



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Departamento de Sistemas Informáticos y Computación
Universidad Politécnica de Valencia

Simulación de un entorno real utilizando holofonías

Trabajo Fin de Máster

**Máster Universitario en Inteligencia Artificial,
Reconocimiento de Formas e Imagen Digital**

Autor:

Andrés Estuardo Prieto López

Directoras:

Dra. M. Carmen Juan Lizandra

Dra. Larisa Dunai Dunai

Valencia, septiembre 2016

Agradecimientos

Agradezco a Dios por ser mi guía, a mi familia por todo su apoyo incondicional. También a todas las personas que de una u otra forma me ayudaron en todo este tiempo de estudio.

Resumen

Hay 39 millones de personas ciegas en el mundo y 246 millones presentan baja visión [1]. La mayoría de las personas ciegas tienen problemas al trasladarse de un lugar a otro, si están en su casa pueden percibir que hay alguien, pero no saben quién es, puede ser un amigo, familiar o un ladrón.

Las personas ciegas por lo general utilizan la memoria para recordar el lugar donde dejan los objetos que usan, si viven con más personas que pueden ver y éstas mueven esos objetos. Esto podría acabar siendo un problema para ellas, ya que les resultaría difícil saber dónde están dichos objetos.

El objetivo principal del presente Trabajo Fin de Máster es el desarrollo y validación de una aplicación, que, mediante realidad aumentada o sonido aumentado, permita a personas ciegas, apreciar el entorno real por medio de holofonías. La metodología que se utiliza es visión por computador para poder representar el entorno con mapas de profundidad y luego reproducir sonidos dependiendo de la distancia a la que se encuentren los objetos.

La aplicación se centra en la detección de mapas de profundidad, para poder determinar la distancia que hay entre la cámara del móvil y los objetos que ésta capta. De modo que, dependiendo de la distancia a la que se encuentren los objetos, se reproducirán sonidos holofónicos, que permitirán apreciar si se encuentran cerca o lejos, a la izquierda o a la derecha, arriba o abajo y así, poder obtener una representación auditiva de lo que se tiene enfrente.

La aplicación usa realidad aumentada y funciona en dispositivos móviles. Se desarrolló en un motor de videojuegos multiplataforma, Unity 3D, con la integración de RealSpace3D para la gestión de holofonías y OpenCV para el análisis de las imágenes.

La aplicación está dividida en tres partes, la primera para la adquisición de las imágenes. La segunda para determinar el mapa de profundidad y la tercera para reproducir el sonido dependiendo de la imagen procesada del mapa de profundidad.

Con la aplicación se hicieron pruebas para determinar la eficiencia de los mapas de profundidad utilizando una cámara y los sensores (giroscopio y acelerómetro) del móvil. Las pruebas para los sonidos holofónicos se desarrollaron en un videojuego de entorno virtual, para determinar la percepción de los sonidos. Las pruebas que se hicieron basadas en la metodología y los resultados obtenidos de la investigación realizada, han servido para vislumbrar el potencial de la aplicación, para que ésta funcione en tiempo real, y se pueda utilizar en múltiples campos, como en educación, videojuegos, diseño y arquitectura, entre otros.

Contenido

| | |
|---|----|
| 1. Introducción | 9 |
| 1.1. Motivación | 9 |
| 1.2. Objetivos | 10 |
| 1.3. Estructura de la Memoria | 10 |
| 2. Estado del Arte | 12 |
| 2.1. Introducción..... | 12 |
| 2.2. Ceguera | 12 |
| 2.3. Aplicaciones para personas invidentes | 13 |
| 2.3.1. Posibilidades de la holofonía | 13 |
| 2.3.2. Mapas de profundidad | 15 |
| 2.3.3. Proyectos dedicados a personas invidentes..... | 17 |
| 3. Desarrollo de la Aplicación | 20 |
| 3.1. Arquitectura del sistema..... | 20 |
| 3.2. Hardware..... | 21 |
| 3.3. Software | 22 |
| 3.3.1. Unity | 22 |
| 3.3.2. RealSpace3D..... | 23 |
| 3.3.3. OpenCV..... | 24 |
| 3.3.4. OpenAL Soft | 25 |
| 3.4. Implementación..... | 26 |
| 3.5. Procedimiento | 36 |
| 3.6. Validaciones..... | 39 |
| 4. Conclusiones y trabajos futuros..... | 45 |
| 4.1. Conclusiones..... | 45 |
| 4.2. Trabajos futuros | 46 |
| 5. Bibliografía..... | 47 |
| 6. Anexos..... | 50 |

Lista de Figuras

| | |
|--|----|
| Figura 1 Clasificación de los dispositivos electrónicos para invidentes..... | 9 |
| Figura 2 Diseño del Dummy head de Hugo Zuccarelli | 14 |
| Figura 3 Mapa de profundidad representado por colores..... | 15 |
| Figura 4 Dos imágenes tomadas con una cámara desplazada..... | 16 |
| Figura 5 Motion Parallax | 16 |
| Figura 6 Perro lazarillo | 17 |
| Figura 7 The vOICe | 18 |
| Figura 8 The vOICe para Android | 18 |
| Figura 9 EYEMUSIC | 19 |
| Figura 10 Metodología | 20 |
| Figura 11 Huawei P7-L10..... | 21 |
| Figura 12 Desarrolladores registrados en Unity desde 2012 hasta 2016..... | 23 |
| Figura 13 Configuración del sonido en RealSpace3D..... | 24 |
| Figura 14 Círculo representado con los sonidos | 25 |
| Figura 15 License Key RealSpace3D versión de prueba | 26 |
| Figura 16 Flujo óptico a resolución 320 x 240 | 27 |
| Figura 17 Chessboard utilizado para la calibrar de la cámara..... | 28 |
| Figura 18 Imagen Izquierda ojo de pez y a la derecha rectificada | 28 |
| Figura 19 Imagen rectificada Huawei P7-L10..... | 29 |
| Figura 20 Imagen de un fotograma | 29 |
| Figura 21 Imágenes estabilizadas..... | 30 |
| Figura 22 Regiones de la imagen..... | 31 |
| Figura 23 MatchTemplate función de openCV | 31 |
| Figura 24 Líneas epipolares | 32 |
| Figura 25 Imágenes alineadas | 32 |
| Figura 26 Comparación de píxeles..... | 33 |
| Figura 27 Mapa de profundidad | 33 |
| Figura 28 Representación visual de la circunferencia | 34 |
| Figura 29 Representación del área circular..... | 35 |
| Figura 30 RealSpace3D demo en Unity | 35 |
| Figura 31 Ejes de rotación del móvil | 36 |
| Figura 32 Datos obtenidos del acelerómetro..... | 37 |
| Figura 33 Datos obtenidos del giroscopio | 38 |
| Figura 34 Flujo de las imágenes mostrando la orientación | 39 |
| Figura 35 Circunferencia de puntos representada con los sonidos..... | 40 |
| Figura 36 Línea de puntos vertical representada con los sonidos..... | 40 |
| Figura 37 Línea de puntos horizontal representada con los sonidos | 41 |
| Figura 38 Punto representado con los sonidos | 41 |

Lista de Tablas

| | |
|---|----|
| Tabla 1 Información de los participantes y su percepción de los sonidos | 42 |
| Tabla 2 Sonidos que se escuchan similar | 43 |
| Tabla 3 Relación de la forma con el sonido | 43 |
| Tabla 4 Sonidos que la persona tenía que reconocer | 44 |

1. Introducción

1.1. Motivación

La independencia de una persona es uno de los factores que define el carácter de la misma. Sentirse capaz de dar un paseo por los parques, por la calle, o simplemente poder ir de compras o al trabajo, poder reconocer a las personas de su entorno, reconocer los colores, los materiales de los productos, o simplemente no equivocarse entre un refresco de Coca Cola zero y una Coca Cola sin cafeína, sin mencionar los demás productos necesarios para la alimentación, es el deseo más apreciado por las personas con discapacidad visual.

Con la mejora en diversas tecnologías, los investigadores tratan de desarrollar dispositivos para mejorar la vida diaria de las personas invidentes, basándose en los sentidos de percepción humana, como son el tacto y los sonidos [2 y 3].

Hasta la última década había alrededor de 45 dispositivos electrónicos que servían de ayuda a las personas invidentes [4]-[13], pero en los últimos 5 años el uso de estos dispositivos se ha incrementado, para tratar de diseñar el mejor dispositivo electrónico no invasivo capaz de reemplazar la vista, que a su vez sea simplemente un dispositivo o un sistema de apoyo que complemente los sentidos humanos, sin interferir con ellos.

En la figura 1 se representa una clasificación de los dispositivos electrónicos que sirven de ayuda a las personas invidentes según el principio de funcionamiento o, mejor dicho, según su objetivo.

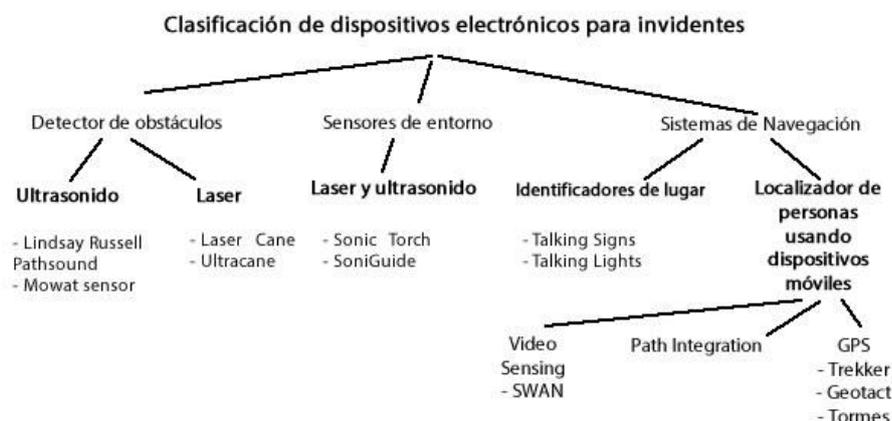


Figura 1 Clasificación de los dispositivos electrónicos para invidentes

A pesar de la cantidad de equipos desarrollados hasta la fecha que están en continuo crecimiento, actualmente no hay en el mercado alguno que cumpla con los requisitos, tanto técnicos como ergonómicos, económicos y como no, los requisitos de los usuarios que se sobrepone a los requisitos de desarrollo de todos los productos: fácil de manejar, con buen aspecto, que sea resistente, no ser conductor térmico ni eléctrico, poseer un peso pequeño y un diseño moderno, adaptable o que se pueda actualizar, etc.

1.2. Objetivos

Las personas tenemos la necesidad de poder realizar las actividades comunes del día a día de forma independiente. Es por eso que, el objetivo principal de este proyecto es implementar una metodología que permita representar el entorno real de una manera no visual para las personas invidentes.

Para lograrlo representaremos el entorno usando sonidos holofónicos, que permiten una percepción de profundidad, distancia y posición de los sonidos, para que puedan moverse con mayor seguridad y a su vez, sentirse autosuficientes, mejorando así su estilo de vida.

Los objetivos específicos del presente trabajo son:

- Representar el entorno real para poder determinar la distancia a la que se encuentran del dispositivo móvil.
- Comparar el entorno real con los mapas de profundidad, para comprobar que éstos sean correctos.
- Reproducir los sonidos holofónicos, utilizando los mapas de profundidad para determinar la posición y la distancia a la que se encuentran los objetos del móvil.
- Corregir posibles errores en la metodología utilizada para poder realizarla en tiempo real, como siguiente paso.

1.3. Estructura de la Memoria

1. Introducción:

En este capítulo se explica cuál ha sido la motivación para el desarrollo del proyecto y los objetivos que se pretenden alcanzar.

2. Estado del arte:

En este capítulo se revisa la literatura relevante sobre holofonía, mapas de profundidad, realidad aumentada y aplicaciones enfocadas para personas invidentes.

3. Desarrollo:

En este capítulo se explica cómo se ha desarrollado la aplicación. Exponiendo las consideraciones que se tuvieron en cuenta, explicando el funcionamiento y las innovaciones de la misma con respecto a otros sistemas similares. También se dan a conocer detalles de su estructura y posible ampliación de la aplicación para futuras investigaciones.

4. Validaciones y análisis de resultados:

En este capítulo se expone el análisis de los datos obtenidos, las pruebas que se realizaron para obtener dichos datos y la evaluación de los mismos. Todo esto con el fin de cumplir el objetivo principal.

5. Conclusiones y trabajo futuro:

En este capítulo se concretan todos los aspectos del proceso de investigación. Además, se mencionan posibles ampliaciones del mismo trabajo para implementarlas en el futuro.

2. Estado del Arte

2.1. Introducción

En este capítulo se hace un repaso sobre la ceguera y aplicaciones para personas invidentes que se han llevado a cabo hasta la fecha. Primero se aborda el tema de la ceguera y sus consecuencias, posteriormente se muestran las metodologías y aplicaciones que se han utilizado para dar una solución a este problema.

2.2. Ceguera

La ceguera es la pérdida total o parcial del sentido de la vista. Existen varios tipos de ceguera parcial dependiendo del grado y tipo de pérdida de visión. La función visual se subdivide en cuatro niveles:

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

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

La ceguera se divide en 4 tipos:

Ciegos: son aquellos sujetos que tienen sólo percepción de luz, sin proyección, o aquellos que carecen totalmente de visión. Desde el punto de vista educativo, ciego es aquel que aprende sistema Braille y no puede utilizar su visión para adquirir ningún conocimiento, aunque la percepción de la luz pueda ayudarle para sus movimientos y orientación.

Ciegos parciales: son aquellos sujetos que mantienen unas posibilidades visuales mayores, tales como capacidad de percepción de la luz, percepción de bultos y contornos, algunos matices de color, etc.

Personas con baja visión: son los que mantienen un resto visual que les permite ver objetos a pocos centímetros. A estos no se les debe llamar nunca ciegos ni se les debe educar como tales, aunque tengan que aprender procedimientos "táctiles" para aumentar sus conocimientos.

Limitados visuales: son los que precisan, debido a sus dificultades para aprender, una iluminación o una presentación de objetos y materiales más adecuadas, utilizando lentes, aumentando la iluminación, etc.

2.3. Aplicaciones para personas invidentes

2.3.1. Posibilidades de la holofonía

El sonido holofónico, u holofonía es una técnica de grabación binaural que fue desarrollada por Hugo Zucarelli en los años 80's [14].

