Voice Converter on Android

To cite this article: W Y Lee et al 2018 J. Phys.: Conf. Ser. 1049 012082

View the [article online](https://iopscience.iop.org/article/10.1088/1742-6596/1049/1/012082) for updates and enhancements.
Voice Converter on Android

W Y Lee, N Murli, N Wahid, R A Hamid, M F Othman
Department of Multimedia, Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, 86400 Batu Pahat, Johor, Malaysia
E-mail: hanifah@uthm.edu.my

Abstract. Voice Converter is an application developed for Android mobile platform that can help user to change his voice the way he likes. Existing voice changer mobile application provides the functions for users to change their voice into the voice of limited preset templates, this cannot see to the needs of film or animation producers, because the templates are limited and the uniqueness of the production cannot be secured. With the Voice Converter mobile application, film or animation producers can easily have access to a number of sound models that can be used to convert audio files which can be used in production later. Besides that, film makers can custom the sound of the conversion by inputting desired values of playback rate to change the audio sound in the way they would like to. The application uses speaker and microphone of the phones to play and record audio. For the recording process, the ‘MediaRecorder’ APIs is used to allow the device to capture audio, save the audio and play it back. The recordings captured can be converted into various sound models and also a model with custom parameters. The SoundPool Library is adopted to alter the playback rate of the audio. A mobile application named Voice Converter which is compatible for Android platforms is successfully developed. The functions of the application include recording and saving audio, converting input audio into existing template sound and converting recorded audio with custom parameter

1. Introduction
Sound can be referred to as vibration that propagates through a transmission medium such as gas, liquid or solid matter and can be heard by human’s or animal’s ear. It is generated from a sound source, where it creates vibrations and these vibrations travels away from the source to the destination in the form of sound wave. Voice in the other hand is the sound uttered through the mouth of living creatures, especially human being in speaking, singing, and shouting. The mechanism for generating human voice can be divided into three parts: the lungs, the vocal cord within the larynx, and the articulators. In the lungs, air is exhaled to create an airstream in the trachea and across the larynx, which is often called the voice box.

When air passes through, the vocal cords vibrate very quickly to produce sound [1]. The articulators act as filters that to some degree can interact with the airflow to control the strength of the sound source. Every human being has their unique voice due to a number of factors. The frequency of the vibration of the vocal cords, the length and width of the vocal cords, the shape and size of the resonance chambers, are among the factors that make a person’s voice unique [2]. A voice changer program is a program or an application that alters a person’s voice to either make them sound like someone else or disguise their voice. Voice changer applications work by changing the tone or pitch, to change the nature of the voice into that of another.

Film and animation production involve various post-production activities, and one of them is the voice dubbing process. With the aid of the Voice Converter mobile application in which audio file can be converted into model with custom parameters, the work will be made easy without having to worry much about the voice actor’s condition.
This paper is organized as follows. In Section 2, the related works are presented. In Section 3, the development of the Voice Converter is discussed. In Section 4 discusses the implementation of voice converter application and finally, future works and conclusion of this project is summarized in Section 5.

2. Related Works
Voice plays an important role in movie and film production, which makes up a big part of entertainment for people around the world. In some drama series or animated series (anime), dubbing is a core post-production process. However, problem arises as drama or anime series filmed over years, and the children characters in the series remain children, while the voice actor has already aged by a few years and may have undergone puberty and have their vocal line changed. In these cases, a new voice actor with similar vocal line is needed but in most cases the new voice actor will not provide a voice identical to the previous voice actor and more or less will affect the quality of the series.

Existing voice changer mobile application provides the functions for users to change their voice into the voice of very limited pre-set templates such as Voice Changer Mobile Application [3], Voice Changer with Effects Mobile Application [4] and Best Voice Changer Mobile Application [5]. This cannot fulfil the needs of professional film or animation producers, because the templates are limited and the uniqueness of the production cannot be secured. With this voice changer application, users are allowed to record and convert human voice into a custom sound. Some software on the computer provides similar functions in professional approaches but users have to be in a particular location to access the hardware. This is inconvenient when the voice actors have difficulties to present on the spot. The Voice Converter mobile application however provides the convenience to the users where they can access the function anytime and anywhere.

The following section will discuss about the concepts and techniques that is employed in the development of the Voice Converter mobile application.

