International Journal of Electronics and Electical Engineering

Volume 2 | Issue 4 Article 9

April 2014

A HIGH SPEED VLSI ARCHITECTURE FOR DIGITAL SPEECH WATERMARKING WITH COMPRESSION

RAJA SREE. AVIRNENI
E.C.E Department, Shri Vishnu Engineering College for Women, Bhimavaram, India, rajasree@svecw.edu.in

K. PADMA VASAVI
E.C.E Department, Shri Vishnu Engineering College for Women, Bhimavaram, India,
padmavasaviece@svecw.edu.in

A HIGH SPEED VLSI ARCHITECTURE FOR DIGITAL SPEECH WATERMARKING WITH COMPRESSION

RAJA SREE. AVIRNENI[1] & K. PADMA VASAVI[2]

[1,2]E.C.E Department, Shri Vishnu Engineering College for Women, Bhimavaram, India
Email: rajasree@svecw.edu.in , padmavasaviece@svecw.edu.in

Abstract- The need to provide a copy right protection on digital watermarking to multimedia data like speech, image or video is rapidly increasing with an intensification in the application in these areas. Digital watermarking has received a lot of attention in the past few years. A hardware system based solely on DSP processors are fast but may require more area, cost or power if the target application requires a large amount of parallel processing. An FPGA co-processor can provide as many as 550 parallel multiply and accumulate operations on a single device, but FPGAs excel at processing large amounts of data in parallel, as they are not optimized as processors for tasks such as periodic coefficient updates, decision- making control tasks. Combination of both the FPGA and DSP processor delivers an attractive solution for a wide range of applications. A hardware implementation of digital speech watermarking combined with speech compression, encryption on heterogeneous platform is made in this paper. It is observed that the proposed architecture is able to attain high speed while utilizing optimal resources in terms of area.

Keywords*: Speech watermarking, Heterogeneous platform, Compression, DWT, latency, Throughput.*

I. INTRODUCTION

With the rapid development of the speech, audio, image, and video compression methods, it is not a difficult task to spread digital multimedia over internet. This makes the protections of digital intellectual property rights and content authentications a serious problem. Hence the technology of digital watermarking has received a large deal of attention. Generally, digital watermarking techniques are based on either spread spectrum methods or changing the least significant bits of selected coefficients of a certain signal transform. For speech watermarking, to ensure the embedded watermark is indiscernible, the audio masking phenomenon is considered together with these conventional techniques. In addition, a speech watermarking system should be robust to various speech compression operations.

Digital watermarking is the technique of inserting specific information (watermark) into audio signal, data, image or video. The specific information is known as watermark and usually used for ownership verification, identification, authentication purpose and protection. Due to rapid progress in wireless communication systems, mobile systems, and smart card technology, information is more vulnerable to abuse. For these reasons, it is important to make information systems secure to protect data and resources from malicious acts. The watermark is extracted or detected at any point where identification or integrity is concerned. Traditional digital watermarking schemes are mainly based on spatial domain or transform-domain, such as discrete cosine transform and discrete wavelet transform.

Speech watermarking system should be robust to various speech Compression operations. Speech compression is a field of digital signal processing that focuses on reducing bit-rate of speech signals to enhance transmission speed and storage requirement of fast developing multimedia. Speech compression is the technology of converting human speech into an efficiently encoded representation that can later be decoded to produce a close approximation of the original signal. Wavelet analysis is the breaking up of a signal into a set of scaled and translated versions of an original *(or mother)* wavelet. Taking the wavelet transform of a signal decomposes the original signal into wavelets coefficients at different scales and positions. These coefficients represent the signal in the wavelet domain and all data operations can be performed using just the corresponding wavelet coefficients.

