Project Enhanced Audio Recognition

Joshuan Jimenez

July 7th, 2019

Contents

[Summary 2](#_Toc13434054)

[Introduction 3](#_Toc13434055)

[Needs/Problems 4](#_Toc13434056)

[Goals/Objectives 4](#_Toc13434057)

[Procedures/Scope of Work 4](#_Toc13434058)

[Timetable 5](#_Toc13434059)

[Budget 5](#_Toc13434060)

[Evaluation 5](#_Toc13434061)

[Conclusion 6](#_Toc13434062)

[Appendix 7](#_Toc13434063)

# Summary

Randy Rivas and I, Joshuan Jimenez, are two musicians who write music and are interested in music technology. The purpose of this proposal is to enhance audio recognition by creating an application that allows users to identify songs without having to play it. This would allow the user to hum a melody, answer questions, and listen to a clip of the resulting song to see if it is what the user was looking for.

When the user hums the melody, the application will analyze the acoustic fingerprint of the user’s humming and compare it to the song files on the database. Based on this, the results will be narrowed down. Then, by implementing an algorithm, questions concerning the song’s genre, rhythm, and more will be asked and the user will reply what they remember. During this questioning, the application may play audio clips of certain rhythms, beats, or lyrics for the user to listen to and decide if it is familiar. In the end, a clip of the end result song will play, and the user will say if the song was found or not.

From this, we would like to see that the application has been downloaded at least 10,000 times and has a success rate of 95% in finding the desired song during a live-testing session. This live-testing session will show if the application serves its function in the world and is performing well after the development and testing process. All that is needed for this project to be undertaken is a budget of approximately $82,000 to cover wages and advertising. This would allow the application to be developed, as tools and software are already taken care of. After developing and testing on our own, we would then be able to market the application to be live tested.

# Introduction

Me and my partner Randy Rivas would like to enhance audio recognition as we are musicians who work together on writing music and performing it as well. I am studying computer science and have developed some applications since 2014. I have been performing and writing music since 2010 as well. Meanwhile, Randy Rivas studies Music Composition, writes music literature, and assists in music technology projects since 2011.

We strive to help the music industry in any way possible, due to our admiration for the arts. As musicians, we actively perform and work on projects to write music for whenever we have the time for. Me and my partner have known each other for a few years, but they have been a wonderful few.

Over our history, we have written music for a variety of genres, such as Rap, Classical, Jazz, and Latin. We also listen to a multitude of artists to gather inspiration from and try to fuse genres to create unique music fit for any project.

We use a variety of tools to write our music, some of which are Sibelius, Finale, and FL Studio. Using these tools daily makes us more familiar with the quirks and benefits of each software, which is why we use a huge variety of them. As musicians, we make sure to provide high quality music with only the best software that can be used, while checking and making sure to use appropriate software based on genre.

# Needs/Problems

A common issue seen by many teenagers and adults every day is hearing a song on the radio or at an event and not being able to identify it at the moment. Applications exist to do this for us, but many times we are not able to open it due to not being able to pull out our phones in time. Then, we are stuck wondering what the name of that song was or who was the artist. A common issue that has been around since music has started to be streamed over radios and the internet.

Attempts to remedy this problem are available on the internet but are not really made for use on a person’s phone. Websites like Midomi and Musipedia assist with this, but Midomi requires users to sift through various song guesses, which can be faulty at times, as well as being outdated at times. Musipedia does the same but is more inclined for musicians who know how to use a piano or spell out a rhythm.

To fix this issue, me and Mr. Rivas have decided to create an application that enhances audio recognition and allows people to find songs that they listened to, but do not know information about. For this project to be undertaken, we would need to address what to focus on:

• Allows users to hum a melody

• Allow users to be asked questions about song and artist

• Allow users to hear clips of sounds and compare to desired song

• Tell user information about song that is being looked for

# Goals/Objectives

For this project, we strive to have around 10,000 downloads within 3 months of releasing the application. Along with this, we are striving for a 95% success rate in finding the correct song for the user after 3 months of releasing the application.

# Procedures/Scope of Work

Based on the needs mentioned before, I have devised an outline of what we would need to be able to meet them:

• Create or use an existing database full of songs, containing information and acoustic fingerprint of the songs

• Create an application with an interface that allows users to hum a melody

• Take user input (humming) to then link to similar sound files in database through acoustic fingerprinting

• Create a decision tree algorithm with a list of questions to narrow down search results based on song’s genre, style, and singer

• Provide audio clips when necessary to user to narrow down possible choices

• Play audio clip after asking enough questions to see if the song is the right one

• Allow user to say if the result is the right song

# Timetable

The project will span over six months and two weeks, split into three phases as shown:

|  |  |  |
| --- | --- | --- |
|  | **Description of Work** | **Start and End Dates** |
| **Phase One** | Planning and Design | August 1st 2019 – August 15th 2019 |
| **Phase Two** | Coding, Development, and Testing | August 16th 2019 – November 16th 2019 |
| **Phase Three** | Release, Marketing, and Live Testing | November 16th 2019 – February 16th 2020 |

# Budget

When it comes to funding, minimal funding is required for the salaries for the personnel and advertisement of the application. It is broken down in the tables below:

|  |  |
| --- | --- |
| Wages and Salaries | |
| Software Engineer (Joshuan Jimenez) | $40,000 |
| Software Developer (Randy Rivas)  Subtotal | $40,000  $80,000 |

|  |  |
| --- | --- |
| Advertisment during Phase 3 (3 months) | |
| Social Media (Twitter, Instagram, Facebook) | $1,000 |
| TV Advertisment  Subtotal | $1,000  $2,000 |

|  |  |
| --- | --- |
| Total | $82,000 |

# Evaluation

Throughout this project, progress will be documented through notes. This will be useful during Phase 1 and 2, as we will make a weekly log concerning how far into the development of the application we are. Documenting will include notes on planning, designing, coding, current errors and bugs, and improvements.

Most of the documentation will be done by me mainly, unless noted to be written by Randy Rivas. We will provide key fragments of the code to show progress in the notes while going into depth as to why it is needed. Along with this, we will occasionally send out test videos to display the functionality of the application along with the notes.

During Phase 3, upon releasing the application, we will be doing the same thing, documenting on a weekly basis the statistics concerning the program. Statistics can include total downloads, downloads per day, success rate with finding songs, and advertising analytics. This will allow us to visualize and predict whether we will meet our desired goals or not. These predictions will be shown on graphs and charts when possible.

Positive indicators of progress being fulfilled are high downloads, projected downloads showing goal of 10,000 downloads being made, and meeting the success rate when finding songs (above 95%). Negative indicators are low downloads, projected downloads showing goal of 10,000 downloads not being made, and not meeting the success rate when finding songs (below 95%)

# Conclusion

This six month and two-week project is aimed towards improving audio recognition while fixing a common issue people face. By being able to address this issue, we will be able to assist the music industry and possibly come up with a new application that can change how people search for music.

A project like this may seem to be heavy, but rest assured, if the funding is provided, me and Mr. Rivas will be able to work on this project and come up with a product that will impress. We please request that you consider funding for our project and allow us the ability to work on this for the time period established. We would highly appreciate it.

# Appendix

Programs discussed in Needs/Problems:

* [Musipedia](https://www.musipedia.org/)
* [Midomi](https://www.midomi.com/)

Research Material:

* [Acoustic Fingerprinting](https://blog.chirp.io/audio-fingerprinting-what-is-it-and-why-is-it-useful/)
* [Decision Tree Algorithms](https://medium.com/deep-math-machine-learning-ai/chapter-4-decision-trees-algorithms-b93975f7a1f1)