamente, sino que se van almacenando un conjunto de datos en un chip del sensor para luego ser procesados y obtener la distancia definitiva.

El sistema sensor 3D CMOS procesa los pulsos de luz basados en el método de medida TOF, utilizando pulsos de luz infrarrojos de corta distancia (NIR), en un rango de 20 a 200 ns.

Como muestra la figura 2.3.1.1.1, el sensor emite mediante el módulo de emisión pulsos de luz. Cuando estos pulsos son reflejados desde la superficie del objeto, son devueltos con un retraso de tiempo y disminución de amplitud a la cámara del sensor. Las lentes de la cámara transmiten parte de la dispersión de la luz recibida al chip del sensor. El chip contiene píxeles

sensibles a la luz que funcionan como integradores electrónicos de una corta ventana de duración. Estos integradores están sincronizados con los pulsos emitidos por parte del láser, donde finalmente se aplica el método de cálculo de distancia TOF al pulso del láser. Después, se aplican una serie de algoritmos de evaluación de distancia: Método gradiente [Magnus, 1952], método CSI y DCSI, para finalmente llegar a calcular el  $\tau_{TOF}$  (tiempo de retardo del impulso) de un conjunto de medidas en bloque. La distancia al objeto se calcula como indica la fórmula (1).

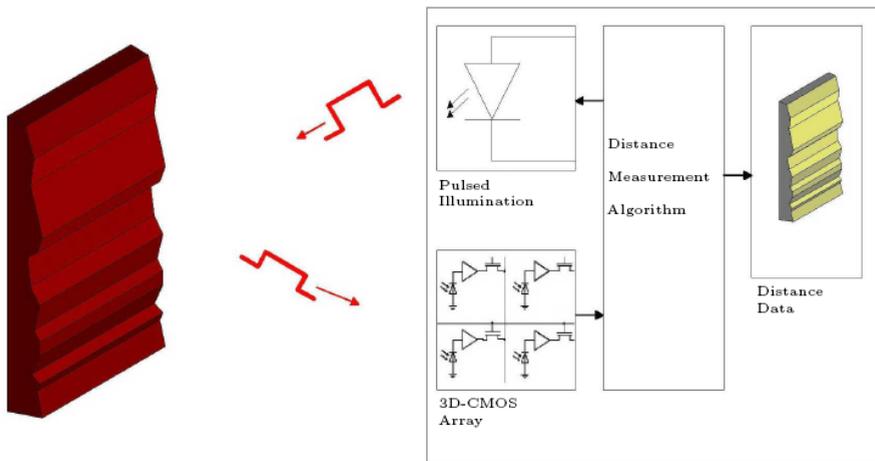


Figura 2.3.1.1.1 Principio de medida 3D CMOS

Las lentes de la cámara están formadas por una lente de proyección de imagen de distancia focal, por un plano del sensor paralelo al plano de la lente y por una ranura limitando el diámetro efectivo de la lente  $D$ .

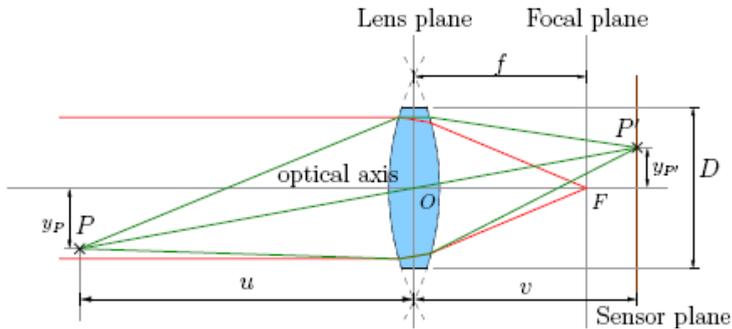


Figura 2.6 Método de proyección de la luz a través de las lentes ópticas [Bernhard, 2008]

Donde  $P$  representa la luz entrante,  $P'$  la luz tratada,  $u$  y  $v$  representa las distancias de  $P$  y  $P'$  con respecto al plano de las lentes.

### 2.3.1.2 Sensor 3D CMOS

En la figura 2.7 se puede ver todos los componentes del sistema 3D CMOS. a) Representa la electrónica, incluyendo la matriz de sensores; b) es el módulo de iluminación laser; c) es la óptica y d) es la batería del sistema.

#### a) Electrónica

La matriz de sensores 3D CMOS se compone de  $1 \times 64$  pixeles basados en el sistema 3D-CMOS de  $4 \times 64$  pixeles desarrollado por Fraunhofer IMS dentro del proyecto UseRCams. El sistema permite obtener medidas indirectas de los impulsos TOF del láser en magnitudes de ns.

En el trabajo realizado, por cuestiones de espacio y tamaño del sistema, se ha utilizado solo una línea de 64 puntos.

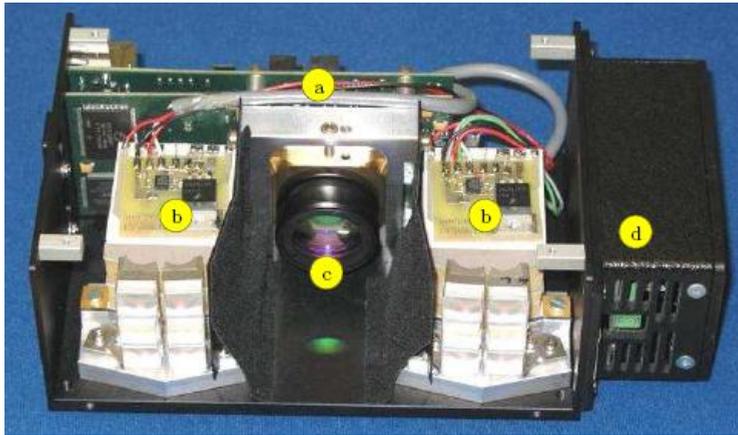


Figura 2.7 Sistema Sensor 3D CMOS

Según los parámetros de la figura 2.8, se puede calcular el área de sensibilidad del funcionamiento del pixel.

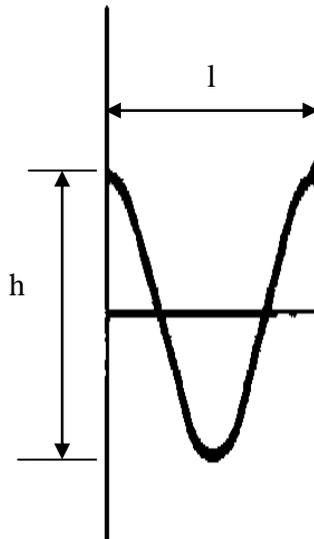


Figura 2.8 Principio de cálculo de las propiedades del pixel:  $x = nx \times mx \times h$

Donde  $n$  es el número de impulsos (píxeles),  $m$  es el número de niveles (matriz de píxeles) en elevación o altura,  $l$  representa la longitud de la onda y  $h$  es la amplitud de la onda.

El área del píxel es  $130 \mu\text{m}$  en horizontal y  $300 \mu\text{m}$  en dirección vertical. De este modo obtenemos el área de sensibilidad de  $64 \times 130 \mu\text{m} \times 1 \times 300 \mu\text{m} = 2,496 \text{mm}^2$ .

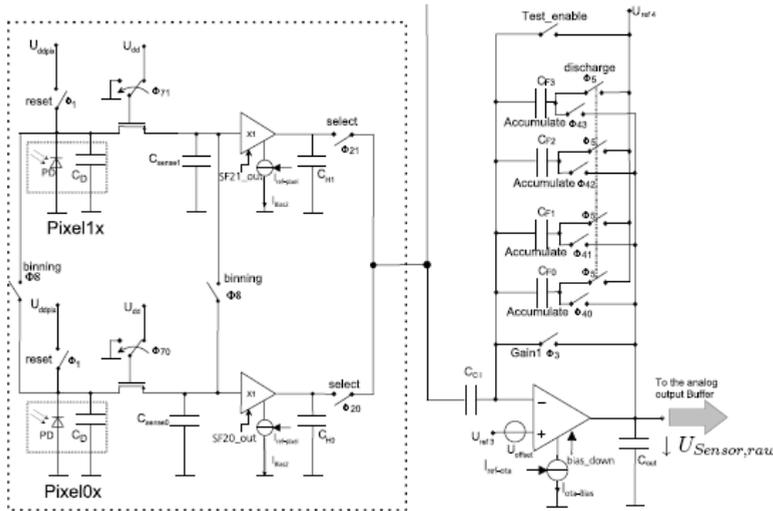


Figura 2.9 Circuito del píxel CDS readout stage [Elkhalili et al]

El diseño mostrado en el circuito de la figura 2.9, reduce considerablemente la complejidad del hardware.

La operatividad del circuito se basa en el reajuste periódico de la capacidad  $C_D$  y la capacidad  $C_{sense}$  del fotodiodo para llegar a tener un potencial fijo  $U_{ddpix}$ .

### b) Módulo de iluminación láser

El módulo de iluminación láser juega un papel decisivo en el método del cálculo de distancias de las cámaras 3D CMOS, debido a que emite señales de luz que son las que actúan como emisoras-receptoras de la luz en el espacio y los objetos. El elemento básico del módulo de iluminación son los diodos láser, y está montado sobre una placa de cerámica.

### c) Óptica

Ha sido diseñada para cubrir un ángulo de  $60^\circ$  en posición horizontal, con una resolución mejor que la del tamaño el pixel  $1x64$  3D-CMOS. Además, no provoca distorsión en el funcionamiento del 3D-CMOS.

El tamaño de la imagen es de  $8,32 \text{ mm} \times 2,40 \text{ mm}$  y representa el área activa del sensor.

Está formada por una capa anti reflejante para longitudes de onda del láser de  $905 \text{ nm}$  y reflejar el resto de señales.

Antes de poner en funcionamiento las cámaras que forman la óptica se han realizado unas operaciones de ajuste, entre las que se encuentra su calibración offset y la calibración en distancias.

La calibración en distancias, como punto referencia para posteriores cálculos, se basa en los parámetros de calibración por pixel utilizados en el proceso de calibración de las cámaras. Esta calibración consiste en eliminar deformaciones que la interpretación de los pixeles pueda introducir y asegurar la coherencia de la información capturada.

Finalmente, el sensor 3D CMOS fue diseñado e integrado en una gafas tal como se muestra en la figura 2.10.

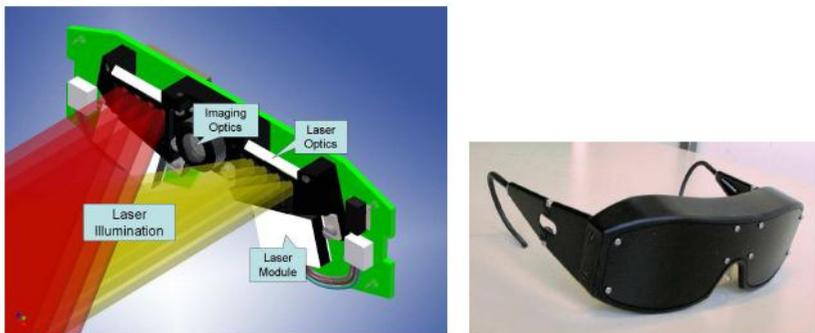


Figura 2.10 Componentes del sensor 3D CMOS y su integración en gafas

### 2.3.2 Prototipo de visión: M2

La segunda parte del proyecto ha consistido en desarrollar un segundo prototipo denominado M2 (Figura 2.3.2).

A diferencia de M1, y aunque el objetivo final es el mismo, es decir, capturar información del mundo real y sonificarla [Dunai L. et al., 2009], [Dunai I. et al., 2009], ahora la información es capturada mediante cámaras de estereovisión y con formato 3D. El sistema segmenta la imagen en diferentes áreas de interés, basándose concretamente en la detección de un objeto (en movimiento con prioridad sobre los estáticos, o el más cercano en caso de varios objetos estáticos) y de pasillos o áreas libre de paso, por las cuales el invidente puede circular sin peligro alguno. El sistema funciona en tiempo real (10 fps) por lo que se puede utilizar mientras se camina por la calle a ritmo normal. Este tiempo viene fijado por la cantidad de imágenes que se pueden procesar por segundo, ya que el conjunto de captura de imágenes, la sincronización, su cacheo y el filtrado requiere un tiempo entre imágenes de  $1/10 \text{ s} = 0,1 \text{ s} = 100 \text{ ms}$ .

El primer paso ha consistido en desarrollar mecanismos de tratamiento de imagen para, mediante la obtención de diferentes mapas de profundidad reconocer objetos estáticos y en movimiento, además de pasillos libres. Tras la obtención de los mapas de profundidad de la escena, los diferentes niveles de gris son indicativos de las diferentes distancias a los objetos de la misma.



Figura 2.11 Mapa de Profundidad

Tras este análisis de distancias, se localiza el objeto móvil más próximo a la posición de las cámaras, o bien la de aquel que se esté moviendo a mayor velocidad hacia éstas, al igual que se aplican diferentes algoritmos para la detección de vías libres de obstáculos mediante el análisis de los mapas de profundidad.

Toda esta información ha sido integrada mediante lenguajes de alto nivel (visual estudio y C++), en un interfaz para traspasar la información al módulo encargado de sonificar las imágenes. Este módulo trabaja de manera similar al descrito en el prototipo M1, pero con diferentes bancos de sonido y para distancias comprendidas entre 5 y 15 metros.

Paralelamente existe un sistema basado en localización GPS para, de forma independiente al prototipo M2, ofrecer al usuario información de su situación y ruta a seguir para llegar a un punto concreto.

Las diferentes investigaciones llevadas a cabo durante el desarrollo del proyecto han desembocado en la unión e integración de diferentes tecnologías, con el objetivo común de unir información procedente del mundo real 3D y convertirla en sonidos para poder guiar a personas con deficiencias visuales. A continuación se describen por separado cada una de ellas.

### **2.3.2.1 Módulo de visión artificial**

La parte de visión artificial está compuesta por cámaras de estéreo visión y algoritmos de procesamiento de imagen, y su objetivo ha sido analizar las escenas que el usuario tiene en frente y extraer características sobre los obstáculos que la forman, tanto objetos fijos como en movimiento, además de llegar a interpretar el índice de peligrosidad que estos obstáculos tienen para el usuario. Sobre todo, la investigación se centró en la detección de objetos con movimiento rápido, para obtener su posición relativa, su dirección y la velocidad.

No fue objetivo del sistema de visión el reconocer los objetos, pero si determinar si son potencialmente peligrosos aquellos objetos que, sin saber de qué se tratan, poder determinar que por sus características pueden llegar a dañar la integridad física del usuario, y debemos avisar al usuario para restringir su movimiento o dirección. Todo ello mediante sonidos, atacando la parte cognitiva de la persona, es decir, que sea ella la que pueda tomar decisiones y aprender a manejarse por el espacio.

Dada la complejidad que tienen los sistemas y entornos al aire libre, esta interpretación de las escenas sigue siendo tarea muy difícil, no sólo por cuestiones de iluminación, ambigüedades de objetos, etc. sino por cuestiones de movimiento. También es muy importante hacer frente a los requisitos y

necesidades de procesar y calcular en tiempo real. Dos cámaras de estéreo visión y un sensor de inercia 6 grados de libertad se han utilizado como componentes clave del sistema de visión. Las cámaras están dotadas con un control de iris automático con el fin de adaptarse a los cambios de iluminación.

El hardware utilizado en la captura de imágenes ha sido:

- 2 cámaras Fireware Flea 2 fabricadas por Point Gray Research [Flea2 Firewire camera] (ver Figura 2.13). Esta opción fue seleccionada por sus prestaciones en ancho de banda y alta resolución.
- 1 sensor inercial con acelerómetro con el fin de ofrecer información del movimiento independientemente de las cámaras y poder evitar riesgos con mayor eficacia.



Figura 2.12 Cámara Fireware Flea 2 color

### **2.3.2.2 Estimación de profundidad y segmentación de imágenes**

Para poder estimar la distancia de los objetos desde la referencia del punto de vista del usuario y con una precisión aceptable, se han usado un conjunto de dos cámaras para trabajar en modo estéreo visión. Para la estimación de la profundidad es requisito imprescindible el uso de dos cámaras. La segmentación de las imágenes se puede realizar con una o dos cámaras indistintamente. Su velocidad depende de los algoritmos utilizados.

Se eligieron las cámaras con interfaz Fireware para permitir tasas de tiempo de procesado on line y lo más próximas posible a la captura real de la imagen y los objetos.

La estimación de la profundidad se basa en la correlación y la búsqueda de correspondencia en las deformaciones para su posterior reconstrucción. Este tipo de procesos es muy eficiente desde el punto de vista computacional. Sin embargo, no es muy óptimo utilizar estos métodos en nuestra aplicación puesto que podemos obtener regiones aisladas e imágenes poco coherentes para asegurar que puedan pertenecer a un solo objeto, lo cual hace muy difícil su representación acústica y sonificación [Scharstein, D. and Szeliski, R., 2002]. Varios métodos se experimentaron, tales como el de la visión temprana [Felzenszwalb P. and Huttenlocher D., 2006], la programación dinámica [Birchfield, S. and Tomasi, C., 1999], la suma de la diferencia absoluta, y la suma de los cuadrados de las diferencias con agregación iterativa [Zitnick C. and Kanade T., 2000]. No se puede decir de forma global que unos sean mejores o peores que otros, ya que depende de la escena que se está analizando y del contexto de la imagen, es decir, si está en movimiento, si la imagen de ha capturado en interiores o exteriores, etc. Manejando todos estos parámetros y tras varios ensayos previos, el método que finalmente fue utilizado en el proyecto CASBLiP fue el de programación dinámica.

La figura 2.14, refleja la relación existente entre los mapas de profundidad y la segmentación dentro de una secuencia de imágenes.



Figura 2.13 Fusión de mapas de profundidad y segmentación de imágenes

### 2.3.2.3 Detección de obstáculos y estimación de movimiento

Los objetos más cercanos al ángulo de visión del usuario son los potencialmente más peligrosos y deben ser detectados con precisión. Esto puede solucionarse eficientemente mediante técnicas de *thresholding* sobre los mapas de profundidad, es decir, con el umbral bien calculado en base a parámetros de estereo visión y diferencias podemos representar las distancias de los objetos, tal como se muestra gráficamente en la figura 2.14. Un procesamiento morfológico de imágenes, basado en la extracción de los componentes de las imágenes para la representación de límites o bordes, nos ayuda a eliminar regiones aisladas y facilitarnos la sonificación de la imagen. Toda esta información extraída de los mapas de profundidad es utilizada por el módulo de sonificación para mejorar sus decisiones además de la proporcionada por el sensor CMOS descrito anteriormente, ya que este solo llega a dar información sobre objetos inmediatos situados a distancias de hasta 5 metros. Además, toda esta información del módulo laser y del módulo de visión será sumada a la generada de forma natural por el bastón tradicional.

Este método propuesto es bastante sencillo y puede avisar al usuario de obstáculos por encima del nivel del suelo.



