

# **UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA**

## **CODIFICACIÓN DEL CONTORNO DE LA IMAGEN VISUAL A PATRÓN ACÚSTICO Y SU DECODIFICACIÓN REALIZADA POR INVIDENTES**

TESIS

PARA OBTENER EL GRADO DE:

**DOCTOR EN ELECTRÓNICA Y COMPUTACIÓN**

PRESENTA:

**JOSÉ RODOLFO MARTÍNEZ Y CÁRDENAS**

DIRECTORES DE TESIS

**DRA. VIRGINIA BERRÓN LARA**

**DR. FELIPE ROLANDO MENCHACA GARCÍA**

HUAJUAPAN DE LEÓN, OAXACA, DICIEMBRE DE 2011.



Tesis presentada en Diciembre de 2011  
ante los sinodales:

Dr. Salvador Álvarez Ballesteros

Dr. Pablo Lizana Paulin

Dr. Carlos Martínez Álvarez

**Directores de tesis:**

Dra. Virginia Berrón Lara

Dr. Felipe Rolando Mechaca García



Dedico este esfuerzo, el mayor de mi vida a mi compañera, amiga y amada esposa, Lupita, quien me ha apoyado en todos los desafíos que me he propuesto en mi vida. Quien ha soportado todos los sinsabores de mis días de ausencia física y mental. A quien en ocasiones le he dejado toda la responsabilidad del hogar y los hijos. A ella quien siempre estuvo presente y creyendo en mi.

A mis hijos, Jennifer, Rachman y Luzita, a quienes en momentos importantes de su vida no los acompañé debido al desafío que me propuse, pero finalmente les puedo decir que sí se pueden lograr las metas que nos imponemos.

A mis nietos: Rebeca, Arturo Rodolfo, Román, Sarah Fernanda, Lourdes Guadalupe y Rodolfo, les dejo este testimonio de la vida y trayectoria de su abuelo, quien los ama.

A mis amigos por brindarme su sincera e incondicional amistad; porque con ello han llenado en gran parte mi existencia cotidiana. Sabemos, que los amigos apenas los podemos contar con los dedos de nuestras manos, para esos pocos y que me conocen, les dedico y comparto esta vivencia de mi esfuerzo personal.



## **Agradecimientos**

Agradezco a la Universidad Tecnológica de la Mixteca por haberme dado la oportunidad de llevar a cabo mis estudios de Doctorado en Electrónica y Computación.

Al Dr. Modesto Seara Vázquez, Rector de la Universidad Tecnológica de la Mixteca, por haber aceptado mi ingreso a tan prestigiada Casa de Estudios.

Al Ing. Gerardo S. García Hernández, Vice-Rector Académico de la UTM durante todos estos años, gracias por su amistad, apoyo y paciencia.

Al Dr. Felipe Rolando Menchaca García por haber compartido conmigo su visión de la Ingeniería y su amplio conocimiento en Acústica; por haber creído en mí y en este proyecto. Además, de brindarme su confianza y amistad que estoy seguro que perdurará por mucho tiempo.

A la Dra. Virginia Berrón Lara, le agradezco que haya creído siempre en mí; estuvo siempre a mi lado durante todo el desarrollo de mi trabajo doctoral, y además vivió junto conmigo las angustias y los logros. Gracias a este apoyo incondicional logramos finalizar este proyecto de investigación.

Al Ing. Moisés Ramírez con quien logramos implementar el algoritmo basado en conceptos de Topología Digital, lo que nos permitió lograr nuestro objetivo de obtener el contorno de la imagen.

Al Dr. Santiago Omar Caballero Morales, profesor de la materia "Metodología de la Investigación" por su orientación en la redacción de la tesis.

Al Dr. Ramón Brena Pinedo, con quien inicié en el año de 1982 las primeras teorías del Videófono.

Al Dr. Carlos Martínez Álvarez, en particular por su ayuda en la comprensión de los conceptos de aprendizaje y construcción del conocimiento.

A los revisores que hicieron observaciones de vital importancia para esta tesis: Dr. Salvador Álvarez Ballesteros, Dr. Pablo Lizana Paulin, Dr. Maximino Peña Guerrero, Dr. Juan Manuel Ibarra Zannatha y Dr. Ricardo Barrón Fernández.

A la M. en C. María de Lourdes Robles Martínez, gran amiga y compañera, le agradezco por revisar la estructura y estilo del primer borrador.

Al M. en C. Fidel Diego Nava, amigo y compañero quien en su momento me animó para llevar a cabo mi doctorado, y siempre me incitó a obtener mi grado.

Al Dr. Cirenio Escamirosa Tinoco, exdirector del CIIDIR Unidad Oaxaca, durante cuya gestión obtuve el apoyo para iniciar mis estudios de doctorado.

Así mismo, les estoy muy agradecido a todos aquellos catedráticos de la UTM que intervinieron en mi formación académica, por su paciencia y el brindarme sin reserva lo mejor de sí mismos. De todos ellos me llevo un recuerdo que perdurará durante toda mi existencia.

Son muchos los amigos y compañeros que contribuyeron a que no haya flaqueado en el camino, por cierto largo y difícil: ¡A todos ellos mil gracias!

Agradezco tanto al Instituto Politécnico Nacional como al ANUIES por todo el apoyo recibido.

# Índice

Índice .....	I
Índice de Figuras.....	V
Índice de Tablas .....	VII
Resumen .....	IX
Abstract .....	XI
Capítulo 1.....	1
Introducción .....	1
1. Motivación del trabajo .....	1
2. Planteamiento y justificación del problema.....	5
3. Hipótesis .....	6
4. Objetivo general .....	7
5. Objetivos particulares.....	7
6. Delimitaciones .....	8
7. Metodología .....	9

---

8. Reseña de los capítulos .....	10
Capítulo 2.....	13
Marco Teórico (Estado del arte).....	13
1. Revisión crítica de trabajos y propuestas anteriores .....	13
Capítulo 3.....	29
Técnicas utilizadas .....	29
1. Introducción.....	29
2. La forma, los contornos visuales y los objetos. ....	29
3. El contorno y la comunicación.....	30
4. Técnicas usadas para el procesamiento de imágenes.....	33
5. Lenguaje de programación CSound.....	42
Capítulo 4.....	51
Desarrollo de la propuesta .....	51
1. El Modelo propuesto de conversión de imagen a sonido .....	51
2. El taller .....	57
Capítulo 5.....	61
Experimentos y resultados .....	61
1. Procesamiento de la imagen .....	61
2. Sistema básico experimental propuesto .....	64
3. Cuadrados latinos y estadística .....	66
4. Taller de capacitación .....	66
5. Resultados de las pruebas .....	70
Capítulo 6.....	75
Discusión, Conclusiones y Trabajo Futuro.....	75
1. Discusión.....	75
1.1. Procesamiento de la imagen .....	76
1.2. Diseño del patrón acústico .....	80
1.3. Experimentación con seres humanos.....	85

---

2. Conclusiones.....	89
3. Contribuciones.....	91
4. Trabajo futuro.....	92
Referencias .....	95
Anexo I: Lista de publicaciones y presentaciones en congresos .....	103
Congresos Nacionales.....	103
Congresos Internacionales .....	103
Publicación en Revistas Nacionales.....	104
Publicación en Revistas Internacionales .....	104
Anexo II: Figuras y patrones acústicos .....	125
Anexo III: Guión del curso de capacitación e instructivo para participantes.....	127
I. Unidad didáctica inicial: Sensibilización táctil .....	127
II. Unidad de entrenamiento acústico.....	128
III. Unidad de decodificación de patrones acústicos .....	129
IV. Unidad de evaluación.....	129
V. Recursos materiales .....	130
VI. Instructivo para el participante.....	131
Anexo IV. Topología Digital.....	133
Anexo V Ejemplos de archivos orquesta y partitura .....	139



## Índice de Figuras

Figura 1.1. Arquitectura de las propuestas de dispositivos de sustitución de la visión por el oído. ....	5
Figura 2.1. Se muestra el principio de barrido de la imagen visual y su transformación a patrón acústico. Tomada de [Meijer, 1992]. ....	21
Figura 2.2. Representación tridimensional de una imagen. Tomada de [The vOICe at CNS, 2005]. ....	22
Figura 3.1. Contornos básicos, imagen tomada de [Dondis, 2002]. ....	31
Figura 3.2. Un cuadrado, imagen tomada de [Dondis, 2002]. ....	31
Figura 3.3. Un círculo, imagen tomada de [Dondis, 2002]. ....	32
Figura 3.4. Un triángulo equilátero, imagen tomada de [Dondis, 2002]. ....	32
Figura 3.5. Con el círculo, cuadrado y triángulo, podemos obtener cualquier imagen, imagen tomada de [Dondis, 2002]. ....	33
Figura 3.6. Convenio de ejes utilizado para la representación digital de imágenes (Fotografía tomada por el autor). ....	35
Figura 3.7. Operación individual, imagen tomada de [Pajares, 2004] ....	36
Figura 3.8. Aquí se muestran los puntos encendidos (en negro), imagen tomada de [Martínez, 1984]. ....	38
Figura 3.9. Tramo de línea horizontal, imagen tomada de [Martínez, 1984]. ....	39
Figura 3.10. Imagen formada por una pendiente positiva, imagen tomada de [Martínez, 1984]. ....	39
Figura 3.11. Imagen formada por puntos que forman una matriz rectangular, imagen tomada de [Martínez, 1984]. ....	40
Figura 3.12. Las 50 frecuencias correspondientes a las 50 notas del piano escogidas, imagen tomada de [Tremaine, 1969]. ....	41
Figura 3.13. Encabezado del archivo CSound [Boulangier, 2000]. ....	43

---

Figura 3.14. Código orquesta para el instrumento 101. ....	45
Figura 3.15. Diagrama a bloques del <i>inst 101</i> [Boulanger, 2000]. ....	46
Figura 4.1. Se muestra el principio de barrido de la imagen visual y su transformación a patrón acústico del sistema de conversión de imagen a sonido, figura elaborada por el autor. ....	52
Figura 4.2. Diagrama representativo de todos los algoritmos utilizados en este estudio...	54
Figura 4.3 Interface gráfica para convertir el contorno a patrón acústico .....	56
Figura 4.4 Se observa cómo se lleva a cabo la asociación del patrón acústico con la percepción táctil. ....	59
Figura 5.1 Se presenta el procesamiento de la imagen siguiendo los algoritmos propuestos.....	62
Figura 5.2 Ejemplificación del resultado de algunas imágenes procesadas con los algoritmos propuestos.....	63
Figura 5.3 Contornos de las imágenes sobre las tarjetas.....	69

# Índice de Tablas

Tabla 4.1. Asignación de las frecuencias en la columna <i>l</i> .....	55
Tabla 5.1. Sistema básico propuesto de 33 figuras geométricas, el cual representó los patrones acústicos utilizados en nuestros experimentos. ....	65
Tabla 5.2. Porcentaje de identificación de los patrones acústicos. ....	71
Tabla 5.3. Dibujos realizados por los invidentes interpretando los patrones acústicos. Primera parte.....	72
Tabla 5.4. Dibujos realizados por los invidentes interpretando los patrones acústicos. Segunda parte.....	73
Tabla 6.1 Análisis comparativo de las diferentes etapas de procesamiento de la imagen.	78
Tabla 6.2 Tamaños de la imagen a procesar, según la línea de investigación. ....	79
Tabla 6.3 Número de tonalidades de gris utilizados en la imagen a sonificar. ....	80
Tabla 6.4 Diseño del Patrón acústico. ....	82
Tabla 6.5 Tolerancia al ruido .....	85
Tabla 6.6 Experimentación con seres humanos.....	86
Tabla 6.7 Imágenes utilizadas en las pruebas .....	87
Tabla 6.8 Ambiente de experimentación .....	88
Tabla 6.9 Duración del entrenamiento.....	89



# Resumen

En el diseño de un dispositivo de sustitución de la visión por la audición usando una cámara fotográfica digital, una de las preguntas clave es construir un esquema de codificación de la imagen a sonido que sea “de fácil decodificación” por parte de los invidentes.

En este trabajo se propone un esquema de codificación de este tipo para ayudar a los invidentes a obtener información sobre la forma que poseen los objetos. Se le ha llamado “Visor Acústico”. Es diferente a las propuestas presentadas anteriormente pues toma en cuenta ideas de la Psicología Cognitiva. Se describen los algoritmos de procesamiento digital de la imagen utilizados para la extracción del contorno y los de su conversión a sonido mediante el uso de CSound.

Para evaluar la idoneidad de la propuesta se seleccionaron 33 contornos (básicos) formados por líneas rectas y curvas y sus respectivos patrones acústicos, se diseñó e implementó un curso de entrenamiento en su decodificación como parte de un experimento en el que participaron dos grupos de personas, uno de invidentes y el otro de sujetos con visión normal, se les entrenó y evaluó su capacidad adquirida usando los cuadrados latinos. Los resultados muestran que ambos grupos desarrollaron la habilidad de decodificar los patrones en igual medida, apoyando la idea de que el “Visor Acústico” podrá ayudar a los invidentes en un futuro.

Palabras clave: patrones acústicos y ceguera, prótesis no invasivas para personas ciegas, sustitución sensorial, visión acústica.



# Abstract

In order to design a sensory substitution device, which would substitute vision through audition, a key issue is the definition of a coding scheme of (captured by a digital photographic camera) images into acoustic pattern, such that they are easily decoded by blind or visual impairment people.

In this dissertation a coding scheme for digital images into acoustic patterns is proposed, which eventually could be part of a non-invasive prosthesis, which would substitute vision by audition. This coding, called "Acoustic Viewer", is based on a conceptually different approach from previous papers' method, partially because it is based on Cognitive Psychology. The algorithms for the contour extraction and for the conversion to sounds by using CSound are presented here.

Thirty-three geometric figures, each formed by straight lines and curves and their respective acoustic patterns, were selected to assess the suitability of the proposed coding scheme. A training course specially designed was implemented as part of experiment involving two groups of people: one of eight blind subjects and the other with eight individuals with normal sight. Their acquired ability was assessed using latin squares. The results show that both groups developed the ability to decode the patterns in an equal measure, supporting the idea that the "Acoustic Viewer" may help the blind people to recognize many objects.

Key words: sensory substitution, sonification, Acoustic Viewer, CSound and acoustics patterns, acoustic vision.



# Capítulo 1

## Introducción

### 1. Motivación del trabajo

La Organización Mundial de la Salud (dependiente de la ONU), OMS, ha estimado que hay 161 millones de personas visualmente incapacitadas en el mundo, de las que 124 millones tienen poca visión y 37 millones son ciegas [OMS, 2009]. De acuerdo con el Instituto Nacional de Estadística, Geografía e Informática (INEGI) en México hay aproximadamente 600 000 habitantes con algún tipo de discapacidad visual [Cano, 2011], [Aponte, 2009].

La mayoría no son económicamente activos [Goertz, 2010], tienen problemas para tener acceso a la educación, nuevas tecnologías y la cultura en general<sup>1</sup>. Ver la introducción de [Sánchez, 2010] y [Aycinema, 2008].

---

<sup>1</sup> Por ejemplo en el estado de Oaxaca, hasta donde sabemos, hay una sola biblioteca pública para ciegos.

En México los invidentes solamente tienen acceso a unos pocos elementos de ayuda, como son: El sistema Braille, los perros guías y el bastón blanco. El Sistema Braille es un sistema de escritura que consiste en signos dibujados en relieve para poder leer con los dedos.<sup>2</sup> Los perros guías y el bastón blanco los ayudan a tener mucha más movilidad, pero no todos pueden conseguir un perro<sup>3</sup> y el bastón les ayuda a sentir los obstáculos en un rango de más o menos 1 metro de distancia. Para mayores distancias dependen de otras personas y de su sentido del oído, de acuerdo a la definición de espacio extra personal de Hall<sup>4</sup> [Strumillo, 2010]. Debido a que la cantidad de ciegos en el mundo está en crecimiento a causa de, entre otras cosas, la mayor expectativa de vida, tenemos un grupo de personas a las que valdría la pena pensar en ayudar [Greenemeir, 2008].

Investigadores en diferentes partes del mundo han trabajado y están trabajando en propuestas para ayudarlos<sup>5</sup>. Estas propuestas se dividen en invasivas<sup>6</sup> y no invasivas. Una revisión de algunas propuestas invasivas se puede encontrar en [Weiland, 2008]: implantes retínales, implantes en el nervio óptico, implantes en la corteza cerebral han ayudado a que los ciegos reconozcan objetos simples y detecten movimiento, pero todavía falta mucho por hacer para que puedan reconocer caras y leer.

---

<sup>2</sup> Definición tomada de la versión en línea del Diccionario de la Real Academia Española, (<http://www.wordreference.com/es/en/frames.asp?es=braille>, consultada el 24 de septiembre de 2011).

<sup>3</sup> En comunicación privada, el Dr. V. A. Osorio, entrenador de perros, nos informó que no todos los perros de una camada sirven para ser guías; el costo de entrenamiento de un perro guía es de alrededor de \$50,000 M. N., su "vida útil" es de unos 6 años, y su entrenamiento los capacita para un determinado vecindario.

<sup>4</sup> Para distancias menores o iguales que un metro, llamadas por Hall espacio personal, el ser humano depende de los sentidos de la visión, el tacto, el oído y olfato; para el espacio cercano, que comprende distancias entre 1 y 4 metros dependemos de la visión y del oído, y para el espacio lejano también.

<sup>5</sup> EUA, Inglaterra, Holanda, Canadá, Japón, Italia, Malasia y China, entre otros, ver [Liu, 2010].

<sup>6</sup> Procedimiento Invasivo es aquel procedimiento por el cual el cuerpo es "invadido" o penetrado con una aguja, una sonda, un dispositivo o un endoscopio. (<http://www.wordreference.com/es/en/frames.asp?es=invasivo>, consultada el 24 de septiembre de 2011).

Con respecto a las no invasivas, la mayoría comparten la idea prevaleciente desde el siglo XVIII que los ciegos (de nacimiento) tenían el sentido del tacto y del oído mucho más desarrollados para compensar la falta de la visión<sup>7</sup>, [Morgan, 1999]. Por lo tanto las dos principales líneas de estudio del enfoque no invasivo se centran en el uso del tacto o en la audición; aunque recientemente se empieza a experimentar con una mezcla de los dos [Edward, 2006].

El ya fallecido investigador norteamericano Paul Bach-y-Rita es uno de los pioneros en las investigaciones con el enfoque no invasivo, en particular en la línea que se centra en el tacto [Bach-y-Rita, 1998], [Tyler, 2003],y [Sampaio, 2001].

Otro grupo de investigadores privilegia la audición como canal alternativo para “proveer algún equivalente de la visión” a los invidentes. Uno de los pioneros más reconocidos mundialmente es Peter B. L. Meijer [Meijer, 1992]. Las razones que se esgrimen para usar el sentido del oído en “sustitución” de la visión son, entre otras, que el sistema auditivo humano es capaz de manejar patrones de sonido complejos que cambian rápidamente incluso en un ambiente ruidoso, tiene muy buenos umbrales de discriminación de frecuencias e intensidad, además de que la interfaz de salida sería muy simple, audífonos [Auvray, 2007].

A grandes rasgos las propuestas de la sustitución de la visión por el oído tienen la arquitectura mostrada en la Figura 1.1<sup>8</sup>. La arquitectura propuesta consta de 5 fases: 1.- Dispositivos para la adquisición de la imagen y la adquisición de la misma. 2.-

---

<sup>7</sup> Aunque hasta hace poco, más o menos desde finales del siglo XX, se empezó a analizar a profundidad esta aseveración, por ejemplo: [Peretz, 2003], [Rauschecker, 2003].

<sup>8</sup> La primera mención de una arquitectura de este tipo la encontramos en [Capelle, 1998]. La aquí presentada es una generalización de ésta.

Procesamiento de la imagen. 3.- Sonificación de la imagen. 4.- Generación del sonido y 5.- Interfaz de salida. La parte “dispositivos para adquisición de la imagen y la adquisición de la misma” depende de la elección del dispositivo de entrada de la imagen; aquí se resuelven los problemas directamente relacionados con la adquisición de la imagen, (muestreo de imagen, cambio de formato, etc.), la etapa “procesamiento de imagen” busca preparar la imagen para su sonificación, puede contener algoritmos para recorte y limpieza de la imagen, los algoritmos para encontrar el “objeto central”, etc.

La etapa “sonificación de la imagen” es la que transforma la información de los píxeles de la imagen obtenida en los datos necesarios para la generación del sonido. La etapa “generación del sonido”, como su nombre lo indica se preocupa por los problemas relacionados con la generación del sonido. La interfaz de salida son los audífonos.

Todas las propuestas tienen en común que buscan que en la señal de salida esté codificado lo que se ve en la imagen o en parte de ella. Constan de desarrollos de hardware y software. En este trabajo nos restringiremos, porque nos parece prioritario, a buscar una solución a algunos de los problemas que plantearía el diseño del software de una propuesta en esta dirección. De estos problemas nos interesan los relacionados con la codificación de imagen a sonido. Por una codificación entenderemos un proceso por el que la información de la imagen es transformada en un patrón sonoro, aunque esto conlleve la pérdida de información contenida en la imagen original. Esta codificación involucra a las etapas de procesamiento de la imagen (parcial o totalmente), sonificación de la imagen y generación del sonido (total o parcialmente).

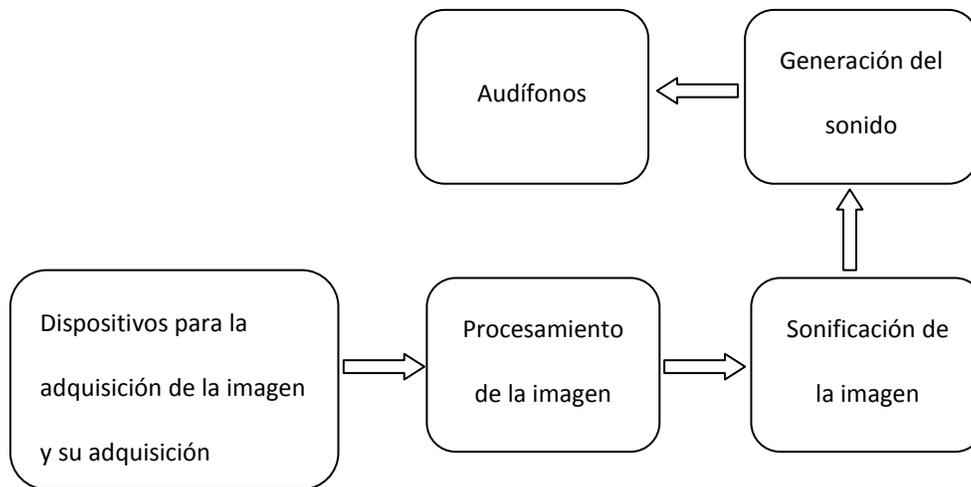


Figura 1.1. Arquitectura de las propuestas de dispositivos de sustitución de la visión por el oído.

En el siguiente capítulo mencionaremos las codificaciones de imagen a sonido más importantes en la literatura científica. En la primera parte del último capítulo haremos un análisis crítico de las tres más citadas junto con la que hemos desarrollado. En general se puede decir que los tiempos de aprendizaje son de algunos meses; los usuarios admiten tener dificultades para decodificar las señales; daremos más detalles de esto en los Capítulos 2 y 6 de este trabajo. La evaluación del aprendizaje de los patrones básicos de cada propuesta no parece hacerse de manera muy rigurosa (en trabajos de hasta el 2005). A partir del análisis de la bibliografía encontramos que hace falta una propuesta que contemple la manera en la que reconocemos formas. Dicha propuesta facilitaría el aprendizaje y reconocimiento de patrones acústicos. También hemos observado que hace falta llevar a cabo una evaluación más rigurosa de sus resultados.

## 2. Planteamiento y justificación del problema

Nuestra meta es definir un *modelo* de una codificación de imagen a sonido (patrón acústico) que sea de más fácil aprendizaje y reconocimiento por parte de los invidentes;

además, diseñar un curso para el aprendizaje de los patrones generados y evaluar su aprovechamiento de una forma rigurosa. Con esto, probaríamos la factibilidad de nuestro modelo.

La codificación buscada vendría a formar parte del software de un dispositivo de sustitución sensorial, al que le llamaremos “Visor Acústico”. En esta etapa del proyecto el nombre de tal dispositivo será el nombre de la codificación. Este nombre lo hemos escogido con la intención de no olvidarnos de nuestro objetivo personal a largo plazo: el diseño y construcción de un dispositivo de sustitución sensorial.

Pensamos que la solución exitosa de este problema sería decisiva para la continuación del proyecto de diseño de este dispositivo. La novedad de nuestra propuesta será mostrada en el Capítulo 6.

### **3. Hipótesis**

Se plantean las siguientes hipótesis<sup>9</sup> de investigación:

1. Los usuarios potenciales del Visor Acústico serán personas visualmente discapacitadas cuyos ojos o las conexiones neurológicas directas de éstos con el cerebro no funcionan (adecuadamente) pero cuyo sentido del oído funcione normalmente.
2. Los objetos situados enfrente del usuario son los objetos relevantes para él.

---

<sup>9</sup> Hipótesis.- Del lat. Hypothesis: 1. f. Suposición de una cosa posible o imposible para sacar de ella una consecuencia; hipótesis de trabajo 2. La que se establece provisionalmente como base de una investigación que puede confirmar o negar la validez de aquella.

3. El contorno del objeto relevante es invariante bajo la conversión de una imagen de color a una imagen en tonos de gris. Nos concretaremos a trabajar imágenes en tonos de gris<sup>10</sup>.
4. Los seres humanos reconocen a los objetos por su contorno<sup>11</sup>.
5. Los usuarios pueden (aprender a) reconocer objetos a partir del patrón acústico asociado a su contorno.
6. Para ayudar a la memorización de los patrones acústicos usaremos las frecuencias de algunas teclas del piano, pues se sabe que los sonidos musicales son más fáciles de memorizar [Trehub, 2003].
7. Sólo usaremos monovisión<sup>12</sup>.

#### **4. Objetivo general**

Proponer un modelo de una codificación de imagen a sonido a partir del objeto en primer plano visual usando frecuencias de notas musicales, estudiar su factibilidad y evaluar el aprendizaje de la codificación por un grupo de personas ciegas o vendadas, lo más rigurosamente posible.

#### **5. Objetivos particulares**

1. Encontrar un proceso para obtener el contorno de la imagen.

---

<sup>10</sup> Todos los investigadores sobre esta temática han elegido imágenes en tonos de gris, ver Capítulo 2 y 6 de este documento.

<sup>11</sup> Discusión en el Capítulo 6 de este documento.

<sup>12</sup> Las codificaciones más citadas utilizan una sola cámara [Meijer, 1992], [Capelle, 1999] y [Sairayanan, 2007].

2. Encontrar una codificación de las imágenes en patrones acústicos usando las frecuencias de algunas notas del piano.
3. Con el fin de evaluar la idoneidad de la codificación encontrada en el punto anterior: Selección de un sistema básico de patrones<sup>13</sup> y diseño de un curso basado en éstos, su impartición a individuos discapacitados y la evaluación de su desempeño.

## 6. Delimitaciones

La información relevante para los ciegos es la de la imagen de los objetos que están justo enfrente de la cámara (o de ellos). Dicha cámara estaría situada en la cabeza, montada sobre una diadema o sobre unos lentes.

Si el ciego quiere ver algo que esté a los lados o hacia abajo, moverá la cabeza de manera correspondiente. Es decir, no se busca que el usuario tenga visión periférica, ni tampoco que “sense” en ambientes con poca luminosidad. Esta hipótesis subyace a todos los trabajos presentados hasta la fecha, pues las imágenes de sus publicaciones presentan a los ciegos en un espacio interior con buena iluminación o en una escena con luz de día (días soleados). Nosotros al igual que los investigadores anteriores, usaremos únicamente imágenes en tonos de gris<sup>14</sup>.

---

<sup>13</sup> Nosotros lo entendemos como un conjunto de imágenes de prueba. Si los ciegos pueden decodificarlo, entenderemos que vamos por buen camino. Al principio le llamamos alfabeto, pero decidimos llamarlo un sistema básico porque no estamos seguros de que sea el mínimo conjunto con el cual podríamos formar cualquier otra figura.

<sup>14</sup> En apoyo de nuestra decisión mencionaremos que el Dr. V. A. Osorio Cruz nos comentó que los ojos de los perros guías se vuelven los ojos del invidente y que los perros ven tonos de gris.

Los resultados de evaluaciones posteriores al año 2000 de vOICe<sup>15</sup> (el sistema propuesto por Meijer) hacen pensar que basta pedir monovisión. Por el momento buscaremos que los usuarios sean capaces de aprender a reconocer un conjunto básico de formas geométricas sencillas<sup>16</sup>.

## 7. Metodología

Primeramente se hará una recolección y análisis crítico de los trabajos en donde se documentan las codificaciones más importantes de imagen a sonido implementadas en dispositivos de sustitución sensorial de la visión por medio de la audición.

A continuación se determinará cuales son los elementos más importantes de la imagen relevantes para la codificación.

- Se implementará un proceso algorítmico para extraer la información relevante de la imagen con miras a usarla en la codificación.
- Se propondrá la estrategia para determinar el tipo de sonidos que se van a usar en la codificación.
- Se buscará encontrar la manera óptima de implementación de los sonidos encontrados para después hacer las primeras pruebas con posibles usuarios.

---

<sup>15</sup> Ir al Capítulo 6; ver [Ward, 2010].

<sup>16</sup> Los psicólogos afirman que reconocemos las formas más complejas, descomponiéndolas en partes más sencillas, [Hoffman, 1985].

- Seguidamente se diseñará la forma de entrenamiento y evaluación de un grupo de sujetos. Después de llevarse a cabo lo anterior, se evaluarán los resultados.

