2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing

A reversible self watermarking framework for Integral Images

N. P. Sgouros[1], J. N. Ellinas[2], M. S. Sangriotis[1]

[1]Department of Informatics and Telecommunications, University of Athens,
Panepistimiopolis Athens, P.C.15784, Greece

[2]Department of Electronic Computer Systems, Technological Educational Institute of Piraeus
Athens, P.C. 12244, Greece

***Abstract*—Integral Imaging (InIm) is a promising three-dimensional (3D) content creation and delivery technique. The simplicity of the method constitutes a straightforward and compact way for acquiring and displaying 3D data. However issues like integrity and authentication of a transmitted InIm are not discussed in the literature. In this paper we propose a jointly optimized InIm fragile self watermarking and compression framework which can be used in order to protect sensitive InIm data from tampering. The main concept is based on a disparity based residual embedding framework that combines increased data embedding capacity with improved compacting properties for the encoded image. The effectiveness of the proposed encoder is evaluated using objective quality metrics over a large InIm dataset.**

I. Introduction

Over the past few years, autostereoscopic display systems [1] have greatly evolved and are now widely used for both specialized high quality applications and everyday 3D viewing needs. Since in autostereoscopic display systems all required optical components are integrated in the display device they allow viewing from multiple simultaneous viewers without the use of additional gadgets.

One of the most promising autostereoscopic display systems is based on the principle of Integral Imaging (InIm) initially proposed by Lippman [2]. As InIm provides two-dimensional (2D) parallax to the viewer it is currently considered as a near ideal multiview technique [3]. The operational principle of an InIm capturing setup is based on the acquisition of images of small portions of an object which are called elemental images (EIs), through a lens array (LA) placed in front of a Charged Coupled Device (CCD) as shown in Fig. 1(a). On the display side an appropriate LA is placed over an LCD display to reconstruct the 3D captured scene to the observer as shown in Fig. 1(b). Two characteristic InIms of a computer generated fish [4] and an optically acquired dice along with the original objects are shown in Fig. 2(a),(b) respectively.

The main bottlenecks in a stereoscopic system pipeline include the required large transmission bandwidths, storage capacities, and computational power needed to process huge data volumes. In addition, authentication is an obligatory task for many applications during the decoding stage to ensure that the original data are not tampered during transmission or storage operations.

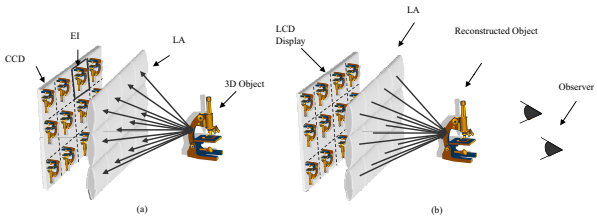

Fig. 1. An InIm (a) acquisition and (b) display setup.

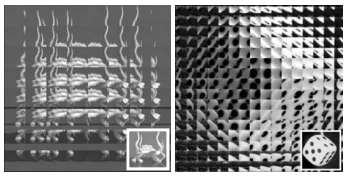

Fig. 2. The InIms of (a) a fish and (b) a dice.

It should be noted that the large amount of redundant data contained in autostereoscopic image sets enhances their ability to store extra information which can be used for authentication purposes. The proposed method utilizes the information capacity and inherent redundancy of an InIm to embed a portion of its EIs to the rest of the EIs in the InIm, along with additional authentication data like checksums during the encoding process, in order to detect tampering, hence providing a joint framework for InIm compression and authentication. The encoding process utilizes a standard approach for InIm coding based on a 1D rearrangement of the EIs and a video encoder as proposed in [5]. The watermarking method is based on a reversible fragile self watermarking scheme, using the image data itself and the disparity information among neighboring EIs for watermarking purposes. The specific algorithm uses Pixel Value Differencing (PVD) and combines data embedding efficiency with implementation simplicity at both the encoder and decoder sides [6].

In section II, we provide an overview of the encoding scheme and the reversible watermarking technique which are used. In section III we present the implementation details of the proposed framework. In section IV we evaluate the

978-0-7695-4712-1/12 $26.00 © 2012 IEEE
DOI 10.1109/IIH-MSP.2012.71

performance of the proposed framework and provide useful conclusions.

