FPGA implementation of DWT for Audio Watermarking Application

Naveen.S.Hampannavar [1], Sajeevan Joseph [2], C.B.Bidhul [3], Arunachalam V [4]
[1, 2, 3] M.Tech VLSI Students, [4] Assistant Professor Selection Grade
VLSI Division, School of Electronics Engineering, VIT UniversityVellore, India
[1]hmpnr.naveen@gmail.com, [2]sajeevanjoseph88@gmail.com, [3]cbbidhul@gmail.com
[4] varunachalam@vit.ac.in

***Abstract*—Digital water marking is a technique of embedding extra information into the multimedia content, which can be extracted to prove the copy rights. Compared to human visual system, audio system is more sensitive. As a result very few audio watermarking algorithms have been robust and imperceptible.**

In this paper we are implementing audio watermarking using discrete wavelet transform (DWT). An audio signal in the form of .wav file is decomposed into multi level DWT coefficients. A watermark signal is embedded in the final level coefficients. The audio is reconstructed from the embedded co-efficient using inverse DWT. The simulation results will be verified by comparing watermarked audio with the original audio for its perceptibility. The watermarked audio will be tested for its robustness towards retaining watermark. Computation of DWT involves large number of arithmetic operations. Hence, a hardware chip for the same would help in achieving real time performance, low power consumption and lesser area utilization. This hardware implementation through FPGA eases the integration of watermarking feature with the existing audio electronic systems.

Keyword - DWT, discrete wavelet transform, audio watermarking, digital watermarking, copyright protection

I. INTRODUCTION

The advancement in computer applications and internet has lead to the increase in production and transmission of multimedia contents. The ownership and copyright of these multimedia data must be protected. In order to do this, the distributed multimedia content must contain some information that proves the ownership of the data. This technique is called digital watermarking which means inserting imperceptible or invisible information into the data. Digital watermarking can be divided into video watermarking, image watermarking and audio watermarking. Many effective and robust watermarking schemes are available for image and video watermarking. Only a few algorithms have developed for audio watermarking, mainly because the human auditory system is more sensitive than the human visual system.

The audio watermark should be imperceptible, robust to signal processing attacks and statistically undetectable to prevent its unauthorized removal. Imperceptibility means that the watermark should not produce any audible distortion to the original audio.

Audio watermarking techniques can be grouped into two namely, time-domain and frequency-transform domain techniques. In time-domain technique, the watermark is directly embedded onto the amplitude of the original audio signal. Time domain audio watermarking is relatively easy but less robust to signal processing attacks like compression, shifting and filtering. In frequency-transform technique, while embedding the watermark information, the amplitude of the frequency transformed coefficients is changed. Commonly used frequency transforms are Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT).

Another popular technique is the spread-spectrum watermarking. This method embeds the watermark signal into audio signal over the entire audible range. Watermark is detected by correlating the watermarked audio signal and the watermark signal. The implementation steps for this technique are very complex.

In this paper, we deal with an effective audio processing algorithm based on 1-D DWT and its VLSI implementation. The DWT decomposes the original audio signal into sub-bands and a binary recognizable image is embedded on to the high resolution sub-bands. Here, DWT is implemented using lifting based architecture [3] in order to reduce hardware complexity.

II. DWT AND LIFTING SCHEME

In this section we brief about discrete wavelet transform and computation of DWT coefficients using one of the methods called lifting scheme.

A. Discrete Wavelet Transform

Wavelets are a particular kind of basal functions analogous to sinusoidal functions used in Fourier analysis, used to represent the signals. They are used in multi-resolution analysis of time domain non stationary signals. The discrete wavelet transforms provide time domain coefficients that represent different band of frequencies, thus giving time localization information. Mathematically, the DWT coefficients can be calculated by the pyramidal algorithm proposed by Mallat [4]. The algorithm involves decomposition of the signal into a low pass and a high pass band followed by sub-sampling in the time domain, to get next level coefficients.

