

Nº de proyecto: 2020-1-EL01-KA201-079157

Producción intelectual

O2: Software auditivo para discapacitados visuales

Socio: PANEPISTIMIO PATRON (E10209090, GR)



Erasmus+



Información sobre el proyecto

Acrónimo del proyecto:	G4ViD
Título del proyecto:	Geología para ciegos y deficientes visuales
Número de acuerdo:	2020-1-EL01-KA201-079157
Programa de la UE:	Erasmus
Página web del proyecto:	http://www.g4vid.eu/

Preparado por

Organización asociada: PANEPÍSTIMIO PATRÓN (E10209090, GR)

© G4ViD - Geología para ciegos y deficientes visuales

Descargo de responsabilidad:

El apoyo de la Comisión Europea a la elaboración de esta publicación no constituye una aprobación de su contenido, que refleja únicamente las opiniones de los autores, y la Comisión no se hace responsable del uso que pueda hacerse de la información aquí difundida.



Índice

1. Introducción

1.1 Acerca de esta producción intelectual

2 Colores y sonidos

2.1 RGB y HSL

2.2 El sonocromatismo

2.3 Igualación de colores en las frecuencias

3 Hardware y software

3.1 Sistema operativo del dispositivo portátil inteligente

3.2 Dispositivo portátil inteligente

3.3 Unidad

4 Diseño y desarrollo de la aplicación

4.1 Punto de contacto de la pantalla

4.2 Función de vibración

4.3 Colores y conversión a sonido

4.4 Sonido

4.5 Algoritmo final

5 Interfaz de usuario de la aplicación

5.1 Menú principal

5.2 Modo tren

5.3 Modo Play

5.4 Instrucciones vocales



1 Introducción

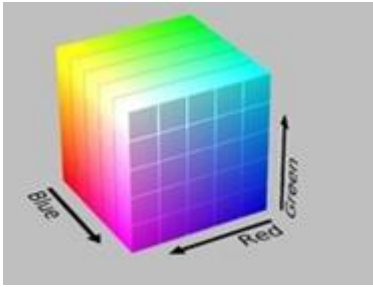
1.1 Acerca de esta producción intelectual

La aplicación de conversión visual a auditiva es el segundo resultado intelectual del proyecto. Esta aplicación utiliza parte de la base de datos geosite que contiene imágenes, pero también metadatos sobre las imágenes que ayudan al funcionamiento de la traducción visual a aural. La aplicación está diseñada para sistemas operativos Android. La interfaz es fácil de usar y consiste en sencillos comandos táctiles y retroalimentación por vibración. Todas las pantallas están asistidas por voz para los discapacitados visuales y muestran imágenes reales para que puedan ser utilizadas por usuarios con poca o ninguna discapacidad visual. Al iniciarse, la aplicación presenta una pantalla de selección de imágenes guiada por voz. El usuario sólo tiene que tocar la imagen correcta, que se traducirá en sonido. La aplicación también incluye una función de formación. El público puede descargar la última versión en Google Play Store, a la que se accede desde el sitio web del proyecto.

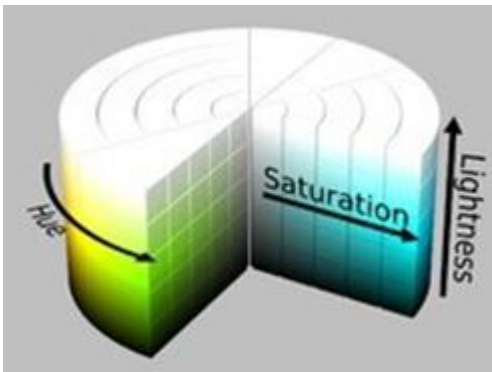
2. Colores y sonidos

2.1 RGB y HSL

RGB es un patrón de color aditivo en el que los tres colores primarios rojo, verde y azul se combinan para crear todos los demás colores. Un color en el estándar RGB, esencialmente describe la cantidad de rojo, verde y azul que contiene. Si los tres colores primarios están al mínimo, el color representado es el negro. Si estos tres colores están al máximo, el color representado es el blanco. Los valores RGB son tres números en los que el primero representa la cantidad de rojo, el segundo la cantidad de verde y el tercero la cantidad de azul. Las representaciones más famosas de los colores RGB están en la escala de 0 a 1 o en la escala de 0 a 255. Por ejemplo, el rojo máximo puede representarse como (1,0,0) o (255,0,0). Matemáticamente hablando, el patrón RGB, es una forma de describir un color en un cubo donde el rojo, el verde y el azul están cada uno en un eje diferente como se muestra en la imagen de abajo.



