EFFICIENT COMPRESSION METHOD FOR INTEGRAL IMAGES USING MULTI-VIEW VIDEO CODING

Shasha Shi [1, 2], Patrick Gioia [1] and Gérard Madec[2]

[1] Orange Labs. France Télécom
[2] Image and Information Processing Department. Télécom-Bretagne

ABSTRACT

Integral imaging is an attractive auto-stereoscopic three dimensional (3D) technique for next generation 3DTV. To improve its video quality, new techniques are required to effectively compress the huge volume of integral image (II) data. In this paper, a new compression method implemented by multi-view video coding (MVC) is provided and used for sub-images (SI). SI is an alternative form of 2D image transformed from original II. Each SI represents the 3D scene from parallel viewing directions and contains superior compression capabilities than original captured elemental images (EI). For this reason, we consider arranging the group of SIs as the format of multi-view video (MVV) and then encode the generated MVV by MVC standard. Experimental results show that our proposed compression approach improves the compression efficiency when compared to the traditional MPEG-4/AVC compression method for II.

Index Terms— Integral imaging, image compression, sub-images, multi-view video coding (MVC)

1. INTRODUCTION

Integral imaging was first invented by G. Lippmann in 1908. Due to its extraordinary advantages in terms of free-viewpoint visualization, it is regarded as a promising technique for future 3D TV and has become a vigorous research topic for the past two decades. Better than conventional stereoscopic technologies, it is able to provide high quality 3D images with full parallax and continuous viewing angle without any need of additional devices such as dedicated glasses [1]. In spite of its obvious advantages, the current integral imaging system still contains some inborn drawbacks such as limited image resolution, depth range and viewing angle [1].

Different from traditional binocular vision systems which simulate the human 3D perception procedure by transmitting correspondence images into left and right eyes separately, the integral imaging system allows the generation of true 3D images by reproducing the light rays of the 3D scene with different colors and intensities in different direction. For this reason, it is theoretically capable to eliminate the inconsistency between convergence and accommodation in human viewing system, supply a real 3D experience for the viewers without any uncomfortable feelings such as eye strain or headache.

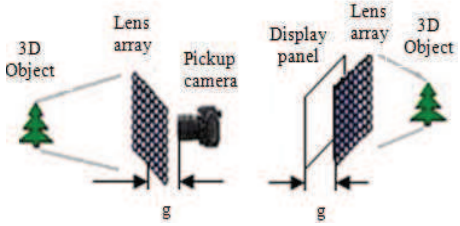

Fig. 1-1 Optical setup for generating (left) and displaying (right) 3D scene in integral imaging system

The acquisition and display process of the integral imaging system can be performed optically in the way illustrated in Figure 1-1. In the pickup setup, the light rays emanated or reflected from the 3D object are captured through an array of micro convex lenses or pinholes and then recorded by a conventional 2D image sensor such as charged-couple device (CCD), as shown in the left picture of Figure 1-1. The captured II is composed of numerous elemental images (EIs) recorded through different lenses, each EI represents a slightly different view of the 3D scene. While in the visualization side, the recorded II will be displayed behind the same lens array to reproduce the light rays in the reverse direction to the captured ones so as to reconstruct the 3D scene, see the right picture of Figure 1-1.

In a typical integral imaging system, the gap between the CCD and the lens array in pickup part is equal to the distance between the displaying panel and the lens array in displaying device, denoted by '*g*' in Figure 1-1. According to this gap, the integral imaging system can be classified into three modes, namely 'real mode', 'virtual mode' and 'focal mode' [4]. In this paper, we will focus on the last case. In this situation, there is no limitation in viewing depth for the displayed 3D scene but its viewing resolution is constrained by the density of EIs i.e. the number of micro lenses in the applied lens array [5]. For example, to display a 3D image with resolution $m*n$, the required resolution of II is $(m*w)*(n*h)$ where w and h are the number of pixels in one EI, and m, n are the number of EIs. Therefore, in order

978-1-4577-1302-6/11/$26.00 ©2011 IEEE

to produce a high quality 3D image, the integral imaging system demands an extremely high definition II with significant amount of data, thus an efficiently compression approach is needed for practical implementation of integral imaging system.

Since integral images obviously exhibit high correlation between different EIs, it is not surprising that previous compression attempts sought to remove these redundancies using JPEG or MEG-4 for EIs [2, 8, 9]. In a different approach, [6] and [7] proposed a new idea to analyze the integral image data, transforming the original II into an alternative format named sub-image (SI), and demonstrated that the SI contains several better characteristics for compressing than EIs. However, their work is still constrained on the single integral image with static scene, and does not involve the moving picture instances. In our research, with the purpose of further exploiting the special properties of SI for video sequence compression, we propose to reorganize the SIs as multi-view video and encode them by multi-view video coding (MVC) technique. According to the simulation, our proposed approach shows an outperformance result compare to the precedent methods.