Figura 2.14 Detección de obstáculos basados y mapas de profundidad

### 2.3.2.4 Detección de objetos usando mapas de profundidad

El primer objetivo respecto a la detección de objetos es identificar los objetos en movimiento independientemente de la escena, ya que lo más probable es que sean los que representen el mayor peligro para el usuario. Esto se consigue mediante el seguimiento de un conjunto de puntos de función [Hannuna S., 2007], lo que implícitamente etiqueta a los objetos en movimiento y segmentan sus características intrínsecas, con el fin de compararlas con las que no concuerdan con el movimiento de las cámaras.

El método de Kanade-Lucas-Tomasi (KLT) [Shi, J. and Tomasi, C., 1994], es el utilizado para generar este conjunto de puntos dispersos, en paralelo con el proceso de estimación de la profundidad. Este algoritmo de *tracking* se decanta preferentemente por identificar zonas de alta entropía y muy dispersas. Además de facilitar el seguimiento, esto también asegura que los valores obtenidos de los mapas de profundidad en las proximidades de los puntos KLT son relativamente fiables, ya que es probable que las correspondencias sean coherentes en estas regiones.

Los puntos correspondientes a los objetos en movimiento son segmentados de forma independiente utilizando MLESAC [Torr, P. and Zisserman Mlesac, A., 2000], basado en el supuesto de que el movimiento aparente generado por el movimiento en 3D de una cámara dentro de un ambiente (egomotion) es dominante. En concreto, técnicas de similitud, considerando el tamaño y la orientación del contorno rectangular, son utilizadas para deducir un modelo provisional sobre la base de la trayectoria de los cuatro puntos seleccionados al azar de una ventana deslizante temporal. Es entonces cuando el modelo más

probable es retenido. Los valores extremos corresponden, en general y de forma independiente, a los objetos en movimiento (pero a veces las regiones cercanas a las cámaras están aparentemente en movimiento debido al "efecto paralaje").

Además de facilitar la localización y seguimiento de los objetos, esta metodología también asegura que los valores extraídos de los mapas de profundidad en las cercanías de los puntos KLT sean relativamente fiables, ya que es probable que las correspondencias hayan sido las óptimas para estas regiones. En concreto, la homografía y detección de contornos es utilizada para deducir un modelo provisional sobre la base de la trayectoria de los cuatro puntos seleccionados al azar, de una ventana deslizante temporal, siendo finalmente el modelo más probable el seleccionado.

Los valores extremos corresponden a los objetos en movimiento, excluyendo los movimientos producidos por el efecto de paralaje de las cámaras, como se ha dicho anteriormente.

### **2.3.2.5 Sensor inercial**

La cámara está en constante movimiento, lo que hace muy difícil calcular el movimiento de los objetos, y sobre todo en tiempo real. Con el fin de conocer de manera más eficiente y precisa la estimación del movimiento de la cámara, un sistema inercial se integró en el sistema de estéreo visión. El sistema inercial está compuesto por un sensor giroscópico 3D y un acelerómetro 3D.

Esto implicó un desarrollo adicional para conseguir, tanto a nivel de hardware como de software un entendimiento y correcto funcionamiento del sistema. El sensor inercial se colocó junto con el equipo de estéreo visión, lo más próximo posible a la cámara izquierda (cámara configurada como cámara principal y como base de la captura de imágenes), la cual se encarga de proporcionar ángulos de rotación y parámetros de aceleración.

### **2.3.2.6 Sistema GPS**

En condiciones de visibilidad óptimas, los receptores estándares de GPS proporcionan los datos con un error medio de 15-10 m. Hoy en día, los sistemas GPS se utilizan principalmente para la navegación de automóviles en

carreteras, ayudando a los conductores a llegar a su punto final simplemente introduciendo en el sistema la dirección deseada.

El principal problema de esta tecnología está cuando nos encontramos en entornos en los que la señal de los satélites es débil o inexistente, por ejemplo cuando los usuarios se encuentran en un túnel. Para resolver estos problemas, se propuso el uso de un dispositivo llamado Dead Reckoning Compass (DRC), fabricado por Vectronix (figura 2.15). Se trata de un sensor inercial compuesto por tres acelerómetros, tres giroscopios y tres magnetómetros, el cual es extremadamente eficaz cuando la señal de GPS no es óptima.

Su funcionamiento se basa en detectar la posición de los sensores inerciales y calcular el desplazamiento de los usuarios desde un punto inicial. Este dispositivo se combina con el sistema GPS con el fin de mejorar la precisión de la posición y reducir el efecto de la visibilidad de los satélites y los errores parciales de las rutas a seguir.



Figura 2.15 Dead Reckoning Compass

El objetivo del uso de esta tecnología en el proyecto CASBlIP, fue realizar un sistema de posicionamiento para la navegación de los usuarios con discapacidad visual. Los parámetros más importantes a tener en cuenta fueron la desviación en el posicionamiento de los usuarios y la posibilidad de operar en condiciones adversas dentro de los entornos urbanos. También se tuvo muy en cuenta, la portabilidad, la ergonomía y la posibilidad de instalar este sistema en los usuarios finales de forma cómoda y estética.

Finalmente, se montó un sistema GPS basado en procedimientos para el aumento de señal EGNOS combinándolo, como se indica en el punto 2.3.2.5, con un sensor inercial que pueda suministrar datos de posición cuando la señal procedente de los satélites sea de mala calidad.

El sistema se diseñó para proporcionar un alto grado de funcionalidad y portabilidad, y se basa en el uso de un receptor GPS conectado a una PDA que integra una plataforma de software de generación de voz para poder vocalizar los datos de posicionamiento.

El sistema fue probado en un espacio abierto usando tecnología de alta precisión con D-GPS<sup>1</sup>, obteniendo un mapa con una precisión de  $\pm 1$  cm.

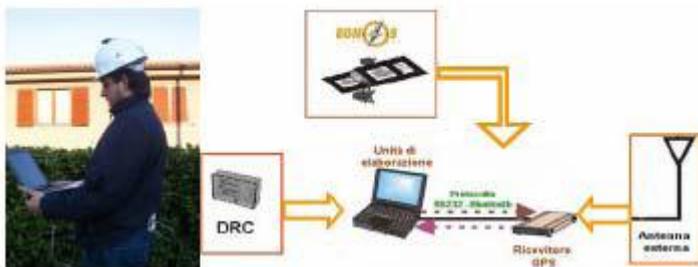


Figura 2.16 Versión experimental del sistema GPS

### 2.3.3 Interface de audio

Esta parte del proyecto consistió en el desarrollo de un módulo electrónico, denominado audio interface, cuya misión era relacionar la información procedente de los sensores externos (laser y video) para transformarla en una señal de audio y transmitirla a través de unos auriculares. El primer paso se llevó a cabo con el diseño del módulo estratega. Consiste en definir las reglas

---

<sup>1</sup> DGPS (Differential GPS), o GPS diferencial, corrige sobre los GPS tradicionales los datos recibidos por los satélites para proporcionar mayor precisión de las medidas.

y los protocolos de cómo se van a representar los datos y la información procedente de los subsistemas sensoriales. Las estrategias de sonificación son descritas en el punto 2.3.3.3. Todo esto se integró en una electrónica FPGA (Field Programmable Gate Array) como la descrita en el apartado 2.3.1.

Se desarrollaron dos prototipos, el primero fue una versión inicial, el cual dio un paso al definitivo para poderse utilizar en el sistema final. Estaba compuesto por un sistema de adquisición de datos mediante laser, mientras que el segundo prototipo se diseñó para poder atender los requerimientos del subsistema de estéreo visión.

Referente a la información 3D procedente de los sistemas de captura de información, se desarrollaron dos módulos de audio, uno para el sistema 3D-CMOS y otro para el sistema basado en estéreo visión.

En caso del primer prototipo, el módulo de audio tiene dos componentes, el estrategia y el mezclador. El estrategia, que define desde el punto de vista lógico que reglas se van a seguir para convertir el espacio capturado por las cámaras en sonidos, se ejecuta sobre un microprocesador, y el mezclador trabaja sobre la FPGA para poder gestionar el volumen del sonido generado. Todo está integrado y montado en una caja portable, interconectada directamente al sistema 3D-CMOS.

En la primera versión del prototipo, el estrategia se ejecutaba en un procesador independiente incluido en la misma caja de la FPGA, para cumplir con las especificaciones de portabilidad y peso. En el prototipo definitivo, donde se incluyó un ordenador portátil para la gestión de la información sensorial, el estrategia se ejecutó en el propio ordenador por razones de mayor flexibilidad y potencia, así como para facilitar los cambios y modificaciones en la representación e interpretación de la información. Por supuesto, en la versión definitiva la información procedente de las cámaras de estéreo visión y del GPS es gestionada por un sistema de audio basado en PC.

Todo el diseño de los módulos de audio se realizó siguiendo especificaciones de alta portabilidad, bajo peso, control de volumen por potenciómetro independiente y modularidad y autonomía.

### 2.3.3.1 Medida de HRTF

El objetivo de este subcapítulo es definir la metodología seguida para la obtención de los filtros espaciales llamados *Head Related Transfer Function* (HRTF). Los HRTFs son necesarios para obtener la colección de sonidos espaciales o virtuales que se utilizarán en el capítulo 4 para definir y representar la información sensorial y cognitiva de la percepción del entorno.

El sistema de medida de los HRTFs está compuesto por un sistema robótico y un sistema de adquisición de señales y HRTF. El sistema robótico se compone de un brazo automatizado con un altavoz y un sillón según se muestra en la figura 2.17.

El principio de medida del sistema consiste en que el brazo robotizado va moviendo el altavoz a lo largo y ancho de diferentes puntos espaciales, registrando los diferentes sonidos emitidos por el altavoz. Estos, son capturados por dos micrófonos situados a la altura de la cabeza de un maniquí simulado (KEMAR). Es entonces cuando se asocia la sensación del sonido generado a unas coordenadas en distancia, azimut y elevación. Cada punto obtenido se procesa y se almacena en un sistema informático. Este proceso se realiza por cada uno de los puntos espaciales entre  $60^\circ$  ( $30^\circ$  izquierda y  $30^\circ$  derecha respecto a un punto central) y 5 metros de distancia (desde 0,3 m a 5 m). Dentro de esta área, se genera con la metodología expuesta, un total de 64 sonidos, distribuidos cada  $0,93^\circ$ , debido a que el láser utilizado y explicado en el capítulo 2.3.1, está formado por un total de 64 píxeles distribuidos en el azimut de  $60^\circ$ .



Figura 2.17 Sistema de obtención de HRTFs

La base matemática para la generación de los modelos HRTF se basa en la transformada de Fourier según se explica a continuación:

Tenemos  $x_1(n)$  que representa el sonido digital a reproducir en un altavoz, y  $y_1(n)$  que hace lo propio con el sonido que resulta de grabar con un micrófono colocado en el canal de audio de una persona la respuesta de  $x_1(n)$ . De la misma forma, tenemos  $x_2(n)$  que representa el sonido a reproducir por un auricular, e  $y_2(n)$  la respuesta grabada del mismo. El objetivo que buscamos es, que para un determinado  $x_1(n)$ , generar un  $x_2(n)$  de tal forma que la resultante  $y_2(n)$  sea igual a  $y_1(n)$ , y de esta forma conseguiremos que una sensación sonora producida desde un altavoz, y otra generada desde un auricular, produzcan la misma respuesta en el canal auditivo de las personas, con el fin de llegar a conseguir la misma respuesta acústica y espacial. Para ello, se diseña un filtro digital para transformar  $x_1(n)$  en  $x_2(n)$ .

El filtro lo definimos en el dominio de la frecuencia. Sea  $X_1$  la representación de  $x_1(n)$  en el dominio transformado de la frecuencia, e  $Y_2$  la representación de  $x_2(n)$ . Al reproducir  $x_1(n)$ , la respuesta que se detecta se puede representar como:

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

donde  $L$  representa la función de transferencia del altavoz con toda la cadena de audio de reproducción,  $F$  la función de transferencia del espacio libre entre el altavoz y el canal auditivo, es decir, los HRTF, y  $M$  es la función de transferencia del micrófono y toda la cadena de audio de grabación. La respuesta grabada por el micrófono al reproducir  $x_2(n)$  por el auricular se representa como:

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

donde  $H$  es la función de transferencia del auricular al canal auditivo. Si igualamos  $Y_1 = Y_2$  y despejamos  $X_2$ , obtenemos:

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

por lo que el filtro digital que queremos conseguir se define por:

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

Tras todos estos cálculos, si aplicamos el filtro construido a la señal  $x_1(n)$ , y reproducimos la señal  $x_2(n)$  por los auriculares, se espera que la señal grabada por el micrófono sea  $y_1(n)$ , es decir, sería la misma que la que se obtendría al reproducir  $x_1(n)$  por el altavoz.

En realidad, se necesitarían dos filtros, uno para cada oído, para conseguir que la sensación de la fuente sonora que le llega al usuario simule perfectamente la posición espacial.

El siguiente paso es medir las funciones de transferencia  $Y_1/X_1$ , a las que denominaremos “funciones de transferencia altavoz-micrófono” ( $G_{LM}$ ). Esto lo haremos para todas las posiciones espaciales que nos interese y para los dos oídos a la vez:

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

A continuación se ha de medir a la vez, y para los dos oídos, las funciones de transferencia  $Y_2 / X_2$ , y las denominaremos “funciones de transferencia auricular-micrófono” ( $G_{HM}$ ):

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

Para conseguir los filtros digitales definitivos para simular sonidos procedentes de las posiciones del espacio que nos interesen (filtros totales), se debe multiplicar las funciones  $G_{LM}$  o altavoz-micrófono de cada oído, por la inversa de la función  $G_{HM}$  o auricular-micrófono del mismo oído. Esto se ha de hacer para ambos oídos. Por tanto, según ecuación (4):

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

Para la medida de las dos funciones de transferencia, altavoz-micrófono y auricular-micrófono, se utilizó la técnica de medida de respuestas impulsionales a través de secuencias de máxima longitud (MLBS), correlacionando de forma cruzada la salida del sistema y la entrada en secuencia MLBS.

Por tanto, si sometemos al sistema con una secuencia de máxima longitud, que llamamos  $s(n)$ , y se mide el resultado de la señal de salida  $y(n)$  durante el intervalo de tiempo que dura la MLBS, la respuesta impulsional  $h(n)$  sería:

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

donde  $\Phi$  es la correlación cruzada circular o periódica y  $h(n)$  es una respuesta impulsional periódica, contaminada con aliasing<sup>2</sup> temporal y no pura. Solo

---

<sup>2</sup> Aliasing, es el efecto por el cual señales continuas distintas, a través del tiempo, se convierten en señales indistinguibles al muestrearlas digitalmente.

cuando se escoja una secuencia temporalmente larga se podrá desprejar el aliasing.

Para implementar la ecuación (8) en secuencias largas se necesitaría mucho tiempo de cálculo. Por ello, se utilizaron equivalencias entre correlación y convolución, pasando al dominio frecuencial. La operación de convolución se traduce en una multiplicación de vectores, y su equivalencia queda así:

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

A pesar de esto, el tiempo de cálculo es excesivo por las FFTs (Fast Fourier Transform) utilizadas. Por ello, y finalmente, se redujo el tiempo de cálculo usando la FHT (Fast Hadamard Transform), y la ecuación que definitivamente queda para calcular 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 el sistema se estimuló con una secuencia de máxima longitud  $L$ . Las matrices  $P$ , son matrices de permutación y las  $S$  matrices de redimensionamiento.  $H_{L+1}$  es la matriz de Hadamard<sup>3</sup> de orden  $L+1$ .

### 2.3.3.2 Señal de sonido utilizada

Una vez generados los sonidos HRTF tal como se describe en el apartado 2.3.3.1, es necesario aplicar una convolución con un sonido externo Delta con el fin de obtener un banco de sonidos definitivo y óptimo que pueda ser usado por la persona invidente. En estos casos, las convoluciones son necesarias

---

<sup>3</sup> Matriz de Hadamard o matriz de diagonal estrictamente dominante. Son aquellas matrices regulares cuyo determinante es no nulo.

para poder dar al sonido resultante sensación de espacio, es decir, poder espacializar el sonido final y dar sensación, no de escuchar, sino de sentir el mundo real y lo que el usuario tiene delante.

Para realizar la convolución se utilizó la transformada de Fourier discreta. El resultado final fueron 1024 sonidos con las siguientes características (ver tabla 2.1):

<b>PARAMETRO</b>	<b>VALORES</b>
Dimensiones Azimut del campo de visión del dispositivo de audio (PFAI – Perceptual Field of the Audio Interface)	60° (programable hasta 360°)
Distancia del PFAI	Entre 0,5 y 5 metros (Programable a valores entre 0,25 y 6 metros)
Mínimo intervalo de tiempo entre interclicks	8 ms (programable a 7 ms)
Número de pixeles sonificados por imagen	64 (programable entre 1 y 64)
Frecuencia de refresco	2 frames/s (programable hasta 8,77 frames/s)
Máxima longitud de los sonidos	2048 samples (programable hasta 16.384 samples)
Frecuencia de Sampling	44.100 Hz
Tamaño de la colección de sonidos (N° Azimut * N° distancia * longitud del sonido * 2 oídos * n° byte/simple)	Hasta 16 MB
Generación de las secuencias de sonido	Aleatoria
Control del nivel de sonido del sistema de audio	8 niveles

Tabla 2.1 Características de los sonidos

La correspondencia entre los puntos medidos y la distancia fue (ver tabla 2.2):

<b>PUNTO EN PROFUNDIDAD</b>	<b>DISTANCIA AL USUARIO EN CM</b>
1	0,500
2	0,583
3	0,680
4	0,792
5	0,924
6	1,077
7	1,256
8	1,464
9	1,707
10	1,991
11	2,321
12	2,706
13	3,155
14	3,678
15	4,288
16	5,000

Tabla 2.2 Correspondencia de puntos

El total de sonidos obtenidos en la base de datos es de 64 (puntos de laser) \* 16 distancias o puntos en profundidad = 1024 sonidos.

### **2.3.3.3 Estrategia de sonificación**

Un objeto en movimiento se representó como un vector continuo de un sonido de muy corta duración, muy similar a la que se utiliza para representar la información con láser, pero con un timbre diferente, construyendo un efecto de manera que al usuario le dé la sensación de “objeto en movimiento”. Concretamente, este sonido fue de 250 ms de longitud.

Un ejemplo del aspecto de la señal, tanto en el dominio del tiempo, como en el espectro de frecuencia, se puede ver en la figura 2.18.

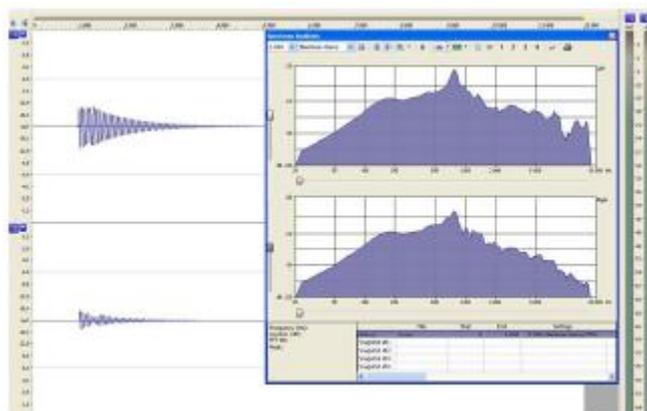


Figura 2.18 Sonido utilizado en el estrategia

Tres fueron los parámetros que se sonificaron [Dunai L. et al., 2009] por cada trama de imagen obtenida: la **posición 3D** del objeto, la **dirección del objeto** en relación al usuario y la **velocidad** respecto al usuario.