La representación HSL, se basa en Tono, Saturación y Luminosidad del color. HSL es una descripción del color basada en la geometría cilíndrica. El matiz viene dado por la dimensión angular sobre el cilindro, comenzando con el matiz rojo a 0 grados, continuando con el verde a 120 grados, el azul a 240 grados y cerrando de nuevo con el rojo a 360 grados. En el eje vertical central de esta geometría se representa la luminosidad, que es prácticamente la gama de los colores neutros. Interpretando los valores del eje de abajo hacia arriba, con la luminosidad en la parte inferior teniendo el valor 1, se cumple el color Negro, mientras que en la parte superior del eje la luminosidad tiene el valor 0 y se cumple el color blanco. Por último, en el eje central horizontal, la saturación del color se representa con el valor 0 situado en el centro del cilindro donde se encuentran los colores no saturados, mientras que el valor máximo 1 se encuentra en la superficie del cilindro donde se encuentran los colores saturados.










2.2 El sonocromatismo

El sonocromatismo es un fenómeno neurológico en el que los colores se perciben como sonidos. El término fue acuñado por Neil Harbisson para diferenciar su propia experiencia de percepción del color



de la empatía del sonido y el color. La empatía del sonido y el color es el resultado de una combinación de sentidos, en la que el sonido crea inadvertidamente una experiencia de color, forma y movimiento. La principal diferencia con la coloración sonora es que, en lugar de combinar dos sentidos, se crea una sensación completamente distinta. Quienes experimentan la coloración sonora no perciben necesariamente el color como una experiencia visual y son capaces de sentir, a través de este sentido, colores que están fuera del espectro visual, como el infrarrojo y el ultravioleta. Para realizar el fenómeno de la coloración sonora se utiliza una forma específica de transformar las frecuencias de color en frecuencias de sonido, la escala sonido-color de Harbisson. Esta escala tiene dos enfoques, la escala de sonido y color de la música y la escala de sonido y color puros.

PURE SONOCHROMATIC SCALE		
(invisible)	Ultraviolet	Over 717.591 Hz
	Violet	607.542 Hz
	Blue	573.891 Hz
	Cyan	551.154 Hz
	Green	478.384 Hz
	Yellow	462.023 Hz
	Orange	440.195 Hz
	Red	363.797 Hz
(invisible)	Infrared	Below 363.797 Hz

2.3 Correspondencia de colores en las frecuencias

La escala que se ha elegido utilizar es la escala de colores del sonido musical, mediante la cual cada color identificado en la imagen se asigna a la frecuencia deseada y se reproduce el tono correcto. En detalle, la correspondencia se muestra en la tabla siguiente.

Ángulo de matiz Frecuencia ajustada (Hz)

0°-30° 349,23

31°-60° 369,99

61°-90° 392,00

91°-120° 415,30

121°-150° 440,00



151°-180° 466,16

181°-210° 493,87

211°-240° 523,25

241°-270° 554,37

271°-300° 587,33

301°-330° 622,25

331°-359° 659,25

3 Hardware y software

3.1 Sistema operativo del dispositivo portátil inteligente

En primer lugar, hay que seleccionar el sistema operativo al que irá dirigida la aplicación. Esencialmente, es el sistema operativo que se requiere que esté instalado en el dispositivo del usuario que va a instalar y utilizar la aplicación. El sistema operativo seleccionado es Android. Android es un sistema operativo para dispositivos móviles que ejecuta el núcleo del sistema operativo Linux⁴. Fue desarrollado originalmente por Google y posteriormente por Open Handset Alliance. Android está diseñado principalmente para dispositivos con pantalla táctil, como teléfonos inteligentes y tabletas, con diferentes interfaces de usuario para televisores (Android TV), automóviles (Android Auto) y relojes de pulsera (Android Wear).