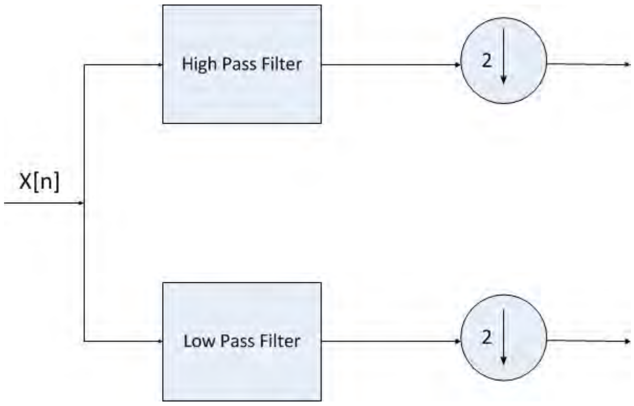

Fig. 1. Decomposition of the signal in the wavelet domain

Computing multi level DWT coefficients would use the above filters recursively. This increases the computation time.

B. Lifting Scheme

Various other techniques to compute DWT coefficients by using basic building blocks that could be arranged in a lattice have been proposed. One such method is lifting scheme [5]. According to this scheme the coefficient are obtained by splitting the original signal into its odd and even samples, and then follows a series of predicting and updating steps. The results are scaled to get the detailed and approximation coefficients.

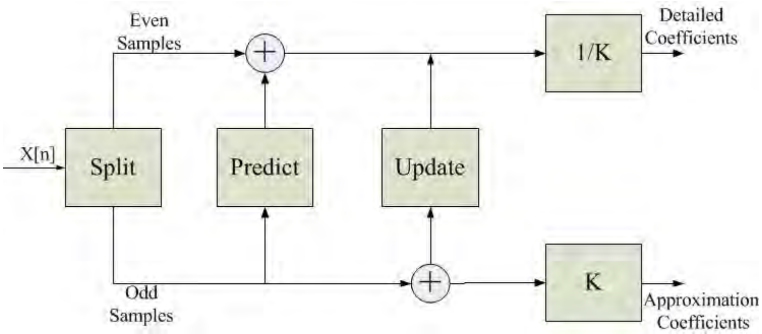

Fig.2. Decomposition Implementation of DWT using lifting scheme

The nature of lifting steps depends on the coefficients of low pass and high pass filters used in pyramidal algorithm. The filter coefficients are determined by the wavelet being used.

III. ARCHITECTURE

A. DWT computation

The wavelet used in computing DWT coefficients of audio samples here is a Daubechies2 (db2). The corresponding lifting scheme is as follows.

The step (t (z)) to predict odd coefficients using even coefficients is called as a dual. The step (s (z)) to update even coefficients using recent odd coefficient is called as a primal. The series of dual and primal steps for db2 wavelet DWT is given as:

$$t_1(z) = -\sqrt{3}$$
$$s_1 = \frac{\sqrt{3}}{4}+\frac{(\sqrt{3}-2)z}{4}$$
$$t_2(z) = z^{-1}$$
$$k = \frac{\sqrt{3}+1}{\sqrt{2}}$$

$$k^{-1} = \frac{\sqrt{3}-1}{\sqrt{2}}$$

Using the above lifting scheme the architecture for one level DWT computation would be given as below:

