Speech compression and decompression using DWT and DCT

Harmanpreet Kaur, Ramanpreet Kaur

Abstract—speech compression is the digital signal which is compressed by using various techniques for transmission. This paper explains a transform methodology for compression of the speech signal. In this paper speech is compressed by discrete wavelet transform technique afterward compressed signal is again compressed by discrete cosine transform afterward compressed signal is decompressed by discrete wavelet transform. The performance of speech signal is measured on the basis of peak signal to noise ratio (PSNR) and mean square error (MSE) by using different filters of wavelet family.

Keywords—DCT, DWT, SPEECH COMPRESSION AND DECOMPRESSION

I. INTRODUCTION

Speech is a very basic way for humans to convey information to one another. With a bandwidth of only 4kHz, speech can convey information with the emotion of a human voice. People want to be able to hear someone’s voice from anywhere in the world—as if the person would be in the same room. Speech can be defined such as the response of the vocal tract to one or more excitation signals.

Compression of signals is based on removing the redundancy between neighbouring samples and/or between the adjacent cycles. In data compression, it is desired to represent data by as small as possible number of coefficients within an acceptable loss of visual quality. Compression techniques can be classified into one of two main categories: lossless and lossy.

Compression methods can be classified into three functional categories:

- Direct Methods: The samples of the signal are directly handled to provide compression.
- Transformation Methods: such as Fourier Transform (FT), Wavelet Transform (WT), and Discrete Cosine Transform (DCT).
- Parameter Extraction Methods: A preprocessor is employed to extract some features that are later used to reconstruct the signal.

A discrete wavelet transform is defined as a “small wave” that has its energy concentrated in time to provide a tool for the analysis of transient, non-stationary, or time-varying phenomena. It has the oscillating wave like properties but also has the ability to allow simultaneous time and frequency analysis. Wavelet Transform has emerged as a powerful mathematical tool in many areas of science and engineering, more so in the field of audio and data compression.

Wavelet transform decomposes a signal into a set of basic functions. These basis functions are called wavelets. Wavelets are obtained from a single prototype wavelet y(t) called mother wavelet by dilations and shifting:

\[ \psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \]

Where a is the scaling parameter and b is the shifting parameter.

The wavelet transform is computed separately for different segments of the time-domain signal at different frequencies. Multi-resolution analysis: analyzes the signal at different frequencies giving different resolutions. MRA is designed to give good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. Good for signal having high frequency components for short durations and low frequency components for long duration e.g. images, video frames, and speech signal.

The Discrete Cosine Transform (DCT) is a transform that is very common when encoding video and audio tracks on computers. Many “codec” for movies rely on DCT concepts for compressing and encoding video files. The DCT can also be used to analyse the spectral components of images as well. The DCT is very similar to the DFT, except the output values are all real numbers, and the output vector is approximately twice as long as the DFT output. It expresses a sequence of finite data points in terms of sum of cosine functions.

The discrete cosine transform (DCT) is similar to the discrete Fourier transform (DFT), but using only real values. In speech processing, dct

The 1-d discrete cosine transforms

\[ y(k) = w(k) \sum_{n=1}^{N} x(n) \cos\left(\frac{(2n-1)(k-1)}{2N}\right) \quad k = 1, 2, \ldots, N \]

Where...
\[ u(k) = \begin{cases} \frac{1}{N} & k = 1 \\ \frac{2}{N} & 2 \leq k \leq N \end{cases} \]

\( N \) is the length of \( x \), and \( x \) and \( y \) are the same size. If \( x \) is a matrix, \( \text{dct} \) transforms its columns. The series is indexed from \( n = 1 \) and \( k = 1 \) instead of the usual \( n = 0 \) and \( k = 0 \) because MATLAB vectors run from 1 to \( N \) instead of from 0 to \( N-1 \).

It often reconstructs a sequence very accurately from only a few DCT coefficients, a useful property for applications requiring data reduction.

II. IMPLEMENTATION

In this paper hybrid model which is using dwt and DCT techniques for speech compression. First set the compression and decompression flag at zero level then define the Fs of 8000 samples in 1sec which is using for recoding of the signal.

Record the signal which is to be compress. Fig 1. Shows the recoded signal which is to be compressed.

![fig.1 recorded signal](image)

The above Fig. 1, represents that the recoded signal which is to be compressed. Before compression set the filter through which signal is to be passed and define the signal time. Afterwards compress the signal using dwt technique.

Matlab code:

\[ [C,L] = \text{wavedec(frame,level,wavelet)}; \]

\%wavelet decomposition of the frame

\[ [\text{thr,sorh,keeppapp}] = \text{ddencmp('cmp','wv',frame)}; \]

\%wavelet compression scheme

\[ [XC,CXC,LXC,PERF0,PERFL2] = \text{wdencmp('gbl',C,L, wavelet,level,thr,sorh,keeppapp)}; \]

\[ C=CXC; \]

\[ L=LXC; \]

In the above code, first analyze wavelet decomposition of the frame by using “wavedec” command in MATLAB. Then, set the default values for threshold value for one dimension wavelet compression by using “ddencmp” command in MATLAB. Then, “wdencmp” command in MATLAB using for 2-D compression of the signal. Check the PSNR and MSE of the compressed signal.