2.1. MediaRecorder API
MediaRecorder is a set of Application Program Interface (API) for Android Mobile Platform. The set of programming algorithms allows the recording of media such as video and audio and for us to access the microphone and storage on an Android device. The Voice Converter mobile application will adopt this algorithm as the basic technique to allow recording of media on Android devices as it is simple to use and is compatible on most Android mobile devices.

2.2 SoundPool Library
SoundPool library is a Java class that manages and plays audio resources for applications. It is a collection of samples that can be loaded into memory from a resource inside the APK. The SoundPool library uses the MediaPlayer service to decode the audio into a raw 16-bit PCM mono or stereo stream. This allows applications to ship with compressed streams without having to suffer the CPU load and latency of decompressing during playback. SoundPool is capable of setting the audio to loop and also change the playback rate of an audio when playing it. The SoundPool class allows loading a number of sound clips into a device’s memory where they can be played on demand. It even allows playing a number of sound clips at the same time.

3. Methodology
The development of this application consists of 3 phases: system design, interface design and application development. The process model adopted in the development of the Voice Converter mobile application is the Waterfall Process Model. It allows for departmentalization and control of the project. A schedule can be set with deadlines for each stage of development to ensure that a phase is completed before moving to the next. Table 1 shows the phases and activities involved in the development of the project.

This paper is organized as follows. In Section 2, the related works are presented. In Section 3, the development of the Voice Converter is discussed. In Section 4 discusses the implementation of voice converter application and finally, future works and conclusion of this project is summarized in Section 5.

2. Related Works
Voice plays an important role in movie and film production, which makes up a big part of entertainment for people around the world. In some drama series or animated series (anime), dubbing is a core post-production process. However, problem arises as drama or anime series filmed over years, and the children characters in the series remain children, while the voice actor has already aged by a few years and may have undergone puberty and have their vocal line changed. In these cases, a new voice actor with similar vocal line is needed but in most cases the new voice actor will not provide a voice identical to the previous voice actor and more or less will affect the quality of the series.

Existing voice changer mobile application provides the functions for users to change their voice into the voice of very limited pre-set templates such as Voice Changer Mobile Application [3], Voice Changer with Effects Mobile Application [4] and Best Voice Changer Mobile Application [5]. This cannot fulfil the needs of professional film or animation producers, because the templates are limited and the uniqueness of the production cannot be secured. With this voice changer application, users are allowed to record and convert human voice into a custom sound. Some software on the computer provides similar functions in professional approaches but users have to be in a particular location to access the hardware. This is inconvenient when the voice actors have difficulties to present on the spot. The Voice Converter mobile application however provides the convenience to the users where they can access the function anytime and anywhere.

The following section will discuss about the concepts and techniques that is employed in the development of the Voice Converter mobile application.

2.1. MediaRecorder API
MediaRecorder is a set of Application Program Interface (API) for Android Mobile Platform. The set of programming algorithms allows the recording of media such as video and audio and for us to access the microphone and storage on an Android device. The Voice Converter mobile application will adopt this algorithm as the basic technique to allow recording of media on Android devices as it is simple to use and is compatible on most Android mobile devices.

2.2 SoundPool Library
SoundPool library is a Java class that manages and plays audio resources for applications. It is a collection of samples that can be loaded into memory from a resource inside the APK. The SoundPool library uses the MediaPlayer service to decode the audio into a raw 16-bit PCM mono or stereo stream. This allows applications to ship with compressed streams without having to suffer the CPU load and latency of decompressing during playback. SoundPool is capable of setting the audio to loop and also change the playback rate of an audio when playing it. The SoundPool class allows loading a number of sound clips into a device’s memory where they can be played on demand. It even allows playing a number of sound clips at the same time.

3. Methodology
The development of this application consists of 3 phases: system design, interface design and application development. The process model adopted in the development of the Voice Converter mobile application is the Waterfall Process Model. It allows for departmentalization and control of the project. A schedule can be set with deadlines for each stage of development to ensure that a phase is completed before moving to the next. Table 1 shows the phases and activities involved in the development of the project.
3.1 System Design
The Voice Converter mobile application consists of only two modules, which are the “Convert” and “Custom” interfaces. The “Convert” module provides the function of recording the voice of user and the conversion of the recorded voice into another voice, while the “Custom” module allows user to change their voice with the tempo and pitch-shift values input by themselves. Figure 1(a) shows the interface of the Convert Module – Audio Input. This interface provides functions for user to record audio, playback the recorded audio, and redo the recording if they would like to before moving to the next interface. This allows user to record their voice for many attempts before they proceed with the one they are satisfied with. Figure 1(b) shows the second interface of Convert Module Interface – Model Selection. This interface provides different sound models for the users to choose from (e.g. Slow-mo, Speed up, Helium, Witch, Ghost, Rhino, Chipmunk, Male, Female, Demon, or Robot). By pressing on the button, the recorded audio will be altered and played. Figure 1(c) shows the record interface of the Custom module. This is the first interface of the module that allows the user to record and playback their voice. Figure 1(d) shows the second interface of Custom module. In this interface, users are required to input the playback rate of the audio. Users then press the convert button to convert and play the audio.