## **8. Reseña de los capítulos**

A continuación se hace un resumen del contenido de los siguientes capítulos de este documento.

En el Capítulo 2 se presenta el marco teórico del problema (el estado del arte), en donde se hace una presentación de los antecedentes de nuestra investigación

En el Capítulo 3 se hace una breve descripción de las técnicas de procesamiento de imágenes utilizadas, la estrategia de conversión de imagen a sonido y de la herramienta de software utilizada para llevarla a cabo

El Capítulo 4 corresponde al desarrollo de la propuesta. Se informa acerca de la secuencia de pasos que se siguió para alcanzar nuestro objetivo.

En el Capítulo 5 se presentan los detalles de los experimentos efectuados para llevar a cabo este trabajo de investigación.

En el Capítulo 6 se presenta una discusión rigurosa de los resultados obtenidos comparándolos con los de las codificaciones más importantes encontradas en la literatura, las contribuciones de nuestro trabajo y finalmente, mencionamos algunas de las direcciones de trabajo futuro.

---

Se presentan cinco anexos: la lista de las publicaciones y presentaciones en congresos que se hicieron sobre este trabajo, lo cual ha permitido ir estableciendo los avances que se han ido logrando a través del tiempo; se anexa además un disco compacto con: ejemplos de figuras y sus correspondientes patrones acústicos, una comparación entre algunos patrones acústicos generados por la propuesta de Meijer y los generados por el Visor Acústico, el guión del curso de capacitación y el instructivo para participantes y un archivo de texto de los archivos necesarios para que el software generador de los patrones acústicos que generan el patrón del cuadrado.



# Capítulo 2

## Marco Teórico (Estado del arte)

### 1. Revisión crítica de trabajos y propuestas anteriores

Nuestros sentidos juegan un papel muy importante en nuestra vida, siendo la vista el más importante de ellos. ¿Qué pasaría si lo perdiéramos? Basta con cerrar por algunos momentos nuestros ojos y que tratemos de deambular alrededor de nuestra posición actual para empezar a intuir los problemas a los que nos enfrentaríamos.

¿Qué se podría hacer por las personas discapacitadas visualmente en estas situaciones?

A través del tiempo se han propuesto prótesis de todo tipo, invasivas y no invasivas, con el fin de poder integrar a estos seres humanos al mundo cotidiano. Muchas propuestas tecnológicas se han realizado para ayudar a los invidentes a tener una mejor calidad de vida, proponiendo diferentes sistemas de ayuda: en [Baskett, 2005] se

presentan diferentes propuestas y técnicas que pueden ser usadas por los ciegos manejar sus miedos cuando están navegando en forma autónoma; en [Bialistock, 2005] se informa de una propuesta de técnicas que ayudan a los ciegos a tener un mejor lenguaje corporal lo que les facilitará la comunicación con las personas con visión normal; en [Cimarolli, 2006] se investiga que tipo de ayuda se les presta a los ciegos y visualmente discapacitados y cómo se les ayuda (en EUA); se concluye que la ayuda se restringe principalmente a ayudarlos a tener una mayor movilidad e independencia; [Hardwick, 1998] es un trabajo donde se habla sobre el uso de una interfaz háptica para proveer acceso a imágenes virtuales en tridimensionales; [Harris, 1993] presenta un sistema para ayudar a los ciegos y gente con visión defectuosa para encontrar y usar muchas máquinas públicas y sistemas de información; en [Lorenzo, 2005] se informa sobre un conjunto de lineamientos (claros) e información práctica para arquitectos, diseñadores y urbanistas para la planeación y diseño de ciudades, edificios, transportes, productos y servicios para seres humanos de todo tipo en especial invidentes; [Oberdofer, 2004] se sugieren diferentes posibilidades de líneas de investigación en la rehabilitación de ciegos y personas con poca visión, [Studebaker, 2004] describe brevemente la historia y evolución de los servicios de la rehabilitación de la visión en los EUA, entre otros. En otras publicaciones se informa del avance del desarrollo de las prótesis visuales con enfoques que incluyen metodologías neuro-quirúrgicas [Pezaris, 2007], [Pezaris, 2009].

En 1990, se desarrolló una prótesis visual para invidentes, [Cha, 1990], que estimulaba eléctricamente la corteza visual de sujetos con visión normal mediante un arreglo de electrodos. Ya que al estimular eléctricamente la corteza visual se evoca la

sensación de un chispazo de luz (un fosfene), se cree que con la implantación de un arreglo de electrodos en la corteza visual, se podría lograr crear una imagen (imagen de fosfenos) [Greenemeir, 2008]. Los implantes retinales<sup>17</sup> basados en la microelectrónica tienen problemas con la resolución<sup>18</sup> de la imagen, y se estima que el cerebro necesitará meses para llegarse a adaptar a los nuevos patrones sensoriales [Greenemeir, 2009]. Más aún, para poder reconocer la imagen de una cara humana se necesitará al menos 1000 pixeles<sup>19</sup>: problema a resolver en el futuro según los autores consultados. Otros problemas en esta línea de investigación es el consumo de potencia y el empaque electrónico [Weiland, 2008]. Obsérvese, que en caso de existir un trastorno neurofisiológico a nivel de corteza cerebral del área visual, poco se podrá hacer por el invidente.

En Hoydal [Hoydal, 1991] se presenta un bastón electrónico portátil ultrasónico y algunas pruebas de su utilidad; [Kaczmarek, 1991] se discuten las limitaciones de las tecnologías hápticas y se sugieren áreas que requieren mayor investigación para que los sistemas de sustitución sensorial sean más prácticos; en [Long, 1992] se presenta un desarrollo de un dispositivo que provee de información a los viajeros ciegos sobre su posición y su destino deseado usando mapas computarizados y aprovechando los sistemas de posicionamiento global actuales; [Langtao, 1993] presenta una brújula de bajo costo y fácil de usar que informa acerca de las direcciones que están tomando en su travesía;

---

<sup>17</sup> Retina.- Del b. lat. retina, y este del lat. rete, red. 1. f. Anat. Membrana interior del ojo de los vertebrados y de otros animales constituida por varias capas de células de forma y función muy variadas, y de la cual parten las fibras componentes del nervio óptico. En ella se reciben las impresiones luminosas y se representan las imágenes de los objetos.

<sup>18</sup> Rehabilitar.- tr. y prml. Habilitar de nuevo o restablecer a una persona o cosa en su antiguo estado.

<sup>19</sup> Pixel.- (voz i.) m. Punto mínimo de luz que entra a formar parte de una imagen: la resolución gráfica del monitor de una computadora depende de los píxeles.

[Shoval, 1998] informa de una ayuda para viaje que transforma las señales de sensores ultrasónicos en patrones acústicos; en [Arno, 2001] se describe cómo entrenan ciegos y sujetos con vista normal para reconocer patrones visuales (sencillos) desplegados en la pantalla de una computadora recorriéndolos con la ayuda de una tableta gráfica y su pluma, usando los patrones de sonido generados según el equipo de Capelle [Capelle, 1999], con esto afirman que el reconocimiento de patrones en un ambiente computarizado es posible; en [Moustakas, 2007] se da un método para convertir modelos en 3D en campos de fuerzas con la intención de ayudar a gente visualmente discapacitada a navegar en los espacios físicos; [Bartyna, 2005] es una propuesta de un sistema de ayuda para la navegación para gente invidente usando GPS y GIS<sup>20</sup>; [Horgan, 2005] reseña el trabajo de José Delgado, pionero de la implantación de chips en el cerebro; [McCreath, 2005] propone un sistema para mejorar el acceso a materiales de enseñanza para alumnos de todos los niveles en particular para ciegos y discapacitados visuales; [Renier, 2005] presenta un panorama de las líneas de investigación surgidas o relacionadas con la sustitución sensorial, demuestra como su estudio profundiza nuestro entendimiento de la percepción humana y de la plasticidad cerebral; [Snyder, 2005] propone una manera de ayudar a personas con poca o nula visión a comprender y apreciar eventos visuales (funciones de teatro, filmes, videos, exhibiciones de museos, y otros eventos multimedia) o libros; [Srivilailuck, 2005] intenta resolver la pregunta de qué pueden hacer los arquitectos y los ingenieros para ayudar a los ciegos, en particular en ambientes de

---

<sup>20</sup> GIS.- quiere decir sistemas de información geográfica.

aprendizaje; [Zelek, 2005] presenta un sistema háptico basado en una cámara que provee información para encontrar las trayectorias de su camino y evitar obstáculos usando un dispositivo háptico; [Jihong, 2006] hace un recuento del desarrollo de interfaces y técnicas de procesamiento de información que hacen posible que los invidentes obtengan una mejor percepción de su medio ambiente; [Noordzij, 2006] investiga la creación de modelos espaciales mentales en ciegos y personas normales, concluyendo que los ciegos se desempeñan mejor con descripciones de rutas, mientras que las personas normales lo hacen con otro tipo descripciones; en [Wall, 2006] se presentan experimentos sobre la percepción táctil relacionada con el entendimiento de grafos y visualizaciones simples usando un dispositivo de arreglos táctiles disponibles en el mercado, y [Forster, 2007] presenta el hallazgo de que el uso más intensivo de un sentido para compensar la carencia de otro conlleva a alteraciones de los mecanismos de selección de la atención dentro de sus áreas específicas del cerebro. Con esta recopilación de trabajos queremos convencer al lector del gran interés y trabajo realizado para generar ayudas para mejorar la calidad de vida de los ciegos.

Como se vio en el párrafo anterior, se pensó en “sustituir” una modalidad sensorial por el uso de otra, aunque esta interpretación ha sido objeto de controversias. Históricamente, primero se desarrollaron líneas de investigación en las que prevaleció la “sustitución” del sentido de la vista por el del tacto [Bach-y-Rita, 1998].

Para la mayoría de los investigadores el segundo sentido en importancia es el del oído [Pérez, 1991]. Basándonos en esta relación de los sentidos, se ha planteado el desarrollo de propuestas de sistemas de ayuda para las personas visualmente

incapacitadas usando el sentido del oído. Así comenzó el desarrollo de **dispositivos<sup>21</sup> de ayuda para la visión, basados en la audición para invidentes** (auditory-based vision assistive devices for blind). Dentro de estos dispositivos ha sobresalido el sistema *vOICe*<sup>22</sup> [Meijer, 1992].

El sistema de sustitución sensorial planteado por Meijer, contempla el uso de una computadora conectada a una cámara de televisión tipo estándar. La representación del sonido multiplexado<sup>23</sup> en el tiempo resulta de la conversión de una imagen a sonido por columnas; las imágenes con las que se trabaja tienen un tamaño de 64 x 64 píxeles con 16 tonos de gris. Sin embargo, la preocupación del autor es demostrar que se puede construir el sistema que él propone, dejando la evaluación por seres humanos para un futuro [Meijer, 1992].

En este sistema, la imagen es muestreada, digitalizada y almacenada como una matriz<sup>24</sup> de píxeles  $P$  constituida por renglones y columnas,  $M \times N$ , en donde  $M$  simboliza los renglones y  $N$  las columnas respectivamente. La conversión de una imagen a sonido se inicia tomando la primera columna de la izquierda  $j = 1$ , recorriendo la imagen columna por columna hacia la derecha hasta que el final, después de lo cual se genera un pulso de sincronización, aviso de que una nueva imagen será transformada. Es

---

<sup>21</sup> Dispositivo.- adj. Que dispone: medidas dispositivas. m. Mecanismo dispuesto para obtener un resultado.

<sup>22</sup> Para ver más sobre *vOICe* ver siguiente liga: <http://www.seeingwithsound.com/>. Página accesada por última vez el 24 de septiembre de 2011.

<sup>23</sup> Multiplexado.- adj. y com. [Sistema de comunicación] que permite combinar múltiples mensajes simultáneamente en el mismo soporte de transmisión físico o lógico: los sistemas dúplex y múltiplex son una técnica derivada del teléfono; m. Sistema que permite transmitir en directo varias señales de radio o televisión desde diferentes lugares.

<sup>24</sup> Matriz.- adj. Mat. Perteneciente o relativo al cálculo con matrices. En matemáticas, una matriz es una tabla bidimensional de números consistente en cantidades abstractas que pueden sumarse y multiplicarse. Las matrices se utilizan para describir sistemas de ecuaciones lineales, realizar un seguimiento de los coeficientes de una aplicación lineal y registrar los datos que dependen de varios parámetros. Pueden sumarse, multiplicarse y descomponerse de varias formas, lo que también las hace un concepto clave en el campo del álgebra lineal.

decir, la señal de la imagen es la sucesión (serie) ininterrumpida de señales de igual duración en el tiempo correspondientes a cada columna. Cada píxel en una columna, según el renglón en el que esté, tiene asociada una onda cuya frecuencia indica su posición; la mayor frecuencia está en la parte de arriba, la menor en la de abajo. La amplitud está determinada por la brillantez (tono de gris) del píxel. Los píxeles más brillantes producen ondas con la amplitud más grande; los píxeles negros tienen asignada amplitud cero. La frecuencia corresponde al tono y la amplitud al volumen de la señal. La señal de toda la columna es la suma de todas las ondas asociadas a sus píxeles (y sus respectivos tonos de gris).

En la Figura 2.1 se muestra el procedimiento de conversión para la columna 6 de una imagen de  $8 \times 8$  píxeles con tres tonos de gris ( $M = N = 8$ ,  $G = 3$ ). En realidad el prototipo de Meijer es un sistema de  $M = N = 64$  y  $G = 16$ , en el cual utilizó 16 tonos de gris con la idea de “asegurar la preservación de suficiente información”. Su sistema estuvo constituido por una cámara, una laptop adaptada, audífonos y un sistema que emula la conducta de 64 osciladores senoidales por medio de cómputos digitales. Una de las debilidades del estudio realizado en esta etapa es el de haber utilizado sólo a dos sujetos humanos en sus experimentos.

En la Figura 2.2, se puede observar la representación tridimensional de una imagen resultante después de haber llevado a cabo el barrido<sup>25</sup> correspondiente, los ejes

---

<sup>25</sup> Barrido.- 1. m. Acción de barrer; 2. Fís. Proceso por el que un dispositivo explora sistemática y repetidamente un área o un espacio reconociéndolos punto por punto para transformar la imagen de cada uno de ellos en señales eléctricas transmisibles a distancia, que, a su recepción, por otro proceso inverso y similar, se convierten en imágenes. Es el fundamento de la televisión, el radar, el microscopio de barrido, etc.

cartesianos están constituidos por la frecuencia, el tiempo y la amplitud de la onda. El propósito es demostrar que la asociación propuesta es biyectiva<sup>26</sup>.

Se encontró en la revisión bibliográfica otro método de barrido para patrones acústicos que representan información visual [Wang, 1996]. Los parámetros físicos de la imagen tomados en consideración para estos patrones son: el tamaño, la resolución, diferencia de nivel, curvatura y contraste<sup>27</sup> de color (¡aunque trabajan exclusivamente con imágenes binarias!). El rastreo<sup>28</sup> de la imagen se lleva a cabo de manera semejante a un monitor de televisión, de arriba a la izquierda en forma horizontal, línea por línea hasta llegar finalmente abajo a la esquina inferior derecha. El sonido es sintetizado por una tarjeta de sonido comercial.

---

<sup>26</sup> Byectiva.- Viene de la palabra inglesa Bijection . En matemática, una función  $f: X \rightarrow Y$  es biyectiva si para todo elemento  $y$  de  $Y$ , existe un único elemento  $x$  en  $X$  tal que  $f(x)=y$ . En otras palabras, una función biyectiva aparea cada elemento de  $X$  con un (único) elemento de  $Y$  y viceversa. Las funciones biyectivas son muy importantes en Matemáticas porque sirven para clasificar, por ejemplo, dos conjuntos tienen la misma cardinalidad (tamaño) si existe una función biyectiva entre ellos.

<sup>27</sup> Contraste.- 1. m. Acción y efecto de contrastar; 2. Oposición, contraposición o diferencia notable que existe entre personas o cosas; 3. Relación entre el brillo de las diferentes partes de una imagen; 4. Relación entre la iluminación máxima y mínima iluminación de un objeto; 5. En la imagen fotográfica o televisiva, inexistencia o escasez de tonos intermedios, de tal manera que resaltan mucho lo claro y lo oscuro.

<sup>28</sup> Rastreo.- 1. tr. Seguir el rastro o buscar alguna cosa por él; 2. fig. Inquirir, indagar, averiguar una cosa, discurriendo por conjeturas o señales.

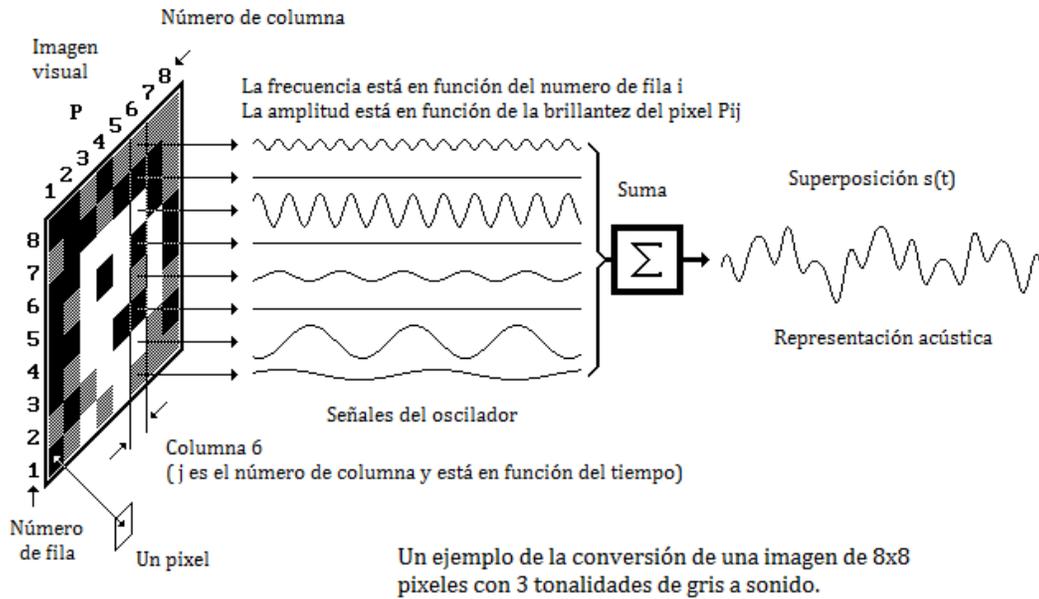


Figura 2.1. Se muestra el principio de barrido de la imagen visual y su transformación a patrón acústico. Tomada de [Meijer, 1992].

Se encontró en la revisión bibliográfica otro método de barrido para patrones acústicos que representan información visual [Wang, 1996]. Los parámetros físicos de la imagen tomados en consideración para estos patrones son: el tamaño, la resolución, diferencia de nivel, curvatura y contraste<sup>29</sup> de color (¡aunque trabajan exclusivamente con imágenes binarias!). El rastreo<sup>30</sup> de la imagen se lleva a cabo de manera semejante a un monitor de televisión, de arriba a la izquierda en forma horizontal, línea por línea hasta llegar finalmente abajo a la esquina inferior derecha. El sonido es sintetizado por una tarjeta de sonido comercial.

<sup>29</sup> Contraste.- 1. m. Acción y efecto de contrastar; 2. Oposición, contraposición o diferencia notable que existe entre personas o cosas; 3. Relación entre el brillo de las diferentes partes de una imagen; 4. Relación entre la iluminación máxima y mínima iluminación de un objeto; 5. En la imagen fotográfica o televisiva, inexistencia o escasez de tonos intermedios, de tal manera que resaltan mucho lo claro y lo oscuro.

<sup>30</sup> Rastreo.- 1. tr. Seguir el rastro o buscar alguna cosa por él; 2. fig. Inquirir, indagar, averiguar una cosa, discurriendo por conjeturas o señales.

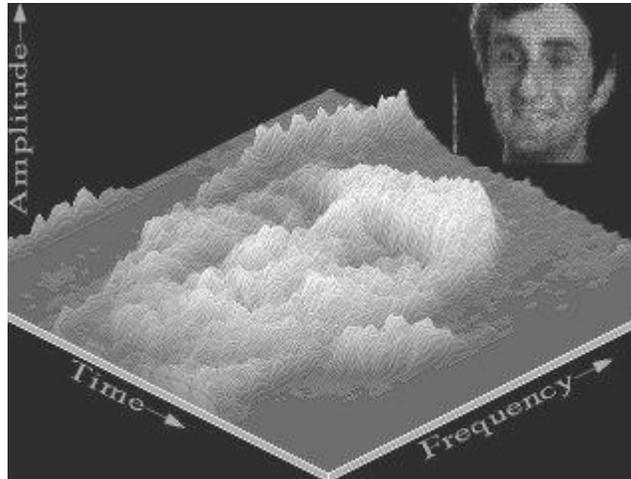


Figura 2.2. Representación tridimensional de una imagen.  
Tomada de [The vOICe at CNS, 2005].

La intensidad de la brillantez del pixel se transforma en nivel de sonoridad<sup>31</sup>. Sus experimentos se enfocan exclusivamente a la determinación de patrones de letras. El trabajo se concentra únicamente en la conversión de imágenes (binarias) en “imágenes auditivas” (auditory image) sin presentar un sistema de hardware parecido al presentado por Meijer. En su artículo no hace mención si los diez voluntarios participantes en sus experimentos son invidentes, débiles visuales o personas de vista normal. [Wang, 1996].

A finales de los 90's surge un prototipo experimental [Capelle, 1998] que consta esencialmente de una cámara de video, colocada en la cabeza del usuario, conectada a una tarjeta de video, una computadora y una tarjeta de audio dedicada a la generación del sonido conectada esta última a unos audífonos; basándose para la generación de la señal en el hecho de que el número máximo aceptable de frecuencias para la conversión de imagen a sonido es de 600 tonos de frecuencia, porque según Zwicker la banda audible de

---

<sup>31</sup> Nivel de Sonoridad.- La sonoridad es una medida subjetiva de la intensidad con la que un sonido es percibido por el oído humano. Es decir, la sonoridad es el atributo que nos permite ordenar sonidos en una escala del más fuerte al más débil. La unidad que mide la sonoridad es el decibel. La sensación sonora de intensidad (sonoridad) se agudiza para sonidos débiles, y disminuye para sonidos fuertes, lo que se debe a que la audición humana no es lineal, sino logarítmica.

los seres humanos está dividida en 600 bandas de frecuencia; los sonidos dentro de una misma banda difícilmente se pueden diferenciar, no así dos sonidos de frecuencias en bandas distintas [Zwicker, 1957] y [Zwicker, 1975].

En este caso las imágenes a convertir tienen una resolución mayor en la parte central de la imagen, de manera parecida a como lo hace la retina humana. A cada uno de sus píxeles se le asocia un tono senoidal; la amplitud de cada onda es modulada por el tono de gris del píxel correspondiente. El sonido correspondiente a toda la imagen es la suma sobre todos los píxeles de la imagen de los tonos modulados. Las frecuencias de las ondas asociadas a los píxeles en la parte central se determinan usando una fórmula, mientras que en la parte restante, otra. Se propusieron 50 patrones basados todos ellos en líneas rectas; se afirma que el modelo permitió un reconocimiento de los 50 patrones visuales.

Experimentaron con un grupo de 24 sujetos con visión normal pero vendados, con los que formaron dos grupos, uno de control y el otro de experimentación. El entrenamiento duró 10 sesiones de una hora cada una, más cuatro sesiones de evaluación. Según los autores este modelo de sustitución sensorial resultó ser pertinente y de utilidad, ya que los experimentos resultaron exitosos dentro de la evaluación.

En [Meijer, 2001] se afirma que técnicamente es posible convertir la imagen de video a audio, preservando una cantidad significativa de información visual en el patrón de sonido resultante. Si los invidentes pudieran aprender a reconstruir el contenido visual de estos sonidos, entonces ellos podrían ser capaces de "ver" por el oído. Durante la fase experimental de ese proyecto, la primera invidente voluntaria empezó a explorar los

límites de ese sistema a través del uso continuo del mismo: portó una cámara montada en la cabeza y una computadora con el software que convertía las imágenes de la cámara en “panoramas de sonido<sup>32</sup>” (patrones acústicos). La opinión personal de la voluntaria después de casi un año de uso fue: *"El usar este sistema es como caminar desde la total oscuridad a la luz"* [Meijer, 2001].

Un sistema mejorado es el **vOICe** que consta de una pequeña cámara, una computadora personal portátil (laptop) y audífonos. La cámara se monta sobre la cabeza y la laptop toma la imagen de entrada transformándola en información auditiva (patrones de sonido) [Motluck, 2005]. La escena que se encuentra enfrente de los sujetos es barrida en estéreo: oyéndose los objetos de la izquierda a través del oído izquierdo y los objetos de la derecha a través del oído derecho. La forma de transformar las imágenes en sonidos sigue siendo la misma que ha utilizado Meijer, actualizándose la imagen cada segundo.

En un análisis del sistema **vOICe** se concluye que, para hacer una interpretación correcta de una escena, se requiere tanto del conocimiento del contexto como de la intuición<sup>33</sup> del usuario, debido a que los sujetos no pueden fácilmente decodificar el patrón acústico por varios factores: la imagen puede ser borrosa, incluso antes de ser comprimida en forma drástica a pocos píxeles y la manera en que es recorrida durante la conversión a sonido. La participación de dos sujetos, un sujeto masculino y uno femenino, ambos con una edad de 23 años con audibilidad y visión deficientes. El sujeto masculino

---

<sup>32</sup> Panoramas de sonido.- 1. Paisaje muy dilatado que se contempla desde un punto de observación y que es representado mediante un patrón acústico; 2. fig. Aspecto acústico de conjunto de una cuestión visual.

<sup>33</sup> Intuición.- Del lat. mediev. intuitio, -onis: 1. f. Fil. Percepción íntima e instantánea de una idea o una verdad, tal como si se tuviera a la vista; 2. Facultad de comprender las cosas instantáneamente, sin razonamiento.

participó en el entrenamiento y pruebas de segmentos, mientras el sujeto femenino solamente participó dentro de la prueba de segmentos seleccionados como un control.

Los sujetos fueron sometidos a un entrenamiento de 20 horas, concluyéndose al final que el número de horas no fue suficiente. Hicieron la recomendación, que para optimizar la situación hay que usar el sistema por periodos más largos [Storer, 2006] y [Meijer, 2007].

Existe un sistema conocido como NAVI (Navigation Assistance Visually Impaired), que incluye un sistema de procesamiento implementado en una sola tarjeta (Single Board Processing System, SBPS), un sensor de visión montado en unas gafas de sol y audífonos. El sensor de visión capta la información visual enfrente del usuario, la imagen capturada es procesada para identificar el objeto en la forma de la imagen. Esto se hace usando algoritmos de redes neuronales difusas<sup>34</sup> basadas en LVQ (Learning Vector Quantization) y un algoritmo de agrupamiento (clustering) FCM (Fuzzy C Means). La imagen procesada es partida por la mitad para generar dos (nuevas imágenes con sus correspondientes) patrones de sonido. La imagen de la mitad izquierda es barrida columna por columna de extrema izquierda hacia el centro, simultáneamente la mitad derecha es barrida de la columna extrema derecha hacia el centro, los patrones acústicos generados se mandan al oído correspondiente a través de los audífonos. La forma en la que hacen generar los patrones acústicos de las mitades de imagen es similar a la de Meijer. No es muy claro a

---

<sup>34</sup> Algoritmos difusos, lógica difusa.- La lógica difusa o lógica heurística se basa en lo relativo de lo observado. Este tipo de lógica toma dos valores aleatorios pero contextualizados y referidos entre sí. Así, por ejemplo, una persona que mida 2 metros es claramente una persona alta, si previamente se ha tomado el valor de persona baja y se ha establecido en 1 metro. Ambos valores están contextualizados a personas y referidos a una medida métrica lineal. Un algoritmo difuso es un algoritmo que en su diseño incorpora conceptos de lógica difusa.

cuantos invidentes han entrenado para usar su prototipo, ni tampoco como han llevado a cabo su entrenamiento [Nagarajan, 2003], [Nagarajan, 2004], [Jihong, 2006].

El Videófono<sup>35</sup> [Martínez, 1984] tenía la capacidad de generar patrones acústicos, los cuales representaban imágenes geométricas, formadas exclusivamente con líneas rectas; sus creadores demostraron la viabilidad de decodificar patrones acústicos, después que los sujetos fueran sometidos a un entrenamiento orientado a desarrollar la capacidad de decodificación de los patrones propuestos.

En el Videófono se cambiaron respectivamente los ejes  $(x,y)$  de la imagen bidimensional por unos nuevos ejes, los cuales correspondieron a la frecuencia y al tiempo  $(f,t)$ . Se utilizaron bandas de frecuencia de ruido rosa, donde estas bandas de frecuencia formaron los patrones de sonido en función del tiempo, describiendo así los contornos de las imágenes formadas por líneas rectas. Se lograron generar un total de 127 patrones acústicos. Sin embargo, es importante mencionar, que nunca se generaron patrones acústicos que involucrasen curvas. Según [Liu, 2010] estas son todas las propuestas que anteceden a la propuesta presentada en este documento.

Algunas de las deficiencias que observamos en la revisión bibliográfica fueron: no está claro cómo se midió el resultado de las evaluaciones del aprendizaje en los invidentes o no se establece claramente qué tipo de imágenes geométricas se utilizaron en los entrenamientos o bien informan que usan dibujos que recuerdan letras pero no figuras geométricas, o no se reportan con claridad en qué consistieron los cursos de

---

<sup>35</sup> Videófono.- m. Sistema antecesor del Visor Acústico cuyo objetivo es generar patrones acústicos a partir de una a ocho líneas rectas.

entrenamiento, ni las evaluaciones de los mismos: [Meijer, 1992], [Meijer, 2007], [Nagarajan, 2003], [Nagarajan, 2004], [Matta, 2004], [Balakrisman, 2005].

