Advanced User Interface

2018/19

Public speech in Virtual Reality SpeechVR

Author

Alberto Patti
Yu Zhu

Abstract

This project offers people that have fear of speaking in public an instrument that allows them to try their speech in a virtual environment. This basic idea is expanded by using a biosensor to manipulate the environment the user is put in: the amount of people that the user sees in the audience changes based on the anxiety level determined by reading the values from the biosensor.

Contents

1 Introduction

1.1 Purpose

The purpose of this document is to give information about the "WIVR games for stress relief" project developed for the Advanced User Interface course.
This document aims to explain:

- The needs, goals and requirements for the targeted users;
- Previous researches and projects on the same topic;
- The choices made throughout the development of the project;

1.2 Scope

"Public speech in Virtual Reality" (SpeechVR) is a VR application that tries to give an instrument to people that have fear of speaking in public to improve their ability to speak to an audience.
The application offers to the user a virtual theatre where he/she can try a speech in front of an audience that can react based on his/her performance. The main functionality of the application is given by a biosensor (Empatica E4) that allows the tracking of the heart rate and the galvanic skin response of the user to evaluate the state of mind of the subject and decide how the environment should change: whether thee amount of people that the user sees in the audience can be changed or, in case the application consider that the user is in a situation of high stress, block the test.

1.3 Definitions, acronyms and abbreviations

- VR: Virtual Reality
- HMD: Head Mounted Display
- WIVR: Wearable Immersive Virtual Reality
- HR: Heart Rate
- GSR: Galvanic Skin Response

2 NGR

2.1 Target

The main target of the project are people that have fear of speaking in public. This kind of fear can be categorized as part of social phobia, i.e. "persistent fears of situations involving social interaction or social performance or situations in which there is the potential for scrutiny by others" [1].

2.2 Context and Needs addressed

"In social/evaluative situations, the primary threat stimulus is an audience and the primary threatening outcome is negative evaluation from the audience" [1]. This means that the idea of being evaluated by the audience is enough to start a loop that keeps fuelling the anxiety of the subject as shown in figure 1.

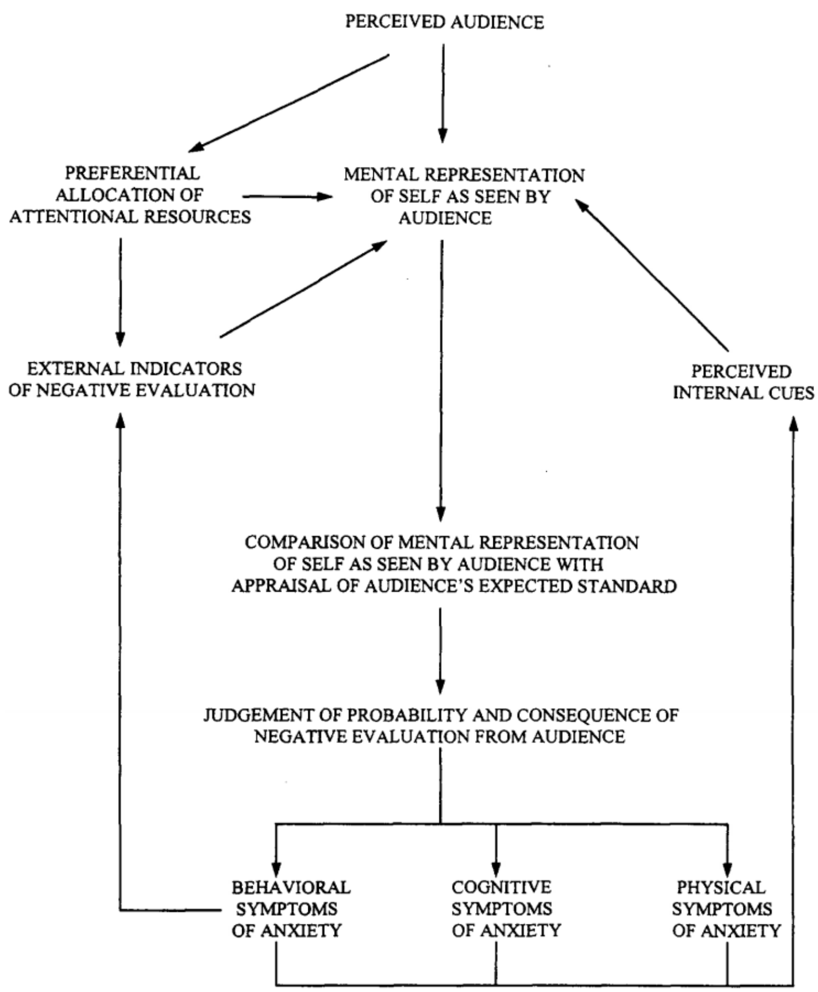

Figure 1: A model of the generation and maintenance of anxiety in social/evaluative situations [1].

Knowing this, the needs that were identified are:

- Have more confidence around people during the speech
- Listen to the speech after the performance

2.3 Goals

- Improve the ability to speak in public
- Allow the subject to be less anxious before and during the speech

2.4 Requirements

- The applications should provide an environment where the subject can try his/her speech in front of a virtual audience.
- The application should progressively change the number of people that the user can see in the audience based on his/her state of mind.
- The application must stop the test in case the subject doesn't feel well.
- The application should reward the user at the end of his/her speech.
- The application should record the speech and play it if needed.

2.5 Constraints

2.5.1 Base

- HMD;
- A smartphone running Android Jellybean or higher (4.1.x+);
- Empatica E4;
- Microphone;
- Headphones;
- Comfortable place where the user can sit down and rest the arm;
- An internet connection.

2.5.2 Pc Client

- A pc (Windows 7 or higher) with Visual C++ Redistributable Package installed;
- Bluegiga Bluetooth Smart Dongle;
- E4 Streaming Server installed.

2.5.3 Android Client

- A smartphone running Android.

3 State of the art

3.1 Applications

There are many application with the same target and objectives as this project that were developed and are nowadays available:

- Virtual Orator
- Speech Center VR
- VirtualSpeech
- #BeFearless
- Public Speaking Simulator VR

These are some of the available applications but, even though the premises are the same, each of them present different features and are available on different platform as shown in the tables below.

Virtual Orator	Oculus Rift / HTC Vive	3D Environment
Speech Center VR	Oculus Rift	3D Environment
VirtualSpeech	Android	360° video
#BeFearless	Android	3D Environment
Public Speaking Simulator VR	Android	3D Environment

	Multiple Environment	Upload documents	Record your performance	Question from the audience	Speech analysis	Distractions	Variable number of people during the speech	Biosensor	Lectures	Evaluation of the performance
Virtual Orator	X	X	X	X		X				
Speech Center VR	X	X	X			X			X	X
VirtualSpeech	X	X	X		X	X		X	X	X
#BeFearless	X	X	X		X			X		X
Public Speaking Simulator VR						X	X			

The base of this project is the same as the application listed before: giving the user an environment where he/she can freely try his/her speech. What makes SpeechVR unique is the usage of a biosensor as a mean to control the environment the user is put in. In fact, the only app that uses a biosensor are VirtualSpeech and #BeFearless but they use it as another parameter to give a score to the overall performance.

3.2 Research

There are many researches about public speech anxiety (and social phobia in general) but the most relevant for the sake of this project are:

- Slater, M., Pertaub, D. P., & Steed, A. (1999). Public speaking in virtual reality: Facing an audience of avatars. [2]

  The focus of this paper is to analyze how people evaluate themselves while in front of an audience with different reactions using VR.

- Pertaub, D. P., Slater, M., & Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. [3]

  This is a more thorough analysis of the previous research.

- Chollet, M., Sratou, G., Shapiro, A., Morency, L. P., & Scherer, S. (2014, May). An interactive virtual audience platform for public speaking training. [4]

  The focus of this research is to design a way to let people learn how to behave in front of a fake audience that reacts to the user actions. This research doesn't use VR but instead works with screens and audiovisual sensors (Kinect) to analyze the user behaviour.

- Poeschl, S., & Doering, N. (2012, March). Virtual training for Fear of Public Speaking—Design of an audience for immersive virtual environments. [5]

  This research explains how to develop an audience that shows realistic behaviour.

- McKinney, M. E., Gatchel, R. J., & Paulus, P. B. (1983). The effects of audience size on high and low speech-anxious subjects during an actual speaking task. [6]

  This research studies how people react during a speech in front of different amount of people hearing.

4 UX design

4.1 General

SpeechVR is a VR application on smartphones: it allows the user to be put into a virtual environment thanks to the usage of only a HMD and the user phone. It also uses a biosensor as a way to manipulate such environment.
The application offers two modes to the user:

- Free talk;
- Interview.