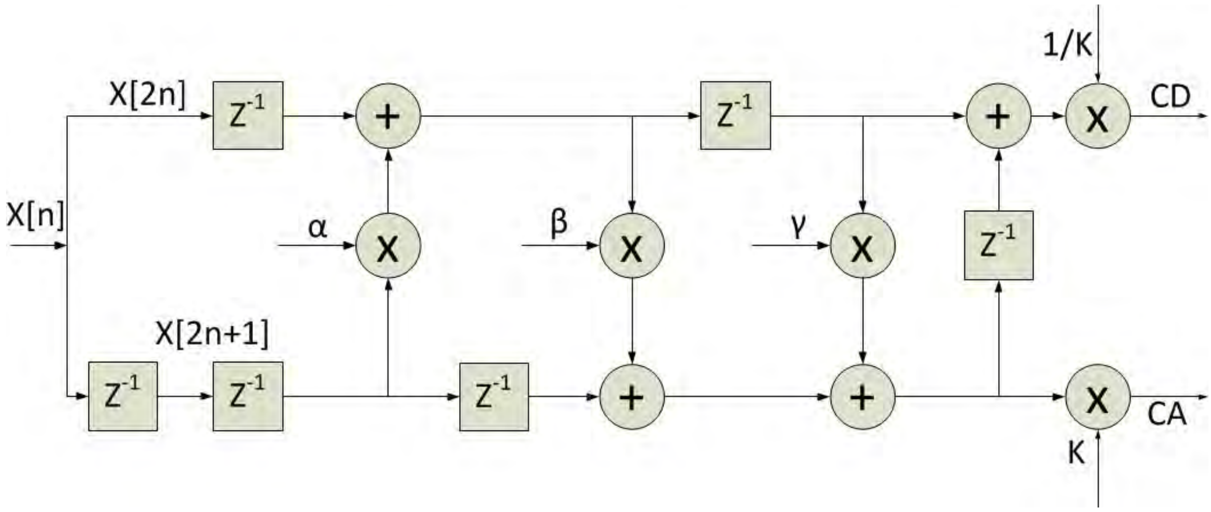

Fig.3. Decomposition Architecture for DWT using lifting scheme

where, $\alpha = -\sqrt{3}$ $\beta = \frac{(\sqrt{3}-2)}{4}$ $\gamma = \frac{\sqrt{3}}{4}$

Similarly, for inverse DWT computation architecture is as follows:

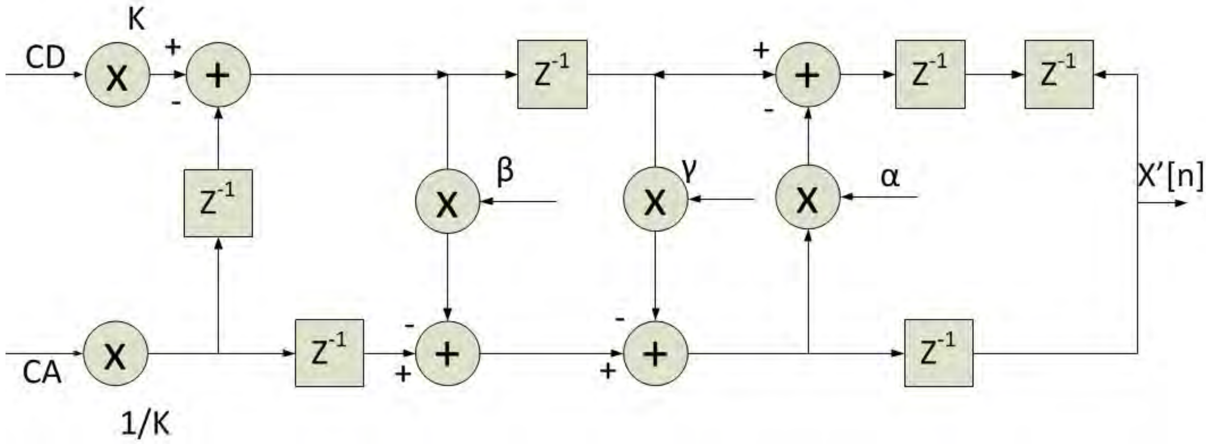

Fig.4. Architecture for IDWT

B. Watermark Embedding Algorithm

The watermarking module takes as input the watermark image, decomposed audio sub-bands (using 2 levels 1D-DWT) and embeds each image pixel onto each DWT coefficient of audio sample as discussed below.