![Image of the interfaces]

Figure 1. (a) Convert Module Interface – Audio Input, (b) Convert Module Interface – Model Selection, (c) Custom Module Interface – Audio Input; and (d) Custom Module Interface – Convert

3.2 Implementation of Convert Module
The Convert Module allows users to record their voice, to layback, to convert their voice into a list of sound models, and to save and to playback the converted audio. The recording function is implemented in Android Studio using the MediaRecorder API which is a Java-based API. It allows the application to access the device microphone to record audio. Shown in the appendix are the code segments for the recording function in Android.

For the purpose of declaration, the recorder bit-rate is set to 16-bit and the recordings format is set to .wav. A recording temporary file will be created during the recording process. The sample rate and channels of the recordings are also declared. Figure 2 shows the code segment for the implementation of the startRecording function method, in which a new MediaRecorder is introduced. Next the recorder function is called to start the recording and the state of recording is set to True. Finally, the writeAudioDataToFile function is called to write recorded audio data to storage, and to save the recorded audio into the device. Figure 3 shows the code segment of stopRecording function, the recording state is first set to False, the recorder is called to stop recording and the whole recorder function is called to End. In the next function interface of the “Convert” module as shown in Figure 1(b), user can choose from a number of sound models that he or she wants to convert to. A SoundPool library is also introduced to the class, so that the alteration of playback rate of audio can be implemented. The default
value for the parameters of soundpool explosion is set to 0. The buttons for different sound models are declared with onclicklisterner that allows them to implement the functions when it is clicked. A new SoundPool function is then declared and the function is set to load the recorded audio that is saved to the device. Figure 4 shows the code segment of the core method of the audio alteration. For different sound models, different parameters for the SoundPool function are defined. The declaration of SoundPool explosion method allows a few parameters to be manipulated when an audio is played. First are the left and right volume for the audio file when played. 1.0 is the maximum value for the playback volume. Next the priority of audio file can be set, with 0 is the lowest priority; the higher the number, the higher the priority. Followed by is the loop value of the audio file: 0 indicates no loop and -1 will set the audio to loop continuously. And finally the playback rate or frequency of the audio. For instance 0.65f indicates a playback rate of 65% of the original audio for the Demon sound model.

```
private void startRecording(){
    recorder = new AudioRecord(MediaRecorder AudioSource.MTC, RECORDER_SAMPLERATE, 
                                RECORDER_CHANNELS, RECORDER_AUDIO_ENCODING, bufferSize);

    int i = recorder.getState();
    if(i==1)
        recorder.startRecording();

    isRecording = true;
    recordingThread = new Thread(new Runnable(){
        @Override
        public void run(){
            writeAudioDataToFile();
        }
    },"AudioRecorder Thread");
    recordingThread.start();
}
```

**Figure 2. Code Segment for startRecording Method**

```
private void stopRecording(){
    if (null != recorder){
        isRecording = false;
        int i = recorder.getState();
        if(i==1)
            recorder.stop();
        recorder.release();
        recorder = null;
        recordingThread = null;
    }
}
```

**Figure 3. Code Segment for stopRecording Method**

```
public void onClick(View v){
    switch (v.getId()){  
        case R.id.imageButton_demon:
            sp.play(explosion, 1, 1, 0, 0, 
                     0.65f);
            break;
        case R.id.imageButton_chipmunk:
            sp.play(explosion, 1, 1, 0, 0, 
                    2.2f);
            break;
        }
}
```

**Figure 4. Code Segment for stopRecording Method**

### 3.3 Implementation of Custom Module

The recording function of the Custom module is implemented using the same codes and functions as in the Convert module. Therefore, this section will discuss about mainly the implementation of converting audio into a custom model. An EditText symbol is used to get the frequency value from the user. A new SoundPool function is then declared and the function is set to load the recorded audio that is saved to the device. In Figure 5 we can see that the SoundPool function to alter the frequency of audio is implemented with a button click. The parameters are predefined except for the playback rate, it is defined by the value input by the user in the application.
3.4 Development of Sound Models

