ENCS4310 (DSP) Course Project

Pierre Backleh [1], *Christina Saba*[2], Joud Hijaz[3]

[1] Electrical and Computer Engineering, Birzeit University
[2] Electrical and Computer Engineering, Birzeit University
[3] Electrical and Computer Engineering, Birzeit University

1201296@student.birzeit.edu[1], 1201255@student.birzeit.edu[2], 1200342@student.birzeit.edu[3]

1. Introduction

The project's objective is to design a system for encoding and decoding information in an audible format by dealing with the interface of digital signal processing and communication technologies. The project explores the unique intersection between communication technologies and digital signal processing with the goal of offering a new perspective on information transfer.

2. Problem Specification

The main technological issue this project attempts to solve is how to use digital signal processing to convert textual data into audio signals i.e. Transcribing music into well-defined technical goal. One way to achieve this translation from text to audio is to build an English character encoder which generates a signal representing each character through a combination of low, middle, and high-frequency components, and ensure each character's duration is approximately 40ms, this can be dealt with by designing a User-interface allowing the user to enter the English string in the txt box and visualize the corresponding signal. The user is also given the choice to play the generated signal or save it as .wav file. The system also proved a reversed method to translate an audio saved in .wav file into text representation, it constructs two decoders to recover the original text from the encoded multi-frequency signal using two different approaches – frequency analysis such as FFT and a bandpass filter.

3. Data

In this project, the data set can be categorized into two phases, the encoder phase and decoder phase. The encoder the dataset utilizes strings in English characters each with their own voice-band frequencies which are then transformed into sine signals and they can be transform into real-time audio signal saved as .wav files by the user. The second phase represents decoder, the data consists of a collection of .wav files uploaded by the user and containing spoken English words or phrases to learn the mapping between voice-band frequencies and English characters.

4. Evaluation Criteria

The performance of the project can be measured using various criteria aiming to measure the system's effectiveness, functionality, and accuracy. These criteria are specified as objective, quantitative, discriminatory, and qualitative.
After building the GUI interface, it was tested for being responsive, straightforward, and user-friendly. We verified that users can easily open the interface, input strings, encode them, and save the output for the encoder phase, and for the decoder phase, users could upload .wav files as required.
The accuracy and adaptability of encoding and decoding phases were also measured by testing the correctly encoded input text to signals and decode the signals back to text. We used a list of entered strings with English alphabets and compared the output with the original input. Also, the audio signals were analyzed by visualizing the waveforms and the encoded frequencies.
The system's robustness and reliability were tested by ensuring handling unexpected inputs such as numbers or special characters to control the error handling process. We also compared the analysis o FFT and bandpass filter decoding approaches by monitoring the signal distortion, handling noise, and processing time. this also helped us make some computational calculations by checking the time complexity for both encoding and decoding algorithms in both phases through testing the scalability of the system by providing varying lengths string inputs and recording the total execution time for processing them.
Following these criteria supported us with various methods to improve the system's overall performance theoretically and practically.

5. Approach

Our approach for solving the problem of encoding and decoding English characters into voice-frequency signals was based on two phases, Phase 1 and Phase 2.

- Phase 1: Encoder Designing: Graphical User-Interface (GUI) was built to enable the user to easily interact with the system, enter input string and receive the encoded waveform signals, and give them the option to save the signals as .wav files.
  We designed the encoder by mapping each English alphabet to a unique set of frequencies (low, middle, high). By using python programming language and libraries i.e. numpy, scipy.fft, matplotlib for signal visualization, PyAudio for real-time audio playback, tkinter, etc. we used sine wave generation method to map the characters.

- Phase 2: Decoder Designing: the GUI was built with a button for uploading the .wav files and display the decoded string and two buttons indicating the method to be used for decoding. As a result, two methodologies have been utilized for decoding process. Fast Fourier Transform FFT was used to analyze the frequency components of the audio signals, the other one is band-pass filters to isolate the frequency bands depending on each letter. Both approaches aimed to match the frequencies from the audio input with the previously defined character frequency map saved (low, middle, high) in order to reconstruct the original text and display it.