Afterwards, compression by dwt now compresses the signal by DCT technique. In this 2-D discrete cosine transform using for the compression of the signal. In this compressed signal is further compress by 2-D invariant DCT. This DCT compression is done to improve the PSNR and MSE level of the signal.

Again check the PSNR and MSE level of the signal.

When the compression of the signal is complete then decompress the signal by dwt technique.

In decompression of the signal, firstly set the decompression value=1. Then, with the help of “waverec” command i.e. for 1-d wavelet function is analysed for reconstructing the signal. However, removing distortion from decompress signal and then reconstruct the signal same as the original signal.

The below fig. 2, represents the compressed signal analysed by DWT and DCT, it also represents the decompressed signal by DCT,

III. RESULT

Speech compression of various signals is achieved by using different filters of wavelet with DWT and DCT transform techniques. Here table 1. Shows the compression of ‘hello’
signal using various filters of Debauchees (db), Symlets (sym), Biorthogonal (bior) and Coiflet (coif) wavelets respectively. It shows that performance of compression on the basis of PSNR and MSE of signal at each filter.

Table 1. Compression result of hello signal

<table>
<thead>
<tr>
<th>Signal</th>
<th>Filters</th>
<th>PSNR after DWT</th>
<th>PSNR after DCT</th>
<th>MSE after DWT</th>
<th>MSE after DCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hello</td>
<td>Db1</td>
<td>84.98</td>
<td>87.33</td>
<td>0.04319</td>
<td>0.03430</td>
</tr>
<tr>
<td>Hello</td>
<td>Db2</td>
<td>83.49</td>
<td>85.93</td>
<td>0.05362</td>
<td>0.02218</td>
</tr>
<tr>
<td>Hello</td>
<td>Db10</td>
<td>83.02</td>
<td>85.98</td>
<td>0.05987</td>
<td>0.03026</td>
</tr>
<tr>
<td>Hello</td>
<td>Coif1</td>
<td>79.92</td>
<td>81.70</td>
<td>0.12198</td>
<td>0.08110</td>
</tr>
<tr>
<td>Hello</td>
<td>Coif2</td>
<td>81.02</td>
<td>82.71</td>
<td>0.09479</td>
<td>0.06417</td>
</tr>
<tr>
<td>Hello</td>
<td>Sym2</td>
<td>78.28</td>
<td>80.02</td>
<td>0.17821</td>
<td>0.11928</td>
</tr>
<tr>
<td>Hello</td>
<td>Sym8</td>
<td>79.79</td>
<td>81.61</td>
<td>0.12588</td>
<td>0.08273</td>
</tr>
<tr>
<td>Hello</td>
<td>Sym10</td>
<td>81.60</td>
<td>83.61</td>
<td>0.08303</td>
<td>0.05148</td>
</tr>
<tr>
<td>Hello</td>
<td>Bior1.1</td>
<td>82.60</td>
<td>84.73</td>
<td>0.06588</td>
<td>0.03994</td>
</tr>
<tr>
<td>Hello</td>
<td>Bior2.2</td>
<td>82.19</td>
<td>84.05</td>
<td>0.07243</td>
<td>0.05302</td>
</tr>
<tr>
<td>Hello</td>
<td>Bior3.1</td>
<td>78.19</td>
<td>81.16</td>
<td>0.18200</td>
<td>0.09303</td>
</tr>
</tbody>
</table>

Table 2. Shows the result of speech signal “blackberry” using different filters of wavelet transform.

<table>
<thead>
<tr>
<th>Signal</th>
<th>Filters</th>
<th>PSNR after DWT</th>
<th>PSNR after DCT</th>
<th>MSE after DWT</th>
<th>MSE after DCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blackberry</td>
<td>Db1</td>
<td>86.18</td>
<td>88.39</td>
<td>0.02891</td>
<td>0.06385</td>
</tr>
<tr>
<td>Blackberry</td>
<td>Db2</td>
<td>85.23</td>
<td>87.53</td>
<td>0.03527</td>
<td>0.02187</td>
</tr>
<tr>
<td>Blackberry</td>
<td>Db10</td>
<td>82.19</td>
<td>87.84</td>
<td>0.07236</td>
<td>0.04319</td>
</tr>
<tr>
<td>Blackberry</td>
<td>Coif1</td>
<td>87.54</td>
<td>89.20</td>
<td>0.02827</td>
<td>0.05302</td>
</tr>
<tr>
<td>Blackberry</td>
<td>Coif2</td>
<td>86.47</td>
<td>88.19</td>
<td>0.02701</td>
<td>0.06680</td>
</tr>
</tbody>
</table>

Table 2. Compression result of blackberry signal

IV. CONCLUSIONS

Speech compression is for transmission or storage, possibly to an unintelligible state, with decompression used prior to playback. The speech compression is achieved by representing each sample of digitized data by lesser number of bits this paper shows the key advantageous features of the Wavelet filters in the field of speech Signal processing. It is found that the wavelet filters significantly improves the reconstruction or fidelity assessments of the compressed speech signal. The PSNR and MSE of signal show the high efficiency of the compression by using various filters with hybrid model. It is concluded that it can be very effectively used for the speech signal compression.

REFERENCES

[8] Russel Lloyd Lim “Speech Compression using the Discrete Wavelet Transform”