

Project No: 2020-1-EL01-KA201-079157

Intellectual Output

O2: Aural Software for the Visual Impaired

Partner: PANEPISTIMIO PATRON (E10209090, GR)



Erasmus+



Project information

Project acronym: G4ViD
Project title: Geology for the Blind and the Visually Impaired
Agreement number: 2020-1-EL01-KA201-079157
EU programme: Erasmus+
Project website: <http://www.g4vid.eu/>

Prepared by

Partner Organisation: PANEPISTIMIO PATRON (E10209090, GR)

© G4ViD – Geology for the Blind and the Visually Impaired

Disclaimer:

The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



Table of Contents

1. Introduction

1.1 About this Intellectual Output

2 Colours and Sounds

2.1 RGB and HSL

2.2 Sonochromatism

2.3 Matching colors at frequencies

3 Hardware and Software

3.1 Operating System of the Smart Portable Device

3.2 Smart Portable Device

3.3 Unity

4 Design and Development of the Application

4.1 Screen Contact Point

4.2 Vibration function

4.3 Colors and Conversion to Sound

4.4 Sound

4.5 Final algorithm

5 Application User Interface

5.1 Main Menu

5.2 Train Mode

5.3 Play Mode

5.4 Vocal Instructions



1 Introduction

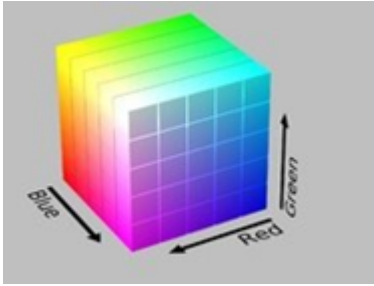
1.1 About this Intellectual Output

The visual to aural conversion application is the 2nd Intellectual Output of the project. This application uses part of the geosite database containing images but also metadata about the images that assist the operation of the visual-to-aural translation. The application is built for Android operating systems. The interface is user friendly and consists of simple touch commands and vibration feedback. All screens are voice-assisted for the visually impaired whilst they also depict real images so that they can be used by users with no or little visual impairment. When initially launched, the application presents a voice-guided image selection screen. The user needs only to touch the right image which will be translated to sound. The application also includes a training feature. The public can download the latest version from the Google play store as linked in the project website.

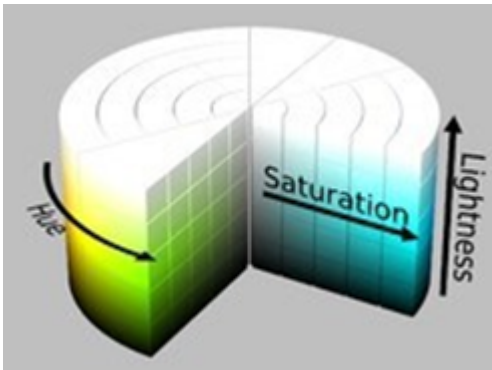
2. Colors and Sounds

2.1 RGB and HSL

RGB is an additive color pattern in which the three primary colors red, green and blue are combined to create all other colors. A color in the RGB standard, essentially describes the amount of red, green and blue it contains. If all three of these primary colors are at minimum, then the color represented is black. If these three colors are maximum, the color represented is white. RGB values are three numbers where the first number represents the amount of red, the second the amount of green and the third the amount of blue. The most famous representations of RGB colors are either on the scale from 0 to 1 or on the scale from 0 to 255. For example, the maximum red can be represented as either (1,0,0) or (255,0,0). Mathematically speaking, the RGB pattern, is a way to describe a color in a cube where red, green and blue are each on a different axis as shown in the image below.



The HSL representation, is based on Hue, Saturation and Lightness of color. HSL is a color description based on cylindrical geometry. The hue is given by the angular dimension on the cylinder, starting with the red tint at 0 degrees, continuing with the green tint at 120 degrees, the blue tint at 240 degrees and closing again to the red one at 360 degrees. On the central vertical axis of this geometry, brightness is represented, which is practically the range of neutral colors. Interpreting the values of the axis from the bottom up, with the brightness at the bottom having the value 1, The Black color is met, while at the top of the axis the brightness has the value 0 and the white color is met. Finally, on the horizontal central axis, color saturation is represented with the value 0 located in the center of the cylinder where the unsaturated colors are located, while the maximum value 1 is located on the surface of the cylinder where the saturated colors are located.