The watermark is extracted or detected at any point where identification or integrity is concerned. Traditional digital watermarking schemes are mainly based on spatial-domain or transform-domain, such as discrete cosine transform and discrete wavelet transform. Recently, several researchers have proposed effective watermarking schemes, which are more robust against compression, by considering the parameters and steps in compression techniques. In [1] Z. M. Lu, et al. proposed a novel VQ-based watermarking algorithm. After that, various VQ-based watermarking algorithms emerged, such as the algorithms based on codebook partition [2], codebook expansion [3], code words labeling [4], indices statistical characteristics [5], predictive VQ [6], lattice VQ [7], finite state VQ [8], and MSVQ [9], etc. Robust, fragile and semi fragile approaches [10],

[11] are the classification of recent watermarking techniques. The robust watermarking approach protects the copyright identifier of data in which watermarks are not easily removed by attacks.

The fragile watermarking approach confirms content integrity. Recently watermark is also used for quality evaluation of speech [12]. Most watermarking algorithms have been developed for software implementations due to ease of use and upgrading. However, the software implementations have the limitations, speed problem and are vulnerable to the off-line attacks. The high density techniques of current FPGAs will be a highly attractive solution for hardware implementation of the software-based watermarking algorithm as they provide flexibility, easy implementation and high performance [13],[14], [15].

In this paper, hardware implementation of wavelet based watermarking schemes, where the watermarking and compression are jointly considered, is presented.

This paper is organized as follows. Section 2 gives Software implementation of speech watermarking. Section 3 discusses Hardware implementation of speech water marking. Section 4 discusses Implementation on Heterogeneous Platform. Section 5 draws conclusion.

II. SOFTWARE IMPLEMENTATION

Speech compression is used in wireless voice communication, speech encryption, and storage for portable audio devices since the wavelet decomposition offers both time and frequency domain localization, it is more efficient than the traditional Fast Fourier Transform (FFT) technique to decompose and reconstruct the non-stationary signals. The translated and scaled mother wavelets are capable to provide multi-resolution of the non-stationary speech signal through low pass and high pass filter banks.

The heterogenous implementation of digital watermarking of compressed speech signal can be split into two parts:

1. Software Implementation
2. Hardware Implementation.

The flow diagram for the software implementation part is shown in Figure1. To implement the speech compression, a db3 wavelet is considered in this paper with three levels of decomposition.

The speech signals used for testing are .WAV files. This signal is framed into segments and each frame consists of 16 samples as shown in Figure 2. In DWT (discrete wavelet transform) right selection of decomposition levels is significant. For processing speech signals low level decomposition gives better compression ratio compared to higher level decomposition because at higher levels, for example beyond 5 levels, approximating the input becomes poor. Hence a 3 level decomposition is taken into consideration. Energy retained by the signal is important criterion for selecting optimal wavelets. The retained signal energy for different wavelets is shown in Figure 3**.** Based on this experiment Daubechies 3 wavelet preserves perceptual information better than all the other wavelets tested. After the wavelet decomposition of the speech signal, compression involves truncating wavelet coefficients below a threshold.