This paper is organized as follows. The generation process and properties of sub-image declared in paper [6] and [7] are introduced in section 2. Section 3 presents our proposed compression method based on MVC. Simulation results are reported in section 4. Finally, conclusions and discussion are drawn in section 5.

2. SUB-IMAGE CHARACTERISTIC

Figure 2-1 illustrates the concept of generating sub-images from integral image. The pixels at the same position in the EIs are extracted to compose the corresponding sub-image [3]. Since there are 16 elemental images in Figure 2-1, each EI consists 3x3 pixels so 9 different sub-images will be generated with size 4x4, repeat this work for every II and finally generate a sequence of SIs for the entire II series. According to the optical property of convex lenses, pixels from the same location of the EIs record the intensity and color of parallel light rays. For this reason, all the pixels in a given sub-image represent the perception of a 3D scene from a specific viewing angle and different SIs possess different viewing directions.

Due to their particular structure, SIs possess two helpful features for efficient compression than original EIs, which are explained by Figure 2-2. The first one is the fact that the perception angle of each sub-image is a constant value that is independent from the 3D position of the object, and the corresponding SIs from successive integral images always possess the same viewing angle [7]. For example, the sub-images collecting the information of slope light rays in Figure 2-2 record the 3D scene from the viewing angle as follows:

$$\theta_i = \arctan\left({}^{y_i}/_{f}\right) \qquad (1)$$

where f is the focal length of the convex lens, y_i is the distance between the *i-th* pixel and the optical axis in the EI

The second useful characteristic of SI is that the object sizes from corresponding SIs with same viewing angle are always invariant and also independent from the object 3D location [7]. For example, the size of the 3D object represented by the horizontal light rays in Figure 2-2 is observed regardless of its actual location.

The valuable properties presented above can be exploited to compress the series of SIs using MVC, and the entire compression process will be explained detailed in next section.

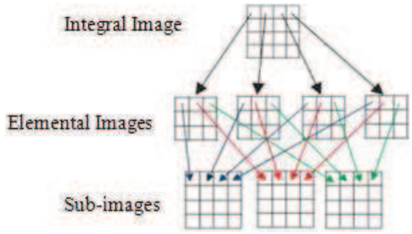

Fig. 2-1 Sub-image generation scheme

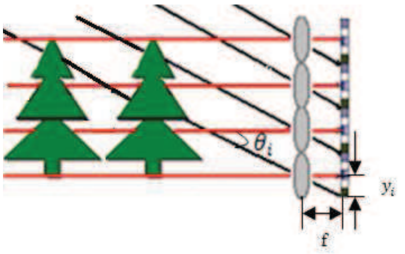

Fig. 2-2 Viewing angle and object size are invariable for corresponding sub-images

3. ENCODE SCHEME WITH MULTI-VIEW VIDEO CODING

Figure 3-1 illustrates the scheme of the proposed integral images compression approach. The whole procedure of this method can be divided into three steps. At first, generate the sub-image through captured II by the regular transformation introduced in section 2. Secondly, arrange the generated SIs as a multi-view video (MVV). Finally, encode the organized MVV by multi-view video coding (MVC). Correspondingly, in the reconstruction part, we need to rearrange the decoded SIs into II and then reproduce the 3D scene with the reconstructed II.

To remove all the redundancy among SIs both in spatial and temporal domain, we considered organizing the sequence of SIs as MVV. As illustrated in Figure 3-2, the SIs from different viewing angles are regarded as the multi-view images in spatial domain and the successive SIs with the same viewing angle constitute the video sequence in temporal domain. Sub-images from different viewing angles are scanned in spiral order so as to minimize the motion compensation between them [8]. For example, in Figure 3-3 the SI made up of the center pixels from each elemental image will serve as the reference view denoted as V_0. The remaining SIs, denoted as $V_1, V_2 \ldots V_8$, are composed of the other pixels of EIs. These 9 different sub-images are scanned successively in spiral order, and then the correlation among them is exploited by inter-view prediction of MVC.

Once the sub-images are rearranged as MVV, the multi-view video coding standard is used to encode this video sequence. We employed the JMVC_2_1 software from the Multi-view Video Coding (MVC) project of the Joint Video Team (JVT) and the ITU-T Video Coding Experts Group (VCEG) to encode our SIs sequence. The structure of employed MVC is shown in Figure 3-4, in this coding algorithm we combine the hierarchical B prediction in temporal domain with the inter-view disparity estimation in spatial domain and apply the inter-view prediction for all the pictures in the video.