II. Encoding and Watermarking

A. Encoder overview

The efficiency of the proposed InIm encoding scheme depends on the exploitation of the different types of redundancies present due to the inherent properties of an InIm. The proposed encoder is based on the 2D Hilbert traversal scheme to initially rearrange the EIs in a one-dimensional (1D) stream retaining the strong correlation of originally adjacent EIs, in the reordered 1D stream of EIs. It has been shown [5] that the high locality preservation property of the Hilbert curve increases the redundancy of the neighboring EIs providing more efficient encoding with respect to other traversal schemes characterized by strong directionality. The locality preservation property of the Hilbert curve is then utilized within a video encoding manner to maximize prediction efficiency between adjacent EIs in the 1D stream and hence reduce the distortion of the watermarked InIm. In order to demonstrate the proof of concept we utilize a simple MPEG-2 [7] like approach where the EIs in each group are marked as I-EIs which are the reference or cover frames and B-EIs which are bidirectionally predicted from their preceding and following I-EIs. The resulting stream is formulated as a motion stream IBIB… and an equally weighted prediction of the B-EIs is performed from their adjacent I-EIs. The embedding unit hides the disparity compensated B-EIs into their preceding I-EIs, while this can be expanded to embed arbitrary B-EIs in each reference I-EI to enhance the authentication objective.

B. InIm Watermarking

According to the human visual system (HVS) properties, distortions near edges are less noticeable than those in smooth areas, which signifies that the pixels near edges can hide more secret bits than those in smooth areas. Thus, different amounts of data can be embedded in different image blocks according to the degree of smoothness in the cover image. However different approaches can be used for determining the embedding capacity of an image area like the one proposed in [8].

In the context of the PVD algorithm the cover image is divided into a number of non-overlapping two-pixel blocks p_i and p_{i+1}. The difference $d = g_{i+1} - g_i$ in the intensities g_i, g_{i+1} of these two pixels is classified into one of a number of predetermined ranges. The number of bits that can be embedded in this block of pixels is $n = \log_2(u_k - l_k + 1)$ where u_k and l_k are the upper and lower bounds of the corresponding intensity range. A new difference d' is computed by

$$d' = \begin{cases} l_k + b & \textit{for } d \geq 0 \\ -(l_k + b) & \textit{for } d < 0 \end{cases} \quad (1)$$

where b is the decimal value of the embedded n bits from the whole binary stream. The new gray values of the two pixels are defined by

$$(g'_i, g'_i) = \begin{cases} g_i - \lceil m/2 \rceil, g_{i+1} + \lfloor m/2 \rfloor & \textit{if d odd} \\ g_i - \lfloor m/2 \rfloor, g_{i+1} + \lceil m/2 \rceil & \textit{if d even} \end{cases} \quad (2)$$

where $m = d' - d$. The predetermined ranges are selected so that for small differences, a small number of bits is embedded in the pixel block as the block belongs to a smooth intensity area where intensity variations are viewable. For large differences, more bits are embedded since the pixel block belongs to a textured area where intensity variations are less perceptible. At the detector, the difference between two consecutive pixels is classified into one of the predetermined ranges. This difference value minus the lower limit l_k of the range implies the decimal value of the n embedded bits. In the simplest setup, let f_1 and f_2 be the cover image and the image to be embedded respectively. The residual R of the two images is calculated as:

$$R = f_1 - f_2 \quad (3)$$

Once the residual is recovered, the two images are readily available at the decoder side. If the residual is embedded into the cover image without any loss, both images may be exactly recovered at the decoder side. If this is not feasible, the residual is subjected to lossy compression. In that case, the cover image is exactly recovered while the embedded image is recovered with some distortion. This type of reversible watermarking process used in [9] produces residuals which still contain a lot of energy, since the method does not take into account the inter-image redundancy between the cover and embedded images. Instead of using Eq.3 for calculating the residual image, in our framework we employ a disparity compensation procedure for calculating the residual of the images that will be embedded in the cover images. This approach further reduces the information content of the residual images and hence improves the quality of the reconstructed image. In detail the image that will be embedded is segmented into blocks and the corresponding blocks in the cover image are identified using disparity estimation. The residual or the Disparity Compensated Difference (DCD) for a block b_{ij} is defined as:

$$r(b_{ij}) = \sum_{(x,y) \in b_{ij}} [b^1_{ij}(x + dv_x, y + dv_y) - b^2_{ij}(x, y)] \quad (4)$$

where b^1_{ij}, b^2_{ij} denote the cover and embedded image intensities respectively and dv_x,dv_y are the disparity vector components. The disparity vector for a corresponding block is estimated by:

$$\mathbf{dv}(b_{ij}) = [dv_x dv_y]^T = \arg \min_{(dv_x, dv_y) \in S} |r(b_{ij})| \quad (5)$$

where S is the search area window size and the metric used in the disparity estimation process is the minimum Mean Absolute Difference (MAD). The performance gain between the direct difference of two EIs and the disparity compensated ones, is visualized by the respective residuals in Fig. 3(a) and (b) respectively. In Fig. 3(c) the disparity vector field of the

compensated difference between the two EIs is also shown. It is clear that the energy of the residual EI is much lower and therefore less data are needed to be embedded in the cover EI. Obviously, the two disparity vector fields for the bidirectional prediction case must be also embedded without any loss, since they ensure the accurate reconstruction of the disparity compensated EI at the decoder.

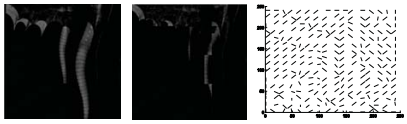

Fig. 3. (a) Difference, (b) Disparity Compensated Difference, (c) Disparity Vectors.

It should be noted at this point that on the average case the residual after JPEG compression with a high quality factor of 95%, results in a file of 56,920 bits. In addition each of the two disparity vector fields, resulting from the bidirectional disparity compensation in a 1D stream of EIs with size of 256x256 pixels when 16x16 pixel blocks are used and a search area window size of 7 pixels, are 512 signed numbers which are expressed as a binary string of 4,096 bits. On the other hand the calculated mean embedding capacity of the cover EI, implementing the PVD embedding scheme, approximates the value of 99 kbits for the InIms evaluated in our dataset. Apparently, the cover image has enough space to carry the residual and the disparity vectors that will be used to recover the embedded EI at the decoder.

III. Implementation

The proposed scheme is illustrated in Fig. 4. The EIs of the target InIm are reordered in a 1D stream using the Hilbert traversal scheme, so that consecutive EIs are highly correlated. The rearranged EIs are considered as a group of consecutive video frames and may be processed as a video stream. In what follows the implementation details for the disparity estimation and encoding procedure are elaborated for an InIm where the EIs form an IBIBIB... stream. The I-EIs of the stream are used as cover images for embedding the information needed for the B-EIs recovery. The B-EIs are bidirectionaly predicted residuals from their adjacent I-EIs, using equal weighting factors. The quality of the JPEG compressed B-EI which will be embedded in its preceding I-EI can be controlled so that it is compatible with the information capacity of the I-EI. In all cases the disparity vectors are losslessly compressed. The selection of the weighting factors is arbitrary and their tuning for every group of frames could provide better disparity estimation at the cost of higher computational complexity. However this choice suffices for the purposes of this paper and is made in order to improve the sub-optimal equal weight case.

Disparity compensation is performed on 16x16 pixels blocks using exhaustive search over an area of 7 pixels. Since pixel accuracy is used in our framework the disparity

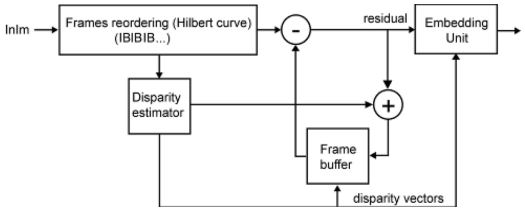

Fig. 4. Block diagram of the encoder for the proposed reversible watermarking scheme.