2.2 Sonochromatism

Sonochromatism is a neurological phenomenon in which colors are perceived as sounds. The term was coined by Neil Harbisson to differentiate his own experience of color perception from sound and color



empathy. Empathy of sound and color is the result of a combination of senses, in which sound inadvertently creates an experience of color, shape and movement. The main difference with sound coloring is that instead of combining two senses, a completely different sensation is created. Those who experience sound coloring do not necessarily perceive color as a visual experience and are able to feel, through this sense, colors that are outside the visual spectrum such as infrared and ultraviolet. A specific way of transforming color frequencies into sound frequencies, the Harbisson sound-color scale, is used to realize the phenomenon of sound coloring. This scale has two approaches, the Music Sound and Color scale and the Pure Sound and Color scale.

PURE SONOCHROMATIC SCALE		
(invisible)	Ultraviolet	Over 717.591 Hz
	Violet	607.542 Hz
	Blue	573.891 Hz
	Cyan	551.154 Hz
	Green	478.394 Hz
	Yellow	462.023 Hz
	Orange	440.195 Hz
	Red	363.797 Hz
(invisible)	Infrared	Below 363.797 Hz

2.3 Matching colors at frequencies

The scale chosen to be used is the musical sound color scale, by which each color identified in the image is assigned to the desired frequency and the correct tone is reproduced. In detail the matching is shown in the table below.

Hue Angle Matched Frequency (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 and Software

3.1 Operating System of the Smart Portable Device

First, the operating system to which the application will be addressed must be selected. Essentially, it is the operating system that is required to be installed in the device of the user who is going to install and use the application. The selected operating system is Android. Android is an operating system for mobile devices which runs the kernel of Linux4 operating system. It was originally developed by Google and later by Open Handset Alliance. Android is primarily designed for touch screen devices, such as smartphones and tablets, with different user interfaces for TVs (Android TV), cars (Android Auto) and wristwatches (Android Wear).

The main reason why this operating system was selected, is the fact that it is the most widely used software in the world since it is estimated that devices with Android software have more sales than all Windows, iOS and Mac OS X devices combined. Also, although it has been developed for touch-screen devices, it has been used in game consoles, digital cameras, ordinary PCs (P. Allah. HP Slate 21) and other electronic devices, which allows the developed application to be compatible with a variety of devices besides mobile phones and tablets, by making little corrections. Also, it often receives updates that contain useful features for the user, and there is a large community that constantly deals with the development and optimization of this operating system. In the operating selection process, we were asked to choose between Android and iOS leaving the rest operating systems for smart devices out



because the selection of the operating system was put as a key axis in how widespread it is among users, i.e. people with visual disabilities. The decision on Android was quick, as other benefits offered by this operating system are easily perceived.

Something very important, related to the ease of the upload process of the application in the application-Store of Android, the well-known Google Play, as opposed to the iOS AppStore that takes quite a bit of time, due to the many and complex controls by Apple staff dealing with the distribution of applications. Equally important is the ease of installation of an application that has just been created on an Android device but also testing in real conditions, simply having the developer in possession of an Android device. For example, the developer can upload it to a cloud.apk file just created, and then download it to his device and choose the installation of the application. Another way is to connect the device to the computer of the developer who created and saved the .apk file, copy the file within the device via a data cable, for example, usb to usb mini, usb to usb- c, usb-c to usb-c, and select the installation of the application. The only prerequisite is to allow the device, through the Privacy menu in its settings, to install third-party applications. This feature we described is not given to a device that uses iOS operating system.

3.2 Smart Portable Device

A Lenovo Tab 4 10" LTE was selected for the app presentations and tests with users. First of all, this tablet uses Android operating system, has a large 10.1" screen and vibration function.

3.3 Unity: Platforms and Frameworks for the Development of the Software

Unity is a program used to develop games, and its most important feature is being cross- platform i.e. either by changing only the way the application is built or by changing a few lines of code at the same time, applications can be created for more than one operating system. For example, the app can also be used in Windows environments by only changing the way the application is built. "Build" is the last process that is done on the platform when creating the application, leading to the creation of the executable file. An important feature of the platform is that it has a very good user interface, also



known as the User Interface (UI), making it easy to structure a project and at the same time it helps the developer to create exactly what he wants without big time consumption.