Las posiciones espaciales de los objetos se representaron por las coordenadas X (azimut) e Y (elevación), correspondientes con el centro de sus “bounding boxes”<sup>4</sup>. La distancia del objeto, respecto al usuario, fue representada por la coordenada Z.

La **posiciones espaciales 3D** de los objetos fueron representadas por sonidos que identificaban tales posiciones. Además, el tono del sonido es cambiante de forma progresiva en función de la distancia. Tras las encuestas pasadas a las personas ciegas, y después de los experimentos realizados, se llegó a la conclusión de que la información espacial auditiva que de forma coherente se podía interpretar estaba entre 0,5 y 15 metros. Más allá de 15 m, el objeto se representó por un sonido espacial equivalente a la distancia mayor del rango mencionado, es decir, 15 metros. Con esto, se pretendía que el usuario detectara la presencia de objetos en movimiento.

La **dirección del objeto** en relación al usuario, es decir, si el objeto se mueve o no respecto al usuario y en qué proporción lo hace, se codificó con la frecuencia del tren de sonidos cortos. La duración del interclick variaba desde

---

<sup>4</sup> Bounding Box representa el rectángulo más pequeño y comprimido que incluye al objeto dentro de la imagen.

los 100 milisegundos, cuando el objeto se alejaba del usuario, hasta los 25 milisegundos cuando el objeto se acercaba al usuario. El ICI<sup>5</sup> es una característica que indica si existe, o no, riesgo de colisión o peligro para el usuario.

La **velocidad** del objeto se codificó por un cambio en la velocidad del tono al pasar el objeto de una distancia a otra respecto a la posición de referencia del usuario.

También se diseñó una estrategia para la sonificación de los **pasillos libres**. Esta información representa aquellas áreas libres que el usuario ha de interpretar como zonas sin riesgo y a través de las cuales se puede navegar sin obstáculos. La información necesaria para representar el pasillo libre es la posición azimuth de los bordes del camino. Estas posiciones están representadas con una continuación aleatoria de dos sonidos espaciales que coinciden con los límites horizontales de la ruta a representar. Este sonido es de 250 ms de largo y tiene un timbre particular, que es claramente diferenciable y reconocido por el usuario como “borde de un camino libre”. Su gráfica se representa en la figura 2.19.

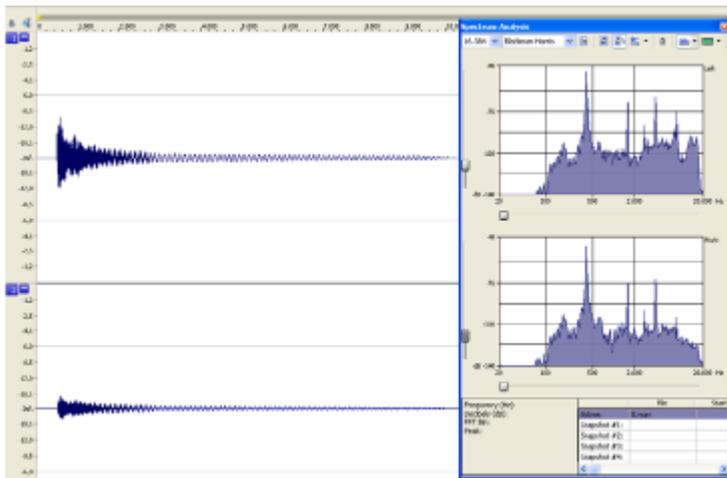


Figura 2.19 Representación acústica de vías libres

<sup>5</sup> ICI: Interclick Interval (Intervalo de tiempo entre dos sonidos, o clics, consecutivos).

Para la implementación y diseño de los estrategias de sonificación, una serie de estudios se llevaron a cabo y fueron fundamentales para la definición de esta representación auditiva, [Mora J. L. et al, 2006]. Algunas de ellas han arrojado resultados muy interesantes, tanto desde el psicofísico, como desde el punto de vista perceptivo y psicoacústico. Varios diseños experimentales se llevaron a cabo para demostrarlo:

- Simultanear las condiciones dinámicas de aprendizaje y azimut con la percepción de la distancia de una imagen auditiva con HRTFs generales y no individuales.
- Para los estudios sobre la resolución espacial en el azimut y la distancia, se aumentó la resolución espacial en torno a 1° sobre el azimut además de variar los centímetros en distancias cortas. Estudios preliminares demostraron que una resolución espacial de unos 3°- 4° en azimut puede ser representado y resuelto con HRTFs no individuales.
- Se trabajó con la variación del ICI. Este parámetro tiene consecuencias importantes en la imagen auditiva espacial, ya que parece tener un efecto importante sobre la solidez de la imagen en lo que se refiere a percibir el objeto como un todo, y no como sonidos individuales.
- Por otro lado, el efecto de la variabilidad del ICI afecta claramente la sensación de muy agradable o desagradable que el usuario percibe, de tal manera que en los sonidos con ICI muy bajo se mezcla el ruido de fondo con la información espacial.
- Sobre el efecto de reverberación de la imagen espacial auditiva, cuando la reverberación es baja parece que afecta a la percepción espacial de la distancia en condiciones estáticas. Sin embargo, este efecto se invirtió en condiciones dinámicas.
- El uso de diferentes tonos para diferentes objetos, ayuda a diferenciar mejor dichos objetos.
- El tipo de sonido jugó un papel importante en la percepción de la distancia y la ubicación lateral, es decir, se demostró que este parámetro afecta a la distancia en lo que se refiere a la percepción de las fuentes del sonido.
- Los cambios de tono también afectan en la percepción final de la distancia de los objetos.
- Estudios referentes al efecto aprendizaje en la percepción de la elevación de la fuente de sonido sugieren que un corto periodo de

entrenamiento puede mejorar el rendimiento tanto en condiciones estáticas como dinámicas.

## **2.4 Conclusiones**

En este capítulo se han comentado las diferentes tecnologías aplicadas para el estudio de los objetivos de esta tesis, además de cómo nació la idea de diseñar una herramienta para ayudar a las personas invidentes y cuales han sido los objetivos, tanto tecnológicos como científicos del proyecto.

También se ha creído conveniente describir a los socios tecnológicos del proyecto y la aportación general que entre todos ellos se ha llevado a cabo para la consecución de los objetivos.

Se ha visto, que el dispositivo final está compuesto por dos interfaces de captura de datos, una mediante sensores CMOS y otra por estereovisión. La información capturada por estos dos sistemas referente al mundo 3D por el que nos movemos, es enviada simultáneamente a un PC y una FPGA para ser procesados y, posteriormente, transformarla en señales acústicas. El sistema CMOS compuesto por un láser, trabaja entre 0,5 y 5 metros, mientras que el sistema de visión opera en un rango de entre 5 y 15 metros.

El primer sistema es capaz de detectar obstáculos tanto estáticos como en movimiento, y lo hace en un azimut de 60°. El sistema de visión complementa al anterior hasta 15 metros, siendo capaz de detectar pasillos libres y objetos estáticos y en movimiento.

Cuando la información es capturada por el sistema, los datos del láser son enviados al PC para ser procesados e interpretados, sacando resultados mediante algoritmos de medida y efectos rebote (Time of Flight) para obtener distancias. Por otra parte y en paralelo, los datos procedentes del sistema de visión son procesados a través de algoritmos de filtrado de imagen para obtener mapas de profundidad y a partir de ellos aplicar criterios de escala de grises, tracking [Vadakkepat P. et al, 2008] y segmentación y conocer las distancias de los objetos y pasillos libres a sonificar.

Los mapas de sonidos se obtienen partiendo de modelos HRTF y convolucionando con sonidos acústicos para mejorar la calidad de los mismos y dar mayor sensación de bienestar y entendimiento de la señal. Los HRTFs representan la Transformada Fourier de la Head Related Impulse Respons

(HRIR), es decir, representa los sonidos creados mediante coordenadas espaciales en ambos oídos, el derecho y el izquierdo.

La conclusión a la que se ha llegado es que mezclando diferentes tecnologías de diferentes campos se puede conseguir un sistema de sonificación [Dunai L. et al., 2009], [Dunai I. et al., 2009] que pueda ayudar a entender el espacio a las personas invidentes, y que se ha realizado una importante labor y un trabajo exhaustivo para conseguir generar un prototipo operativo partiendo de una teoría de investigación.

Finalmente, queda implícito que lo mejor para poder validar los resultados de una investigación, pasa por la construcción de dispositivos reales que nos permitan llevar a cabo experimentos, en nuestro caso con personas invidentes, y demostrar su eficacia. En este caso, se ha demostrado que la tecnología y la ingeniería pueden mejorar la calidad de vida de las personas.



## **3. INTERFAZ DE COMUNICACIÓN E INTEGRACIÓN**

### **2.1 Resumen**

### **2.2 Integración de los módulos**

### **2.3 Conclusiones**

#### **3.1 *Resumen***

Por último, y como prototipo final resultado de toda la investigación llevada a cabo durante este proyecto, se ha llegado a integrar y fabricar un prototipo formado por la unión de los dos prototipos, M1 y M2.

El trabajo de integración se ha llevado a cabo coordinando todas las variables procedentes de los subsistemas periféricos, hasta llegar a diseñar un sistema capaz de ser útil para ser evaluado por personas invidentes totales, independientemente de su edad y de la causa de la enfermedad. Los resultados obtenidos se comentan en el punto 4 (Validación) de la presente tesis.

El diseño de la integración fue dividida en dos partes; la integración desde el punto de vista físico (hardware) y la integración desde el punto de vista lógico (software). Desde el punto de vista físico, se diseñó un sistema que cumple con los requerimientos básicos de tamaño, peso y ergonomía. El resultado final se comenta en el apartado 3.1. Desde el punto de vista lógico, se implementó una aplicación para unir todas las variables que intervienen en el proceso y poder, a partir de toda la información real capturada, convertirla en un mapa de sonidos acústicos para ayudar a las personas invidentes a interpretar el espacio y poder navegar de manera más segura y autónoma.

## 3.2 Integración de los módulos

La figura 3.1 representa un esquema con todos los componentes que componen el prototipo desarrollado, así como las conexiones entre ellos.

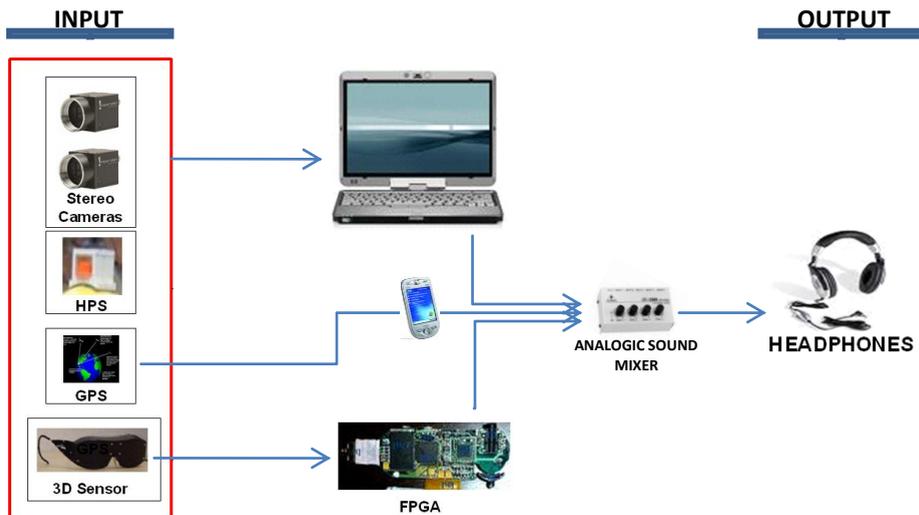


Figura 3.1 Integración física del prototipo definitivo

Como se observa, se ha centralizado todo el proceso en un PC portátil con características estándares que actúa como central de proceso. El sistema está formado por: 1 ordenador portátil, 2 cámaras de visión, 1 PDA, 1 sensor 3D CMOS, 1 FPGA, 1 mezclador analógico y auriculares.

Como entradas al sistema tenemos las 2 cámaras de estereo visión, encargadas de capturar las imágenes responsables del prototipo M2. Estas imágenes, en tiempo real son transmitidas al PC, el cual, por cada trama capturada de la imagen real extrae los mapas de profundidad, los procesa para extraer los objetos finales y aplica los algoritmos necesarios para extraer los pasillos libres. Todos estos datos son almacenados y, mediante protocolos de comunicación son enviados al módulo de sonido para, finalmente ser sonificada por el estratega de

sonido y enviada a los auriculares a través de la tarjeta de sonido del PC. Al mismo tiempo, y en paralelo con la captura de imágenes, existe un sistema GPS integrado en una PDA la cual funciona de manera independiente, y conectada al mezclador emite la información sonora necesaria para guiar al usuario a su destino final.

El sensor 3D CMOS captura información sobre los objetos próximos y la envía a la FPGA, la cual conectada al banco de sonidos del PC genera información sonora para ser enviada finalmente a los usuarios a través de los auriculares.

Se puede observar que la información que llega al usuario final es la mezcla de M1, M2 y GPS, siendo posible discriminar cualquiera de las tres mediante el mezclador analógico integrado en el sistema.

Resumiendo, el prototipo que actualmente existe operativo dispone de las siguientes características:

- De 0 a 15 metros el sistema es capaz de capturar información real y transformarla a mapas acústicos de un objeto simultáneamente.
- De 5 a 15 metros el sistema puede interpretar información de pasillos libres.
- Toda esta información es sonificada y enviada a los auriculares que finalmente se coloca el usuario.
- Para instalar y ejecutar todo el proceso sólo es necesario un PC portátil y una PDA.



Figura 3.2 Unión de M1 y M2

Con el fin de verificar los resultados de la investigación e integración del prototipo final, se reunió a un grupo de usuarios finales y se realizaron diferentes ejercicios.

Las pruebas básicas de aprendizaje efectuadas basadas en BLP (Basic Learning Protocol) con usuarios finales respecto al prototipo M1, estuvieron preparadas para trabajar con siete ejercicios diferentes y con el fin de localizar distintos tipos de objetos para medir sus anchuras y alturas correctamente mientras caminaban a través de ellos sin golpearlos.

Todos los sujetos fueron capaces de identificar correctamente las dimensiones de los objetos en un corto espacio de tiempo y no tuvieron excesiva dificultad para discernir entre la altura y la anchura de los objetos, incluso en dar la vuelta y seguir el camino andado en orden inverso.

Sí que hubo diferencias sin embargo en lo que se refiere al tiempo. El mínimo estuvo en 3:53 minutos y el máximo en 10:35 minutos. Esto puede estar vinculado al nivel de experiencia y de entrenamiento que se reciba, ya que el mejor resultado obtenido correspondía con el más experimentado. Todas estas pruebas obtenidas gracias a la integración de este sistema son comentadas con más detalle en el punto 4 de la tesis.

Respecto a las pruebas del prototipo M2 y siguiendo las normas de la percepción angular, 60 grados de ancho y 20 m de largo, se realizaron diferentes ejercicios basados en 18 posiciones diferentes. Los sujetos fueron situados en el vértice del ángulo de percepción y se les pidió identificar los cambios en las posiciones que una persona realizaba en el campo de visión. El orden de los cambios que se iban produciendo no estaba pactado y se iba realizando al azar.

Antes de realizar las pruebas referentes a M2, se planificó y se explicó con detenimiento la manera en que se iba a trabajar. Las pruebas estaban basadas en posicionar al usuario en una posición fija y sin mover la cabeza. Los resultados fueron bastante prometedores. No se cometieron errores y todos los sujetos fueron capaces de identificar rápidamente y con facilidad la evolución de las posiciones de las personas involucradas en el test, además de indicar los cambios en los movimientos que se estaban realizando.

Sin embargo, fue bastante evidente que los resultados no hubieran sido los mismos en una situación dinámica con el usuario en movimiento o con varias personas participando en el plano de visión.

Los módulos integrados en el proyecto corresponden a todos los algoritmos implementados encargados de realizar la captura de datos del mundo real (3D) desde el prototipo M1 y M2.

En este punto de la presente tesis, se explica que una de las partes más importantes para conseguir demostrar que sistemas compuestos por estereovisión y audio facilitan la vida a las personas invidentes, es el poder haber generado protocolos de comunicación entre los diferentes investigadores que han colaborado en el proyecto, para obtener un prototipo final y tangible que sea capaz de demostrar la eficacia y la funcionalidad del trabajo que se ha llevado a cabo académicamente desde laboratorio. Además del trabajo técnico y de investigación, ha habido un importante esfuerzo en coordinar a todas las partes involucradas en el proyecto para llegar a conseguir la integración de la plataforma definitiva.

Los diferentes módulos que intervienen en la plataforma proceden de: Información del HPS (Head Position System) y de GPS desarrollados por la Universidad Politécnica delle Marche, los algoritmos de procesado de imagen para la detección de objetos estáticos y en movimiento desarrollados por la Universidad de Bristol, y los bancos de sonidos generados por investigadores de la Universidad de La Laguna. Su intercambio de datos y el flujo de comunicación se muestran en la figura 3.3.

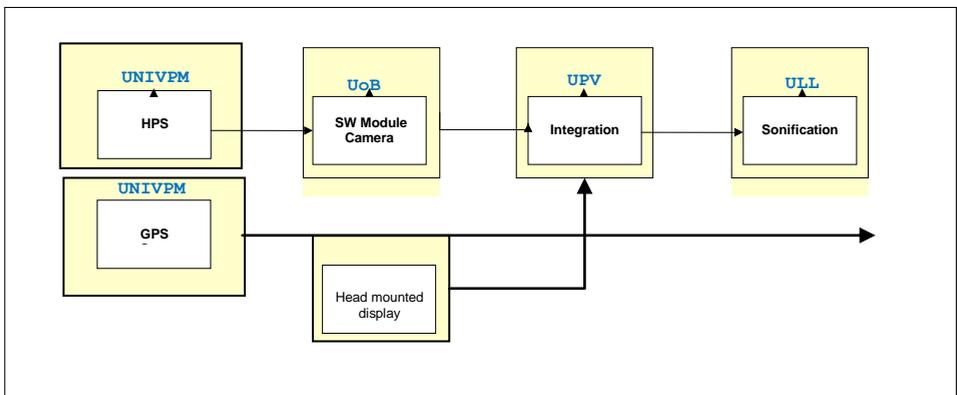


Figura 3.3 Flujo de comunicación entre módulos

Como se puede observar en la figura 3.3, para la recogida global de datos, han intervenido herramientas desarrolladas por diferentes equipos de investigación.

La investigación y desarrollo de la Università Politecnica delle Marche proporciona datos referentes al sistema de posicionamiento de cabeza (HPS). Estos datos, han sido necesarios para conocer la orientación de la cabeza del usuario y la aceleración, con el fin de enviárselos directamente al sistema de estéreo visión con el objetivo de disponer de referencias y poder calcular la velocidad de los objetos en movimiento, así como su posición y calibración.