En general, podemos decir de los trabajos reseñados que no muestran una preocupación explícita por generar un sonido “fácil” de interpretar, es decir, no parece haber ningún criterio en este sentido en la elección de las ondas senoidales para la generación del sonido.

Por otro lado, a partir del año de 2004 se introduce el concepto de **sonificación**: La transformación de datos<sup>36</sup> en sonidos, con la intención de transmitir información (no sólo a personas con problemas de visión) [Matta, 2004]. Este concepto viene a ayudar a clarificar y generalizar las ideas de los trabajos arriba mencionados para un posible uso por personas con vista normal.

Como se verá en el Capítulo 6, los trabajos [Meijer, 1992], [Capelle, 1998] y [Sainarayanan, 2007] son los que más impacto han tenido en la comunidad científica, aunque cada uno tiene su propia repuesta y su propia forma de evaluar su utilidad. Sin embargo, no parece haber una preocupación por el aspecto cognitivo de la visión, ni de la audición, ni una participación “activa” de los posibles usuarios en el diseño del dispositivo, ni las opiniones de los mismos, ni tampoco se aprovecha toda la experiencia de la comunidad de desarrolladores de software libre para generar sonido. Esto último es lo que tratamos de tomar en cuenta en nuestra propuesta.

---

<sup>36</sup> Datos en general, no solamente formas geométricas, sino también datos abstractos.



# Capítulo 3

## Técnicas utilizadas

### 1. Introducción

En este capítulo primero damos una explicación de la razón por la que se consideró el contorno de los objetos para producir los patrones acústicos. Así mismo, se hace una breve descripción de las técnicas de procesamiento de imágenes utilizadas y de la herramienta de software utilizada para la síntesis de sonido CSound.

### 2. La forma, los contornos visuales y los objetos.

*Los contornos presentes en el estímulo visual son los que proporcionan la información necesaria para la percepción de la forma.* La percepción de la forma requiere tiempo, incluso al presentarse una figura muy sencilla (círculo, triángulo, cuadrado) a un sujeto. Si la presentación tiene lugar con luz débil y es muy breve, el sujeto sólo tendrá conciencia<sup>37</sup>,

---

<sup>37</sup> Conciencia.- Del lat. conscientia: 1. f. Propiedad del espíritu humano de reconocerse en sus atributos esenciales y en todas las modificaciones que en sí mismo experimenta; 2. Conocimiento Conocimiento exacto y reflexivo de las cosas.

muy vagamente, de haber visto algo, sin saber bien qué. En este tipo de experiencia de laboratorio, todo sucede como si hubiese una *construcción* de la forma por el sistema visual; así la prolongación de la presentación hace identificable la forma [Berramendi, 2000].

En una forma figurativa, la percepción de la forma es inseparable, no sólo de la percepción de los bordes, sino de la de los objetos figurados. En particular, en la inmensa mayoría de las imágenes fotográficas y videográficas, el problema de la percepción de la forma es el de la percepción de los objetos visuales. La percepción de la forma sólo se hace más difícil, o menos habitual, cuando la imagen se vuelve más abstracta, más simbólica o cuando la profundidad se representa de manera incorrecta.

### **3. El contorno y la comunicación**

En [Dondis, 2002] se describe al contorno como un medio de comunicación y a las líneas, lo que conforma un contorno. En la terminología de las artes visuales se dice que la línea articula la complejidad del contorno. Hay tres contornos básicos: el cuadrado, el círculo y el triángulo equilátero. Cada uno de ellos tiene su carácter específico y rasgos únicos, Figura 3.1. A cada uno se le atribuye gran cantidad de significados, unas veces mediante la asociación, otras mediante una adscripción arbitraria y otras a través de nuestras propias percepciones.

Todos los contornos básicos son figuras planas y simples que pueden describirse y construirse fácilmente, ya sea por procedimientos visuales o verbales. Las descripciones verbales de los contornos básicos son:

Un cuadrado es una figura de cuatro lados con ángulos rectos exactamente iguales en sus esquinas y lados que tienen exactamente la misma longitud, Figura 3.2.

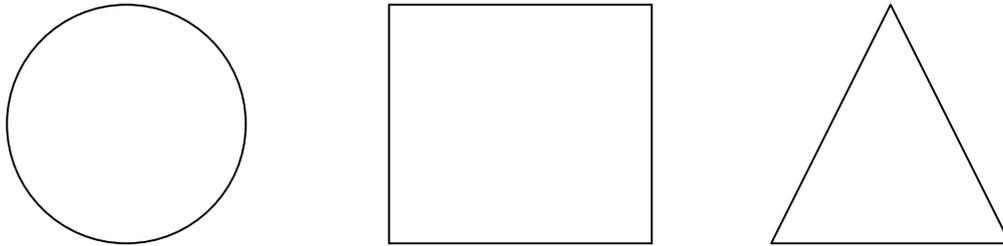


Figura 3.1. Contornos básicos, imagen tomada de [Dondis, 2002].

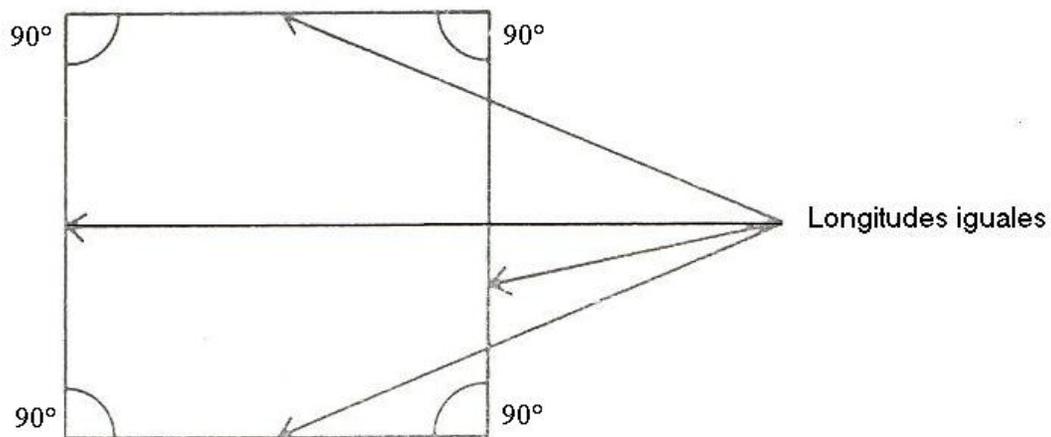


Figura 3.2. Un cuadrado, imagen tomada de [Dondis, 2002].

Un círculo es una figura continuamente curvada, en la que los puntos de su perímetro equidistan del centro, Figura 3.3.

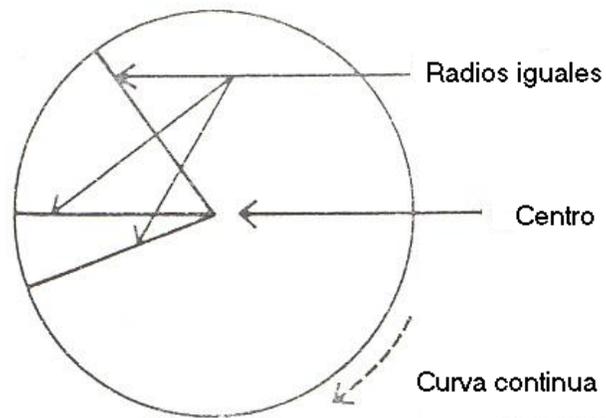


Figura 3.3. Un círculo, imagen tomada de [Dondis, 2002].

Un triángulo equilátero es una figura de tres lados cuyos ángulos y lados son todos iguales, Figura 3.4.

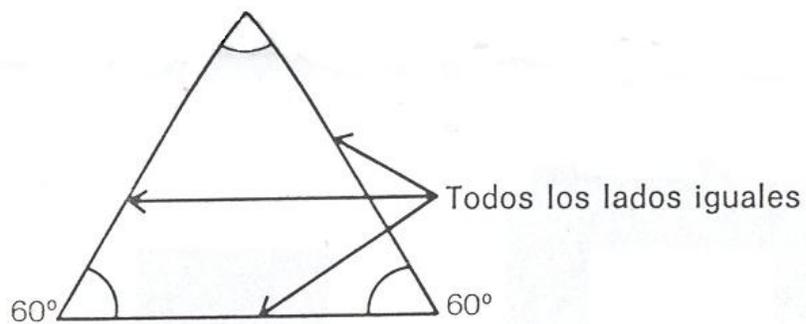


Figura 3.4. Un triángulo equilátero, imagen tomada de [Dondis, 2002].

Dondis afirma también que a partir de estos contornos básicos y las líneas rectas con diferentes orientaciones podemos derivar mediante combinaciones y variaciones incabables todas las formas físicas de la naturaleza y de la imaginación, Figura 3.5.

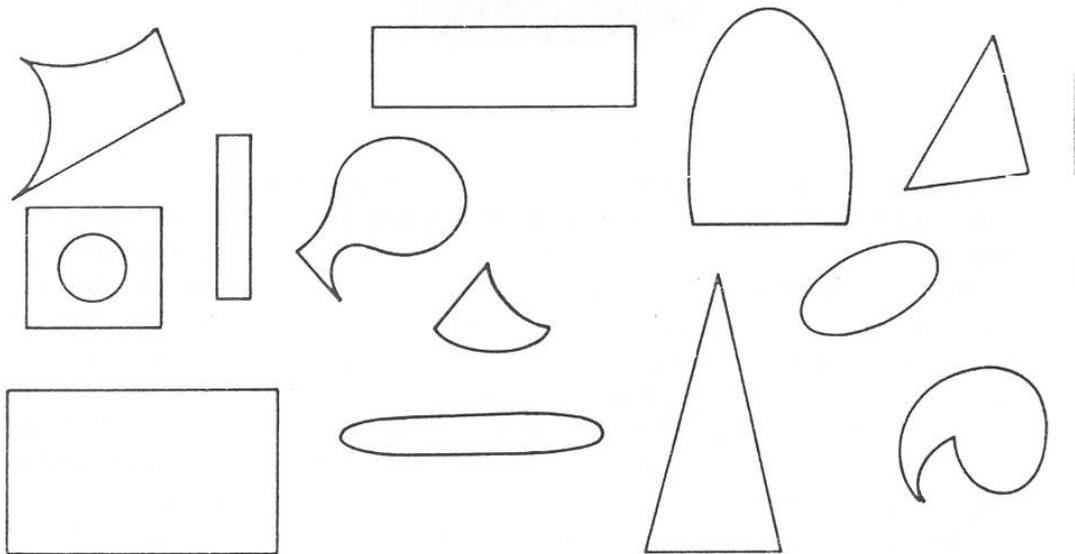


Figura 3.5. Con el círculo, cuadrado y triángulo, podemos obtener cualquier imagen, imagen tomada de [Dondis, 2002].

#### 4. Técnicas usadas para el procesamiento de imágenes

La imagen digital es cualquier imagen, fija o en movimiento, que se capture en un medio electrónico y que se represente como un archivo de información basado en un sistema binario (ceros y unos), esta última acción se le conoce como digitalización. A cada pixel le corresponde un color y un lugar en una imagen, conformando así una representación bidimensional (matricial), también llamada mapa de bits, bitmap, que es una estructura o archivo de datos matricial de pixeles.

En una imagen en escala de grises cada punto de la imagen se almacena en un Byte<sup>38</sup>, donde su valor numérico representa su tono, que puede oscilar entre el blanco

---

<sup>38</sup> Byte.- Palabra en lengua inglesa (pronunciada [bait] o ['bi.te]) muy usada en la literatura técnica, aunque la Real Academia Española ha aceptado como su equivalente a octeto (es decir a ocho bits). En esta tesis entenderemos por un byte una secuencia de ocho bits contiguos, cuyo tamaño depende del código de información o código de caracteres en que

(255) y el negro (0). En este trabajo consideraremos dos tipos de imagen digital: la binaria y en tonos de grises.

En la imagen binaria, sólo existen 2 tonos, el blanco y el negro, es una imagen con un bit por pixel. En una imagen en escala de grises cada punto de la imagen se almacena en un Byte, donde su valor numérico representa su tono, que puede oscilar entre el blanco (255) y el negro (0). El formato de las imágenes digitales con las que trabajaremos será el formato BMP<sup>39</sup>. Una *imagen* (en tonos de gris) la podemos describir como una función bidimensional de intensidad de luz  $f(x, y)$ , donde  $x$  e  $y$  representan las coordenadas espaciales y el valor de  $f$  en un punto cualquiera  $(x, y)$  es proporcional al brillo (*o nivel de gris*) de la imagen en ese punto. La Figura 3.6 muestra la convención de ejes utilizada.

#### **4.1 Transformaciones elementales**

El propósito del procesamiento de imágenes es proporcionar un conjunto de técnicas que se aplican a las imágenes digitales con el objetivo de mejorar la calidad o facilitar la búsqueda de información [González, 1996].

El **procesamiento** de imágenes puede ser visto como una transformación de una imagen en otra imagen, es decir, a partir de una imagen, se obtiene otra imagen modificada. Por otro lado, el **análisis** es una transformación de la imagen en algo distinto a una imagen [Pajares, 2004].

---

sea definido. Se usa comúnmente como unidad básica de almacenamiento de información en combinación con los prefijos de cantidad.

<sup>39</sup> BMP.- es el formato propio del programa Microsoft Paint, que viene con el sistema operativo Windows. Consultado en: [http://es.wikipedia.org/wiki/Windows\\_bitmap](http://es.wikipedia.org/wiki/Windows_bitmap), el 24 de septiembre de 2011.

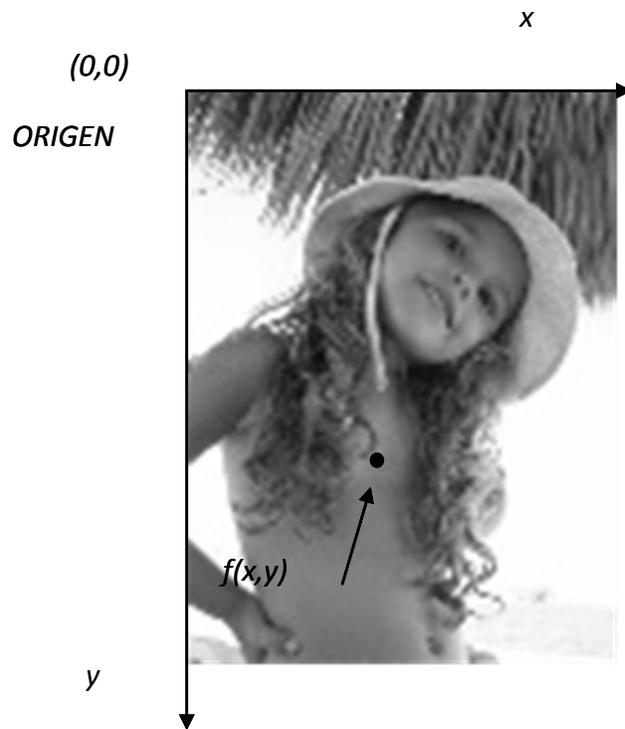


Figura 3.6. Convenio de ejes utilizado para la representación digital de imágenes (Fotografía tomada por el autor).

El procesamiento digital de la imagen en el espacio puede enfocarse desde dos perspectivas:

1. Operaciones de vecindad basadas en varios puntos.
2. Alteración individuales pixel a pixel.

La generación de un nuevo pixel en una nueva imagen, será una función, bien del valor de cada pixel en su localización individual, o bien de los valores de los pixeles en la vecindad de un pixel dado.

Las operaciones individuales implican la generación de una nueva imagen modificando el valor de cada pixel basándose en una regla global aplicada a cada uno de la imagen, modificándolo por una operación lineal o no lineal, Figura 3.7.

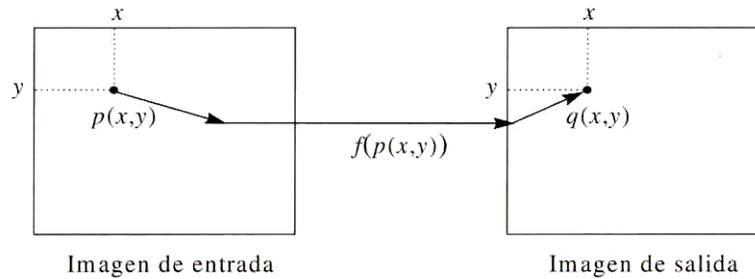


Figura 3.7. Operación individual, imagen tomada de [Pajares, 2004]

### 4.3 Filtrado en el dominio del espacio

El proceso de filtrado en el dominio del espacio se realiza utilizando matrices denominadas máscaras, que son aplicadas sobre la imagen. La imagen resultante de la aplicación de un filtro es una nueva imagen. Los filtros sirven para suavizar o resaltar los detalles de la imagen, o inclusive minimizar efectos de ruido.

Las operaciones de filtrado se llevan a cabo directamente sobre los píxeles de la imagen. El píxel objetivo se verá afectado por el conjunto de píxeles vecinos con la finalidad de obtener una información útil.

Los filtros en el dominio del espacio pueden clasificarse en:

- Filtros lineales (también conocidos como máscaras de Convolución).
- Filtros no lineales.

El concepto de máscara se entiende como una matriz de números reales, en nuestro caso de  $3 \times 3$ , la que se multiplicará con la matriz de  $3 \times 3$  con entradas con los valores correspondientes a una subimagen de  $3 \times 3$  centrada en el píxel  $(x, y)$  para obtener un valor  $g(x, y)$ . Esto también se puede hacer de manera similar con matrices de  $5 \times 5$ .

Dependiendo de la implementación, en los límites de la imagen se aplica un tratamiento especial (se asume un marco exterior de ceros o se repiten los valores del borde) o no se aplica ninguno. El tamaño de las máscaras utilizadas en este trabajo es de 3x3 y supusimos en los bordes un marco exterior de ceros.

#### ***4.4 Obtención del contorno de la imagen con topología digital***

Después de los primeros pasos de simplificación de las imágenes, existen todavía objetos en ellas que no son útiles y que deben ser eliminados de la misma, tomándose en consideración su tamaño. Con base en esto se podrán eliminar los objetos que no cumplan con un tamaño establecido de pixeles. Los objetos que pueden ser eliminados pueden ser de dos tipos: hoyos dentro de las imágenes y objetos que no cumplan con un cierto tamaño establecido. Para poder eliminar estos grupos de pixeles se utilizan algunos de los algoritmos de topología digital [Martínez, 2005], [Rodríguez, 1997]. Para el filtrado de las imágenes digitalizadas se usaron dos herramientas de software, la primera escrita en Java conocida como ATIC la que proporciona el autor Pajares [Pajares, 2004] y la segunda es la desarrollada en el lenguaje Visual C por Ramírez y Martínez [Martínez, 2005].

#### ***4.5 Técnicas de conversión a sonido***

La sonificación que se utiliza en la codificación (del contorno) de la imagen, se le conoce como código BIMUX<sup>40</sup> [Martínez, 1984], una técnica similar será usada por Meijer en su

---

<sup>40</sup> BIMUX quiere decir Bidimensional-Multiplex.

trabajo de 1992 [Meijer, 1992]. El patrón acústico obtenido con el código contiene, en principio, toda la información representativa de la imagen en tratamiento, dejando totalmente al oyente la tarea de efectuar la interpretación; dicha interpretación corresponde a la decodificación de los datos, o bien a lo que sería la lectura del patrón acústico, algo similar al concepto de la lectura de un texto impreso.

La conversión a sonido se basa en la siguiente idea:

Imaginemos una imagen como una pantalla de televisión blanco y negro formada por una multitud de líneas horizontales. Supongamos que a cada línea horizontal se le asocia una banda de frecuencias dentro del rango audible. Estas frecuencias se ordenarán en forma ascendente (de abajo hacia arriba) de la imagen, de modo que las frecuencias más altas queden en la parte superior, Figura 3.8.

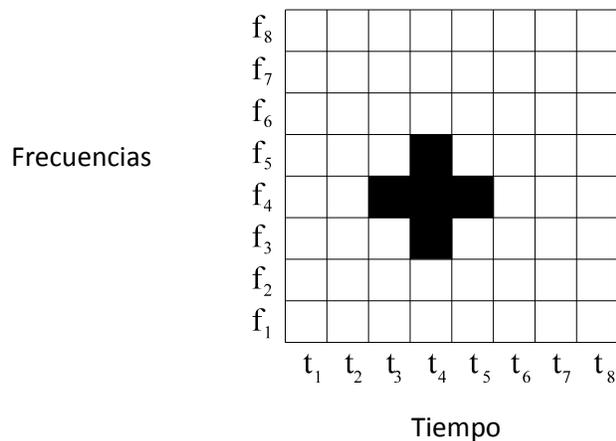


Figura 3.8. Aquí se muestran los puntos encendidos (en negro), imagen tomada de [Martínez, 1984].

Supongamos que una línea vertical imaginaria, llamada línea de barrido, inicialmente colocada en el extremo izquierdo de la imagen se desplaza hacia la derecha con una rapidez uniforme explorando la imagen en su totalidad.

La línea vertical haría sonar las frecuencias de los píxeles negros que vaya encontrando en su recorrido. En caso de que haya varios píxeles negros en un tiempo  $t_i$ , la salida del sistema en ese momento estará compuesta por la ejecución de sus respectivas frecuencias.

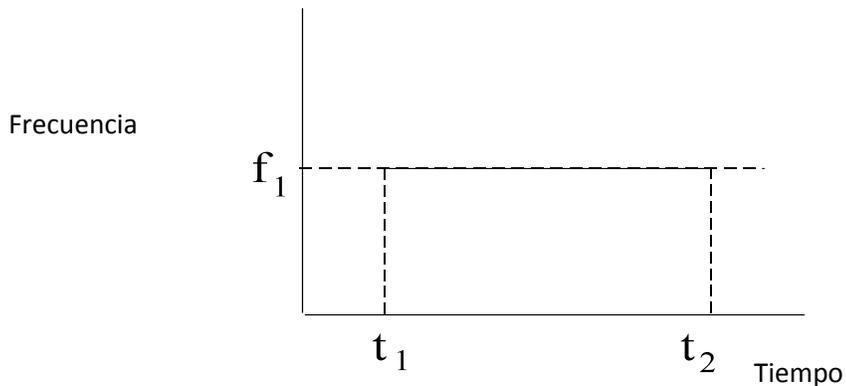


Figura 3.9. Tramo de línea horizontal, imagen tomada de [Martínez, 1984].

Por ejemplo, si tenemos una línea horizontal, Figura 3.9, ésta sonará como sonido de una frecuencia constante mientras se efectúe el barrido. Si la imagen es una línea inclinada de pendiente positiva (ascendente) sonará como un sonido cada vez es más agudo, mientras avanza el tiempo, Figura 3.10.

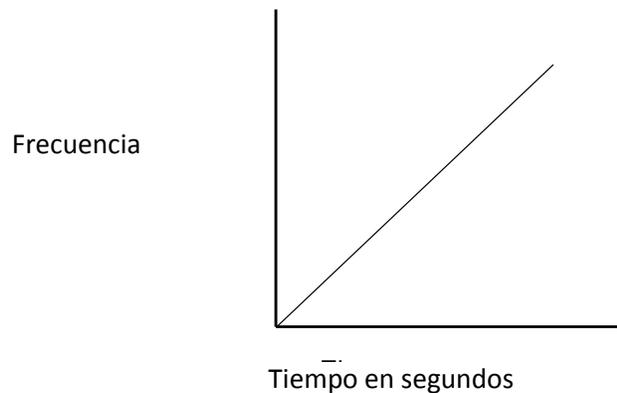


Figura 3.10. Imagen formada por una pendiente positiva, imagen tomada de [Martínez, 1984].

Supongamos ahora que la imagen está formada por puntos que forman una matriz rectangular, Figura 3.11. Para mandar la imagen por el canal (de información) auditivo, cuando la línea de barrido toca varios pixeles negros de una misma columna se ejecutan sus frecuencias correspondientes (simultáneamente) de manera paralela. Dicho de una manera más técnica decimos que los puntos de una misma columna se multiplexan en la frecuencia, mientras que las distintas columnas se multiplexan en el tiempo.

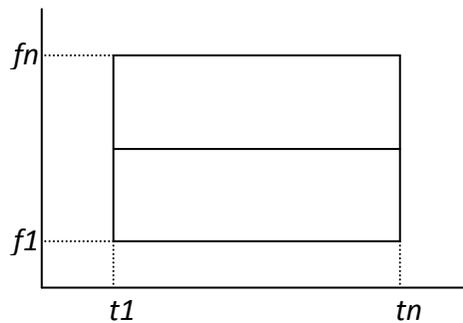


Figura 3.11. Imagen formada por puntos que forman una matriz rectangular, imagen tomada de [Martínez, 1984].

Para escoger las frecuencias se pensó en las de la escala bien temperada, Figura 3.12. De éstas fueron seleccionadas las correspondientes a las primeras 50 teclas blancas del piano a partir de la nota más grave, de izquierda a derecha del teclado. Dichas frecuencias están en el área central de las curvas de igual sonoridad, en dicha región es donde el oído es más sensible [Fletcher, 1933]. El oído a frecuencias más bajas o más altas ya no las percibe fácilmente [Giacomin, 2010].

Como parte de nuestro trabajo futuro queremos emular el timbre acústico del piano como parte del patrón sonoro ya que el piano es considerado como uno de los

instrumentos musicales más completos en lo que respecta a la generación de frecuencias audibles y armoniosas [Adli, 2007] y [Nave, 2005].

Tecla del piano	Nota del piano*	Frecuencia en Hz.
1	A <sub>0</sub>	25.5
2	B <sub>0</sub>	30.868
3	C <sub>1</sub>	32.703
4	D <sub>1</sub>	36.708
5	E <sub>1</sub>	41.203
6	F <sub>1</sub>	43.654
7	G <sub>1</sub>	48.999
8	A <sub>1</sub>	55.00
9	B <sub>1</sub>	61.735
10	C <sub>2</sub>	65.406
11	D <sub>2</sub>	73.416
12	E <sub>2</sub>	82.407
13	F <sub>2</sub>	87.307
14	G <sub>2</sub>	97.999
15	A <sub>2</sub>	110.0
16	B <sub>2</sub>	123.47
17	C <sub>3</sub>	130.81
18	D <sub>3</sub>	146.83
19	E <sub>3</sub>	164.81
20	F <sub>3</sub>	174.61
21	G <sub>3</sub>	196.00
22	A <sub>3</sub>	220.00
23	B <sub>3</sub>	246.94
24	C <sub>4</sub>	261.63
25	D <sub>4</sub>	293.66
26	E <sub>4</sub>	329.63
27	F <sub>4</sub>	349.23
28	G <sub>4</sub>	392.0
29	A <sub>4</sub>	440.0
30	B <sub>4</sub>	493.88
31	C <sub>5</sub>	523.25
32	D <sub>5</sub>	587.33
33	E <sub>5</sub>	659.26
34	F <sub>5</sub>	698.46
35	G <sub>5</sub>	783.99
36	A <sub>5</sub>	880.00
37	B <sub>5</sub>	987.77
38	C <sub>6</sub>	1046.5
39	D <sub>6</sub>	1174.7
40	E <sub>6</sub>	1318.5
41	F <sub>6</sub>	1396.9
42	G <sub>6</sub>	1568.0
43	A <sub>6</sub>	1760.0
44	B <sub>6</sub>	1975.5
45	C <sub>7</sub>	2093.0
46	D <sub>7</sub>	2349.3
47	E <sub>7</sub>	2637.0
48	F <sub>7</sub>	2793.8
49	G <sub>7</sub>	3136.0
50	A <sub>7</sub>	3520.0

Tabla obtenida de la Audio Cyclopedia, by : Howard M. Tremaine, Sams, 1969.

Figura 3.12. Las 50 frecuencias correspondientes a las 50 notas del piano escogidas, imagen tomada de [Tremaine, 1969]<sup>41</sup>.

<sup>41</sup> A corresponde a la nota La, B a la nota Si, C a Do, D a Re, E a Mi, F a Fa y G a Sol. Los subíndices de las notas que a parecen en la Figura 3.12 se refieren a las octavas del teclado del piano.

## 5. Lenguaje de programación CSound

Para llevar a cabo la transformación de la imagen a un patrón acústico se puede necesitar la ejecución de al menos dos frecuencias al mismo tiempo. Este problema se resolvió utilizando el lenguaje de programación de síntesis musical CSound [Boulanger, 2000].

Según el profesor Boulanger, [Boulanger, 2000], CSound es un software con la licencia GNU Lesser General Public License (Licencia Pública Menos General) en el que ha venido trabajando una gran comunidad interesada en la intersección de la música y de las ciencias de la computación. Originalmente fue desarrollado en los años 80's por Barry L. Vercoe del grupo Cognición y Música (Music and Cognition Group of the M.I.T., Media Laboratory) del Laboratorio de Medios del M.I.T. CSound transforma una computadora personal en una estación de trabajo de audio digital de alto rendimiento; crea un ambiente común para el diseño de sonido, investigación acústica, la producción de audio digital y la composición musical. De tiempo en tiempo se liberan nuevas versiones de CSound, para este trabajo se usó la 4.23f08, de enero de 2004.

Para crear los sonidos, en nuestro caso los patrones acústicos, CSound trabaja básicamente con dos tipos de archivos texto, independientes y complementarios. Uno llamado (de tipo) *orquesta* (en inglés, *orchestra*) y el otro (de tipo) *partitura* (en inglés, *score*). Estos archivos pueden ser nombrados como se desee. Normalmente, a ambos archivos se les da el mismo nombre, para que CSound reconozca de qué tipo de archivo se trata, se le asigna la extensión *.orc* para el archivo *orquesta* y *.sco* para el archivo *partitura*.