The former allows the users to try their own speech while the latter allows them to answer questions sent by other people.

4.2 Audience

The audience is divided in three different group:

- Kind classmates & colleagues: with causal wearing and friendly smile, they will listen with smile, nod and praise the user during the speech.
- Indifferent people: with strange clothes, they pay no attention to the speaker and will look around, speak to others and may fall asleep during the speech.
- Experts: with formal suit and serious expression, they will listen without emotion, shake head and get angry during the speech.

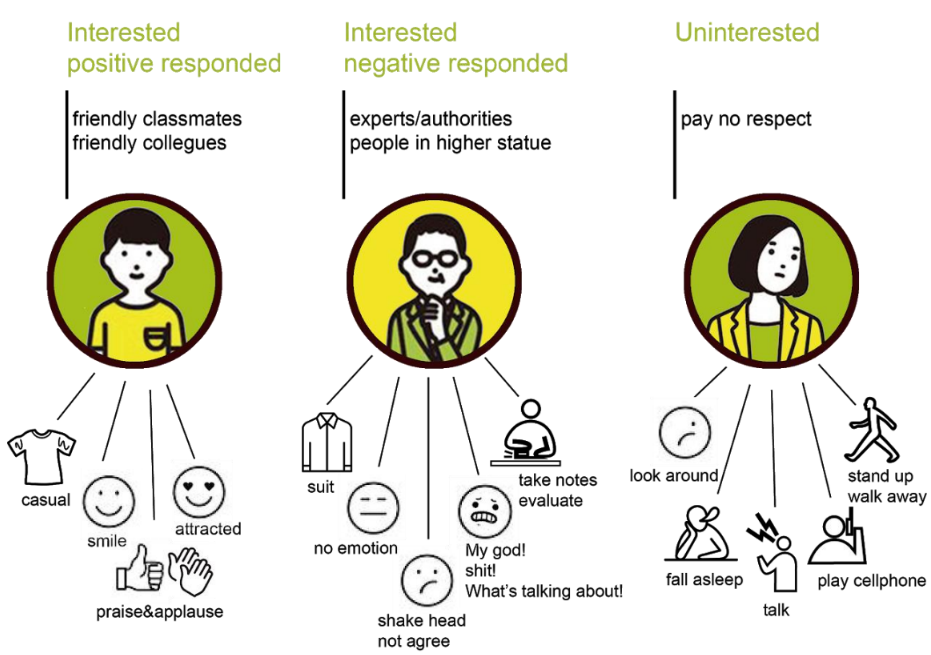

Figure 2: Audience characteristics

The composition of the audience can be chosen before the speech so that the user can choose the challenge he/she want to take on.

Figure 3: Example of compositions

4.3 Biosensor

The biosensor is used as a mean to track the level of anxiety of the user and it allows to manipulate the environment based on the values read:

- The number of people in the audience will change throughout the speech. This is achieved by using lights to hide or show more people in the audience;
- The speech can be interrupted if the anxiety level determined by the biosensor increased too much.

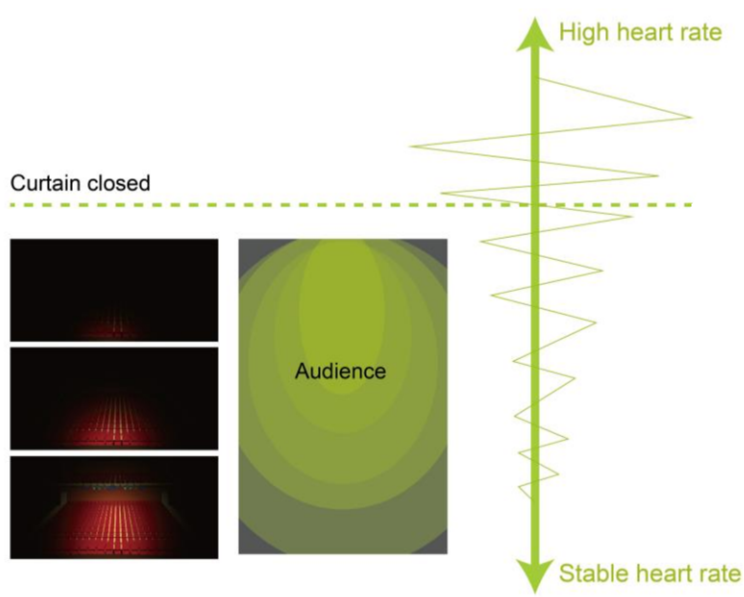

Figure 4: How the lights change based on the biosensor values.

4.4 Process

The application is divided into three stages:

- Menu where the user can decide the settings for the speech:
  - Biosensor (through a QR Code scanner)
  - Speech mode
    - Free talk
    - Interview
  - Audience composition
- Heart rate test where a baseline value for the heart rate is searched: it will be used as reference to determine the state of mind of the user;
- VR activity where the user is put into the virtual environment he/she can try the speech.

