Audio Compression using Wavelets in MATLAB

Harshit Chaudhary 2016151
Siddhant Singh 2016199

INDRAPRASTHA INSTITUTE *of*
INFORMATION TECHNOLOGY **DELHI**

Introduction:

- Audio is common in all entertainment applications .
- Audio/video compression frees up space substantially, which can then be utilised for other purposes.
- There are several algorithms to compressing them
- We intend to compare their efficiency of two algorithms
- A **wavelet** is a wave-like oscillation with an amplitude that begins at zero, increases, and then decreases back to zero
- Wavelets can be combined with known portions of a damaged signal to extract information of the unknown portions

Haar Algorithms

1. Selects an audio
2. Creates frame
3. Decomposes the signal spectrum into wavelet
4. Creates psychoacoustic model
5. Inspects the spectrum and finds tones maskers
6. Applies mu-law of compression
7. Finds and corrects offset
8. Finds the size of compressed signal

Daubenches Algorithms

1. Selects an audio
2. Chooses a block size
3. Changes compression percentages (different window sizes)
4. Initialises compressed matrix
5. Does compression using inverse discrete cosine transform (IDCT)
6. Finds the size of compressed signal

Parameters of Comparison

1. PSNR : Peak signal-to-noise ratio
2. NRMSE: Normalised root-mean-square error
3. Compression ratios

Formulas Used

1.PSNR = $$10 \log_{10} \frac{NX^2}{\|x-r\|^2}$$

N is the length of the reconstructed signal, X is the maximum absolute square value of the signal x and is the energy of the difference between the original and reconstructed signals.

2. MSE = $$\sqrt{\frac{\{x(n)-r(n)\}^2}{\{x(n)-\mu_A(n)\}^2}}$$

x(n) is the speech signal, r(n) is the reconstructed signal, and μx(n) is the mean of the speech signal.

3. Compression ratio: $$\frac{Length\,(x(n))}{Length\,(cWc)}$$

WC is the length of the compressed wavelet transform vector.

3. μ-law compressor:

$$y = \frac{V \log(1+\mu\,|x|\,/\,V)}{\log(1+\mu)} \operatorname{sgn}(x)$$

Where V is the maximum value of signal xμ is the μ-law parameter.

Matlab Inbuilt functions Used

- audioread
- Audiowrite
- length
- ceil
- Sqrt
- DCT
- IDCT
- wavedec
- ddencmp
- wdencmp
- fft
- compand
- waverec