Archivo orquesta

El archivo orquesta está constituido de dos partes: la *sección del encabezado* y la sección de los instrumentos.

Sección de encabezado

Dentro de la sección del encabezado se define el muestreo<sup>42</sup> y se determina cómo el instrumento será ejecutado; se especifica también el número de canales de la salida. Un encabezado orquestal se muestra en la Figura 3.13.

```
sr      =      44100
kr      =      4410
ksmps  =      10
nchnls =      1
```

Figura 3.13. Encabezado del archivo CSound [Boulanger, 2000].

El código en este encabezado asigna:

*sr*, la relación de muestreo, igual a 44.1K (44100),

```
sr      =      44100
```

*kr*, la relación de control, igual a 4410,

```
kr      =      4410
```

y

*ksmps*,  $ksmps = sr/kr$  a 10,

```
ksmps  =      10
```

El encabezado también indica que esta orquesta deberá de generar un archivo de sonido monoaural; esto se indica en el renglón

---

<sup>42</sup> Muestreo.- 1. m. Acción de escoger muestras representativas de la calidad o condiciones medias de un todo; 2. Técnica empleada para esta selección; 3. Selección de una pequeña parte estadísticamente determinada, utilizada para inferir el valor de una o varias características del conjunto.

$$nchnls = 1$$

que quiere decir, el número de canales (*nchnls*) igual a 1,

### Sección de los instrumentos

Un archivo orquesta puede llegar a contener virtualmente cualquier número de instrumentos. En CSound los instrumentos se definen (y también se diseñan) por medio de la interconexión de módulos o bien de *opcodes*, en donde éstos generan o modifican las señales. Estas señales son representadas por medio de *Símbolos, etiquetas, o nombres de variables* que pueden ser modificadas de un opcode a otro.

En un archivo de CSound del tipo *orquesta*, la sintaxis de la declaración de un opcode es:

Output      Opcode      Argumentos      Comentarios(opcional)

Puede haber un número variable de argumentos, aunque ellos llevan un orden (intrínseco) en la lista correspondiente. En nuestro caso usaremos el mínimo número de argumentos, tres: amplitud, frecuencia y número de función.

En el caso del opcode **oscil** se tiene:

Salida	Opcode	amplitud,	frecuencia,	función #	;COMMENTARIOS
a1	<b>oscil</b>	10000,	440,	1	; OSCILATOR

Esto se entiende de la siguiente manera: el opcode llamado **oscil** generará una onda senoidal, lo indica la función # 1, de frecuencia 440 Hz con una amplitud de 10,000. Los datos generados por *oscil* se guardarán en la variable *a1*.

Observe que los argumentos de la entrada están a la derecha del nombre del opcode. Las salidas están a la izquierda. Las palabras que siguen al punto y coma (;) son consideradas como comentarios y son ignoradas.

Los instrumentos individuales están dados por un número único de instrumento y están delimitados por el establecimiento de los comandos *instr* y *endlin*. Es decir, el formato de un instrumento en un archivo orquesta es:

```

instr #

Declaración del opcode 1

.....

Declaración del opcode n

endlin

Por ejemplo,

      instr    101                ; OSCILADOR SIMPLE
a1    oscil   10000,  440,    1
      out     a1
      endlin

```

Figura 3.14. Código orquesta para el instrumento 101.

En la Figura 3.15 se muestra un diagrama a bloques del *inst 101*.

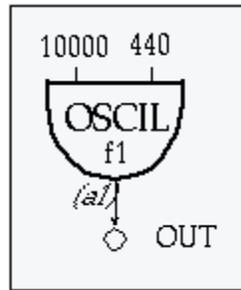


Figura 3.15. Diagrama a bloques del *inst 101* [Boulanger, 2000].

El diagrama a bloques del instrumento 101 muestra claramente como la salida del oscilador, etiquetado como *a1*, es pegado a la entrada de la opcode out, la que *escribe* la señal en el disco duro. El valor de *a1* puede también ser leído por los opcodes subsiguientes, como una entrada.

Csound ejecuta los instrumentos línea por línea, de arriba hacia abajo. Las señales de audio y control pueden ser virtualmente canalizadas a dondequiera dentro de un instrumento y usadas para establecer un parámetro a un nuevo valor, controlar dinámicamente un parámetro (como girar una perilla), o servir como una entrada de audio en algún procesamiento opcode.

Csound traduce primero el archivo de tipo orquesta en una estructura de datos residente en la máquina. Después, para ejecutar estos instrumentos necesita leer respectivo archivo de tipo partitura.

### Archivo de tipo partitura

El archivo *partitura* de Csound, es el que maneja los instrumentos de la orquesta. Tal y como los archivos orquesta, el archivo partitura posee dos partes: *tablas* y *notas*.

Tablas

En CSound se pueden usar un poco más de 20 subrutinas (ya predeterminadas) de funciones matemáticas de CSound llamadas **GENS** para generar tablas de funciones (tablas *f*) y/o llenarlas por medio de la lectura de archivos de sonido desde el disco duro. Por ejemplo, **GEN10** es una subrutina que calcula (aproximaciones) de ondas formadas por sumas pesadas de ondas senoidales. Cuáles y cómo las usarán los instrumentos de archivo orquesta es algo que queda a juicio del programador.

La sintaxis de la declaración de la función (en inglés *f-statement*) donde se utilizan las GENS es:

f	número	tiempo de carga	Tamaño de la tabla	GEN	parámetros de la rutina (con su orden)	;comentarios
---	--------	-----------------	--------------------	-----	--	--------------

Por ejemplo, si deseamos generar una onda senoidal con 4096 puntos para el instrumento 1 que se cargue al tiempo 0 escribimos:

f	1	0	4096	10	1	; 1 le indica a GEN10 que sólo tiene que calcular una onda senoidal.
---	---	---	------	----	---	--

Notas

En la segunda parte, se deberá escribir en las *notas de estados*. Estas notas de eventos controlan los instrumentos y pasan estos parámetros de comportamiento tales como las

frecuencias establecidas, niveles de amplitud, relaciones de vibrato<sup>43</sup> y tiempos de ejecución.

Cada declaración de nota en el archivo partitura ocupa una sola línea. Estas declaraciones activan un instrumento a un tiempo específico y tienen una duración específica. Puede tener un número muy grande de parámetros que determinar para el instrumento en cuestión y estos parámetros pueden variar de nota a nota.

El archivo partitura, igual que una partitura de orquesta, es leído línea por línea y nota por nota; las notas pueden tener los mismos tiempos de inicio y distintas duraciones.

La sintaxis más sencilla de una declaración de nota es:

i	Número de instrumento	Tiempo de inicio	Duración
---	-----------------------	------------------	----------

Por ejemplo:

i	102	0.03	0.09
---	-----	------	------

Dependiendo de la velocidad de la computadora (y de la complejidad del instrumento en el archivo *orquesta*) la ejecución de los archivos orquesta y partitura pueden generar sonido en tiempo real, o escribir datos directamente en un archivo en el disco duro.

---

<sup>43</sup> Vibrato.- Del lat. vibratum, supino de vibrare, vibrar. 1. adj. Que vibra o es capaz de vibrar. 2. El vibrato (it. vibrato = vibración) es una variación periódica de la frecuencia o amplitud de un sonido. En el canto se produce espontáneamente a través de un tremor nervioso en diafragma o laringe. 3. El vibrato de los instrumentos de cuerda y de viento es una imitación de esa función vocal. Si sólo la amplitud (fuerza) del sonido se cambia, se habla de trémolo. En el órgano, se imita el vibrato a través de la interferencia de dos notas cercanas.

---

Una vez hecho esto, se escuchará el resultado del archivo de sonido abriéndolo con el editor de sonido favorito; será ejecutado ya sea a través del convertidor digital analógico (DAC) de la tarjeta madre o bien por medio del DAC de la tarjeta de sonido de la PC. En Anexo V el lector interesado puede encontrar los archivos orquesta y partitura del patrón correspondiente a un cuadrado.



# Capítulo 4

## Desarrollo de la propuesta

### 1. El Modelo propuesto de conversión de imagen a sonido

A continuación presentamos las diferentes etapas del modelo propuesto:

Primeramente se extrae tan sólo el contorno del objeto relevante de la imagen en tonos de gris, el que se corresponderá con el conjunto de píxeles negros de una imagen binaria. Cualquier otro píxel que no pertenezca al contorno tiene asignado el color blanco en ésta. Al leer la imagen binaria, los píxeles negros son los que proveerán la información necesaria para construir el patrón acústico, compare con [Figura 4.1].

Usando el trabajo de Ramírez [Ramírez, 2002], y su colaboración se desarrolló un programa de software basado en métodos de topología digital para eliminar los objetos no deseados que estuvieran dentro y fuera del contorno de la imagen obtenido. Así fue

cómo se logró obtener un contorno limpio de la imagen, adecuado para ser transformado a patrón acústico. Este trabajo fue publicado en *Temas de Ciencia y Tecnología* [Martínez, 2005].

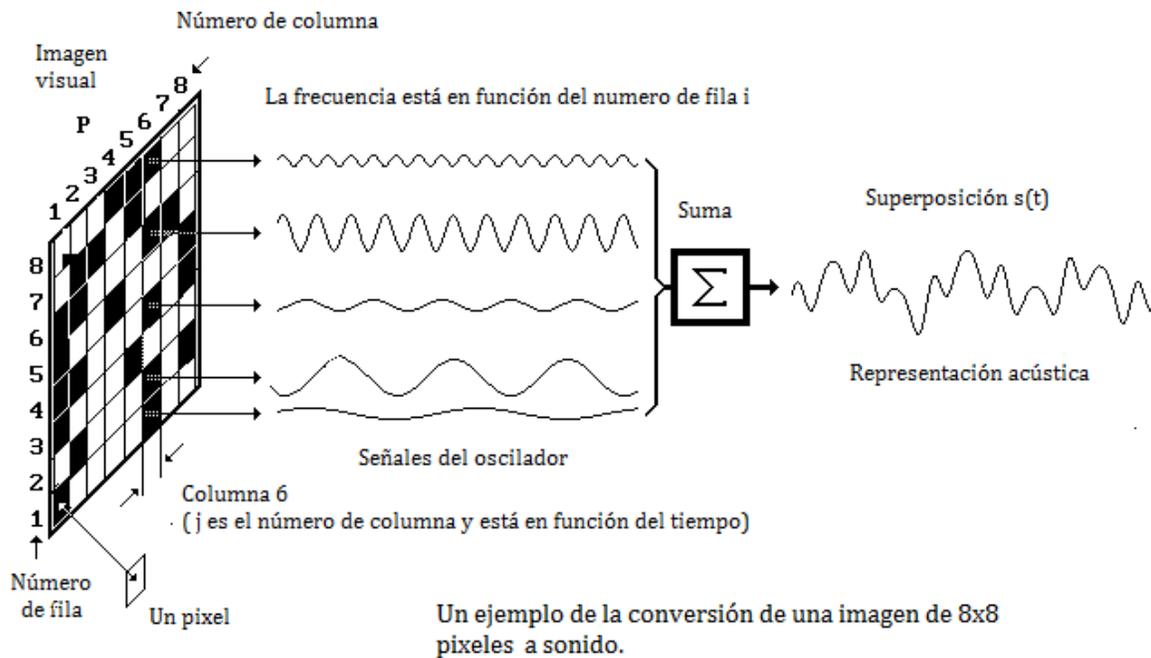


Figura 4.1. Se muestra el principio de barrido de la imagen visual y su transformación a patrón acústico del sistema de conversión de imagen a sonido, figura elaborada por el autor.

Con respecto a la parte de la obtención del contorno de un objeto se utilizaron los siguientes algoritmos [Pajares, 2004]:

1. **Operador umbral** (imagen binaria). Crea una imagen de salida binaria a partir de una imagen de grises, donde el nivel de transición está dado por un parámetro de entrada.
2. **Operador intervalo de umbral binario**. Crea una imagen de salida binaria a partir de una imagen de grises, donde los valores de gris cuyo nivel está dentro de un intervalo definido se transformarán en negros, los restantes en blancos.

3. **Operador de umbral de la escala de grises.** Crea una imagen de salida con los únicos valores de nivel de gris comprendidos dentro de un intervalo definido, los restantes se convierten en pixeles blancos.

Todos estos tres algoritmos pertenecen al grupo de transformaciones elementales, las que transforman una imagen en otra.

4. **Filtro pasa alto.** Este tipo de filtros están basados en el uso de una máscara, en nuestro

caso usamos la máscara  $\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$ . Este filtro deja pasar las frecuencias altas de la señal; estas frecuencias se asocian a cambios bruscos de intensidad en pequeños intervalos espaciales, es decir bordes. Uno de los efectos que se logra sobre la imagen es acentuar los bordes de los objetos que aparecen en ella.

Estos algoritmos pertenecen al grupo de transformaciones de vecindad que consisten en transformar el valor de un pixel  $p$  en la posición  $(x,y)$  teniendo en cuenta los valores de los pixeles vecinos, usando una máscara con valores constantes.

5. **Modificación del histograma.** Se buscó desaparecer pixeles con valores extremos, ya sean muy blancos o muy negros, utilizando el histograma de la imagen. La modificación del histograma permite variar la distribución de los niveles de intensidad. Dichas técnicas tratan de eliminar efectos no deseados tales como sombras, reflejos, etc., a la vez que aumentan el contraste.

La jerarquía de los algoritmos usados se encuentra en la Figura 4.2.

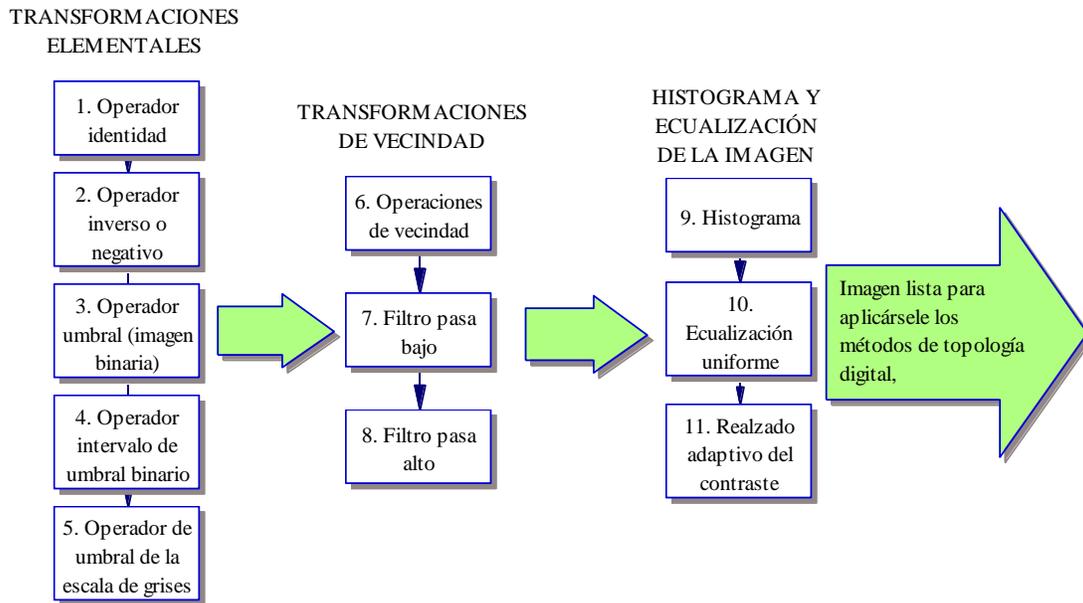


Figura 4.2. Diagrama representativo de todos los algoritmos utilizados en este estudio.

A continuación se eliminaron objetos (agujeros) por afuera y por dentro del contorno tomando en consideración su tamaño. Para poder eliminar estos grupos de píxeles se utilizaron ideas de la topología digital [Rosenfeld, 1989]: primero se ejecutó un algoritmo de etiquetado de las 8-componentes negras y luego uno análogo para el etiquetado de las 4-componentes blancas, [Ramírez, 2002]. Después de etiquetar las componentes se procedió a realizar la eliminación de aquellas que no cumplieron con el tamaño definido. El paso final para la obtención del contorno fue la eliminación de los píxeles con índice 0; las definiciones relacionadas están en el Anexo IV.

A continuación, el patrón acústico se construye asignándole una frecuencia a cada pixel según su posición y valor dentro de la matriz de la imagen binaria obtenida. La asignación de frecuencias a los píxeles de la columna  $i$  de la imagen se muestra en la Tabla 4.1.

Tabla 4.1. Asignación de las frecuencias en la columna *I*.

Pixel	Frecuencia	Pixel	Frecuencia	Pixel	Frecuencia	Pixel	Frecuencia	Pixel	Frecuencia
(I, 0)	3520.00	(I, 10)	1318.50	(I, 20)	493.38	(I, 30)	174.61	(I, 40)	65.406
(I, 1)	3136.00	(I, 11)	1174.70	(I, 21)	440.00	(I, 31)	164.81	(I, 41)	61.735
(I, 2)	2793.80	(I, 12)	1046.50	(I, 22)	392.00	(I, 32)	146.83	(I, 42)	55.00
(I, 3)	2637.00	(I, 13)	987.77	(I, 23)	349.23	(I, 33)	130.81	(I, 43)	48.999
(I, 4)	2349.30	(I, 14)	880.00	(I, 24)	329.63	(I, 34)	123.47	(I, 44)	43.654
(I, 5)	2093.00	(I, 15)	782.99	(I, 25)	293.66	(I, 35)	110.00	(I, 45)	41.203
(I, 6)	1975.50	(I, 16)	698.46	(I, 26)	261.63	(I, 36)	97.999	(I, 46)	36.708
(I, 7)	1760.00	(I, 17)	659.26	(I, 27)	246.94	(I, 37)	87.307	(I, 47)	32.703
(I, 8)	1568.00	(I, 18)	587.33	(I, 28)	220.00	(I, 38)	82.407	(I, 48)	30.868
(I, 9)	1396.90	(I, 19)	523.25	(I, 29)	196.00	(I, 39)	73.416	(I, 49)	25.50

Durante la lectura de la imagen binaria sólo las frecuencias de los pixeles negros de cada columna contribuyen a la construcción del patrón acústico. Usando esta asignación generamos el patrón acústico usando CSound. Esto se logró mediante la construcción de una aplicación de software realizada en el ambiente de desarrollo rápido Delphi que se encarga de la traducción de la imagen binaria de 50 x 50 pixeles en los archivos necesarios para que CSound genere su patrón acústico.

Para construir el archivo orquesta se buscan todos los renglones que tienen al menos un pixel negro. Por cada uno de estos renglones se define un instrumento usando el opcode oscil con la amplitud de 10000 y la frecuencia correspondiente a ese renglón, esta definición hará que en su momento se genere un patrón acústico de una onda senoidal de amplitud 10000 y la frecuencia dada.

Nuestra meta es generar un patrón acústico de duración de a lo más 3s, usando multiplexación en el tiempo. Esto se traduce en que cada columna de la imagen determina un subpatrón de duración de 0.06s. En el archivo partitura se define el orden en el que los instrumentos generarán el patrón, el tiempo en el que empezarán a ejecutar su frecuencia está determinado por la columna en el que esté el pixel asociado. Por ejemplo, un pixel

negro con coordenadas  $(x,y)$  es asociado al instrumento que tiene la frecuencia asignada a la pareja  $(l,y)$  de la Tabla 4.1. La coordenada  $x$  indica que el instrumento y se comenzará a ejecutar al tiempo  $(0.06)x$  con una duración de 0.06s. Al procesar cada columna se “tocarán” tantos instrumentos como pixeles negros haya en esa columna, al pasar a la siguiente tocarán los instrumentos asociados a los pixeles de esa segunda columna sin importar si ya han sido utilizados antes. Esto se verá reflejado en la parte de notas del archivo partitura en donde un mismo instrumento puede aparecer varias veces en la lista.

Para nuestros experimentos grabamos previamente en el disco duro los patrones de nuestro sistema básico en el formato .wav<sup>44</sup>. En la Figura 4.3 se puede ver la interfaz de nuestra aplicación.

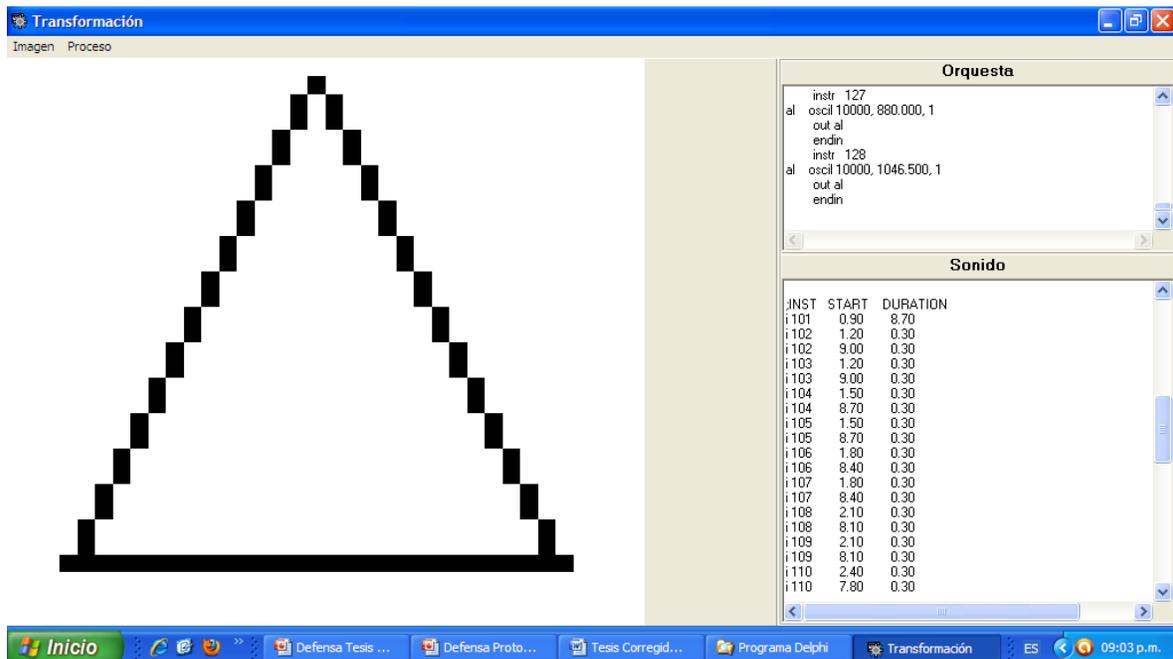


Figura 4.3 Interface gráfica para convertir el contorno a patrón acústico

<sup>44</sup> Wav es un formato de archivos de audio en forma de onda, también llamado formato de audio para Windows. Es el principal formato utilizado en sistemas Windows para audio no comprimido.

## 2. El taller

Haciendo las primeras pruebas nos dimos cuenta de que se necesitaba entrenar a los sujetos en el uso de nuestra codificación. Con base a los planteamientos hechos por Dondis [Dondis, 2000], basados en las técnicas de la psicología Gestalt<sup>45</sup>, partimos de que la forma de cualquier imagen es la combinación de figuras geométricas simples (sencillas), las que a su vez están constituidas por líneas y curvas. Con estos principios se procedió a establecer un primer código básico<sup>46</sup> que estuviera formado por líneas, curvas y figuras geométricas simples. La finalidad de éste fue probar que la codificación propuesta satisfacía nuestros objetivos. En un futuro estableceremos un conjunto de patrones e imágenes a partir de las cuales pudiéramos generar interactivamente (casi) todas las formas.

Apoyándonos en este sistema básico de formas se diseñó un curso de entrenamiento para el desarrollo de competencias cognitivas y de razonamiento [Bolívar, 2002]. Usamos un enfoque constructivista<sup>47</sup> junto con la teoría de Murphy [Murphy,

---

<sup>45</sup> Psicología Gestalt.- Corriente de la psicología moderna, surgida en Alemania a principios del siglo XX, y cuyos exponentes más reconocidos han sido los teóricos Max Wertheimer, Wolfgang Köhler, Kurt Koffka y Kurt Lewin. Es importante distinguirla de la Terapia Gestalt, terapia exponente de la corriente humanista, fundada por Fritz Perls, y que surgió en Estados Unidos. El término Gestalt proviene del alemán y fue introducido por primera vez por Christian von Ehrenfels. No tiene una traducción única, aunque se lo entiende generalmente como "forma". Sin embargo, también podría traducirse como "figura", "configuración" e, incluso, "estructura" o "creación".

<sup>46</sup> En un principio pensamos llamarle al conjunto básico *alfabeto*, sin embargo, nos decantamos por el nombre de *Sistema Básico*, después de revisar la definición de Alfabeto que encontramos en los diccionarios. Alfabeto: El alfabeto de una lengua o idioma es el conjunto ordenado de sus letras. Es también la agrupación, con un orden determinado, de las grafías utilizadas para representar el lenguaje que sirve de sistema de comunicación, (recuperado 1 de octubre de 2011 de: <http://es.wikipedia.org/wiki/Alfabeto>). En ciencias de la computación un alfabeto es un conjunto finito no vacío.

<sup>47</sup> Constructivista.- Con el nombre de constructivismo se denominan distintas corrientes surgidas en el arte, la filosofía, la psicología, la pedagogía y las ciencias sociales en general: 1. En Filosofía de la Ciencia (epistemología) se denomina constructivismo a una corriente de pensamiento surgida a mediados del siglo XX, de la mano de investigadores de disciplinas muy diversas (psiquiatras, físicos, matemáticos, biólogos, psicólogos, etc.); 2. En Pedagogía el constructivismo es una corriente de la didáctica que para sus postulados didácticos supone una teoría del conocimiento constructivista; 3. En Matemáticas el constructivismo es una filosofía que afirma que es necesario encontrar (o "construir") un objeto matemático para poder probar su existencia; 4. En Psicología el constructivismo está basado en los postulados piagetianos en los que el desarrollo de la inteligencia es construido por el propio individuo a través de la interacción con el medio. En

1995]. Dicha teoría establece que los sentidos se apoyan unos a otros para incrementar su eficiencia sensorial. Por esta razón basamos la primera parte del curso de entrenamiento en una mezcla del sentido del tacto y el auditivo.

Partiendo de que la experiencia se traduce en conocimiento, hacemos que el invidente a través del tacto perciba en forma resaltada los contornos de las figuras con la finalidad de que adquiera la experiencia táctil y la consciencia de las formas, posteriormente se hace una exposición al patrón acústico correspondiente a la forma presentada; lo que al *asociarlos* en su mente resulta en conocimiento, es decir, a una forma detectada por el tacto se aprende que le corresponde un código específico de sonido [Figura 4.4].

Un punto importante al que tenemos que hacer referencia es el de la asociación, ésta no solamente incluye la integración de objetos provenientes de propiedades sensorias individuales, sino también el encadenamiento de objetos ya integrados, la percepción o representación de ordenaciones espaciales o secuencias temporales y la vinculación de una experiencia con otra. Al codificar el contorno de una imagen en patrón acústico, queda implícito el encadenamiento de ordenaciones espaciales en función de la frecuencia y tiempo, las cuales al percibir las el invidente serán decodificadas y con base a su conocimiento podrá representar en su mente la forma del objeto correspondiente, como una manera de comprobarlo se les hará dibujar a algunos invidentes las formas correspondientes a algunos patrones acústicos elegidos al azar.

Cuando hablamos de conocimiento (idea) de la forma, se hace referencia a la hipótesis básica de que el sujeto debe construir objetos conceptualmente antes de que pueda hacer algo con ellos: “Saber quiere decir poder comprender”. El saber conceptual consiste en modelos que nos permiten orientarnos en el mundo de la experiencia, prever situaciones y a veces manejar incluso nuevas experiencias. Finalmente, podemos decir que el mundo de la experiencia es exclusivamente un mundo que construimos con conceptos que producimos “según el proyecto de nuestra razón” [Walzlawick, 1994].

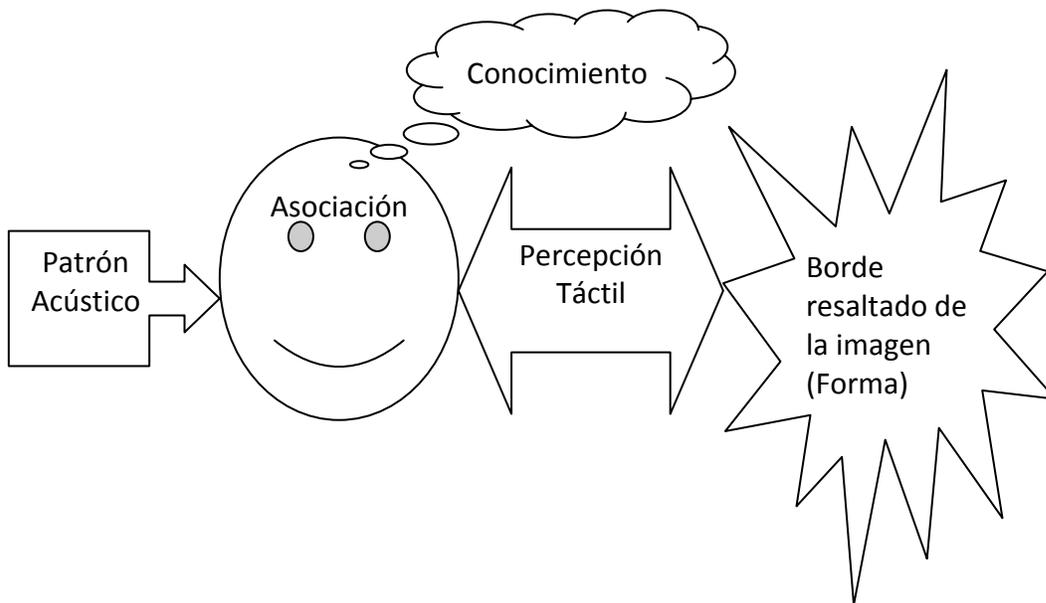


Figura 4.4 Se observa cómo se lleva a cabo la asociación del patrón acústico con la percepción táctil.



