IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)
Volume 6, Issue 6, Ver. I (Nov. - Dec. 2016), PP 25-30
e-ISSN: 2319 – 4200, p-ISSN No. : 2319 – 4197
www.iosrjournals.org

Fpga Based Implementation of Video Authentication using Sensor Pattern Noise

[1]Urvi M. Mittal,[2]Rohita P. Patil
Elect. & Telecom. Department, SKNCOE, Vadgaon (BK), Pune.

Abstract: *Video authentication plays an important role in identifying camera spoofing attacks occurred. Video Authentication can be accomplished by different methods on various platforms. Sensor Pattern Noise (SPN) is used for identifying tempered video or source identification of the video. The source is either authentic or non-authentic is identified using Forward and Inverse wavelet transform such as DWT and IDWT for performing reliable operations on frames of video extracted. MLE and MMSE estimations are calculated for denoising of the frame and for extraction of PNU noise which is treated as the fingerprint. Extraction of PNU is the basis of identification. The system is implemented on ZYNQ Xc7Z020 (ZedBoard) ARM/FPGA SoC development board using Xilinx ISE as well as VIVADO software for designing in VHDL language. The hardware resource utilization is reduced by implementing modified filter bank for DWT and IDWT.*
Keywords: *Video authentication; Sensor Pattern Noise;DWT; Denoising;Hardware Resource.*

I. Introduction

Due to development and rapid innovation in digital technologies, video applications are penetrating into our lives at excessive speed from conventional television to modern accessible communication media such as Internet/Intranet, wireless communication and products such as VCD/DVD.To identify, preserve, collect and examine information which is stored or encoded, usually to provide proper evidence of a specific activity, digital forensics is used. Common video forgery operations are used for interrupting the video sequences by inserting or deleting objects within the frame.

Video surveillance, forensicinvestigations, law enforcement and content ownership are applications where video authenticity is essential.In scenario such as in court of law, it is vital to prove the trustworthiness of any video that is used as an evidence. Whereas in another scenario, imagine astationary video recorder for surveillance is placed on the pillar of a railway platform to investigate every activity, it would be very easy to defame an individual or certain activity from the video sequences. In other way, it would also be possible to insert objects into this video which are sourced from different camera. Thus, video authentication is a process which confirms that the given content in a video is authentic and similar to the picture when captured.

Video authentication is essential to stop illegal copying problem and redistribution of movies i.e. to stop piracy over internet. A typical video authentication system proves the integrity of video and shows that whether the given video has been tampered or not. The process starts with extracting frames from video and identifying signature embedded into the video which is nothing but watermark. The video integrity is computed by generating new authentication data for the present video. When both the new authentication data and original data are compared, there is conclusion that the video is treated as authentic or tampered.

The most feasible method for detecting video forgery such as piracy is source camera identification because it points out the device from which the video is captured. This method gives information about the association of image with characteristics of camera such as model and brand. Thus, Video authentication using sensor pattern noise is the reliable method as it is inherent for camera.

The system implements video authentication using a pixel-non uniformity noise scheme on a hardware platform. For DWT and IDWT, a modified filter bank (FB) based implementation is designed. The paper presents array-based implementation of an image denoising algorithm.Implementation on FPGA is preferred because it provides a means for rapid implementation of proposed architecture and supports functional parallelism.

The paper is organized as follows: The techniques of video authentication, related problems are discussed in section in I. Literature review is presented in section II. In section III, implemented system is explained. In section IV, Results are discussed. Conclusion and set of remarks presented at the end of the brief is described in section VI.

II. Related Work

Various methods for video authentication are utilized for implementation of Noise Extraction from video for identification which are review in following section.

Dateand time, camera settings, or the serial number of the camera such tags can be written to the image as metadata captured by digital cameras in EXIF(EXchangeable Image Format) or XMP (eXtensible Metadata Platform). But this metadata is manipulated or removed easily from image present. Therefore, alternatives should be designed for unique identification [1].

A novel idea is proposed for camera identification based on supervised learning. To find camera model, image features are computed in spatial and wavelet domain and a support-vector-classifier is trained. An accuracy of 78-85% is achieved from five different cameras for identifying and classifying images and for which a multiclass classifier is used [1].

In a digital camera pipeline, imageprocessing stages leave footprints as forensic signatures. The left footprints consist of various pixel defects, of inappropriate responses in the charge-coupled device sensor, black current noise, and may be originated from proprietary interpolation algorithms involved in color filter array. Depending on these artifacts, imaging device identification methods are based. To identify the source camera with SVM classifiers three sets of forensic features such as binary similarity measures, image-quality measures, and higher order wavelet statistics in conjunction are explored [8].