La Universidad de Bristol captura la información procedente del sistema de posicionamiento HPS y la integra en sus algoritmos de proceso de imagen para generar mapas de profundidad y extraer de ellos información referente a objetos, tanto estáticos como en movimiento, para distancias superiores a 5 metros.

Al mismo tiempo, los mapas de profundidad obtenidos como consecuencia del filtrado de las imágenes procedentes de las cámaras de estéreo visión, es enviada al siguiente módulo, desarrollado por el Centro de Investigación en Tecnologías Gráficas, para la extracción, mediante técnicas de procesado de imagen, de pasillos y vías libres.

Finalmente, toda la información generada en los módulos anteriores, y antes de ser enviada al módulo de sonificación, es encapsulada y empaquetada para enviarla a los algoritmos de sonido generados por la Universidad de La Laguna.

La secuencia de pasos definitiva que el sistema sigue para la generación de los mapas acústicos es:

1. Captura de imágenes en tiempo real de las cámaras de estéreo visión.
2. Análisis de dichas imágenes y creación de mapas de profundidad.
3. Obtención de información acerca de objetos de los mapas de profundidad.
4. Obtención de información sobre espacios libres de los mapas de profundidad.
5. Integración de toda esta información.
6. Envío al módulo de sonificación.

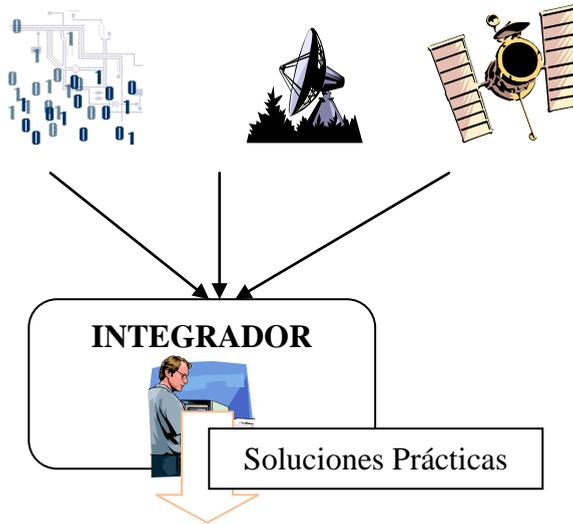
El prototipo se ha diseñado para ser ejecutado en Windows XP, y todo el software encargado de integrar las variables y datos del total del proyecto se ha desarrollado utilizando Microsoft Visual C++ 2005 y Microsoft Visual Basic 2005 de Visual Studio 2005.

### **3.3 Conclusiones**

Como resultado final de la unión de todas las variables procedentes de los subsistemas del prototipo construido, se ha conseguido crear una plataforma sencilla, pero robusta y eficaz para unir información y conseguir que lo que los usuarios finales reciben sea información coherente acerca de lo que se necesita que perciban. Desde el punto de vista humano, como se verá en el punto 4 de esta tesis, los resultados han sido especialmente satisfactorios muy halagüeños como punto de partida a futuras investigaciones, e incluso sistemas definitivos y no prototipos, pero desde el punto de vista tecnológico, se ha demostrado una vez más la complejidad de lo que conlleva un trabajo de integración entre diferentes tecnologías. Si se me permite un comentario personal tras mi dilatada experiencia profesional en proyectos de desarrollo, comunicaciones, control de procesos y de investigación, siempre he pensado que el integrador es la persona que más difícil lo tiene, ya que es el responsable de la planificación, la especificación de interfaces y el seguimiento del proyecto entre otras cosas. Este comentario lo justifico diciendo que, mientras cada desarrollador y analista se centra única y exclusivamente en su parte de trabajo y lo ve como una caja de entradas y salidas autónoma la cual tiene que hacer funcionar, un integrador de varias tecnologías o proyectos ha de conseguir que todas las cajas procedentes del resto de desarrolladores e investigadores se unan y generen una caja global con unas entradas y salidas que funciones, y le hagan totalmente transparente al usuario final lo que el sistema haga internamente. Además, el integrador ha de conseguir otros objetivos:

1. Que el sistema sea robusto y no se bloquee ni deje de funcionar
2. Conseguir equilibrar la funcionalidad con la ergonomía
3. Minimizar el tamaño y el peso
4. Asegurar que el uso sea lo más sencillo e intuitivo posible
5. Pensar siempre en el punto de vista de los usuarios
6. Cubrir la mayor parte de las necesidades del público al que va dirigido el producto final

Por a estos requerimientos, tanto humanos, tecnológicos y funcionales, vuelvo una vez más a justificar la importante tarea de los integradores y la capacidad que hay que tener para mezclar y resolver distintos tipos de problemas, y sobre todo en ser capaces de adelantarse a lo que pueda ocurrir y pensar en todas las posibilidades, sobre todo centrándose en el caso peor.



## 4. VALIDACIÓN

### 4.1 *Resumen*

En este apartado se describe las estrategias experimentales, pruebas y resultados obtenidos tras la validación llevada a cabo en usuarios invidentes. En la primera parte del capítulo se analiza el test llevado a cabo para detectar y localizar objetos con el fin de validar el sistema. La percepción del sonido externalizado por el usuario, la navegación y los errores en la detección de los objetos también es descrita en el presente capítulo.

Estas pruebas experimentales se han podido llevar a cabo gracias a la integración, la tecnología empleada y el prototipo fabricado según se explica en los capítulos anteriores. Las señales espaciales son generadas para ofrecer al usuario ciego una herramienta cognitiva para que pueda interpretar y crear su propio mundo cuando anda a través de un laberinto de obstáculos fabricado para la ocasión. La información ofrecida por el sistema indica en tiempo real la localización precisa de los obstáculos, tanto en distancia como en azimut, traduciéndose en mapas acústicos.

Gracias al sensor de precisión y sus características, los resultados obtenidos fueron bastante significativos, y pueden ser muy relevantes para tenerlos como base a la hora de testear otros sistemas de navegación con distintas tecnologías en un futuro próximo. Tras las primeras pruebas, se mejoró el sistema acústico en lo que se refiere a parámetros de la señal y el banco de sonidos, concretamente decrementando el interclick entre sonidos y parametrizando las propiedades base de los sonidos.

Los experimentos se basaron en tres test distintos: el test de BLP (Basic Learning Protocol), durante el cual los participantes se familiarizaron con el sistema, las señales espaciales, la externalización del sonido y el reconocimiento de objetos, navegando a través del laberinto de pruebas. Los usuarios obtuvieron resultados excelentes, tanto en tiempo de recorrido, como en número de errores. Un tiempo mínimo de 41 minutos fue necesario para completar todo el recorrido en las pruebas de laboratorio, y el número de errores y tropiezos fue disminuyendo hasta 0 la segunda vez que se recorrió de nuevo el escenario.

Tras la introducción del usuario en el manejo del sistema, se llevaron a cabo las pruebas del laberinto y de movilidad, las cuales adquirieron un mayor nivel de complejidad debido a la interacción con el mundo real.

Los resultados de estos experimentos basados en señales acústicas y en laser con rebote en tiempo de vuelo de la señal, demostraron que el sistema es capaz de ofrecer información real de la localización de los obstáculos tanto, como he comentado antes, en lo que se refiere a distancia y azimut. Todos los resultados tuvieron implicación humana y, los test de navegación estuvieron guiados constantemente por personas videntes.

Además del sistema de adquisición de datos y del entorno, se comprobó que un importante factor para las personas ciegas es la representación de la información adquirida, por ejemplo, como informar al ciego de que el sistema ha detectado un objeto que está en frente de él, aun lado, o al otro.

## ***4.2 Test de navegación***

### **4.2.1 Diseño del experimento**

Las fases llevadas a cabo con las personas invidentes para llegar a las conclusiones finales se pueden clasificar en tres tipos:

- a) Basic Learning Protocol (BPL)
- b) Prueba del laberinto
- c) Test de movilidad

#### **a) Basic Learning Protocol (BPL):**

Diferentes experimentos fueron desarrollados para comprobar las características del sistema. Estas pruebas fueron llevadas a cabo sobre la base de un protocolo de aprendizaje. El entorno de pruebas en interior, fue diseñado con una distancia de 115 metros de longitud y 10 metros de ancho. Las pruebas en exterior se llevaron a cabo en un parque abierto. El intervalo de tiempo entre los sonidos fue de 153 ms.

A continuación se describen los ejercicios llevados a cabo para la demostración de los experimentos.

### **Ejercicio 1: Detección de un obstáculo**

En este ejercicio, el usuario se colocó en frente de un objeto estático (puede ser una persona o un objeto similar con un diámetro alrededor de 30cm) a una distancia de 1 metro. Inicialmente, el punto de visión del usuario esta desviado hacia uno de los lados, de manera que la intención sea que los ojos del usuario no estén apuntando a ningún objeto, en este caso el sistema no emite ningún sonido. A continuación se gira la cabeza del usuario hacia el objeto y comenzará escuchar unos sonidos a través de los auriculares.

Para conseguir el efecto de externalizar el sonido, es decir, que el sonido provenga del objeto, el usuario debe mover la cabeza hacia la derecha o a la izquierda muy despacio hasta que el sonido desaparezca, y volverlo a hacer en sentido contrario alternativamente. A continuación, cuando la cara del usuario apunta directamente al objeto, debe acercarse y alejarse del mismo. Entonces, explicamos al usuario que el sonido aparece solo cuando en el área de visión del sistema hay un objeto, y el sonido procede exactamente de donde se encuentra el objeto. Ver figura 4.1.

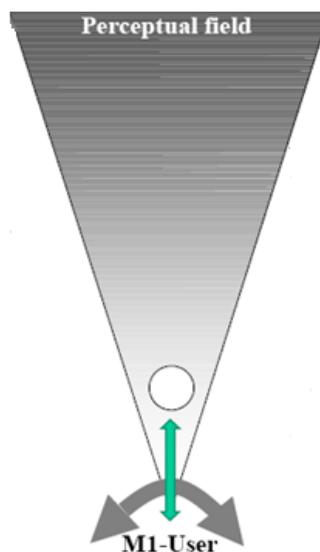


Figura 4.1 Ejercicio con una columna estática

## Ejercicio 2: Detección de dos obstáculos

En este caso, se colocó al usuario en frente de dos objetos estáticos (pueden ser personas u objetos similares con un diámetro alrededor de 30cm y posicionados a ambos lados del campo visual), a una distancia de 2 metros.

Se le quita el volumen al sistema y el usuario es colocado en frente de las columnas. Se activa el sistema y el usuario empieza a percibir el sonido a través de los altavoces. Esta vez, para conseguir la externalización del sonido, es decir, dar la sensación de que el sonido procede del objeto, se indicó al usuario que moviera la cabeza despacio de derecha a izquierda. Si no se consigue que la externalización sea sentida por el usuario, el ejercicio se ha de repetir cambiando de posición las columnas, llevándolas hacia delante o detrás tal como se muestra con las flechas verdes en la figura 4.2.

Se debe continuar con el ejercicio hasta que el usuario detecte y diferencie las dos columnas. Una vez se consiguió, el ejercicio se dio por válido.

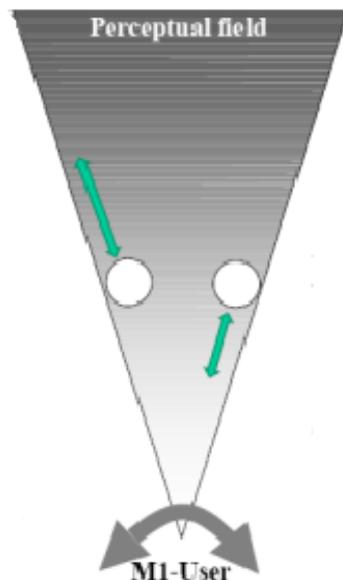


Figura 4.2 Ejercicio con dos columnas estáticas

### **Ejercicio 3: Dos columnas juntas**

En nuestro tercer caso, se colocó al usuario en frente de dos objetos estáticos juntos, es decir, uno al lado de otro estáticos (pueden ser dos personas o un objeto similar con un diámetro alrededor de 30 cm y posicionados juntos en el dentro del campo visual).

Inicialmente, el sistema esta desconectado y se colocó al usuario hacia las columnas. A continuación se conecta el sistema.

Asumiendo que la externalización del sonido ya es sentida por el usuario (sino los ejercicios 1 y 2 han de ser repetidos de nuevo), cada pocos minutos se ha de mover uno de los objetos de sitio tal como indica la flecha de la figura 4.3. El fin de esta prueba es ir desplazando una de las columnas para que el usuario pueda sentir que el sonido procede de un área más ancha o más estrecha (según pongamos las dos columnas juntas o las separemos).

La prueba continuó hasta que el usuario fue capaz de diferenciar uno o dos objetos señalándolos con sus manos.

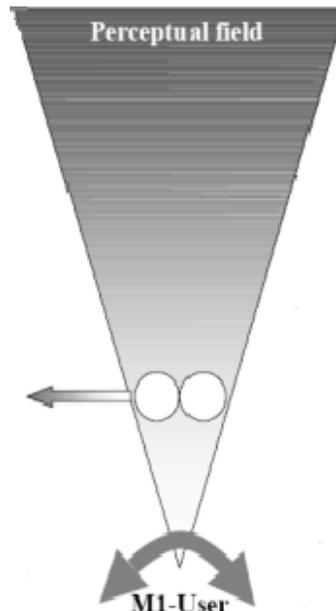


Figura 4.3 Ejercicio con dos columnas juntas

#### **Ejercicio 4: Superficie estática**

A continuación se realizaron ejercicios con objetos estáticos pero de mayor superficie, tales como un muro, una mesa grande de 2 metros de longitud, etc. y situados a una distancia de 1,5 metros, más o menos, del usuario. Ver figura 4.4.

Como en las pruebas anteriores, inicialmente el sistema está desconectado y en silencio, y orientamos al usuario hacia la superficie. Se le explicó al usuario que el sonido ha de sentirse como si viniera de una pared y se conectó el sistema.

Dado por hecho, como en la prueba 3, que el sonido se percibe de manera externalizada, sino habría que repetir los ejercicios 1 y 2, en este caso se ha de continuar hasta que el usuario nos informa correctamente de la anchura de la superficie que tiene delante, señalando los bordes con sus manos. Una vez conseguido, el ejercicio finalizó.

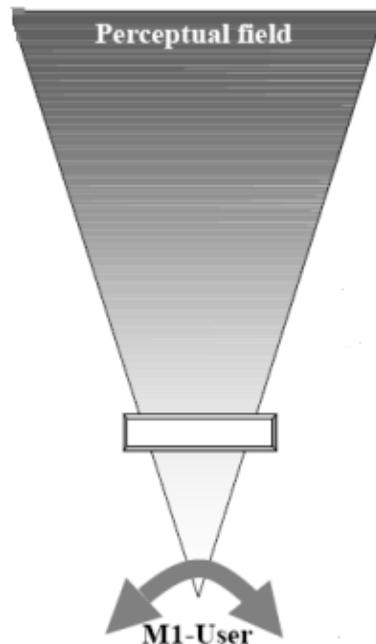


Figura 4.4 Ejercicio con superficie estática

### **Ejercicio 5: Pasillo libre (puerta abierta)**

En este quinto ejercicio, se colocó al usuario invidente enfrente de un muro (pared o similar) con un hueco o puerta abierta de 80 cm de ancho más o menos y situada a una distancia de 1,5 metros. Inicialmente el sistema está en silencio y, cuando el usuario es posicionado hacia la superficie es cuando se conecta el dispositivo.

Como en los casos anteriores y suponiendo que se siente correctamente la externalización, se le indica al usuario que mueva la cabeza lentamente de izquierda a derecha y de forma repetitiva. Al mismo tiempo, cuando se posicione enfrente del hueco el usuario debe andar hacia el mismo. Nosotros le explicamos que el sonido ha de sentirse como si viniera de un marco circundando una puerta que se encuentra enfrente de él.

El ejercicio ha de continuar hasta que el usuario nos indique correctamente la anchura del hueco del muro, haciéndonoslo saber indicando los bordes con sus brazos (figura 4.5).

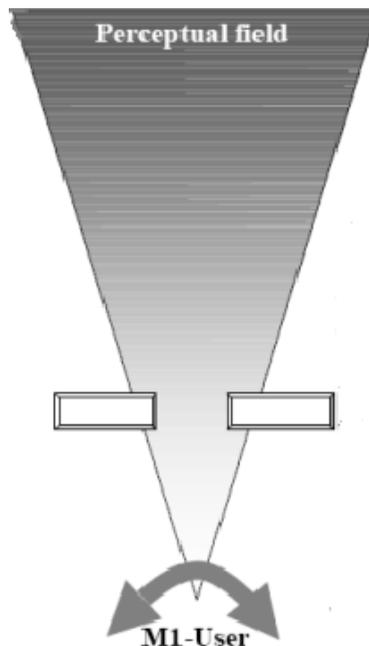


Figura 4.5 Ejercicio con hueco

### **Ejercicio 6: Superficie detrás de un objeto**

En nuestra prueba número 6 el usuario fue colocado enfrente de un muro a una distancia entre 1,5 y 2 metros.

El dispositivo estuvo en silencio hasta que se orienta el punto de visión del usuario hacia la superficie. Es entonces cuando se conecta el sistema, y tras una espera de pocos minutos, se colocó una columna enfrente del usuario a una distancia de 1 o 1,5 metros, manteniendo la superficie detrás de la columna.

Como en todos los ejercicios, y asumiendo la percepción de la externalización del sonido, el objetivo de este experimento fue detectar la presencia del objeto delante de una pared o similar.

La prueba termina cuando el usuario siente la columna sonar cerca de él, al mismo tiempo que tiene otra superficie enfrente. Ver figura 4.6.

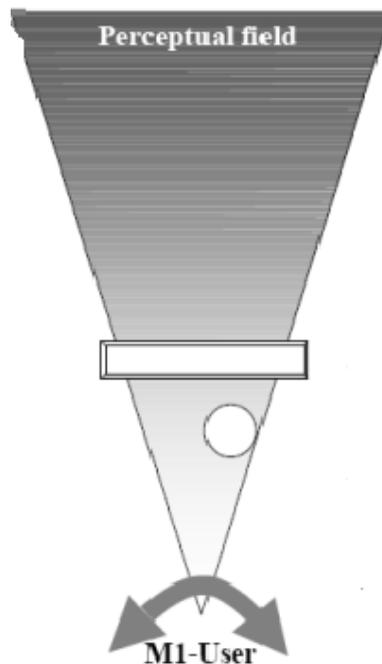


Figura 4.6 Ejercicio con una superficie detrás de una columna

### **Ejercicio 7: Movimientos hacia una superficie**

En este caso, el usuario fue colocado enfrente de una pared a una distancia de entre 2,5 y 3 metros.

Repitiendo los pasos anteriores, el dispositivo estaba desconectado hasta que el invidente fue orientado hacia la superficie a detectar. Después se conectó el sistema, y tras permitir al invidente sentir el objeto durante algunos segundos, comenzó a andar lentamente hacia la columna, volviendo hasta la posición inicial. Esto se repitió durante varias veces.