Holofonía es el equivalente fonético de la holografía visual, es una tecnología con la capacidad de registrar el sonido exactamente como es. La dimensión espacial, el ambiente, la emoción y el espíritu de vida de la fuente sonora se presentan de tal manera que nuestros cerebros no tienen otra opción que decirnos que el contenido de la grabación es, de hecho, la vida real. No hay tratamiento, no hay efectos especiales, sólo una señal pura, mezclada con una buena dosis de ingenio [15].

Para la grabación de holofonía se usa el mismo sistema con cabeza de dummy que para el sonido binaural, solo que introduciendo dentro de ésta un emisor de sonido de baja intensidad. Zuccarelli le agregó a su cabeza, llamada "Ringo", una emisión de sonido interna de referencia, así los micrófonos graban las interferencias creadas entre ambos sonidos (exterior y el propio), consiguiendo la tridimensionalidad del sonido [16].

La holofonía se diferencia de la técnica de Head-Related Transfer Function (HRTF), en que esta última trata de simular mediante el procesamiento de la señal lo que la holofonía registra de manera directa. HRTF representa de forma matemática el tipo de ecuación o transformación que aplica el conjunto de nuestro cuerpo (hombros, cabezas, etc.) a un sonido en el aire libre cuando éste entra en nuestro canal auditivo para ser interpretado por nuestro cerebro [17].

En abril de 1982 en Italia, Hugo Zucarelli de origen argentino, y con la ayuda del profesor Alessandro Mascioli, publica la aplicación, en la patente europea EP0050100, de su "Sistema para la codificación espacial de sonidos", Figura 2.

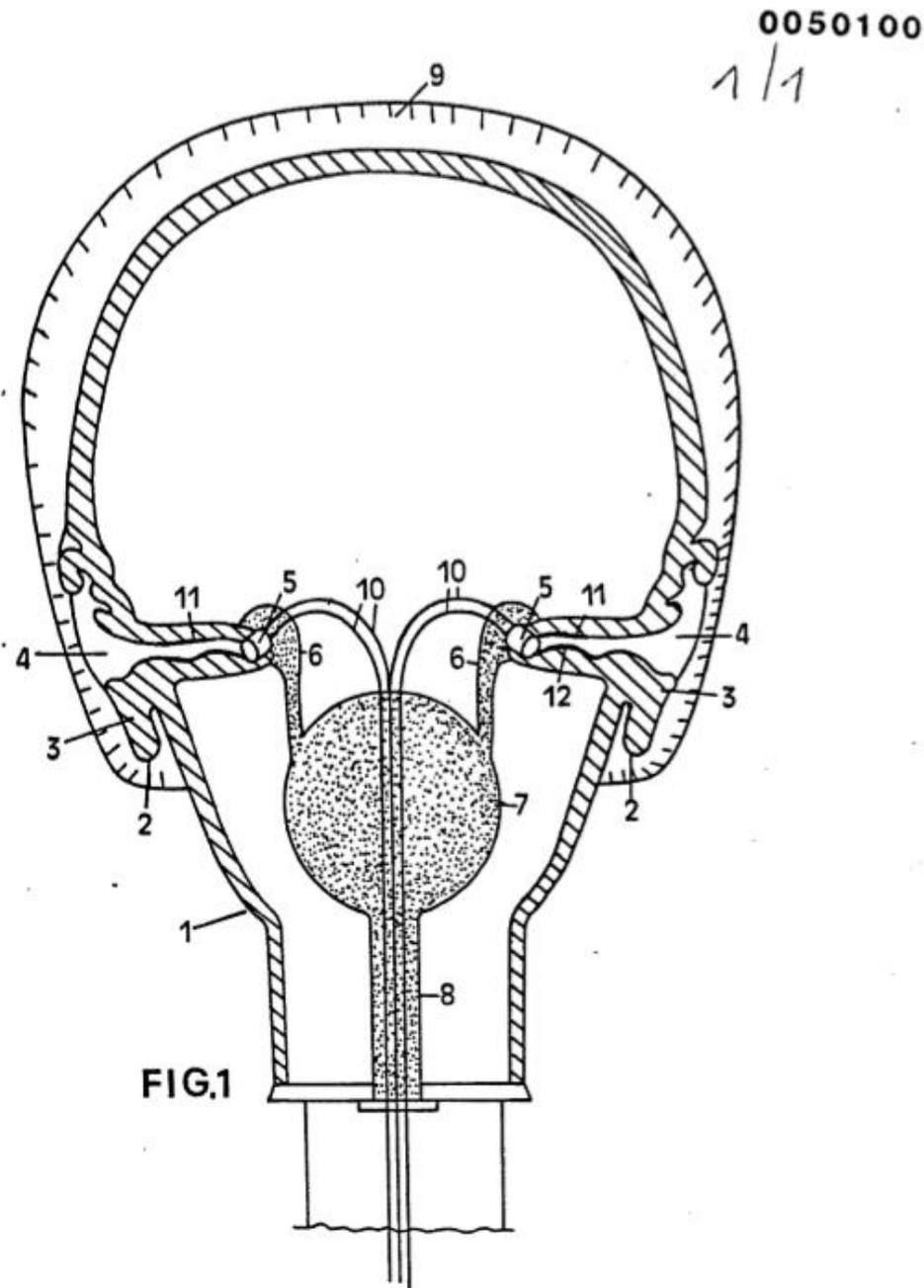


Figura 2 Diseño del Dummy head de Hugo Zuccarelli

Debido a que la patente tiene más de 20 años es de dominio público.

En la actualidad existen micrófonos binaurales, que se pueden colocar dentro de las orejas. No se necesita de la dummy head para simular los obstáculos de la onda ya que nosotros mismos somos los obstáculos naturales [18].

En esta última década se ha venido popularizando la espacialización sonora, ya que se está utilizando en la industria del entretenimiento, como en videojuegos y realidad virtual, ya que crean un ambiente más realista, con las Oculus Rift se están empezando a utilizar.

Teniendo en cuenta que las personas podemos percibir la posición y distancia de los sonidos, las posibilidades para implementarlo en una representación tridimensional del entorno real utilizando estos sonidos son interesantes, ya que con esto se podría apreciar el entorno escuchándolo.

2.3.2. Mapas de profundidad

La reconstrucción de la profundidad de una imagen es el proceso por el que a partir de imágenes bidimensionales se puede obtener información que permita una recreación de la distribución espacial real. En la figura 3 se muestra diferentes tonalidades de colores, cada color representa la distancia a la que se encuentran de la cámara.

A pesar de que actualmente se están dedicando muchos esfuerzos para desarrollar algoritmos de reconstrucción de la profundidad a partir de dos o más imágenes, el análisis de características de una única imagen (como variaciones de la textura o el color) también puede aportar información sobre su profundidad.



Figura 3 Mapa de profundidad representado por colores

Para determinar la profundidad se utiliza la técnica de motion parallax, que describe el desplazamiento de los objetos entre las dos imágenes, este desplazamiento es la disparidad, como se muestra en las figuras 4 y 5. En este ejemplo, el árbol tiene mayor disparidad que la casa y esto quiere decir, que el árbol está más cerca que la casa [19].

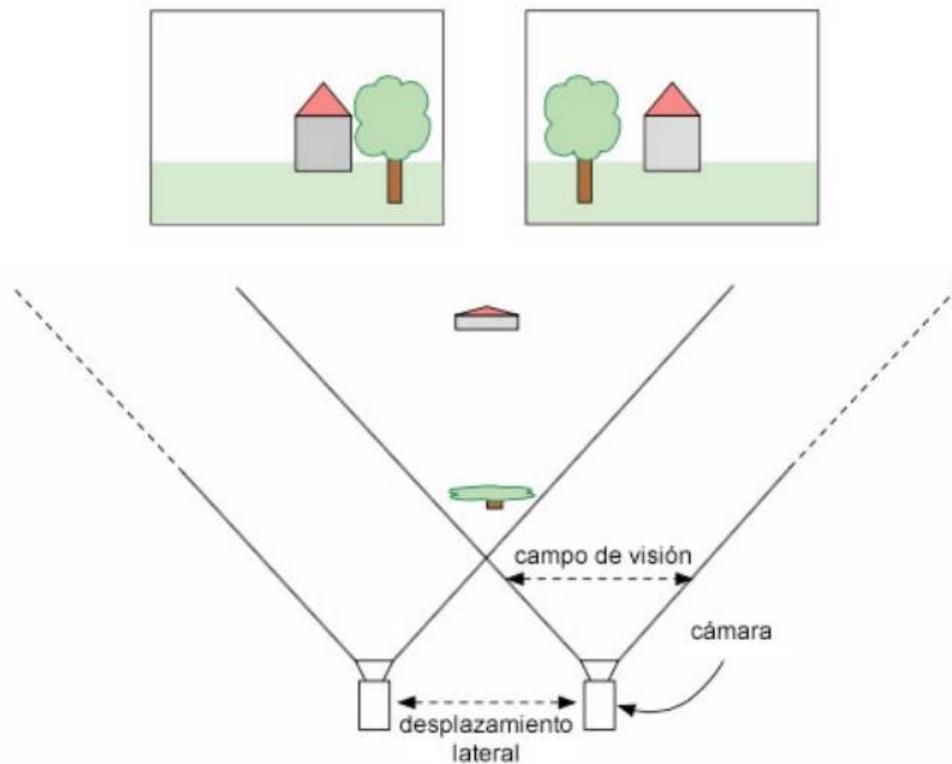


Figura 4 Dos imágenes tomadas con una cámara desplazada

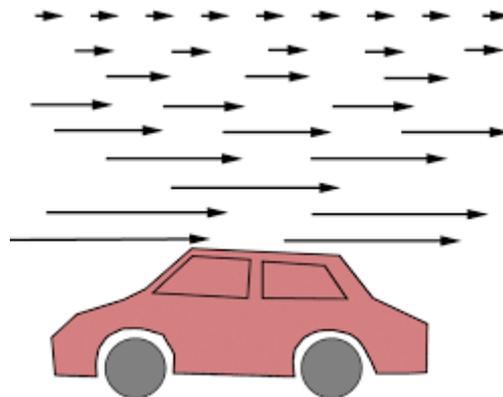


Figura 5 Motion Parallax

El desplazamiento de los objetos en la imagen nos da información de profundidad. La visión binocular, también llamada visión estereoscópica, permite a los humanos percibir el entorno en tres dimensiones, como resultado de la combinación de las imágenes de ambos ojos. La disparidad entre las imágenes es calculada y usada para la reconstrucción de la escena en tres dimensiones.

Esta técnica se utiliza para determinar la profundidad del entorno, la cual se utiliza para sistemas de navegación en robots, drones, coches autónomos, robots lazarillos, entre otros.

En sistemas de visión estereoscópica, la distancia entre las cámaras se conoce, pero en este proyecto se van a utilizar los sensores del móvil (giroscopio y acelerómetro), para determinar la posición de la cámara.

2.3.3. Proyectos dedicados a personas invidentes

Algunos proyectos para personas invidentes están teniendo gran auge en estos últimos años, desde la aparición del Kinect que saca mapas de profundidad, ha sido utilizado para diversos proyectos.

La compañía NSK en 2011 desarrolló un robot cuadrúpedo, con pequeñas ruedas, puede moverse sobre superficies plana y escaleras, utiliza Kinect para reconocer el entorno, figura 6. El robot cuenta con una barandilla para sujetarse la persona, que es regulable, haciendo que el usuario mantenga la posición erguida todo el tiempo, incluso para bajar las escaleras [20].



Figura 6 Perro lazarrillo

Un proyecto para personas ciegas, es The vOICe en combinación con Kinect fue presentado en 2015 por Giles Hamilton-Fletcher y Jaime Ward de la Universidad de Sussex en la BBC, figura 7. The vOICe permite detectar objetos cercanos y sus formas. Está conectado a un ordenador portátil que se encuentra dentro de la mochila que lleva la persona en su espalda [21].



Figura 7 The vOICe

The vOICe para Android convierte las vistas en vivo de la cámara a paisajes sonoros, brindando de esta manera realidad aumentada para ciegos. También incluye detector de colores, brújula y GPS.

The vOICe utiliza tonos graves o agudos para señalar la altura de lo captado y el volumen alto o bajo para señalar la claridad de éste en la cámara, que barre de izquierda a derecha cada segundo.

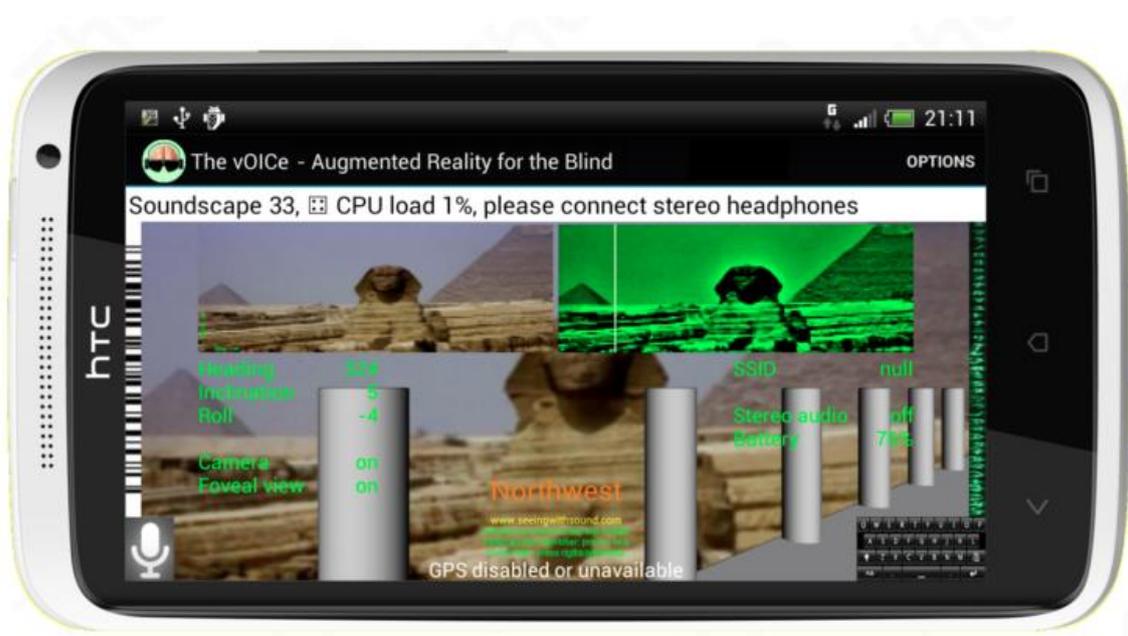


Figura 8 The vOICe para Android

EYEMUSIC capta formas, que los representa de un mismo color y los traduce en sonidos [22]. Los colores se representan mediante diferentes instrumentos musicales, mayor cantidad de píxeles del mismo color en la imagen se traducen en notas más altas, mientras menor cantidad de píxeles en la imagen lo hacen en notas más bajas, los píxeles de la izquierda suenan primero que los píxeles de la derecha, figura 9.