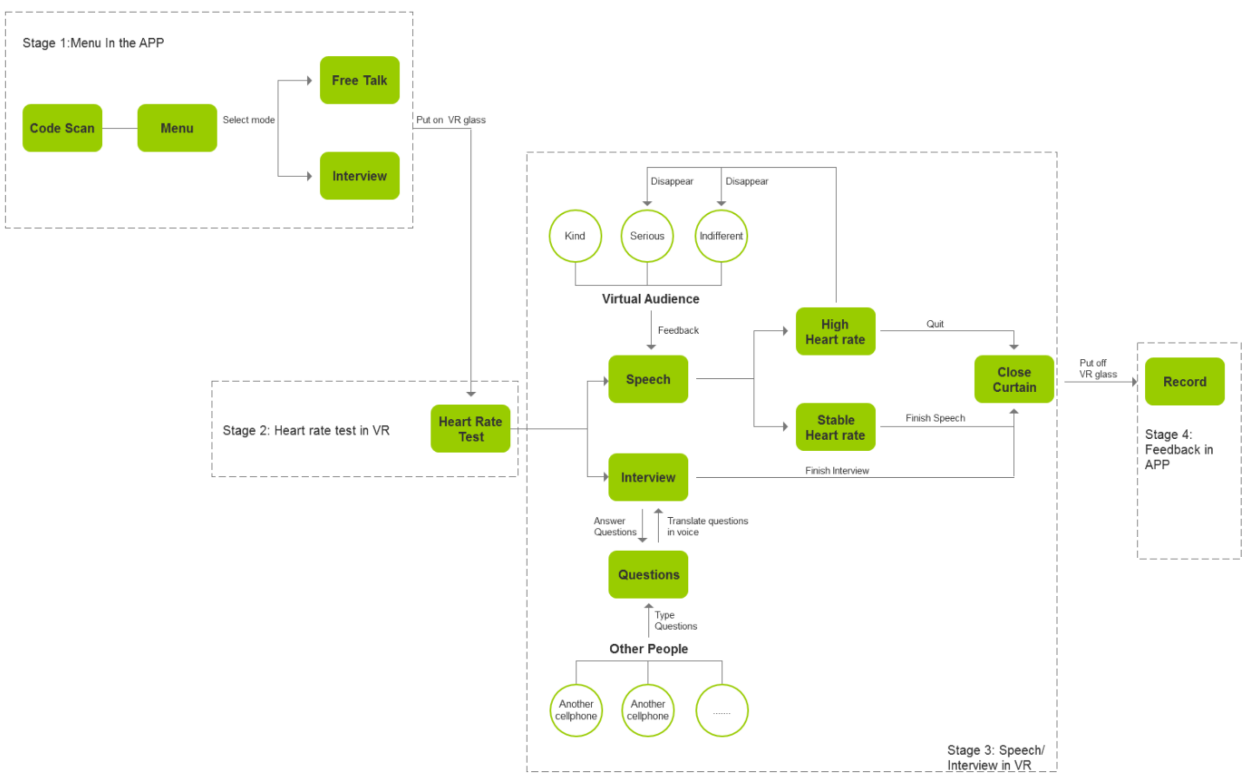

Figure 5: Flow of the application

4.5 Interface

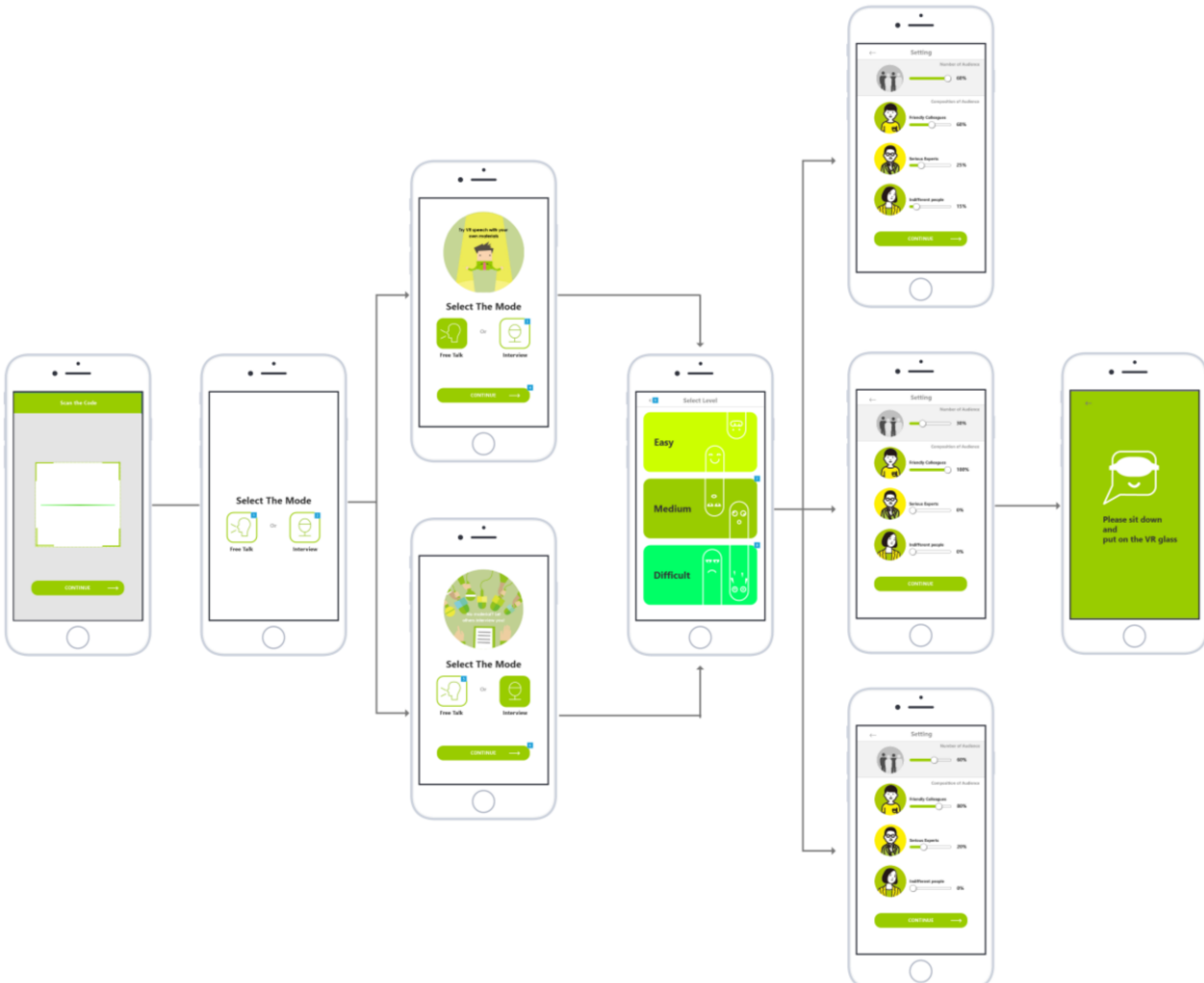

Figure 6: Stage 1: App

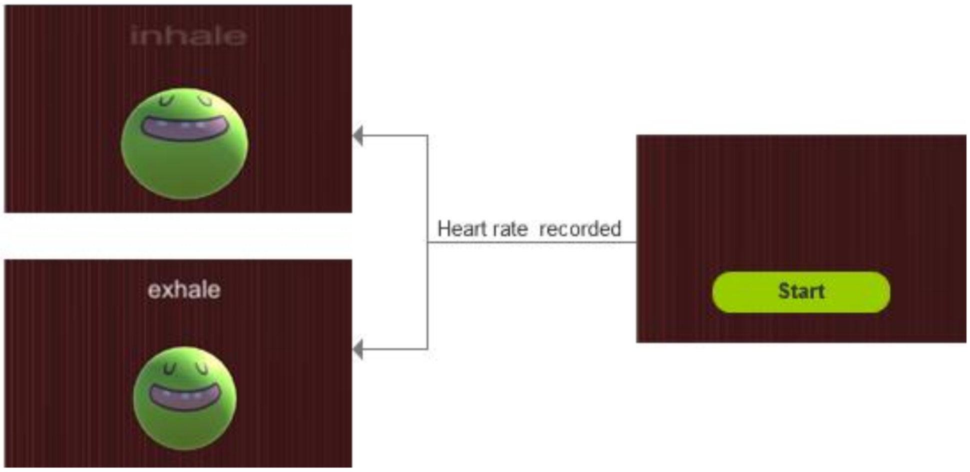

Figure 7: Stage 2: Heart rate test

Figure 8: Stage 3: VR activity

4.6 Scenario

4.6.1 Scenario 1

Paolo has been working on a project for a course at the university. He has to present it to the professors at the end of the semester together with his teammate Roberto. He is really nervous about it and so it decides to try a new application he found recently. Paolo asks Roberto to help him coming up with possible questions that the professor could ask. Paolo puts on the Empatica E4 on his wrist and connect it to Roberto's phone and he starts the main application on his smartphone. He then follows the instruction: he scans the QR code that appears on Roberto's phone, selects the interview mode and then selects the easy difficulty without changing the suggested value.
He puts on the HMD on and starts with the first activity suggested to him. When the activity finishes, he starts answering the questions that Roberto sent.

4.6.2 Scenario 2

Giulia has to hold a speech in front of an audience for an important event and she is really afraid to not be able to do it correctly. A new application that came out recently was suggested to her by her friend to try her speech and see how well it goes.
He decides to follow the suggestion, and so she downloads the app, she puts on the Empatica E4 and follows the instruction on screen: she scan the QR code, she select Solo mode and select the easy difficulty but still change some of the preset values. She puts on the HMD on and starts with the first activity. When it finishes, she starts her speech.

5 Implementation

5.1 Introduction

The whole project is divided into different applications:

- Android application built on Unity
  It's the core application of the project that is going to run all the operation required to show and manage the virtual environment shown to the user;

- Android application
  It's one of the possible way to read the data from the biosensor.

- Pc application
  It's one of the possible way to read the data from the biosensor.

- Firebase Server
  Allows the communication between the pc client and the Firebase database.

Unfortunately, Unity doesn't offer a simple way to retrieve the data directly from the biosensor and this is the reason that the two alternative application had to be made. In order to make to communication as simple as possible, it was decided to take advantage of Firebase as a way to store and retrieve data thanks to its database.
So, as shown in figure 9, the information from the biosensor are read first by a Computer or an Android device and then sent to a Firebase server that stores the values.

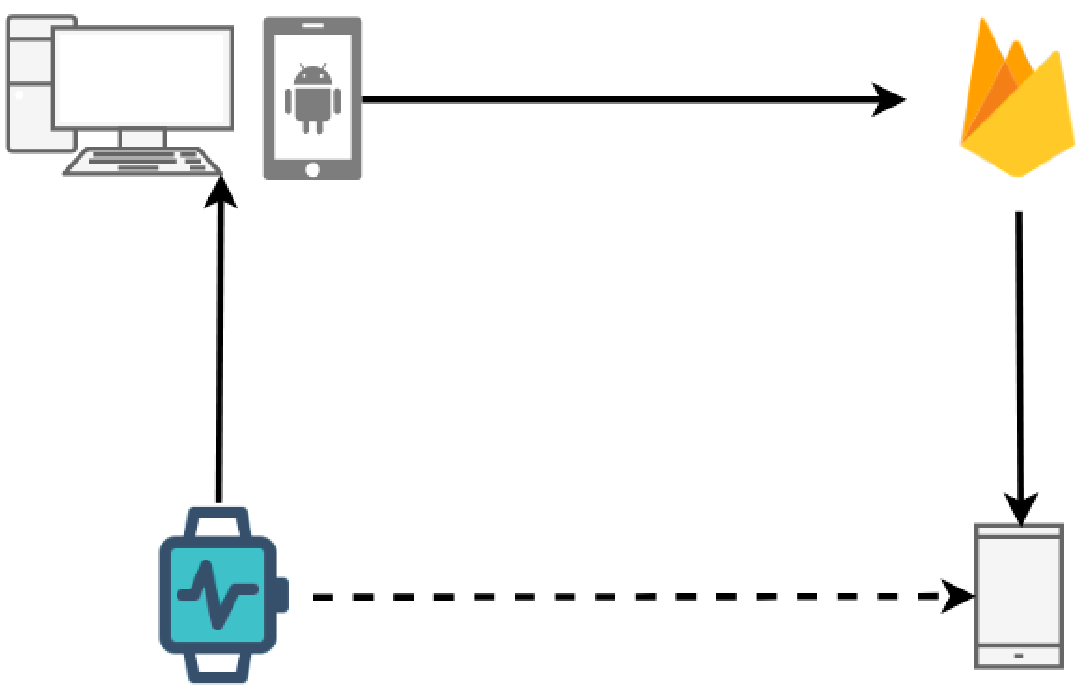

Figure 9: Diagram that shows how the communication from the biosensor to the smartphone works.