Como en los pasos anteriores, y estando seguros de que el efecto de externalización era el adecuado, se obtuvo el objetivo deseado de este experimento, que era detectar la distancia y el ancho de la superficie que estaba enfrente mientras el usuario está en movimiento. No solo cambios en la intensidad del movimiento, sino sensaciones de espacio. Cuando estas condiciones se cumplían se daba por finalizada la prueba (figura 4.7).

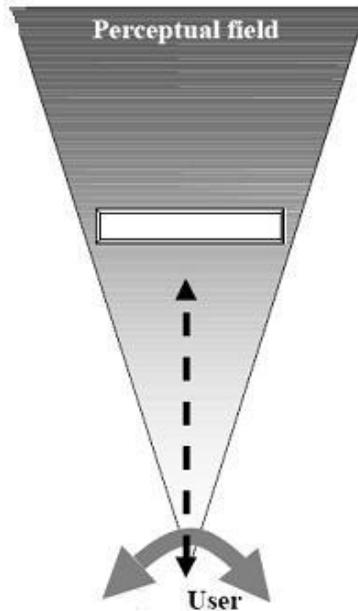


Figura 4.7 Ejercicio en movimiento

Tras realizar todas las pruebas con el BPL, se abrió una fase para que los usuarios finales preguntaran dudas, así como para aclarar cualquier comentario que necesitaran los participantes en las pruebas.

Todos los parámetros referentes al tiempo de externalización del sonido, la detección de los objetos y los errores en distancia y azimut fueron registrados.

### **b) Prueba del laberinto:**

Después, cuando los participantes adquirieron más experiencia y destreza en los siete ejercicios, se pasó a la prueba más importante del experimento: la localización de obstáculos y la navegación simultánea, a través del laberinto diseñado para los experimentos en un test de laboratorio para andar a través de pasillos libres.

Los participantes fueron guiados hasta un espacio diáfano y libre donde no había obstáculos a la vista en un área de 15 metros cuadrados. Tras conectar y preparar el sistema, el usuario se colocó en el escenario. Se les pidió a los participantes que anduvieran a través de un laberinto de unos 14 metros de largo con 4 pares de obstáculos sin riesgo de unos 180 centímetros de altura. Las columnas fueron colocadas con una separación horizontal de 2 metros y una distancia de 2,5 metros entre el último par de columnas y el muro (ver figura 4.8).

Dos rutas fueron diseñadas para realizar el experimento. La primera tenía 4 puntos de inflexión, incluyendo dos de 20° a la izquierda y a la derecha y dos de 40° a derecha e izquierda. La segunda ruta tenía 6 puntos de inflexión, incluyendo dos de 25° de izquierda a derecha, otros dos de 40° de derecha a izquierda y, finalmente, otros dos de 45° de izquierda a derecha.

Lo primero que hizo el usuario fue analizar el contexto y adaptarse al mismo, detectando el primer par de columnas. Anduvo entre ellas hacia el siguiente paso, en el que era necesario detectar el siguiente par de columnas y pasar entre ellas, y así sucesivamente hasta finalizar con la ruta marcada por el laberinto.

Cuando un objeto era detectado por el sistema, a través de los auriculares los invidentes escuchaban sonidos espacializados, indicando la presencia de un objeto, permitiéndoles desviarse y pasar sin impactar. Los sonidos utilizados



obstáculos, sino el definir sus características. Un escenario de este tipo se puede ver en la figura 4.9.



Figura 4.9 Test de movilidad

Durante este test, se utilizaron los métodos comentados en este capítulo pero con un grado más de complejidad, ya que las tareas de los usuarios en este caso eran múltiples: localizar los objetos, comentar las diferencias entre los objetos percibidos, analizar el entorno y elegir el camino libre para caminar sin problemas. También se les preguntó sobre la sensación que tenían cuando detectaban objetos con sonidos diferentes, incluyendo sonidos procedentes del propio entorno.

En esta prueba de movilidad, primero se realizó el experimento en un escenario artificial. Esta prueba fue diseñada para aproximarse lo más fielmente posible al entorno real. Más tarde, se realizó la misma prueba en un entorno real, abierto y desconocido para los usuarios.

Cuatro escenarios diferentes se utilizaron para esta prueba. El grupo A, formado por ciegos de Alemania, utilizó dos escenarios externos. Los dos escenarios se diseñaron en el patio de la escuela de ciegos de Berlín (ABSV). Los ciegos tuvieron que navegar a través de 29 metros de longitud con una gran variedad de obstáculos en el camino.

El test arrancó con el usuario mirando hacia un espacio amplio donde no había objetos ni obstáculos en el punto de visión del invidente.

Al no aparecer objetos en el punto de visión del usuario, el sistema no envió ninguna señal acústica a través de los auriculares. Así, el usuario sabía que no existe objeto alguno en una distancia de 5 metros. Por otra parte, el usuario sabía que cuando escuche sonido, existirán objetos en su área de visión, pero quedaba bien claro que, cuando el sistema esté en silencio total, no había que preocuparse y creer que el sistema no funciona.

Para el siguiente grupo (grupo B) también se utilizaron dos escenarios diferentes, y las pruebas se llevaron a cabo en un patio diáfano y amplio. El escenario utilizado era muy parecido también al ilustrado en la figura 4.9.

Ambos grupos fueron evaluados durante dos sesiones cada uno y en días diferentes. La ubicación de los obstáculos artificiales fue modificada, manteniendo la estructura real de los obstáculos propios del contexto (columnas, bancos, muros, etc.). Todos los obstáculos artificiales eran seguros y no ofrecían peligro alguno para los participantes.

Durante el experimento, los usuarios tenían plena libertad para ajustar el volumen del sonido según su comodidad y necesidad, pero no se les dejaba desconectar el dispositivo. Además, tampoco se les daba instrucciones o pistas de donde se encontraban o de cuál era su posición, es decir, eran completamente autónomos. Sí que, cuando el usuario se encaminaba hacia una ruta o dirección errónea, se le redirigía por parte del entrenador a la posición buena.

Lo que era cierto es que los participantes tenían experiencia previa en el sistema (por los test realizados y comentados con anterioridad), pero que ellos eran los que tenían que interpretar y sentir el nuevo medio por el que se estaban moviendo. Por lo tanto, antes de comenzar con la prueba real, los usuarios estaban familiarizados con los estímulos anteriores y con la navegación. Se les pidió que prestaran especial atención a una serie de comentarios y características que se les iba diciendo verbalmente. Estos comentarios fueron referentes a:

- El sonido era externo, es decir, se les repitió que debían sentirlo como procedente del medio ambiente y no de los auriculares.
- Prestar atención a la detección y localización de objetos.

- Intentar sentir y percibir las distancias.
- Ser capaces de identificar la anchura de los objetos.
- Cómo sentían los sonidos externos procedentes del mundo real, ¿molestan estos sonidos o son incompatibles con la percepción de los sonidos enviados por el sistema?
- Se incidió en la práctica psicoacústica, es decir, ¿percibe el usuario los objetos de forma instintiva?, ¿notan o sienten la temperatura?, ¿la densidad?, etc.
- Reflexionar sobre las sensaciones del usuario.

Lógicamente y de forma general, los comentarios de los usuarios sobre estas cuestiones diferían de uno a otro, pero si se consideraron todos ellos importantes para la evaluación final. También es cierto, que cada participante se concentraba y sacaba más rendimiento de unos atributos que de otros, según su concentración, su capacidad y su forma de ser.

Después de todos los preparativos, sin ninguna directriz, se les permitió a los usuarios mirar libremente a su alrededor para detectar el espacio. Tras este tiempo se realizó una sesión de ensayo. Hacer importante mención, que todo el procedimiento completo, incluyendo la sesión de familiarización fue bastante largo, con un promedio de 2 horas, y solo se evaluó una sola muestra por cada oyente. Tras terminar cada prueba se les permitió un descanso de 15 minutos.

### **Test 2:**

El test de movilidad 2, es muy parecido al realizado en durante el test 1 con pequeñas diferencias. Los experimentos consistieron en una ruta que atraviesa una zona peatonal con una importante intersección de calles, es decir, se realizó en espacios reales con gente y objetos cotidianos. La prueba se dividió en dos tipos. El primero de ellos se realizó con los miembros del grupo A, y consistió en realizar el experimento en una acera de unos 145 metros de largo próxima a la escuela para ciegos de Berlín. La segunda de ellas se hizo con los miembros del grupo B y consistió en navegar por un camino de unos 500 metros cercano al instituto. Las pruebas y los datos de los dos grupos se registraron y analizaron por separado.

En la prueba de pasear por la acera de 145 metros de largo existía poco tráfico, y era una zona tranquila. Los participantes debían recorrer la ruta entera intentando no colisionar con ningún obstáculo (pilares, mobiliario

urbano, sillas, mesas, etc.), lo que para ellos representaba un reto constante. Era una concurrida zona comercial con una amplia variedad de obstáculos, y al final confluía con una compleja intersección de cuatro calles con un paso de peatones por el que se debía cruzar.

## 4.2.2 Usuarios

Los usuarios con los que se hicieron los experimentos fueron 10 personas ciegas con edades comprendidas entre los 26 y los 69 años y de ambos sexos. Todos ellos fueron alemanes de nacionalidad.

El total de ellos poseían una buena capacidad auditiva y eran capaces de estimar la distancia de los objetos, por tanto esto nos aseguraba que podían percibir correctamente las señales acústicas a través de los auriculares.

También se hicieron pruebas en Italia replicando el mismo sistema que el utilizado en Alemania, y los dos grupos estaban formados por 10 personas de similares características.

### Características de los participantes:

<b>EDAD</b>	Entre 26 y 69 años
<b>SEXO</b>	Masculino: 6 Femenino: 4
<b>SITUACIÓN PROFESIONAL</b>	Trabajando: 2 En paro: 4 Jubilados: 4
<b>TIPO DE CEGUERA</b>	Nacimiento: 2 Adquirida: 8

<p><b>APARICIÓN DE LA CEGUERA</b></p>	<p>Entre 0 y 5 años = 4                  Entre 6 y 18 años = 1                  Entre 6 y 18 años = 5</p>
<p><b>HABILIDAD ESTIMAR DISTANCIAS</b></p>	<p>Buena: 6                  Pobre: 4</p>

**Tabla 4.1 Perfil de los participantes**

Todos los participantes tenían buena audición, eran usuarios avanzados en el bastón y uno de ellos tenía perro guía.

### 4.2.3 Resultados

Los resultados de todos los test muestran que los distintos ejercicios presentan diferentes niveles de complejidad y dificultad, como se refleja en los resultados de las pruebas, tanto individuales como colectivas. Es obvio que los primeros ejercicios fueron mucho más sencillos de realizar por parte de los usuarios, ya que, mientras que el ejercicio 1 (columna estática) fue muy sencillo, el número 6 (superficie detrás de la columna) y el ejercicio número 7 (Moviéndose de y hacia la superficie) presentaron un grado mayor de dificultad.

Por ello, el tiempo promedio de todos los usuarios para la prueba realizada en el ejercicio 1 fue de casi 12 minutos, mientras que el tiempo total se elevó a 62 minutos y 33 segundos para el ejercicio número 6, y 44 minutos en el ejercicio número 7. Los resultados se mantuvieron estables a lo largo de los test.

Ocurre lo mismo si nos fijamos en los parámetros individuales. La persona TP01 completó el ejercicio número 1 en 1 minuto 48 segundos la primera vez, mientras que aumento el tiempo en 4 minutos para el ejercicio 6 y en 4 minutos y 24 segundos para el ejercicio 7.

Otras diferencias importantes fueron:

Las personas jóvenes se manejan con mayor destreza y necesitan mucho menos tiempo que las personas mayores. Por dar un ejemplo, las personas TP09 y TP10 son mayores de 60 años. Sus puntuaciones absolutas en tiempo en la primera prueba fueron de 32 minutos y 49 segundos, y de 32 minutos y 26 segundos respectivamente. Si se compara con la persona TP01, una mujer de 26 años, su tiempo fue de 18 minutos y 24 segundos. Se ve claramente que es casi el doble de tiempo el que emplean las personas mayores.

Este patrón se mantuvo estable durante las tres pruebas. Por lo tanto, podría haber una correlación entre la edad y el tiempo necesario para terminar las pruebas, pero no se llegó a estudiar este parámetro por la poca cantidad de información que se poseía.

### **ANALISIS DE LAS PUNTUACIONES MEDIAS DEL EXPERIMENTO:**

Los resultados se plantean según el tipo de ejercicio. A continuación se indican los resultados obtenidos en las pruebas BPL, en el laberinto y en los test de movilidad.

#### a) Resultados de las pruebas BPL:

Es interesante analizar con detenimiento las puntuaciones medias individuales obtenidas en todas las pruebas con tres ejercicios, mostradas en la figura 4.10:

### BLP Results: Individual Test Scores from three Test Runs

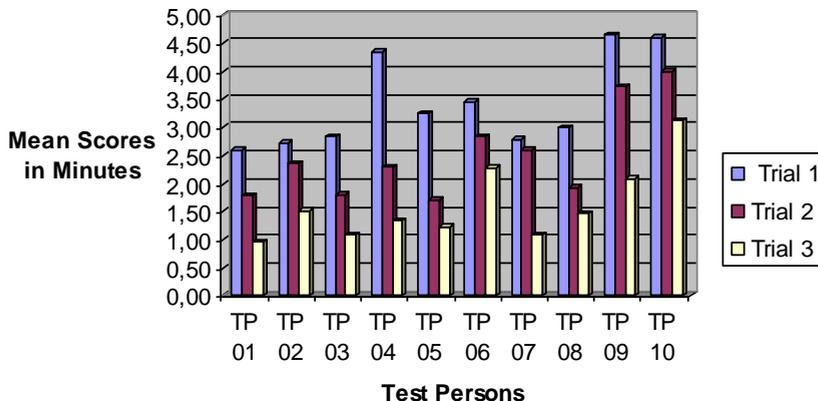


Figura 4.10 Resultados BPL de tres pruebas consecutivas

En esta figura, se refleja las puntuaciones medias obtenidas tras realizar en tres pruebas el test BPL a 10 personas de diferentes perfiles. Las alturas de las columnas representan los diferentes niveles de dificultad y complejidad del ejercicio. El eje horizontal indica los usuarios y el vertical el tiempo necesario para realizar la prueba en minutos. Al aumentar el entrenamiento con más repeticiones, los resultados disminuyen considerablemente, lo que nos muestra la importancia y el impacto que tiene el formar y entrenar a los usuarios con el sistema.

Es obvio decir, que a medida que los usuarios iban mejorando sus resultados, cada vez se sentían más motivados e ilusionados con el sistema, ya que empezaban a entender lo que tenían a su alrededor y se sentían más seguros. Cada vez necesitaban menos tiempo para terminar sus ejercicios, y una vez más en los más jóvenes las puntuaciones absolutas en tiempo cayeron casi 1/3, mientras que para usuarios de tercera edad fue más moderada, pero no menos importantes. En resumen, las puntuaciones promedio de las tres pruebas para TP01 están en 2,60 minutos, 1,78 minutos y 0,95 minutos respectivamente. Para el usuario TP03, el tiempo de la primera prueba fue de 2,83 minutos, el de la segunda de 1,80 minutos y la tercera duró 1,08 minutos. Si nos fijamos en el usuario TP9, un usuario de 67 años totalmente ciego, los

resultados fueron de 4,64 minutos para la primera prueba, 3,73 minutos para la segunda y 2,09 minutos para la tercera.

Este patrón confirma su mejora si reflexionamos de manera general sobre los test realizados. Mientras que la primera prueba tardó en ejecutarse sobre todo el grupo un total de 34,27 minutos, este resultado se redujo a 16,16 minutos en la tercera vez que se realizó la prueba.

El objetivo de la estrategia BPL es mostrar los resultados de las pruebas en minutos. En la figura 4.11 se reflejan los resultados generales de todos los usuarios por ejercicio, no como antes que hablábamos de resultados particulares por usuario. El tiempo necesario por el grupo para completar el ejercicio disminuye desde la primera prueba hasta la tercera. Se ve claramente que el ejercicio más complicado para los usuarios fue el 6, aunque desde la primera repetición hasta la tercera mejoró considerablemente el rendimiento. La clara disminución del tiempo medio necesario que las personas ciegas necesitan para pasar por una serie de ejercicios BPL, es decir, la habilidad para sentir y medir la anchura y la altura de los objetos, identificar áreas libres, y caminar a través del espacio puede ser justificado por un constante entrenamiento. Algunas personas que participaron en las pruebas, y siempre de acuerdo a sus características y habilidades, pasaron de 30 a 60 minutos mejorando su reconocimiento y rendimiento de la percepción.

**BLP Results: Collective Scores  
from three Test Runs**

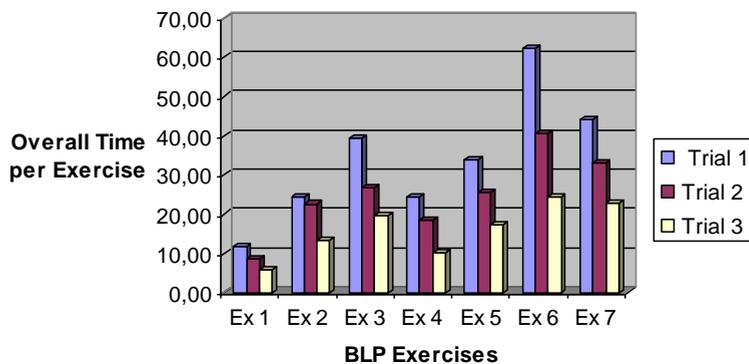


Figura 4.11 Resultados totales por ejercicio

b) Resultados de las pruebas en laberinto:

Tiempo de recorrido

Como segundo paso dentro de las etapas de evaluación, los usuarios navegaron a través del laberinto construido con cuatro pares de columnas y por los escenarios objeto de nuestro test. No usaron bastón, por lo tanto debían confiar únicamente en el sistema para identificar los obstáculos y detectar zonas libres a través de las cuales se podían mover sin tocar ni golpearse con los objetos.

Durante el ejercicio debían mover la cabeza lentamente de izquierda a derecha y viceversa, con el fin de identificar las columnas y medir la anchura del pasillo por el que se podían mover hasta llegar al siguiente par de columnas o vía libre. Aunque no lo parezca, estas personas estaban acostumbradas a moverse con el bastón, por lo tanto sin el, estos ejercicios fueron todo un reto para ellos. Lo que sí que era notable en todos los ejercicios es que los resultados mejoraban proporcionalmente al número de veces que se repetía la prueba, así como al tiempo de entrenamiento.

Analizando el tiempo que los usuarios necesitaban para caminar y recorrer la ruta planteada, se notaba que al principio tuvieron grandes dificultades al no utilizar el bastón.

Todas las personas que intervinieron en estas pruebas tenían un buen nivel de movilidad, incluidas las más mayores. El rango de PWT varió de 1 a 2 segundos por metro, con una media de 1,40 segundos por metro. De esta forma, la persona más lenta de la prueba empleó casi el doble de tiempo que tardó la más rápida para completar la ruta.

Se puede pensar en justificar esta diferencia con la edad, pero no existen evidencias estadísticas que lo demuestren ya que las muestras tomadas no fueron suficientes.

En la figura 4.12, se muestra una comparativa del “Relative Walking Time” (RWT), o tiempo relativo empleado durante el recorrido en este test de movilidad, y el “Preferred Walking Time” (PWT) o tiempo de recorrido con bastón.

