AUDIO AND SPEECH COMPRESSION USING DCT AND DWT TECHNIQUES

Isaac Dzakpata

Febraury 2026

1 INTRODUCTION

Speech is a basic way for humans to convey information and communicate effectively. The change in the telecommunication infrastructure, in recent years, from circuit switched to packet switched systems has also reflected on the way that speech and audio signals are carried in present systems [1]. Audio compression has since become an important concept in the new multimedia age with a goal of coding audio and speech signals at the lowest possible data rates. The main objective of speech compression is to process human speech signals into an efficient encoded form that can be decoded back to produce a close approximation of the signals [1]. Storage and transmission of uncompressed speech data will be extremely costly and impractical, so we try to reduce the size of the audio signals whiles still maintaining an acceptable quality. Balance is key for audio compression. To compare the efficiency of audio compression methods, this study investigates specialized transform techniques known as Discrete Cosine Transform (DCT) and the Discrete Wavelet Transform (DWT). Discrete Cosine Transform (DCT) is often described as a specialized or "low-level" version of the Discrete Fourier Transform (DFT/FFT). It is frequently used for data compression because it concentrates the energy of a signal (like an image) into a small number of coefficients more effectively.

2 THEORY

There are various techniques for speech compression like waveform coding and parametric coding. This paper focuses on the transform coding techniques which mainly works by converting the signals into the frequency domain and isolating the dominant features only taking out any extra noise which comes off as less dominant peaks or features. In transform method we have used discrete wavelet transform technique and discrete cosine transform technique. When we use wavelet transform technique, the original signal can be represented in terms of wavelet expansion [1]. Similarly in case of DCT transform, speech can be represented in terms of DCT coefficients. Wavelet transform is the latest method

of compression because of its ability to describe any type of signals both in time and frequency domain [2]. Transform techniques do not compress the signal, they provide information about the signal and using various encoding techniques like Run-length encoding and Huffman encoding, we then compress the signals with the information deduced. In both methods, the transform coefficients provide an alternative representation of the signal. Many of these coefficients have very small values and contribute little to the overall signal. By removing these small coefficients, significant compression can be achieved while maintaining acceptable signal quality.

References

[1] MV Patil, Apoorva Gupta, Ankita Varma, and Shikhar Salil. Audio and speech compression using dct and dwt techniques. *International Journal of Innovative Research in Science, Engineering and Technology*, 2(5):1712–1719, 2013.

[2] K.A. Subramanian and R. Karthigeyan. Wavelet transform and fast fourier transform for signal compression: A comparative study. *International Journal of Engineering Research & Technology (IJERT)*, 2(1), 2014. IFET–2014 Conference.