One problem that arises with this choice is the need for authentication so that the Unity application is always connected to the right biosensor.
To solve this problem, the authentication is done by using QR codes so that the device connected to the biosensor generates a QR code that allows a smartphone to "connect" to it.
This solution works with the assumption that the user is near the device that displays the QR code. This is a safe assumption as the biosensor is connected by bluetooth to that device and so the user shouldn't be too far from it.

5.2 Main Android Application

Language used: C#
Plugins & Libraries:

- ZXing
- Android Runtime Permissions
- Speech Engine
- FirebaseDatabase

5.2.1 Description

This is the core of the project and it's split in three parts:

1. QR code reader that allows to connect the application to the right "room" where it can retrieve the required data as stated before;
2. Main menu that contains all of the settings available for the application;
3. VR application that represent the main component.

In order to accomplish all the task required, the application takes advantages of the functionalities offered by Unity and by the additional libraries to:

- Read QR codes thanks to ZXing;
- Communicate with Firebase Realtime Database thanks to the respective plugin;
- Convert Text-to-Speech thanks to Speech Engine;
- Ask the required permissions to execute operations thanks to Android Runtime Permissions;
- Track the activity of the microphone thanks to the functions available on Unity.

In order to keep the algorithm as simple as possible, most of them use a counter to decide whether the have to perform an action or not. Usually the algorithm is:

1. Read value;
2. Analyze the value read;
3. Increase or decrease the counter based on the analysis;
4. Check whether the counter is greater or lower than a set threshold;
5. Perform the action if the threshold was exceeded.

This system is used in many algorithm (as shown in the next section) as a way to manage the environment and decide which action should be performed.

5.2.2 Algorithm design

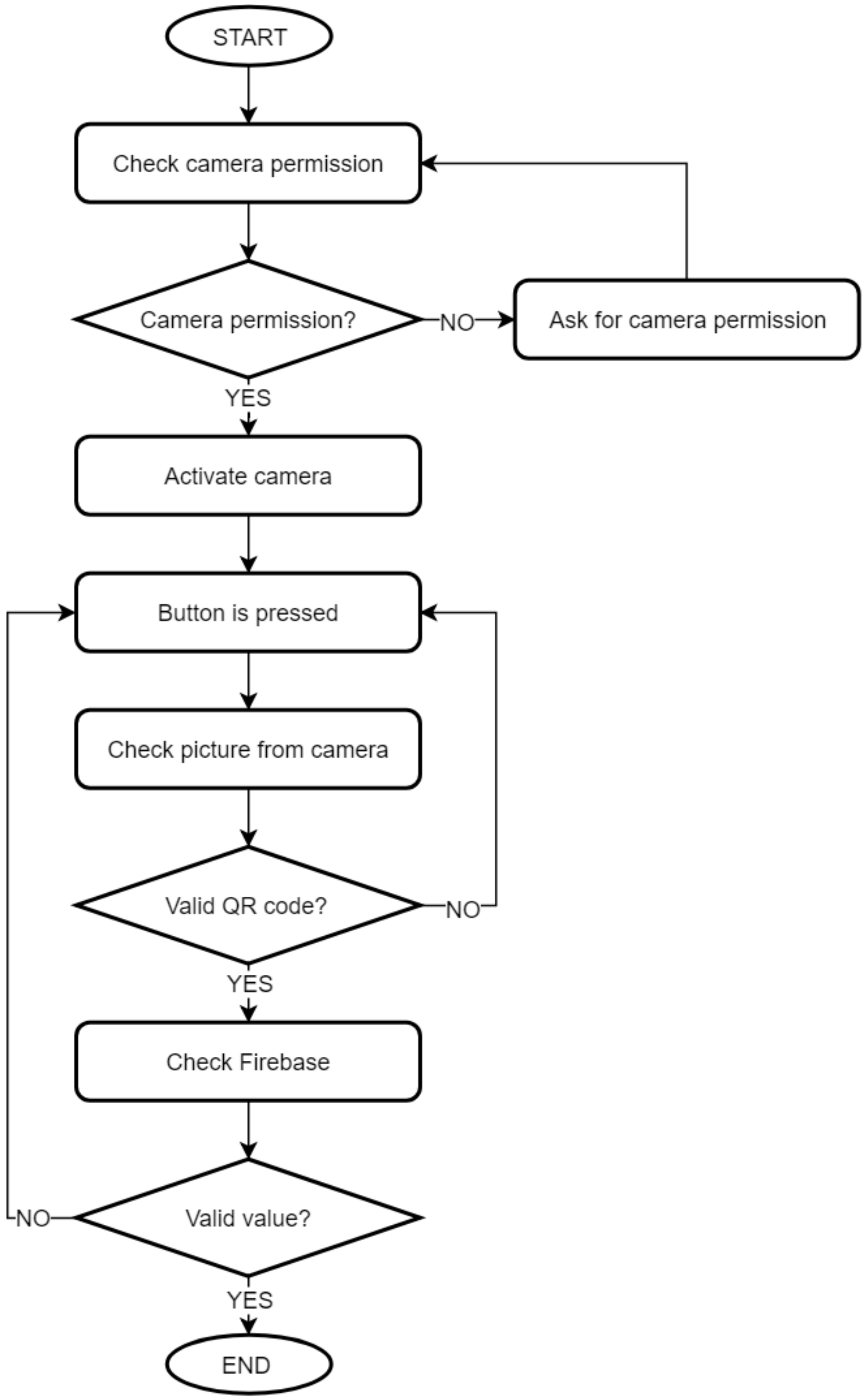

Figure 10: Flowchart that describes how the connection to the right "room" in the Firebase server works

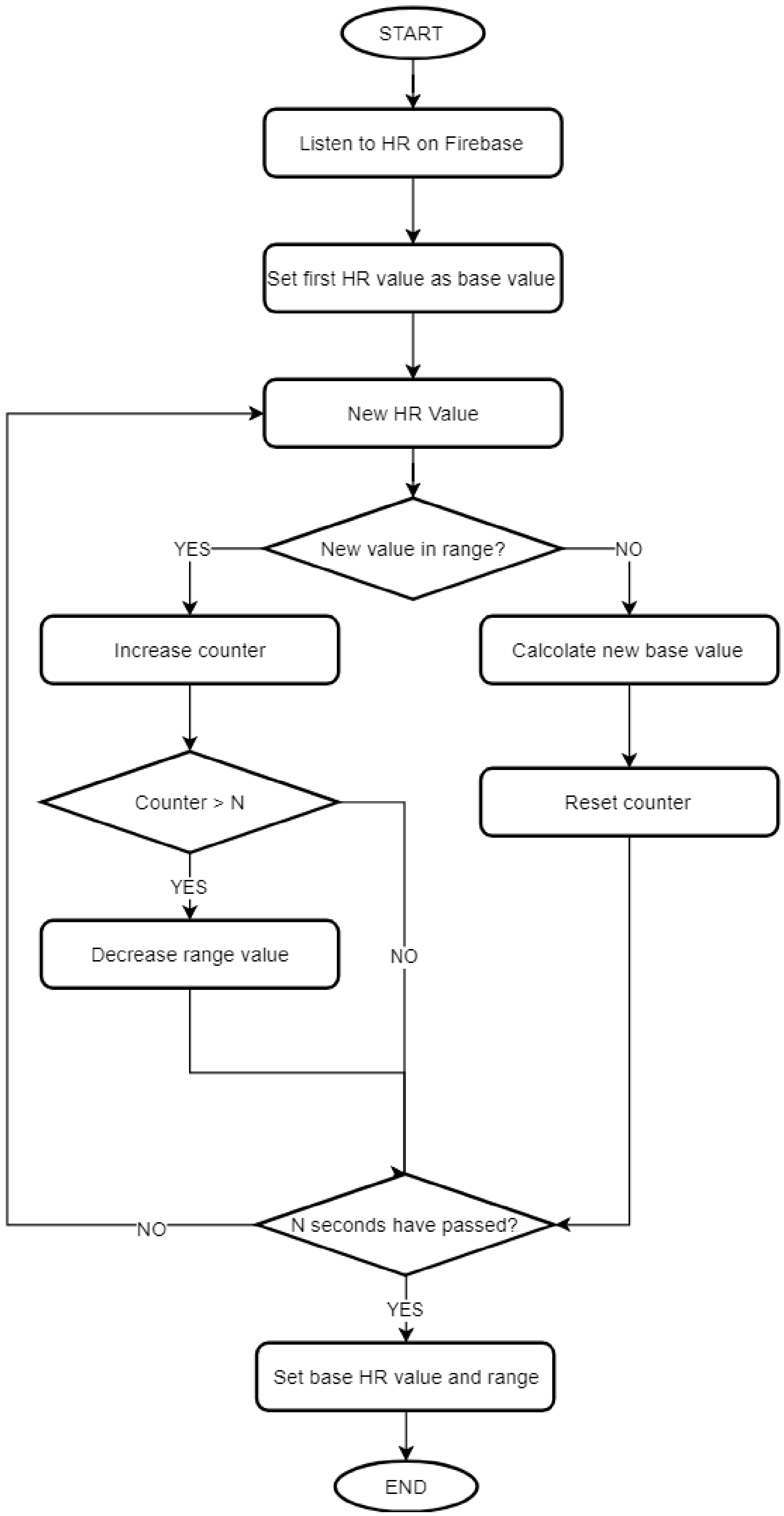

Figure 11: Flowchart that describes how the search of the base value for HR and GSR works. The two operations are executed in two different coroutines.