### LAB Results: Walking Time Scores from Two Test Runs

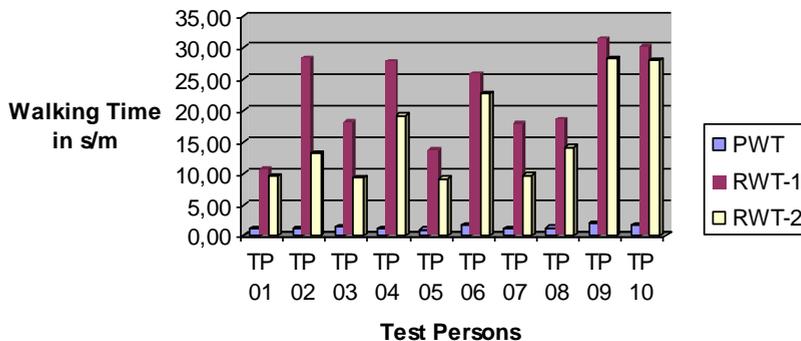


Figura 4.12 Comparativa entre tiempos RWT y PWT

En la figura 4.12, PWT refleja el valor resultante de dividir el tiempo absoluto de paseo necesario para recorrer toda la ruta con el bastón, entre la distancia del recorrido. RWT indica la cantidad de tiempo que una persona necesitó para realizar todo el recorrido durante las pruebas sin el bastón y apoyándose únicamente en el sistema, y se calculó dividiendo este tiempo entre la distancia del recorrido. PWT indica tiempo de recorrido utilizando bastón y RWT tiempo de recorrido sin bastón y solo utilizando el sistema. Se realizaron dos pruebas por persona, y por eso aparecen dos tiempos: RWT-1 y RWT-2.

Los resultados de las pruebas reflejadas en la tabla 4.12, revelan que el tiempo necesario para realizar el recorrido aumentó considerablemente cuando los usuarios tuvieron que hacerlo sin bastón y utilizando solo el sistema. Mientras que el RWT-1 para la persona más rápida, TP01, fue casi 9 veces mayor que su PWT, 10,50 segundos/metro (PWT) frente a 1,20 segundos/metro (RWT), para la persona más lenta, TP09, la diferencia fue de casi de 16 veces más, 31,21 segundos/metro frente a 2 segundos/metro. Esto no es casualidad, ya que todas las personas que participaron en las pruebas tenían mucha experiencia en el uso del bastón, y ninguna debido a su novedad, en el uso del sistema. Navegando sin el bastón y utilizando solo señales acústicas para interpretar los detalles del espacio provocó cierta incomodidad en los usuarios, pero desde el punto de vista de la investigación nos da la

oportunidad de analizar la eficacia y potencial de los mapas acústicos para la representación del espacio, y evaluar si el entrenamiento en este tipo de sistemas puede mejorar la calidad de vida y el rendimiento en la movilidad de las personas invidentes.

También hay una gran diferencia entre los resultados obtenidos entre los usuarios de manera individual. El RWT-1 varía desde 10,50 segundos/metro en el mejor de los casos (usuario TP01) y 31,32 segundos/metro en el peor de los casos (usuario TP09). Por lo tanto, la persona más lenta necesitó alrededor de tres veces más de tiempo que la más rápida. Pensamos que la edad también podía ser un factor importante para los resultados, pero la poca cantidad de datos no fue suficiente para confirmarlo.

Si nos centramos en los resultados de la segunda prueba RWT-2, volvemos a confirmar la teoría de que el entrenamiento mejora la capacidad y seguridad de navegación. En algunos casos es pequeña la diferencia, como en el caso de TP01 donde la mejora de tiempo es de tan solo 0,95 segundos/metro, pero en otros casos, como por ejemplo en el caso de TP02 la diferencia es exagerada, donde el tiempo varía de 28,20 segundos/metro en el primer ejercicio a 13,11 segundos/metro en el segundo.

Respecto al grupo en general, y comparando las diferencias de tiempo de la primera prueba respecto la segunda, se observa que las medias oscilan entre 22,14 segundos/metro y 16,28 segundos/metro.

Por lo tanto, podemos concluir diciendo que un entrenamiento adecuado tras el primer contacto del usuario con el sistema implica una mejora importante en la detección de los obstáculos y en el tiempo de recorrido, reforzando la confianza de los usuarios en el sistema e incrementando su seguridad.

### Análisis de las colisiones

Además de los tiempos empleados en recorrer los escenarios de test para medir la mejora que el sistema aporta a la navegación de las personas invidentes, también se analizó el número de colisiones NUH (Number of Unintended Hits) no deseadas para evaluar el rendimiento de movilidad. En la figura 4.13 se muestra el resultado de las pruebas. Este nuevo parámetro nos da información adicional indicándonos como los invidentes son capaces de esquivar obstáculos.

La metodología consistió en que tras la primera prueba, se les informó a los usuarios la manera de como poder evitar los obstáculos utilizando el sistema, es decir, volvemos otra vez más a la importancia del entrenamiento y el entendimiento del dispositivo. Tras estas explicaciones los usuarios realizaron un segundo intento con una previa reorganización de las columnas. Estas pruebas se llevaron a cabo en el escenario del laberinto.

Las puntuaciones representadas en la figura 4.13 revelan grandes diferencias entre los usuarios, ya que algunos lo hicieron mucho mejor que el resto. El rango de NUH varia de 1 a 5 accidentes no deseados en la primera prueba y de 0 a 3 en la segunda. Lo mismo que en ejercicios anteriores, el modelo refleja que el número de colisiones disminuyó de forma general tras una fase de entrenamiento y entendimiento del sistema.

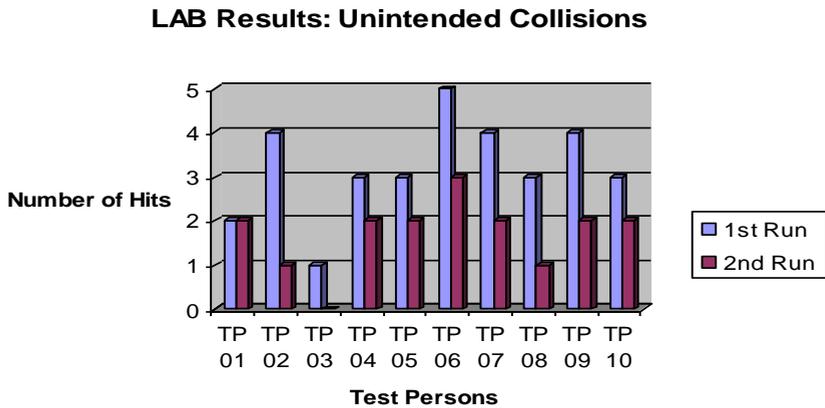


Figura 4.13 Colisiones no deseadas

Además, si nos fijamos en los resultados generales del parámetro NHU, también se hace evidente que el entrenamiento reduce de forma importante el número de colisiones no deseadas. El resultado general de NUH en la primera prueba es de 32 golpes, mientras que en la segunda fue de 17 golpes, es decir, en términos de detección y evasión de obstáculos, el rendimiento de movilidad mejoró en casi un 50% para el grupo.

Correcciones sobre los usuarios

La capacidad de mantener el cuerpo y la cabeza en posición erguida y estable durante la navegación para que lo que perciben sea coherente con la realidad, es uno de los desafíos clave para la movilidad. Esto puede provocar desorientación y pérdida de la dirección correcta. Cuando esto ocurría, no se permitió que los usuarios emplearan un tiempo ilimitado para volverse a colocar bien en caso de pérdida de la dirección correcta, sino que se les corrigió con ayuda de personas videntes para colocarlos en la ruta correcta.

En la figura 4.14 se muestra el número de correcciones realizadas por personas videntes a los usuarios debido a pérdida de la orientación durante el ejercicio.

**LAB Results: Number of Corrections**

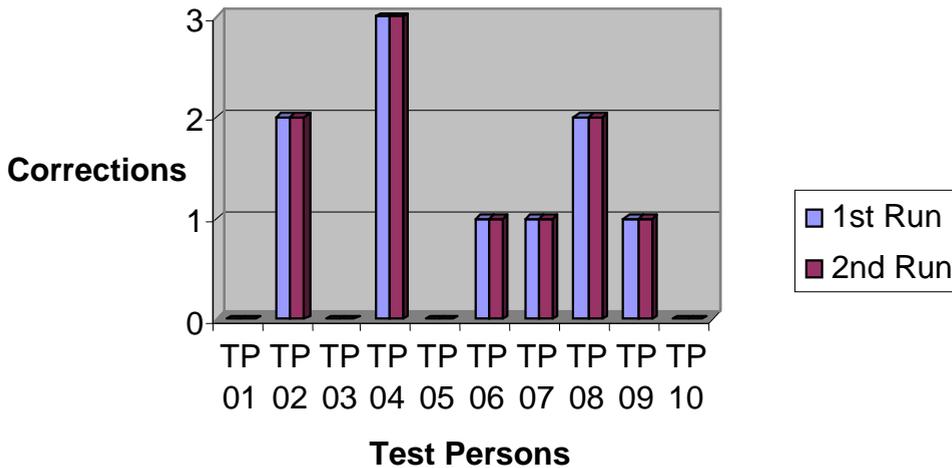


Figura 4.14 Número de correcciones

Al contrario que en pruebas anteriores, los resultados revelan que la formación no mejoro mucho la movilidad en lo que se refiere a las correcciones debido a la desorientación de los usuarios, pero el resultado no fue tan malo después de todo. Mientras que el 40% de los usuarios de la

prueba fueron capaces de navegar sin ningún tipo de corrección, el 30% de ellos no necesitó más de una corrección.

c) Resultados de las pruebas de movilidad:

Dividimos los escenarios de movilidad en dos tipos: MOB-A y MOB-B.

**MOB-A**

El escenario MOB-A fue el primero de los dos test de movilidad en exteriores. La ruta que los usuarios debían hacer, consistió en partir de las instalaciones de la asociación de ciegos de Berlín y recorrer un camino de unos 29 metros de largo hasta un punto donde tenían que tocar una puerta.

En la ruta existían diversos obstáculos tales como un coche, ruedas, cubos de basura, un banco y un par de columnas. Para evaluar la eficacia del sistema en lo que se refiere al reconocimiento de la altura de los objetos, se colocó una cinta de plástico a la altura de la cabeza, obstruyendo el camino.

En general, los resultados obtenidos reflejaron que esta prueba de movilidad fue bastante más complicada para la mayoría de usuarios que la prueba del laberinto, tal como lo indica el índice de tiempo necesario para recorrer todo el camino. El rango de tiempos obtenido respecto al parámetro PWT estuvo entre 1,07 segundos/metro y 2 segundos/metro en la prueba del laberinto, y entre 1,72 segundos/metro y 3,58 segundos/metro en la prueba MOB-A. Obviamente, los usuarios iban con mucha más precaución en este tipo de escenarios que en las pruebas anteriores. Esto demuestra que, incluso con bastón, el tiempo necesario para recorrer escenarios del tipo MOB-A era también mayor, casi el doble, que el necesario para recorrer las pruebas de laberinto, con lo cual para los usuarios la complejidad de las pruebas de escenario es mayor que las pruebas de laboratorio.

Tiempo de recorrido

Al comparar los resultados RWT con los PWT (figura 4.15), se observa que el recorrer un escenario del tipo MOV-A sin el bastón fue un gran desafío para todos los usuarios del sistema. El rango de valores de PWT fue desde 1,72 segundos/metro hasta 3,58 segundos/metro, mientras que los valores de RWT se elevó a valores comprendidos entre 10,61 segundos/metro y 31,61

segundos/metro. Por lo tanto, mientras que la persona más lenta de la prueba necesita casi el doble de tiempo que el más rápido para hacer el ejercicio con el bastón, la cantidad de tiempo necesario fue de casi tres veces más sin el bastón.

Como se puede deducir de la figura 4.15, los resultados obtenidos al comparar los datos PWT y RWT ponen de manifiesto que todo el grupo necesitó 9 veces más de tiempo para desplazarse por el escenario MOB-A sin el bastón y con el sistema, que solo con el bastón. Estos tiempos medios son de 2,40 segundos/metro sólo con bastón y de 18,70 segundos/metro con el sistema. El tiempo que necesitaron los usuarios para realizar la prueba sin el bastón fue considerablemente mayor. Como siempre, es bastante significativo que el tiempo necesario para realizar la prueba disminuye después del entrenamiento, tanto individual como colectivamente.

### MOB-A: Walking Time Scores

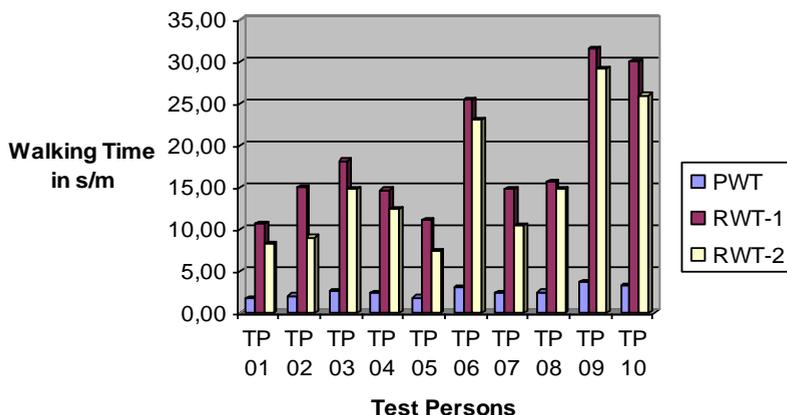


Figura 4.15 PWT y RWT para escenario MOB-A

¿Es por tanto importante el entrenamiento para mejorar los resultados? Al fijarnos en RWT-2 (tiempo que las personas necesitaron para recorrer el escenario tras el entrenamiento) está claro que la respuesta es SI.

Sin embargo, siguen existiendo importantes diferencias entre los propios usuarios. En algunos casos, como el del usuario TP08, la diferencia de tiempo entre RWT-1 y RWT-2 es mínima (de 0,81 segundos/metro). En otros casos, como en el ejemplo de TP02 (14,97 y 8,97 segundos/metro) y TP07 (14,75 y 10,46 segundos/metro) la diferencia es mayor. La mejora de tiempo media para el grupo fue de 3,2 segundos/metro.

En lo que se refiere a la edad, vuelve a hacerse evidente que las personas mayores tuvieron más dificultades para realizar las pruebas que las más jóvenes. Por comentar algún caso, podemos fijarnos en el usuario TP05, un hombre de 28 años. TP09 es el caso contrario, ya que se trata de un ciego de 67 años. Los valores RWT-1 y RWT-2 demuestran que las cantidades de tiempo de TP09 son tres y cuatro veces más que las de TP05.

### Análisis de las colisiones

Revisando de nuevo el número de veces que los usuarios colisionaron durante la ejecución de esta prueba, se confirma otra vez que el entrenamiento y la formación fueron decisivos para el manejo del sistema de forma segura.

Como se puede ver en la figura 4.16, en la primera prueba sin formación previa, hubo gran diferencia entre los diferentes usuarios en lo que respecta al choque y colisión con los obstáculos del recorrido. Los valores del parámetro NUH para esta prueba de 7 obstáculos variaba de 1 a 4. Se puede ver claramente que tras una capacitación se redujo bastante el número de colisiones.

### Results MOB-A: Unintended Collisions

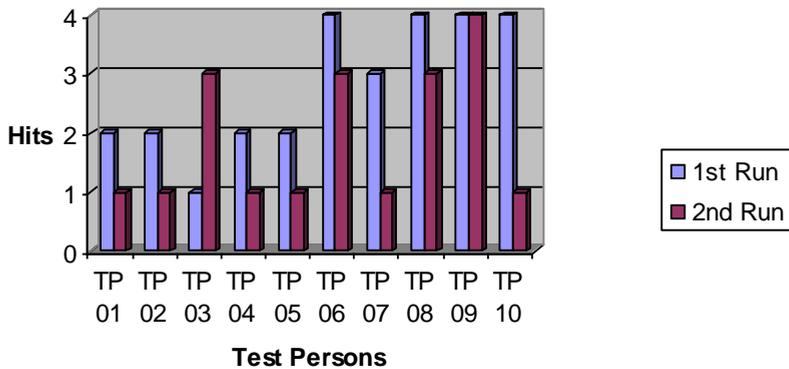


Figura 4.16 Colisiones no deseadas en la prueba MOB-A

En 8 de los 10 casos el número de colisiones se redujo hasta un 50% e incluso un 75% en el usuario TP10.

Ninguna de las personas que se sometieron a la prueba fue capaz de detectar la banda de plástico situada a la altura de la cabeza, lo que indica que el sistema no es de gran ayuda en la detección de objetos en elevación, a no ser que el usuario tenga la paciencia y sea consciente de explorar el contexto moviendo la cabeza de arriba abajo para poder identificar los objetos por debajo o por arriba del centro de visión.

#### Correcciones sobre los usuarios

En la figura 4.17 aparece un análisis del número de veces que se tuvo que reorientar a los usuarios durante las pruebas para seguir su ruta. En 3 de cada 10 casos no se produjeron mejora tras el entrenamiento. Solo en dos casos se produjo un aumento del número de colisiones tras el entrenamiento. La gráfica muestra que el 50% (5 de 10) de las personas involucradas en la prueba fue capaz de, tras el entrenamiento, recorrer la ruta con menos correcciones que al principio.

### Results MOB A: Number of Corrections

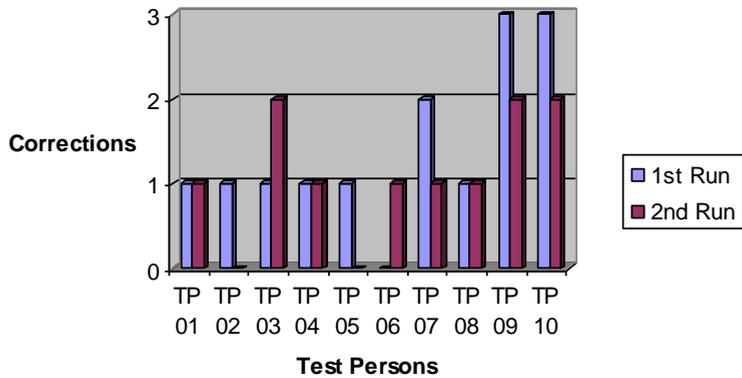


Figura 4.17 Correcciones en el test MOB-A

### MOB-B

El escenario MOB-B se configuró en una zona comercial. Era la prueba más complicada de las realizadas hasta ahora, pues existían una gran cantidad de obstáculos, tales como mobiliario urbano, cabinas, carteles publicitarios de diferentes tamaños, coches estacionados en las aceras, motos, pilares, bolardos de acero, etc... Los usuarios de la prueba no solo tuvieron que hacer frente a estos obstáculos, sino también a los numerosos peatones que paseaban por el centro comercial y a los ciclistas que había alrededor de ellos.

Sorprendentemente, y a pesar de la complicaciones añadidas en este ejercicio, las personas ciegas que hicieron la prueba fueron más capaces de sortear obstáculos y navegar de forma más coherente que en la prueba MOB-A. Esto se demuestra, entre otras cosas, porque la diferencia entre los parámetros PWT y RWT han disminuido.

