ELectronics EXpress

LETTER

IEICE Electronics Express, Vol.11, No.9, 1–7

Hardware architecture for security improved Fallahpour audio watermarking scheme

Claudia Feregrino-Uribe[1a)], **Ernesto Aparicio-Díaz**[1a)], **José Juan García-Hernández**[2b)], **Alejandra Menendez**[1a)], **René Cumplido**[1a)], **and Alicia Morales-Reyes**[1a)]

[1] *Instituto Nacional de Astrofisica, Optica y Electronica (INAOE), Sta. Ma. Tonantzintla, Pue., Mexico*

[2] *Information Technology Laboratory, CINVESTAV, Tamaulipas, Mexico*

a) *cferegrino@ccc.inaoep.mx, e_aparicio@ccc.inaoep.mx, m.menendez@ccc.inaoep.mx, rcumplido@ccc.inaoep.mx, a.morales@ccc.inaoep.mx*

b) *jjuan@tamps.cinvestav.mx*

Abstract: One of the audio watermarking schemes with highest payload published to date is Fallahpour scheme, achieving about 3 kbps [1]. A keybased security improvement to this algorithm is proposed in [2] while significatively maintaining the payload. In this letter, a high throughput and compact hardware architecture for the Security Improved Fallahpour Audio Watermarking Scheme is proposed, demonstrating the suitability of this algorithm for real-time applications.

Keywords: audio watermarking, hardware architectures, FPGA

Classification: Electron devices, circuits, and systems

References

[1] M. Fallahpour and D. Megias: IEICE Electron. Express **6** (2009) 1057. DOI: 10.1587/elex.6.1057

[2] J. J. García-Hernández, C. Feregrino-Uribe, R. Cumplido and R. Parra-Michel: IEICE Electron. Express **7** (2010) 995. DOI:10.1587/elex.7.995

[3] A. Sghaier, S. Areibi and R. Dony: ACM Trans. Reconfigurable Technol. Syst. **3** [3] (2010) 1057. DOI:10.1145/1839480.1839482

[4] J. J. García-Hernández, C. Feregrino-Uribe, R. Cumplido and C. Reta: J. Signal Process. Syst. **64** (2011) 457. DOI:10.1007/s11265-010-0503-8

[5] P. Karthigaikumar, K. Jaraline Kirubavathy and K. Baskaran: Microelectron. J. **42** (2011) 778. DOI:10.1016/j.mejo.2011.01.008

[6] J. J. García-Hernández, C. Reta, C. Feregrino-Uribe and R. Cumplido: IEICE Electron. Express **6** (2009) 1064. DOI:10.1587/elex.6.1064

[7] D. M. Ballesteros L. and J. M. Moreno: Comput. Electr. Eng. **39** (2013) 1192. DOI:10.1016/j.compeleceng.2013.02.006

DOI: 10.1587/elex.11.20140223
Received March 12, 2014
Accepted March 25, 2014
Publicized April 17, 2014
Copyedited May 10, 2014

1 Introduction

Today's research in digital watermarking has increased rapidly as a solution to the piracy problem that has proliferated mainly due to the quick growth of the Internet. To alleviate the problem, the use of watermarking techniques allows to add an imperceptible and statistically undetectable signature to digital media (such as images, video or audio) for content protection. In order to preserve the original signal quality, the modifications after the watermark embedding process must be imperceptible. It is also highly desirable that, depending on the application, the watermark should be robust to media manipulations through signal processing operations; such as lossy compression, filtering, resampling, noise corruption, among others. At the same time, there has been an increased demand for watermarking techniques that may provide high data rates.

Field Programmable Gate Arrays (FPGAs) offer the possibility of fully exploiting the algorithmic inherent parallelism for more demanding applications, such as massive content distribution, real-time military communications and online audio-clips trade. FPGAs are reconfigurable, flexible and physically secure devices with high computational capabilities and offer a fast design cycle [3]. Besides, FPGA-based implementation of data hiding systems seems to be an interesting option since its capacity for parallel processing could allow multi-channel processing. Although several hardware architectures for watermarking algorithms have been proposed during the last decade, it is not a deeply researched topic. These architectures aim to solve both necessities: security and processing speed [4, 5, 6, 7]; where the increased hardware flexibility has been demonstrated through the use of FPGA devices as well as the suitability of data hiding algorithms to be implemented in hardware.