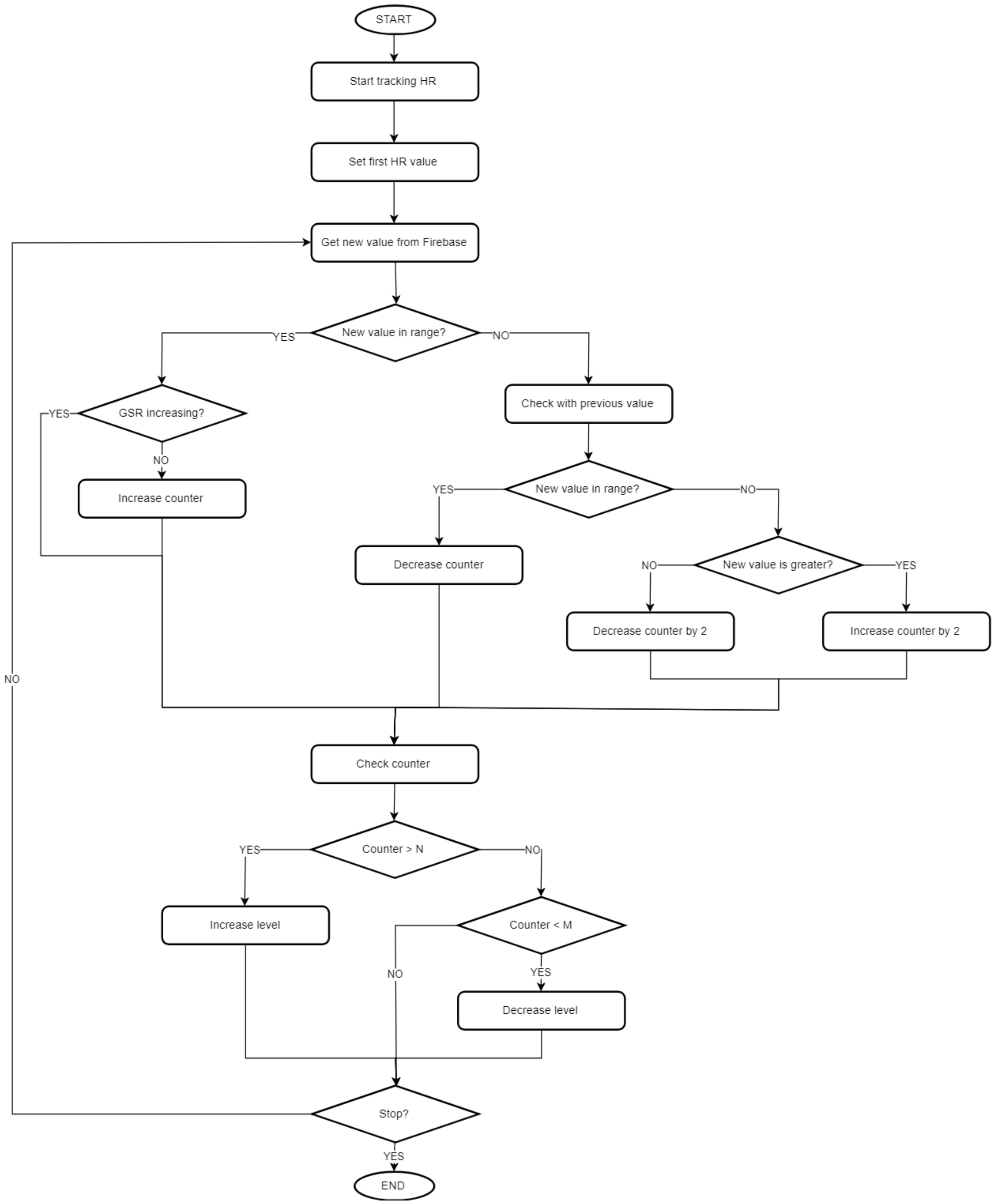

Figure 12: Flowchart that describes how the HR and GSR are used to determine whether to increase the level or not.

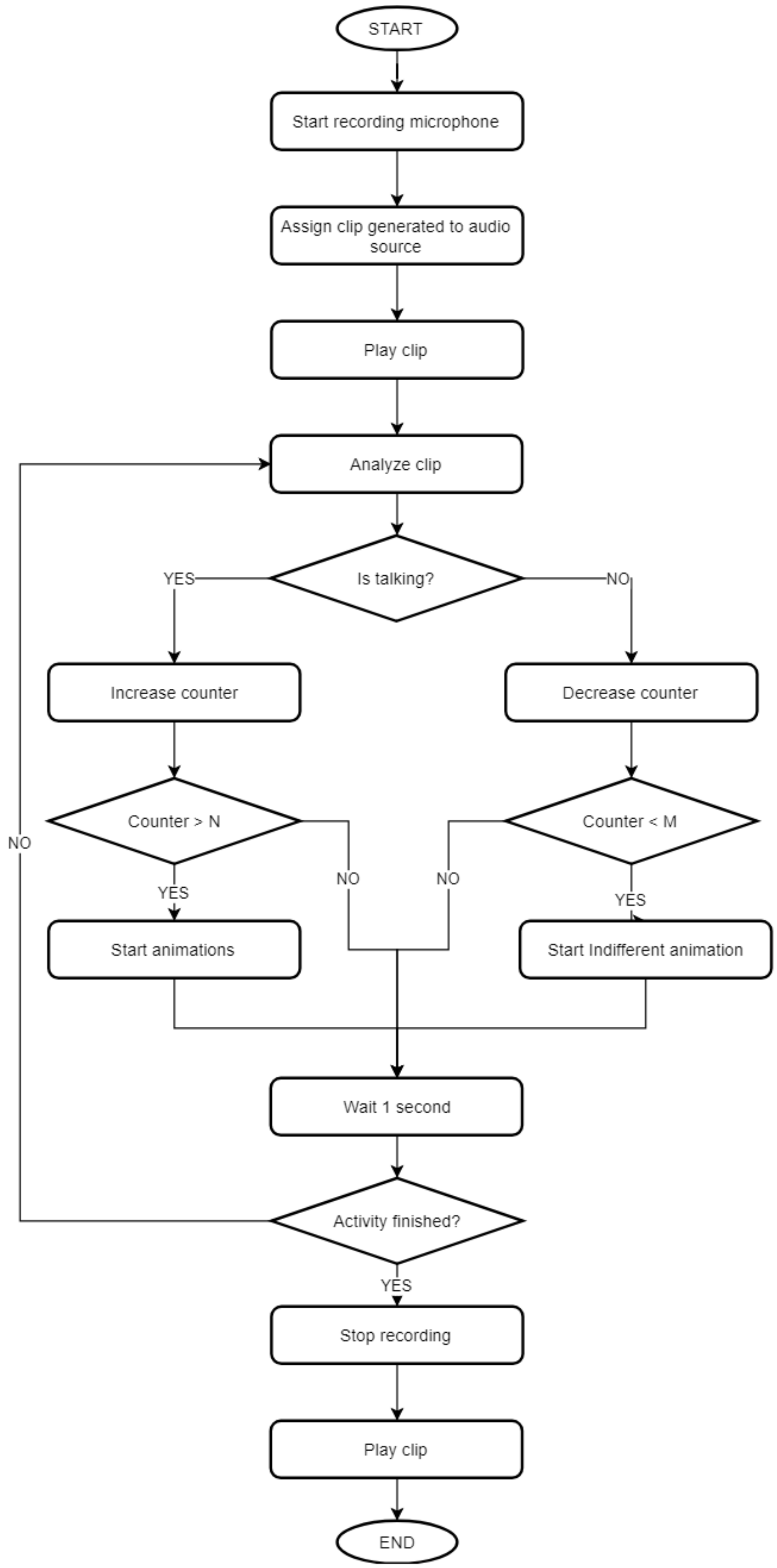

Figure 13: Flowchart that describes how the microphone is used to determine whether the animations should be played or not.