vector components are integers which are further losslessly compressed. The residual B-EI is JPEG compressed to reduce its size and subsequently embedded into the immediately preceding I-EI, along with the corresponding two disparity vector fields. The embedding unit employs the PVD algorithm for hiding the residual B-EI along with its two vector fields. The embedding capacity of each I-EI is pre-estimated and the compression quality of the residual of the B-EI is adjusted, so that it can be incorporated in the cover I-EI. The watermarked I-EI contains the information needed to recover each group of IB EIs. At the decoder, the compressed residual B-EI is extracted along with its respective disparity fields. Thus, the initial I-EIs are recovered and the B-EIs are further recovered from I-EIs using the disparity vectors and the decompressed residuals. The I-EIs are recovered without distortion, as watermarking is reversible. The recovered B-EIs are distorted due to the compression of the residuals, which is at that stage unavoidable to match the embedding bit-rate of the cover I-EI. Obviously, the embedding bit-rate ability of the cover I-EI is InIm dependent. Images with textured regions provide increased embedding capacity and thus allow the use of high quality factors for the embedded images and vice versa.

IV. Results and conclusions

A. Experiments and Results

The proposed framework was tested on a number of synthetic and physically acquired InIms including the ones shown in Fig. 2(a),(b). The typical dimensions of the computationally generated InIms were 2048x2048 pixels in two sets, containing 8x8 or 16x16 EIs of 256x256 or 128x128 pixels in each EI respectively. For optically acquired InIms we utilized the simple capturing setup described in[10]. The dimensions of the physically acquired InIms varied in size ranging from approximately 1280x1280 to 2048x2048 with a total of 16x16 EIs in each image.

There are primarily two different approaches for evaluating the performance of the image encoding techniques, based on subjective or objective evaluation metrics. The most used metric in the literature for objective 2D image quality assessment is the peak-signal-to-noise-ratio (PSNR). However, this measure estimates the overall image quality and does not take into account the special structure of the displayed InIm. In detail, as neighboring EIs are simultaneously displayed to produce the 3D representation, they should exhibit similar

quality characteristics in order to produce viewable stereoscopic results. In this case large variations of the PSNR values can greatly reduce the 3D effect while PSNR values remain high. It is therefore necessary to calculate the fluctuation of the PSNR value over the InIm. In this work the mean PSNR value ($\overline{PSNR}$) and its standard deviation ($\sigma PSNR$) are calculated using the PSNR values of each of the EIs in the InIm. The values for the $\overline{PSNR}$ and $\sigma PSNR$ for an InIm assembled of $K \times L$ EIs are given by Eqs.6 and 7 respectively.

$$\overline{PSNR} = \frac{1}{K \cdot L} \sum_{s=1}^{K} \sum_{t=1}^{L} PSNR_{s,t} \tag{6}$$

$$\sigma PSNR = \sqrt{\frac{\sum_{s=1}^{K} \sum_{t=1}^{L} PSNR_{s,t}^2}{K \cdot L} - \left(\frac{\sum_{s=1}^{K} \sum_{t=1}^{L} PSNR_{s,t}}{K \cdot L} \right)^2} \tag{7}$$

The performance of the proposed reversible watermarking method is finally evaluated using the above defined quality measures. The fluctuation of the embedding capacity of the cover EIs shows that it is depended on the image content. The embedding capacity is also affected from the effectiveness of the used PVD data hiding algorithm. The quality factor for the compression of the residual B-EIs is selected so that their total resulting size along with the data resulting from the lossless compression of the two disparity vector fields is slightly less than the corresponding capacity of the cover image in each group. The total amount of the embedded bits is less than the capacity of the cover image, leaving enough space for overhead information like the length of the stream, the dimensions of the EIs and an authentication sequence.

The objective quality of the watermarked I-EI and the reconstructed B-EI in each group for the InIm in Fig. 2(a) is illustrated in Fig. 5. The mean PSNR of the I-EIs is about 41.85 dB showing that the embedded information is imperceptible throughout the stream. The mean PSNR of the B-EIs is about 41.31 dB revealing that the average quality of the recovered B-EIs is also high. The relative standard deviation, given as $\sigma PSNR/\overline{PSNR}$ which quantifies the fluctuation of the $\overline{PSNR}$ value is 0.11. It should be noted that similar values for the PSNR and the relative standard deviation, which is crucial for the stereoscopic effect were attained throughout the InIm dataset used in our experiments, validating the robustness and effectiveness of the proposed framework.