A good candidate to be implemented in hardware is the security improved Fallahpour audio watermarking scheme [2], that besides achieving a high insertion capacity and keybased security, it provides robustness against the most common audio signal processing operations such as echo, added noise, filtering, resampling and MPEG compression (MP3).

The goal of this work is to design a hardware architecture that demonstrates real-time applicability of this security improved and high capacity audio watermarking scheme.

The remainder of this article is organized as follows: Section 2 describes Fallahpour's watermarking scheme as well as the improvement proposed by Garcia et al., Section 3 shows the hardware architecture proposed for this scheme. Section 4 discusses the implementation results and Section 5 concludes.

2 Fallahpour's audio watermarking scheme

The scheme proposed by Fallahpour and Megias [1] is based on the difference between the original and the interpolated amplitudes of the DFT samples as obtained by spline interpolation. A sample is selected for embedding secret information if the difference is lower than a given fraction of the interpolated

DOI: 10.1587/elex.11.20140223
Received March 12, 2014
Accepted March 25, 2014
Publicized April 17, 2014
Copyedited May 10, 2014

value. To obtain the marked DFT samples, the interpolated value is changed according to the secret bit. The embedding steps are as follows:

$k = 1$
for $i = low_{band}$ to $high_{band}$
  if $\mod(i, 2) == 0$
    $e_i = f_i - I_i$;
    if $|e_i| > 2\alpha I_i$
      $f'_i = f_i$;
    elsif $b_k == 0$
      $f'_i = I_i;\ k = k + 1$;
    elsif $\{(b_k == 1) \wedge (e_i \geq 0)\}$
      $f'_i = I_i(1 + \alpha);\ k = k + 1$;
    else
      $f'_i = I_i(1 - \alpha);\ k = k + 1$;
    end;
  end;
end;

where f_i is the magnitude of the i^{th} DFT spectrum sample, low_{band} and $high_{band}$ are the lower and higher limits of a selected band for embedding secret information, I_i is the interpolated value of f_i, α is a threshold, b_k is the k^{th} bit of the secret bit stream, and f'_i is the watermarked value of f_i. The extraction process is as follows:

$k = 1$;
for $i = low_{band}$ to $high_{band}$
  if $\mod(i, 2) == 0$
    $e'_i = f'_i - I_i$;
    if $|e'_i| < 0.5\alpha I_i$
      $b'_k = 0;\ k = k + 1$;
    elsif $(|e'_i| \geq 0.5\alpha I_i) \wedge (|e'_i| \leq 1.5\alpha I_i)$
      $b'_k = 1;\ k = k + 1$;
    end;
  end;
end;

where b'_k is the k^{th} bit of secret bit sequence.

In [2] we proposed a security improvement that consisted in adding a Pseudo-Random Number Sequence (PRNS) in the frequency domain to data samples before applying the insertion algorithm. This improvement kept the perceptual transparency and robustness to all attacks reported by Fallah-

DOI: 10.1587/elex.11.20140223
Received March 12, 2014
Accepted March 25, 2014
Publicized April 17, 2014
Copyedited May 10, 2014