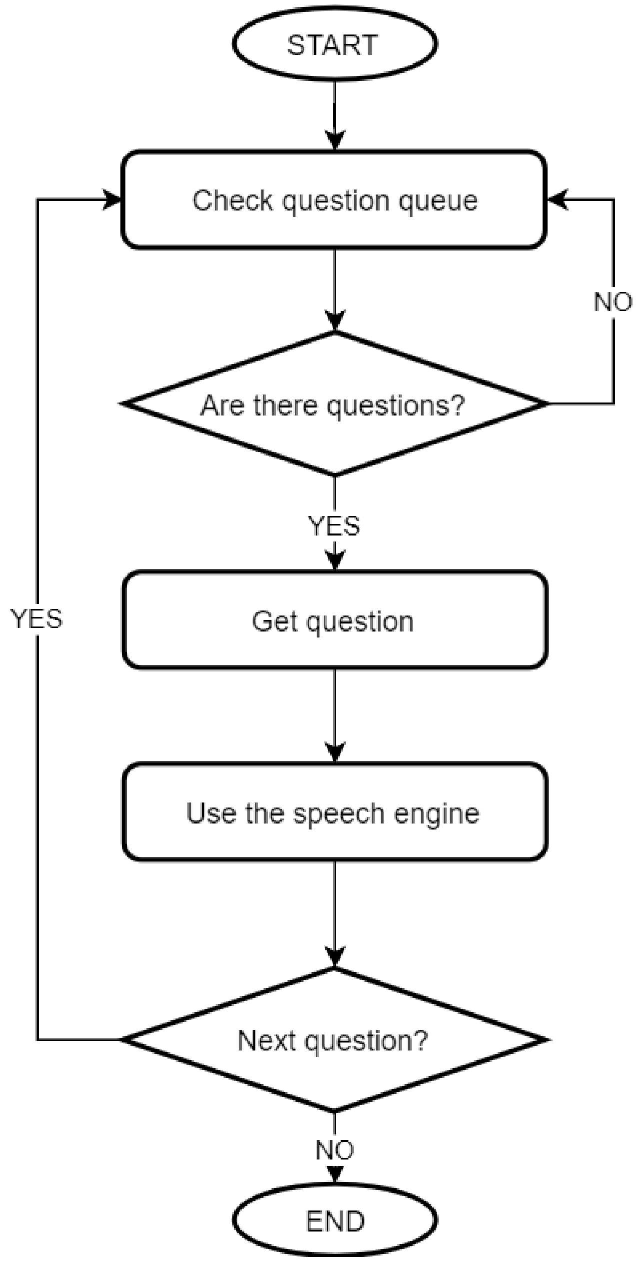

Figure 14: Flowchart that describes how question system works when the user doesn't want to get the questions from the audience.

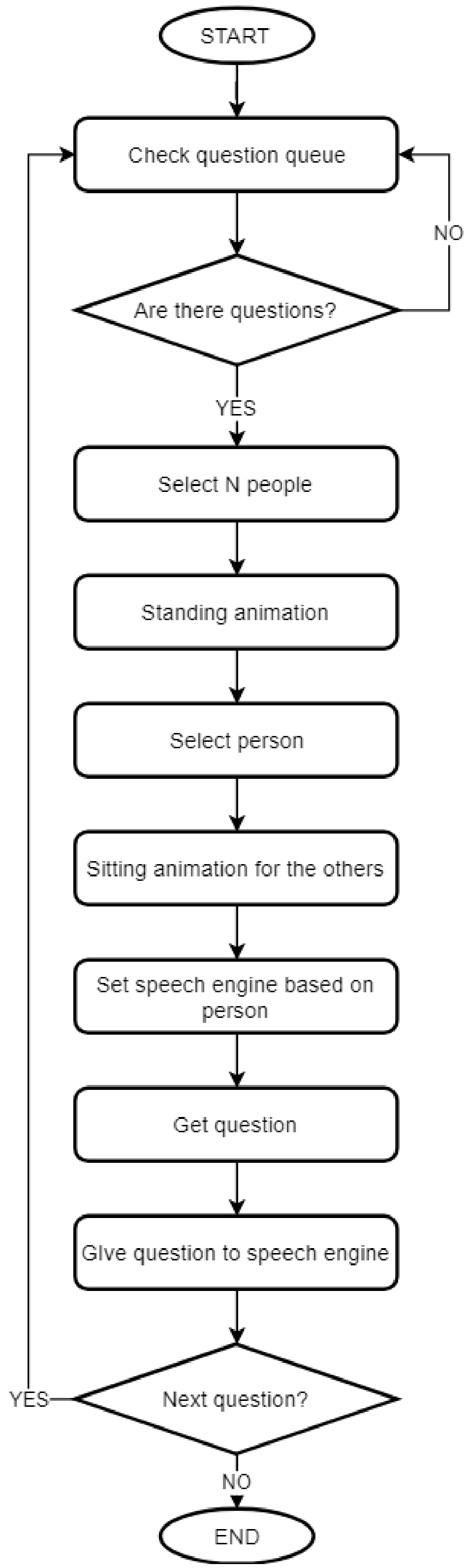

Figure 15: Flowchart that describes how question system works when the user want to get the questions from the audience.

5.3 Computer Client

Language used: Java
Plugins:

- ZXing
- JavaFX

5.3.1 Description

The computer client main task is to retrieve data from the Empatica E4 and send the values to the Firebase server. In order to do this, it communicates with E4 streaming server, an application that allows to forward realtime data of multiple Empatica E4 devices to multiple TCP socket connections.
The E4 Streaming server works through a message protocol where client request are in the following format:

COMMAND ARGUMENT_LIST

Messages from server containing responses to commands are in the following format

COMMAND ARGUMENT_LIST

Messages from server containing data from device are in the following format

STREAM_TYPE TIMESTAMP DATA

The commands used from the client are:

- device_list
  requests the list of Empatica E4 devices to the E4 Streaming server
- device_connect DEVICE_ID
  sends a connection request to a specific device
- device_subscribe STREAM STATUS
  start or stop receiving data from a given stream.
- device_disconnect
  sends a device disconnection request

5.3.2 Algorithm design

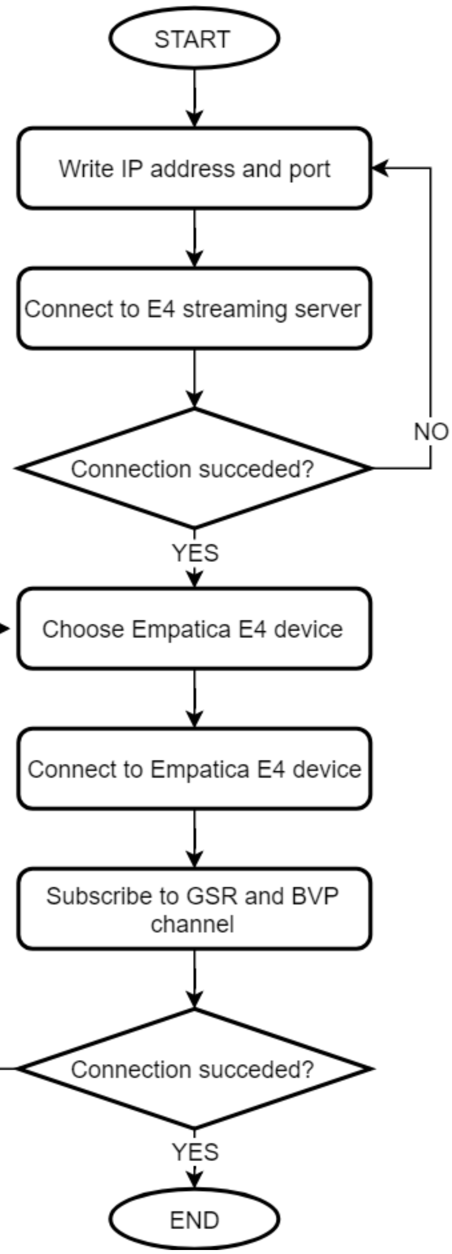

Figure 16: Flowchart that describes how the connection to the Empatica E4 device works