4 Design and Development of the Application

The design and development of the application were based on two main functions, the user training mode, in which the user enters by selecting the "Train" option from the main menu of the application, and the game mode, in which the user enters by selecting the "Play" option from the same menu. In the training mode, the user learns how different shapes through vibration and colors through sound are rendered by the application, being first informed about what is currently being displayed on the screen of the device. In the game mode, the user is asked, using the knowledge and experience gained from education, to identify in photos, shapes and colors of which he is not aware in terms of content, which, however, he knows that they belong to the broader field of a photographic exhibition, i.e. photos containing geological objects. An important feature of the application which makes it easy for the user to access and navigate it to a large extent is the support of voice commands.

4.1 Screen Contact Point

The extremely necessary function of the application is the correct recognition of the user's contact. That is, the application to recognize exactly where the user touches the screen and what exactly is at that point of the screen. The first check to be made is whether the user touches the screen or not. As long as the Update function, which runs once per frame, does not recognize contact from the user the application does nothing and waits until the user receives the input he wants to operate, i.e. the contact on the screen of the tablet. From the moment the function receives the user's contact, the identification of the contact point begins which is practically a pixel of the screen. Next, it must be checked at which point of the photo the coordinates of the user's contact point correspond. The first thing to check is whether at the point of contact, there is a photo or not. This is achieved by using the Raycast function of the Physics2D Class of Unity. While the user does not touch a point of the screen where the photo is, then the beam does not detect a "collision" with an object, so the application



waits until the user touches a point of interest. When the user touches the photo, which is the point of interest, Raycast returns that there was a collision and now the coordinates of the point of the screen that the user touched and there is a photo are available.

4.2 Vibration function

One of the most important features of the application, is the function of vibration used to help the user to identify the shape of objects presented in the photo. After assessing the requirements of the users, we concluded that it will be easier to recognize the shape if it is cut off from the process of recognizing the colors by reproducing the sound. For this reason, a toggle has been created which, when enabled, puts the user in a shape recognition environment, while when disabled, the shape recognition function and the other functions related to color recognition and sound reproduction begin. Each photo examined in the application is accompanied by its contour, which is a white photo of the same dimensions that captures the contours of the objects contained in the photo. These contours have either been created algorithmically by tracing and coloring the edges of objects or have been created manually. Since the contour has the same dimensions as the corresponding photo, then their coordinate systems are identical, that is, each pixel of the photo with coordinates (x, y) is identical to each pixel of the contour with the same coordinates.

4.3 Colors and Conversion to Sound

The main feature of the application is the recognition of the colors of the image, the correct abstraction of the characteristics of the colors that interest us and the correct conversion of them into sound. So, each pixel has its own color, so therefore it corresponds to a certain sound, where the frequency of the note corresponds to the hue of the color, while the luminance of the color, corresponds to the intensity of the note. Following the universal requirement of users, the restriction of the number of sounds in each photo was qualified. The next step is to extract the color characteristics of interest from the RGB scale, before processing them and converting them to HSL. The last stage is the calculation of the elements of the color of interest. First we calculate the average RGB value of the entire block and then we calculate the Hue, Saturation and Luminance, which are



elements of the color for choosing the right sound to play. Then, purely algebraic calculations follow, and the result is a unique value for Hue, Saturation and Luminance based on which the sound corresponding to the block of the contact point will be generated.