La principal razón por la que se seleccionó este sistema operativo, es el hecho de que es el software más utilizado en el mundo ya que se estima que los dispositivos con software Android tienen más ventas que todos los dispositivos Windows, iOS y Mac OS X juntos. Además, aunque se ha desarrollado para dispositivos con pantalla táctil, se ha utilizado en videoconsolas, cámaras digitales, PC ordinarios (P. Allah. HP Slate 21) y otros dispositivos electrónicos, lo que permite que la aplicación desarrollada sea compatible con una gran variedad de dispositivos además de teléfonos móviles y tabletas, realizando pequeñas correcciones. Además, a menudo recibe actualizaciones que contienen características útiles



para el usuario, y existe una gran comunidad que se ocupa constantemente del desarrollo y optimización de este sistema operativo. En el proceso de selección del sistema operativo, se nos pidió que eligiéramos entre Android e iOS dejando fuera el resto de sistemas operativos para dispositivos inteligentes ya que la selección del sistema operativo se ponía como eje clave en la difusión que tiene entre los usuarios, es decir, las personas con discapacidad visual. La decisión por Android fue rápida, ya que se perciben fácilmente otras ventajas que ofrece este sistema operativo.

Algo muy importante, relacionado con la facilidad del proceso de carga de la aplicación en la aplicación-Store de Android, la conocida Google Play, a diferencia de la AppStore de iOS que lleva bastante tiempo, debido a los muchos y complejos controles por parte del personal de Apple que se ocupa de la distribución de aplicaciones. Igualmente importante es la facilidad de instalación de una aplicación que acaba de ser creada en un dispositivo Android, pero también su prueba en condiciones reales, simplemente teniendo el desarrollador en posesión un dispositivo Android. Por ejemplo, el desarrollador puede subirlo a una nube.apk archivo que acaba de crear, y luego descargarlo en su dispositivo y elegir la instalación de la aplicación. Otra forma es conectar el dispositivo al ordenador del desarrollador que creó y guardó el archivo .apk, copiar el archivo dentro del dispositivo a través de un cable de datos, por ejemplo, usb a usb mini, usb a usb- c, usb-c a usb-c, y seleccionar la instalación de la aplicación. El único requisito previo es permitir que el dispositivo, a través del menú Privacidad de sus ajustes, instale aplicaciones de terceros. Esta característica que hemos descrito no se da a un dispositivo que utiliza el sistema operativo iOS.

3.2 Dispositivo portátil inteligente

Se seleccionó una Lenovo Tab 4 10" LTE para las presentaciones de la aplicación y las pruebas con los usuarios. En primer lugar, esta tableta utiliza el sistema operativo Android, tiene una gran pantalla de 10,1" y función de vibración.

3.3 Unidad: Plataformas y marcos para el desarrollo del software

Unity es un programa utilizado para desarrollar juegos, y su característica más importante es ser



multiplataforma, es decir, que cambiando sólo la forma en que se construye la aplicación o cambiando unas pocas líneas de código al mismo tiempo, se pueden crear aplicaciones para más de un sistema operativo. Por ejemplo, la aplicación también se puede utilizar en entornos Windows cambiando únicamente la forma en que se construye la aplicación. "Build" es el último proceso que se realiza en la plataforma al crear la aplicación, que lleva a la creación del archivo ejecutable. Una característica importante de la plataforma es que tiene una muy buena interfaz de usuario, también conocida como Interfaz de Usuario (UI), lo que facilita la estructuración de un proyecto y al mismo tiempo ayuda al desarrollador a crear exactamente lo que quiere sin gran consumo de tiempo.