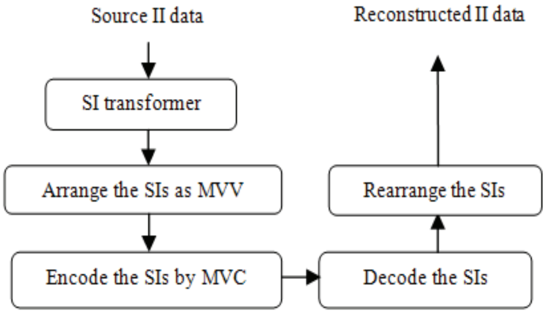

Fig. 3-1 Flow chart of the proposed integral images compression system

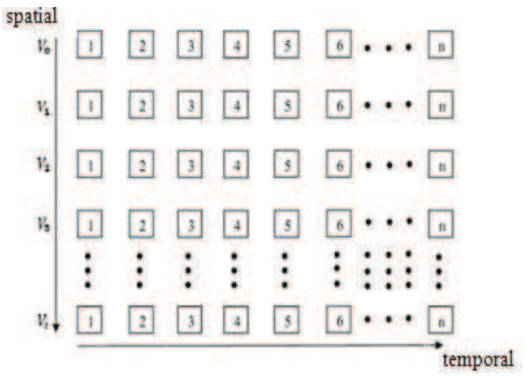

Fig. 3-2 Converting SIs into MVV

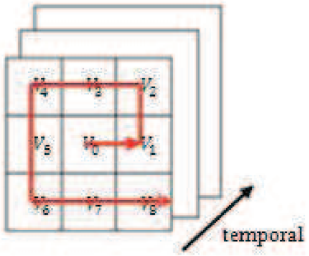

Fig. 3-3 Spiral scanning topology for SIs

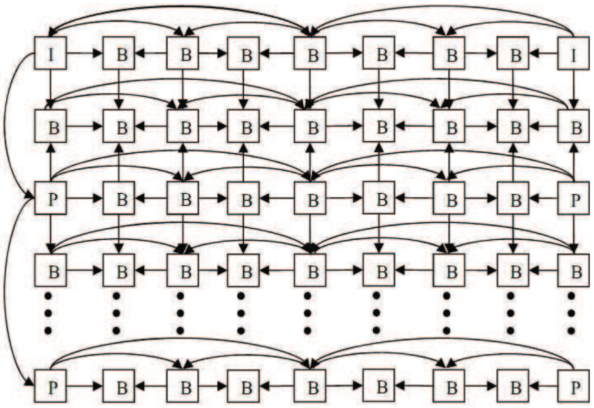

Fig. 3-4 Scheme of MVC

4. EXPERIMENTS AND RESULTS

The tests were carried out through a sequence of 30 frames integral images with resolution 1920x1440. Along with the 30 frames, the object moves both in depth and planar direction. The number of elemental images in each II is 640x480 where each EI contains 3x3 pixels. As a result, we obtained 3x3 sub-images with distinct viewing angles from each II with the resolution 640x480. Therefore, the final constructed multi-view video is composed by 9 sequences from different viewing angle, each one possess 30 frames.

Figure 4-1 shows an example of integral image made up of 640x480 elemental images.

Figure 4-2 shows the generated sub-images from the integral image in Figure 4-1. As we mentioned before, 9 SIs with different viewing angles were constructed in our simulation.

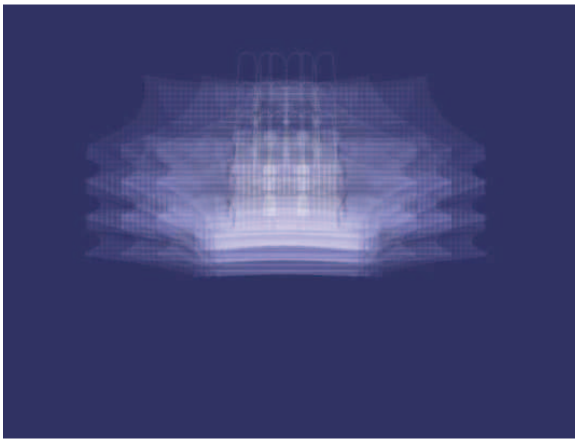

Fig. 4-1 Original integral image

Figure 4-3 shows the comparison results between various algorithms. Three algorithms were tested and compared. The first one is the traditional method that encodes the sequence of integral images by MPEG-4/AVC. The second approach is also based on MPEG-4/AVC but specifically implemented for sub-images. In this method, we arrange the transformed sub-images as a sequence of moving pictures and then use MPEG-4/AVC to encode it. The last compression method is our proposed MVC algorithm. In this approach, the transformed sub-images are arranged as multi-view video and encoded by standard MVC.