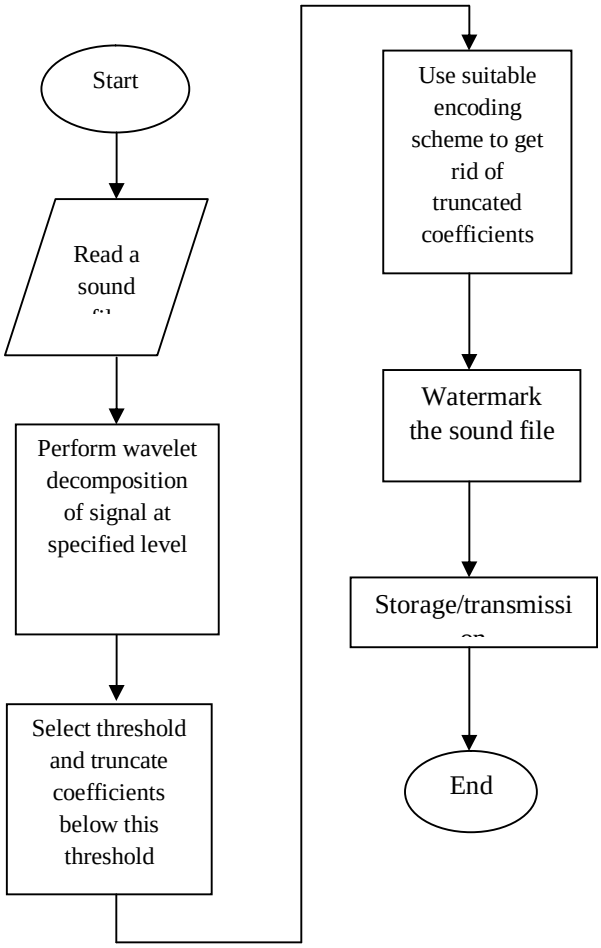


Figure 1. Flow chart for Software implementation

A Frame of 16 Samples

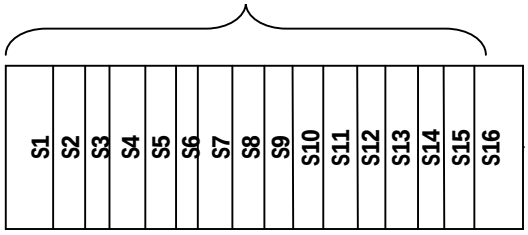


Figure 2. A Frame of 16 samples

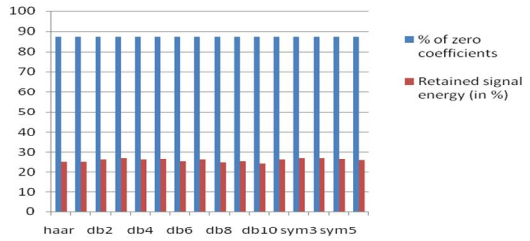


Figure 3. Performance comparison of different wavelets

Most of the speech energy is in the high-valued coefficients, thus the small valued coefficients can be condensed or zeroed and then be used to reconstruct the signal. Two different approaches are available for calculating thresholds:

1. Global Threshold
2. Level dependent Threshold

Global threshold involves taking the wavelet expansion of the signal and keeping the largest absolute value coefficients. Level dependent threshold consists of applying visually determined level dependent thresholds to each decomposition level in the Wavelet Transform. Here global thresholding considered. Signal compression is achieved by first truncating small-valued coefficients and then efficiently encoding them. Four bit Shift register is used at both transmitter and receiver sections to encode the speech signal. After encoding process watermark of random number of bits are inserted into the speech signal then it will be transmitted. Figure 4,5 shows original, 3-level speech compressed signal by using the db3 DWT which gives a compression ratio of 8:1, and the compressed sound can be heard very clearly.

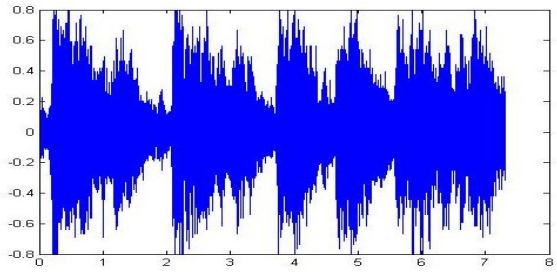


Figure 4. Original speech signal

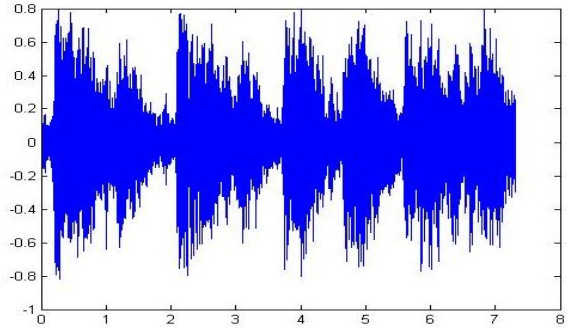


Figure 5. 3 level Compressed speech signal

III. HARDWARE IMPLEMENTATION

Digital speech watermarking is implemented on Heterogeneous platform. Hardware Implementation is as shown in Figure 6. In wavelet analysis, the signal is decomposed into approximate and detail signals. The approximations are the high-scale, low-frequency components of the signal. The details are the low-scale, high-frequency components. The speech signal is passed through high pass and low pass filters. High pass filter which gives detailed component of the speech data. Low pass filter which gives approximation component of speech signal.