Steps for operations to be done on the image:

- Create a 1dimensional matrix from a 1dimensional gray scale image of size ROW *X* COL.

$$\mathrm{Im} = \{\mathrm{Im}(i,j), 0 < i \le ROM, 0 < j \le COL\}(2D\ image\ matrix)$$

$$\omega = \{\omega(k) = \mathrm{Im}(I,j), 0 < i \le ROM, 0 < j \le COL, 0 < K \le ROM * COL\}$$

- Normalize the 1-dimensional matrix by dividing each element by 255.

$$\omega_N(i) = \{\omega(i)/255, 0 < i \le ROW * COL\}$$

- Watermark is embedded on to the high-resolution sub-band of the 2nd level DWT coefficients [1]. Low resolution sub-band (Approximation coefficients) corresponds to the low frequency components of the audio signal and high resolution sub-band (Detail coefficients) is that of the high frequency components.

C. Watermark embedding steps[2]:

- Using DWT, decompose the original audio signal (X) at level two into low and high resolution sub-bands.

CA2	*CD2*	*CD1*

Fig.5. DWT coefficients after two levels of decomposition

- CD1 is the 1st level high resolution sub-band. CA2 and CD2 are the 2nd level low and high resolution sub-bands respectively. To make the watermark in the embedded audio signal imperceptible, we choose the high resolution sub-band for embedding the binary watermark image.
- The watermark is embedded according to the following equation: $CD2_W(i) = \{CD2(i) + \alpha\omega\}, 1 \le i \le N$

CA2	*$CD2_W$*	*CD1*

Fig.6. DWT coefficients after embedding the watermark image

- Below shown is the architecture for the watermark embedding module

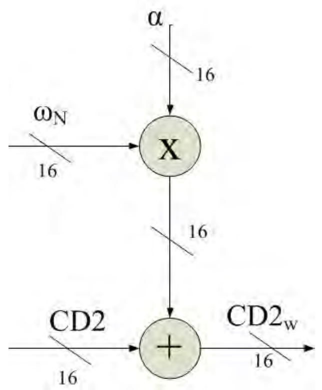

Fig.7. Watermark embedding module

- $CD2_W$ is the 2nd level detail coefficients after embedding process. N is the length of the CD2 sub-band or number of pixels in the watermark image. ω is the watermark signal. α is the scaling factor which determines the strength of the watermark image strength($0 \le \alpha \le 1$). To make the watermark inaudible, use small values for α.
- DWT is performed on segments of audio sample. The segment size is chosen in such a way that the 2nd level DWT sub-band length is equal to the size of the image (number of pixels in the image).

*Length of segment = N * (2^L), L: number of levels of DWT*

- Using IDWT (Inverse DWT), reconstruct the audio signal, $Y = \{CD2_W, CA2, CD1\}$ from the coefficients.

D. Watermark Extraction Algorithm

In order to prove the ownership of the signal, the embedded watermark needs to be extracted. This requires the DWT sub-band (or the original audio signal) onto which the image was embedded, value of the image scaling factor (α) and size of the watermark image.

- Decompose the audio signal to two levels and take the 2nd level Detail coefficients (CD2'). Now, the watermark image can be extracted from CD2' using the following equation:

$$\omega' = \{[cd2_W(i) - CD2(i)]/\alpha\}; 1 \le i \le N$$

IV. IMPLEMENTATION DETAILS

A .wav file of around *10s* length with *44.1 kHz* of sampling rate (CD quality) was used for testing. Each audio sample is represented by *16-bits* in the *.wav* file. A gray scale image of size *160*160* was used as a watermark. Each pixel of the image is coded by *8-bits*. The audio signal is divided into a number of segments. Two level DWT coefficients are computed for each audio segment. Watermark image information was formed into a one dimensional sequence. Each element of the watermark data are embedded into the second level detailed coefficients. The length of the segment N_s, length of watermark N_w and length of second level detailed coefficients N_{CD2} are related as: $N_s = 2^2 * N_{CD2}$ and $N_{CD2} = N_w$.