6. Results and Analysis

Fast Fourier Transform (FFT) Analysis

"The key impact of FFT is it provides an efficient way to compute the Fourier Transform of real-world data". (Talebi, 2020)

The FFT method was implemented to convert the time-domain signal into the frequency domain. The fundamental equation used is:

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-\frac{i2\pi}{N}kn}$$

Where X(k) is the frequency component at frequency k, x(n) is the signal amplitude at time n, and N is the total number of samples. Using FFT, we analyzed chunks of the signal to identify the dominant frequencies. The accuracy in ideal conditions was high, but noise significantly affected the FFT's ability to accurately identify character frequencies.

The key impact of FFT is it provides an efficient way to compute the Fourier Transform of real-world data.

Band-pass Filter Analysis

The band-pass filter allows a specific range of frequencies to pass while attenuating others outside the accepted range. The Root Mean Square RMS value indicates the average power of a signal calculated by taking the square root of the average of the squared values of a signal over a specified time period. Later, the RMS value of the filtered signal represents the amplitude of the filtered frequencies. Here's a formula of RMS value for continuous signals is calculated as:

$$X_{\text{rms}} = \sqrt{\frac{1}{T}\int_0^T x(t)^2\,dt}$$

For each character frequency set, bandpass filters were applied to isolate specific frequencies. The filtered signal y(t) is obtained by:

$$y(t) = \int_{-\infty}^{\infty} x(t) \cdot h(t)\,dt$$

Where x(t) is the input signal and h(t) is the impulse response of the band-pass filter. The band-pass filter showed better noise resilience compared to FFT. Although slower, it was more robust in noisy conditions, indicating its suitability for real-world applications.

We present the outcomes of our signal processing program in this field. In order to evaluate the effectiveness of our encoding and decoding system, we ran a number of experiments. For clarification, screenshots of the program's interface are included.

6.1 Normal Sentence Input

6.1.1 Scenario:

An ordinary sentence served as the encoding's input. There were no numbers or prohibited characters in the sentence.

6.1.2 Input Sentence:

"Joud Pierre Christina"

When the input sentence "Joud Pierre Christina" is given, Figure 1 shows the user interface of the program.

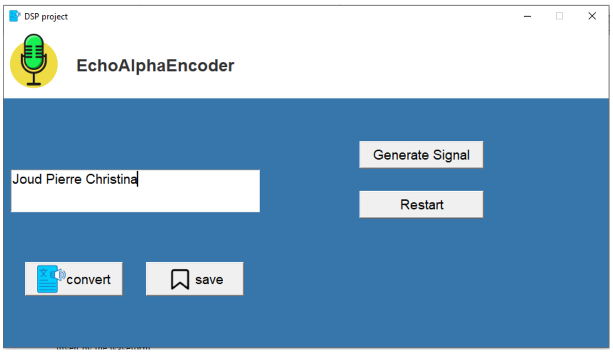

Figure 1- input sentence

6.1.3 Program Output:

The encoderOutput1.wav file contains the result of the first example, which utilized an ordinary phrase as the input. For a thorough examination of the encoded audio, please see the Appendix section.

Figure 2 shows the generated signal in the time domain that was taken after selecting the 'Generate Signal' button. A visual picture of the signal properties during the encoding process is given by the waveform.

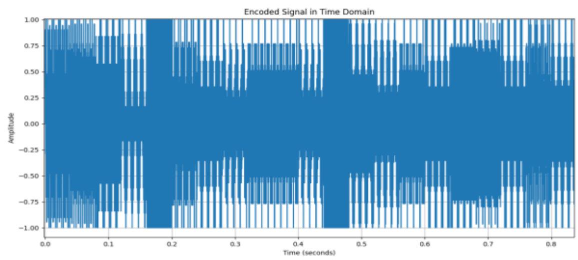

Figure 2 - generated signal

6.2 Sentence with Numbers and Forbidden Characters

6.2.1 Scenario:

As input for encoding, an entered sentence with numbers and prohibited characters was utilized.

6.2.2 Input Sentence:

"Testing 123"

When the input sentence "Testing 123" is entered, the system should when non-alphabetic letters and numerals were present, the system recognized them correctly. The following outcomes were shown in the figure below:

6.2.3 Program Output:

Figure 3 - Invalid Input

The system correctly identified and reported the presence of incorrect characters in the numeric segment when the input sentence "Testing 123" was submitted. It also properly encoded the word "Testing."
This case's output is saved in the "Invalid_Sentence.wav" file.