4 Diseño y desarrollo de la aplicación

El diseño y desarrollo de la aplicación se basó en dos funciones principales, el modo de entrenamiento del usuario, en el que éste entra seleccionando la opción "Entrenar" del menú principal de la aplicación, y el modo de juego, en el que el usuario entra seleccionando la opción "Jugar" del mismo menú. En el modo de entrenamiento, el usuario aprende cómo la aplicación representa las diferentes formas a través de la vibración y los colores a través del sonido, siendo informado en primer lugar de lo que se está visualizando en ese momento en la pantalla del dispositivo. En el modo de juego, se pide al usuario que, utilizando los conocimientos y la experiencia adquiridos en la formación, identifique en las fotos formas y colores cuyo contenido desconoce, pero que, sin embargo, sabe que pertenecen al ámbito más amplio de una exposición fotográfica, es decir, fotos que contienen objetos geológicos. Una característica importante de la aplicación que facilita en gran medida el acceso y la navegación del usuario es la compatibilidad con comandos de voz.

4.1 Punto de contacto de la pantalla

La función extremadamente necesaria de la aplicación es el reconocimiento correcto del contacto del usuario. Es decir, que la aplicación reconozca exactamente dónde toca la pantalla el usuario y qué hay exactamente en ese punto de la pantalla. La primera comprobación que hay que hacer es si el usuario toca la pantalla o no. Mientras la función Actualizar, que se ejecuta una vez por fotograma, no reconozca



el contacto del usuario la aplicación no hace nada y espera a que el usuario reciba la entrada que quiere operar, es decir, el contacto en la pantalla de la tableta. A partir del momento en que la función recibe el contacto del usuario, comienza la identificación del punto de contacto, que es prácticamente un píxel de la pantalla. A continuación, hay que comprobar a qué punto de la foto corresponden las coordenadas del punto de contacto del usuario. Lo primero que hay que comprobar es si en el punto de contacto hay foto o no. Esto se consigue utilizando la función Raycast de la clase Physics2D de Unity. Mientras el usuario no toca un punto de la pantalla donde está la foto, entonces el rayo no detecta una "colisión" con un objeto, por lo que la aplicación espera a que el usuario toque un punto de interés. Cuando el usuario toca la foto, que es el punto de interés, Raycast devuelve que hubo una colisión y ahora están disponibles las coordenadas del punto de la pantalla que el usuario tocó y hay una foto.

4.2 Función de vibración

Una de las características más importantes de la aplicación, es la función de vibración utilizada para ayudar al usuario a identificar la forma de los objetos presentados en la foto. Tras evaluar los requisitos de los usuarios, llegamos a la conclusión de que será más fácil reconocer la forma si se interrumpe el proceso de reconocimiento de los colores mediante la reproducción del sonido. Por este motivo, se ha creado un conmutador que, cuando está activado, sitúa al usuario en un entorno de reconocimiento de formas, mientras que, cuando está desactivado, comienza la función de reconocimiento de formas y las demás funciones relacionadas con el reconocimiento de colores y la reproducción del sonido. Cada foto examinada en la aplicación va acompañada de su contorno, que es una foto blanca de las mismas dimensiones que capta los contornos de los objetos contenidos en la foto. Estos contornos se han creado algorítmicamente trazando y coloreando los bordes de los objetos o se han creado manualmente. Dado que el contorno tiene las mismas dimensiones que la foto correspondiente, sus sistemas de coordenadas son idénticos, es decir, cada píxel de la foto con coordenadas (x, y) es idéntico a cada píxel del contorno con las mismas coordenadas.

4.3 Colores y conversión a sonido

La principal característica de la aplicación es el reconocimiento de los colores de la imagen, la correcta



abstracción de las características de los colores que nos interesan y la correcta conversión de los mismos en sonido. Así, cada píxel tiene su propio color, por lo que corresponde a un determinado sonido, donde la frecuencia de la nota corresponde al tono del color, mientras que la luminancia del color, corresponde a la intensidad de la nota. Siguiendo la exigencia universal de los usuarios, se matizó la restricción del número de sonidos en cada foto. La siguiente etapa consiste en extraer las características de color de interés de la escala RGB, antes de procesarlas y convertirlas a HSL. La última etapa es el cálculo de los elementos del color de interés. En primer lugar, calculamos el valor RGB medio de todo el bloque y, a continuación, calculamos el Tono, la Saturación y la Luminancia, que son elementos del color para elegir el sonido adecuado que se reproducirá. A continuación, se realizan cálculos puramente algebraicos y el resultado es un valor único de Tono, Saturación y Luminancia en función del cual se generará el sonido correspondiente al bloque del punto de contacto.