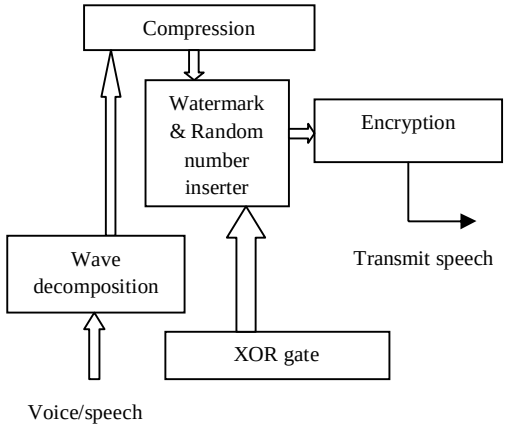


Figure 6. Arrangement for transmitting the speech signal on Heterogeneous platform

TABLE I: DB3 QUANTIZED WAVELET COEFFICIENTS

Filter coefficients	
Low pass filter Coefficients scale up by 2^3	***High pass filter coefficients scale up by 2^3***
a3_in = 4	b3_in= 1
a2_in = 8	b2_in = 2
a1_in = 2	b1_in = -8
a0_in = -1	b0_in = 4

For hardware implementation, the floating-point operations are more complex and speed is very low so, db3 coefficients need to be quantized into binary digits and the resulting less accuracy not that much effect the quality of wavelet transform. Table 1 shows the quantized filter coefficients for db3 wavelet transform. Hardware implementation of low pass filter is shown in Figure 7.

To construct a low pass filter in hardware initially speech signal is divided into frames. Here 16 samples are considered as a frame. 4-stage 16-bit pipeline registers are used to store four continuous speech signals. They are multiplied with the filter coefficients a3_in, a2_in, a1_in and a0_in and then added together. Up-scaled db4 wavelet coefficients a3_in, a2_in, a1_in and a0_in are applied to the digital speech sample **w_in**, the low pass filter results

will be scaled up by 8. So, the adder result needs to be shifted right by 3 bits to get the correct value **y_out**. Here the shifting is signed right shifting. Similar process for highpass filter but filter coefficients will change. The output of the filter is down sampled by 2. The decomposition process can be iterated, with successive approximations being decomposed in turn, so that one signal is broken down into many lower resolution components.

