

Compression of 3D color integral images

Sekwon Yeom, Adrian Stern,* and Bahram Javidi

Electrical and Computer Engineering Department, University of Connecticut
371 Fairfield Road, Unit 1157, Storrs, CT 06269-1157

*Electro Optical unit, Ben-Gurion University of the Negev,
B.O. Box 653, 84105 Beer-Sheva, Israel
bahram@engr.uconn.edu

http://www.engr.uconn.edu/ece/ece_facu_javidi.htm

Abstract: In this paper, we discuss the compression results of full color 3D Integral Images (II) by MPEG-2 (Motion Picture Experts Group). II is a popular three-dimensional image video recording and display technique. The huge size of II data has become a practical issue for storing and transmitting of 3D scenes. The MPEG is a standard coded representation of moving pictures. We model the elemental images in II as consecutive frames in a moving picture. Therefore, MPEG scheme can be applied to take advantage of the high cross-correlations between elemental images. We also introduce several scanning topologies along the elemental image sequences and investigate their performance with different number of pictures in GOP (Group of Picture). Experimental results are presented to illustrate the image quality of the MPEG-2 and the baseline JPEG with the same compression rate. We show that a well-known and widely-available MPEG-2 scheme can be a good alternative for II compression.

©2004 Optical Society of America

OCIS codes: (999.9999) Image compression, (100.2000) Digital image processing, (100.6890) Three-dimensional image processing, (110.3000) Image quality assessment, (110.6880) Three-dimensional image acquisition.

References and links

1. F. Okano, H. Hoshino, J. Arai, and I. Yuyama, "Real-time pickup method for a three-dimensional image based on integral photography," *Appl. Opt.* **36**, 1598-1603 (1997).
2. F. Okano, H. Hoshino, J. Arai, and I. Yuyama, "Three-dimensional video system based on integral photography," *Opt. Eng.* **38**, 1072-1077 (1997).
3. S. A. Benton, ed., *Selected Papers on Three-Dimensional Displays* (SPIE Optical Engineering Press, Bellingham, WA., 2001).
4. D. H. McMahon and H. J. Caulfield, "A technique for producing wide-angle holographic displays," *Appl. Opt.* **9**, 91-96 (1970).
5. T. Okoshi, "Three-dimensional displays," in *Proceedings of IEEE* **68**, 548-564 (1980).
6. G. Lippmann, "La photographie intergrale," *C. R. Acad. Sci.* **146**, 446-451 (1908).
7. C. B. Burckhardt, "Optimum parameters and resolution limitation of integral photography," *J. Opt. Soc. Am.* **58**, 71-76 (1968).
8. M. Martinez-Corral, C. Ibáñez-López, and G. Saavedra, "Axial gain resolution in optical sectioning fluorescence microscopy by shaded-ring filters," *Opt. Express* **11**, 1740-1745 (2003).
9. P. Ambs, L. Bigue, Y. Fainman, R. Binet, J. Colineau, J.-C. Leheureau, and J.-P. Huignard, "Image reconstruction using electrooptic holography," in *Proceedings of IEEE Conference on the 16th Annual Meeting of the IEEE Lasers and Electro-Optics Society*, **1** (IEEE, Piscataway, NJ., 2003), pp. 179-180.
10. H. Arimoto and B. Javidi, "Integral three-dimensional imaging with digital reconstruction," *Opt. Lett.* **26**, 157-159 (2001).
11. J. Jang and B. Javidi, "Three-dimensional synthetic aperture integral imaging," *Opt. Lett.* **27**, 1144-1146 (2002).
12. A. Stern and B. Javidi, "3-D computational synthetic aperture integral imaging (COMPSAI)," *Opt. Express* **11**, 2446-2451 (2003), <http://www.opticsexpress.org/abstract.cfm?URI=OPEX-11-19-2446>.