4.4 Sonido

Un punto crítico de la aplicación G4ViD es la reproducción del sonido correcto, ya que es la característica en la que se basa la correcta renderización de los elementos de color de la foto al usuario con discapacidad visual. La reproducción se realiza desde Audio Source, un componente unity que se adjunta a un GameObject y cuya función es reproducir audio o sonidos. En el caso de esta aplicación, este componente se adjunta al GameObject cargado y la foto es examinada y reproduce un seno cuya frecuencia viene dada por el emparejamiento color-tono según la escala cromático-acústica de Harbisson, mientras que la intensidad del tono viene dada por el brillo del color(Luminancia). La correspondencia de la frecuencia con el color se realiza mediante el Tono del bloque. El matiz se mide y calcula en grados, ya que HSL es la representación de los colores en la llamada Rueda de Color. Cada vez que el dedo del usuario sobre la pantalla cambia de bloque, se produce un seno cuya frecuencia varía en función del valor de Hue. En cuanto a la segunda característica del sonido, su volumen, varía en función de lo brillante que sea el color en el bloque que se está examinando. La luminosidad se normaliza en un rango de valores de 0 a 100 después de que su cálculo inicial se hiciera en una escala de 0 a 1. El volumen que es una variable de AudioSource, toma valores de 0 a 1. Ahora, al igualar los valores de brillo con los niveles de intensidad, se hizo una aproximación experimental, comenzando con



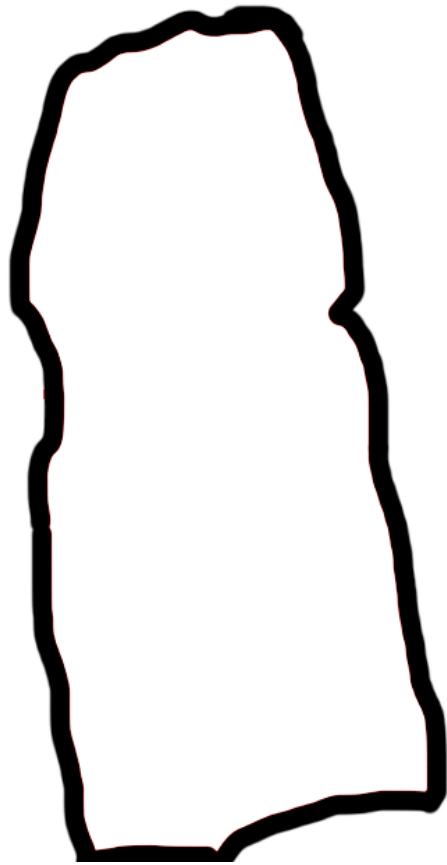
la lineal, ya que el brillo sube y la intensidad sube con 0 brillo = 0 intensidad (completamente oscuro) subiendo linealmente hasta 100 brillo = 1 intensidad (completamente brillante).