The method calculating similarity by minimum average correlation energy (MACE) filter is proposed. MACE filter is robust to noise and have high identification accuracy for compressed videos without denoising. Extracts the RPRNU (Reference PRNU) & TPRNU (Test PRNU) from the reference video and text video respectively. For enhancing the detection accuracy, RPRNU is transformed to the form of MACE filter. The MACE filter is made for minimizing the energy of correlation plane and ensuring a sharp correlation peak at the origin [4].

The method transforms the RPRNU to the form of MACE filter correlation function and the correlation plane by calculating cross-correlation between the TPRNU and the transformed RPRNU and calculate peak to correlation energy (PCE) in the correlation plane. The PCE calculated by the MACE filter is higher than that calculated by NCC.

III. System Implementation

The camera is inherently having a PNU footprint (P) into the captured image or video, because of defects in manufacturing of sensor. PNU extraction is done using algorithms such as Maximum Likelihood estimate (MLE) and MMSE estimation for denoising of image. The operations are performed on subbands of the images which are converted using Wavelet Transforms such as Forward and Inverse Discrete wavelet Transform. The denoised frame obtained is subtracted with the original frame for PNU estimate. The system diagram implemented is shown in Fig.1.

For identifying whether image is authentic or tampered, the estimated PNU is compared with the reference PNU of the camera. The reference PNU is obtained by averaging the past values of PNU’s not directly by camera.

$$\boldsymbol{P} = \frac{\mathbf{1}}{\boldsymbol{N}}\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{N}} \boldsymbol{P}^{\mathbf{1}} \qquad (1)$$

Where, P is the reference PNU estimated by averaging the PNU values and P^1 is the PNU values of the past samples of the camera used.MMSE estimation is used for Wavelet-domain denoising as it is efficient and natural image boundaries are taken into consideration compared to other denoising algorithms. The original frame and PNU extracted frames are correlated and value of correlation is computed. Depending on this correlation value there is a conclusion that whether the frame is tampered or not.

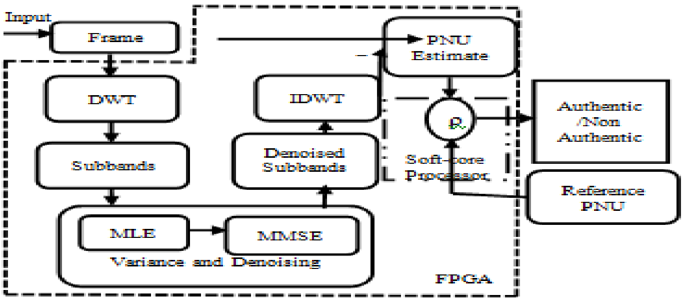

Fig.1 System Diagram

If video is taken as an input it is defined in MATLAB. The frames are extracted from the video in MATLAB. Before conversion into subbands, the input image is resized into 256 x 256 pixel size and converted into grayscale. The processed frame obtained is given as an input to DWT.

A. DWT

In the system design,the wavelet coding techniques calculates the co-efficient of a transform which can be used to decorrelate the pixels of an image and are very efficient for coding than the original pixels.Decorrelation of the pixels means the calculation of subband values from the pixel values for easier computation.

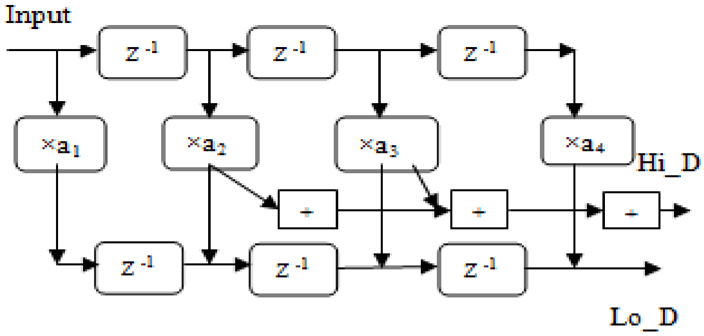

Fig.2 Modified Filter Bank Implementation

The Modified Filter Bank is implemented using db2 wavelet transform which redues area and computational requirement for hardware implementation.. The coefficients obtained for db2 wavelet are given in Table I.

The implemented Filter bank has reduced number of multipliers and adders which reduces the hardware required for the system.The 2-D DWT is computed up to three levels and the converted subbands are passed for variance estimation and denoising leaving only LL3 band i.e. LL band of 3rd level DWT.