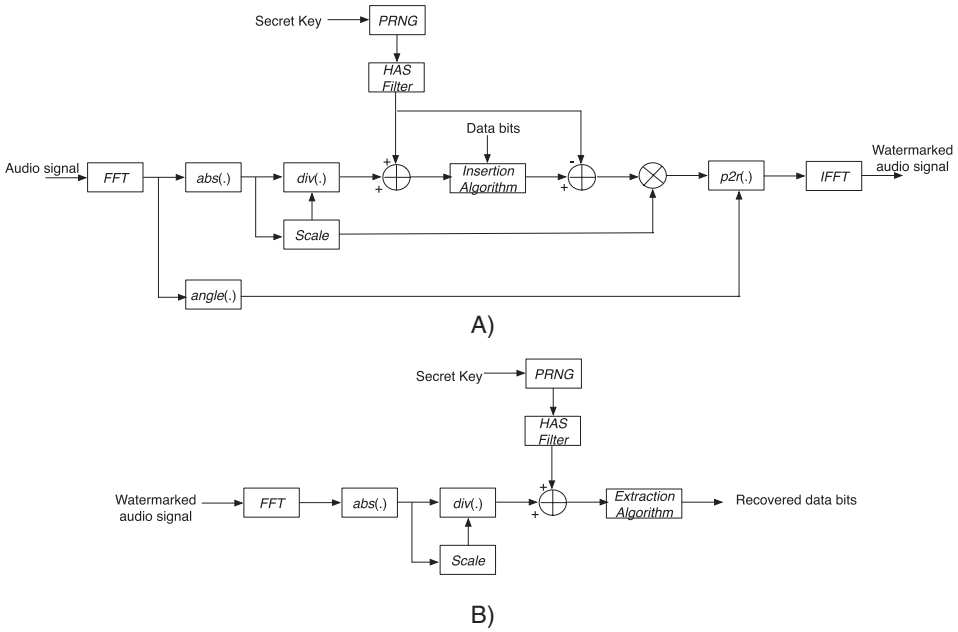

Fig. 1. The key-based improved Fallahpour's audio watermarking system. A) Insertion, B) Extraction.

pour's scheme varying just slightly the payload. Audio processing is performed by blocks of size M, which allows for multiple applications including streaming. The scheme of this secure algorithm is shown in Fig. 1.

As it can be seen from the figure, the keybased security improved algorithm includes the use of a PRNG, a divisor and a multiplier block, and maintains the complexity of the entire system similar to the original scheme while adding the security characteristic. The next section shows the applicability in real-time of this security improved and high capacity audio watermarking scheme.

3 Hardware architecture for security improved Fallahpour scheme

The proposed architecture consists of three main modules for watermark insertion, as it can be seen in Fig. 2. The architecture inputs *Data* and outputs *Marked Data*, audio signals codified with 16 bits of precision (CD quality). The Pre-Insertion module prepares the signal for insertion. Fallahpour insertion module implements the Secure Fallahpour Algorithm. The Post-Insertion module reconverts the watermarked vector to an audio signal. Fig. 2 shows a block diagram of the architecture.

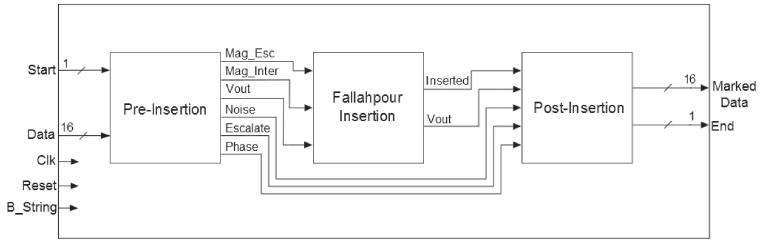

Fig. 2. Proposed architecture for Secure Fallahpour Algorithm.

DOI: 10.1587/elex.11.20140223
Received March 12, 2014
Accepted March 25, 2014
Publicized April 17, 2014
Copyedited May 10, 2014

The Pre-Insertion module is shown in Fig. 3, it gets the FFT transform of a n-samples block using a FFT predesigned Xilinx IPCore block, Rec_to_Pol submodule converts cartesian coordinates to polar ones and it is built in a CORDIC block from Xilinx, the Scalator submodule builds a vector from *Mag* and generates *Escalate* and *Division*, the former is built by placing the first position equal to *Mag*, second and third position equal to the first value, the fourth and fifth values equal to the third one, and so on. *Division* is the result of dividing *Mag* with *Escalate*. The Noise_Add submodule generates a pseudorandom sequence, *Noise*, using a predesigned PRNG block from Xilinx, which is added to *Division* to generate *Div_Noise*. The Interpolation submodule takes *Div_Noise* vector for downsampling and linear interpolation. *Noise*, *Escalate* and *Phase* vectors are used by the Post-Insertion module.