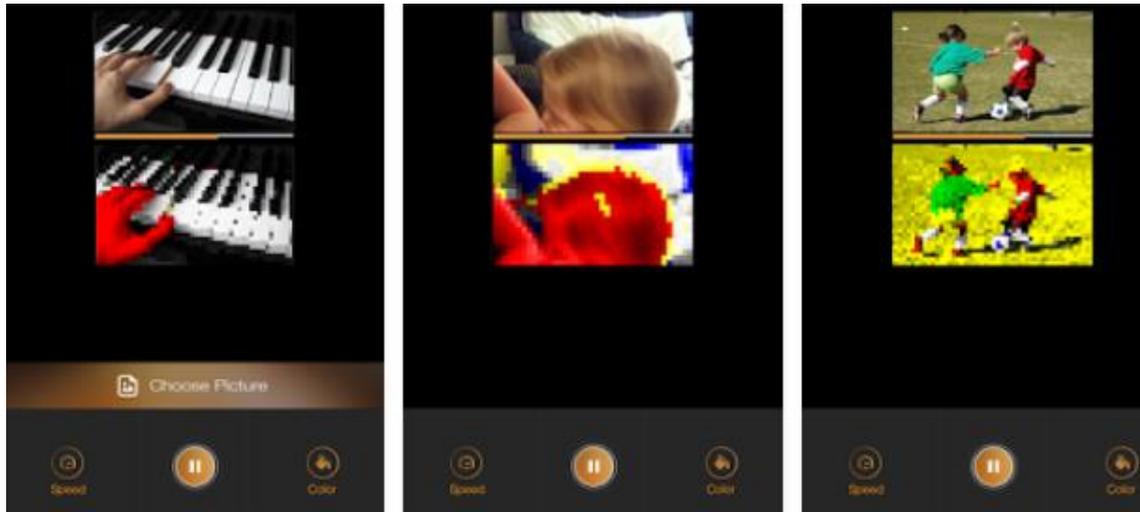


Figura 9 EYEMUSIC

3. Desarrollo de la Aplicación

3.1. Arquitectura del sistema

El sistema propuesto en este trabajo combina: las prestaciones de hardware que ofrecen los dispositivos móviles de última generación, con un motor de video juegos muy potente como es Unity. Además de combinar OpenCV para el análisis y procesamiento de las imágenes, se utiliza RealSpace3D que es un plugin para Unity que permite escuchar sonidos en un espacio tridimensional.

Tras revisar el estado del arte, se ha definido la metodología a seguir. Uno de los primeros pasos, es conseguir un mapa de profundidad a partir de la cámara del móvil, utilizando los datos del acelerómetro y giroscopio para poder determinar la orientación de la cámara y así poder obtener una imagen lo más estabilizada posible.

Finalmente, se generan los sonidos que se van a escuchar dependiendo del mapa de profundidad, los cuales están dados en escala de grises, desde el color más claro (blanco), hasta el color más oscuro (negro), que representarán si está más cerca o lejos respectivamente. Es decir, cuanto más blanco sea el color de la imagen del mapa de profundidad, el sonido que se escuche será más fuerte, mientras que, si el color es más oscuro, el sonido se escuchará más débil. En la figura 10 se puede apreciar la metodología utilizada.



Figura 10 Metodología

3.2. Hardware

Para las pruebas realizadas se utilizó un Smartphone Huawei P7-L10, figura 11, dentro de las especificaciones técnicas tenemos:

- Tamaño
 - Altura 139.8 mm
 - Ancho 68,8 mm
 - Profundidad 6.5 mm
- Peso
 - Sobre 124g
- CPU
 - Hisilicon Kirin 910T
 - 1,8 GHz de cuatro núcleos
- Funcionamiento del sistema
 - Android [™] 5.1.1 lollipop / Huawei Emoción UI 3.1
- Memoria
 - RAM: 2 GB
 - ROM: 16 GB
- Sensores
 - G-sensor
 - Acelerómetro
 - Brújula
- Cámara
 - La cámara trasera: 13MP BSI AF F2.0, con Flash
 - Frente Cámara: 8MP FF
- Audio
 - Formato de archivo de audio: Mp3, mid, amr, 3gp, mp4, m4a, wav, ogg, ra
- Auriculares de 3,5 mm



Figura 11 Huawei P7-L10

Las prestaciones que ofrece este dispositivo son aceptables, pero no ha cumplido con las expectativas en términos de rendimiento, debido a su lentitud en cuanto al procesamiento de las imágenes se refiere.

3.3. Software

3.3.1. Unity

La plataforma elegida para el desarrollo es Unity (<http://unity3d.com>). Unity es un motor de video juegos multiplataforma, permite crear entornos virtuales en 3D. A su vez, ofrece una versión gratuita y una versión profesional de pago. Se ha utilizado la versión gratuita 5.3.5, la cual ofrece todas las prestaciones para poder desarrollar la aplicación.

Unity permite trabajar para distintas plataformas con una sola herramienta. Abstrae la mayoría de las diferencias entre las plataformas y además, se puede tener un control más preciso utilizando el comando `#ifdef`.

Las ventajas de Unity son las siguientes:

- Soporta OpenGL ES (Android)
- Es multiplataforma
- Permite llamar a funciones escritas en C/C++
- Proporciona un API para acceder a datos de entrada del dispositivo Android
- Documentación muy completa en la página de Unity
- Dispone de una licencia gratuita

Para el desarrollo de la aplicación con Unity (para la plataforma Android), además de los requisitos generales del sistema es necesario instalar:

- Kit de desarrollo para Android SDK y Java.
- Los equipos Android donde se ejecute han de tener las siguientes características:
 - Android OS 2.3.1 o posterior.
 - Dispositivo con CPU ARMv7 (familia Cortex).
 - Se recomienda soporte GPU para OpenGL ES 2.0.

Unity está a la vanguardia del creciente mercado de realidad virtual, aproximadamente el 90% de los juegos de Samsung Gear VR y el 53% de Oculus Rift fueron creados con Unity [23]. Cada vez son más los desarrolladores registrados y siguen en aumento, Figura 12.

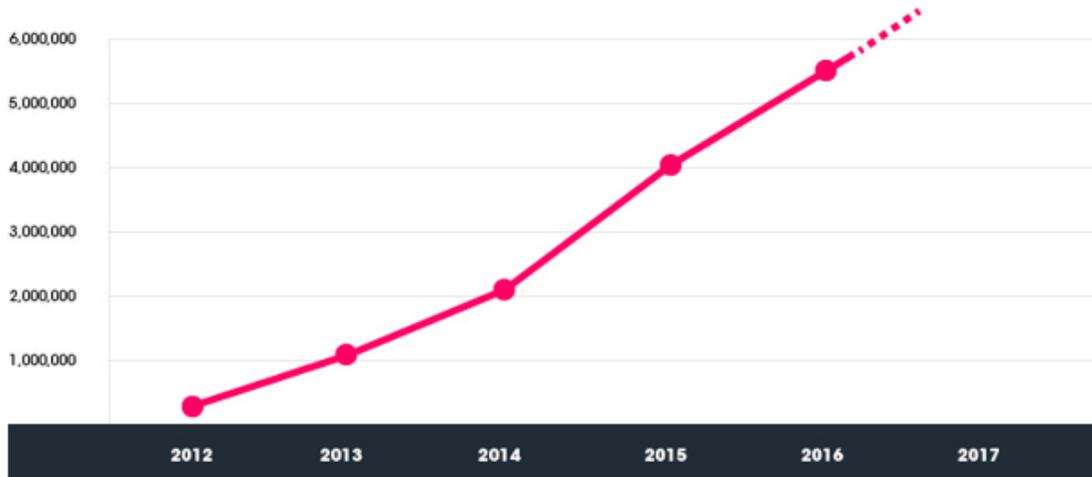


Figura 12 Desarrolladores registrados en Unity desde 2012 hasta 2016

Podemos decir que es el software líder a nivel mundial en la industria de aplicaciones en realidad virtual. Se crean más aplicaciones con Unity que con cualquier otra herramienta [23].

3.3.2. RealSpace3D

RealSpace 3D (<http://realspace3daudio.com/>) ofrece una solución completa para reproducir sonidos binaurales, en videojuegos, realidad virtual y películas. Crea una percepción real de la dirección dónde se encuentra el sonido, distancia y profundidad, creando una completa experiencia de audio.

Los avances en audio 3D, permiten procesar el sonido para que parezca como si la persona estuviera en el mismo lugar real. RealSpace 3D utiliza HRTF genéricos, que funcionan bien, pero no tienen en cuenta las diferentes formas y tamaños de las orejas y cabeza de cada individuo, que también afectan en la percepción del sonido, pero sí tienen una opción manual para configurarlos, figura 13.

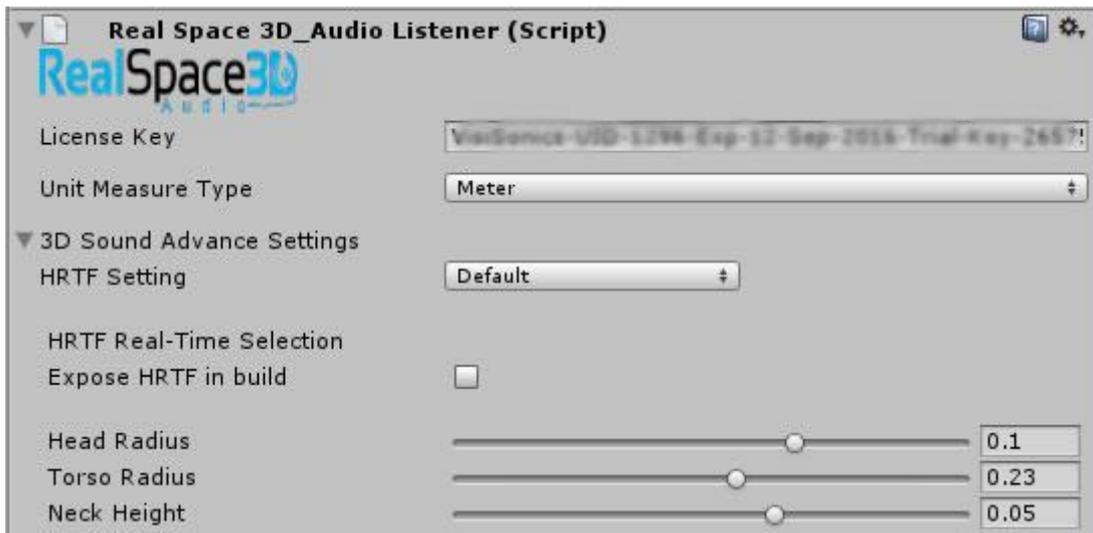


Figura 13 Configuración del sonido en RealSpace3D

RealSpace3D tiene soporte multiplataforma, para Windows, MacOS, Android, IOS, Unity y Unreal Engine. Este plugin se está dando a conocer de forma reciente, gracias al plugin que tienen para Unity, el cual permite hacer aplicaciones para Oculus Rift, para crear ambientes virtuales más realistas utilizando sonidos binaurales.

3.3.3. OpenCV

Para el análisis de las imágenes utilizamos OpenCV. Se trata de una librería de código abierto para visión artificial, donde existen funciones para el procesamiento digital de imágenes. Esta librería permite analizar las imágenes de la cámara, hacer rotaciones, buscar puntos característicos en la imagen, determinar los bordes de la imagen, usar operaciones matemáticas, crear nuestras propias funciones, así con esto tener una variedad de opciones para poder realizar lo que queremos.

La versión que utilizamos para el proyecto es la versión 1.2.3 de OpenCV para Unity (<https://www.assetstore.unity3d.com/en/#!/content/21088>), este plugin utiliza la versión 2.4.11 de OpenCV para java (<http://docs.opencv.org/java/2.4.11/>), es una adaptación de esta librería para que funcione en Unity. Cabe recalcar que este plugin requiere Unity 5.0.0 o superior. Trabaja con Unity y Unity Pro, soporta IOS, Android, Windows, Linux y Mac.

3.3.4. OpenAL Soft

OpenAL Soft (<http://kcat.strangesoft.net/openal.html>) tiene de base OpenAL (<http://www.openal.org/>). Es una API (Application Programming Interface) de audio multiplataforma, para el renderizado eficiente de audio posicional y multicanal en tres dimensiones. Esta librería de sonido tiene soporte y documentación actualmente, se utilizó la versión 1.13 de OpenAL Soft

Se hicieron pruebas para tratar de representar un círculo con los sonidos, la representación se puede ver en la figura 14. Cada punto representa un sonido (un bip) en el espacio tridimensional, así como en una imagen utilizamos muchos píxeles para formar dicha imagen, aquí utilizamos muchos sonidos para representar el círculo.

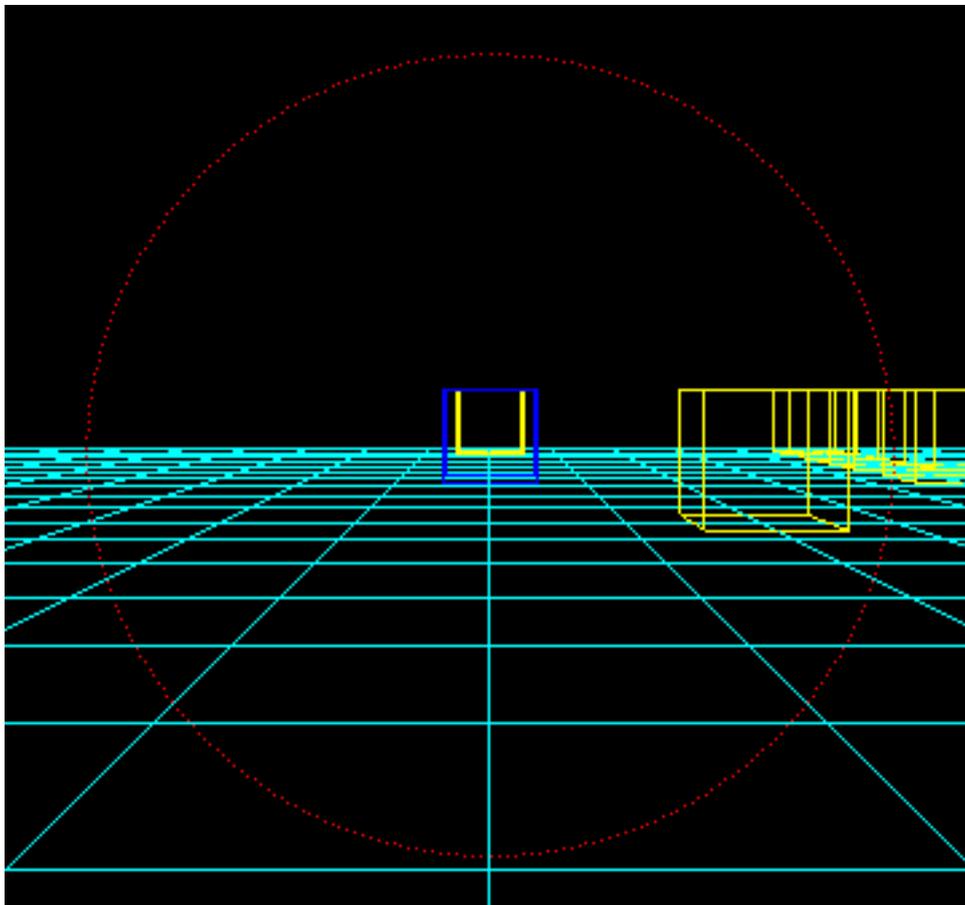


Figura 14 Círculo representado con los sonidos

No se utilizó con Unity porque no existe un plugging, por eso se buscó otra opción que tiene similares resultados como lo es RealSpace3D.