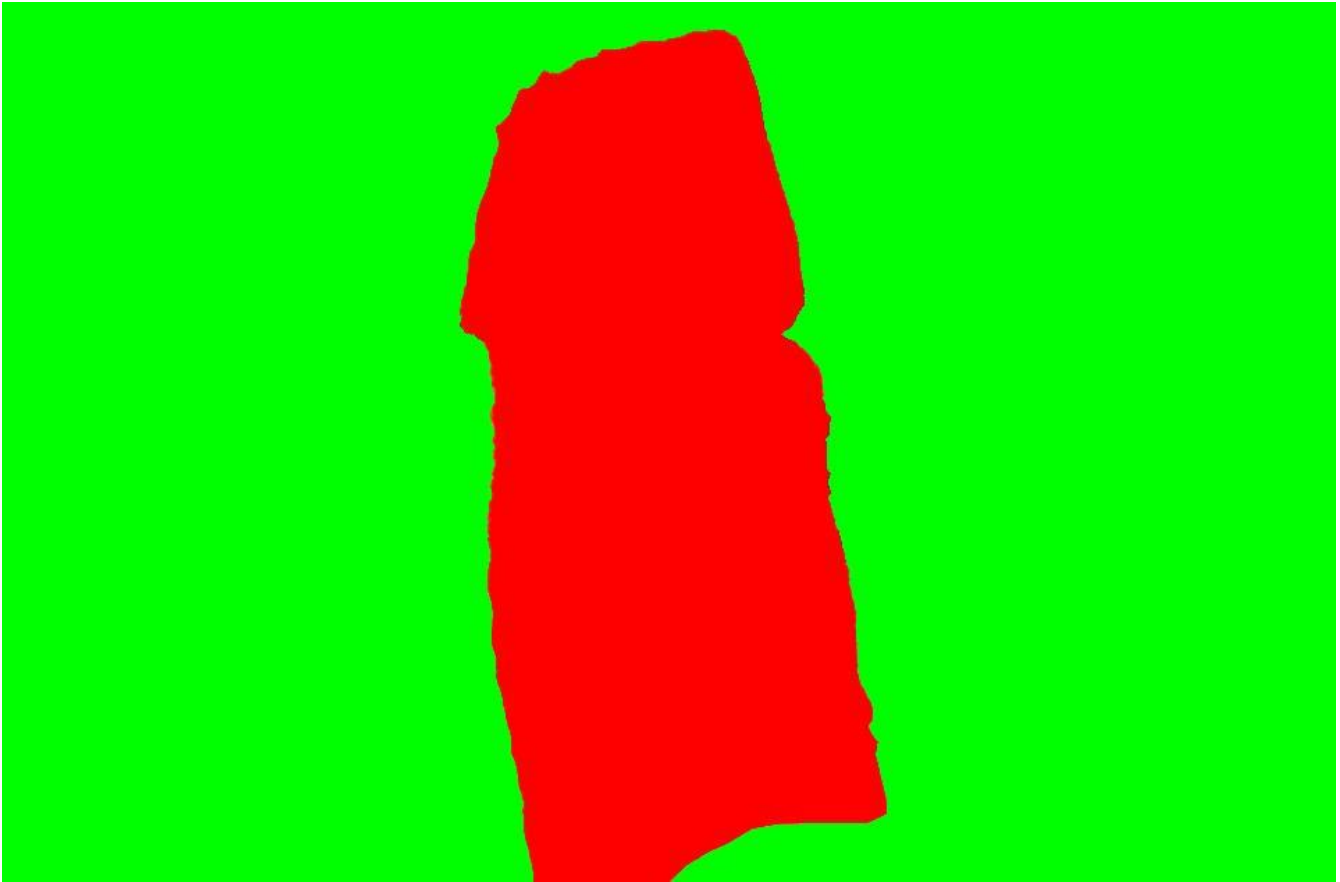
4.5 Algoritmo final

Cada foto examinada en la aplicación va acompañada de su contorno, creado algorítmicamente trazando los bordes de los objetos. Los usuarios sienten vibrar el dispositivo táctil cuando tocan los bordes de la imagen. Pasando al modo en color, cuando el usuario toca la pantalla, un algoritmo detecta en cuál de ellas

bloquea las coordenadas del dedo. A continuación, se calcula algebraicamente el valor RGB medio de esta zona, se almacena en una matriz y se traduce mediante la escala sonocromática de Harbisson en tono, luminosidad y saturación. Estos valores se utilizan para generar la frecuencia y el volumen de la onda sinusoidal, que se crea y reproduce en combinación con el nivel de interés que tiene cada parte de la imagen (primer plano, objeto principal, fondo) mediante un mapa de calor de interés.







