LBYEC4A – EK1

Signals, Spectra and Signal Processing Laboratory

Final Project Proposal

Voice Identification using Pitch pattern recognition through Digital Signal Processing

Renato Angelo DL. De Vera
Abigail Joy T. Leoncio
Ethan James T. Lucmayon

PROJECT DESCRIPTION (Describe what your project is all about and its intended application. Include your research showing how your intended application can be achieved by your project. Also, provide theoretical concepts that will be utilized.)

With the modern world overrun by technological advancements, certain devices, and machines have adapted to gathering and analyzing information inputted by humans whether it may be in an audial or visual manner. Audially, humans can omit sound naturally with their mouths as they communicate with each other. With that being said, recognizing voice by sound and pattern is one of the few ways machines can analyze the information omitted by humans.

This project will revolve around a voice recognition system that is trained to recognize specific speech by comparing amplitudes set in the database by the user. Using a recorded voice of one member of the group as the trained variable. For multi-variable testing, the other members of the group will be creating a similar audio file to test the efficiency of the program in terms of the trained variable that is stored. Wherein the program will attempt to compare all the recordings to the trained variable via the amplitude and pitch of the audio signal. The program must be able to match the trained variable to the audio recording with the same voice.

The waveforms formed from the amplitudes and frequencies found in speech will be used as an analysis for identifying pitch patterns. With numerous possible transfer functions to use such as discrete Fourier and fast Fourier transform, the group has decided to focus on using discrete cosine transform.

Discrete Cosine Transform

Similar to the discrete Fourier transform, the discrete cosine transform manipulates a signal to be converted into the frequency domain. The discrete cosine transform is represented by the equation [1]:

$$F(u) = (\frac{2}{N})^{\frac{1}{2}} \sum_{k=0}^{N-1} \Delta(k) cos[\frac{\pi * u}{2N}(2k + 1)] f(i)$$

The discrete cosine transform utilizes cosines in which are components for real vectors. Discrete cosine transforms are classified as sinusoidal unitary transforms. Such transforms are linear and defined by discrete cosine functions [2].

METHODOLOGY (How are you going to do it? Included an overall system flowchart of how your project should work as well as initial draft of schematic diagram. Include description of digital signal processing concepts that will be used to develop the project.)

The group plans on executing the voice recognition by relying on the amplitude obtained from each audio file. Since the group will be having a directory for the audio, they will be loading the input audio and loading the directory from its file location. They will then plan on differentiating the amplitude of the audio using the discrete cosine transform function. The maxima of the signals are to be acquired and to be matched.

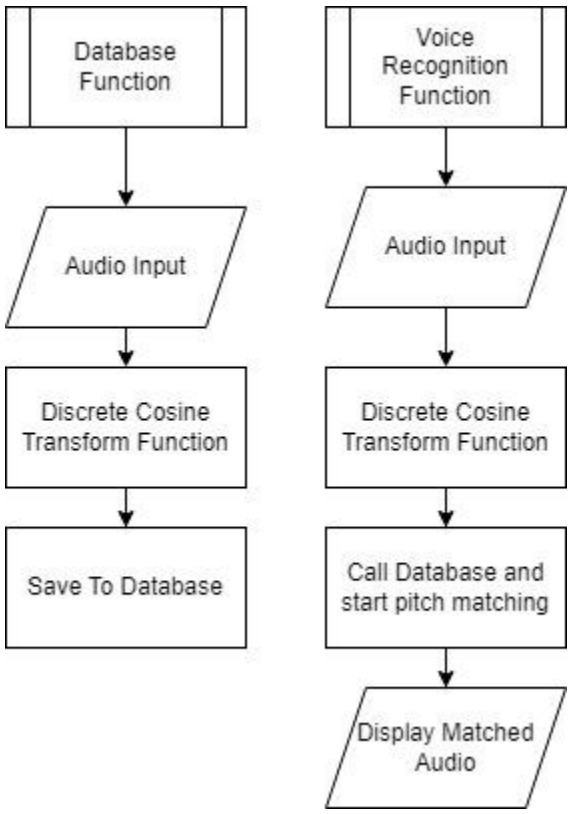

Figure 1. Method Flowchart

Training

In training, a set of voice recordings from each member are run into the database function and saved into the database to have reference audios for the test input. By saving multiple reference data, the program is trained to have a variety of choices when the operation eventually runs the actual feature matching.