3.4. Implementación

Para la implementación se integró, OpenCV para Unity 3D, este plugging se puede descargar desde la Asset store de Unity y lo utilizamos para el análisis de las imágenes. Además, accedemos a la página web de RealSpace3D, nos registramos, descargamos el plugging para Unity 3D, y solicitamos una clave para que poder utilizarlo, les escribimos un correo y te envían una licencia de prueba, figura 15.



Figura 15 License Key RealSpace3D versión de prueba

Primero se realizaron las pruebas para determinar el mapa de profundidad y para ello se probaron tres métodos:

- Profundidad de campo
- Flujo óptico
- Visión Estereoscópica

Profundidad de campo

Para realizar las pruebas utilizando la técnica de profundidad de campo, se integró Vuforia con Unity 3D, porque tiene una opción para controlar el enfoque de la cámara. Unity 3D y OpenCV no permiten controlar el foco de la cámara, solamente permiten un enfoque automático, pero no manual. El problema que se encontró, es que sólo se puede hacer un enfoque macro a unos 15 cm de distancia, todo lo demás lo deja desenfocado a una misma proporción. Si fuera un desenfocado gradual se podría utilizar esa información para poder determinar la profundidad en la imagen, por eso se descartó esta forma de hacerlo.

Flujo óptico

Se utilizó como base, el ejemplo de Optical Flow de OpenCV y se utilizó el acelerómetro y giroscopio para determinar la orientación. Se definieron las dimensiones de la resolución de la imagen. Se trabajó con 640 x 480 y 320 x 240, para tener un mejor rendimiento. Al realizar las pruebas, para determinar la distancia que varía del flujo, entre las imágenes, cuando la imagen es de resolución baja, no se obtienen muy buenos resultados, y aunque se aprecia la dirección, la distancia que recorren es casi la misma para todos los puntos, y no se puede sacar información precisa, Figura 16.



Figura 16 Flujo óptico a resolución 320 x 240

Cuando se aumenta la resolución a 640 x 480 los fotogramas por segundo bajan a 10, es muy lento, solo para determinar la dirección, por eso también se descartó esta opción [24 al 27].

Visión Estereoscópica

Este método fue con el que mejores resultados se obtuvieron, las imágenes utilizadas para llevar a cabo la prueba tienen una resolución de 640 x 480 y 320 x 240.

Se utilizó el acelerómetro y el giroscopio del móvil, para determinar la orientación y así poder tener las imágenes un poco más alineadas, porque cuando una persona camina, se mueve mucho la cámara, por eso se utilizan los sensores para obtener una mejor estabilización de la imagen.

Lo primero que se hizo fue rectificar la imagen. Para ello se utilizó la función de OpenCV, `findChessboardCorner`, figura 17, porque las lentes de las cámaras generan más o menos distorsión a la imagen. Por ejemplo, un lente ojo de pez genera mucha distorsión a una imagen, figura 18.



Figura 17 Chessboard utilizado para la calibrar de la cámara

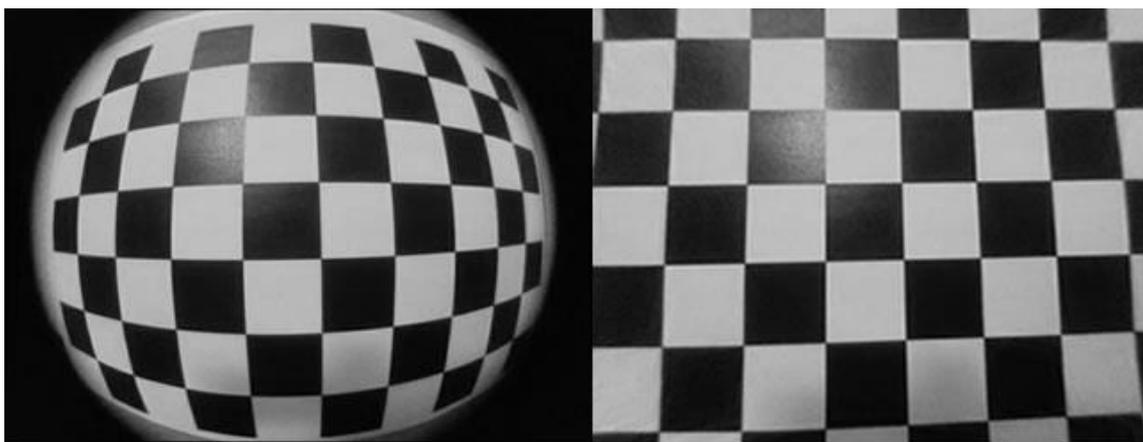


Figura 18 Imagen Izquierda ojo de pez y a la derecha rectificada

La lente que utiliza el móvil Huawei P7-L10 no tiene mucha distorsión, pero igual se aplicó la rectificación a la imagen, como se muestra en la figura 19.



Figura 19 Imagen rectificada Huawei P7-L10

Para realizar el análisis de las imágenes, se guardaron los datos del acelerómetro y giroscopio generados por el móvil, en un archivo de texto y para esto se creó una clase en Unity, `AdquisicionGyA.cs`. Los datos se guardan en el directorio de la aplicación. El problema que surgió fue que no se podía acceder a estos directorios. Por eso es necesario tener acceso root en el móvil, que da permisos para poder acceder a cualquier carpeta del dispositivo. Además, se instaló una app para poder acceder a las carpetas desde el móvil, Root Browser que se descargó de Google Play.

Se agregó la librería `System.IO` para guardar en el móvil el archivo de texto con los datos obtenidos del acelerómetro y giroscopio. Para ello se utilizó la función `Application.persistentDataPath`, para obtener la ruta absoluta de dónde se guarda nuestra aplicación en el móvil. En la figura 20 se muestra en la parte izquierda los datos que se van guardando, se muestran los fotogramas por segundo, los valores del acelerómetro y giroscopio en el segundo 2.



Figura 20 Imagen de un fotograma

Con los datos del acelerómetro se puede estabilizar la imagen, conociendo el ángulo de inclinación del móvil, lo cual, ayuda a tener siempre la imagen horizontal con respecto al suelo. Se muestra un ejemplo en la figura 21. A pesar de no ser muy preciso, ayuda a tener en la imagen a analizar, los objetos casi en la misma posición. Con las líneas verdes paralelas se puede observar la corrección y éstas tienen una resolución de 320 x 240 píxeles.

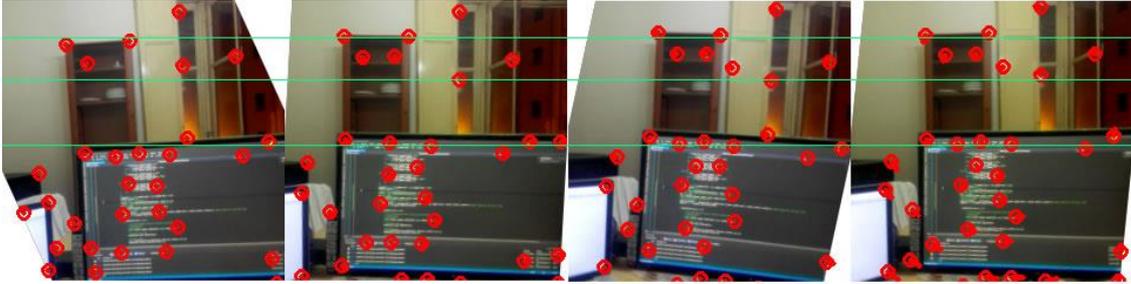


Figura 21 Imágenes estabilizadas

Después de tener las imágenes estabilizadas, se buscan puntos comunes entre las dos imágenes, para poder determinar su matriz fundamental. Encontramos 9 puntos en cada imagen, estos puntos deben ser similares y para encontrarlos se divide la imagen en 24 regiones iguales. Los puntos de la imagen de la izquierda siempre van a tener la misma posición, lo que se hace es buscar la región de la imagen objetivo, en la imagen 2. La Región objetivo es la unión de 4 regiones como se puede observar en el rectángulo de color rojo. De modo que, cuando se haga la búsqueda de la imagen objetivo, en la imagen 2, no se tiene que buscar en toda la imagen, sino en una región más pequeña permitiendo así hacer el cálculo más rápido. En la figura 22 se muestra el diseño de las regiones.

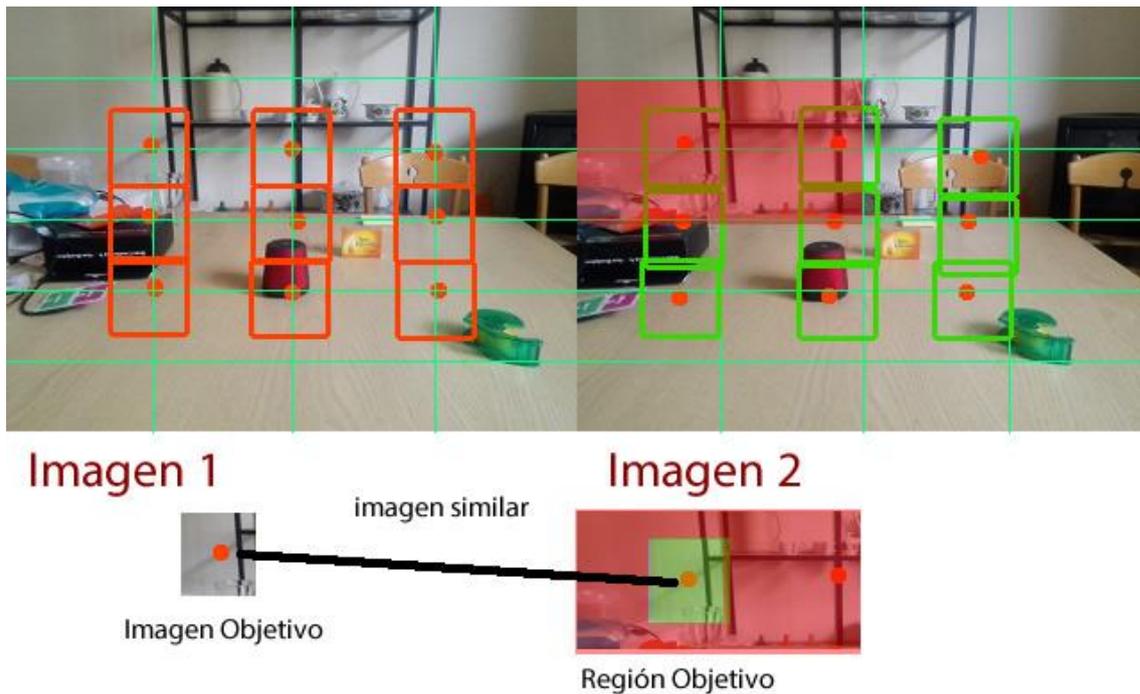


Figura 22 Regiones de la imagen

Para hacer la búsqueda del área de la imagen 1 en la imagen 2, se utiliza la función de OpenCV, `matchTemplate`, que hace una búsqueda de una imagen lo más similar posible, figura 23.

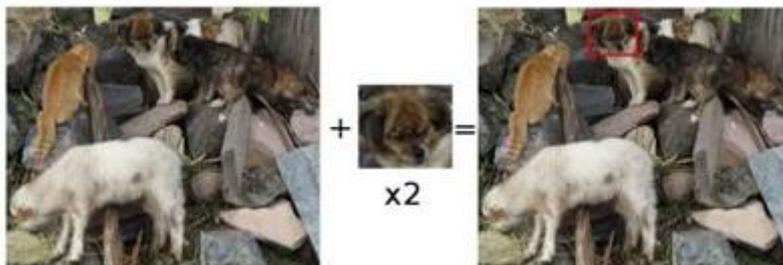


Figura 23 MatchTemplate función de openCV

Se hicieron pruebas utilizando histogramas, para llevar a cabo la comparación de las regiones similares de la imagen, pero había mucho error cuando tenía que encontrar el área. Con los puntos encontrados, se puede calcular la matriz fundamental, para esto utilizamos la función de OpenCV `findFundamentalMat`, que nos devuelve una matriz de 3x3. Con esta matriz fundamental se pueden calcular las líneas epipolares, como se muestra en la figura 24.

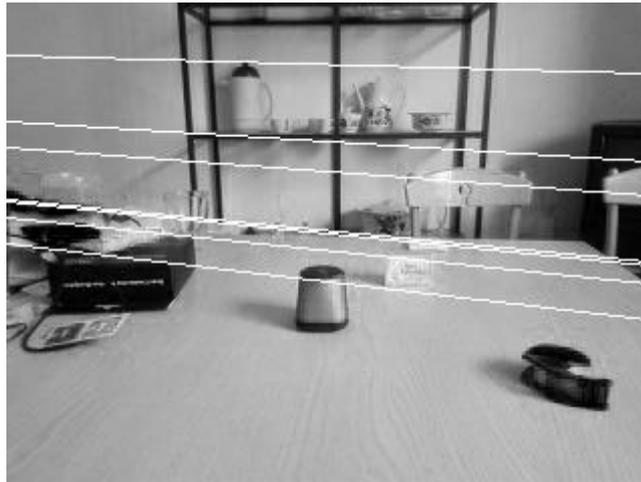


Figura 24 Líneas epipolares

Después de obtener las líneas epipolares de las imágenes, se aplica la transformación de perspectiva, utilizando la función de OpenCV `warpPerspective`, para alinear las imágenes. Se puede observar una corrección para que queden alineadas las dos imágenes, pero no se tiene un resultado aceptable porque cuando se quiere sacar la disparidad entre las dos imágenes no funciona correctamente [28-29], figura 25.