#### Tiempo de recorrido

Los resultados de esta prueba revelan que existen bastantes diferencias entre los usuarios sometidos al test MOB-B, tal como lo indican los valores RWT. Tomando como ejemplo RWT-1, la diferencia va desde 2,72 segundos/metro

en el usuario TP01 hasta 8,40 segundos/metro en el usuario TP09, lo que indica una diferencia entre el más rápido y el más lento de 3 veces mayor.

De nuevo podemos pensar una vez más que existe una estrecha relación entre la edad y la calidad de la prueba, ya que los valores de tiempo más altos se relacionan con los usuarios de mayor edad, aunque en este caso las diferencias no son tan grandes como en las pruebas anteriores (laberinto y MOB-A).

También, y como en casos anteriores, se debe recalcar la importancia del entrenamiento. Tras completar la primera prueba (RWT-1), se debatieron con los usuarios cuestiones relacionadas en cómo se podría mejorar los resultados obtenidos, así como cualquier observación que los usuarios quisieran abordar. Obviamente, y como se puede observar en los resultados finales de la figura 4.18, esto ayudó a mejorar el rendimiento de la movilidad, reflejándose objetivamente en el descenso de los valores de RWT-2.

### MOB-B Results: Walking Time Scores

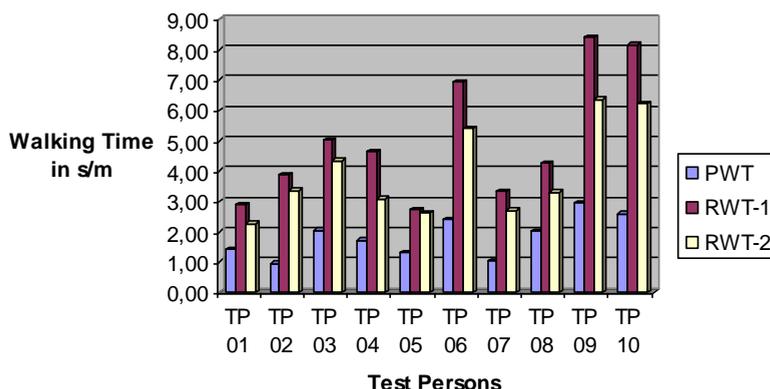


Figura 4.18 PWT y RWT para escenario MOB-B

La figura 4.18 muestra claramente la diferencia existente entre acciones individuales entre TP01 y TP10, así como la importancia de la formación en este tipo de pruebas. El tiempo necesario para completar el ejercicio disminuyó en la segunda prueba en el 100% de los casos.

### Análisis de las colisiones

Los resultados generales indican que los usuarios fueron capaces de evitar colisiones y choques con los obstáculos o personas del recorrido de diferente forma. Algunos usuarios llegaron a tener hasta 7 y 8 colisiones durante la prueba, mientras que otros no pasaron de 2. Sin embargo, volvemos a resaltar la importancia que tiene el entrenamiento si comparamos los valores del primer recorrido con los del segundo. En todos los casos, salvo en dos, el número de colisiones en un escenario tan complicado como este disminuyó, salvo en uno de los casos (TP03) que se obtuvo el mismo resultado, y en otro (TP04) que incluso aumentó.

Además, excepto en el usuario TP09 (el más mayor de todos), el número de colisiones durante la prueba no subió de 5, cantidad que durante el primer recorrido estuvo por encima de 5 en casi todos los casos.

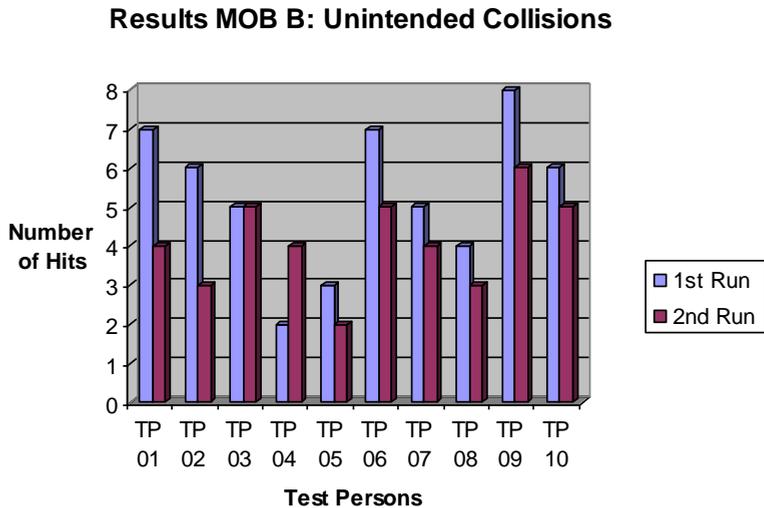


Figura 4.19 Colisiones no deseadas en la prueba MOB-B

En la figura 4.19 se resumen el número de colisiones que los 10 usuarios de la prueba tuvieron durante el recorrido en el escenario MOB-B, antes y después de la formación. Se puede observar, que tras el entrenamiento, la mayoría de

ellos fue capaz de hacer el recorrido con un menor número de colisiones, ya sea con objetos o personas en movimiento.

Con el fin de evaluar el rendimiento individual de movilidad, se añadió el parámetro “NUH\_Ind”. Este índice se define como la relación entre la longitud del recorrido y el número de colisiones no deseadas. Se mide en metros de distancia recorridos sin golpes. Cuanto mayor sea el valor del índice, mayor es la distancia recorrida sin tropezar con obstáculos o personas. Los resultados obtenidos demostraron que todas las personas fueron capaces de recorrer distancias más largas sin colisiones tras el entrenamiento.

La persona que más distancia recorrió sin colisionar con ningún obstáculo, lo hizo durante 72,50 metros, mejorando en 24,17 metros respecto al primer ejercicio. El índice NUH\_Ind medio del grupo pasó de 27,35 metros sin chocar en el primer ejercicio a 35,36 metros en el segundo. La diferencia es de 8,01 metros.

Es interesante mencionar que el índice NUH\_Ind refleja que, en general, las 10 personas ciegas que se sometieron a la prueba fueron capaces de realizar el ejercicio en el escenario MOB-B con menos colisiones que en el caso de MOB-A, aunque lógicamente las condiciones y el contexto fueron mucho más complejos. Es decir, anduvieron más metros sin colisionar en el test de MOB-B que en el MOB-A (ver figura 4.20).

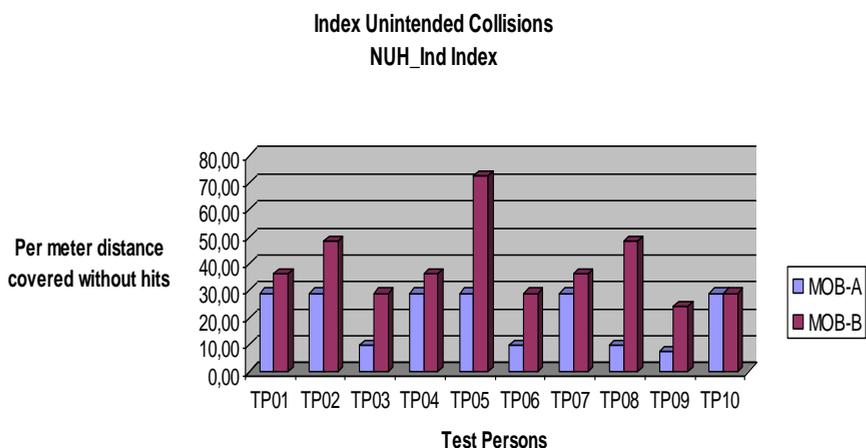


Figura 4.20 Comparativa del índice NUH\_Ind entre MOB-A y MOB-B

En resumen, si nos fijamos en el grupo, es evidente que durante el recorrido MOV-A, el grupo fue capaz de andar sin chocar aproximadamente 15 metros seguidos tras la formación, mientras que en el caso del escenario MOB-B, la distancia aumentó 35 metros.

Mi reflexión final, es que como la prueba MOB-B se realizó después de la MOV-A, y tras todo el periodo previo de entrenamientos y ejercicios de laboratorio, los usuarios estaban mucho mejor formados y entendían mucho mejor la filosofía del sistema y la interpretación del espacio mediante los mapas acústicos generados por el dispositivo. Es decir, una buena formación y entrenamiento en estos sistemas, puede asegurar la mejora de la navegación en las personas invidentes, así como mejorar su calidad de vida.

### Correcciones sobre los usuarios

Como a lo largo de toda la tesis hemos estado demostrando, es evidente que la necesidad de formación y capacitación en el manejo del sistema mejora considerablemente el rendimiento en la navegación de las personas invidentes. También se cumple en la cantidad de correcciones que fueron necesarias para permitir que los usuarios, durante la realización del ejercicio MOB-B, siguieran el rumbo de su camino. Durante este ejercicio, el rango numérico del número de correcciones realizadas fue de 2 a 5 en la primera prueba (1st Run) y de 1 a 5 en el segundo recorrido (2nd Run), tal como se muestra en la figura 4.21. Ninguna persona ciega fue capaz de hacer el ejercicio sin tener que ser ayudado en algún momento del mismo, es decir, todos necesitaron alguna corrección para poder retomar el sentido del recorrido. Si comparamos los resultados de la primera y segunda prueba, podemos darnos cuenta de que en 6 de los 10 casos (TP02, TP04, TP06, TP07, TP08 y TP09) se produjo una mejora disminuyendo el número de correcciones, mientras que en dos de los casos no hubieron cambios (TP01 y TP05), y en otros dos casos (TP03 y TP10) incluso se incrementaron.

El número total de correcciones se redujo de 34 a 27 después del entrenamiento.

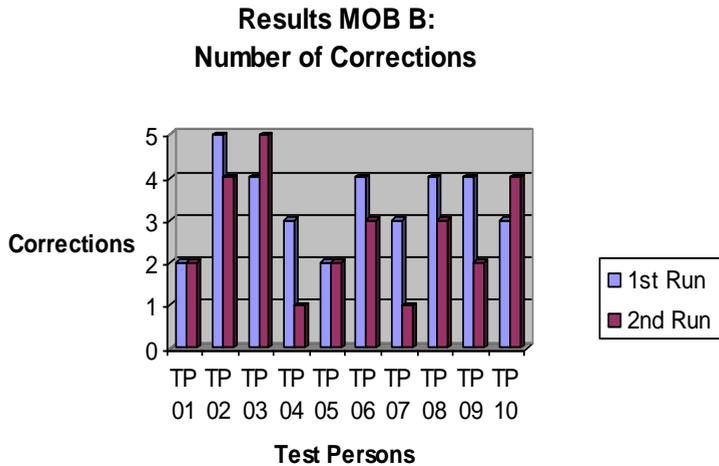


Figura 4.21 Correcciones en el test MOB-B

Otra variable importante analizada respecto al número de correcciones, fue el índice de correcciones obtenido comparando las veces que los usuarios necesitaron ayuda durante las pruebas de los escenarios MOB-A y MOB-B. Este índice relaciona la longitud de la ruta con el número de correcciones que la persona necesitó hasta llegar a su destino. Las unidades de este índice son “metros por corrección”, y prácticamente nos indica la cantidad de metros que los usuarios pudieron andar como máximo de forma independiente y continuada sin necesidad de ser corregidos y atendidos durante su recorrido por el escenario.

Por tanto, la figura 4.22 refleja en buena medida la comparativa del rendimiento de movilidad entre los dos escenarios, MOB-A y MOB-B. Cuanto más alta es la columna (puntuación del índice) mayor es la distancia andada sin ayuda de los instructores.

No se incluyen los resultados de los casos TP02 y TP05 porque no necesitaron ayuda durante el recorrido. Su número de correcciones fue 0.

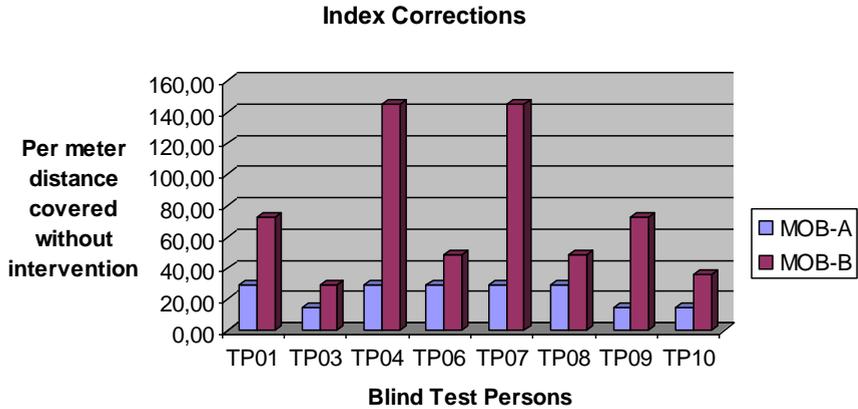


Figura 4.22 Índices de corrección entre MOB-A y MOB-B

Se puede observar con claridad que, aunque la prueba diseñada en el escenario MOB-B fue mucho más complicada que la del MOB-A, la cantidad de metros andados de forma continuada es mucho mayor en el escenario MOB-B en todos los casos. La causa se puede achacar al entrenamiento tan intenso que ya habían realizado los usuarios y a lo familiarizados que ya estaban con el sistema. Como comparativa final, la cantidad total de metros recorridos por todo el grupo de forma continua y sin correcciones durante la prueba del escenario MOB-A fue de alrededor de 130 metros sin correcciones. El mismo dato en el escenario MOB-B fue de 540 metros, lo que indica un aumento de 4 veces más.

## 4.3 Conclusiones

Tras haber estado durante varias sesiones de entrenamiento con personas invidentes de diferentes edades, perfiles, capacidades y tipo de enfermedad aprendes a entenderlos y puedes imaginar un poquito más como pueden llegar a pensar e interpretar el espacio, aunque nunca llegaremos a poder sentir lo mismo que ellos o, ni mucho menos, llegar a conocer como exactamente manejan su cerebro. Sí que es verdad que, únicamente desde el punto de vista científico y experimental e independientemente de intentar entender su cerebro, todas las pruebas realizadas y comentadas durante este capítulo de la tesis demuestran que la tecnología bien aplicada puede ayudar a la navegación de estas personas.

Sin entrar en detalles tecnológicos, pues ya se abordaron en capítulos anteriores, desde el punto de vista experimental se han realizado diferentes pruebas, tanto en entornos cerrados como abiertos, en laboratorio, en lugares con pocos obstáculos e incluso en lugares públicos muy concurridos. En todos ellos se analizaron los mismos parámetros: tiempo de recorrido, número de colisiones y número de correcciones. Además, se compararon recorridos con bastón, sin bastón, con ayuda y sin ayuda y cada vez más se fueron complicando los ejercicios.

Tras describir en este capítulo las estrategias experimentales, pruebas y exponer los resultados obtenidos tras la validación llevada a cabo en usuarios invidentes, objetivamente, las principales conclusiones a las que puedo llegar son:

1. Cada vez que pasaba el tiempo, los resultados positivos fueron incrementándose, es decir, en ningún momento tuvimos la sensación de retroceso y de frustración científica. Además, ellos mismos se motivaban cada vez más y se sentían más independientes. La ilusión que le ponían era maravillosa, e incluso llegaron a comentar varias veces que les dejáramos más tiempo para estar con el sistema, ya que estaban muy involucrados y felices.
2. Tantas pruebas, experimentos y ejercicios me han llevado a certificar rotundamente que la base del éxito de estas pruebas está en la formación y entrenamiento. No es un sistema invasivo o incómodo en absoluto, pero sí que, dependiendo de la capacidad del usuario final, y

sin un tiempo exagerado, la capacitación aumenta y mejora sin ninguna duda el rendimiento de movilidad y el éxito del sistema.

Además, ha quedado claro que los usuarios pueden ser capaces de, mediante señales acústicas, detectar y localizar objeto de forma autónoma. Todo esto lo consiguen de forma cognitiva, es decir, aprenden de manera constante sin comparar nada con nada, sino repitiendo las pruebas y autoalimentándose de información procedente del mundo real, nuestro mundo, que algunos de ellos nunca han llegado a conocer, siendo capaces de tener cierta autonomía para andar y esquivar obstáculos.

Estas pruebas experimentales se han podido llevar a cabo gracias a la integración, la tecnología empleada y el prototipo fabricado según se explica en los capítulos anteriores, mezclándola con diferentes protocolos y metodologías de aprendizaje para analizar el comportamiento de la unión entre el usuario invidente final y la tecnología integrada en este proyecto.

Gracias a esta unión y a los resultados conseguidos, podemos casi asegurar que, si seguimos investigando en esta línea, se pueden conseguir aún mejores resultados con otros sistemas de navegación en un futuro próximo. Además, se ha demostrado que el sistema es capaz de ofrecer información real de obstáculos, en lo que respecta a la distancia del objeto según la referencia del punto de visión del usuario portador del sistema y al azimut, o posición en el plano horizontal del objeto.

Por otra parte, queda constancia de que es imprescindible la implicación de otras personas videntes en la realización de las pruebas, tanto para la formación de estas personas como para ofrecer ayuda en caso de pérdida o desorientación.

Por último, desde el punto de vista estrictamente humano, y sin mezclar ciencia y tecnología con sensaciones, me gustaría decir que, por encima de todo, hasta de la propia ciencia, está el bienestar que provoca ver a personas que padecen alguna incapacidad física contentos y con esperanzas de que algún día, algunos científicos puedan ayudarles a sentirse más autónomos, mas integrados en la sociedad y, en definitiva, un poquito más ellos.



## 5. CONCLUSIONES GENERALES

Las conclusiones generales de esta tesis las voy plantear desde tres puntos de vista: el profesional, el científico y tecnológico y el humano.

### PROFESIONAL:

Aprovecho para decir que he estado durante 12 años de mi vida en la empresa privada participando en el diseño y ejecución de proyectos de ingeniería directamente relacionados con tecnología en general y aplicada a muy diversos sectores. Así, puedo hablar de líneas de producción aplicadas a la automoción, cerámica, siderurgia, alimentación, entre otros, y aplicando técnicas de robótica, visión artificial, automatismos, etc..., y la conclusión a la que he llegado tras estar involucrado muy directamente en un trabajo de este tipo ha sido que, aunque todo lo hecho por supuesto tiene su valor, la sensación que he experimentado al cambiar el trabajar con máquinas, a hacerlo con personas, es mucho más gratificante y se lo aconsejo a todo ingeniero o investigador que pueda y quiera experimentarlo.

### CIENTÍFICO Y TECNOLÓGICO:

Se ha demostrado que aún nos queda mucho que hacer en lo que respecta a entender cómo funciona el cerebro humano y que pensamos y como lo hacemos. Además, cada vez está más claro que cada uno de nosotros vemos e interpretamos el mundo a nuestra manera. No hay más que ver que, entre videntes, muchísimas veces y durante nuestro día a día, no nos entendemos en la forma de ver, hacer o interpretar las cosas. Eso es porque el prisma con el que nuestros cerebros interpretan las cosas varia y, cada uno de nosotros nos creamos nuestro propio mundo.