# Capítulo 5

## Experimentos y resultados

### 1. Procesamiento de la imagen

En esta sección se describen y presentan los detalles restantes de los experimentos efectuados para llevar a cabo este trabajo de investigación. Ya dimos previamente algunos detalles de éstos en los capítulos 3 y 4. A continuación, se presentan los resultados obtenidos mediante la combinación de algoritmos referidos en el Capítulo anterior, Figura 5.1.

Primero convertimos la imagen a color en una imagen en tonos, de gris después con el filtro pasa alto acentuamos los bordes de todos los objetos de la imagen, alteramos la distribución de los niveles de intensidad mediante el uso del histograma, finalmente usamos un operador umbral, eliminamos hoyos y objetos demasiados pequeños para ser

tomados en cuenta, Figura 5.1. Más ejemplos de los resultados de estos procedimientos se muestran en la Figura 5.2.

#### Etapas del procedimiento

			
Cambio de imagen de color a tonos de grises.	Filtro pasa alto, con la máscara: $\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	Histograma, Max. 218, Min. 38	Operador umbral 125 (imagen binaria). Eliminación de Hoyos con tamaño de a lo más 1,200 píxeles y eliminación de objetos en negro de tamaño menor o igual a 4,054.
	Acentuación de bordes	Variación de la distribución de los niveles de intensidad.	Creación de imagen binaria.

Figura 5.1 Se presenta el procesamiento de la imagen siguiendo los algoritmos propuestos.

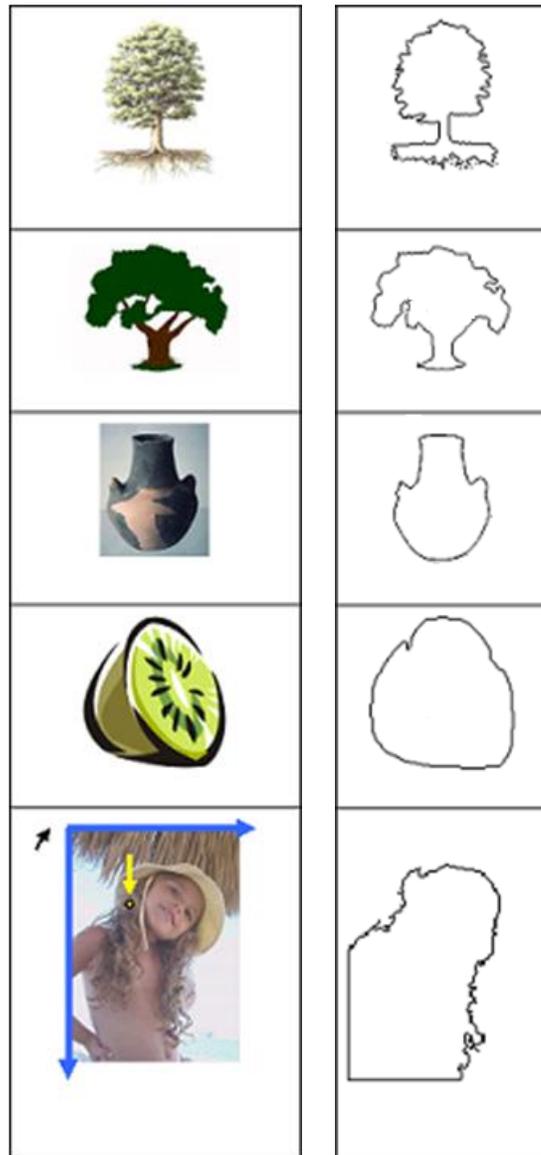


Figura 5.2 Ejemplificación del resultado de algunas imágenes procesadas con los algoritmos propuestos.

## **2. Sistema básico experimental propuesto**

En un principio observamos que los patrones acústicos que representaban imágenes correspondientes a cosas de nuestro entorno no tuvieron ningún significado (de entrada) para los invidentes; para ellos fue como escuchar un lenguaje desconocido por primera vez. Fue entonces, que nació la idea de proponer patrones acústicos que representaran imágenes muy simples.

Como se mencionó en el Capítulo 3, partimos de la idea de que todos los contornos en la naturaleza están contruidos a partir de líneas y (trozos de) curvas sencillas. Con base en esto, propusimos el siguiente sistema básico constituido por 33 figuras geométricas, Tabla 5.1.

Acto seguido desarrollamos los 33 patrones acústicos correspondientes. Las imágenes con las que trabajamos en esta etapa poseen un tamaño 50 por 50 pixeles; el patrón acústico que generamos con cada una de ellas tiene una duración de 3 segundos. La duración del patrón fue consensada con algunos usuarios que manifestaron que así les era más fácil de percibir y aprender.

Tabla 5.1. Sistema básico propuesto de 33 figuras geométricas, el cual representó los patrones acústicos utilizados en nuestros experimentos.

Patrones con una línea recta							
							
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Patrones con dos líneas rectas							
							
(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
							
(17)	(18)	(19)	(20)				
Patrones con tres líneas rectas							
							
(21)	(22)	(23)	(24)	(25)	(26)		
Patrones con cuatro líneas rectas							
							
(27)	(28)	(29)	(30)				
Patrones con líneas curvas							
							
(31)	(32)	(33)					

Nos dimos a la tarea de conformar dos grupos:

### **3. Cuadrados latinos y estadística**

En cualquier experimento, la variabilidad que surge de un factor perturbador puede afectar los resultados. En general, un factor perturbador puede definirse como un factor del diseño que probablemente tenga un efecto sobre la respuesta. El cuadrado latino se usa para eliminar fuentes de variabilidad perturbadora [Mongomery, 2007], [Little, 2002]. Por esta razón para evaluar el éxito en el aprendizaje de la decodificación de los patrones acústicos por parte de los invidentes se decidió usar los cuadrados latinos.

Se utilizaron dos cuadrados latinos de 8x8, uno conformado por sujetos con visión normal y el otro por invidentes. Desde un principio nuestra meta fue determinar si ambos grupos desarrollaban la misma habilidad para decodificar los patrones acústicos o, si esta habilidad era diferente debido a sus experiencias previas y el desarrollo de conceptos relativos a su visión.

Para esto se utilizó la prueba *t-student*, una prueba estadística que se utiliza para muestras pequeñas menores a 30 datos, que permite decidir si existe o no una diferencia estadísticamente significativa entre dos grupos, y entonces poder afirmar que las dos muestras corresponden a distribuciones de probabilidad de media poblacional distinta, o por el contrario afirmar que la diferencia de medias puede deberse a oscilaciones estadísticas aleatorias.

### **4. Taller de capacitación**

Para la implementación del taller de capacitación se formaron dos grupos de 8 miembros cada uno, de los cuales fueron 4 hombres y 4 mujeres. Los miembros del grupo

con visión normal se caracterizaron por tener una educación musical; 3 de ellos profesionales de la música (una mujer y dos hombres): un alumno de Bellas Artes, perteneciente a la Universidad Autónoma Benito Juárez (UABJO), un músico de Tlahuitoltepec, Oaxaca y uno de la Escuela Nacional de Música de la Ciudad de México; el resto toca o tocó en algún momento de su vida un instrumento musical, sus edades están dentro del rango de 17 a 43 años, y su escolaridad varió desde estudios de preparatoria hasta licenciatura.

La finalidad de esta elección fue asegurarnos que el centro auditivo musical de la corteza cerebral estuviera acostumbrado a los tonos musicales para que no estuvieran en desventaja con los invidentes, ya que se supone que los invidentes, de alguna forma, han desarrollado su sistema auditivo para ambientarse mejor en los espacios acústicos que los circundan, ver la definición de espacio extrapersonal de Hall [Strumillo, 2010].

Para tener contacto con el grupo de invidentes se visitó al DIF, donde nos canalizaron a la Biblioteca para Ciegos “Jorge Luis Borges”, en la ciudad de Oaxaca<sup>48</sup>. Las edades del grupo de invidentes fluctuaron entre 12 y 23 años; 7 de ellos ciegos de nacimiento y el octavo perdió la vista a los 4 años, éste aseguró que ya no recordaba que era ver, la vista la perdió debido a una enfermedad de la que no supo darnos el nombre. Son usuarios habituales de la biblioteca para ciegos; usan el bastón blanco como dispositivo de ayuda para navegación, ninguno tiene perro guía. Para la selección de los sujetos invidentes el criterio fue “son constantes, cooperativos y leen en Braille”.

---

<sup>48</sup> Esta biblioteca depende de la Biblioteca del Instituto de Artes Gráficas de Oaxaca (IAGO) en la misma ciudad. El IAGO tiene un directorio de 60 personas ciegas en el estado de Oaxaca, la mayoría de ellas jóvenes. Aunque mencionan que el número de personas ciegas en todo el estado se estima en una cifra cercana a 500.

Tanto el taller de entrenamiento como las pruebas con los patrones acústicos realizadas con los invidentes se llevaron a cabo en las instalaciones de la Biblioteca “Jorge Luis Borges” y en las del CIIDIR Oaxaca del IPN. Todos los sujetos invidentes así como los de visión normal asistieron en forma individual al entrenamiento durante 3 sesiones cada uno, con una duración de 45 minutos aproximadamente por sesión.

En la primera sesión los invidentes palparon cada una de las tarjetas, en donde cada una de ellas tenía resaltada con fomi la imagen correspondiente (Figura 5.3). Los sujetos con visión normal fueron vendados de los ojos con una pañoleta para que no vieran las tarjetas, estuvieran en condiciones similares a las de los invidentes y lo hicieron durante todas sus sesiones. Conforme el instructor les proporcionaba cada tarjeta, les repetía el nombre de la imagen correspondiente. Después de palpar las 33 tarjetas dentro de esa misma sesión, se repetía la exploración táctil de cada tarjeta pero ahora escuchando el patrón acústico correspondiente.

En la segunda sesión se repitieron las mismas acciones que en la primera y en la tercera sesión, ya no se usaron las tarjetas, sólo los patrones acústicos. En la posterior etapa de evaluación se utilizaron de los 33 patrones solamente 8 seleccionados aleatoriamente, siguiendo la secuencia específica para cada sujeto, según lo dictó el cuadrado latino al que estuvieron asociados. Cada uno de los estímulos se expuso tres veces consecutivas, con la finalidad de darle oportunidad al sujeto para que decodificara el patrón acústico<sup>49</sup>.

---

<sup>49</sup> Cuadrado latino.- Un cuadrado latino es una matriz de  $n \times n$  elementos, en la que cada casilla está ocupada por uno de los  $n$  símbolos de tal modo que cada uno de ellos aparece exactamente una vez en cada columna y en cada fila. Los cuadrados



Figura 5.3 Contornos de las imágenes sobre las tarjetas.

Para dos invidentes hubo una cuarta y última prueba. La de dibujo, esta prueba se planteó con la finalidad de demostrar la habilidad adquirida por el invidente de poder decodificar los patrones acústicos, para ello se les pidió que dibujaran sobre una hoja de papel la interpretación mental del patrón.

Para la experimentación se utilizó una laptop común; para la reproducción de los patrones acústicos se usó una tarjeta de sonido Sound Blaster Creative modelo 580270 y un juego de altoparlantes marca Zodiac Multimedia Speaker. Los sujetos involucrados en el taller de capacitación, durante las pruebas con los patrones acústicos, fueron situados a 1 metro de distancia de la fuente de sonido. Durante el taller nos hicimos conscientes de todos los problemas que enfrentan los ciegos para poder navegar autónomamente.

## 5. Resultados de las pruebas

Los resultados obtenidos después de llevar a cabo las pruebas en los dos grupos se presentan en la Tabla 5.2. Finalmente, nos propusimos hacer una última prueba (subjetiva) con los invidentes; les pedimos a algunos invidentes que dibujaran los patrones acústicos que en forma aleatoria les serían expuestos. Esto fue con el fin de lograr demostrar la capacidad lograda por los invidentes para decodificarlos y comprobar que efectivamente construían en sus mentes las formas presentadas, Tabla 5.3 y 5.4. Los dibujos hechos por nuestros invidentes voluntarios son muchos más complejos que los que se pueden ver en [Arno, 2001], o en [Rovira, 2010]. Después de llevar a cabo esta meta, consideramos que los resultados mostrados en la Tabla 5.3 y 5.4 constatan la capacidad del cerebro humano para recibir y procesar información visual suministrada por el oído.

Al final del entrenamiento y de la etapa evaluativa, los invidentes manifestaron su inquietud por aprender más, ya que habían adquirido nuevos conceptos y eso los había motivado mucho ... y formalizaron la siguiente petición: ¿Qué más sigue?

Tabla 5.2. Porcentaje de identificación de los patrones acústicos.

PA	Descripción de la imagen	Porcentaje de identificación de sujetos con visión normal	Porcentaje de identificación de sujetos invidentes
A	Vertical. izquierda y horizontal alta	100	100
B	Círculo	50	50
C	Vertical. izquierda y horizontal alta y vertical derecha	62.5	100
D	Vertical	100	100
E	Cuadrado	75	100
F	Vertical. izquierda y horizontal baja	100	100
G	Horizontal baja	100	87.5
H	Ascendente y descendente	87.5	50

Tabla 5.3. Dibujos realizados por los invidentes interpretando los patrones acústicos. Primera parte.

Tabla 5.4. Dibujos realizados por los invidentes interpretando los patrones acústicos. Segunda parte.



# Capítulo 6

## Discusión, Conclusiones y Trabajo Futuro

### 1. Discusión

Para situar nuestro trabajo en contexto haremos un análisis del Visor Acústico y otras tres codificaciones de imágenes a sonido: la desarrollada por Meijer, publicada en [Meijer, 1992], que forma parte de vOICe; la desarrollada por el equipo conformado por Capelle, Trullemans, Arno y Veraart, presentada en [Capelle, 1998], parte del PSVA (Prosthesis Substituting Vision by Audition) y la desarrollada por el equipo de Nagarajan, Sainarayanan y Yacoob, tal y como está presentada en [Sainarayanan, 2007], parte de un sistema llamado NAVI (Navigation Assistance for Visually Impaired).

La razón por la que escogimos las dos primeras es que, hasta donde sabemos, son las dos más citadas en la literatura sobre el tema, mientras que la tercera ha tenido un

desarrollo paralelo (en el tiempo) a nuestro trabajo aunque supimos de ésta ya cuando estábamos en el proceso de escritura de esta tesis; ha despertado interés pues tiene ya algunas citas.<sup>50</sup>

A continuación haremos un análisis de estas codificaciones y la del Visor Acústico. En las tablas de este capítulo, la columna rotulada como vOICe se refiere a el trabajo de Meijer, la rotulada como PSVA a la propuesta de Capelle et al., NAVI a la de Sainarayanan et al. y VA a la propuesta presentada en esta tesis.

En el disco compacto que acompaña este trabajo, la presentación de Power Point “Comparación de patrones acústicos” presenta los patrones acústicos generados por la codificación de Meijer y la nuestra de una colección de 5 figuras. El lector podrá apreciar la diferencia entre ambas.

## **1.1. Procesamiento de la imagen**

En la Tabla 6.1 se puede observar la diferencia existente entre las diferentes líneas de investigación en lo que respecta al tamaño de la imagen y cómo se procesa. Todas estas propuestas parten de la idea de que la imagen es capturada por una cámara (de video o fotográfica), Figura 1.1. Como atinadamente se observa en [Sainarayanan, 2007], página

---

<sup>50</sup> El artículo de Meijer tiene 251 citas según Google Académico y 22 según IEEExplorer; el de Capelle et al. tiene 77 y 19 respectivamente; el de Sainarayanan et al. tiene 7 según Google Académico y 9 según Scopus. Es de comentarse que de las 22 citas del trabajo de Meijer reportadas por IEEExplorer, 20 son trabajos publicados del 2005 en adelante, de las 19 del trabajo del equipo de Capelle, 17 son posteriores al 2005. (Información recabada desde las respectivas páginas de internet el 6 de agosto de 2011.) Estos datos sugieren un aumento notable en el interés sobre el tema a partir del 2005. Por cierto, la primera presentación en un congreso de nuestro trabajo data del 2002.

258, el principal problema al que nos enfrentamos en esta primera etapa es que no tenemos un conocimiento a priori del objeto a identificar.

El equipo que propone NAVI y el de VA tienen en común la propuesta de la supresión del fondo de la imagen con la intención de generar un patrón acústico más sencillo. Inspirado esto último, en el hecho de que el sistema de la visión humana crea el mayor enfoque de la visión en la región de interés (fóvea), mientras que las otras regiones son consideradas como fondo y se les da menos consideración [Sainarayanan, 2007], página 259. Esto es algo que no hace Meijer, ni Capelle et al. La diferencia entre NAVI y VA es que mientras NAVI sonifica el objeto en primer plano, nosotros proponemos la sonificación de su contorno y la utilización de algunas frecuencias de la escala bien temperada. Estas últimas decisiones determinan que nuestro patrón acústico sea más sencillo.

La primera publicación en la se presentaron los algoritmos de procesamiento de imágenes de nuestro trabajo es [Martínez, 2005]. Meijer es el único de los autores que justifica el tamaño escogido para las imágenes que sonifica, [Meijer, 1992], página 115. A partir de la fórmula (8) de [Meijer, 1992] el autor afirma que la imagen debe de tener un tamaño máximo de 50 píxeles. Sin embargo, se usa una imagen de 64x64 alegando que es más conveniente para una implementación digital. Los demás no dan ninguna información al respecto. Nosotros escogimos el tamaño máximo propuesto por [Meijer, 1992] aunque, como se verá más adelante, nuestros patrones son más largos, Tabla 6.2.

Tabla 6.1 Análisis comparativo de las diferentes etapas de procesamiento de la imagen.

vOICe	PSVA	NAVI	VA
La cámara proporciona una imagen de la que se toma una parte central de 256 por 256, de la que a su vez se escoge una de cada cuatro columnas (y renglones) para obtener una imagen de 64x64 pixeles, la que finalmente se sonificará.	La imagen original es de 256x256 pixeles, se recorta de ella la parte central de tamaño 64x64 pixeles.  La matriz de la imagen primero se convoluciona con un detector de aristas y después se convierte en una imagen con multiresolución para simular una retina. (El trabajo presenta dos posibilidades para la multiresolución.)	El tamaño de la imagen es de 32x 32 pixeles.  Mediante el entrenamiento de una red neuronal para la clasificación FLVQ (Fuzzy based Learning Vector Quantization) se genera una imagen con cuatro tonos de gris.  En blanco aparece el objeto a sonificar, los pixeles en otros colores no intervienen en la sonificación.	Extracción del contorno de un objeto situado enfrente de la cámara (primer plano), mediante el uso de algoritmos de procesamiento de la imagen en el espacio y de la topología digital, dando como resultado final una imagen binaria.

En el diseño del VA se buscó simplificar la imagen con la condición de conservar características que faciliten el reconocimiento de los objetos. Se tomó la decisión de buscar un procedimiento que extrajera el contorno del objeto que aparezca en primer plano pues en [Hoffman, 1985], página. 65, se puede leer “...su sistema visual está equipado para describir la forma de un objeto y para conjeturar cuál objeto es a partir de su contorno.... esta conjetura puede no ser exactamente correcta; puede simplemente

limitar las coincidencias potenciales y provocar cálculos visuales para limitar aún más estas coincidencias.”<sup>51</sup> (Las cursivas son del autor de este trabajo.)

Tabla 6.2 Tamaños de la imagen a procesar, según la línea de investigación.

vOICe	PSVA	NAVI	VA
64x64 píxeles	Una imagen de 64x64 píxeles se transforma en una con 124 píxeles en total (retina artificial con dos niveles de resolución) o una con 208 (retina artificial con cuatro niveles de resolución)	32x32 píxeles.	50 x 50 píxeles.

En el diseño del PSVA se hace una simulación de la fisiología de la retina, lo que no estamos seguros pudiera ayudar a reconocer el objeto presente en la imagen. Ni en el proyecto vOICe, ni en el de NAVI, ni en el nuestro se encuentra una preocupación acerca de modelar la fisiología del ojo humano.

Cada propuesta tiene su propia elección de tonos de gris de la imagen a sonificar. En general el tono de gris de un píxel determinará la amplitud de la onda sonora asociada a éste. En nuestro caso, trabajaremos con imágenes binarias, Tabla 6.3.

<sup>51</sup> En inglés: “...your visual system is equipped to describe the shape of an object and to guess what the object is from the outline. This guess....might not be exactly correct; it may simply narrow the potential matches and trigger visual computations designed to narrow them further.”

Tabla 6.3 Número de tonalidades de gris utilizados en la imagen a sonificar.

vOICe	PSVA	NAVI	VA
16 tonos de gris. No se especifica cómo son escogidos.	256	4 tonos de gris: negro, blanco, gris oscuro y gris claro.	Blanco y negro

## 1.2. Diseño del patrón acústico

Nuestra propuesta de diseño y generación del patrón acústico se presentó por primera vez en las décima y onceava ediciones del Congreso Internacional Mexicano de Acústica (2003 y 2004 respectivamente). La estrategia de sonificación que utilizamos se mencionó en el resumen publicado en la revista *The Journal of the Acoustical Society of America* en el 2010 [Martínez, 2010] y en el artículo [Martínez, 2011] publicado en agosto de este año.

Algunos investigadores en el área de Interacción Humano-Computadora han encontrado que agregando patrones sonoros a algunas interfaces se puede obtener una mejora de su usabilidad [Frauenberger, 2007]. Para esto han diseñado, entre otras cosas, *earcons*, melodías musicales abstractas que simbolizan objetos o tareas [Power, 2010]. Estas melodías se restringen a muy pocas notas, frecuentemente no más de cuatro [Hankinson, 1999].

Siguiendo este orden de ideas podemos afirmar que la propuesta del VA incluye la construcción de patrones acústicos asociados a imágenes binarias similares a los *earcons* aunque utilizando 50 frecuencias de la escala bien temperada, asociadas a las correspondientes teclas blancas del piano, dichas frecuencias están en la zona de las

curvas de igual sonoridad, en donde el oído es más sensible [Giacomin, 2010] y [Fletcher, 1933]. Dichos patrones acústicos describirían el contorno de los objetos que el invidente tiene en un primer plano y enfrente de él.

El uso de las frecuencias de la escala bien temperada nos pareció ventajoso por su relación con la música, de la que los investigadores han encontrado que los adultos pueden reconocer miles de melodías y que incluso los niños pequeños son capaces de reconocer una sucesión de tonos [Trehub, 2003].

En consecuencia, divergiendo de las otras líneas de investigación el rango de las frecuencias usadas varía de los 25.5 Hz a los 3520 Hz. En el caso del PSVA sólo se hace (a grandes rasgos) un análisis de frecuencia de la información auditiva que llega al oído interno. En general podemos afirmar que la preocupación principal de los diseñadores de todas las propuestas, excepto de la del VA, es que las diferentes frecuencias que conforman los patrones sonoros sean reconocibles pero no que los patrones, como un todo, sean fáciles de aprender o recordar o de decodificar, Tabla 6.4. La estrategia seguida en este trabajo no es usada en ninguna de las otras codificaciones de las que hemos tenido noticia.

En la propuesta del NAVI y en la de PSVA se generan dos patrones sonoros distintos, uno para cada oído. Las demás propuestas sólo contemplan la generación de un patrón que será enviado a ambos oídos simultáneamente.

En cada propuesta las frecuencias están asignadas de forma tal que las más bajas corresponden a la parte inferior de la imagen incrementándose progresivamente hasta tener las frecuencias más altas en la parte superior de la misma.

Tabla 6.4 Diseño del Patrón acústico.

vOICe	PSVA	NAVI	VA
Cada pixel tiene asociado un oscilador senoidal con frecuencia elegida dentro del rango de las audibles en una distribución exponencial.	Las frecuencias se escogieron en el rango de 50Hz-15kHz. La asignación de las frecuencias a los pixeles en la retina artificial con dos niveles de resolución se asigna con dos reglas de correspondencia distintas, una correspondiente a la zona periférica y otra a la fovea. Ambas dependen del número del pixel.  En el caso de la retina con cuatro niveles son cuatro las reglas de correspondencia.	Las frecuencias del sonido producido fueron diseñadas dentro del rango audible humano de 20Hz a 20kHz, de éste se seleccionó una banda de frecuencias en el rango de frecuencias bajas.	25.5 Hz - 3520 Hz. De aquí se escogieron las frecuencias de la escala bien temperada correspondientes a 50 teclas del piano.

En el caso de la sonificación del PSVA, cada uno de los patrones sonoros se genera, en principio, de una forma que toma en cuenta TODOS los pixeles de la imagen simultáneamente. En las restantes propuestas el patrón se genera columna por columna. A cada pixel se le asocia una frecuencia según su posición vertical, su posición horizontal

determina el tiempo (dentro del patrón sonoro) y su tono de gris las amplitudes de las ondas senoidales.

En la propuesta NAVI el patrón para el oído izquierdo se genera leyendo la imagen desde su extremo izquierdo hasta el centro de la misma y el del oído derecho leyendo del extremo derecho hacia el centro.

Con respecto a la duración del patrón sonoro en vOICe es de 1.05 s, [Meijer, 1992], página 116, mientras que la del VA es de 3.0 s. Esta última se escogió así porque al hacer las primeras pruebas con los sujetos videntes e invidentes, estos sugirieron que era el tiempo adecuado para poder lograr decodificar el patrón acústico. Dicho se ha de paso, nuestra propuesta es la única en la que se informa de la interacción entre el autor y los posibles usuarios en esta etapa del diseño. Las duraciones de los patrones generados por las otras propuestas no se mencionan claramente.

Todos los proyectos mencionados, excepto el VA, generan sus patrones trabajando directamente con las tarjetas de sonido u osciladores senoidales. Por su parte, el VA aprovecha todo el trabajo y la experiencia detrás de CSound [Boulander, 2000].

Según Boulander, CSound es un lenguaje de programación y paquete de software libre orientado a crear, editar, analizar y componer música y sonido. Puede funcionar en computadoras con diversos sistemas operativos. Esto se traduce en que no nos tenemos que preocupar (al menos en este momento) por las especificaciones de la tarjeta de sonido, lo que nos facilitó la generación de los patrones pero nos forzó a construir las interfaces necesarias para la sonificación de los contornos por CSound.

En todas las propuestas aquí presentadas el tamaño de la imagen y el tiempo de barrido de la imagen son constantes y su transformación en sonido es acotada en el tiempo. Los usuarios utilizan audífonos en todos los sistemas hasta ahora creados, aislándolos parcialmente del ambiente acústico normal. Por lo general se espera que el usuario tendrá acceso al control de la intensidad del sonido.

La decisión de sonificar una imagen binaria ayuda a definir, explícitamente y para todas las imágenes, una codificación con un volumen constante. Esto puede ayudar en el caso del VA, a que la señal sea “resistente” al ruido ambiental. Si el ruido ambiental es fuerte el aumento del volumen del patrón acústico puede enmascarar a este último. En una codificación con volumen variable se podría tener mayor sensibilidad al ruido, como en el caso de los otros trabajos que estamos considerando. Por otro lado, en los trabajos aquí considerados se definen codificaciones con volumen variable pero, sólo se hacen pruebas (en el trabajo del NAVI y en el PSVA) con imágenes binarias que generan señales con ¡volumen constante!, Tabla 6.5.

Por último, NAVI y vOICe presentan patrones acústicos de un poco más de un segundo de duración por cada imagen. En el trabajo del equipo de Capelle no se es muy preciso con respecto a la duración del patrón, pareciera que se genera un patrón de alrededor de una centésima de segundo de duración por cada cuadro de video. En el caso del VA el patrón tiene una duración de 3 segundos, este tiempo fue encontrado como el óptimo para que nuestros sujetos experimentales pudieran aprender y reconocer los patrones propuestos. Somos el único equipo que trabaja en esta etapa con los posibles usuarios.

Tabla 6.5 Tolerancia al ruido

vOICe	PSVA	NAVI	VA
No es relevante en este momento (1992) para el autor.	La intensidad del sonido puede ser ajustada por los usuarios.  Preocupación sobre la calidad de la imagen original es inexistente. No contemplan el posible enmascaramiento de los sonidos.	No hay una preocupación explícita por este problema.	Con respecto al sonido ambiental: el patrón tiene un volumen constante que podría ser elevado manualmente por el usuario en una versión posterior.  Con respecto a la imagen: Mientras el contorno del objeto se preserve, el ruido en la imagen no es relevante.

### 1.3. Experimentación con seres humanos

El trabajar con seres humanos es realmente difícil y más si son invidentes, pues la mayoría no son independientes. En nuestro caso se trabajó con dos grupos, el primer grupo fue conformado por 8 sujetos con visión normal (con los ojos vendados durante todas las sesiones) y el segundo grupo por 8 invidentes de nacimiento, Tabla 6.6.

Tabla 6.6 Experimentación con seres humanos.

vOICe	PSVA	NAVI	VA
<p>Se deja para un trabajo futuro, la preocupación del autor es la “factibilidad de un sistema de bajo costo que preserve mucha información visual en los patrones de sonido generados”.</p> <p>Se han publicado algunos trabajos que documentan las experiencias de los ciegos al usar vOICe. El más reciente es [Ward, 2010], que reporta 2 usuarios.</p>	<p>Dos grupos de 12 <u>seres humanos con los ojos vendados</u>: Uno de ellos fue el grupo de control. El entrenamiento constó de 10 sesiones de una hora más cuatro sesiones de evaluación.</p>	<p>Se menciona el haber hecho pruebas con NAVI pero no se menciona ni el número de sujetos, ni detalles acerca del entrenamiento.</p>	<p>16 voluntarios divididos en dos grupos.</p> <p>Un primer grupo formado por 8 invidentes de nacimiento y el segundo grupo formado por 8 sujetos con visión normal.</p>

Las figuras utilizadas en las pruebas varían. En nuestro caso creamos un código con 33 figuras geométricas que incluye no sólo rectas sino también algunos polígonos y un círculo. Ninguno de los otros tres primeros equipos de investigadores justifica la elección del conjunto de imágenes de prueba. Nosotros nos apoyamos en la idea presentada en [Hoffman, 1985]: “para reconocer una forma, el sistema visual la descompone en partes (más simples)”. Esto nos animó a proponer un conjunto de formas “sencillas”.