Table I Coefficients of db2 Wavelet Filter

	Lo_D	Hi_D
1	-0.129409522550 91	-0.482962913144 61
2	0.224143868041857	0.83651630373746
3	0.836516303737465	-0.22414386804185
4	0.482962913144691	-0.12940952255091

The filter Coefficients for db2 filter are represented as

$$\boldsymbol{Lo_D(z) = a_1 + a_2 z^{-1} + a_3 z^{-2} + a_4 z^{-3}}$$
$$\boldsymbol{Hi_D(z) = b_1 + b_2 z^{-1} + b_3 z^{-2} + b_4 z^{-3}} \qquad (2)$$

B. Variance and Denoising

The subbands pixels *S(i,j)*are denoised using MLE and MMSE estimation. The variance is estimated depending on local neighborhood of the pixel in the wavelet domain. The denoising is performed with a square mask of 3×3 for uniformity. The MMSE estimation calculates the variance σ using the relation given below:

$$\min_{W=\{1,2,3,4\}}\{\widehat{\sigma_w}^2\} \qquad (3)$$

The denoising model assumes that component fields are having zero mean and unknown variance distribution. The model is highly robust and effective.

$$\overline{S(i,j)} = \frac{S(i,j) \times \widehat{\sigma^2(i.j)}}{\sigma^2(i,j) + \sigma_0^2} \qquad (4)$$

The above equation shows the relation for a denoised subband. The value for σ_0 is taken as 5 for eight bit pixels. Thus, using this process all denoised subbands are recovered and after that inverse DWT (IDWT) is applied to obtain the denoised frame ($\bar{I}$) from the original image (I).

$$\bar{I} = IDWT(all\ subbands\ \bar{S}) \qquad (5)$$

The obtained frame is subtracted from the original frame to get the PNU estimate.
The PNU estimate is obtained as P^1

$$P^1(i,j) = I(i,j) - \overline{I(i,j)} \qquad (6)$$

The refrence PNU and estimated PNU are correlated to obtain correlation value which is denoted as ρ.

$$\rho = corr(P, P^1) = \left(\frac{\left(P^1 - \overline{P^1}\right)(P - \bar{P})}{\left\|P^1 - \overline{P^1}\right\| \|P - \bar{P}\|} \right) \qquad (7)$$

Here, $\bar{P}, \overline{P^1}$ denotes mean value of pixel in P and P^1.
Respectively. The value of ρ above a tresholde value of 0.3 is denoted as authenticated video of correct camera.

IV. Results and Discussion

This section discusses about performance of the system and implemented architecture. For experimentation purpose, two cameras are used for identification. The frames were extracted from video for the experiment. The sensor pattern noise extraction algorithm is used to calculate the correlation coefficient.

A.MATLAB simulation of Sensor pattern noise Algorithm.

In this simulation, first the input image of random size is taken and then resized into size of 256 × 256. If the image is color image then it is converted into the grayscale image by using the rgb2gray function. Then the conversion of the image into subbands is done by using the orthogonal db2 filter for easy functioning. After DWT of frame is calculated, minimum variance is estimated by squaring the subband pixel value and taking variance 5 for eight bit pixel for denoising of frame. This denoised subbands are used for performing IDWT of the signal which are then using reconstruction coefficients to obtain denoised image as shown in Fig.3. This denoised image is again compared with original frame. The PNU is extracted from the comparisonwhich is a pixel by pixel computation.

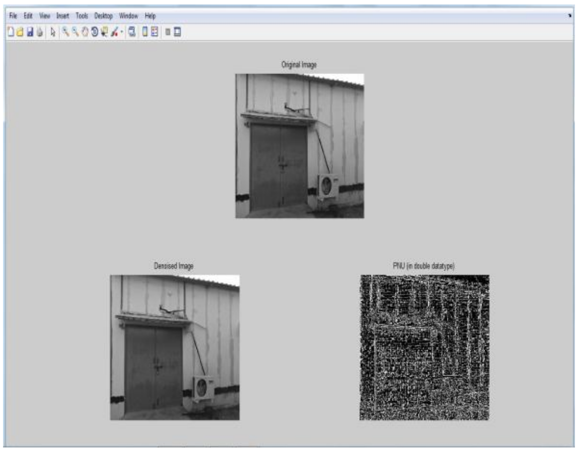

Fig. 3 MATLAB simulation results

For the experiment the resolution to the frame is set as 256×256. The Design is written in VHDL and synthesized using Xilinx Vivado 2015.3, simulations were performed to test the waveforms. The extracted frame of video can be loaded on to the BRAM for applying the algorithm.