Puesto que esto es así y seguirá siendo así por siempre, humildemente, lo que investigadores como nosotros podemos hacer es intentar utilizar todo el poder de la tecnología y nuestro conocimiento, en diseñar mecanismos para que las personas puedan mejorar su calidad de vida y la coherencia con la que hay que interpretar el mundo. Por ello, tras el trabajo explicado en esta tesis

doctoral, he llegado a la conclusión de que se puede conseguir unir tecnología y personas por:

Se ha definido, integrado y demostrado que un sistema compuesto por láser, por visión puede ayudar a interpretar el mundo a invidentes.

Este dispositivo permite detectar obstáculos y moverse por espacios y pasillos libres.

Aunque los resultados han sido buenos, aun es necesaria la ayuda del bastón, y durante algún tiempo se deberán mezclar ambas ayudas, manteniendo como insustituible el bastón.

El sistema demuestra que es de los pioneros en el mundo científico, avalado por los resultados tanto científicos como prácticos, ya que, entre otras cosas permite su interacción con el contexto externo sin anular las sensaciones procedentes del entorno del usuario.

Confío en que se puede acortar el tiempo de entrenamiento del sistema, siempre dependiendo de la actitud, aptitud y capacidad del usuario final, pero es imprescindible ayudar a entender el sistema y capacitar a los usuarios para conseguir el mayor éxito en su manejo.

De momento, y dependiendo del avance en cuestiones de nano electrónica, el volumen y tamaño del sistema en general llega a ser un poco incómodo y pesado para la persona que lo utiliza.

Por tanto, basándome en los resultados obtenidos y documentados en esta tesis, y tras los experimentos llevados a cabo durante este trabajo de investigación, así como de las opiniones obtenidas de los usuarios, confirmo que la integración de diferentes tecnologías mejora la vida de las personas invidentes ayudándoles a sentir y percibir el espacio que les rodea.

### HUMANO:

Ya lo he comentado en el apartado 4.2, pero me gustaría hacer especial hincapié en ello debido a lo importante que este trabajo ha sido para mí, posiblemente el más importante en mi vida profesional después de participar en decenas de proyecto, y de lo orgulloso que me siento de los resultados, decir que en algunos momentos del proyecto se me llegaron a, literalmente,

poner los pelos de punta al ver a los invidentes contentos, escuchar frases como “Víctor no me quites el sistema, déjame un poco más”, “¡Cuánta gente tengo en frente!, ¿Es un árbol y no una persona lo que hay delante de mí, verdad?, y otras cuantas más que podrían ser dignas de otra tesis pero en la rama de ciencias médicas o sociales.

En definitiva y para terminar, aunque somos científicos y pragmáticos midiendo todo con parámetros objetivos, cuando este proyecto fue evaluado por la comisión europea, al preguntarme un montón de cosas de burocracia y papeleo, lo único que les contesté fue “mirar al invidente como se mueve y preguntarle a él”. Me contestaron: “es verdad Víctor, es increíble”

¡La sensación personal es inexplicable!



## 6. FUTUROS PLANES DE TRABAJO

Como trabajos futuros tras los resultados obtenidos en este trabajo, planteo:

### Investigación Académica

- Optimizar los algoritmos de procesado de imagen para que el tiempo de filtrado y obtención de los mapas de profundidad para después extraer los objetos sea el mínimo posible.
- Generar diferentes bancos de sonidos con distintas estrategias para permitir al usuario final elegir diferentes opciones en función de donde se encuentre: calle, casa, aeropuerto, etc...
- Desde el punto de vista del desarrollo se ha de mejorar la integración de software con técnicas de multiproceso sobre sistemas capaces para ello, para optimizar aún más el tiempo de procesado.
- Por último, el investigar y desarrollar diferentes herramientas para que personas invidentes puedan aún más integrarse en la sociedad y disfrutar del contexto como hacemos el resto de personas: reconocimiento de colores, traducción de textos, reconociendo de personas, ocio, etc...

### Aplicación práctica:

- Miniaturización del sistema en lo que respecta al tamaño, peso y autonomía. Para ello se ha de trabajar en cuestiones concretas sobre: integrar el sistema en micro sistemas informáticos, y disminuir en la medida de lo posible, el cableado de los interfaces.
- Integrar en unas gafas de tamaño normal dos microcámaras que sustituyan al casco y cámaras actuales y al laser, con el fin de que el

usuario final solo tenga que llevar consigo unas gafas de sol y un mini PC.

- Buscar una solución final para conseguir equilibrio y compromiso entre software y hardware para que la captura del mundo 3D, el procesado de esta información y la generación de mapas acústicos sea lo más rápido posible.
- Sobre la plataforma hardware definitiva integrar diversas aplicaciones y utilidades para adaptar el sistema a ciegos parciales.

## 7. REFERENCIAS

[Benham T.A., 1954]

Benham T. A. 1954. Guidance Device for the Blind. *Phys. Today*, Vol. 7, No. 12, pp. 11-14

[Benjamin, 1968]

Benjamin J. M. (1968) A review of the veterans administration blind guidance project. *Bull. Of Prosth. Res.*, pp. 63-90

[Bernhard, 2008]

“Optimized distance measurement with 3D-CMOS image sensor and real-time processing of the 3D data for applications in automotive and safety engineering,” PhD Dissertation, 2008

[Birchfield, S. and Tomasi, C., 1999]

Birchfield, S. and Tomasi, C. 1999. Depth discontinuities by pixel-to-pixel stereo. *Int. J. Comput. Vision*, Vol. 35, No. 3, pp. 269–293

[Brabyn J., 1982]

Brabyn J. 1982. New Developments in Mobility and Orientation Aids for the Blind. *IEEE Transactions on Medical Engineering*, Vol. BME-29, No. 4, pp. 285-289

[Brabyn J., 1997]

Brabyn J. 1997. Technology as a support system for orientation and mobility. *American Rehabilitation*

[Capp y Picton, 2000]

Capp M. and Picton P.D. 2000. The optophone: An electronic blind aid. *Engineering Science and Education Journal*. Vol. 9, No. 3, pp. 137-143

[ChinaDis]

<http://disabilitychina.org/admin/upLoadPic/2009326152139974.pdf>

[CIDAT]

Manual de usuario del Sistema Tormes

[http://webcache.googleusercontent.com/search?q=cache:8dsSvMMZ7c0J:ftp://ftp.once.es/pub/utt/biblioteca/GPS/Orienta/manual\\_orienta.doc+Manual+de+usuario+del+Sistema+Tormes.&hl=es&gl=es](http://webcache.googleusercontent.com/search?q=cache:8dsSvMMZ7c0J:ftp://ftp.once.es/pub/utt/biblioteca/GPS/Orienta/manual_orienta.doc+Manual+de+usuario+del+Sistema+Tormes.&hl=es&gl=es)

[CIE-10]

Clasificación Internacional de Enfermedades.

[http://www.sssalud.gov.ar/hospitales/archivos/cie\\_10\\_revi.pdf](http://www.sssalud.gov.ar/hospitales/archivos/cie_10_revi.pdf)

[Collins C.C., 1985]

Collins C. C. 1985. On mobility aids for the blind. Eds. D. H. Warren & E. R. Strelow, *Electronic spatial sensing for the blind*, pp. 35-64

[DBSV]

Asociación alemana de ciegos y discapacitados visuales

<http://www.dbsv.org/infothek/zahlen-und-fakten/>

[Douglas G. et al., 1997]

Douglas G., Tobin M., Bozic N., Hill E. 1997. MOBIC: a system for facilitating independent mobility and navigation for blind people. *Proceedings of ICEVI's Xth World Conference, Sao Paulo, Brazil*, 3-8

[Dunai L. et al., 2009]

Dunai L., Peris F G., Defez B. G., Ortigosa A.N. 2009. Acoustical Navigation System for Visual Impaired People. *LivingAll European Conference*

[Dunai I. et al., 2009]

Dunai L., Peris F G., Defez B. G., Ortigosa A.N., Brusola S F. 2009. Perception of the sound source position. *Applied Physics Journal*, Vol. 55, No. 3, 448-451

[Dupress J. K., 1963]

Dupress J. K., 1963. The requirements for successful travel by the blind. In L. L. Clark Eds. 1, pp. 7-11

[ELD, 1995]

<http://www.nfbcal.org/nfb-rd/0772.html>

[Farmer L. W., 1975]

Farmer L. W., 1975. Travel in adverse weather using electronic mobility guidance devices. *The new outlook for the blind*, Vol. 69, No. 10, pp. 433-439

[Farmer L. and Smith D., 1997]

Farmer L., Smith D., 1997. Adaptive Technology. Book chapter, *Foundations on Orientation and Mobility*, Eds. 2, pp. 231-260

[Farmer W., 1978]

Farmer W. 1978. Mobility devices. *Bulletin of Prosthetic Research*, pp-47-118

[Fay R. R. and Popper A. N., 2005]

Fay R.R., Popper A.N., 2005. Introduction to sound source localization, *Springer Handbook of Auditory research, sound source localization*, pp. 1-5

[Felzenszwalb P. and Huttenlocher D., 2006]

Felzenszwalb P. and Huttenlocher D., 2006. Efficient belief propagation for early vision. *Int. J. Comput. Vision*, Vol. 70, No. 1, pp. 41–54

[Flea2 Firewire camera]

[http://www.ptgrey.com/products/flea2/flea2\\_firewire\\_camera.asp](http://www.ptgrey.com/products/flea2/flea2_firewire_camera.asp)

[Foulke E., 1971]

Foulke E. 1971. The perceptual basis for mobility. *AFB Res. Bull.* Vol. 23, pp. 1-8

[Freiberger H. A. M., 1974]

Freiberger H. A. M., 1974. Mobility Aids for the Blind. *Bulletin of Prosthetic Research*, pp. 73-78

[Hoyle B. S., 2003]

Hoyle B. S., 2003. The Bathcane – mobility aid for the vision impaired and the blind. *IEE Symposium on Assistive Technology*, pp. 18-22

[HumanWare]

“Braille Note GPS”, [www.humanware.com](http://www.humanware.com)

[Kelly G. W., 1981]

Kelly G.W., 1981. Sonic Orientation and Navigational Aid (SONA). Bulletin of Prosthetics Research. Vol. 28, No. 1, pp. 189.

[Kistler D. J. et al., 1996]

Kistler D.J., Wightman F.L., 1996. A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction a, b). J. Acoust. Soc. Am. Vol. 91, No. 3, pp. 1637-1647

[Klatzky R. L. et al, 2006]

Klatzky R. L., Marston J. R., Guidice N. A., Golledge R. G., Loomis J. M. 2006. Cognitive load of navigating without vision when guided by virtual sound versus spatial language. Journal of Experimental Psychology: Applied. Vol. 12, No. 4, pp. 223-232

[Kuc R., 2002]

Kuc R., 2002. Binaural sonar electronic travel aid provides vibrotactile cues for landmark, reflector motion and surface texture classification. IEEE Transactions on Biomedical Engineering, Vol. 49, No 10, pp. 1173-1180

[Lagace Y., 2005]

Lagace Y., April 2005. HumanWare's Trekker GPS system creates handheld personal guide with HP iPAQ Pocket PC. HP notice

[Loadstone, 2008]

The Loadstone GPS team. November 2008. Open Source Software Helps to Turn a Smartphone into Talking GPS. <http://www.loadstone-gps.com>

[Loomis, 1985]

[http://www.geog.ucsb.edu/pgs/papers/loomis\\_1985.pdf](http://www.geog.ucsb.edu/pgs/papers/loomis_1985.pdf)

[Loomis J. M. et al, 2001]

Loomis J. M., Golledge R. G., Klatzky R. L. 2001. GPS-based navigation systems for the blind. In W. Barfield & T. Caudell Eds. Fundamentals of wearable computers and augmented reality. Mahwah, NJ: Erlbaum. pp. 429-446

[Loomis J. M. et al, 1998]

Loomis J.M., Golledge R.G., Klatzky R.L., 1998. "Navigation system for the Blind: Auditory display modes and guidance", *Journal of Presence*, Vol.7, No.2 pp.193-203

[Loomis J.M. et al, 2005] Loomis, J. M., Marston, J. R., Golledge, R. G., & Klatzky, R. L. (2005). Personal guidance system for people with visual impairment: A comparison of spatial displays for route guidance. *Journal of Visual Impairment & Blindness*, 99, 219-232

[Loomis J.M. et al, 2006] Loomis, J. M., Golledge, R. G., Klatzky, R. L., & Marston, J. R. (2006). Assisting wayfinding in visually impaired travelers. In G. Allen (Ed.), *Applied spatial cognition: From research to cognitive technology*. Mahwah, N. J.: Lawrence Erlbaum Associates.

[Magnus, 1952]

Magnus R. Hestenes, Eduard Stiefel. *Methods of Conjugate Gradients for Solving Linear Systems*. *Journal of Research of the National Bureau of Standards*, Vol. 49, No. 6, December 1952

[Makino H. et al., 1996]

Makino, H., Ishii, I., & Nakashizuka, M., October 31–November 3 1996. Development of navigation system for the blind using GPS and mobile phone connection. In *Proceedings of the 18th anual meeting of the IEEE EMBS Amsterdam, The Netherlands*.

[Mann R.W., 1970]

Mann R.W., 1970. Mobility aids for the blind—An argument for a computer-based, man-device-environment, interactive, simulation system. *Pro Conf on Evaluation of Mobility Aids for the Blind*, Washington, DC: Com. on Interplay of Eng. with Biology and Medicine, National Academy of Engineering. pp. 101–16.

[Meijer P., 1992]

Meijer P. B. L., 1992. An experimental system for auditory image representations. *IEEE Transaction Biomedical Engineering*, BME-39, (2), pp. 112-121

[Meijer P., 2005]

Meijer P. B. L., 2005. A modular synthetic vision and navigation system for the totally blind. World Congress Proposals

[Mora J. L. et al, 2006]

Mora J.L. G., Rodriguez-Hernandez A.F., Martin F., Castellano M.A., 2006. Seeing the world by hearing: virtual acoustic space (VAS) a new space perception system for blind people. Proceedings of the 2nd Information and Communication Technologies Conference, ICTTA'06, IEEE. pp. 837-842

[Morrissette D. L. et al, 1971]

Morrissette D. L. Goddrich G.L., Henesey J.J., 1971. A follow-up-study of the Mowat sensors applications, frequency of use and maintenance reliability. J. Vis. Impairment Blindness, Vol. 75, No. 6, pp. 244-247

[NRC, 1972]

NRC (1972), The cane as a mobility aid for the blind. A report of a conference. September 10-11, 1971

[Nye P. and BlissJ, 1970]

Nye, P. W. and Bliss J. C., 1970. Sensory aids for the blind : A challenging problem with lessons for the future. Proc. of the IEEE, Vol. 58, No. 12, pp. 1878-1898

[OED]

Observatorio Estatal de la Discapacidad

<http://www.observatoriodeladiscapacidad.es>

[OED inf]

[http://sid.usal.es/idocs/F8/FDO25888/Informe\\_contratacion\\_PCD\\_marzo\\_2011.pdf](http://sid.usal.es/idocs/F8/FDO25888/Informe_contratacion_PCD_marzo_2011.pdf)

[OMS Nota descriptiva]

<http://www.who.int/mediacentre/factsheets/fs282/es/index.html>

[Petrie H. et al, 1997]

Petrie H., Johnson V., Strothotte T., Raab A., Michel R., Reichert L., Schalt A., 1997. MoBIC: An aid to increase the independent mobility of blind travellers. *British Journal of Visual Impairment*, Vol. 15, No. 2, pp. 63-66

[Russell, 1965]

Russell L., 1965. Travel path sounder. Proceedings, Rotterdam Mobility Research Conference. N.Y: American Foundation for the Blind

[Sainarayanan G. et al, 2007]

Sainarayanan G., Nagarajan R., and Yaacob S., Jan. 2007. Fuzzy image processing scheme for autonomous navigation of human blind. *Appl. Soft Comput.* Vol. 7, No. 1 pp. 257-264.

[Ortigosa Araque N. et al, 2009]

Ortigosa Araque N., Dunai L., Peris Fajarnes G., Santiago Praderas V., Dunai I., 11 December 2009. Sound map generation for a prototype blind mobility system using multiple sensors. ABLETECH'08 Conference

[Santiago Praderas V. et al, 2009]

Santiago Praderas V., Ortigosa N., Dunai L., Peris Fajarnes G. 2009 10-12 June. Cognitive Aid System for Blind People (CASBlIP). Proceedings of XXI Ingehrat-XVII ADM Congress pp. 31

[Scharstein, D. and Szeliski, R., 2002]

Scharstein, D. and Szeliski, R. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int. J. Comput. Vision*, Vol. 47, No.1-3, pp.7-42

[Shi, J. and Tomasi, C., 1994]

Shi, J. and Tomasi, C., 1994. Good features to track. In CVPR 94 Conference. June

[Shoval S. et al., 2000]

Shoval S., Ulrich I., Borrenstein J. 2000. Computerized obstacle avoidance systems for the blind and visually impaired. Invited chapter in *Intelligent systems and technologies in rehabilitation engineering*. Editors Teodorescu, H.N.L & Jain, L.C., CRC, pp. 414-448

[Hannuna S., 2007]

Hannuna S., 2007. Quadruped gait detection in low quality wildlife video. PhD thesis, University of Bristol

[Starkiewicz W. and Kuliszewsky T., 1963]

Starkiewicz W., Kuliszewski T., 1963. Active energy radiating system: the 80-channel Electroftalm. Proceedings Int. Congress on Technology and Blindness, Vol. 1, pp. 157-166

[Strothotte T. et al, 1997]

Strothotte T., Petrie H., Johnson V., Reichert L., 1997. MoBIC: user needs and preliminary design for a mobility aid for blind and elderly travellers. British Journal of Visual Impairment, Vol. 15, No. 2, pp. 63-66

[Torr, P. and Zisserman Mlesac, A., 2000]

Torr, P. and Zisserman Mlesac, A., 2000. A new robust estimator with application to estimating image geometry. Comp. Vis. and Image Understanding, Vol. 78, pp. 138-156

[Vadakkepat P. et al, 2008]

Vadakkepat P., De Silva L.C., Jing L, Ling L.L., 2008, March. Multimodal Approach to Human-Face Detection and Tracking. IEEE Trans. on Industrial Electronics, Vol. 55, No. 3, pp. 1385-1393

[Versenyi G., 2007]

Versenyi G., 2007. Localization in a Head-Related Transfer Function-based virtual audio synthesis using additional high-pass and low-pass filtering of sound sources. Acoust. Science & Technology, Vol. 28, No. 4, pp. 244-250

[WBU 2009]

White cane safety day. World Blind Union, Press release, October 9, Canada

[Wilson J. et al, 2007]

Wilson J., Walker B. N., Lindsay J., Cambias C., Dellaert F., 2007. SWAN: System for wearable audio navigation. 11th IEEE International Symposium on, pp. 91-98

[Zitnick C. and Kanade T., 2000]

Zitnick C. and Kanade T. 2000. A cooperative algorithm for stereo matching and occlusion detection. IEEE Trans. PAMI, Vol. 22, No. 7, pp. 675–684

[Zitnick, C and Kanade, T.]

A cooperative algorithm for stereo matching and occlusion detection. IEEE Trans. PAMI, 22(7):675–684, 2000.