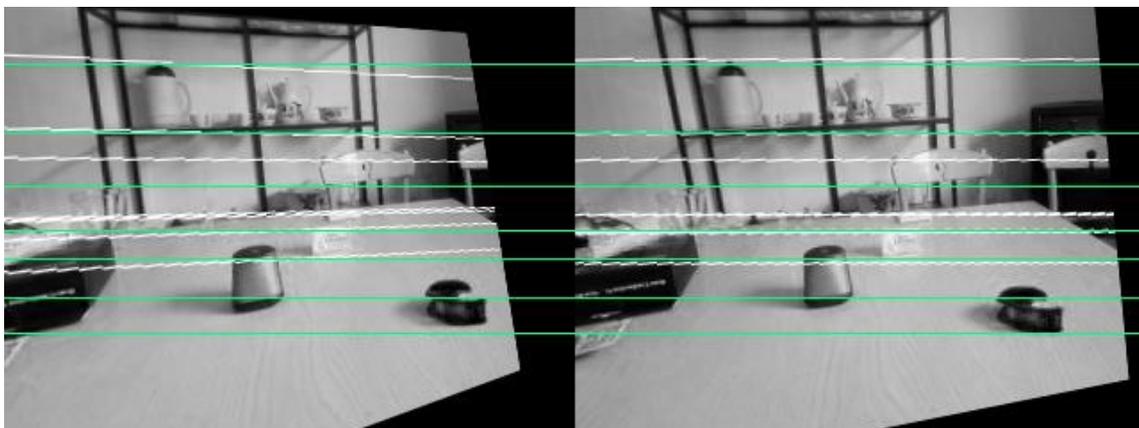


Figura 25 Imágenes alineadas

La imagen debería tener alineados todos los píxeles para poder sacar una disparidad correcta. Lo que analiza el algoritmo, son filas de píxeles, la fila 1 de la imagen 1 con la fila 1 de la imagen 2. En la figura 26 se puede observar un ejemplo de los píxeles que se comparan de la imagen a con la imagen b [30].

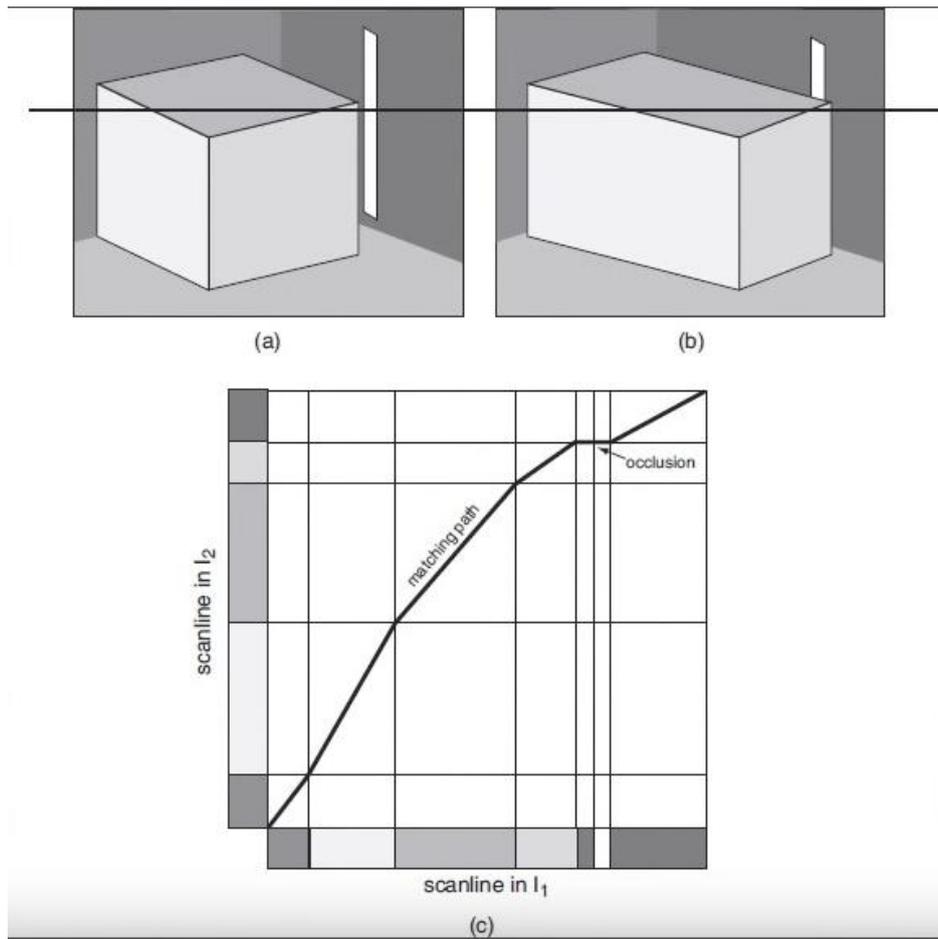


Figura 26 Comparación de píxeles

Como se puede observar en la figura 27 no se puede obtener un mapa de profundidad correcto debido a su imprecisión.

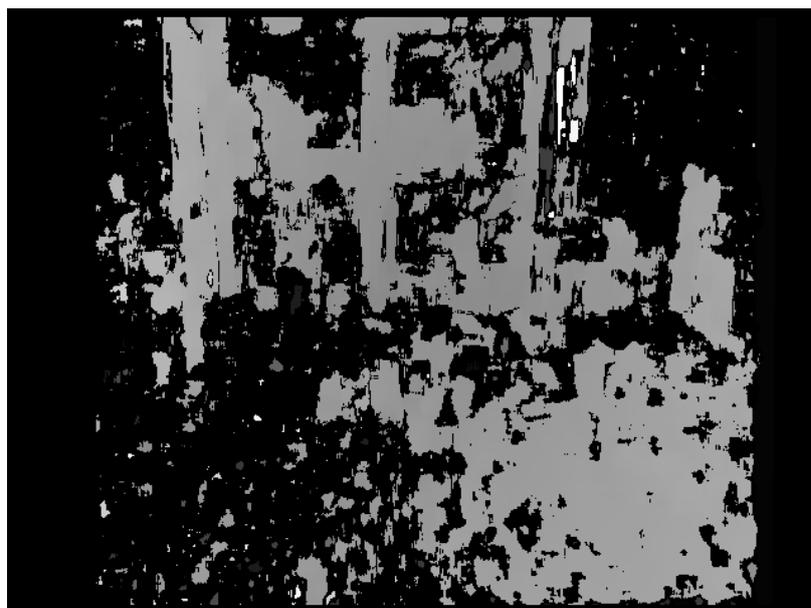


Figura 27 Mapa de profundidad

Reproducción de los sonidos binaurales

Para reproducir los sonidos en Unity utilizamos el plugging de RealSpace3D, con el cual nos permite obtener una percepción de profundidad y distancia.

Los sonidos que se generaron utilizando OpenAL Soft, después de las pruebas de “dibujar” con sonidos una circunferencia, como se muestra en la figura 28, también se hizo una prueba que representa un área en forma circular, figura 29.

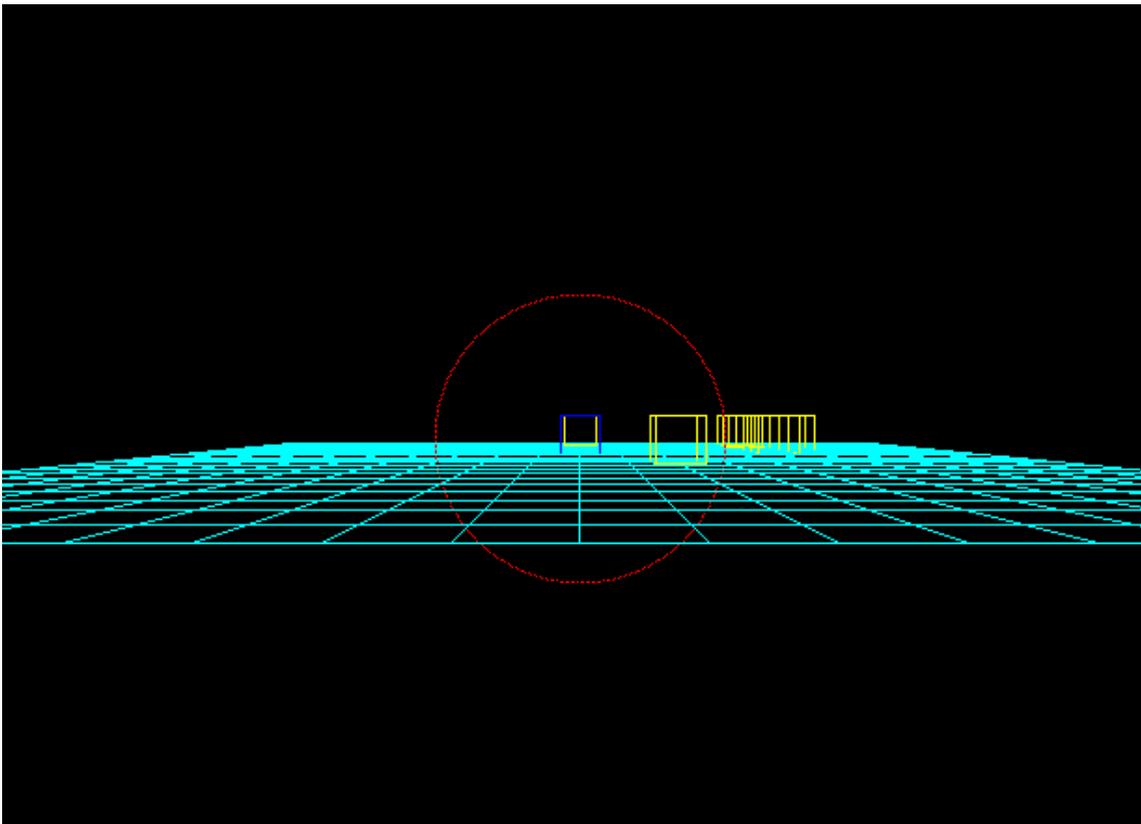


Figura 28 Representación visual de la circunferencia

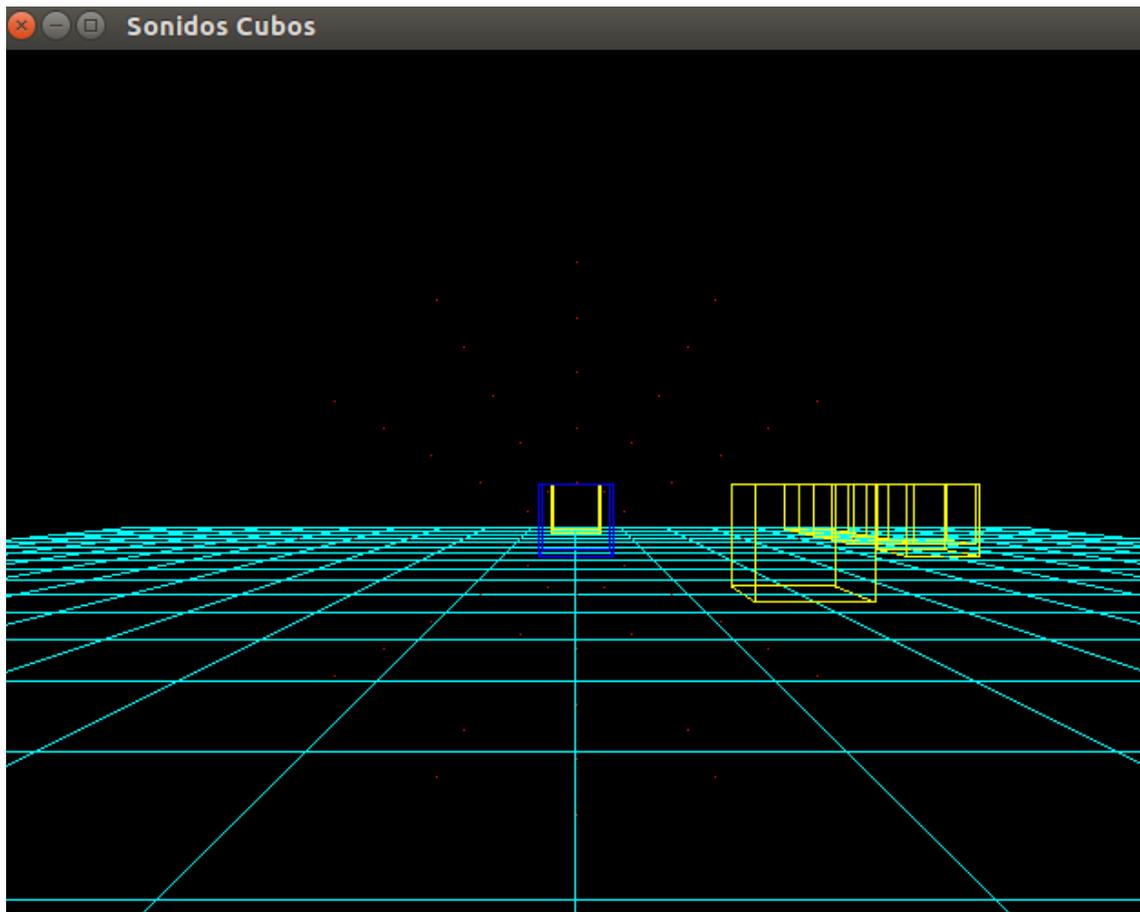


Figura 29 Representación del área circular

Para la demostración con Unity se desarrolló un videojuego en el que se representan 3 cubos. Cada uno de los cuales emite un sonido, según la posición a la que se encuentre el usuario respecto al cubo, el sonido suena más fuerte o más bajo y de este modo se puede determinar si hay un cubo o no frente al usuario. Se puede mover por el escenario con un joystick virtual, que se encuentra en la parte inferior izquierda de la pantalla, figura 30.

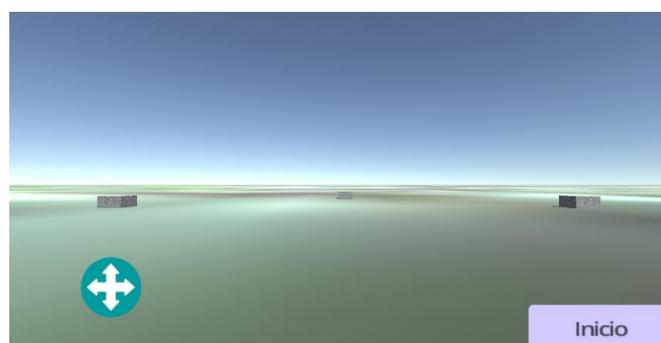


Figura 30 RealSpace3D demo en Unity

Estos sonidos resultan ser tan reales, que al ser escuchados, se percibe la distancia y profundidad a la que se encuentran con mucha precisión; pero esto es relativo, ya que cada persona percibe los sonidos de forma distinta. Es por ello, que existe la necesidad de realizar pruebas con distintas personas, ya, que a su vez, éstas, presentan diferentes grados de sensibilidad auditiva.