5 Interfaz de usuario de la aplicación

5.1 Menú principal

A continuación se muestra el menú principal de la aplicación. Es el menú que el usuario ve cuando entra en la aplicación. Hay tres botones que tienen funciones diferentes mientras que el título de la aplicación aparece en lo alto del centro. El botón "Entrenar" lleva al modo de entrenamiento del usuario, el botón "Jugar" lleva al modo de juego de la aplicación mientras que "Salir" lleva al cierre de la aplicación.

GEOLOGY FOR VISUAL IMPAIRED AND DISABLED G4ViD



SELECT PARK

TRAIN

USER PROFILE

Co-funded by the
Erasmus+ Programme
of the European Union



REPEAT

QUIT

Co-funded by the Erasmus+ Programme of the European Union National Agency IKY - Project: 2017-1-EL01-KA201-036255

5.2 Modo tren

Al pulsar el botón "Entrenar", el usuario entra en el modo de entrenamiento para familiarizarse con el reconocimiento de formas a través de la vibración del dispositivo y el reconocimiento de colores a través de sonidos. El protocolo de entrenamiento prevé que el usuario sea informado por EE.UU. o por su guía sobre lo que aparece en la pantalla y, a continuación, proceda a explorar las funciones, sabiendo lo que debe sentir u oír. La entrada en este modo se realiza con la aparición de un cuadrado. El usuario se encuentra ahora en modo de vibración y reconocimiento de formas. Esto significa que cada vez que su dedo toque cualquiera de los contornos de abajo, el dispositivo vibrará. También vemos tres tipos de botones en este proceso. El "Exit" que lleva a la salida del modo de entrenamiento, el siguiente que lleva a la siguiente fase de entrenamiento y el botón de retroceso que gira a la fase de entrenamiento anterior. Cuando el usuario termina con el Cuadrado y pulsa el botón siguiente, es conducido al triángulo isósceles, mientras que la siguiente fase es un círculo, que es la última fase del entrenamiento de reconocimiento de formas mediante vibración. Pulsando de nuevo Siguiente, el usuario procede al



entrenamiento en colores, a través del sonido. Esta etapa consta también de 3 imágenes. La primera está dividida por la mitad, en la parte izquierda hay color rosa mientras que en la parte derecha hay color verde. El objetivo de esta imagen es que el usuario se entrene en el proceso de percibir el cambio de tonos. Para que esto sea posible, el usuario se mueve de una mitad a la otra y señala el momento en el que cree que hay un cambio de color. Se observa que en toda la parte rosa y en toda la parte verde, el brillo es constante, por lo que allá donde se desplaza en cada mitad el volumen no fluctúa sino que permanece constante. En la segunda imagen hay un Gradiente de color, esencialmente es una transición suave del azul en la parte izquierda al rojo en la parte derecha. El objetivo de esta imagen es entrenar al usuario con los matices de un color en el reconocimiento de las fluctuaciones de volumen al mismo tiempo que el cambio de nota. En esta foto el usuario es conducido de izquierda a derecha desde tonos brillantes de azul, a tonos más oscuros de azul, a púrpura, a tonos oscuros de rojo y al final a tonos brillantes de rojo. En el centro tiene el brillo y la intensidad más bajos, mientras que hacia los extremos tiene un brillo y una intensidad más altos, al tiempo que han cambiado las frecuencias del sonido que se reproduce. La tercera y última imagen, es una secuencia de colores. En ella se explica el orden de los colores y el usuario intenta identificar las notas con los colores. En este gráfico el usuario escuchará las 12 frecuencias que se pueden oír dentro de la aplicación. Estas notas también se oirán por turnos, si empieza de arriba a abajo las oirá desde la frecuencia más alta y terminará en la más baja, mientras que si se mueve al revés las oirá en el orden inverso.

5.3 Modo Play

La segunda función de la aplicación, y la más importante, es la de juego. Se invita al usuario a entrar en ella una vez que ha completado el proceso de formación y se ha familiarizado con la vibración y los sonidos, para experimentar y "sentir" una serie de fotos que representan objetos geológicos. Así, la primera imagen que aparece tras entrar en este modo es un nuevo menú, que incluye una lista de geoparques, cada uno de los cuales contiene fotos geológicas. Por último, en la parte inferior derecha de la pantalla, hay un botón "Atrás" que devuelve al usuario al menú principal de la aplicación al pulsarlo. Pulsando sobre cualquiera de las fotos, el usuario abre la foto y ya está preparado para aplicar lo que se le ha enseñado anteriormente. En la parte superior derecha de la pantalla destaca un

conmutador que alterna los modos de vibración y sonido. Mientras está marcada, el usuario está en modo vibración, al pulsarla pasa a modo sonido.