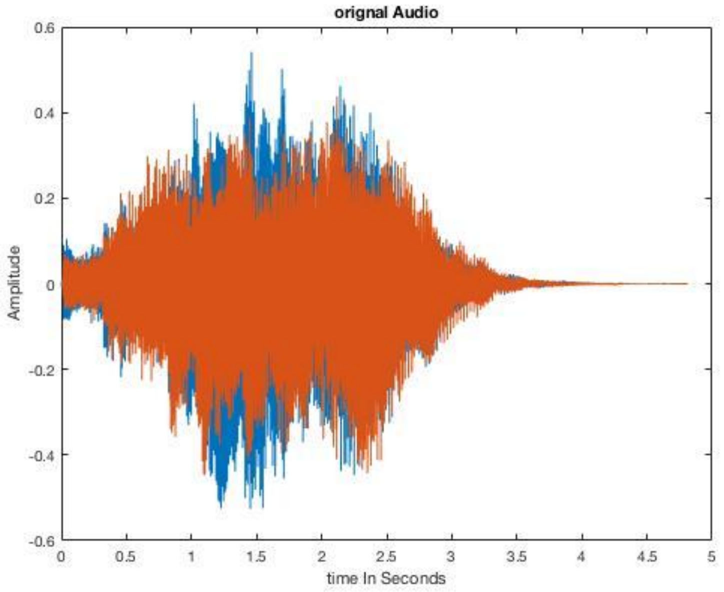
Concept of DSP Used in the Project

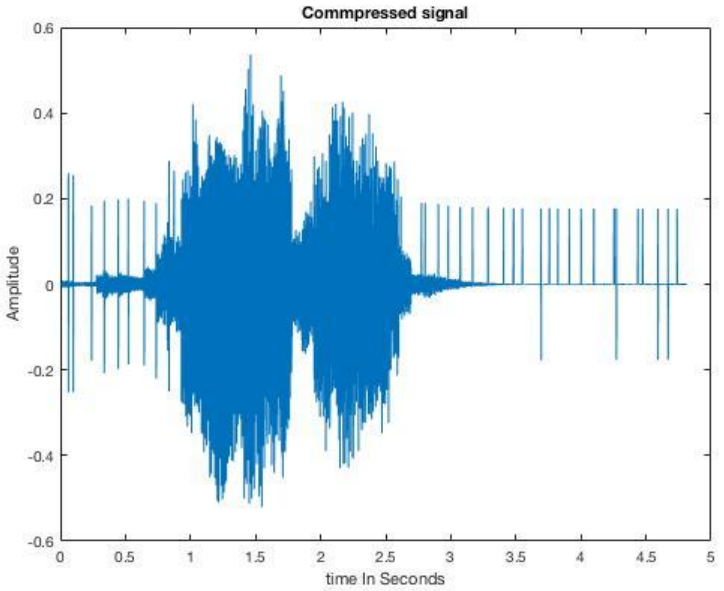
1. Discrete Cosine Transform
2. Inverse Discrete Cosine Transform
3. Windowing

Results

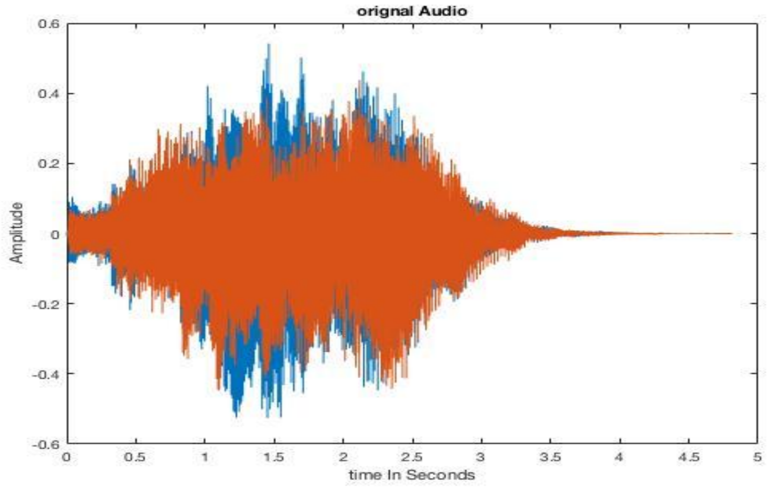
Algorithm	PSNR	MSE	Compression Ratio
Haar	55.4912	0.4285	1.9998
Daubenches	69.7492	0.0830	2.1594

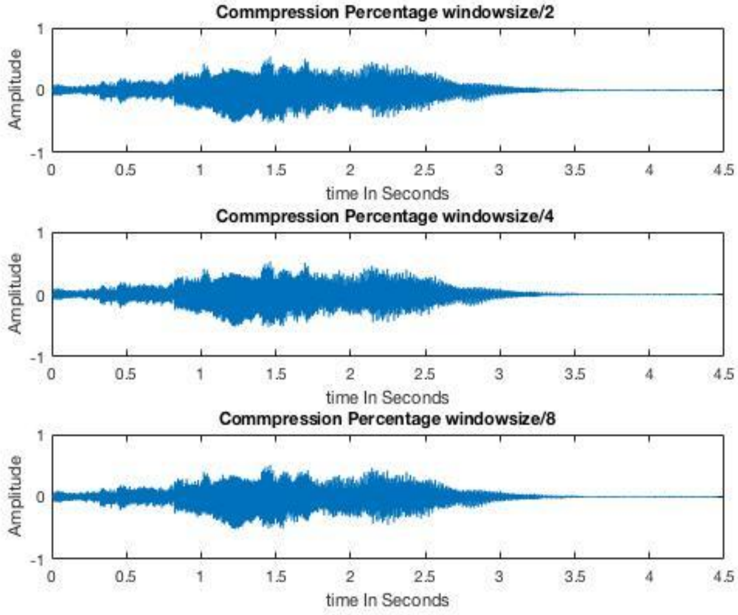
Algorithm 1


orignal Audio
Amplitude
time In Seconds
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5


Commpressed signal
Amplitude
time In Seconds
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5

Algorithm 2


orignal Audio
Amplitude
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
time In Seconds


Commpression Percentage windowsize/2
Amplitude
1
0
-1
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
time In Seconds
Commpression Percentage windowsize/4
Amplitude
time In Seconds
Commpression Percentage windowsize/8
Amplitude
time In Seconds

Conclusion

- After analysing the observations we see that the compression ratio of the haar algorithm is less than daubenches algorithm so the audio compressed by daubenche’s algorithm will require lesser space than that by haar’s algorithm.
- Daubenches algorithm is best suited for lossless compression of speech signals as it has more PSNR and substantially low NMSE.

Thank You