3.5. Procedimiento

En este apartado se muestran los datos obtenidos del acelerómetro, además del cálculo para determinar el ángulo de cada uno de sus ejes. En la figura 31 se representan los ejes de rotación de móvil.

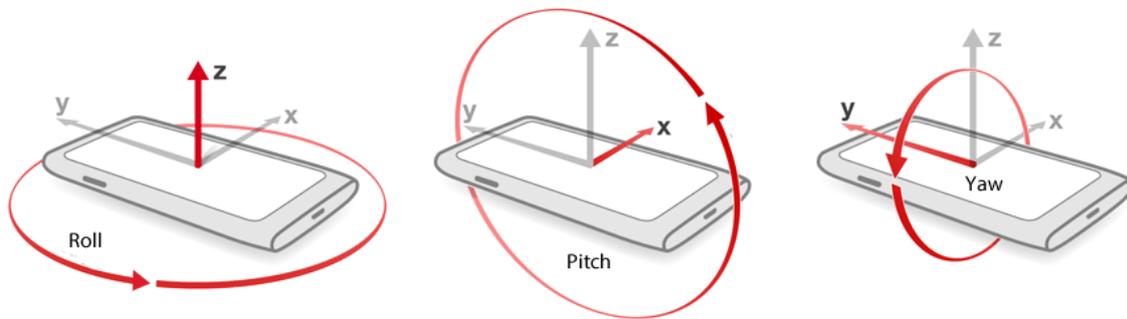


Figura 31 Ejes de rotación del móvil

Para los cálculos de los ángulos se utilizaron las siguientes fórmulas:

$$Roll = \arctan\left(\frac{x}{\sqrt{y^2 + z^2}}\right) \times \frac{180}{\pi}$$

$$Pitch = \arctan\left(\frac{-y}{z}\right) \times \frac{180}{\pi}$$

$$Yaw = \arctan\left(\frac{y}{\sqrt{x^2 + z^2}}\right) \times \frac{180}{\pi}$$

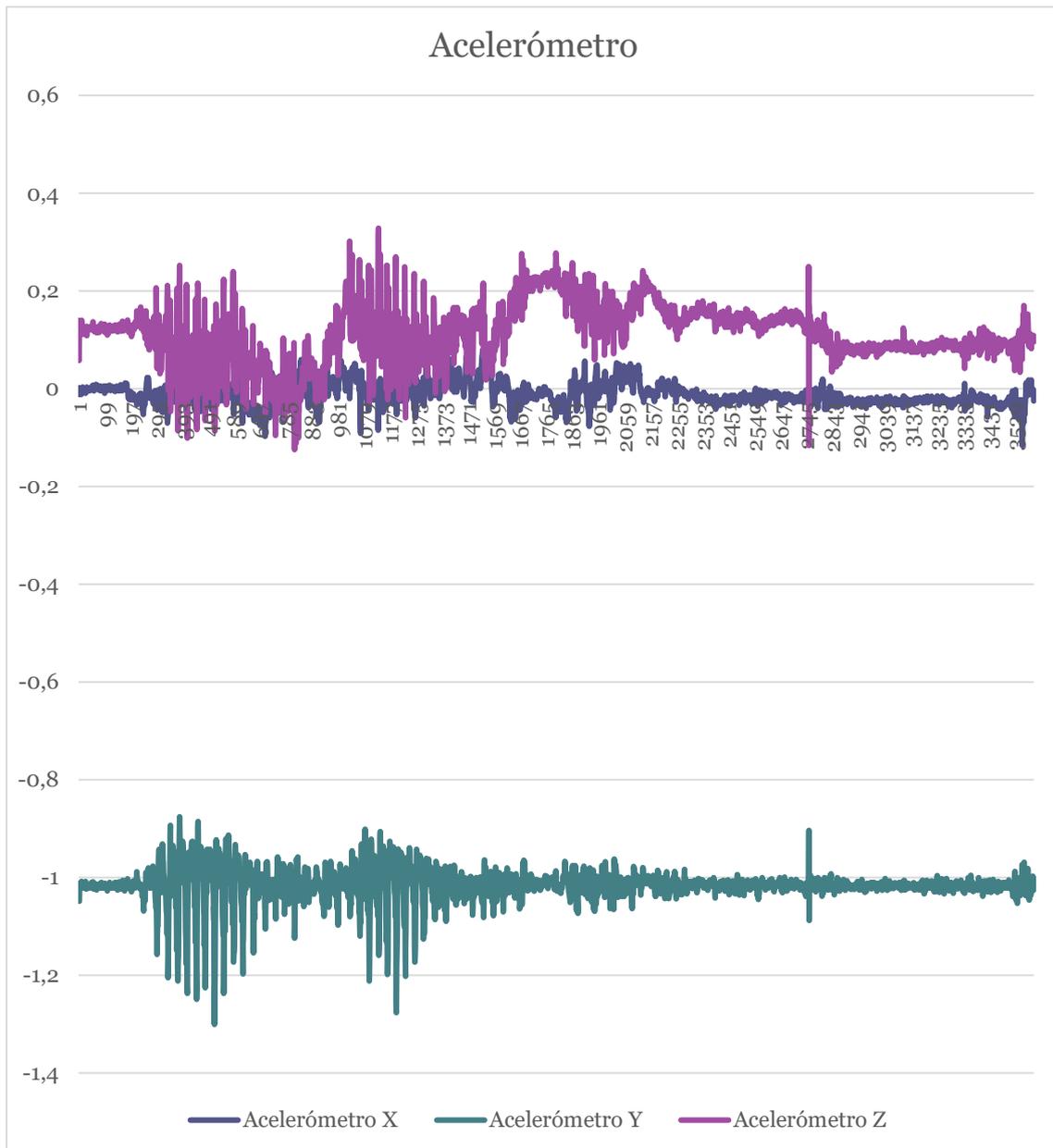


Figura 32 Datos obtenidos del acelerómetro

La figura 32 (datos del acelerómetro) representa las variaciones de los 3 ejes. En el eje horizontal se representa el número de muestras tomadas, a razón de 60 fotogramas por segundo. El eje vertical muestra el valor obtenido del acelerómetro.

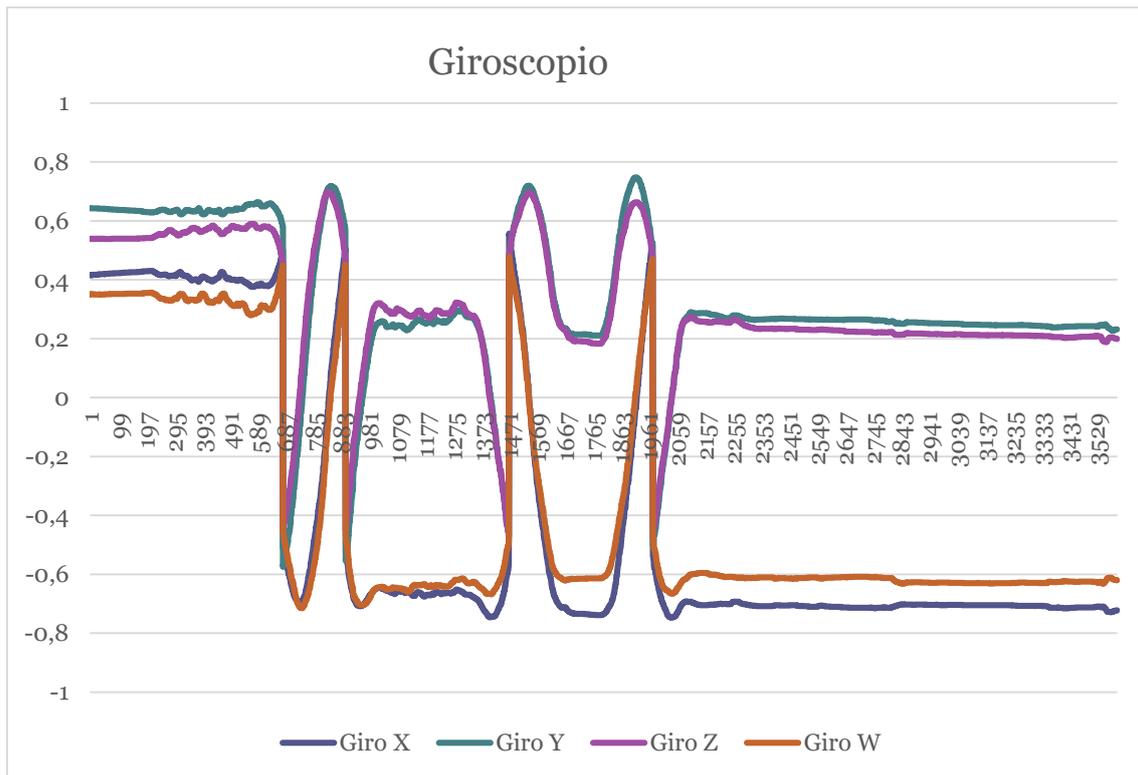


Figura 33 Datos obtenidos del giroscopio

La figura 33 (datos obtenidos del giroscopio) representa la orientación. En el eje horizontal se representa el número de muestras tomadas, a razón de 60 fotogramas por segundo. El eje vertical muestra el valor obtenido del giroscopio.

En la figura 34 se muestra el flujo de las imágenes captadas en el transcurso de la adquisición de los datos del giroscopio y acelerómetro. Se puede observar que:

- Fotograma 189 muestra orientación Norte
- Fotograma 641 sigue con la misma orientación Norte
- Fotograma 731 ha girado a su derecha mostrando orientación Este
- Fotograma 775 ha girado a su derecha mostrando orientación Sur
- Fotograma 827 ha girado a su derecha mostrando orientación Oeste
- Fotograma 883 ha girado a su derecha mostrando orientación Norte
- Fotograma 930 ha girado a su derecha mostrando orientación Este
- Fotograma 1002 ha girado a su derecha mostrando orientación Sur
- Fotograma 1321 sigue con la misma orientación Sur
- Fotograma 1452 ha girado a su izquierda mostrando orientación Este
- Fotograma 1515 ha girado a su izquierda mostrando orientación Norte
- Fotograma 1573 ha girado a su izquierda mostrando orientación Oeste
- Fotograma 1639 ha girado a su izquierda mostrando orientación Sur
- Fotograma 1894 ha girado a su derecha mostrando orientación Oeste
- Fotograma 1952 ha girado a su derecha mostrando orientación Norte
- Fotograma 2121 ha girado a su derecha mostrando orientación Sur

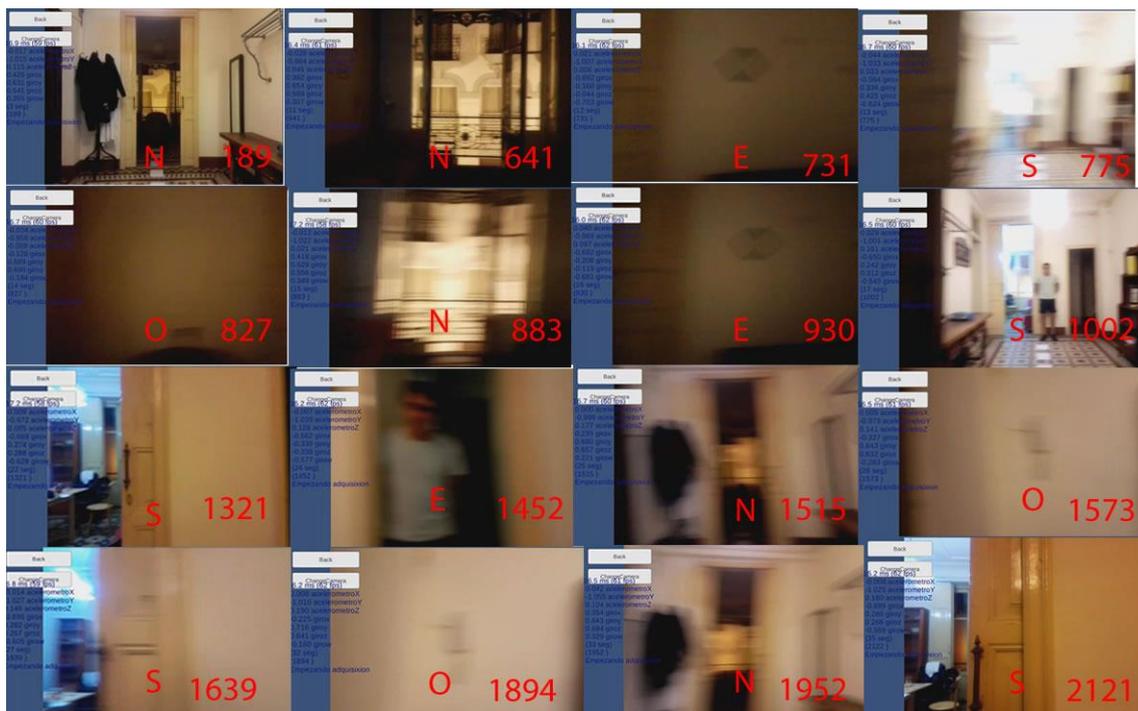


Figura 34 Flujo de las imágenes mostrando la orientación

3.6. Validaciones

Las validaciones se realizaron para determinar si los sonidos que se escuchaban representan las formas que se dibujaban. Los sonidos que se representaron fueron: circunferencia, línea vertical, línea horizontal y punto, figuras 35, 36, 37 y 38, respectivamente.

Las figuras se representaron utilizando 15 puntos (bips de sonido) de resolución para cada una, en la circunferencia, línea vertical, línea horizontal y punto. Los sonidos se reproducían uno después de otro, con un intervalo de 8 milisegundos entre sonido, cada bip o sonido dura 47 milisegundos.