4.4 Sound

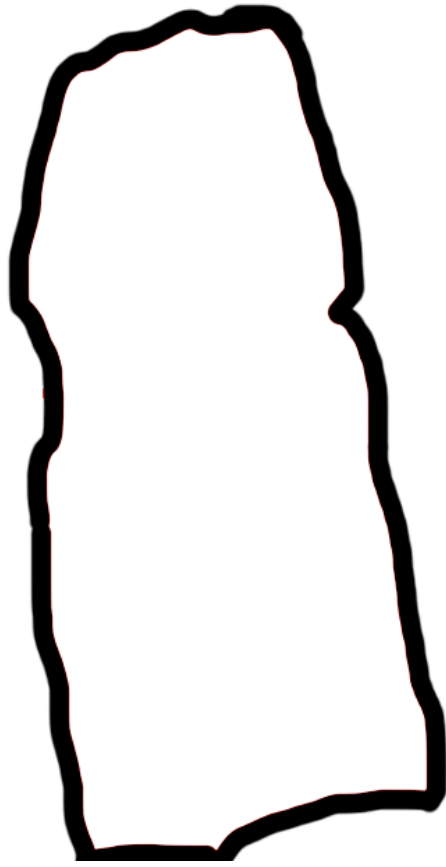
A critical point of the G4ViD application is the reproduction of the correct sound, as this is the characteristic on which the correct rendering of the color elements of the photo to the visually impaired user is based. Playback is done from Audio Source, a unity component attached to a GameObject and its role is to play audio or sounds. In the case of this application, this component is attached to the loaded GameObject and the photo is examined and reproduces a sine whose frequency is given by the color-tone matching according to Harbisson's chromatic-acoustic scale, while the tone intensity is given by the brightness of the color(Luminance). The matching of frequency with color is done by the Hue of the block. Hue is measured and calculated in degrees since HSL is the representation of colors on the so-called Color Wheel. Each time the user's finger on the screen changes block, a sine is produced whose frequency varies according to the value of Hue. As for the second characteristic of the sound, its volume, it varies according to how bright the color is on the block that is being examined. The brightness is normalized in a range of values from 0 to 100 after its initial calculation was made on a scale from 0 to 1. The volume that is a variable of AudioSource, takes values from 0 to 1. Now, in matching brightness values to intensity levels, an experimental approach was made, starting with the linear one, as the brightness goes up and the intensity goes up with 0 brightness = 0 intensity (completely dark) going linearly up to 100 brightness = 1 intensity (completely bright).