The audio signal is reconstructed back from the watermark embedded coefficients by computing inverse DWT. The reconstructed audio is tested for imperceptibility of the watermark. MATLAB is used to extract the embedded watermark. The robustness and imperceptibility of the watermark are traded off with each other by suitable embedding strength. The extracted image is comparable enough with the original image used.

The audio samples are represented in their normalized form to range between +1 and -1. An estimate of the range of values for coefficients up to second level was drawn and was found to be between +2 and -2. Hence, every number in the system has been represented in signed 16 bit form with 1bit for integer and 14 bits of fractional part.

V. RESULTS

The original audio sample and watermarked version are compared using Matlab. The shapes of the two samples were matching which confirms that embedded water mark did not disturb the content of the audio.

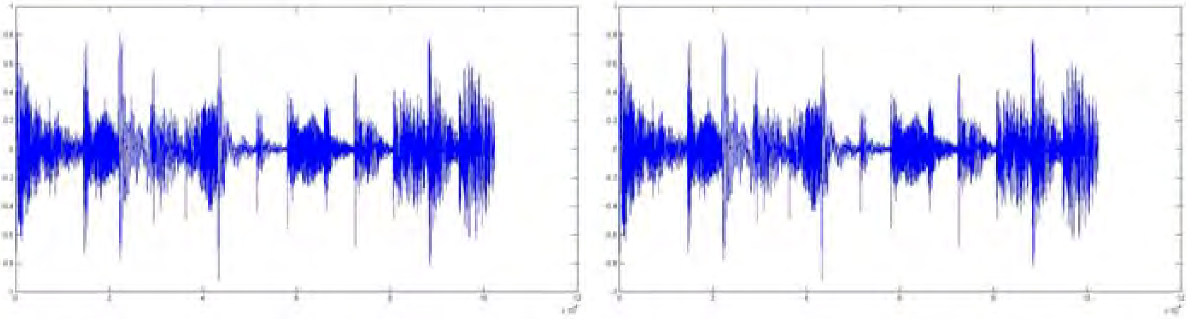

Fig.8 Original audio and its water marked version

The architecture has been synthesized using *Altera Cyclone-II EP2C20F484C7* FPGA device.

TABLE 1
Resource Utilization, power and timing statistics

Total logic elements	Power(mW)	Delay(ns)
2062	72.49	61.01

VI. CONCLUSION

The project shows possibility of hardware implementation for the algorithm to watermark the audio. It briefs the minimum requirements for the implementation. The modules in the architecture have to be pipelined and synchronized in order to make the system real time. While in the current implementation the modules communicated in an asynchronous fashion, the future work would be focused on making it synchronous.

REFERENCES

[1] Ali Al-Haj, Ahmad Mohammad and Lama Bata, “DWT–Based Audio Watermarking”, The International Arab Journal of Information Technology, Vol. 8, No. 3, July 2011.

[2] B. Charmchamras, S. Kaengin, S. Airphaiboon and M. Sangworasil, “Audio Watermarking Technique using Binary Image in Wavelet Domain”, 6th International Conference on Information, Communications & Signal Processing 10-13 Dec. 2007.

[3] Durga Sowjanya, K. N. H. Srinivas and P. Venkata Ganapathi, “Fpga Implementation of Efficient Vlsi Architecture for Fixed Point 1-D Dwt Using Lifting Scheme”, International Journal of VLSI design & Communication Systems (VLSICS) Vol.3, No.4, August 2012.

[4] Mallat, S. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation, "IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 11, 1989, pp.674-693.

[5] Daubechies, I. and W. Sweldens. Factoring Wavelet Transforms Into Lifting Steps. J. Fourier Anal. Appl., Vol. 4, Nr. 3, 1998, preprint.