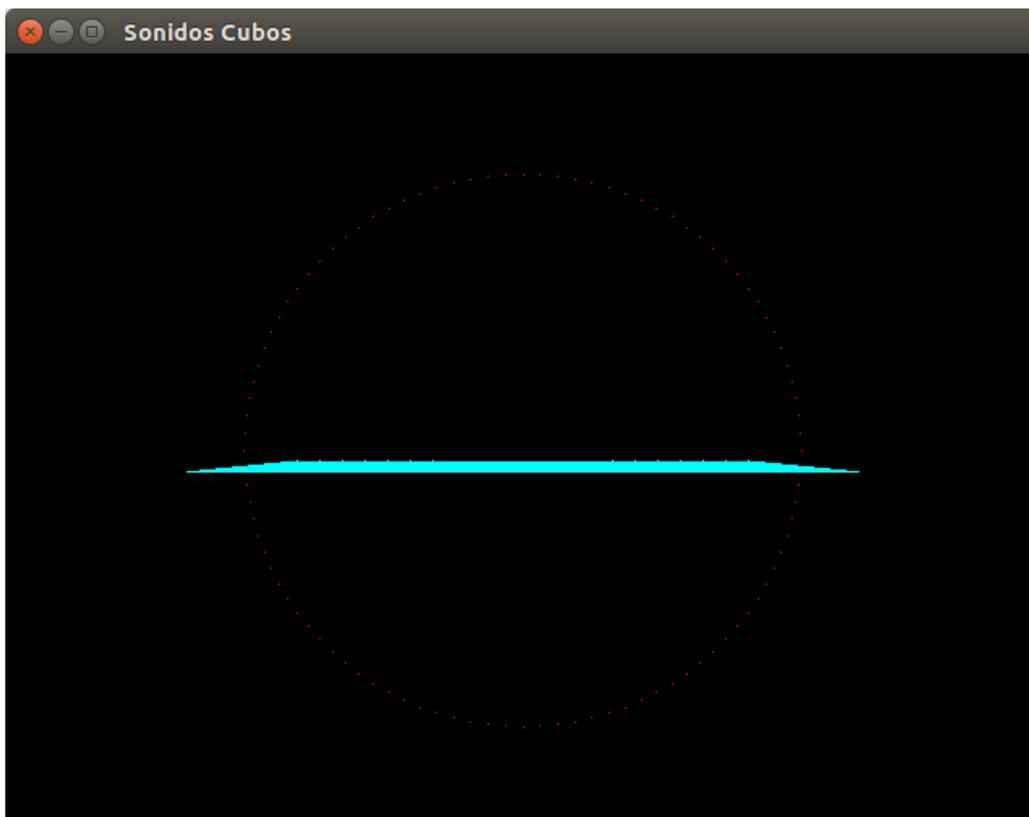


Figura 35 Circunferencia de puntos representada con los sonidos

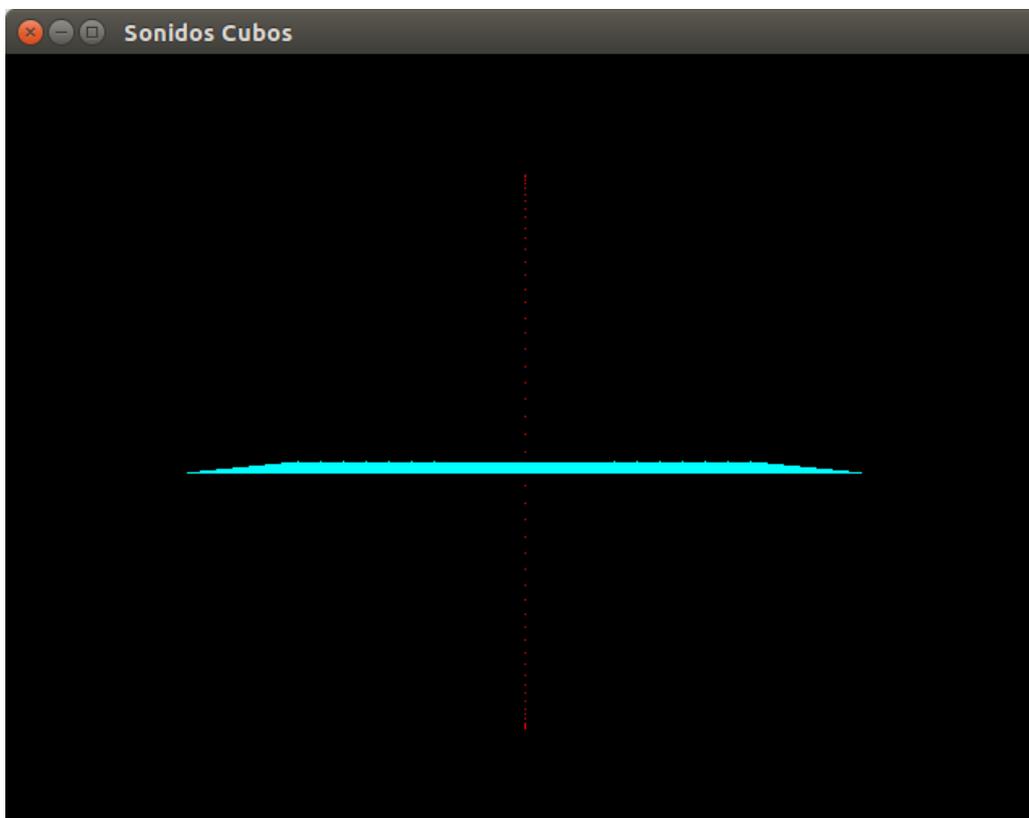


Figura 36 Línea de puntos vertical representada con los sonidos

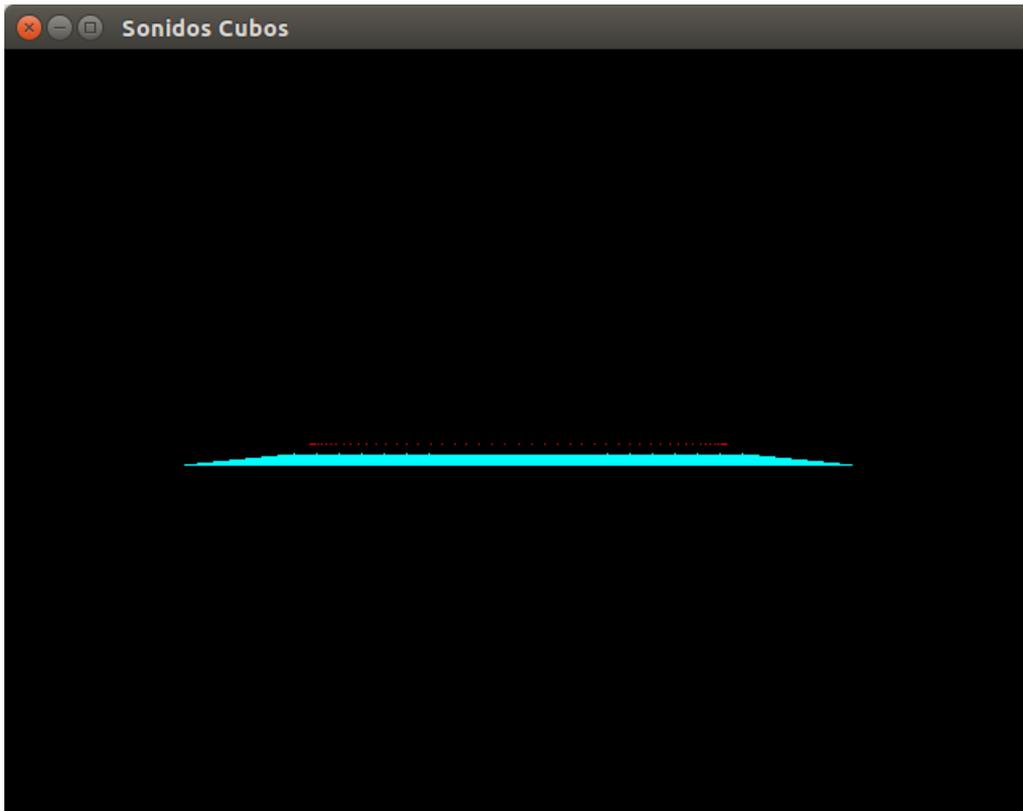


Figura 37 Línea de puntos horizontal representada con los sonidos

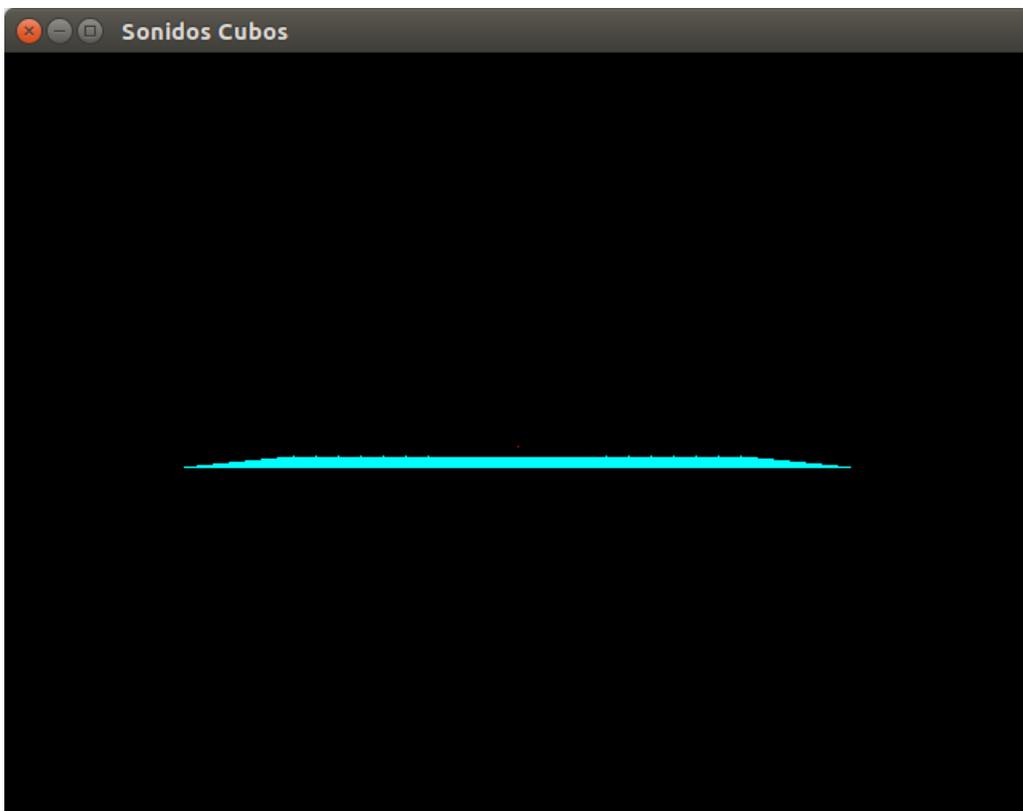


Figura 38 Punto representado con los sonidos

La muestra está compuesta por personas de entre 20 y 42 años. Las personas que hicieron la prueba no tienen problema visual. Se ha tenido en cuenta el género y la edad. Los datos almacenados son anónimos. Concretamente las personas que participaron son 5.

Se realizaron 3 pruebas diferentes:

1.- Si percibía alguna diferencia entre los sonidos, cuántos sonidos parecidos escuchaba y cuáles eran los sonidos. Se reprodujeron 4 sonidos diferentes, uno a la vez, con un silencio de 3 segundos entre cada sonido.

2.- Si percibía la forma de lo que estaban escuchando, en caso de ser no, qué forma percibían.

3.- Se reprodujeron 8 sonidos aleatorios y se debía reconocer qué sonido estaba escuchando.

El cuestionario de las pruebas realizadas se encuentra en la sección 6 de anexos.

Resultados prueba 1

Como se puede observar en la tabla 1, todos los participantes pudieron percibir que los sonidos eran diferentes, pero algunos sonidos se escuchaban similares. La mayoría de los participantes reconocieron dos sonidos que se parecían.

Tabla 1 Información de los participantes y su percepción de los sonidos

| id | Edad | Género (M/F) | Percibe diferencia entre los diferentes sonidos (Sí/No) | ¿Cuántos sonidos parecidos? |
|----|------|--------------|---|-----------------------------|
| 1 | 32 | M | Sí | 1 |
| 2 | 26 | F | Sí | 2 |
| 3 | 26 | M | Sí | 2 |
| 4 | 20 | F | Sí | 2 |
| 5 | 42 | M | Sí | 2 |

En la tabla 2 se muestran los sonidos que se parecen. En la primera prueba los participantes solamente escuchaban los sonidos sin saber cuáles eran, pero decían que se parecía el primer sonido con el segundo, el tercero con el cuarto, por ejemplo.

Los sonidos reproducidos fueron en el siguiente orden: punto, línea vertical, línea horizontal y circunferencia.

Tabla 2 Sonidos que se escuchan similar

| | L. VERTICAL | L.HORIZONTAL | CIRCUNFERENCIA |
|-------|-------------|--------------|----------------|
| PUNTO | 60% | 0 | 80% |
| | L. VERTICAL | 0 | 0 |
| | | L.HORIZONTAL | 40% |

Se puede observar que el sonido del punto se parece al sonido de la línea vertical y al sonido de la circunferencia. También se puede apreciar que la línea horizontal se parece a la circunferencia, en un menor porcentaje.

Resultados prueba 2

En esta prueba se analizó si los sonidos que escuchaban, lo podían relacionar con la forma de punto, línea o circunferencia. En la tabla 3 se observa que el 100% de los participantes pudieron relacionar el sonido que escuchaban con un punto. El 80% de los participantes también reconoció que la línea horizontal tenía una relación con lo que escuchaban.

Tabla 3 Relación de la forma con el sonido

| | RELACIÓN CON EL SONIDO | | ¿Cómo se percibe? |
|----------------|------------------------|------|--|
| | SÍ | NO | |
| PUNTO | 100% | 0 | Muchos puntos juntos |
| L. VERTICAL | 0 | 100% | Similar al punto, punto y rayas |
| L. HORIZONTAL | 80% | 20% | No nota asociación, sonidos a los lados |
| CIRCUNFERENCIA | 0 | 100% | Sonidos a los lados, sonidos dispersos, similar al punto |

También se puede observar que el sonido de la línea vertical no tiene relación con lo que escuchaban y más bien lo asociaban al punto, porque es parecido. La circunferencia no tenía relación con lo que escuchaban, porque percibían como sonidos dispersos o similar al punto, pero un poco más grande.

Resultados prueba 3

En esta prueba se reprodujeron sonidos en diferente orden, para que los participantes pudieran reconocer a qué sonido se referían, de acuerdo con los sonidos que habían escuchado anteriormente.

En la tabla 3 se resalta con color verde el sonido que se reprodujo y los aciertos que obtuvo. La persona podía escoger una o dos opciones. Si la persona escogía dos, se asignaba 0,5 a cada uno de ellos, a menos que la persona decidiera que le daba más valor a una de las dos opciones que había escogido, asignando 0,75 a la que más se parecía y 0,25 a la que menos se parecía.