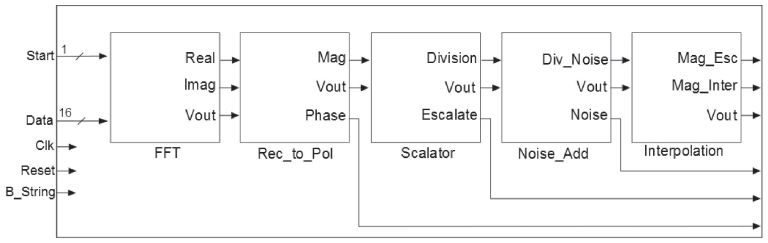

Fig. 3. Pre-Insertion module for the proposed architecture for Secure Fallahpour Algorithm.

In the Fallahpour Insertion module, Fig. 4, it can be seen that signals *Mag_Esc* and *Mag_Inter* are compared, if the error is below a threshold, the *Mag_Inter* vector is multiplied by 1, 1.3 or 0.7, depending on the bit read from *B_String*, according to [2]. The error sign indicates if it should be $+0.3$ or -0.3. If an even sample occurs or it is out of the range [23,700] Hz, then the signal sent out is the scaled one.

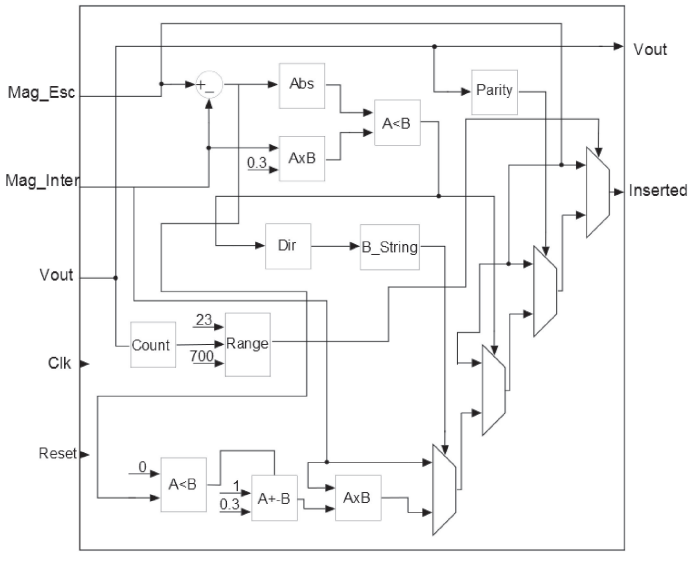

Fig. 4. Fallahpour Insertion module for the proposed architecture for Secure Fallahpour Algorithm.

DOI: 10.1587/elex.11.20140223
Received March 12, 2014
Accepted March 25, 2014
Publicized April 17, 2014
Copyedited May 10, 2014

The Post-Insertion module, shown in Fig. 5, performs the opposite operations of the Pre-Insertion block, except for the Mirror submodule that outputs the second half of Mag_RE as its conjugated complex. This module reconverts the vector to an audio signal.

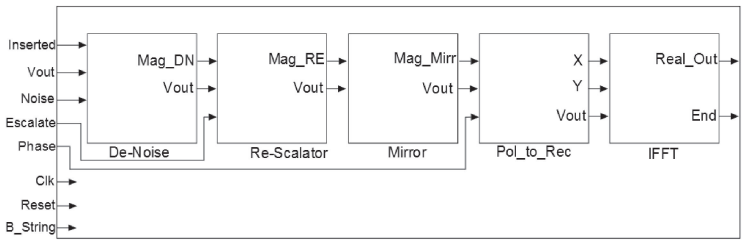

Fig. 5. Post-Insertion module for the proposed architecture for Secure Fallahpour Algorithm.