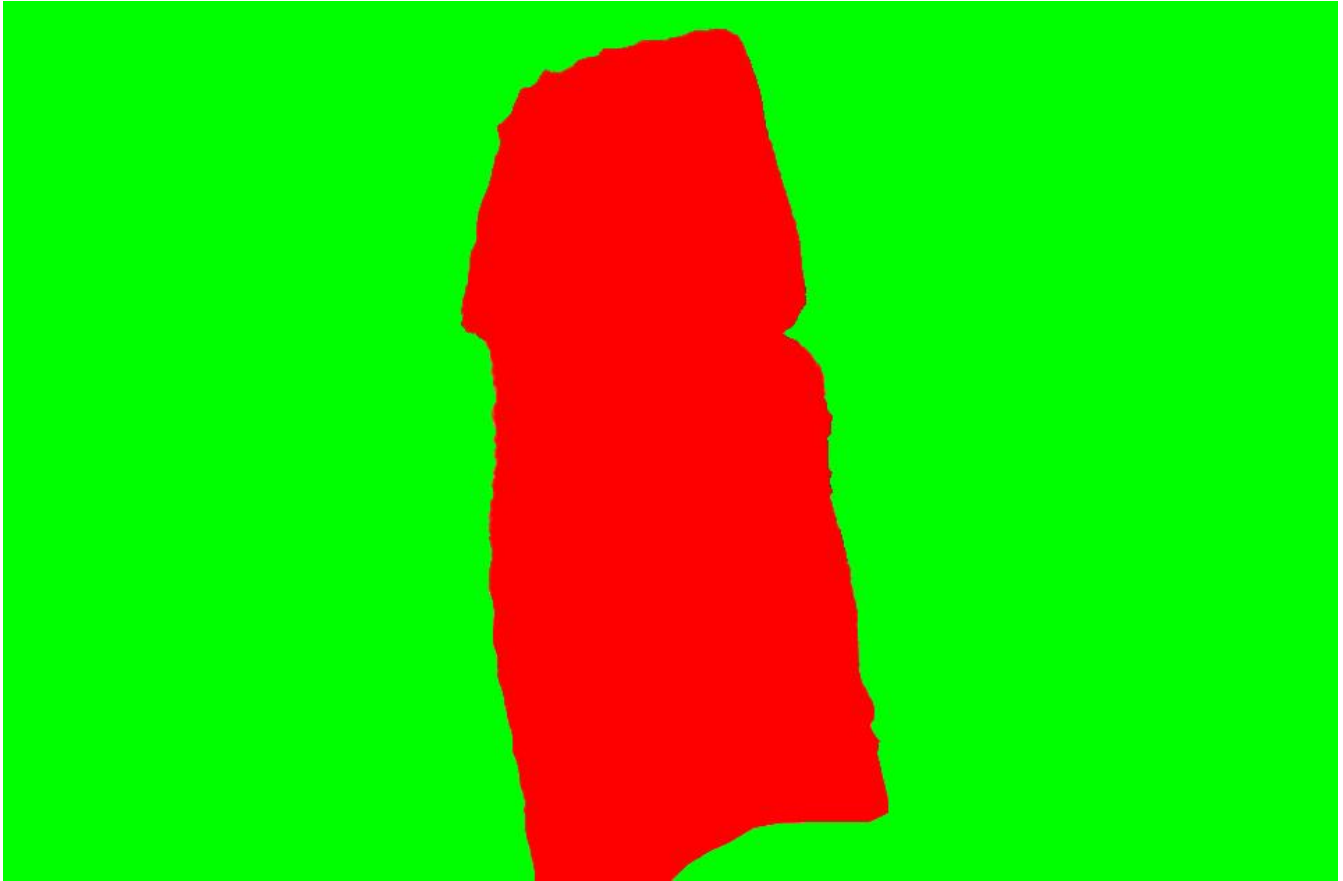
4.5 Final algorithm

Each photo examined in the application is accompanied by its contour, created algorithmically by tracing the edges of objects. Users feel the touch device vibrate when they touch the image borders. Proceeding to colour mode, when the user touches the screen, an algorithm detects on which of these blocks the coordinates of the finger. Then, the mean RGB value of this area is algebraically calculated,

stored in a matrix and translated through the Harbisson's Sonochromatic Scale into hue, lightness and saturation. These values are used to generate the frequency and the volume of the sinewave, which is created and played in combination with the level of interest each part of the image has (foreground, main object, background) through an interest heat map.







5 Application User Interface

5.1 Main Menu

Below is the main menu of the application. It is the menu that the user sees when he enters the application. There are three buttons that have different functions while the title of the application appears high in the center. The “Train” button leads to the user's training mode, the “Play” button leads to the app's play mode while “Quit” leads to the shutdown of the app.

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 Train Mode

By pressing the “Train” button, the user enters the training mode to familiarize himself with the recognition of shapes through the vibration of the device and the recognition of colors through sounds. The training protocol provides for the user to be informed by US or his / her guide about what is displayed on the screen and then proceed to explore the functions, knowing what to feel or hear. Entry into this mode is made with the appearance of a square. The user is now in vibration and shape recognition mode. This means that every time his finger touches any of the contours below, the device will vibrate. We also see three kinds of buttons in this process. The “Exit” that leads to the exit from the training mode, the next that leads to the next training phase and the back button that spins to the previous training phase. After the user finishes with the Square and presses the next button, he is led to the isosceles triangle while the next stage is a circle, which is the last stage in the training of shape recognition through vibration. By pressing Next again, the user proceeds to train in colors, through sound. This stage consists of 3 images as well. The first is divided in the middle, on the left part there is



pink color while on the right part there is green color. The purpose of this image is for the user to be trained in the process of perceiving the change of tones. To make this possible, the user moves from one half to the other and points to the moment when he thinks there is a color change. It is noticed that throughout the pink and throughout the green part, the brightness is constant, so wherever it moves in each half the volume does not fluctuate but remains constant. In the second image there is a Gradient color, essentially it is a smooth transition from blue on the left side to red on the right side. The purpose of this image is to train the user with the shades of a color in the recognition of fluctuations in volume at the same time as the change of note. In this photo the user is led from left to right from bright shades of blue, to darker shades of blue, to purple, to dark shades of red and at the end to bright shades of red. In the middle it has the lowest brightness and intensity, while towards the ends it has a higher brightness and intensity, while at the same time the frequencies of the sound being reproduced have changed. The third and last image, is a sequence of colors. This explains the order of colors and the user tries to identify the notes with the colors. In this graph the user will hear all 12 frequencies that can be heard within the application. These notes will also be heard in turn, if he starts from top to bottom he will hear them from the highest frequency and will end up at the lowest, while if it moves upside down he will hear them in the reverse order.