Feature Matching or Testing

Feature matching or testing is the actual comparison of the input test data to the multiple reference test data from the database. The program should be able to identify through visual means the similarity between the test input data from the counterpart reference data.

SCHEDULE OF ACTIVITIES (Provide a timetable or Gantt chart of your deliverables. Indicate also whom and when the specific deliverables will be accomplished.)

	TASK TITLE	Member Assigned	WEEK 1					WEEK 2					WEEK 3					WEEK 4					WEEK 5					WEEK 6					WEEK 7				
			M	T	W	H	F	M	T	W	H	F	M	T	W	H	F	M	T	W	H	F	M	T	W	H	F	M	T	W	H	F	M	T	W	H	F
1	**Project Proposal**																																				
1.1	Research possible projects	All																																			
1.2	Post project proposal in chat room	Renato																																			
1.3	Create GitHub Repository	Abigail																																			
1.4	Create board on Trello	Ethan																																			
2	**Final Project Proposal**																																				
2.1	Search Review of Related Literature	All																																			
2.2.1	Create Voice Directory	Abigail																																			
2.2.2	Create Voice Pitch Pattern	Renato																																			
2.2.3	Create Voice Recognition	Ethan																																			
2.3	Update files in repository and board	Abigail																																			
2.4	IPO model	All																																			
3	**Review**																																				
3.1	Review files in repository and board	All																																			
3.2.1	Final checking of codes	Renato																																			
3.2.2	Test run	Ethan																																			
3.3	Final paper revision	All																																			
4	**Presentation**																																				
4.1	Project Presentation	All																																			

REFERENCES (Cite the resources that will be used as well as your research regarding your project.)

[1] D. Marshall, “The Discrete Cosine Transform (DCT),” Cf.ac.uk. [Online]. Available: https://users.cs.cf.ac.uk/Dave.Marshall/Multimedia/node231.html. [Accessed: 13-Mar-2023].

[2] V. Britanak, P. C. Yip, and K. R. Rao, “Fast DCT/DST Algorithms,” in Discrete Cosine and Sine Transforms, Elsevier, 2007, pp. 73–140.

[3] S. Tallat, “Voice identification and recognition system,” 2015.

[4] IMPLEMENTATION OF A VOICE-BASED BIOMETRIC SYSTEM, “PROJECT REPORT,” Cornell.edu. [Online]. Available: http://chenlab.ece.cornell.edu/people/adarsh/research/speech_project.pdf?fbclid=IwAR1IoJWjuS5pfHklVtfRtyb2cmOqctexdCxfXX0A9Mn8chex12flupUCYHs. [Accessed: 13-Mar-2023].

[5] S. Bunrit, The authors are with the School of Computer Engineering, SUT, 111 University Avenue, Muang, Nakhon Ratchasima 30000, Thailand, T. Inkian, N. Kerdprasop, and K. Kerdprasop, "Text-independent speaker identification using deep learning model of convolution neural network," Int. J. Mach. Learn. Comput., vol. 9, no. 2, pp. 143–148, 2019.

[6] T. Siddiqui, "Voice Recognition System," Jul. 2020.

[7] R. A, R. Agustina, and Hidayatulloh, "Comparison of discrete cosine transforms (DCT), discrete Fourier transforms (DFT), and discrete wavelet transforms (DWT) in digital image watermarking," Int. J. Adv. Comput. Sci. Appl., vol. 8, no. 2, 2017.