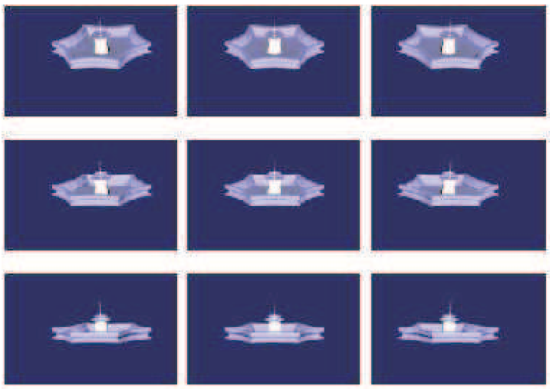

Fig. 4-2 Generated Sub-images

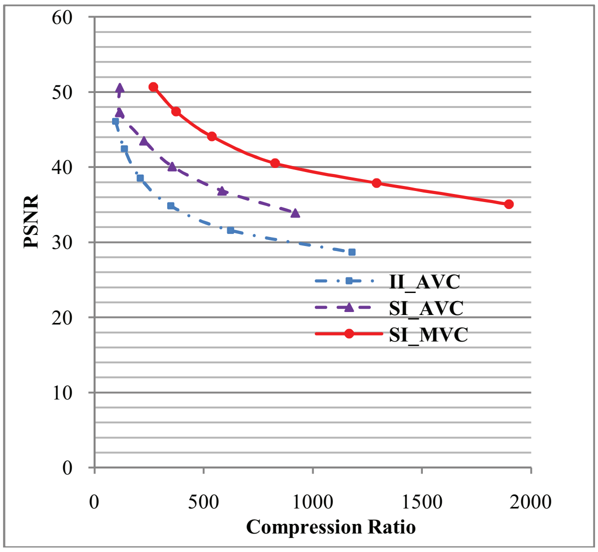

Fig. 4-3 Performance of MVC compression scheme for SIs, AVC scheme for SIs and baseline AVC scheme for integral images sequence

Finally, we evaluate the compression efficiency of the algorithms by the PSNR curves in function of compression ratio. According to Figure 4-3, it can be noticed that the MEG-4 based algorithm for sub-images outperforms the regular MPEG-4 approach for original integral images because the transformed SIs contains more redundancy than initial integral images. We can also find that our proposed MVC algorithm performs better than the other two schemes. This demonstrates the advantage offered by the correlation removing process among inter-view sub-images.

5. CONCLUSIONS

In this paper, a new video compression scheme for integral imaging system was proposed and demonstrated. In this approach, a recently raised image format known as sub-image was studied and employed. According to previous research, sub-image contains several advantages for compressing than original integral image. In order to adopt these useful features for video compression, we proposed to arrange the sub-images as multi-view video and encode the generated MVV by MVC standard. The experiment results show that our proposed algorithm is more efficient than other two compression approaches base on MEG-4/AVC.

6. REFERENCES

[1] Youngmin Kim, Keehoon Hong, and Boungho Lee, “Recent Researches based on Integral Imaging Display Method”, 3D Research Center and Springer 2010, pp. 17-27, 3D Research (2010) 01.

[2] E.Elharar, Adrian Stern, Ofer Hardar, and Bahram Javidi, “A Hybrid Compression Method for Integral Images Using Discrete Wavelet Transform and Discrete Cosin Transform”, Journal of Display Techonlogy, VOL.3, NO.3, September 207.

[3] Dong-Choon Hwangm Jae-Sung Parkm Dong-Hak Shinm Eun-Soo Kim, “Depth-controlled Reconstruction of 3D Integral Image Using Synthesized Intermediate Sub-images”, Optical Communications 281 (2008) 5991-5997, 10 September 2008.

[4] Manuel Maritinez-corral, Raul Martinez-Cuenca; Genaro Savedra; Héctor Navarro, Amparo Pons, and Bahram Javidi, “Progresses in 3D Integral Imaging with Optical Process”, Journal of Physics: Conference Series 139 (2008).

[5] Fumio Okano, Jun Arai, Kohji Mitanni; and Makoto Okui, “Real-Time Integral Imaging Based on Extremely High Resolution Video System”, Proceeding of the IEEE, VOL.94, NO.3, March 2006.

[6] Ho-Hyun Kang, Dong-Hak Shin, Eun-Soo Kim, “Efficient Compression of Motion-compensated Sub-images with Karhunen-Loeve Transform in Three-dimensional Integral Imaging”, Optics Communications 283 (2010) 920-928.

[7] Jae-Hyeung Park, Joohwan Kim, and Byoungho Lee, “Three-dimensional Optical Correlation Using a Sub-image Array”, Optical Express 5116, 27 June 2005.

[8] Sekwon Yeom, Adrian Stern, and Bahram Javidi, “Compression of 3D Color Integral Images”, Optical Express 1632, Vol.12, No.8, 19 Apprial 2004.

[9] R. Zaharia, A. Aggoun and M. McCormick, “Adaptive 3D-DCT Compression Algorithm for Continuous Parralax 3D Integral Imaging”, Signal Processing: Image Communication 17 (2002) 231-242.