B. Hardware Results for frame

First the input image of random size is taken from camera 2 and then resized into size of 256x256. If the image is color image then it is converted into the grayscale image by using the rgb2grayfunction. In Fig.4 original image converted to

Fig. 4 Results obtained for frame (a) Original Image (b) Denoised image (c) PNU Image

grayscale of size 256 X 256 is shown, which on applying algorithm gives denoised image and PNU extracted image. The output is shown on command window which shows that source identified as 2 with correlation calculated for each camera as shown in Fig.5.

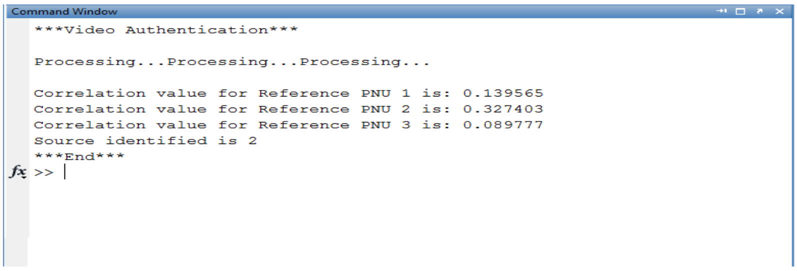

Fig.5 Command Window Output

The approach was evaluated on ZYNQ XC7Z020 (Zedboard) ARM/FPGA SoC development Board which has combination of ARM processing system and series-7 programming logic with AXI interconnections. FPGA was preferred because it reduces time required for processing because of its parallel functioning and its reconfigurable properties. The experimental setup is shown in Fig.6.

Fig.6 Experimental Setup

Table II: Resource Utilization

Resource	Estimation	Utilization (%)
Flip-flop	4608	4
LUT	8328	16
Memory	70	1
BRAM	110	79
DSP48	110	50
BUFG	1	3

Table II shows information related to device utilization analysis. Flip-Flop available are 106400 out of which 4608 are used for implementation which is equal to 4% of available resources. Number of slice LUTs used is 16% which gives 8328 usages out of 51,200. Usage of memory LUT is70 out of 17400 which is equal to 1%. Utilization of BRAM is 110 out of 140 which counts 79% of available. Number of BUFG used are 1 out of which counts 3% of available.

V. Conclusion

In this paper, the architecture for hardware acceleration of video authentication algorithm using pixel non-linearity noise was proposed to identify the correct camera. Our algorithm is able to accurately authenticate source camera with 90-95 % accuracy.

The MFB approach for DWT and IDWT was implemented which reduces the hardware requirements. An architecture for denoising process was implemented which optimized hardware requirements and performance using square mask. Hardware prototyping was implemented on Xilinx Zynq XC7Z020 ARM/FPGA SoC development Board.

Acknowledgment

I would like to thank my beloved and respected Guide Prof. Rohita R. Patil for her encouragement and support. I gratefully acknowledge her for imparting me valuable basic knowledge of Digital Image Processing.

References

[1] Pande*et al.*: "Hardware Architecture For Video Authentication Using Sensor Pattern Noise "in *IEEE transactions on circuits and systems for video technology*, vol. 24, no. 1, January2014.

[2] S. Chen, A. Pande, K. Zeng, and P. Mohapatra, "Video source identification in lossy wireless networks," in Proc. IEEE Int. Conf. Comput. Commun., Mini-Conf., pp. 215–219, Apr. 2013.

[3] X. Kang, Y. Li, Z. Qu, and J. Huang, "Enhancing source camera identification performance with a camera reference phase sensor pattern noise," *IEEE Trans. Inf. Forensics Security,, vol. 7, no. 2, pp. 393–402,* Apr. 2012.

[4] D. Hyun, C. Choi, and H. Lee, "Camcorder identification for heavily compressed low resolution videos," *Comput. Sci. Convergence, pp.* 695–701, 2012.

[5] A. Pande and J. Zambreno, "Poly-DWT: Polymorphic wavelet hardware support for dynamic image compression," *ACM Trans. Embedded Comput. Syst., vol. 11, no. 1, pp.6:1–6:26, Apr. 2012.*

[6] S. Murugesan and D. Tay, "New techniques for rationalizing orthogonal and biorthogonal wavelet filter coefficients," *IEEE Trans. Circuits Syst. Regular Papers, vol. 59, no. 3, pp. 628–637, Mar. 2012*.

[7] C.-T. Li and Y. Li, "Color-decoupled photo response non-uniformityfor digital image forensics," *IEEE Trans. Circuits Syst. Video Technol.,*vol. 22, no. 2, pp. 260–271, Feb. 2012.

[8] M. Kharrazi, H. T. Sencar, and N. Memon, "Blind source camera identification,"in*Proc. ICIP*, Singapore, Oct. 24–27, 2004.