Tabla 6.7 Imágenes utilizadas en las pruebas

vOICe	PSVA	NAVI	VA
En [Ward, 2010] se informa que dos invidentes aprenden a usar vOICe después de haberse familiarizado con los patrones de algunas figuras geométricas (no se menciona cuáles con exactitud). Se prefiere que los invidentes aprendan a distinguir los patrones que se generan en su casa o entorno.	15 patrones básicos junto con los obtenidos por medio de rotaciones de 45°, 90°, lo que hizo un total de 50 patrones. Todos ellos formados por líneas rectas y puntos; ninguna curva.	“algunas formas geométricas básicas” página 263...  No se especifica cuáles.	33 imágenes de formas geométricas: rectas, curvas cerradas y polígonos. Además se obtuvieron contornos de imágenes comunes de nuestro entorno visual.

Los diferentes grupos de investigación no son muy explícitos en describir el ambiente de experimentación. Los proyectos PSVA y VA tienen ambientes de entrenamiento y evaluación similares aunque el VA es el único que (en las primeras fases del entrenamiento) utiliza el tacto para ayudar a los ciegos a aprehender las formas asociadas a los patrones acústicos, ver Tabla 6.8.

Realmente no hay mucha información sobre las experiencias que han obtenido los usuarios de estos sistemas. Los usuarios en el sistema desarrollado por el equipo de Capelle manifiestan cierta incomodidad al principio de su entrenamiento en el uso de su sistema. Una usuaria del sistema de Meijer menciona que al utilizarlo “es como regresar a la luz”; en [Ward, 2010] se informa que otros dos usuarios de vOICe necesitaron alrededor de tres meses para adaptarse y después de un largo período de uso intensivo llegaron a percibir objetos en movimiento (lento) y a desarrollar un sentido de la profundidad de manera espontánea.

Tabla 6.8 Ambiente de experimentación

vOICe	PSVA	NAVI	VA
En [Ward, 2010]. Los usuarios reportan sus experiencias en sus respectivas casas.	Los voluntarios estuvieron con los ojos vendados, ninguno de ellos vio nunca ninguno de los estímulos (su figura); sólo oyeron los correspondientes sonidos. Estuvieron sentados frente a una mesa frente a un pizarrón blanco a una distancia de entre 25 a 30 cm; en consecuencia cada estímulo podía caber en la fóvea de la retina artificial con dos niveles de resolución. Todo esto se llevo a cabo en un laboratorio.	Primero se entrena en el reconocimiento básico de algunas formas geométricas, luego se entrena en la detección de obstáculos en un cuarto, se informa de las primeras pruebas de uso en un ambiente exterior.	Se trabajó principalmente en la biblioteca para ciegos y en un cubículo del CIDIIR (Oaxaca). Ambos grupos en la fase básica utilizaron el tacto para hacer un reconocimiento de las figuras. Los sujetos fueron expuestos a un metro de distancia de la fuente de sonido.

En nuestro caso los usuarios invidentes manifestaron con sorpresa la creación de nuevos conceptos de los cuales ni siquiera tenían idea de su existencia, y mostraron una actitud de ansia por aprender más, entre la exclamaciones que hicieron esta la siguiente: ¿qué más sigue?, queremos aprender más.

Tabla 6.9 Duración del entrenamiento

vOICe	PSVA	NAVI	VA
En [Ward, 2010] no es muy claro. Se habla de meses de entrenamiento.	El entrenamiento constó de 10 sesiones de una hora más cuatro sesiones de evaluación.	No se menciona claramente cuánto tiempo requirió el entrenamiento básico.	Participaron los sujetos primero en tres sesiones individuales de 45 minutos cada una, luego en una cuarta sesión para ser evaluados. Algunos invidentes participaron en una quinta sesión en donde dibujaron en papel los patrones que iban escuchando.

Finalmente queremos hacer notar que la codificación del VA es la que, al parecer, requirió menos tiempo de entrenamiento de las cuatro, ver Tabla 6.9. Con esto pensamos que hemos cumplido nuestro objetivo de desarrollar una codificación más sencilla y fácil de aprender y aprehender para los ciegos o incapacitados visualmente.

## 2. Conclusiones

En conclusión podemos decir que cada grupo ha desarrollado su propia propuesta de codificación. Lo que hemos observado en los otros trabajos es que no hay una

preocupación por considerar la manera en los seres humanos reconocemos las formas de los objetos, ni tampoco por desarrollar una señal tomando en cuenta lo que sabemos sobre la capacidad del cerebro (el normal y el de los invidentes) para procesar, recordar y asociar sonidos. También no hay una preocupación por involucrar a los posibles usuarios en el diseño de la señal sonora, ni se aprovecha el trabajo ya hecho de investigadores en Acústica o en líneas relacionadas con la Música y las Ciencias de la Computación.

En la parte de la evaluación de los resultados no parecen prestar atención a la posible saturación de los evaluados a los estímulos presentados, ni buscan desarrollar un curso formal de capacitación para el uso de su codificación desarrollado a partir de alguna teoría de la Pedagogía que además utilice la fusión del tacto con el oído en una primera etapa. Dicha fusión es de mucha importancia para los seres humanos y particularmente para los discapacitados visuales [Verstegen, 2005].

Por el contrario, todos estos puntos se han tenido en cuenta en el desarrollo de nuestra propuesta. Podemos afirmar que nuestra mayor contribución es la propuesta de un esquema de codificación de imagen a sonido que toma en cuenta los puntos antes mencionados. Creemos que los resultados obtenidos indican que estamos dando pasos en la dirección correcta.

Finalmente habría que comentar que los investigadores de las Ciencias de la Cognición todavía debaten acerca de cómo interpretar las experiencias de los usuarios de los modelos propuestos. Nos parece muy interesante el punto de vista de Auvray, [Auvray, 2009] que las interpreta como una transformación, extensión o aumento de las

capacidades perceptivas, más que algo equivalente o reducible a una modalidad sensorial ya existente.

### **3. Contribuciones**

El diseño de un dispositivo de sustitución de la visión por el oído plantea muchos retos. De estos escogimos considerar el planteado por las preguntas: ¿Cómo codificamos la imagen a un patrón acústico?, Y ¿cómo evaluamos su idoneidad? A esto nos avocamos en este trabajo, a continuación presentamos lo que a nuestro criterio son las contribuciones del mismo:

1. Proponemos una arquitectura para los dispositivos no invasivos de sustitución de la visión por la audición.
2. Nuestra principal contribución es un esquema de codificación de imagen a sonido que toma en cuenta algunas de las ideas de la psicología cognitiva moderna.
3. El sonido es desarrollado usando un lenguaje y paquete de software libre aprovechando la experiencia de investigadores de la informática y de la música. Nuestros patrones proponen una nueva aplicación para un software de diseño musical.
4. Nos preocupamos por elegir entre las frecuencias de sonidos más agradables para el oído humano (occidental).
5. Es la primera propuesta en la que en todo el desarrollo se tomó en cuenta las opiniones de los futuros usuarios.

6. Es la primera propuesta que genera patrones a volumen constante.
7. Se diseñó una evaluación con la ayuda de los cuadrados latinos para evitar la saturación de los sujetos experimentales a los estímulos, este fenómeno no ha sido tomado en cuenta anteriormente.
8. El curso desarrollado propone una construcción de earcons de figuras geométricas abstractas, lo que permite generar una alternativa para la comprensión de dichos conceptos (visuales por excelencia). Esfuerzos en facilitar a los invidentes la enseñanza de conceptos matemáticos (relacionados con el cálculo) usando estimulación táctil y sonificación se presentan en [Van Scoy, 2003] y [Ramloll, 2000].

#### **4. Trabajo futuro**

1. Mejora de la extracción del contorno. Durante la escritura de este trabajo hemos llegado a la conclusión de que vale la pena explorar la posibilidad de una mejora en los algoritmos de procesamiento de la imagen utilizados en nuestra codificación. Una de las mejoras a explorar es que el contorno pueda ser expresado con una aproximación más geométrica, es decir “un poco más abstracta”.
2. Continuar con los experimentos de reconocimiento de formas más complicadas y experimentar con el acortamiento de los patrones acústicos. Ya que diversos investigadores sostienen que el reconocimiento de las formas es básicamente una inferencia, esperamos comprobar experimentalmente que los ciegos con

más familiaridad con la codificación puedan reconocer las figuras con tamaños diferentes y trasladadas (pero dentro la imagen) sin necesidad de entrenamiento previo.

3. Utilizamos las frecuencias del piano, sin embargo el siguiente paso es emular el timbre acústico del mismo.
4. Diseñar e implementar una herramienta de software que haga todo el proceso automáticamente.
5. Construir un prototipo del Visor Acústico portátil que además permita que el usuario elija entre una o dos velocidades de conversión a patrón acústico y del volumen del patrón acústico.
6. Diseñar un curso de entrenamiento que permita a los invidentes reconocer patrones acústicos resultantes de imágenes construidas a partir del pegado de formas más sencillas. Esto conlleva a investigar en la posibilidad de construir o definir un alfabeto a partir del cual los ciegos puedan interpretar cualquier patrón acústico generado por el VA. También evaluar qué tan ventajoso sería agrandar el tamaño de las imágenes a 88 x 88 pixeles, aprovechando todas las frecuencias (todos los sonidos) del piano.



## Referencias

- [Adli, 2007] Adli A., Nakao Z., Yokoda T., Nagata Y., *"Piano sound characteristics: a study on some factors affecting loudness in digital and acoustics pianos"*, Second International Conference on Innovative Computing, Informatio and Control (ICICIC), 2007.
- [Auvray, 2007] Auvray M., Hanneton S., O'Regan J. K., *"Learning to perceive eith visuo-auditory system: Localization and object recognition with The vOICe"*, Séminaire Espaces d'Action, Espaces de Perception, 21-29 January Compiègne, France, 2003.
- [Auvray, 2009] Auvray M., Myin E., *"Perception with compensatory devices"*, Cognitive Science 33, pp. 1036-1058, 2009.
- [Aponte, 2009] Aponte Y.: *"Secundarias para ciegos"*. La Gaceta de la Univ. Guadalajara, 7, Edición 570, Lunes 9 de febrero 16, 2009.
- [Arnhein, 1999] Arnhein R., *"Arte y percepción visual, Psicología del ojo creador"*, Alianza Editorial, 1999
- [Arno, 2001] Arno P., Vanlierde A., Streel E., Wanet-Defalque M., Sanabria-Bohorquez S., Veraart C., *"Auditory substitution of vision pattern recognition by the blind"*, Applied Cognitive Psychology, Vol. 15, pp. 509-519, 2001.
- [Aycinema, 2008] Aycinema P., *"Access for all"*, Communications of the ACM, Vol. 51, No. 8, pp. 12-14, 2008.
- [Bach-y-Rita, 1998] Bach-y-Rita P., Kaczmarek, K. A., Meier, K., *"The tongue as a man machine interface: a wireless communication system"*, Proceedings of the 1998 International Symposium on Information Theory and its Application, Mexico City, Mexico, October 15-16, pp. 79-81, 1998.
- [Balakrishnan, 2005] Balakrishnan, G., Sainarayanan G., Nagarajan R., Yaacob S. , *"Fuzzy Matching Scheme for Stereo Vision based Electronic Travel Aid"*, TENCON 2005 IEEE Region 10, pp. 1-4, Nov. 2005.
- [Bartyna, 2005] Bartyna W., Ambroszkiewicz S., Faderewski M., Jakubowsky S., Kocieliński D., Mikulowsky D., Terlikowski G., *"Blind-enT: Making objects visible for blind people"*, Vision 2005: Proceedings of the International Congress held between 4 and 7 April 2005 in London, UK, International Congress Series vol. 1282, pp. 974-979, September 2005.
- [Baskett, 2005] Baskett L., *"Fear of orientation and mobility"*, Vision 2005: Proceedings of the International Congress held between 4 and 7 April 2005 in London, UK, International Congress Series vol. 1282, pp. 469-473, September 2005.
- [Bialistock, 2005] Bialistock R., *"Towards better communication, from the interest point of view. Or – skills of sight-glish for the blind and visually impaired"*, Vision 2005: Proceedings of the International Congress held between 4 and 7 April 2005 in London, UK, International Congress Series vol. 1282, pp. 793-795, September 2005.
- [Bolivar, 2002] Bolivar C., *"Más allá de la formación: el desarrollo de competencias"* (Consulta: 12 de julio de 2006).  
Disponibile en:

- [http://www.gestionescolar.cl/userFiles/P001/Image/gestion\\_portada/documento](http://www.gestionescolar.cl/userFiles/P001/Image/gestion_portada/documento)
- [Boulanger, 2000] Boulanger R., *"The CSound Book: Perspectives in Software Analysis, Sound Design, Signal Processing, and Programming"*, USA: The MIT Press, 2000.
- [Cano, 2011] Cano M. B. D., Vergara V. O. O., Cruz S. V. V, Ochoa D. H. de J., Maynez L., *"Visual Perception Substitution by the Auditory Sense"*, en el libro Computational Science and Its Applications - ICCSA 2011, Lecture Notes in Computer Science, Vol. 6783, pp. 522-533, Springer Berlin/Heidelberg, 2011.
- [Capelle, 1998] Capelle C., Trullemans C., Arno P., Veraart C., *"A real experimental prototype for enhancement of vision rehabilitation using auditory substitution"*, IEEE Transactions on Biomedical Engineering, Vol. 45, No. 10, pp. 1279-1293, October 1998.
- [Cha, 1990] Cha K., Horch, K.W., Normann, R.A., *"Simulation of a phosphene field based visual prosthesis"*, Proceedings of the 1990 IEEE Conference on Systems, Man and Cybernetics, pp. 921-923, 1990.
- [Cimarolli, 2006] Cimarolli V., Boerner K., Wang S., *"Life goals in vision rehabilitation: Are they addressed and how?"*, Journal of Visual Impairment & Blindness, pp. 343-352, June 2006.
- [Dondis, 2000] Dondis D., *"La sintaxis de la imagen, Introducción al alfabeto visual"*, Ediciones G. Gill, S. A. de C. V., México, 2000.
- [Edwards, 2006] Edwards A. D. N., McCartney H., Fogarolo F., *"Lambda: A multimodal approach to making mathematics accessible to blind students"*, 6 Proceedings of the 8th international, ACM SIGACCESS Conference on Computers and accessibility ACM New York, NY, USA, pp. 48-54, 2006.
- [Fletcher, 1933] Fletcher & Munson, bajado de Wikipedia el 3 de septiembre 2011: [http://en.wikipedia.org/wiki/Equal-loudness\\_contour](http://en.wikipedia.org/wiki/Equal-loudness_contour)
- [Forster, 2007] Forster B., Eardley A. F., Eimer M., *"Alter tactile spatial attention in the early blind"*, Brain Research, Vol. 1131, 2, pp. 149-154, February 2007.
- [Frauenberger, 2007] Frauenberger C., Stockman T., Bourguet M. L., *"Pattern Design in the Context Space. A Methodological Framework for Auditory Display Design"* Proceedings of the 13<sup>th</sup> International Conference on Auditory Display, Montreal, Canada, June 26-29, pag. 513-518, 2007.
- [Giacomin, 2002] Giacomin J., *"Human Factors: Sound Exposure"*, (2002), Recuperado el 1 de octubre de 2011 del sitio Web del Human Centred Design Institute, en Londres, Inglaterra: [http://www.perceptionenhancement.com/docs/human\\_factors\\_course/human\\_factors\\_course\\_lec14\\_sound\\_exposure.pdf](http://www.perceptionenhancement.com/docs/human_factors_course/human_factors_course_lec14_sound_exposure.pdf)
- [Goertz, 2010] Goertz H. Y. H., Lierop G. B., Houkes I., Nijhuis N. F. J., *"Factors related to the employment of visually impaired persons: a systematic literature review"*, Journal of Visual Impairment & Blindness, pp. 404-418, 2010.
- [González, 1996] González R., Woods R., *"Tratamiento Digital de Imágenes"*, USA : Addison-Wesley Iberoamericana y Díaz de Santos, 1996.
- [Gosset, 2010] Gosset W., Beaven E., *"Latin squares in practice and in theory I: Agronomy experiments.- Latin squares in real life"*, American Mathematical Society, 2010  
Bajado de: <http://resources.metapress.com/pdf->

- preview.axd?code=w58771628111t168&size=largerst
- [Greenemeier, 2008] Greenemeier L., *"Scientists set sights on an implantable prosthetic for the blind"*, Scientific American, March 19, 2008.
- [Greenemeier, 2009] Greenemeier L., *"Bionic eye restores visión after three decades of darkness"*, Scientific American, March 4, 2009.
- [Hankinson, 1999] Hankinson J., Edwards A., *"Designing Earcons with Musical Grammars"*, ACM SIGCAPH Computers and the Physically Handicapped, Issue 65, September, pp.16-20, 1999.
- [Hardwick, 1998] Hardwick A., Furner S., Rush J., *"Tactile display of virtual reality from the WorldWideWeb-a potential access method for blind people"*, IEE Colloquium on Developments in Tactile Displays", London , ROYAUME-UNI (21/01/1997) , Vol. 18, Issue 3, pp. 153-161, 15 May 1998.
- [Harris, 1993] Harris D., Whitney G., *"Visually impaired people and public transport information"*, IEEE Colloquium on Public Transport Information and Management Systems, pp. 5/1-5/7, London, UK, 1993.
- [Hoffman, 1985] Hoffman D.D., Richards W., *"Parts of recognition"*, Cognition, 18, pp.65-96, 1985.
- [Horgan, 2005] Horgan J., *"The forgotten era of brain; the work of Jose Delgado, a pioneering star"*, Scientific American, pp. 66-73, October 2005.
- [Hoydal, 1991] Hoydal T. O., Zelano J. A., *"An alternative mobility aid for the blind: the ultrasonic cane"*, Proceedings of the 1991 IEEE Seventeenth Annual Northeast Bioengineering Conference 4-5, pp. 158-159, Apr. 1991.
- [Jihong, 2006] Jihong L., Xiaoye S., *"A Survey of Vision Aids for the Blind."*, Proceedings of the 6<sup>th</sup> World Congress on Intelligent Control and Automation, June 21-23, Dalian, China, pp. 4312-4316, 2006.
- [Kaczmarek, 1991] Kaczmarek K., Webster J. G., Bach-y-Rita P., Tompkins W. J., *"Electrotactile and vibratory displays for sensory substitution systems"*, IEEE Transactions on Biomedical Engineering, Vol. 38, No. 1, Pages 1-16, January 1991.
- [Langtao, 1993] Langtao L., Balachandran W., *"Electronic compass for blind or deaf-blind pedestrians"*, Electronics Letters, 5<sup>th</sup> August 1993, Vol. 29, No. 16, pp. 1462-1463, 1993.
- [Little, 2002] Thomas M. Little y F. Jackson Hills., *"Métodos estadísticos para la investigación en la agricultura"*, Editorial Trillas, 2002.
- [Liu, 2010] Liu J, Xu L., Jin W., *"Electronic travel aids for the blind base don sensory substitution"*, The 5<sup>th</sup> International Conference on Computer Science & Education, Hefei, china, August 24-27, 2010.
- [Long, 1992] Long R., Fazenbaker S., *"Development of an orientation aid for blind individuals"*, Proceedings of the Johns Hopkins National Search for Computing Applications to Assist Persons with Disabilities, pp. 205-207, 1992.
- [Lorenzo, 2005] Lorenzo M., *"Accessibility for blind and visually impaired people"*, Vision 2005: Proceedings of the International Congress held between 4 and 7 April 2005 in London, UK, International Congress Series vol. 1282, pp. 1038-1040, September 2005.

- [Martínez, 1984] Martínez R., Brena R., *"Ver por el oído"*, Contactos, Revista de Educación en Ciencias Básicas e Ingeniería, Vol. 1, No. 2, Universidad Autónoma Metropolitana, pp. 66-71, Junio 1984.
- [Martínez, 2005] Martínez R., Ramírez M., *"Obtención del icono y del contorno por medio de métodos de topología digital"*, Temas de Ciencia y Tecnología Revista de la Universidad Tecnológica de la Mixteca, Vol. 9 No. 25, Enero-Abril pp. 35-41, 2005.
- [Martínez, 2007] Martínez R., Martínez A. C., *"Desarrollo de tecnología educativa para invidentes"*, 2do. Foro de Investigación Educativa "Fomento a la Investigación Educativa" organizado por IPN y CFIE, Ciudad de México, D.F., del 18 y 19 de enero, 2007.
- [Martínez, 2010] Martínez R., Menchaca G. R. F., Berrón L. V., *"Acoustics pattern code to represent the contour of any visual image for people who are blind"*, J. Acoust. Soc. Am., Vol. 128, Issue 2281, 2010.
- [Martínez, 2011] Martínez R., Menchaca G. R. F., Berrón L. V., *"Codificación del contorno de la imagen visual a patrón acústico y su decodificación realizada por invidentes"*, Revista Técnica de la Facultad de Ingeniería de la Universidad de Zulia, Vol. 34, No. 2, Agosto 2011.
- [Matta, 2004] Matta S., Kumar D. K., Yu X., Burry M., *"An Approach For Image Sonification"*, First International Symposium on Control, Communications and Signal Processing, pp. 431-434, 2004.
- [McCreath, 2005] McCreath G., Cuthbertson J., *"Using digital technology to improve access to learning"*, Vision 2005: Proceedings of the International Congress held between 4 and 7 April 2005 in London, UK, International Congress Series vol. 1282, pp. 956-959, September 2005.
- [Meijer, 1992] Meijer P., *"An Experimental system for Auditory Image Representation"*, IEEE Transactions on Biomedical Engineering, Vol. 39, No. 2, pp. 112-121, February 1992.
- [Meijer, 2001] Meijer P., *"Seeing with sound for the blind: Is it vision?"*, VSPA Conference on Consciousness, at the University of Amsterdam, The Netherlands, June 1, 2001. (Abstract y diapositivas de la conferencia descargados desde <http://www.seeingwithsound.com/> el 13/8/08.)
- [Meijer, 2007] Meijer P., *"Vision Technology for the totally blind"*, Seeing with Sound. The vOICe. 2007. Descargado el 14/8/08 desde <http://www.seeingwithsound.com/>.
- [Montgomery, 2007] Montgomery D., *"Diseño y análisis de experimentos"*, Segunda Edición, México Limusa Wiley, 2007.
- [Morgan, 1999] Morgan M., *"Sensory perception: Supernormal hearing in the blind"*, Current Biology, Volume 9, Issue 2, pp. 53-54, 28 January 1999.
- [Motluck, 2005] Motluck A., *"Seeing with your ears"*, The New York Times,, Sunday's Late Edition in print, section 6, page 90, column 3, 11. December 2005.
- [Moustakas, 2007] Moustakas K., Nikolakis, G. Kostopoulos, K. Tzovaras, D. Strintzis, M.G., *"Haptic Rendering of Visual Data for the Visually Impaired"*, IEEE Multimedia, Vol. 14, Issue 1, pp.62-72, 2004.
- [Murphy, 1995] Murphy R., *"Sensor fusion"*, The handbook of brain theory and neural networks, Michel A. Arbid, editor, Bradford Books/The MIT Press, pp. 857-859, 1995.

- [Nagarajan, 2003] Nagarajan R., Yaacob S., Sainarayanan, G., *“Role of object identification in sonification system for visually impaired”*, Conference on Convergent Technologies for Asia-Pacific Region TENCON 2003, Vol. 2, pp.735- 739 Vol.2 , Oct. 2003.
- [Nagarajan, 2004] Nagarajan, R. Sainarayanan, G. Yacoob, S. Porle, R.R., *“An improved object identification for NAVI”*, 2004 IEEE Region 10 Conference TENCON 2004, Vol. A, pp. 455- 458 Vol. 1, Nov. 2004.
- [Nave, 2005] Nave R., *“The piano”*, 2005, bajado de internet el 25 de septiembre de 2011: <http://hyperphysics.phy-astr.gsu.edu/hbase/music/pianof.html#c1>
- [Noordzij, 2006] Noordzij M., Zuidhock S., Postma A., *“The influence of visual experience on the ability to form spatial mental models based on route and survey descriptions”*, Cognition, Vol. 100, Issue 2, pp. 321-342, June 2006.
- [Oberdofer, 2004] Oberdofer M., *“Research opportunities in low vision and blindness rehabilitation”*, Journal of Visual Impairment & Blindness, pp. 583-584, October 2004.
- [OMS, 2009] Organización Mundial de la Salud, Centro de Prensa: *“Ceguera y Discapacidad Visual”*. Nota Descriptiva No. 282, Mayo 2009.
- [Pajares, 2004] Pajares G., De la Cruz M. J., *“Imágenes digitales: Procesamiento práctico con Java”*, Alfaomega- Ra-Ma, 2004.
- [Peretz, 2003] Peretz I., Zatorre R., *“The Cognitive Neurosciende of Music”*, Oxford University Press, 2003.
- [Pérez, 1991] Pérez P., *“Elaboración y uso de material didáctico”*, Módulo I, EA, México: Instituto de Capacitación de la Industria de la Construcción, 1991.
- [Pezaris, 2007] Pezaris J., and Clay R., *“Demonstration of artificial visual percepts generated through thalamic microstimulation”*. 7670-7675, PNAS, May 1, vol. 104, no. 18, 2007.
- [Pezaris, 2009] Pezaris J., *“Getting signals into the brain: visual prosthetics through thalamic microstimulation”*, National Institutes of Health, Neurosurg Focus, Vol. 27, No. 1,Jul, 2009:E6.
- [Power, 2010] Power C., Jürgensen H., *“Accessible presentation of information for people with visual disabilities”*, Universal Access in the Information Society , Volume 9, Number 2, pp. 97-119, 2010
- [Pronin, 2007] Pronin E., *“Perception and misperception of bias in human judgment”*, Trends in Cognitive Sciences, Volume 11, Issue 1, Pages 38-43, January 2007.
- [Ramloll, 2000] Ramloll R., Wai Y.,Brewster S., Riedel B., Burton M., Dimigen G.,*“Constructing sonified haptic line graphics for the blind student: :first steps”*, “ASSETS ´00 Proceedings of the fourth international ACM conference on Assistive technologies”, pp. 17-25, 2000.
- [Ramírez, 2002] Ramírez M., *“Procesamiento paralelo de imágenes de coloides y Ronchigramas”*, Tesis, UTM, 2002.
- [Rauschecker, 2003] Rauschecker J. P., *“Musical Predispositions in Infancy: An Update”* , capítulo del libro *“The Cognitive Science of Music”*, Peretz I., Zatorre R.J. (eds.), Oxford University Press, 2003.

- [Renier, 2005] Renier L., De Volder A., *"Cognitive and brain mechanisms in sensory substitution of vision: A contribution to the study of human perception"*, Journal of Integrative Neuroscience, Vol. 4, No.4, pp. 409-503, 2005.
- [Rodríguez, 1997] Rodríguez M., *"El Concepto de índice de un punto lattice en el análisis homotópico de 2D y 3D imágenes y su realización algorítmica"*, Tesis doctoral, UAP, 1997.
- [Rosenfeld, 1989] Rosenfeld A., Kong T.Y., *"Digital Topology: Introduction and Survey"*, Computer Vision, Graphics and Image Processing, Vol. 48, pp. 357-393, 1989.
- [Rovira, 2010] Rovira k., Gapenne O., Ali A. A., *"Learning to recognize shapes with a sensory substitution system: a longitudinal study with 4 non-sighted adolescents"*, IEEE 9<sup>TH</sup> International Conference on Development and Learning, pp. 1-5, 2010.
- [Sainarayanan, 2007] G. Sainarayanan, R. Nagarajan, S. Yacoob, *"Fuzzy image processing scheme for autonomous navigation of human blind"*, Applied Soft Computing, Vol. 7, Issue 1, pp. 257-264, January 2007.
- [Sampaio, 2001] Sampaio E., Maris S., and Bach-y-Rita P., *"Brain plasticity: 'visual' acuity of blind persons via the tongue"*, Brain Research 908, pp 204-207, 2001.
- [Sánchez, 2010] Sánchez J., *"Autonomous navigation through the city for the blind"*, ASSETS '10 Proceedings of the 12th international, ACM SIGACCESS Conference on Computers and accessibility ACM New York, NY, USA, pp. 195-202, 2010.
- [Shoval, 1998] Shoval S., Borenstein J., Koren, Y., *"Auditory Guidance with the Navbelt- A Computerized Travel Aid for the Blind"*, IEEE Transactions on Systems, Man, and Cybernetics- Part C: Applications and Reviews, Vol. 28, No. 3, pp. 459-467, August 1998.
- [Snyder, 2005] Snyder J., *"Audio description: The visual made verbal"*, Vision 2005: Proceedings of International Congress held between 4 and 7 April 2005 in London, UK, International Congress Series vol. 1282, pp. 935-939, September 2005.
- [Srivilailuck, 2005] Srivilailuck S., Beale P., Murray I., Kidd B., *"Hidden design: An inquiry into the design of inclusive building environments and digital interface design for the vision impaired"*, Vision 2005: Proceedings of the International Congress held between 4 and 7 April 2005 in London, UK, International Congress Series vol. 1282, pp. 1071-1074, September 2005.
- [Storer, 2006] Storer A., *"Visual perception through sensory substitution"*, Department of Cognitive and Neural Systems. Term paper CN730, Boston University, 2006. (Descargado el 19.08.08 desde: <http://cns.bu.edu/~storer/sensorysubstitution/Alex%20Storer%20-%20Final%20Paper%20-%20Updated.pdf>).
- [Strumillo, 2010] Strumillo P., *"Electronic interfaces aiding the visually impaired in environmental access, mobility and navigation"*, Rzeszow, IEEE, pp.17-24, May, 2010.
- [Studebaker, 2004] Studebaker J., Pankow L., *"History and evolution of vision rehabilitation- Parallels with Rehabilitation Medicine, Geriatric Medicine and Psychiatry"*, Topics in Geriatric Rehabilitation, Vol. 20, No. 3, pp. 142-153, 2004.