5.3 Play Mode

The second, and more important, function of the application is the game function. The user is invited to enter there after he has completed the training process and has become familiar with vibration and sounds, to experiment and "feel" a series of photos that represent geological objects. So the first image that comes across after entering this mode is a new menu, which includes a list of geopark each one containing geological photos. Finally, at the bottom right of the screen, there is a "Back" button that returns the user to the main menu of the application by pressing it. By clicking on any of the photos, the user opens the photo and is now ready to apply what he has been trained on earlier. On the top right of the screen a toggle which alternates vibration and sound modes, stands out. While it is checked, the user is in vibration mode, when pressing it passes into sound mode.



Lesvos



Sevilla



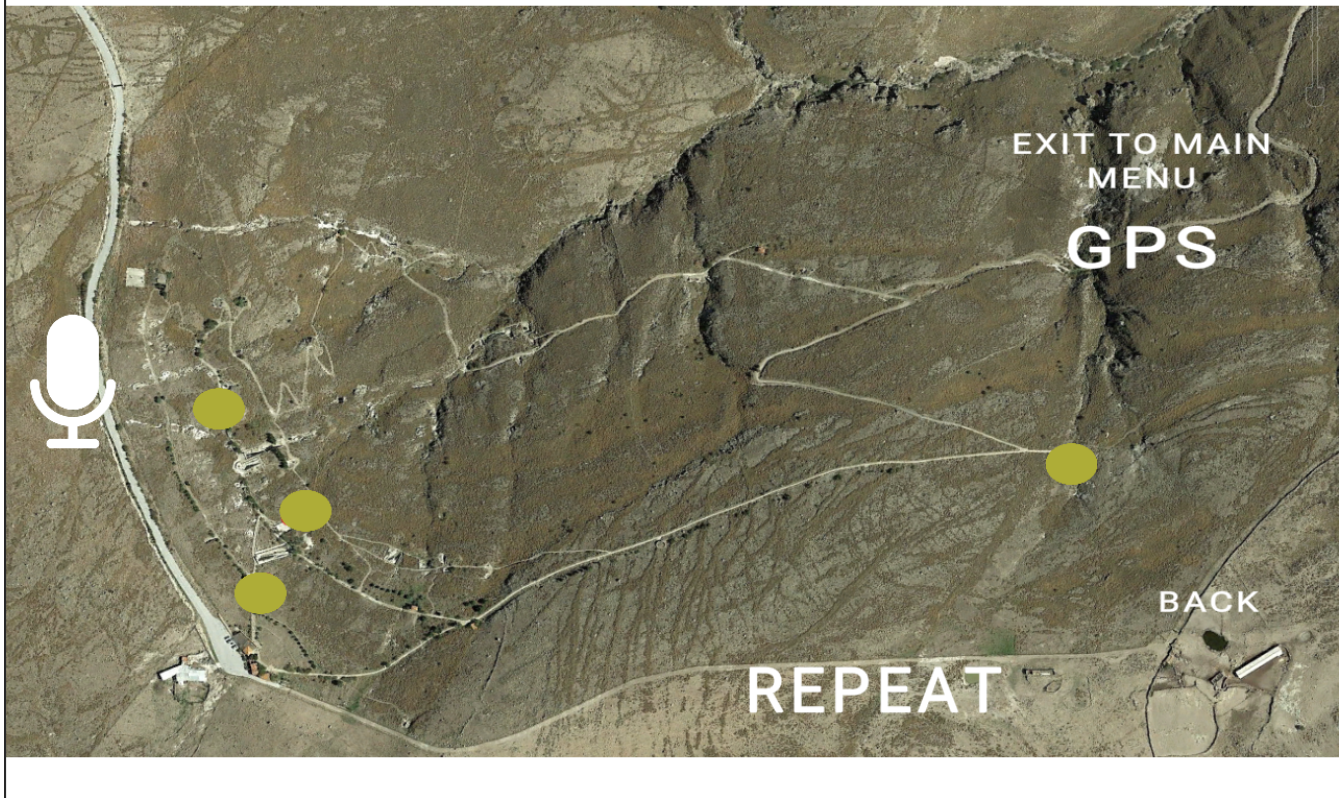
Latvia

EXIT TO MAIN
MENU



Cyprus

REPEAT



5.4 Vocal Instructions

As mentioned earlier, to achieve accessibility, voice instructions and voice commands were integrated. The user after entering the application hears the greeting and each screen has detailed voice instructions.