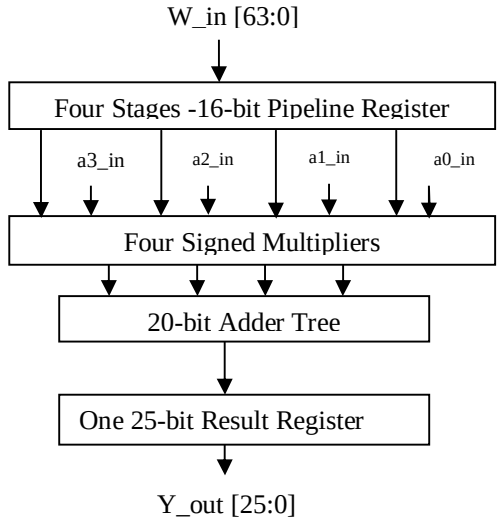


Figure 7. db3 Low Pass Filter

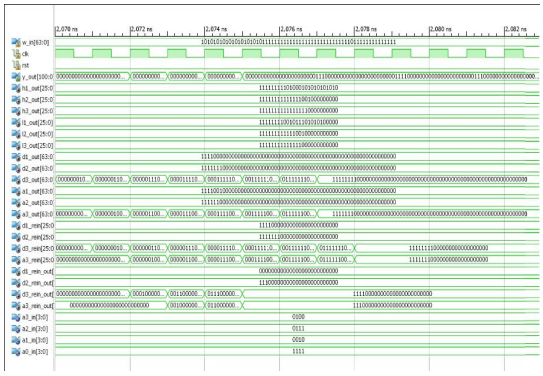


Figure 8. Simulation results of speech watermarking

This is called the wavelet decomposition tree. Here decomposition tree consists of 3 levels. Decomposition tree output is the compressed signal. Speech watermarking means inserting specific information into the speech and it is mainly used for authentication. xoring of the signal is used as a watermark. Finally shift register is used to encrypt to speech signal.

An important part of the logic design process is verifying that the final design is correct and debugging the design if necessary. Logic circuits may be tested either by actually building them or by simulating them on computer. Simulation is generally easier, faster, and more economical. The purpose of simulation is to learn (infer) something about these unknown output distributions, like maybe their expected values, variances. The simulation results of speech watermarking and encryption is shown in Figure.8. Speech watermarking implementation is done using Verilog programming and this Verilog program of speech watermarking is downloaded using ISE12.2 and as per input .wav file is provided as input and simulated on ISim the result is watermarked speech signal. Hardware used is vertex 2 pro board and actual compressed signal is observed.

IV. IMPLEMENTATION ON HETEROGENEOUS PLATFORM

FPGAs and DSP processors have fundamentally different architectures. An algorithm that is well suited for implementation on one device may be very inefficient on the other. A hardware system based solely on DSP processors may require more area, cost or power if the target application requires a large amount of parallel processing or a maximized multichannel throughput. An FPGA co-processor can provide as many as 550 parallel multiply and

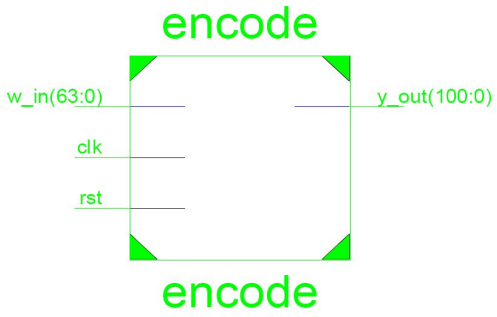


Figure 9. RTL View

accumulate operations on a single device, delivering the same performance with fewer devices and lower power for many applications. Although FPGAs excel at processing large amounts of data in parallel, they are not as optimized as processors for tasks such as periodic coefficient updates, decision- making control tasks, or high-speed serial mathematical operations. It is the combination of both the FPGA and DSP processor that delivers attractive solution for a wide range of applications. For example, a heterogeneous, reconfigurable DSP platform can be ideal for smart cameras that employ pattern recognition. The parallel processing capacity of the FPGA is well suited to image segmentation and feature extraction, while a video and imaging DSP processor is better suited to math-intensive tasks such as statistical pattern classification.

Vertex 2 pro is a heterogeneous platform. An audio CODEC and stereo power amplifier are included on the XUP Virtex-II Pro Development System to provide a high-quality audio path and provide all of the analog functionality in a PC audio system. It features a full-duplex stereo ADC and DAC, with an analog mixer, combining the line-level inputs,

microphone input, and PCM data. The audio system on the XUP Virtex-II Pro Development System consists of a National Semiconductor® LM4550 AC97 audio CODEC paired with a stereo power amplifier made by Texas Instruments.