- [Trehub, 2003] Trehub S., *"Musical Predispositions in Infancy: An Update"*, capítulo del libro *"The Cognitive Science of Music"*, Peretz I., Zatorre R.J. (eds.), Oxford University Press, 2003.
- [Tremaine, 1969] Tremaine H., *"Audio Cyclopedia"*, Sams, 1969.
- [Tyler, 2003] M. Tyler, Y. Danilov, P. Bach-y-Rita, *"Closing an open-loop control system: Vestibular substitution through the tongue"*. *Journal of Integrative Neuroscience*, Word Scientific, Vol. 2, pp. 159-164, 2003.
- [Van Scoy, 2003] Van Scoy F., McLaughlin, Fullmer A., *"Auditory Augmentation of Haptic Graphs: Developing a Graphic Tool for Teaching Precalculus Skill to Blind Students"*, Proceedings of ICAD 05-Eleventh Meeting of the International Conference on Auditory Display, Limerick, Ireland, pp. 1-4, July 6-9, 2005.
- [Wang, 1996] Wang Z., Ben-Arie J., *"Conveying visual information with spatial auditory patterns"*, *IEEE Transactions on Speech and Audio Processing*, Vol. 4, No. 6, pp. 446-455, November 1996.
- [Watzlawick, 1994] Watzlawick P., and Humayun, M., *"El ojo del observador; Contribuciones al constructivismo, Homenaje a Heinz von Foerster"*, Gedisa editorial, pp. 9-31, 1994.
- [Ward, 2010] Ward, J., Meijer, P., *"Visual experiences in the blind induced by an auditory sensory substitution device"*, *Consciousness and Cognition*, Vol. 19, No. 1, pp. 492-500, 2010.
- [Weiland J., 2008] Weiland J. and Humayun M., *"Visual prosthesis"*, Proceedings of the IEEE, Vol. 96, No. 7, pp 1076-1084, July 2008.
- [Zelek, 2005] Zelek J., *"Seeing by touch (haptics) for way finding"*, *Vision 2005: Proceedings of the International Congress held between 4 and 7 April 2005 in London, UK, International Congress Series vol. 1282*, pp. 1108-1112, 2005.
- [Zwicker, 1957] Zwicker E., Flottorp G., Stevens S. S., *"Critical Band Width in Loudness Summation"*, *The Journal of the Acoustical Society of America*, Vol. 29, No. 5, pp. 548-557, 1957.
- [Zwicker, 1975] Zwicker E., *"Scaling"*, *Handbook of Sensory Physiology*, vol. V/2 Auditory System, Physiology (CNS). Behavioral Studies, Psychoacoustics, W.D. Keidel and W.D. Neff, Eds.; Berlin, Germany: Springer-Verlag, 1975.



# Anexo I: Lista de publicaciones y presentaciones en congresos

## Congresos Nacionales

1. Martínez J. R., Brena R., "Codificación auditiva de imágenes", *V Congreso Nacional de Ingeniería Biomédica*, Sociedad Mexicana de Ingeniería Biomédica, del 8 al 13 de noviembre de 1982, Puebla, México.
2. Martínez J. R., Martínez C., "Desarrollo de tecnología educativa para invidentes", *2º. Foro de Investigación Educativa-Fomento a la Investigación Educativa*, en el Centro de Formación e Innovación educativa, IPN., 18 y 19 de enero, 2007.
3. Martínez J. R., Martínez C., Berrón V., "Prótesis visual para invidentes con base en la lectura de patrones acústicos", 8º. Simposio Internacional de Estudios Oaxaqueños, Oaxaca, Oax., del 25-27 de junio, 2009. Instituto Welte para Estudios Oaxaqueños.
4. Martínez J. R., Berrón V., Menchaca F. R., "Codificación de patrones acústicos efectuada por invidentes de la Biblioteca para Ciegos Jorge Luis Borges de la Ciudad de Oaxaca", 9º. Simposio Internacional de Estudios Oaxaqueños, Oaxaca, Oax., del 7 al 9 de Julio, 2011. Instituto Welte para Estudios Oaxaqueños.

## Congresos Internacionales

1. Martínez J. R., Beristain S., "La percepción sensorial auditiva vista desde el punto de vista de la reversibilidad del mapeado de la imagen del mundo", First Pan-american/iberian Meeting on Acoustics, 144th Meeting of the Acoustical Society of America, 3rd Iberoamerican Congress of Acoustics, 9th Mexican Congress on Acoustics , 2-6 DEC. 2002, Cancún, Q. R. México.
2. Martínez J. R., Martínez C., "El concepto de earcon como una posibilidad para el desarrollo de código de texto acústico", *10º. Congreso Internacional Mexicano de Acústica*. 26-28 Nov. 2003, Puebla, Pue.
3. Martínez J. R., Martínez D. S., "Codificación del borde de imágenes geométricas básicas utilizando CSound y las 88 teclas del piano", *11º. Congreso Internacional Mexicano de Acústica*. 6-8 Oct., 2004, Morelia, Michoacán.
4. Martínez J. R., Menchaca F. R., "El concepto de sonificación, el contorno de los objetos, la aculogía y su lectura como texto acústico", *11º. Congreso Internacional Mexicano de Acústica*. 6-8 Oct., 2004, Morelia, Michoacán.
5. Martínez J. R., Menchaca F. R., "Planteamiento de los patrones acústicos utilizados como un primer código de sonificación para invidentes y su análisis de varianza utilizando cuadrados latinos", *12º. Congreso Internacional Mexicano de Acústica*. 26-28 Oct., 2005, Sta. Cruz, Tlaxcala.

6. Martínez J. R., Menchaca F. R., Berrón V., "Resultado del análisis experimental perceptivo de un grupo de patrones acústicos aplicados a un grupo de videntes con educación musical", 13º. *Congreso Internacional Mexicano de Acústica*. 11-14 Oct., 2006, Oaxaca, Oax.
7. Martínez C., Martínez J. R., "Dibujo de la silueta de los objetos como una manifestación de la construcción conceptual a partir de la lectura del texto acústico", 14º. *Congreso Internacional Mexicano de Acústica*. 14-16 Noviembre de 2007, León, Guanajuato, México.
8. Martínez J. R., Berrón V., Menchaca R., "Subjective evaluation in blind humans of the perception of acoustic patterns generated by image contours", 10<sup>th</sup> *School on Acousto-Optics and Applications*, Gdansk-Sopot, Polonia, del 12 al 15 de mayo, pp. 33, 2008.
9. Martínez J. R., Martínez C., Berrón V., "Ensayo sobre subjetividad sensorial humana", 15o. *Congreso Internacional Mexicano de Acústica*, Taxco, Guerrero, México, 2008.
10. Martínez J. R., Martínez C., Berrón V., Menchaca R., "La generación de conceptos de forma de las cosas a partir de la lectura de patrones acústicos por invidentes", 16º. *Congreso Internacional Mexicano de Acústica*, Uruapan, Mich., México, 2009.
11. Martínez J. R., Berrón V., Menchaca R., "Un código de patrones acústicos para representar el contorno de cualquier imagen visual para invidentes", 2nd Pan-American and Iberian Meeting on Acoustics, *Cancún, Q. R. México*, 2010.
12. Martínez J. R., Berrón V., Menchaca R., "El concepto de invariabilidad en los patrones acústicos representativos de imágenes geométricas", 18º. *Congreso Internacional Mexicano de Acústica*, Cholula, Puebla, México, 2011.

### **Publicación en Revistas Nacionales**

5. Martínez J. R., Brena R., "Ver por el oído", *Contactos - Revista de Educación en Ciencias Básicas e Ingeniería*, Vol. I, Num. 2, Abril - Junio 1984, Universidad Autónoma Metropolitana, pp. 66-71.
6. Martínez J. R., Ramírez M., "Obtención del icono y del contorno de la imagen por medio de métodos de topología digital", *Temas de Ciencia y Tecnología*, Revista de la Universidad Tecnológica de la Mixteca. Vol. 9, Número 25, Enero-Abril 2005 pp. 35-41.

### **Publicación en Revistas Internacionales**

1. Martínez J. R., Beristain S., "Artificial environment mapping from acoustic information (A)", *The Journal of The Acoustical Society of America*, Vol. 112, No.5, Nov. 2002, pp. 2245. 2, Revista Internacional Arbitrada por Science Citation Index (SCIExpanded).
2. Martínez J. R., Menchaca R., Berrón V., "A code of acoustic patterns to represent the contour of any visual image for the blind", *J. Acoust. Soc. Am.*, Vol. 128, No. 4,

- Pt. 2, October 2010, pp. 2281. 2, Revista Internacional Arbitrada por Science Citation Index (SCIExpanded).
3. Martinez J. R., Menchaca R., Berrón V., *“Codificación del contorno de la imagen visual a patrón acústico y su decodificación realizada por invidentes”*, Revista Técnica de Ingeniería de la Universidad de Zulia, Vol. 34, pp. 129-138, Agosto 2011, No. 2, Revista Internacional Arbitrada por Science Citation Index (SCIExpanded).

A continuación anexamos las primeras páginas o los resúmenes publicados de algunos de los trabajos mencionados en la lista anterior:

Martínez J. R., Beristain S., "La percepción sensorial auditiva vista desde el punto de vista de la reversibilidad del mapeado de la imagen del mundo", First Pan-american/iberian Meeting on Acoustics, 144th Meeting of the Acoustical Society of America, 3rd Iberoamerican Congress of Acoustics, 9th Mexican Congress on Acoustics , 2-6 DEC. 2002, Cancún, Q. R. México. (Resumen publicado en *The Journal of The Acoustical Society of America*, Vol. 112, No.5, Nov. 2002, pp. 2245 con el título " Artificial environment mapping from acoustic information (A)" ).

101  
2002

## The Journal of the Acoustical Society of America

Vol. 112, No. 5, Pt. 2 of 2, November 2002

<http://asa.aip.org>



### First Pan-American/Iberian Meeting on Acoustics

144th Meeting of the Acoustical Society of America  
3rd Iberoamerican Congress of Acoustics  
9th Mexican Congress on Acoustics

Fiesta Americana Grand Coral Beach Hotel  
Cancun, Mexico  
2-6 December 2002

Table of Contents on p. A5

Published by the Acoustical Society of America through the American Institute of Physics

**PRIMERA REUNION PAN-AMERICANA/IBERICA DE ACUSTICA****2-6 Dec. 2002****Cancún, Q.R. México****La percepción sensorial auditiva vista desde el punto de vista de la reversibilidad del mapeado de la imagen del mundo****R. Martínez (CIIDIR IPN, Oaxaca) y S. Beristain (ESIME ICE)****Resumen**

Este trabajo está fundamentado en la capacidad que poseen los seres vivos para construir un mapeado del mundo en función de sus sistemas de percepción. Mapeado que le permite sobrevivir dentro de una naturaleza compleja y que aparentemente se ve como un simple principio. Sin embargo, cuando tratamos de llevar este concepto a la vida artificial o a la inteligencia artificial, este principio se complica a tal grado que aun se están buscando soluciones a esta problemática. El sistema auditivo normal y la percepción ultrasónica crean mapeados propios del mundo, según sea la especie a la que hagamos referencia.

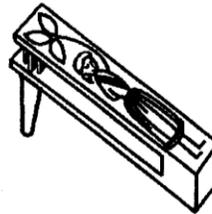
Martínez J. R., Martínez C., "El concepto de earcon como una posibilidad para el desarrollo de código de texto acústico", *10º. Congreso Internacional Mexicano de Acústica*. 26-28 Nov. 2003, Puebla, Pue

17:00 EL CONCEPTO DE EARCON COMO UNA POSIBILIDAD PARA EL DESARROLLO DE CODIGO DE TEXTO ACUSTICO (The EARCON concept as a development possibility for an acoustic text code)  
 RODOLFO MARTINEZ Y CARDENAS, CARLOS MARTINEZ  
 ALVAREZ, rodolfo\_myc@hotmail.com

RESUMEN. El presente trabajo está enfocado a un tema de investigación conocido como earcons, el cual es el acrónimo de ear (oído) y cons, en otro contexto es similar al eye-con cuyo acrónimo es el bien conocido icon. El earcon se usa para transmitir mensajes, ejemplo de ello son los tambores en África, el telégrafo, los cuales son de uso común desde hace años. Esta característica del sonido permite pensar en la posibilidad de desarrollar un código de texto acústico que aplicado a la computadora permita al ciego decodificar las formas de los objetos por medio del uso del earcon.

Martínez J. R., Menchaca F. R., "El concepto de sonificación, el contorno de los objetos, la aculogía y su lectura como texto acústico", *11º. Congreso Internacional Mexicano de Acústica*. 6-8 Oct., 2004, Morelia, Michoacán.

**11º CONGRESO INTERNACIONAL MEXICANO DE ACUSTICA**  
**11th Mexican International Congress on Acoustics**



Morelia, Mich. México

6 – 8 Oct. 2004

**A3-El concepto de sonificación, el contorno de los objetos, la aculogía y su lectura como texto acústico**

José Rodolfo Martínez y Cárdenas\*, Felipe Rolando Menchaca García\*,  
[rodolfo\\_mvc@hotmail.com](mailto:rodolfo_mvc@hotmail.com).

**Resumen**

Se exponen aquí algunas de las bases teórico acústicas con las cuales se sustenta el aspecto cognitivo del proyecto llamado "Visor Acústico". Entre estos conceptos se encuentra la sonificación, el contorno de los objetos como elemento básico para la creación de los patrones acústicos, la aculogía<sup>1</sup> palabra creada por Schaeffer y la lectura de textos acústicos, como un medio de comunicación entre el mundo de los objetos y el invidente. Este tipo de comunicación, un poco atrevida, es tratada por la semiología.

\* CIIDIR IPN Unidad Oaxaca, Calle Hornos 1003, 71230 Santa Cruz Xoxocotlán, Oax., Becario COFAA, IPN y Decano del CIIDIR Oaxaca, Proyecto CGPI: 20040542.

\* Director de CENAC del IPN.

<sup>1</sup> Michael Choin, "El Sonido: música, cine y literatura". Teoría 107, Piados, Comunicación, 1999.

Martínez J. R., Martínez D. S., "Codificación del borde de imágenes geométricas básicas utilizando Csound y las 88 teclas del piano", *11º. Congreso Internacional Mexicano de Acústica*. 6-8 Oct., 2004, Morelia, Michoacán.

## **A4-Codificación del borde de imágenes geométricas básicas utilizando Csound y las 88 teclas del piano**

José Rodolfo Martínez y Cárdenas<sup>♦</sup>, Dalia Silva Martínez<sup>♦</sup>,

[rodolfo\\_myc@hotmail.com](mailto:rodolfo_myc@hotmail.com),

### **Resumen**

Dentro de la búsqueda de nuevos tipos de prótesis para invidentes esta el llamado Visor Acústico, el cual estamos desarrollando. Este sistema en una de sus secciones del proceso necesita de la codificación de los bordes de las imágenes a un patrón acústico. Para lograr esto se está utilizando el lenguaje de programación musical Csound, el cual nos ha ofrecido la posibilidad de generar  $n$  frecuencias dentro del mismo periodo de tiempo. Siendo esto un procesamiento en paralelo y en tiempo real. Las imágenes se adecuaron a 88 filas, las cuales corresponderán a las 88 teclas del piano y se dejó variable la cantidad de columnas, las cuales representarán al tiempo. Una imagen está representada por una matriz de pixels, el problema a resolver fue el de transportar los datos de la imagen representada por filas y columnas a una matriz de frecuencias y tiempo. Otro problema que se enfrentó fue la duración de escaneo total del borde de la imagen. Se utilizaron todas las octavas del piano, por ser uno de los instrumentos más rico en generación de frecuencias. Con el uso de las notas musicales se pretende darle más sustento acústico a los patrones acústicos por las características propias de las escalas musicales.

**Palabras clave:** Prótesis acústica para invidentes, Csound, escalas del piano, visor acústico.

Martínez J. R., Ramírez M., "Obtención del icono y del contorno de la imagen por medio de métodos de topología digital", *Temas de Ciencia y Tecnología*, Revista de la Universidad Tecnológica de la Mixteca. Vol. 9, Número 25, Enero-Abril 2005 pp. 35-41.

## Notas

# Obtención del Icono y del Contorno de la Imagen por Medio de **Métodos de Topología Digital**

## I. Introducción

Sabemos que nuestros ojos están equipados para ver el color de los objetos; pero es menos sabido que también lo están para percibir los límites espaciales entre los objetos, sus **contornos**. La noción de *borde visual* designa la frontera entre dos superficies de diferente luminancia, cualquiera que sea la causa de esta diferencia de luminancia (diferentes iluminaciones, diferentes propiedades reflectantes, etc.) para un punto de vista determinado (por ejemplo, entre dos superficies, una de las cuales está tras la otra, hay un borde visual, pero si cambia el punto de vista, el borde ya no estará en el mismo lugar).

El sistema visual está equipado por *construcción* con instrumentos capaces de reconocer un borde visual y su orientación, una ranura, una línea, un ángulo, un segmento: estos preceptos son como las unidades elementales de nuestra percepción de los objetos y del espacio. El mecanismo extremadamente complejo de esta percepción ha dado buenos resultados, puesto que somos capaces de distinguir bordes visuales de muy pequeña dimensión.

Las propiedades físicas del mundo no dependen de nuestra mirada sobre él. El mundo, a grosso modo, tiene siempre la misma apariencia o, al menos, esperamos encontrar en él, cada día, cierto número de elementos invariantes. La percepción de esos aspectos invariantes del mundo (tamaño de los objetos, formas, orientaciones, propiedades de las superficies, etc.) es lo que se designa mediante la noción de **constancia perceptiva**: a pesar de la variedad de las percepciones, siempre se encontrarán algunos aspectos constantes.

Estas constantes perceptivas invariantes nos facilitan la vida, permitiéndonos crear esquemas mentales del

modelo del mundo. En el presente trabajo se hace referencia a la importancia que tiene el icono y el contorno de la imagen en diferentes áreas. Se presenta la metodología para obtener tanto el icono representativo de una imagen real, así como el contorno de la misma. Para lograr esto, primero la imagen debe de ser preparada, utilizando para ello los filtros espaciales correspondientes (procesamiento de la imagen por medio de máscaras) que nos permitan obtener la imagen binaria. A partir de la imagen binaria se le aplican los métodos de topología digital, logrando con ello la obtención del icono o del contorno deseado.

## II. Sensación y percepción de la imagen

Existe una gran diferencia entre la sensación y la percepción [Matlin y Foley<sup>1</sup>], la primera se refiere a experiencias inmediatas y básicas, generadas por estímulos aislados simples, en cambio, la segunda incluye la interpretación de esas sensaciones, dándoles significado y organización. Por ejemplo, cuando leemos un libro los procesos sensoriales y preceptuales demuestran sus extraordinarias habilidades. Los ojos recorren cada página a paso firme, identificando letras y palabras tan rápido que desafían las explicaciones. Así mismo, cada mañana al despertar y levantarnos ejecutamos varias funciones de sensación y de percepción como es: ver, escuchar, tocar, oler y degustar.

Cuando en la radio escuchamos algunas notas musicales, sus características de volumen y tono son sensaciones. En caso de reconocer las primeras notas que forman parte de una tonada, experimentaremos una percepción.

Matlin y Foley también comentan que cuando tenemos frente a nuestros ojos una imagen, cada

Martínez J. R., Virginia B. L., Menchaca R., "Planteamiento de los patrones acústicos utilizados como un primer código de sonificación para invidentes y su análisis de varianza utilizando cuadrados latinos", *12º. Congreso Internacional Mexicano de Acústica*. 26-28 Oct., 2005, Sta. Cruz, Tlaxcala. ISBN 970-27-0763-3

### TG3 - Planteamiento de los patrones acústicos utilizados como un primer código de sonificación para invidentes y su análisis de varianza utilizando cuadrados latinos

J. Rodolfo Martínez y C.<sup>4</sup>, rodolfo\_mc@yahoo.com;  
Rolando Menchaca García<sup>5</sup>, menchaca@pollux.cic.ipn.mx;

Virginia Berrón Lara<sup>6</sup>, berron@nuyoo.utn.mx

#### **Resumen**

En este trabajo se presentan los patrones acústicos utilizados como un primer código de sonificación con la finalidad de capacitar a un grupo de invidentes, y la metodología de evaluación la cual consiste tanto en el uso de Cuadrados Latinos así como el uso de ítems binarios para la construcción del instrumento psicométrico. Los Patrones Acústicos han sido desarrollados en CSound, para llevar a cabo esto se desarrolló un programa de computación el cual permite codificar la imagen BMP resultante del procedimiento planteado para extraer la silueta representativa de cualquier imagen en un archivo con extensión .wav.

Martínez J. R., Menchaca F. R., Berrón V., "Resultado del análisis experimental perceptivo de un grupo de patrones acústicos aplicados a un grupo de videntes con educación musical", 13<sup>º</sup>. *Congreso Internacional Mexicano de Acústica*. 11-14 Oct., 2006, Oaxaca, Oax. ISBN 970-27-1047-2-

## **Resultado del análisis experimental perceptivo de un grupo de patrones acústicos aplicados a un grupo de videntes con educación musical.**

José Rodolfo Martínez y Cárdenas<sup>1</sup>, [rodolfo\\_mc@yahoo.com](mailto:rodolfo_mc@yahoo.com);

Rodrigo Menchaca García<sup>2</sup>, [menschaca@polimx.cic.ipn.mx](mailto:menchaca@polimx.cic.ipn.mx);

Virginia Berrón Lara<sup>3</sup>, [berron@yahoo.utm.mx](mailto:berron@yahoo.utm.mx)

### ***Resumen***

En este trabajo se presentan los primeros resultados del análisis experimental perceptivo de un grupo de patrones acústicos aplicados a varias personas con visión normal, siendo esta serie de patrones acústicos un primer código de sonificación, serie que se utilizará para capacitar en un futuro a un grupo de invidentes. El grupo humano que se sometió a la fase de aprendizaje y evaluación tuvo por característica común el haber tenido durante su vida una educación musical, este parámetro permite inferir que ya poseen desarrollado sus centros frenológicos orientados hacia la música y que algo similar sucede con los invidentes en su percepción de códigos de sonificación.

Este grupo en estudios posteriores se someterá a ser comparado con el grupo de invidentes en lo que respecta a su capacidad de aprendizaje, ya que ambos poseen un desarrollo especial en su de análisis e interpretación de patrones acústicos complejos.

Martínez J. R., Martínez C., “Desarrollo de tecnología educativa para invidentes”, *2º. Foro de Investigación Educativa-Fomento a la Investigación Educativa*, en el Centro de Formación e Innovación educativa, IPN., 18 y 19 de enero, 2007.

## 2do. FORO DE INVESTIGACIÓN EDUCATIVA

### Desarrollo de tecnología educativa para invidentes

Martínez-Y-Cárdenas José Rodolfo,<sup>1</sup>  
Martínez Álvarez Carlos.<sup>2</sup>

CIDIR OAXACA IPN, Calle Hornos No. 1003, 7 1230. Sta. Cruz Xoxocotlán, Oax.

[rodolfo\\_mo@yahoo.com](mailto:rodolfo_mo@yahoo.com), 01 (951) 5170400 ext. 82739  
Fax 01 (951) 5170400

#### Resumen

En este trabajo se presentan tanto la metodología como las diferencias obtenidas en la evaluación del aprendizaje de videntes vs. invidentes después de que ambos grupos fueron sometidos a un proceso educativo básico de aprendizaje de patrones acústicos. Donde estos patrones acústicos son el resultado de la transformación de imágenes geométricas básicas a imágenes acústicas codificadas, con la finalidad de ofrecer un sistema de ayuda para los invidentes durante su navegación por escenarios complejos.

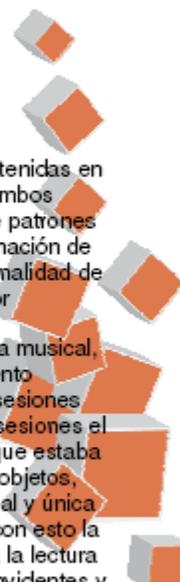
El grupo seleccionado de videntes se caracterizó por tener una experiencia musical, queriendo decir con esto que pueden ejecutar o ejecutaron algún instrumento musical en alguna etapa de su vida. Su entrenamiento consistió en cinco sesiones de aprendizaje con los patrones acústicos, en donde en las tres primeras sesiones el invidente sensó táctilmente el borde (silueta) de un elemento geométrico que estaba plasmado en forma resaltada sobre la tarjeta, formando así un total de 32 objetos, cada una de estas imágenes geométricas posee una correspondencia lineal y única con un patrón acústico además de tener un nombre. Finalmente se logró con esto la creación de conceptos de figuras geométricas acústicas, lo que equivale a la lectura de texto acústico. En forma paralela se siguió la misma metodología con invidentes y en este documento se presentan los resultados de ambos grupos.

Entre los resultados más importantes de este trabajo es que ambos grupos tienen la capacidad para decodificar los patrones acústicos, lo cual significa que se puede desarrollar la lectura del texto acústico, por otro lado se muestra la posibilidad de enviar al cerebro información correspondiente a la forma de un objeto real mediante la obtención de su silueta. Este procedimiento de enviar datos físicos a través del sonido se conoce como sonificación.

**Palabra clave:** Invidentes, prótesis acústica, patrones acústicos, sonificación.

#### Metodología

Fomento a la  
Investigación Educativa



Martínez C., Martínez J. R., "Dibujo de la silueta de los objetos como una manifestación de la construcción conceptual a partir de la lectura del texto acústico", *14º. Congreso Internacional Mexicano de Acústica*. 14-16 Noviembre de 2007, León, Guanajuato, México. ISBN 978-907-27-1324-1

### **SD3 - Dibujo de la silueta de los objetos como una manifestación de la construcción conceptual a partir de la lectura del texto acústico**

Carlos Martínez Álvarez J<sup>3</sup>, José Rodolfo Martínez y Cárdenas<sup>4</sup>

CIIDIR OAXACA IPN, Calle Héroes No. 1003, 71230. Sta. Cruz Xoxocotlán, Oax.

[rodolfo\\_mc@yahoo.com](mailto:rodolfo_mc@yahoo.com), 01 (951) 5170400 ext. 82739

#### ***Resumen***

Para llevar a cabo el análisis de la influencia que tiene un sistema de visión acústica (Visor Acústico) sobre la cognición en invidentes se requiere de una estructura conceptual que sea capaz de aclarar el proceso de construcción de conceptos.

Por medio de la fusión de sentidos y la formación de conceptos, se busca interpretar las respuestas de invidentes con el uso de un sistema de visión acústica en prueba. Los resultados son alentadores según se argumenta en este documento.

Martínez J. R., Berrón V., Menchaca R., "Subjective evaluation in blind humans of the perception of acoustic patterns generated by image contours", *10<sup>th</sup> School on Acousto-Optics and Applications*, Gdansk-Sopot, Polonia, del 12 al 15 de mayo, pp. 33, 2008. ISBN 978-83-7326-522-6

---

Acousto-Optics and Applications Página 1 de 2



**UNIWERSYTET GDAŃSKI**

- Home
- General information
- History
- Organisers & structure
- Scientific topics
- Time & venue
- Programme
- Registration & conference fee
- Important dates
- Abstracts
- Proceedings
- Download
- Tourist information
- Useful links
- Office



**Welcome to the web page of the**

# 10<sup>th</sup>



**SCHOOL on ACOUSTO-OPTICS  
AND APPLICATIONS**

Gdańsk-Sopot, Poland  
12-15 May 2008

*The conference dedicated to  
Professor Antoni Śliwiński*

Honorary patronage



Marszałek  
Województwa Pomorskiego  
Marshal of the  
Pomeranian Voivodeship

Organisers



UNIWERSYTET GDAŃSKI

Partners




PAN  
Komitet Akustyki

Support

*infobase@osa.org*  
*InfoBase Comments*

<http://univ.gda.pl/~school/index.htm> 07/02/2008

---

## Subjective Evaluation of the Perception of Acoustic Patterns Generated by *Image Contours* in the Blind

Rodolfo Martínez<sup>1</sup>, Virginia Berron<sup>2</sup>,  
and Rolando Menchaca<sup>3</sup>

<sup>1</sup> Instituto Politécnico Nacional, CIIDIR Oaxaca, Calle Hornos 1003, Santa Cruz Xaxocotán, Oaxaca, México, C. P. 71230

<sup>2</sup> Universidad Tecnológica de la Mixteca, Carretera a Acatlilma m. 2.5, Huajuapán de León, Oaxaca, México, C.P. 69000.

<sup>3</sup> Instituto Politécnico Nacional, CIC, Av. Juan de Dios Batiz s/n, esq. Miguel Othon de Mendizábal, Unidad Profesional Adolfo López Mateos, Edificio CIC, Col. Nueva Industrial, Delegación Gustavo A. Madero, México, D. F. C.P. 07738

\*Corresponding author: [rodolfo\\_mc@yahoo.com](mailto:rodolfo_mc@yahoo.com)

**Abstract:** An image processing system was developed to extract outlines from common objects and transform them into acoustic patterns. Due to the complexity of the outlines and thus the subjective interpretation by blind subjects, the researchers decided to select several geometrical images and to transform them into acoustical patterns to make the subjective evaluation easier. The final objective of the study was the presentation of acoustical patterns to the blind to see their mental reconstruction of the original object.