Tabla 4 Sonidos que la persona tenía que reconocer

| | PUNTO | LÍNEA VERTICAL | LÍNEA HORIZONTAL | CIRCUNFERENCIA | % ACIERTO |
|----------|-------|----------------|------------------|----------------|-----------|
| SONIDO 1 | 4 | 1 | | | 80% |
| SONIDO 2 | | | 0,75 | 4,25 | 85% |
| SONIDO 3 | 4 | 1 | | | 80% |
| SONIDO 4 | 1 | 3 | 0,5 | 0,5 | 60% |
| SONIDO 5 | | | 3,5 | 1,5 | 70% |
| SONIDO 6 | 3,5 | 1,5 | | | 70% |
| SONIDO 7 | | | 4 | 1 | 80% |
| SONIDO 8 | 1 | 1 | | 4 | 80% |

Un dato que algunos participantes mencionaron, es que a pesar de que algunos sonidos no sonaban como la forma que se proponía (línea vertical y circunferencia), los participantes lo asociaban bien, porque aprendieron que ese sonido lo relacionaban a esa forma.

4. Conclusiones y trabajos futuros

4.1. Conclusiones

Se han desarrollado diferentes aplicaciones para probar la metodología planteada. Para ello, se utilizó el dispositivo móvil Huawei P7 para poder obtener un mapa de profundidad del entorno utilizando su cámara, giroscopio y acelerómetro. A pesar de que no se llegó al resultado esperado, se puede concluir que la mejor opción para sacar profundidad es utilizar estéreo visión. Los nuevos móviles están incorporando doble cámara, así que se podrían utilizar para sacar los mapas de profundidad. Si se descalibran las cámaras, después de un tiempo, se pueden volver a calibrar, sin necesidad de requerir un patrón (tablero de ajedrez). Utilizando características naturales de la imagen se podrían recalibrar las cámaras de nuevo.

Para reproducir los sonidos en la aplicación, se modeló un escenario para moverse en un entorno virtual con unos cubos como obstáculos para poder determinar la distancia y dirección a la que se encuentran. Se utilizaron unos auriculares óseos, auriculares intrauditivos (in-ear) y auriculares de diadema. Los auriculares óseos son con los que mejor se escuchaba. Sin embargo, cuando nos encontramos en un lugar muy ruidoso, concretamente en un lugar cerrado con muchas personas, los auriculares (in-ear) son mejores.

Para concluir, después de realizar pruebas en las diferentes etapas del proyecto, el potencial del proyecto es considerable, con los nuevos móviles que tienen 2 cámaras y en un futuro cercano posiblemente más de 2, se podrán obtener mapas de profundidad casi perfectos.

Generar los sonidos en tiempo real no es la mejor opción, ya que, deberían estar grabados para reproducirlos posteriormente.

La capacidad de procesamiento de los móviles cada vez es mayor, haciendo todo esto posible en tiempo real. Por lo que la propuesta es factible en un futuro cercano.

Para este proyecto, se han realizado 3 pruebas con 5 usuarios. Las conclusiones que se han extraído de las mismas se pueden resumir como sigue. Observando los resultados de las tablas 1 y 2, se puede inferir que los sonidos que se perciben son diferentes, pero se pueden parecer, como la línea vertical y la circunferencia que se parecen al punto.

Con los resultados de la prueba 2, se puede inferir que el sonido del punto y la línea horizontal, la pudieron asociar a la forma del mismo. Para la línea vertical y la circunferencia no ha sido así, se perciben de una manera diferente.

En los resultados de la prueba 3 se reprodujeron 8 sonidos en distinto orden. Los sonidos podían ser: punto, línea vertical, línea horizontal o circunferencia. Los porcentajes de acierto fueron buenos, mayor al 60%. Los participantes asociaban bien, a pesar de que algunos sonidos no sonaban como la forma que se proponía (línea vertical y circunferencia), porque aprendieron que ese sonido lo relacionaban a esa forma.

Sin embargo, considerando el tamaño de la muestra, estos datos solo los podemos considerar orientativos. Se debería hacer una validación más exhaustiva para poder extraer resultados concluyentes. Con estas validaciones se podría determinar cómo se perciben estos sonidos binaurales. Comentar que en los videojuegos se están probando y tienen muy buena aceptación, ya que el sonido es más realista. En los juegos en primera persona, si se usa un casco de realidad virtual como las Oculus Rift, la sensación de inmersión es mayor.

4.2. Trabajos futuros

Algunos trabajos futuros podrían ser los siguientes:

- Realizar las pruebas en un Huawei P9 para sacar los mapas de profundidad utilizando estéreo visión, tiene 2 cámaras y un procesador más rápido.
- Generar los sonidos holofónicos, y guardarlos en un formato de audio, para luego reproducirlos, esto le proporcionaría mayor eficiencia, al tener menor procesamiento.
- Utilizar los mapas de profundidad para poder generar entornos virtuales y poderlos utilizar en aplicaciones de realidad aumentada, como por ejemplo escanear una o varias habitaciones y que pueda ser un nivel del juego.
- Desarrollar una versión para usarla con Cardboard de Google. Se podrían hacer estudios comparativos entre estas versiones.
- Hacer pruebas con personas invidentes y personas con visión normal, así tener una comparativa entre ellos y ver si hay diferencias de cómo escuchan.
- Desarrollar la aplicación para que funcione en tiempo real, para tener una mejor perspectiva del entorno.
- Desarrollar una aplicación para escanear objetos en 3D utilizando el móvil, utilizando la técnica de fotogrametría.

5. Bibliografía

- [1] Organización mundial de la salud. (2014). Ceguera y discapacidad visual. <http://www.who.int/mediacentre/factsheets/fs282/es/>, último acceso:08-09-2016
- [2] Nye, P.W., Bliss, J.C., (1970). Sensory aids for the blind: A challenging problem with lessons for the future. *Proceedings of the IEEE*. 58 (12): 1878-1898. DOI: 10.1109/PROC.1970.8061
- [3] Capp, M., Picton, P.D. (2000). The optophone: An electronic blind aid. *Engineering Science and Education Journal*. 9 (3): 137-143. DOI:10.1049/esej:20000306
- [4] Dakopoulos, D., Bourbakis, N.G. (2010). Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*. 40 (1): 25-35. DOI:10.1109/TSMCC.2009.2021255
- [5] Dunai, L., Peris, G., Santiago, V. (2011). Detección de obstáculos mediante sonidos acústicos virtuales. IV International Congress on Design, Research Networks, and Technology for all. Libro de actas: 133-139. <http://www.discapnet.es/Castellano/areastematicas/tecnologia/DRT4ALL/ES/DRT4ALL2011/Documents/Libro-de-actas-congreso-DRT4all-2011.pdf>, último acceso:08-09-2016
- [6] Kumar, A., Patra, R., Manjunatha, M., Mukhopadhyay, J., Majumdar, A. (2011). An electronic travel aid for navigation of visually impaired persons. https://www.researchgate.net/publication/224219996_An_electronic_travel_aid_for_navigation_of_visually_impaired_persons, último acceso:08-09-2016 DOI:10.1109/COMSNETS.2011.5716517
- [7] Mann, R. (1970). Mobility aids for the blind-An argument for a computer-based, man-device-environment, interactive, simulation system. *Proceedings of Conference on Evaluation of Mobility Aids for the Blind*. 101–116. <https://www.nap.edu/read/20595/chapter/11>, último acceso:08-09-2016
- [8] Loomis, J., Golledge, R., Klatzky, R. (2011). GPS-based navigation systems for the blind. In W. Barfield & T. Caudell Edss. *Fundamentals of wearable computers and augmented reality*. Mahwah. Ed. CRC Press. 1: 429-446.
- [9] Kuc, R. (2002). Binaural sonar electronic travel aid provides vibrotactile cues for landmark, reflector motion and surface texture classification. *IEEE Transactions on Biomedical Engineering*. 49(10): 1173-1180. DOI:10.1109/TBME.2002.803561
- [10] Farmer, L., Smith, D. (1997). Adaptive Technology. Book chapter 8, *Foundations on Orientation and Mobility. Instructional Strategies and Practical Applications*. 2 (3): 231-260.

- [11] Ando, B. (2003). Electronic sensory systems for the visually impaired. *Instrumentation and Measurement Magazine*. Proceeding IEEE Xplore. 6 (2):62-67. DOI:10.1109/MIM.2003.1200287
- [12] Petrie, H., Johnson, V., Strothotte, T., Raab, A., Michel, R., Reichert, L., Schalt, A. (1997). MoBIC: An aid to increase the independent mobility of blind travellers. *British Journal of Visual Impairment*. 15 (2): 63-66. DOI:10.1177/026461969701500205
- [13] Dakopoulos, D., Bourbakis, N. (2010). Wearable Obstacle Avoidance Electronic Travel Aids for Blind: A Survey. *Proceeding IEEE Eplore, IEEE Transactions on Systems, Man, and Cybernetics, Part C* 40 (1): 25-35. DOI:10.1109/TSMCC.2009.2021255
- [14] Zuccarelli, H. (1982). Device for the spatial codification of sounds. EUROPEAN PATENT OFFICE.
<https://worldwide.espacenet.com/publicationDetails/originalDocument?CC=EP&NR=0050100&KC=&FT=E>, último acceso:08-09-2016
- [15] Zuccarelli, H. (1982). ACOUSTIC INTEGRITY. Web oficial de Zuccarelli.
<http://www.acousticintegrity.com/acousticintegrity/Holophonics.html>, último acceso:08-09-2016
- [16] Zuccarelli, H. (1983). Tomorrow's sound is a blast from the past. *New Scientist* 98 (1352): 24.
https://books.google.es/books?id=bXxl8lC4cEwC&pg=PA24&lpg=PA24&dq=Tomorrow's+sound%C2%B4is+a+blast+from+the+past&source=bl&ots=cb-ADjT73b&sig=FUeoca2__jqTKuSSZWfIZ5OYdao&hl=es&sa=X&ved=0ahUKEwjM0oeKoYPPAhUCXRoKHbsCA84Q6AEIJAA#v=onepage&q=Tomorrow's%20sound%C2%B4is%20a%20blast%20from%20the%20past&f=false, último acceso:08-09-2016
- [17] Zuccarelli, H. (1983). Ears Hear by Making Sounds. *New Scientist*. 100 (1383): 438-440
- [18] Moncibays, Y. (2011). Sonido Binaural. Evolución histórica y nuevas perspectivas con los paisajes sonoros. <https://repositorio-aberto.up.pt/bitstream/10216/67186/1/000150433.pdf>, último acceso:08-09-2016
- [19] Holzmann, C., Hochgatterer, M. (2012). Measuring Distance with Mobile Phones Using Single-Camera Stereo Vision. *IEEE 32nd International Conference on Distributed Computing Systems Workshops*. 88-93. DOI:10.1109/ICDCSW.2012.22
- [20] Empresa japonesa NSK. (2011). Perro Guía Robótico para ciegos desarrollado por NSK.
<http://www.roboticastreet.com/perro-guia-robotico-para-ciegos-desarrollado-por-nsk/>, último acceso:08-09-2016

- [21] The vOICe. (2016). Realidad Aumentada para invidentes. <https://www.seeingwithsound.com/>, último acceso:08-09-2016
- [22] EYEMUSIC. (2009). Aplicación para Android que representa con sonidos las imágenes. <http://brain.huji.ac.il/site/em.html>, último acceso:08-09-2016
- [23] Unity. (2016). Software líder a nivel mundial en la industria de los juegos. <https://unity3d.com/es/public-relations>, último acceso:08-09-2016
- [24] Vogel, C., Schindler, K., Roth, S. (2015). 3D Scene Flow Estimation with a Piecewise Rigid Scene Model. International Journal of Computer Vision Marr Prize (ICCV 2013) Special Issue. 115 (1): 1-28. doi:10.1007/s11263-015-0806-0
- [25] Herbst, E., Ren, X., Fox, D. (2013). RGB-D Flow: Dense 3-D Motion Estimation Using Color and Depth. Proceedings - IEEE International Conference on Robotics and Automation. 2276-2282. DOI:10.1109/ICRA.2013.6630885
- [26] Letouzey, A., Petit, B., Boyer, E. (2011). Scene Flow from Depth and Color Images. Proceeding of British Machine Vision Conference, United Kingdom. BMVA Press. 46:1-11. DOI: <http://dx.doi.org/10.5244/C.25.46>
- [27] Vedula, S., Rander, P., Collins, R., Kanade, T. (2005). Three-Dimensional Scene Flow. 7th International Conference on Computer Vision. 27 (3): 475-480. DOI: 10.1109/TPAMI.2005.63
- [28] Loop, C., Zhang, Z. (1999). Computing Rectifying Homographies for Stereo Vision. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 1: 125-131
- [29] Castillo, C., Jacobs, D. (2009). Using Stereo Matching with General Epipolar Geometry for 2D Face Recognition across Pose. IEEE Transactions on Pattern Analysis & Machine Intelligence. 31 (12): 2298-2304. doi:10.1109/TPAMI.2009.123
- [30] Radke, R. (2014). CVFX Lecture 15: Stereo correspondence. ECSE-6969 Computer Vision for Visual Effects. <https://youtu.be/kxsvG4sSuvA?t=24m48s>, último acceso:08-09-2016

6. Anexos

Cuestionario de las pruebas realizadas

Hombre: _____ Mujer: _____ Identificador: _____

Edad: _____

Prueba 1

1. ¿Percibe diferencia entre los sonidos?

SÍ: _____ NO: _____

¿Qué sonidos son parecidos?

| | L. VERTICAL | L. HORIZONTAL | CIRCUNFERENCIA |
|--------------|-------------|---------------|----------------|
| PUNTO | | | |
| | L. VERTICAL | | |
| | | L. HORIZONTAL | |

Prueba 2

2. ¿Percibe el punto?

SÍ: _____ NO: _____

¿Qué forma percibe?

3. ¿Percibe la línea vertical?

SÍ: _____ NO: _____

¿Qué forma percibe?

4. ¿Percibe la línea horizontal?

SÍ: _____ NO: _____

¿Qué forma percibe?

5. ¿Percibe la circunferencia?

SÍ: _____ NO: _____

¿Qué forma percibe?

Prueba 3

| | PUNTO | LÍNEA VERTICAL | LÍNEA HORIZONTAL | CIRCUNFERENCIA |
|----------|-------|-------------------|---------------------|----------------|
| SONIDO 1 | | | | |
| SONIDO 2 | | | | |
| SONIDO 3 | | | | |
| SONIDO 4 | | | | |
| SONIDO 5 | | | | |
| SONIDO 6 | | | | |
| SONIDO 7 | | | | |
| SONIDO 8 | | | | |