6.3 Decoded Stage

6.3.1 Scenario:

For decoding, a pre-recorded.wav file was utilized as input. The sentence that was used to produce the.wav file was known.

6.3.2 Input .wav File:

Filename: input_wave.wav

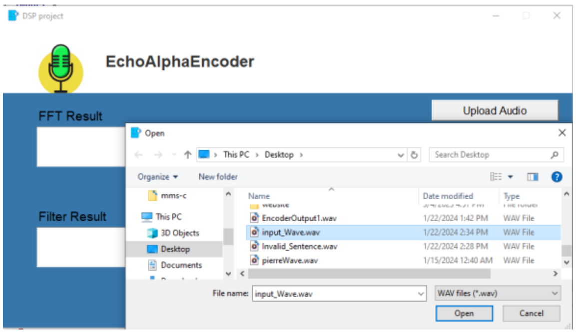

Figure 4 - input wave file

6.3.3 Program Output:
The result is represented visually in figure 5 that shows the text fields for the two decoding techniques used, Band-pass Filter and FFT. Both techniques' decoded results are saved in the related text fields as shown below.

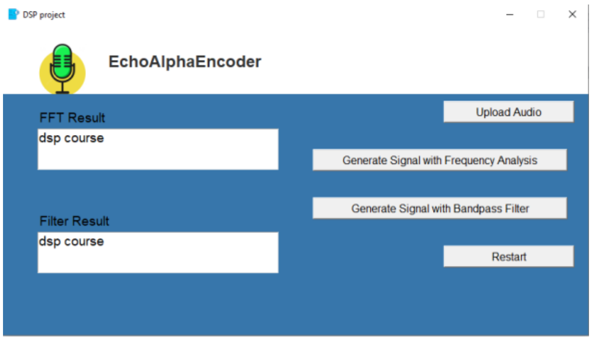

Figure 3 - Decoded Sentence

- FFT Result: dsp course
- Band-pass Filter Decoding Output: dsp course

The input.wav file was successfully decoded by the application, and the text that is shown matches the original message. The encoded frequencies were successfully caught by the band-pass filters, leading to accurate decoding.

The strength of the system can be seen by this visual image, which gives a clear summary of the successful decoding outcomes obtained by both methods for the provided input text, "DSP course."

7. Development

Depending on the evaluation analysis, followed by implementing the encoder and decoder, we divided several categories to enhance the system's effectiveness and guarantee user's experience to handle any unexpected side-effects.
Currently, our system serves only English characters, we plan to expand the set to encompass numbers, punctuation, probably other languages alphabets, and translating universal symbols, becoming a multilateral communication device.
The way the encoder works is by mapping each character with a fixed frequency set, they could be replaced with a dynamic method for assigning frequencies based on the context enhancing the clarity of the signal and system's adaptability.
In order to increase the system's capacity in handling large dataset or real-time audio streaming signals and reduce the processing time, we can use optimized FFT computations and other efficient designs of filters in order to optimize the encoder and decoder algorithms implemented.
It is needed to study the improvements of the advanced noise reduction mechanism and error correction to improve the accuracy and output correctness of decoding phase, especially in realistic environments such as background noise or low-quality recordings.
For future developments, we aim to potentially broadening our scope by integrating our system with bigger ones such as text-to-speech and speech recognition applications. This can give us the ability to test it using larger realistic dataset and receive diverse feedbacks for serious and practical technologies.

8. Conclusion

This report has shown how digital signal processing may be used in practice to encode and decode text into voice frequencies. We have developed a system that successfully decodes audio signals back into text once text has been converted into them. We discovered areas for improvement and learned more about the efficacy of the band-pass filter and FFT decoding methods.

In order to increase system speed and usability, future development will concentrate on algorithm optimization, noise reduction, and user interface enhancements. All things considered, the project shows the potential of DSP applications in communication technologies and lays foundations for further advancements in the area.

9. References

[1] Talebi, S. (2020, Dec 4). *Medium*. Retrieved from The Fast Fourier Transform (FFT): https://medium.com/swlh/the-fast-fourier-transform-fft-5e96cf637c38