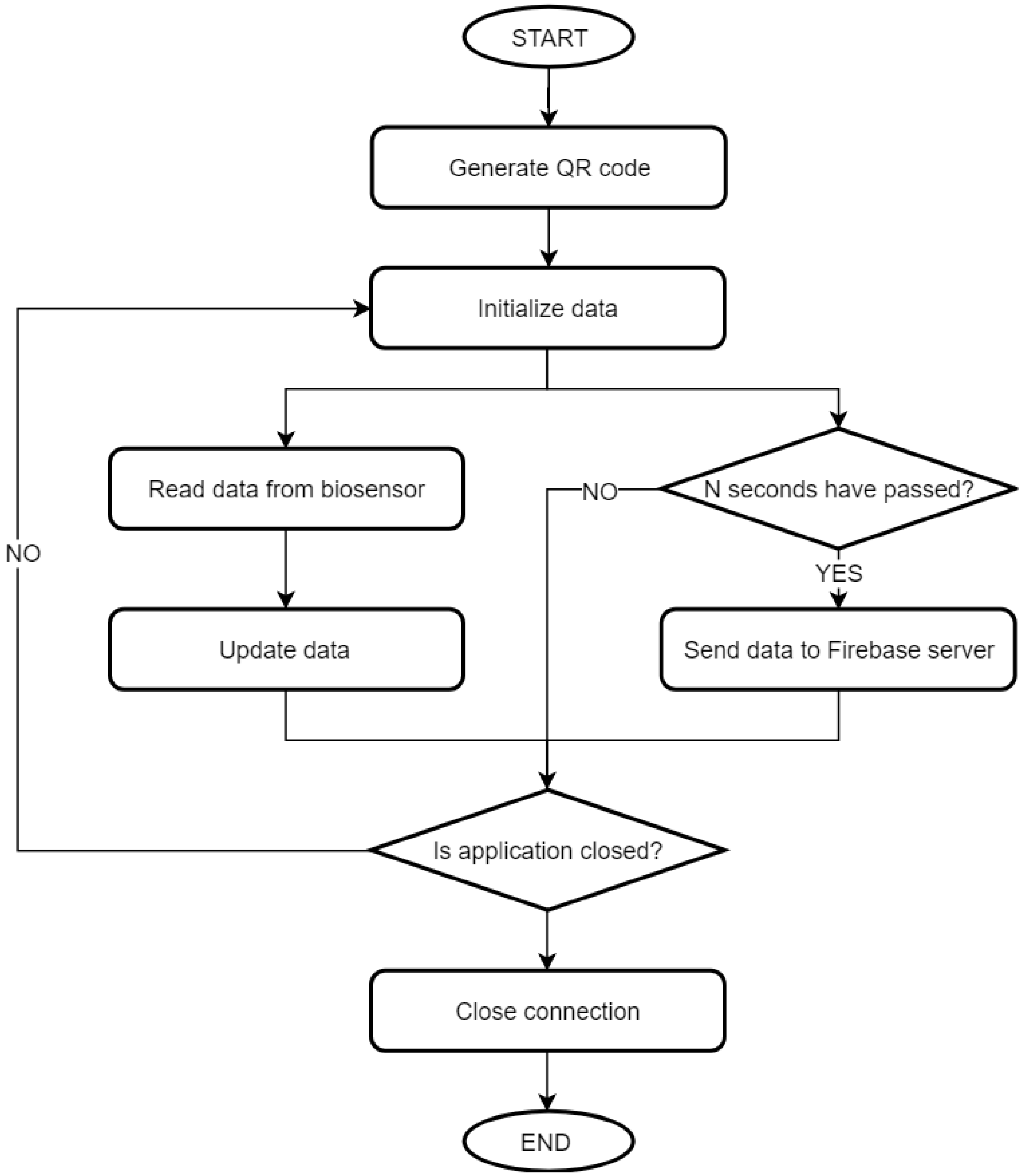

Figure 17: Flowchart that describes how the client retrieve data from the Empatica E4 after the connection has been established

5.4 Firebase server

Language used: Javascript

5.4.1 Description

Firebase is a platform that offers the possibility develop web application easily. It was chosen as the backend for this project because of its ease of use.
The main objective of the Firebase application is to store the data retrieved from the biosensor and then send them when requested.
The services used on Firebase are:

- Realtime Database.
- Cloud functions;

The Realtime Database is a NoSQL database were data is stored as JSON. It is used to store the value read by the biosensor and the questions that have to be sent to the mobile application.

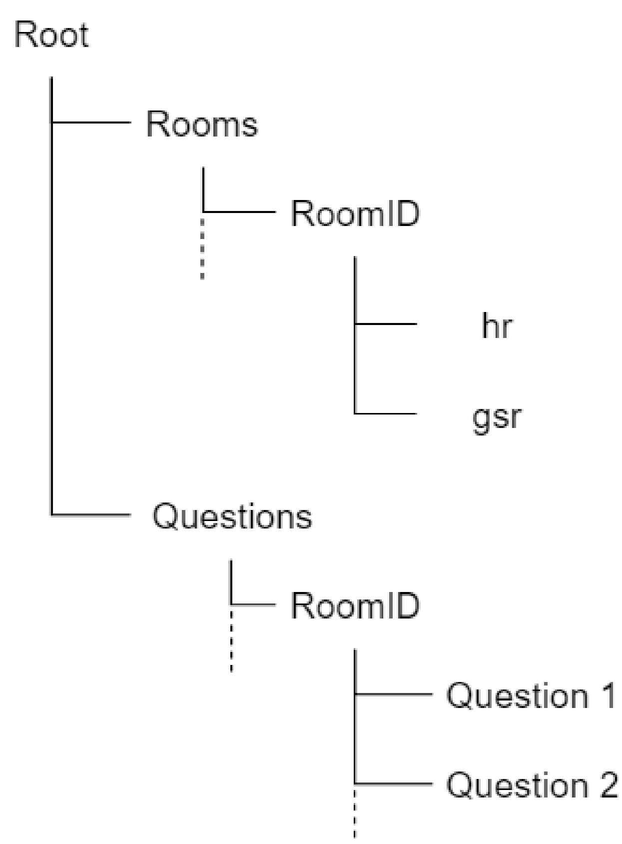

Figure 18: Structure of the Realtime Database

Cloud functions are used in order to let the PC client access the database. The Cloud functions are tasks that can be triggered by HTTP requests. The values required are sent as a JSON string in the body of the request.

- createRoom: creates a room with the given room code;
- updateValues: update the values of the hr and gsr stored in the given room;
- addQuestion: add a question to the list in the given room;
- removeRoom: Closes the room with the given room code.

5.5 Android Application

Language used: Java Plugins & Libraries:

- ZXing
- AndroidImagePopup
- FirebaseDatabase
- Empalink

5.5.1 Description

The Android application main task is to retrieve data from the Empatica E4 and send the values to the Firebase Database. in order to do this, it communicates with the biosensor thanks to the Empalink library developed by Empatica that allows to read the data directly from the Empatica E4 without the need of another application (unlike the pc application that needs the E4 streaming server.

5.5.2 Algorithm design

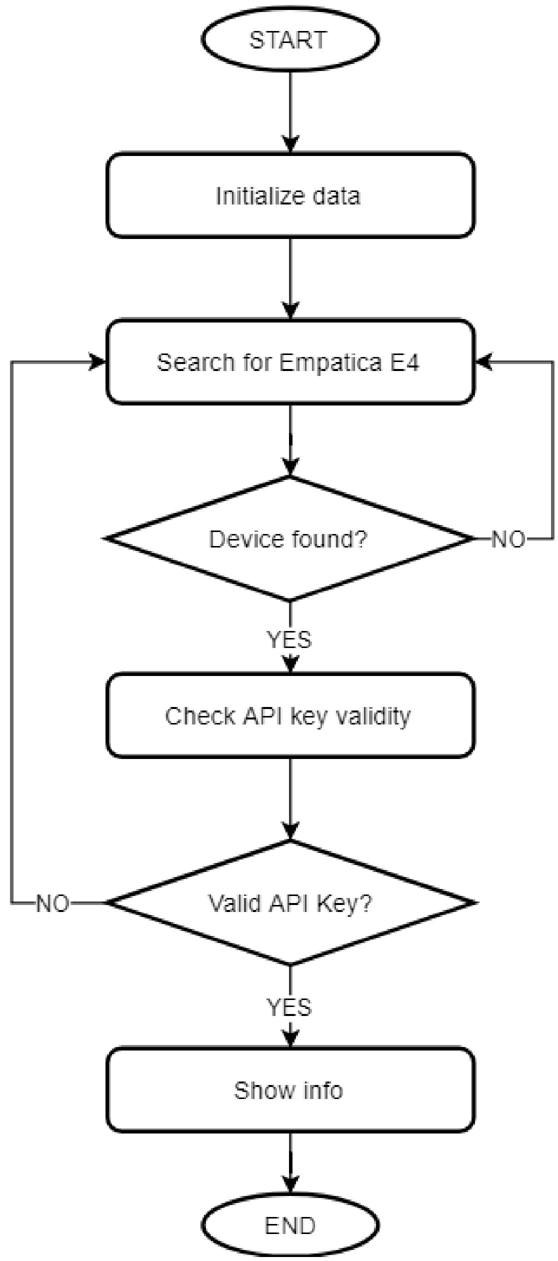

Figure 19: Flowchart that describes how the client connects to the Empatica E4 biosensor