Lesvos



Sevilla



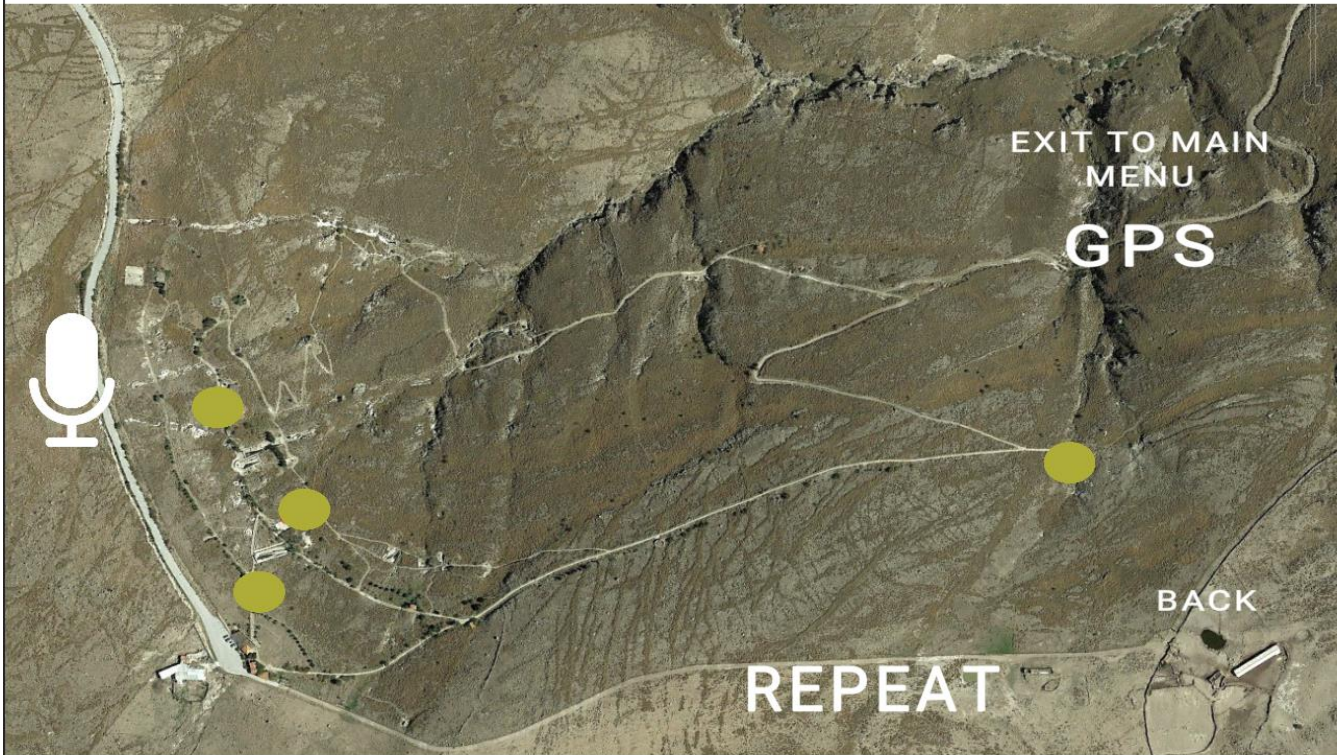
Latvia

EXIT TO MAIN
MENU



Cyprus

REPEAT



5.4 Instrucciones vocales

Como ya se ha dicho, para lograr la accesibilidad se integraron instrucciones y comandos de voz. El usuario, tras entrar en la aplicación, oye el saludo y cada pantalla tiene instrucciones de voz detalladas.