The development of sound models in the Voice Converter mobile application is implemented using Java programming language in Android Studio. The playback rate or frequency of the input audio is altered with different values so that different sound models can be produced. When we increase the playback rate, to 200% of its original value, the audio file will be shortened to half its length and the speech will sound two times faster, the audio will also sound higher by a certain value, which means that the pitch of the audio will also be altered at the same time. In the Voice Converter mobile application, a total of 11 sound models are produced, each of them provides different sound effects and has different parameters. Table 1 below shows the sound models available with their respective playback rate alterations.

| No | Sound Model | Playback rate (% of original audio) |
|----|-------------|-------------------------------------|
| 1  | Speed-up    | 200                                 |
| 2  | Slow-mo     | 60.5                                |
| 3  | Helium      | 180                                 |
| 4  | Male        | 70                                  |
| 5  | Robot       | 120                                 |
| 6  | Female      | 142                                 |
| 7  | Demon       | 65                                  |
| 8  | Ghost       | 40                                  |
| 9  | Witch       | 78                                  |
| 10 | Chipmunk    | 220                                 |
| 11 | Rhino       | 35                                  |

4. Results and Discussions

The testing was carried at Batu Pahat mall targeting open users of age 13-30. A total of 30 respondents were involved regardless of their gender and race. Questionnaires were distributed among them to get feedback on their experience of using the application. The data and feedback collected are then statistically analyzed based on the three categories of questions that are user acceptance, usability and design. The results are tabulated using bar charts as shown in Figure 6 – 8 below. As in Figure 6, almost all respondents agree and strongly agree that the application is fun and easy to use and gives good experience to them. In terms of the overall design (Figure 7), all respondents find that the sound models produced are attractive (100%), the use of buttons are appropriate (90%) and the selection of colours are good and appropriately chosen (70%). And finally, referring to Figure 8, the functionality level of the application is rather at a satisfying level. Most users responded positively, i.e. they strongly agree and agree that the application run well on Android devices and the buttons function properly. Most of them
also agree that the navigation is simple (80%), the application provides quality conversions (70%) and the sound models are fun and useful (60%).

![User Acceptance Level of Audio Converter Application](image1)

**Figure 6.** User acceptance level of Audio Converter Application

![User Acceptance Towards Voice Converter Design](image2)

**Figure 7.** User Acceptance Towards Voice Converter Design

![Functionality Level of Audio Converter Application](image3)

**Figure 8.** Functionality level of Audio Converter Application

5. **Conclusion and Future Works**

The Voice Converter mobile application possess some advantages compared to similar application as the user can input the value to alter the frequency of the audio themselves. On the other hand, the application has also some critical limitations, one of them is that the application does not allow users to save the converted audio file to the device.

The future works of the mobile application involve solving limitations and providing extra features. The Voice Converter mobile application is a production developed in a short timeframe and under limited resources, which gives it a big room for improvement in the future. Main functions like saving the converted audio to the device and also importing existing audio file to be converted should be allowed instead of having to record every time they wish to convert. Instead of having user to input the frequency, sliders shall be created to ease users and avoid invalid input. The application also shall be developed to operate on iOS platform so that it can be made available to more users and larger market. More manipulating properties for the audio other than playback rate, for example altering the pitch and tempo, or adding echo to the audio, this can help to produce more sound models and ensure the application continuously provide fun and usefulness to the users.
Acknowledgement
This study has been supported in part of the Short Term Grant (STG) Vot.U367 under the Malaysia Ministry of Higher Education (MOHE) and Universiti Tun Hussein Onn Malaysia (UTHM). This research also supported by GATES IT Solution Sdn. Bhd. under its publication scheme.

References
[1] The Propagation of sound Retrieved June 2015 from http://pages.jh.edu/~virtlab/ray/acoustic.htm
[2] Kampwirth K 2013 What Determines What Your Voice Sounds Like (Retrieved September 2016) from http://mentalfloss.com/
[3] Voice Changer Mobile Application (Retrieved September 2016) from https://play.google.com/store/apps/details?id=com.androidrocker.voicechanger
[4] Voice Changer with Effects Mobile Application (Retrieved September 2016) from https://play.google.com/store/apps/details?id=com.baviux.voicechanger
[5] Best Voice Changer Mobile Application (Retrieved September 2016) from https://play.google.com/store/apps/details?id=com.scoompa.voicechanger