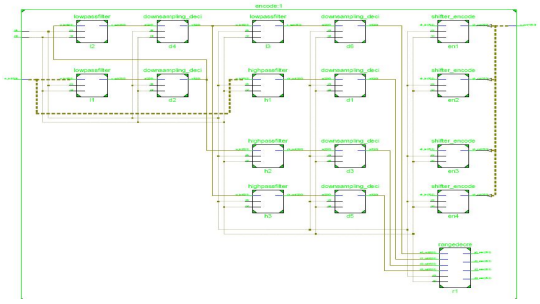
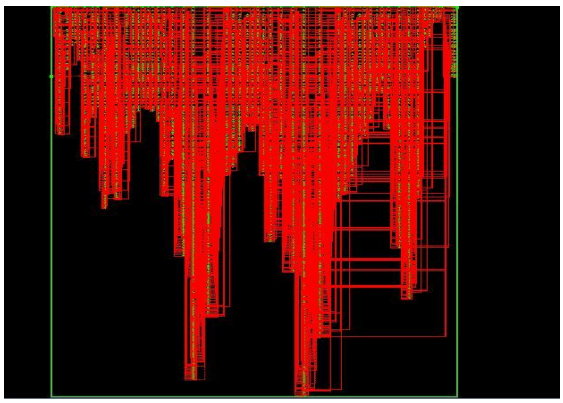


Figure 10. Detailed view of RTL Schematic



Technology level representation of speech watermarking

After the HDL synthesis phase of the synthesis process, the schematic representation of synthesized source file, register transfer logic (RTL) is displayed as shown in Figure 9. This file shows a schematic representation of the pre-optimized design in terms of generic symbols that are independent of the targeted Xilinx device, for example, in terms of adders, multipliers, counters, AND gates, and OR gates etc. Viewing this schematic may help to discover design issues early in the design process. Here w_in is input speech samples.RTL view consists of input speech samples, clock, reset and final watermarked speech signal. The detailed view of RTL Schematic is as shown in Figure 10. After the optimization and technology targeting phase of the synthesis process, the technology view of synthesized source file, as shown in Figure 11.

This schematic shows a representation of the design in terms of logic elements optimized to the target Xilinx® device or "technology," for example, in terms of of LUTs, carry logic, I/O buffers, and other technology-specific components. Viewing this schematic allows seeing a technology-level representation of HDL optimized for a specific Xilinx architecture, which may help discover design issues early in the design process.

Table II gives a summery on hardware resource consumption of speech watermarking on Heterogeneous platform. 368 slices are used out of 10240 available slices.491 slice flip flops are used o

TABLE II HARDWARE RESOURCE UTILIZATION

Logic Utilization	Used	Availabl e	Utilizatio n
Number of Slices	368	10240	3%
Number of Slice Flip Flops	491	20480	2%
Number of 4 input LUTs	1434	20480	7%
Number of bonded IOBs	92	320	28%
Number of GCLKs	1	32	3%
Number of DSP48s	24	128	18%

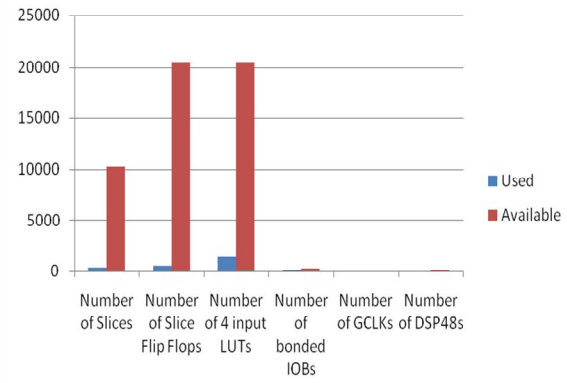


Figure 11. Performance evaluation of proposed hardware architecture

out of 20480, 1434 LUTs used out of 20480,IOB used are 92,available are 320.To increase speech of operation 24 DSP48s are used out of 128 available.

From the Figure 12 utilization of resources is very less in proposed hardware architecture. The Speed of operation is 2.835ns which is very fast compared to previous methods.