It is evident that the quality of the B-EIs may be improved by using a more refined disparity compensation procedure or the residual B-EIs encoding algorithm. Additional gain can be pursued by tuning the forward and backward weighting factors, with a computational overhead, or using a different number of B-EIs between the cover images.

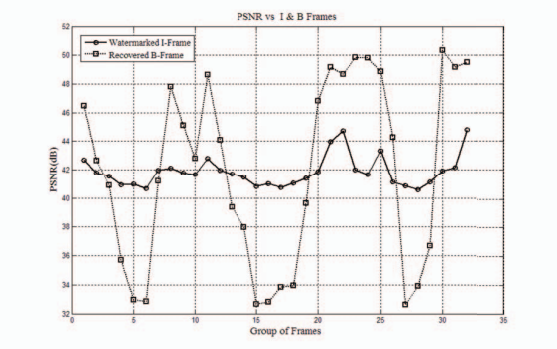

Fig. 5. Objective quality of the watermarked I-EIs and the reconstructed B-EIs, for all IB groups in the 1D stream created using the InIm of Fig. 2.

B. Conclusions

In this work, we propose a joint compression and reversible self watermarking algorithm for InIms and evaluate its efficiency over an InIm dataset. A stream of IBIBIB ... is formed yielding a coding gain of at least 50 % due the number of I-EIs used. In addition the PSNR values achieved throughout the set and the low PSNR fluctuation values indicate the high quality of the reconstructed InIms without compromising stereoscopic effect or the data embedding capability. Data tampering will result in a poor quality of the reconstructed InIm as a failure in reconstructing the B-EIs will cause severe degradation of the stereoscopic effect.

REFERENCES

[1] M. Halle, "Autostereoscopic displays and computer graphics," in *ACM SIGGRAPH 2005 Courses*, ser. SIGGRAPH '05. New York, NY, USA: ACM, 2005. [Online]. Available: http://doi.acm.org/10.1145/1198555.1198736

[2] G. Lippmann, "La photographie integràle," *Comptes-Rendus Academie des Sciences*, vol. 146, pp. 446–451, 1908.

[3] J.-Y. Son and B. Javidi, "Three-dimensional imaging methods based on multiview images," *J. Display Technol.*, vol. 1, no. 1, p. 125, Sep 2005.

[4] S. S. Athineos, N. P. Sgouros, P. G. Papageorgas, D. E. Maroulis, M. S. Sangriotis, and N. G. Theofanous, "Photorealistic integral photography using a ray-traced model of capturing optics," *Journal of Electronic Imaging*, vol. 15, no. 4, p. 043007, 2006.

[5] N. Sgouros, I. Kontaxakis, and M. Sangriotis, "Effect of different traversal schemes in integral image coding," *Appl. Opt.*, vol. 47, no. 19, pp. D28–D37, Jul 2008. [Online]. Available: http://ao.osa.org/abstract.cfm?URI=ao-47-19-D28

[6] D.-C. Wu and W.-H. Tsai, "A steganographic method for images by pixel-value differencing," *Pattern Recogn. Lett.*, vol. 24, no. 9-10, pp. 1613–1626, Jun. 2003. [Online]. Available: http://dx.doi.org/10.1016/S0167-8655(02)00402-6

[7] K. Rao and J. Hwang, *Techniques & standards for image-video & audio coding*. Prentice Hall PTR, 1996.

[8] H.-C. Huang and W.-C. Fang, "Authenticity preservation with histogram-based reversible data hiding and quadtree concepts," *Sensors*, vol. 11, no. 10, pp. 9717–9731, 2011. [Online]. Available: http://www.mdpi.com/1424-8220/11/10/9717/

[9] D. Coltuc and I. Caciula, "Stereo embedding by reversible watermarking: Further results," in *Signals, Circuits and Systems, 2009. ISSCS 2009. International Symposium on*, july 2009, pp. 1 –4.

[10] G. Passalis, N. Sgouros, S. Athineos, and T. Theoharis, "Enhanced reconstruction of three-dimensional shape and texture from integral photography images," *Appl. Opt.*, vol. 46, no. 22, pp. 5311–5320, Aug 2007.