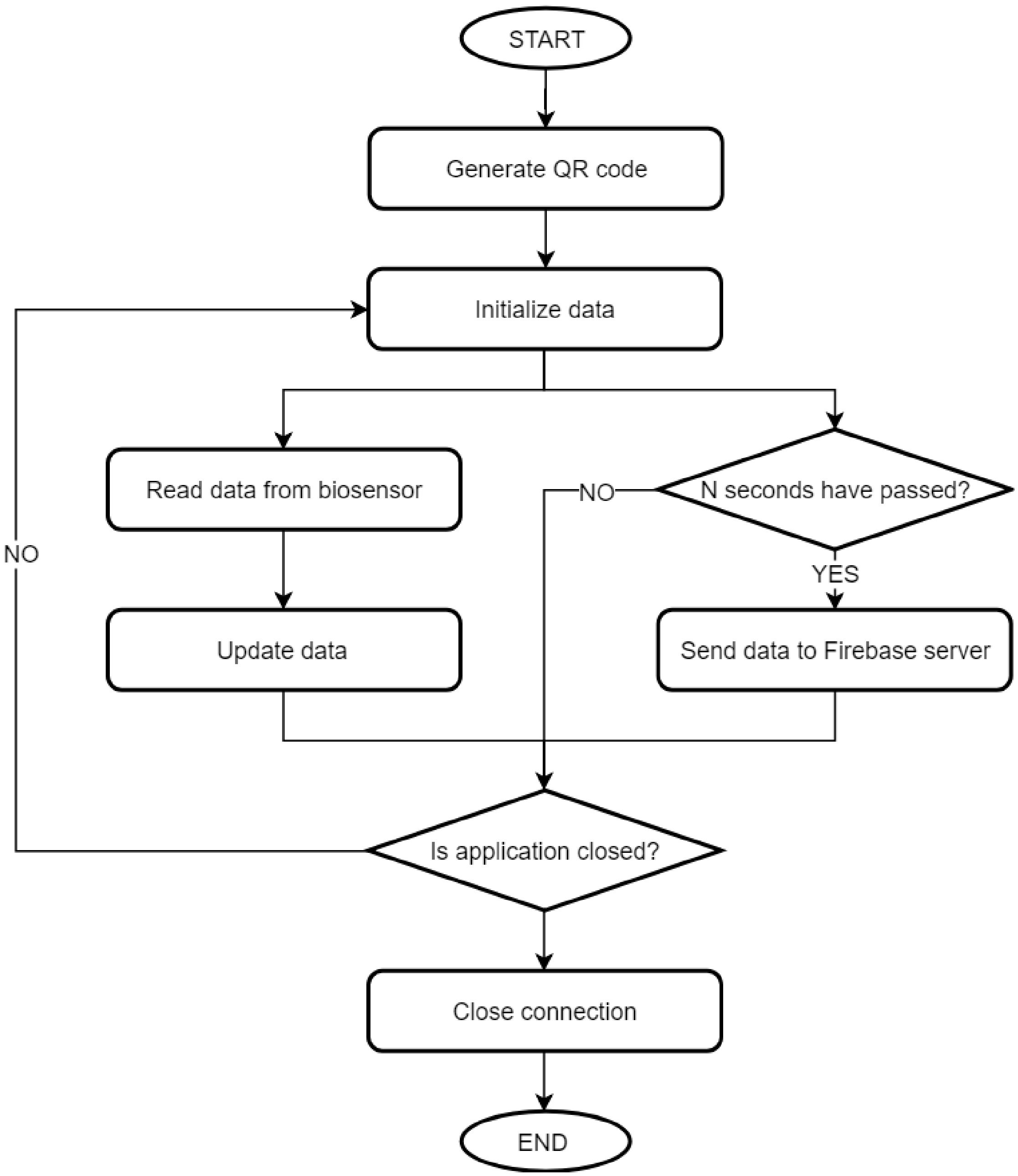

Figure 20: Flowchart that describes how the client retrieve data from the Empatica E4 after the connection has been established (same as the pc client)

6 Value proposition

6.1 Challenges

The main challenges that came up during the development of the project were:

- finding a way to put the subject in a "controlled" stressful situation without leaving him/her in an anxious state;
- giving instruction to the user so that he/she won't stay silent while in front of the virtual audience.

6.2 Main difficulties

- Even though Android devices are able to run VR application, they can only be used to display simple scenery, games with a limited amount of polygons or 360° videos. In fact, the limited resources available on a smartphone makes it difficult to develop VR applications that can run without performance problems. Because of this, during development some choices had to be adjusted or changed in order to make the application run. Unfortunately, this was not enough. In fact, the smartphone used to test the application (Huawei P10 lite) was not powerful enough to handle it correctly.

- The biosensor used (Empatica E4) isn't the best fit for the purpose of this project. It is able to track data in real time but even the smallest movement is enough to disrupt the readings, leading to either wrong values or no value at all. Also, the Empatica E4 doesn't track the HR directly, it needs to take the Inter-Beat Interval (ibi) and convert it into HR.

$$HR = \left\lfloor \frac{60}{ibi} \right\rfloor \tag{1}$$

6.3 Analysis

The effectiveness of VR as a mean to ease the anxiety of people that have the fear of speaking in public is a fact reported in many researches: it doesn't really improve the people's ability to talk to an audience but it helps them gaining enough self confidence so that they can gather the courage to face the audience.
As stated previously, there are many applications that allows the user to deal with this kind of fear but none of them uses a biosensor as a mean to manage the environment, instead they use it as a way to measure the overall score of the performance. This is what makes this project different: even though the base application is the same as the other, it offers an experience that changes based on the values read from the biosensor.

7 Future work

In the future this project could be extended in many different ways:

- The Empatica E4 biosensor could be changed with a more specialized one that could track parameters in a more precise way. One option could be the usage of smart clothing that offer the possibility to track both HR and GSR while also allowing the tracking of the breathing activity and many other parameters. This choice would also allow to have more precise data that should be less influenced by the movement compared to the Empatica E4.

- As stated before, the application is not well optimized to run on an Android device, in fact it runs at less than the required 60 FPS. Unity released a new feature with Unity 2018.3: lightweight rendering pipeline (LW RP). As reported on the Unity blog, "The goal of the LW RP is to provide optimized real time performance on performance constrained platforms by making some tradeoffs with regard to lighting and shading."[1]. This feature is certainly not enough to achieve a good result but it may be a starting point to improve the overall performance of the application.

- All the algorithm related to the biosensor and microphone were made just to show the general functionality of this project and so they are really simple. In the future, these could be improved in order to have a better control over the environment and the situation the user is put in.

[1] https://blogs.unity3d.com/2018/02/21/the-lightweight-render-pipeline-optimizing-real-time-performance/

References

[1] Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. Behaviour research and therapy, 35(8), 741-756.

[2] Slater, M., Pertaub, D. P., & Steed, A. (1999). Public speaking in virtual reality: Facing an audience of avatars. IEEE Computer Graphics and Applications, 19(2), 6-9.

[3] Pertaub, D. P., Slater, M., & Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. Presence: Teleoperators & Virtual Environments, 11(1), 68-78.

[4] Chollet, M., Sratou, G., Shapiro, A., Morency, L. P., & Scherer, S. (2014, May). An interactive virtual audience platform for public speaking training. In Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems (pp. 1657-1658). International Foundation for Autonomous Agents and Multiagent Systems.

[5] Poeschl, S., & Doering, N. (2012, March). Virtual training for Fear of Public Speaking—Design of an audience for immersive virtual environments. In Virtual Reality Short Papers and Posters (VRW), 2012 IEEE (pp. 101-102). IEEE.

[6] McKinney, M. E., Gatchel, R. J., & Paulus, P. B. (1983). The effects of audience size on high and low speech-anxious subjects during an actual speaking task. Basic and Applied Social Psychology, 4(1), 73-87.

[7] Bernardi, L., Wdowczyk-Szulc, J., Valenti, C., Castoldi, S., Passino, C., Spadacini, G., & Sleight, P. (2000). Effects of controlled breathing, mental activity and mental stress with or without verbalization on heart rate variability. Journal of the American College of Cardiology, 35(6), 1462-1469.