A group of blind subjects were trained to be able to interpret the acoustical patterns generated by the system known as Visor Acústico. An experimental model was developed using Latin Squares to achieve a subjective evaluation of the construction of new concepts and of new knowledge in the blind participants. At the end of the evaluation, the participants were asked to draw on paper what they mentally perceived. These drawings provide evidence of the mental construction of the knowledge gained from acoustical patterns, as a representation of the physical shape of real world objects.

### References and links

1. P. Meijer, "An Experimental system for Auditory Image Representation", *IEEE Transactions on Biomedical Engineering*, Vol. 39, No.2 February (1992).
2. P. Meijer, "Vision Technology for the totally blind", *Seeing with Sound. The Voice* (2007), <http://www.seeingwithsound.com/voice.html>.
3. C. Cappella, C. Trullemans, "A real experimental prototype for enhancement of vision rehabilitation using auditory substitution", *IEEE Transactions on Biomedical Engineering*, Vol. 45, No. 10, 1279-1293 October (1998).
4. P. Fletcher, "Seeing with sound: A journey into sight", Lecture abstract. *The Voice at Tucson 2002*, in the plenary opening session on "Sensory Substitution I: Visual Consciousness in Blind Subjects. April 8 (2002).
5. A. Storer, "Visual perception through sensory substitution", Department of Cognitive and Neural Systems, Boston University (2006), Term paper CN73
6. T. M. Little y F. Jackson Hills, "Métodos estadísticos para la investigación en la agricultura", 79-86, Editorial Trillas (2002).

### 1. Introduction

In the search for alternative technologies that offer a better quality of life to the blind and integrate them more efficiently into modern life, the concept of 'seeing' with the ears has been the subject of research.

Martínez J. R., Martínez C., Berrón V., Menchaca F. R., "Ensayo sobre subjetividad sensorial humana", *15o. Congreso Internacional Mexicano de Acústica*, Taxco, Guerrero, México, 2008.

## **PL4 Ensayo sobre la subjetividad sensorial humana**

José Rodolfo Martínez y Cárdenas\*, Carlos Martínez Álvarez\*, Virginia Berrón Lara\*\*  
CIIDIR OAXACA IPN, Calle Hornos No. 1003, C. P. 71230. Sta. Cruz Xoxocotlán, Oax.

[rodolfo\\_mc@yahoo.com](mailto:rodolfo_mc@yahoo.com), 01 (951) 5170400 ext. 82739

\*Investigador del CIIDIR Oaxaca, IPN y Becario de COFAA

\*\*Investigador de la Universidad Tecnológica de la Mixteca

### **Resumen**

Desde el punto de vista epistemológico podemos definir al subjetivismo humano como el equivalente a las verdades personales. La psicobiología se ha volcado al estudio intensivo de la subjetividad de los sentidos en donde esta juega un papel muy importante en la percepción humana. El estudio temático de estos elementos permiten entender como es que se comportan los sentidos primarios como son la visión, el auditivo, el del tacto, el gusto y el olfato: Ahora es bien sabido que los sentidos se apoyan unos a otros con la finalidad de obtener información más sustentada, y así optar por una decisión perceptiva más concreta. Finalmente, la fusión sensorial de información proporciona perspectivas mentales asombrosas sobre el mundo. El objetivo es analizar la concepción de disciplinas cuyo propósito es proporcionar información tangible sobre la realidad del mundo a las personas con habilidades diferentes o especiales.

Martínez J. R., Berrón V., Menchaca F. R. "La generación de conceptos de forma de las cosas a partir de la lectura de patrones acústicos por invidentes", *16º. Congreso Internacional Mexicano de Acústica*, Uruapan, Mich., México, 2009. ISBN 978-607-00-1932-6

### **TG7 - La generación de conceptos de forma de las cosas a partir de la lectura de patrones acústicos por invidentes.**

José Rodolfo Martínez y Cárdenas\*, Virginia Berón Lara\*\*, Felipe Rolando Menchaca García\*\*\*  
CIIDIR Oaxaca IPN, Calle Hornos No. 1003, C.P. 71230. Sta. Cruz Xoxocotlán, Oax.  
[rodolfo\\_mc@yahoo.com](mailto:rodolfo_mc@yahoo.com), 01(951)5170400 ext. 82739

\*Investigador del CIIDIR Oaxaca, IPN y Becario de COFAA

\*\* Investigador de la Universidad Tecnológica de la Mixteca

\*\*\* Investigador del CIC, IPN y Becario de COFAA

#### **Resumen**

Debido a la carencia de la visión en los invidentes a través de su vida no logran desarrollar una gran cantidad de conceptos de las formas que tienen los objetos que son comunes para la mayoría de las personas con visión normal. En este trabajo se presentan los resultados obtenidos con un grupo de invidentes los cuales fueron entrenados para hacer lectura de un grupo de patrones acústicos y se presentan los dibujos efectuados por estos de los patrones a los cuales fueron expuestos. Se hace hincapié en como describen los invidentes la construcción de su nuevo conocimiento y el impacto psicológico que esto tiene en su forma de percibir el mundo con base a la manifestación emocional que esto les produce.

Martínez J. R., Berrón V., Menchaca F. R., "Prótesis visual para invidentes con base en la lectura de patrones acústicos", 8º. Simposio Internacional de Estudios Oaxaqueños", Oaxaca, Oax., del 25-27 de junio, 2009. Instituto Welte para Estudios Oaxaqueños.



## Instituto Welte para Estudios Oaxaqueños

Emilio Carranza 203 • Colonia Reforma • Oaxaca, Oaxaca 68050  
Tel/Fax: (951) 513-8323  
E-mail: welte@prodigy.net.mx • Red mundial: <http://www.welte.org>

# CONSTANCIA

Por este conducto se hace constar que

## MARTÍNEZ Y CÁRDENAS, José Rodolfo

presentó los trabajos de investigación:

**MARTÍNEZ Y CÁRDENAS, José Rodolfo (CIIDIR-Oaxaca), Virginia BERRÓN LARA (UTM). Rolando MENCHACA GARCÍA (CIC-IPN) Prótesis Visual para Invidentes con Base en la Lectura de Patrones Acústicos**

**MARTÍNEZ Y CÁRDENAS, José Rodolfo, Fidel DIEGO NAVA (CIIDIR-Oaxaca), Armando E. de la FIGAL COSTALES (UNAH, la Habana, Cuba). Desarrollo de Sistemas Electrónicos para la Caracterización de los Nuevos Diseños de Aperos para la Labranza**

en el

**OCTAVO SIMPOSIO INTERNACIONAL DE ESTUDIOS OAXAQUEÑOS,  
25-27 de Junio, 2009**

Se extiende la presente constancia el día veintisiete del mes de junio del 2009, para los fines que al interesado convengan.

Dr. Ronald Waterbury, Secretario  
Instituto Welte para Estudios Oaxaqueños

CONACULTA • INAH



**ciesas**

Abstract de la participación en el Octavo Simposio Internacional de Estudios Oaxaqueños, 25-27 de junio, 2009.

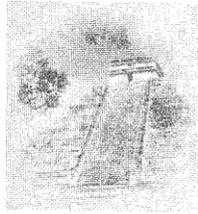
**MARTÍNEZ Y CÁRDENAS, José Rodolfo (CIDIR-Oaxaca), Virginia BERRÓN LARA (UTM), Rolando MENCHACA GARCÍA (CIC-IPN) Prótesis Visual para Invidentes con Base en la Lectura de Patrones Acústicos.** Estamos desarrollando un sistema computacional con la capacidad de poder extraer el contorno del objeto principal que se encuentre frente a un invidente. Este contorno se transforma a un patrón acústico siguiendo un protocolo ya establecido que conforma las bases del lenguaje que deberá de ser apreciado y descifrado por el invidente, permitiéndole esto que logre reconstruir en su mente la forma del objeto. Es bien sabido en la ciencia de la percepción visual que el contorno de un objeto es uno de los parámetros esenciales para la visión. Sin embargo, aunque este modelo de prótesis es primitivo le proporcional al invidente una idea más robusta que la del tacto de la forma de los objetos que están en frente a él. (10)



The screenshot shows a web browser window displaying the Acoustical Society of America (ASA) website. The page is titled "Acoustics pattern code to represent the contour of any visual image for people who are blind." and is part of the "Journal of the Acoustical Society of America / Volume 128 / Issue 4 / SECOND PAN-AMERICAN/IBERIAN MEETING ON ACOUSTICS". The authors listed are José Rodolfo Martínez y Cardenas<sup>1</sup>, Felipe Rolando Menchaca García<sup>2</sup>, and Virginia Berrón Lara<sup>3</sup>. The abstract text reads: "In the search for rehabilitation for people who are partially blind through non-invasive means, an option is by the use of the ear. This paper proposes a conceptual prosthesis model called Acoustics-viewer. The idea is to generate acoustic patterns from contours representing optical images, from frequencies in the range of three octaves from a piano sound source." The page also includes a "Full Text: Download PDF FREE | View Cart" link and a "Permalink" to the article's DOI: <http://dx.doi.org/10.1121/1.3507990>. The website's navigation menu includes Home, About, ASA Publications, Meetings, Membership, Funding Resources, Education/Outreach, About Acoustics, Standards, and Store. The footer of the page indicates the copyright is © 2010 Acoustical Society of America.

## 2nd Pan-American and Iberian Meeting on Acoustics

15 – 19 November 2010



160<sup>th</sup> ASA meeting  
7<sup>o</sup> Congress FIA  
17<sup>o</sup> Congress IMA



## CANCUN – MEXICO

## Session Code:

**Un código de patrones acústicos para representar el contorno de cualquier imagen visual para invidentes**

A code of acoustic patterns to represent the contour of any visual image for the blind

José Rodolfo Martínez y Cárdenas<sup>1</sup>(Becario del SIBE de la COFAA del IPN), Felipe Rolando Menchaca García<sup>2</sup>, Virginia Berrón Lara<sup>3</sup>

<sup>1</sup>Investigador del CIIDIR Oaxaca IPN, Calle Hornos No. 1003, C.P. 71230. Sta. Cruz Xoxocotlán, Oax., México, [rodolfo\\_mc@yahoo.com](mailto:rodolfo_mc@yahoo.com), 01(951)5170400 ext. 82739; <sup>2</sup> Investigador de la ESIME, IPN, México D.F., [fmenchac@gmail.com](mailto:fmenchac@gmail.com); <sup>3</sup> Investigador de la Universidad Tecnológica de la Mixteca, Huajuapán de León, Oaxaca, México, [berron@mixteco.utm.mx](mailto:berron@mixteco.utm.mx)

J. Acoust. Soc. Am., Vol. 128, No. 4, Pt. 2, October 2010

2<sup>nd</sup> Pan-American/Iberian Meeting on Acoustics

**Summary**

In the search for partial rehabilitation of blindness through non-invasive methods is the replacement of vision by ear. Here, we propose a conceptual model of prosthesis, which we call "Acoustic Viewer." The main contribution in this research is the generation of acoustic patterns formed exclusively by the representative of the image contour and the use of three octaves of the piano.

Representing the image through the acoustic patterns, we look the way that the blind can easily decode them. As found in the literature review, this is because blind people so far have encountered difficulties and dissatisfaction in the decoding of the patterns because of the complexity of them.

We propose a basic alphabet of acoustic patterns, through which trained two groups of subjects, a group with normal vision and the other formed by the blind. This was in order to demonstrate that blind people can develop the ability to decode the alphabet proposed acoustic patterns.

**Keywords:** sensory substitution, sonification, sound Viewer, CSound and acoustic patterns, acoustic vision.

## Codification of the contour of the visual image to acoustics patterns and its decoding by blind

José Rodolfo Martínez y Cárdenas<sup>1</sup>, Felipe Rolando Menchaca García<sup>2</sup>,  
Virginia Berrón Lara<sup>3</sup>

<sup>1</sup>Centro Interdisciplinario de Investigación para el Desarrollo Integral Regional (CIIDIR),  
Instituto Politécnico Nacional (IPN), Calle Hornos N° 1003, C.P. 71230,  
Sta. Cruz Xoxocotlán, Oax., México. Telf. 01(951)5170400 ext. 82739. [rodolfo\\_mc@yahoo.com](mailto:rodolfo_mc@yahoo.com)

<sup>2</sup>Centro de Investigación de Cómputo (CIC), IPN, México D.F. [jmenchac@gmail.com](mailto:jmenchac@gmail.com)

<sup>3</sup>Universidad Tecnológica de la Mixteca, Huajuapán de León, Oax., México.  
[berron@mixteco.utm.mx](mailto:berron@mixteco.utm.mx)

### Abstract

In this paper is proposed the development's first phase of a conceptual model for a non-invasive prosthesis for helping blind people to obtain information about the shape of objects. The prosthesis has been called "Acoustic Viewer". It is based on an encoding of digital image contours into an acoustic pattern which is conceptually simpler than those are found in previous papers. The digital imaging processing algorithms and the method for transforming image into sound using *CSound* are described here. In order to evaluate coding's suitability were chosen 33 (basic) figures, which are formed of straight lines and curves. A training course in their decoding was designed and an experiment with two groups of people, the first consisting only of blind people and the second consisting of people with normal vision, who were trained and finally they participated in a random trial to assess their acquired ability to decoding these patterns. Our results show that both groups developed the same ability for decoding the proposed patterns. This paper shows "Acoustic Viewer" could help blind people.

**Key words:** acoustic patterns and blindness, non-invasive prosthesis for blind people, sensorial substitution, acoustic vision.

## Codificación del contorno de la imagen visual a patrón acústico y su decodificación realizada por invidentes

### Resumen

En este trabajo se propone la primera etapa del desarrollo de un modelo conceptual de una prótesis no invasiva para ayudar a los invidentes a obtener información sobre la forma que poseen los objetos. Se ha llamado a esta prótesis "Visor Acústico". Se basa en una codificación del contorno de la imagen digital en un patrón acústico conceptualmente más sencilla que las presentadas en trabajos anteriores; se describen los algoritmos de procesamiento digital de la imagen utilizados para la extracción del contorno y los de su posterior conversión a sonido mediante el uso de *CSound*. Para evaluar la idoneidad de la codificación se propuso la transformación de 33 contornos (básicos) formados por líneas rectas y curvas, un curso de entrenamiento en su decodificación y un experimento en el que a dos grupos de personas, uno de invidentes y el otro de sujetos con visión normal, se les entrenó, y después, sometió a una prueba aleatoria de estímulos con 8 de los patrones propuestos con la finalidad de valorar su capacidad de decodificación ad-

La imagen escaneada de las 4 publicaciones más importantes basadas en nuestro trabajo se encuentran en el archivo Anexo I del CD adjunto.



## **Anexo II: Figuras y patrones acústicos**

Ejemplos de algunas de las figuras y patrones acústicos del sistema básico están en el archivo Anexo II del CD adjunto a este trabajo



# Anexo III: Guión del curso de capacitación e instructivo para participantes

## I. Unidad didáctica inicial: Sensibilización táctil

### a. Objetivo

Al finalizar la unidad el invidente será capaz de reconocer un grupo de figuras geométricas básicas por medio de la percepción táctil del contorno resaltado de las imágenes sobre una tarjeta y podrá relacionarlo con su nombre genérico. Logrando con esto la creación de conceptos de forma de los contornos de las figuras geométricas propuestas gracias a la creación de un acervo de significados, con los que podrá, más adelante, hacer la lectura de textos acústicos.

### b. Contenidos

4. La memorización del contorno de diferentes figuras geométricas básicas con el tacto.
5. La correspondencia de la geometría de las figuras geométricas básicas con su nombre genérico.
6. El reconocimiento e identificación de las figuras geométricas básicas, dada una secuencia de exposición aleatoria por medio del tacto.

### c. Actividades

- Se le proporcionarán al invidente diferentes figuras geométricas; las palpará, se le mencionará el nombre de cada una de ellas. El invidente memorizará táctilmente

tanto la forma que describe todo el contorno de la figura, así como su nombre. Con esta experiencia deberá de crear los conceptos correspondientes.

## **II. Unidad de entrenamiento acústico**

### **a. Objetivo**

Al finalizar la unidad de entrenamiento acústico el invidente será capaz de reconocer una figura por medio de la decodificación (lectura) acústica del patrón acústico; en donde dicho patrón representará el contorno (forma); se buscará que lo relacione tanto con la forma geométrica, como con su nombre genérico.

### **b. Contenidos**

El invidente reconocerá el patrón acústico de diferentes figuras geométricas básicas y lo relacionará tanto con su nombre genérico, así como con su forma geométrica. El orden de la presentación de los eventos acústicos será aleatorio.

En esta unidad el invidente ya no usará las figuras geométricas en forma física, solamente decodificará los patrones acústicos.

### **c. Actividades**

- En la primera fase del entrenamiento se le harán oír diferentes patrones acústicos (archivos .wav) al invidente en forma aleatoria, los escuchará y se le dirá el nombre de cada uno de ellos con el objetivo de que fortalezca sus nuevos conceptos recién adquiridos.
- En la segunda fase del entrenamiento acústico el invidente relacionará cada patrón acústico con una figura y con su nombre genérico. Se evaluará su progreso y se obtendrán datos estadísticos del grupo de invidentes.

### **III. Unidad de decodificación de patrones acústicos**

#### a. Objetivo

Al finalizar la unidad el invidente será capaz de decodificar una secuencia de patrones acústicos (textos acústicos) expuestos en forma continua y deberá de poder leer la secuencia de las imágenes geométricas. Por ejemplo, una secuencia sería: *“círculo, hexágono, triángulo”*.

#### b. Contenidos

El invidente reconocerá una secuencia de patrones acústicos (lectura de texto acústico) y podrá describir verbalmente la secuencia sucesiva de las figuras geométricas. De esta forma se podrá evaluar su capacidad de lectura de un grupo de patrones sucesivos de patrones acústicos.

#### c. Actividades

Se le proporcionarán diferentes secuencias de patrones acústicos al invidente, hará la lectura de texto acústico y describirá verbalmente la secuencia decodificada (leída) por él.

### **IV. Unidad de evaluación**

#### a. Objetivo

Al finalizar esta unidad se evaluará las nuevas capacidades desarrolladas por el invidente

#### b. Contenidos

7. El invidente reconocerá y relacionará patrones acústicos expuestos en forma individual.

8. El invidente leerá textos acústicos y describirá la secuencia de éstos.
9. El invidente después de escuchar el patrón acústico dibujará sobre una hoja de papel la imagen que decodificó en su mente.

c. Actividades

10. Se le proporcionarán diferentes patrones acústicos (archivos .wav) a los sujetos siguiendo una secuencia probabilística proporcionada por la distribución de los cuadrados latinos.
11. Se hará un análisis estadístico entre los dos grupos de sujetos, los invidentes, y los sujetos con visión normal (grupo de control), con la finalidad de observar si existe una diferencia significativa entre ambos grupos.

## **V. Recursos materiales**

12. El instructor le leerá al invidente un instructivo, para darle a conocer la finalidad del curso.
13. Para el reconocimiento táctil de las figuras geométricas básicas, se le proporcionarán al invidente, los 33 contornos de las figuras geométricas propuestas como sistema básico básico sobre un cartón grueso, los contornos fueron resaltados utilizando fomix (hoja de esponja), de tal forma que se pueda palpar y sentir la forma de los contornos de las figuras al ser explorados con la yema de los dedos.

1. El instructor le proporcionará el nombre de cada una de las figuras geométricas al invidente, para que éste vaya creando un conjunto de conceptos basándose en su experiencia presente.
14. El instructor ejecutará los archivos de los patrones acústicos correspondientes a cada una de las figuras.
2. El instructor le proporcionará el nombre de cada una de las figuras geométricas al invidente para que éste vaya creando un conjunto de conceptos basándose en su experiencia presente.
15. Finalmente, el instructor ejecutará los archivos de sonido correspondientes a cada una de las figuras.

## **VI. Instructivo para el participante**

### Objetivos del curso

Este curso tiene como fin el ofrecerle los elementos cognitivos<sup>52</sup> necesarios para poder evaluar nuestros patrones acústicos básicos.

### Instrucciones

1. El instructor le proporcionará de una en una cada tarjeta, con cada una de las imágenes de las figuras geométricas básicas, usted con su tacto explorará cada contorno, al mismo tiempo le será proporcionado el nombre de la figura. Usted deberá relacionar la forma de la figura con el nombre que le proporcionaron.
2. Al finalizar esta primera parte, usted deberá de hacer el esfuerzo por recordar cada figura geométrica con su respectivo nombre.

---

<sup>52</sup> Cognitivo(a).- adj. Perteneciente o relativo al conocimiento.

3. Ahora el instructor le proporcionará de nuevo cada una de las tarjetas, de una en una, asistiéndole si tiene alguna duda en llegar a relacionar la imagen con su nombre.
4. El instructor le proporcionará una por una cada figura, le dirá de nuevo el nombre genérico de la figura; generará con la computadora una secuencia “musical” que describe la forma de la imagen. A esta descripción “musical” de la imagen le hemos llamado patrón acústico.
5. El instructor generará una secuencia de diferentes patrones acústicos y usted tendrá que decidir cuál forma describen, a este ejercicio se le conoce como lectura de patrones acústicos.

## Anexo IV. Topología Digital

Este anexo contiene parte del trabajo [Martínez, 2005] con algunas modificaciones para su mejor comprensión.

**Definición 1.** Por  $\mathbf{Z}^2$  entenderemos el conjunto de todas las parejas ordenadas de números enteros. En símbolos,  $\mathbf{Z}^2 = \{(i, j) \mid i \in \mathbf{Z}, j \in \mathbf{Z}\}$ .

A los elementos de  $\mathbf{Z}^2$  les llamaremos *puntos latís*.

En el plano coordenado  $\mathbf{Z}^2$  corresponde a los puntos del plano cartesiano cuyas coordenadas son números enteros.

**Definición 2.** De dos puntos latís  $p$  y  $q$  distintos con coordenadas  $(i_1, i_2)$  y  $(j_1, j_2)$  respectivamente, decimos que son *8-adyacentes* si,  $0 < |i_1 - j_1| + |i_2 - j_2| \leq 2$  y *4-adyacentes*, si  $|i_1 - j_1| + |i_2 - j_2| = 1$ .

Nótese que en el caso de 4-adyacencia los puntos difieren en exactamente una coordenada, mientras que en la 8-adyacencia pueden diferir en las dos.

Si  $p$  y  $q$  son 8-adyacentes, decimos que *8-vecinos*. Análogamente si  $p$  y  $q$  son 4-adyacentes,  $p$  y  $q$  son *4-vecinos*.

**Definición 3.** Sea  $S$  un subconjunto no vacío de  $\mathbf{Z}^2$  tal que existen dos subconjuntos de  $S$ ,  $B$  y  $W$ , con la propiedad de que  $S = B \cup W$ ,  $B \cap W = \emptyset$ . A los elementos del conjunto  $B$  se les llamará **puntos latís negros**; diremos que dos de sus elementos son vecinos, si son 8-

vecinos y a los elementos del conjunto  $W$  se les llamará **puntos latís blancos**; diremos que dos de sus elementos son vecinos, si son 4-vecinos.

**Definición 4.** Una imagen binaria  $M$  se puede modelar como una matriz  $M=[M_{i,j}]$  en la que las entradas 0 corresponden a los pixeles negros y las entradas con valor 255 corresponden a los pixeles blancos. A esta matriz le hacemos corresponder un conjunto  $S$  de  $Z^2$ , de la siguiente forma: al lugar  $(i,j)$  de la matriz  $M$ , le hacemos corresponder el punto latís  $(i,j)$  de  $Z^2$ .  $S$  tiene tantos elementos como entradas tiene la matriz  $M$ . Los puntos correspondientes a las entradas  $M_{i,j}$  con valor 0 conformarán el conjunto  $B$  y los correspondientes a las que tienen el valor 255, el conjunto  $W$ .

Por una *imagen digital binaria*  $P$  entenderemos una tétroda  $P=(S, 8, 4, B)$ , donde  $B$  es el conjunto de los puntos latís negros junto con 8-vecindad para éstos y 4-vecindad para los blancos.

Con esta correspondencia en mente diremos indistintamente pixeles o puntos latís.

**Definición 5.** Decimos de dos puntos latís  $p$  y  $q$  están **m-conectados** si son del mismo color y, o son iguales o bien, si son distintos, existe una sucesión finita de puntos latís  $p=p_0, p_1, \dots, p_k, q$  en  $S$  tales que  $p_i$  y  $p_{i+1}$  son  $m$ -vecinos (y del mismo color),  $i=0, \dots, k-1$ . Aquí  $m$  es igual a 8, si los puntos son negros y 4, si son blancos.

**Definición 6.** Decimos  $B'$  un subconjunto de  $B$  es una *8-componente conexa* si todos sus puntos están 8-conectados y no hay ningún punto en  $B-B'$  que esté conectado a algún punto de  $B'$ . En término más formales, si  $B'$  es un conjunto conexo maximal.

Análogamente se define una 4-componente conexa para el conjunto de los pixeles blancos.

**Definición 7.** Si  $p = (i, j)$ , entonces  $r = (i, j-1)$  es llamado su *vecino superior*,  $s = (i+1, j-1)$  es llamado el *vecino superior derecho*,  $t = (i-1, j-1)$  el *vecino superior izquierdo* y  $u = (i-1, j)$  el *vecino izquierdo*.

### **Algoritmo de etiquetado de componentes negras de una imagen binaria**

Hágase un barrido de la imagen pixel a pixel de izquierda a derecha y de arriba hacia abajo.

E será un contador que asignará las etiquetas. Se inicializa con, por ejemplo, el valor 2.

**Paso 0.** Se empieza a recorrer la imagen hasta encontrar el primer pixel negro al que se etiqueta con el valor de E. E se incrementa en una unidad.

**Paso 1.** Se busca el siguiente pixel negro p. Si no se encuentra, ir a Paso 3.

Se examinan los pixeles vecinos r, s, t y u correspondientes.

Se tomarán en cuenta de éstos, los que tengan el color negro.

- a) Si ninguno de estos pixeles está etiquetado o no hay, asignarle la etiqueta E e incrementar E en una unidad.
- b) Si todos los vecinos tienen la misma etiqueta, asignarle ésta a p.
- c) Si tienen etiquetas distintas, se elige la etiqueta con menor valor de dichos vecinos reemplazándose en la imagen todas las coincidencias de dichas etiquetas por la etiqueta seleccionada.

**Paso 2.** Ir a paso 1.

**Paso 3.** Fin del algoritmo.

Un algoritmo análogo para las 4-componentes blancas considera únicamente los vecinos  $r$  y  $u$ .

**Definición 8.** Ahora pensemos que una imagen binaria  $P$  es un subconjunto del plano cartesiano  $\mathbf{R}^2$ . A cada punto de la imagen binaria  $P$ ,  $p=(i,j)$  le asignamos los siguientes conjuntos:

1.  $\{(i-\frac{1}{2}, j-\frac{1}{2}), (i+\frac{1}{2}, j-\frac{1}{2}), (i-\frac{1}{2}, j+\frac{1}{2}), (i+\frac{1}{2}, j+\frac{1}{2})\}$ , cuyos elementos serán llamados *vértices*;
2. Sean  $a_1 = \{(i_1, j-\frac{1}{2}) \in \mathbf{R}^2 \mid i-\frac{1}{2} \leq i_1 \leq i+\frac{1}{2}\}$ ,  $a_2 = \{(i_1, j+\frac{1}{2}) \in \mathbf{R}^2 \mid i-\frac{1}{2} \leq i_1 \leq i+\frac{1}{2}\}$ ,  
 $a_3 = \{(i-\frac{1}{2}, j_1) \in \mathbf{R}^2 \mid j-\frac{1}{2} \leq j_1 \leq j+\frac{1}{2}\}$ ,  $a_4 = \{(i+\frac{1}{2}, j_1) \in \mathbf{R}^2 \mid j-\frac{1}{2} \leq j_1 \leq j+\frac{1}{2}\}$ .  
 $a_1, a_2, a_3$  y  $a_4$  (que son subconjuntos del plano) serán llamados *aristas*.
3.  $\{(i_1, j_1) \in \mathbf{R}^2 \mid i-\frac{1}{2} \leq i_1 \leq i+\frac{1}{2}, j-\frac{1}{2} \leq j_1 \leq j+\frac{1}{2}\}$ , un cuadrado sin sus bordes.

La unión de todos estos conjuntos puede verse como un cuadrado (con sus bordes incluidos) de lado 1 centrado en el punto  $(i,j)$  con lados paralelos a los ejes coordenados al que también se le llamará *pixel*. Nótese que en esta definición no hacemos distinción de colores.

**Definición 9.** Sea  $p$  un pixel negro de una imagen binaria  $(S, 8, 4, B)$ . Decimos que  $p$  tiene un *vértice común* a  $B$ , si existe algún otro pixel de  $B$ ,  $q$ , diferente a  $p$ , tal que este vértice también es vértice de  $q$ . Análogamente para las aristas.

El *índice del pixel*  $p$  en relación al conjunto  $B$  de los pixeles negros es, por definición, igual a la diferencia del número de vértices de  $p$  comunes a  $B$  y el número de aristas de  $p$  comunes a  $B$ .

---

Los pixeles negros de índice 0 adyacentes a al menos a otro pixel negro forman el relleno de un objeto, pues tienen “suficientes” pixeles negros a su alrededor; su eliminación aumenta el número de componentes blancas.[Ramírez, 2002], pag. 37.



## **Anexo V Ejemplos de archivos orquesta y partitura**

Ejemplos de los archivos orquesta y partitura del patrón acústico correspondiente al cuadrado; en el CD.