4 Experimental results

Hardware architectures can be designed at different abstraction levels, from RTL to higher ones using CAD tools. One of these tools commonly used in the industry is System Generator [3], which allows fast development in a friendly environment. Although an architecture designed using System Generator is not as efficient as a hand-made one, the difference in performance is not very significant such that System Generator becomes a very attractive design platform [3].

The proposed architecture was implemented using System Generator v14.2 running on a Intel Core i3 at 2.10 GHz with 4 GB RAM. The device used is a Virtex5 (XC5VSX50T-2FF665) from Xilinx.

Table I shows the operational frequency achieved by the proposed architecture is 63.75 MHz, having a medium footprint and consuming 76% of a medium range device. The maximum path delay from/to any node is 3.213 ns and the latency is 16,984 cycles (266 ns), caused by the FFT block.

The insertion hardware architecture requires higher processing power and area than the extraction hardware architecture, however, extraction algorithm was taken into account when designing the architecture in such a way that some blocks can be reused for this later process.

Table I. FPGA's resources utilized for Security Improved Fallahpour Hardware Architecture

Resource	Amount	Available	Percentage
DSP48Es	93	288	32%
RAMB 18X2x	21	132	15%
RAMB 18X2SDPs	8	132	6%
RAMB 36_EXPs	4	32	3%
Slices	6259	8160	76%
Slice Registers	20920	32640	64%
Slice LUTs	19994	32640	61%
Operation Frequency	63.75 MHz		

DOI: 10.1587/elex.11.20140223
Received March 12, 2014
Accepted March 25, 2014
Publicized April 17, 2014
Copyedited May 10, 2014

To the best of our knowledge, this is the first architecture's design of the security improved Fallahpour watermarking scheme, therefore a direct comparison to other works is not possible. A comparison with the implementation of other audio watermarking algorithms would not be fair as they may use other watermarking techniques. However, to provide an overall view of the results obtained in this research, other related research works are discussed. For example, Karthigaikuar architecture [5] reports an audio watermarking chip using chirp spread spectrum. It is mentioned by the authors that the watermarked signal can be transmitted and received at 1800 MHz with a PSNR of 55.67 dB. Another architecture is reported in [7], where authors propose a speech-in-speech hiding scheme in wavelet domain suitable for a real-time implementation. Garcia-Hernandez et al. [4] propose a hardware implementation of a Rational Dither Modulation (RDM) algorithm-based data hiding system in the Modulated Complex Lapped Transform, able to process 819 channels, for embedding, and 682 channels, for detection, of CD-quality audio signals. That system is adequate for watermarking-based multi-broadcasting applications. Also, Garcia-Hernandez et al. presented in [6] an efficient FPGA implementation of an RDM-QIM algorithm that allows real-time multi-channel behaviour. The throughput for insertion and detection stages is 84.8 and 60.2 mega samples per second (Msps), respectively.

Results from the proposed architecture demonstrate its suitability for real-time audio communications such as VoIP. Its high operation frequency stands out its capability for high speed processing such that it could be possible to use the architecture in a complete system working by audio blocks and using multiplexers, so the input could be changed every M number of samples, and at the output the repository changes every M number of samples with a delay equal to the circuit latency. In this way, the speed is enough for processing up to 1460 CD-quality audio signal channels per second and its footprint makes it useful for embedded environments.

The proposed architecture could be used on data streaming. Blocks of 4096 samples cause 92 ms latency approximately, allowing continuos transmission. Therefore, the proposed design is also useful for radio or TV broadcasting.

5 Conclusion

This paper proposed an FPGA hardware architecture for a key-based improved Fallahpour's audio watermarking scheme. The results have demonstrated the suitability of this algorithm to achieve high throughputs, being useful for demanding applications where speed is one of the main requirements. Due to the proposed architecture's compact footprint, it can be used as accelerator in microprocessor-based systems for embedded applications or as a core in custom architectures.

Acknowledgments

Authors wish to thank CONACyT, Mexico, for his support under grants number CB-2010-1-50910, 243961 and 345675.

DOI: 10.1587/elex.11.20140223
Received March 12, 2014
Accepted March 25, 2014
Publicized April 17, 2014
Copyedited May 10, 2014