V. CONCLUSION

The watermarking research is progressing very fast and various researchers from various fields are focusing to develop robust Speech Watermarking schemes. Great advantages are gained due to using hardware based implementation of Speech watermarking algorithms, such as reduce hardware scheme area, decrease power consumption and increase speed of performance. A Heterogeneous platform combines the advantages of DSP processor and FPGA processor. A heterogeneous system allows better exploitation of pipelining and parallel processing, which are essential to achieve high frame rates and low latency for speech water marking attacked by compression. Proposed hardware attains high speech when compared to previous by 10 times as a heterogeneous platform is used.

REFERENCES

[1] Z.M. Lu and S.H. Sun, "Digital Image Watermarking Technique Based on Vector Quantization," Electronic Letters, Vol. 36, No. 4, pp. 303-305,2000.

[2] Lin and J. Pan, "Robust VQ-based Digital Image Watermarking or Noisy Channel," Proc. ICICIC06, pp. 673-676, 2006.

[3] C. Chang, C. S. Shieh, J. S. Pan, B. Y. Liao and L. Hong, "An Improved Digital Watermarking Scheme in the Domain of vector Quantization,"Journal of Computers, Vol. 17, No.2, pp. 19-26, 2006.

[4] Zhe-Ming Lu, et al., "Digital Image Watermarking Method Based on Vector Quantization with Labeled Codewords," EICE Trans. Inf. And Syst., Vol. E86-D, no. 12, pp. 2786-2789, 2003.

[5] H. C. Huang, F. H. Wang, and J. S. Pan, "A VQ-based Robust Multiwa-termarking Algorithm," IEICE Trans. Fundamentals, Vol. E85-A, No. 7, pp. 1719-1726, 2002.

[6] Y. N. Li, C. H. Liu, Z. M. Lu, J. S. Pan, "Robust Image

[7] Watermarking Algorithm Based on Predictive Vector Quantization," Proc. ICICIC06,Vol. 3, pp. 491-494, 2006.

[8] L. Guillemot, J. M. Moureaux, "Indexing Lattice Vectors in a Joint Watermarking and Compression Scheme," Proc. ICASSP06, Vol. 2, pp.11- 329- 11-332, 2006.

[9] Y. C. Lin, Z. K. Huang, R. T. Pong and C. C. Wang, "A Robust Watermarking Scheme Combined with the FSVQ for Images," Proc. CITA05, Vol. 2, pp. 597-602, 2005

[10] Z.M. Lu, D.G. Xu, S.H. Sun, "Multipurpose Image Watermarking Algorithm Based on Multistage Vector Quantization," IEEE Trans. on Image Processing, Vol. 14, No. 6, pp. 822-831, 2005.

[11] C. I. Podilchuk, E. J. Delp, "Digital Watermarking: Algorithms and Applications," IEEE Signal Process. Mag., Vol. 18, no 4, pp. 3-46, Jul.2001.

[12] Izquierdo and V. Guerra, "An Ill-posed Operator for Secure Image Authentication," IEEE Trans. Circuits Syst. Video Technol, Vol. 13, No. 8, pp. 842852, Aug. 2003.

[13] L. Cai, R. Tu, J. Zhao, and Y. Mao, "Speech Quality Evaluation: A New Application of Digital Watermarking," IEEE Transaction On Instrumentation And Measurement, Vol. 56, No. 1,Page: 45-55 February 2007.

[14] Hussain Mohammed Dipu Kabir1, Syed Bahauddin Alam "Hardware Based Realtime, Fast and Highly Secured Speech Communication Using FPGA" 978-1-4244-6943-7/10/$26.00 ©2010 IEEE

[15] J. Pang, S. Chauhan " FPGA Design of Speech Compression by Using Discrete Wavelet Transform" Proceedings of the World Congress on Engineering and Computer Science 2008 WCECS 2008, October 22 - 24, 2008, San Francisco, USA

[16] Jing Pang, Shitalben Chauhan, Jay Maheshkumar Bhlodia "Speech Compression FPGA Design By Using Different Discrete Wavelet Transform Schemes" Advances in Electrical and Electronics Engineering –IAENG Special Edition of the World Congress on Engineering and Computer Science 2008