13. Y. Frauel and B. Javidi, "Digital three-dimensional image correlation by use of computer-reconstructed integral imaging," *Appl. Opt.* **41**, 5488-5496 (2002).
 14. M. C. Forman and A. Aggoun, "Quantisation strategies for 3D-DCT-BASED compression of full parallax 3D images," in *Proceedings of Int. Conf. on Image processing and its applications, 6th*, (Ireland, 1997), pp. 32-35.
 15. R. Zaharia, A. Aggoun, and M. McCormick, "Adaptive 3D-DCT compression algorithm for continuous parallax 3D integral imaging," *Signal Processing: Image Communication* **17**, 231-242 (2002).
 16. J. S. Jang and B. Javidi, "Compression of ray information in three-dimensional integral imaging using the Karhunen-Loeve transform," submitted to *Opt. Lett.* (2004).
 17. V. Bhaskaran and K. Konstantinides, *Image and video compression standards 2nd edition*, (Kluwer Academic Publishers, 1997).
 18. M. Rabbani, *Fundamentals of Wavelet Image compression and the emerging JPEG-2000 standard VT080*, (SPIE Press, Bellingham, WA., 2000).
 19. R. L. Joshi, M. Rabbani, and M. A. Lepley, "Comparison of multiple compression cycle performance for JPEG and JPEG 2000," in *Applications of Digital Image Processing XXIII*, A. G. Tescher, ed., *Proc. SPIE* **4115**, 492-501 (2000).
 20. J. A. Saghri, A. G. Tescher, and A. M. Planinac, "KLT/JPEG 2000 multispectral bandwidth compression with region-of-interest prioritization capability," in *Applications of Digital Image Processing XXVI*, A. G. Tescher, ed., *Proc. SPIE* **5203**, 226-235, (Nov 2003).
 21. T. J. Naughton, Y. Frauel, B. Javidi, and E. Tajahuerce, "Compression of digital holograms for three-dimensional object reconstruction and recognition," *Appl. Opt.* **41**, 4124-4132 (2002).
 22. A. Mahalanobis and C. Daniell, "Data compression and correlation filtering," in *Smart Imaging Systems* (SPIE Press, 2001).
 23. M. Rabbani, *Selected Papers on Image Coding and Compression*, SPIE Milestone Series MS48 (SPIE Press, 1992).
 24. T. A. Welch, "A technique for high performance data compression," *IEEE Computer* **17**, 8-19 (1984).
 25. T. Nomura, A. Okazaki, M. Kameda, Y. Morimoto, and B. Javidi, "Digital holographic data reconstruction with data compression," in *Algorithms and Systems for Optical Information Processing V*, B. Javidi and D. Psaltis, eds., *Proc. SPIE* **4471**, (2001).
 26. MPEG-2 Video Codec (with Source Code), <http://www.mpeg.org/MPEG/MSSG/#source>
 27. MPEG-2 Test Model 5, <http://www.mpeg.org/MPEG/MSSG/tm5>
-

1. Introduction

There has been great interest in three-dimensional (3D) imaging and visualization [1-16]. Integral photography (IP) or integral imaging (II) is one of the technologies considered for 3D image recording and display [1-7]. II uses a pinhole array or a lens array as shown in Fig. 1. It records 3D information of the directions and intensities of optical rays from objects. The 3D scene can be reconstructed by generating a pseudoscopic real image by propagating the elemental images through the lens array. The pseudoscopic real image is formed by the rays of opposite directions and the same intensities from the display device. II provides autostereoscopic images without coherent illumination. It has continuous horizontal and vertical parallax, continuously varying viewpoints, and real perspectives. Recently, computational reconstruction method of II has been developed with various techniques [10,12,13].

II reconstruction method with improved field of view and viewing angle were developed using Synthetic Aperture Integral Image (SAII) optically or digitally [11,12]. In [13], 3D image correlation technique was applied to II for recognition of 3D objects and their locations.

The size of II data can be huge, especially with full color components. Therefore, it has become a critical issue to handle such a large data for practical purposes such as storing on a media device or transmitting in real time. In [14], a compression method based on 3D Discrete Cosine Transform (DCT) was specially developed for unidirectional II. Adaptive quantization strategy was developed based on 3D-DCT in [15]. In [16], Karhunen-Loeve (KL) transform was used to compress ray information for II display.

In this paper, we present II compression by MPEG (Motion Picture Experts Group). The II is converted to a video stream which is compressed by MPEG-2. The MPEG-2 efficiently compresses each elemental image and exploits the correlation between them. Comparison is made with II compression using the JPEG (Joint Picture Experts Group). We also introduce

three types of scanning topologies for elemental images and investigate their compression performance.

The system generates many elemental images. These elemental images can be considered as consecutive frames in a moving picture for the purpose of compression. The high similarity between the elemental images can be used for compression using MPEG. MPEG deals with only the same size of elemental images. In contrast, the large number of pixels may deteriorate the performance of the JPEG. Another advantage of the MPEG is its well-established and fast-developing standard.

The paper is organized as follows. We briefly review the video part of the MPEG-2 encoder in Section 2. In Section 3, we discuss three types of compression topologies and evaluation parameters. Experimental results are presented in Section 4, and conclusions follow in Section 5.

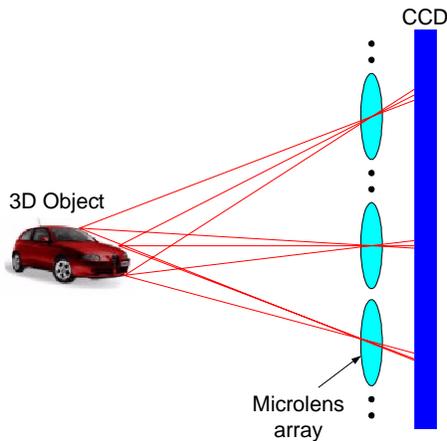


Fig. 1. Optical system for generating integral images of a 3D object

2. Background on MPEG-2

In this section, we briefly review the video part of MPEG-2 encoder. MPEG-2 is the ISO standard 13818 (or ITU-T Recommendation H.262), *Generic coding of moving pictures and associated audio* [17]. It is a popular coding technique for moving pictures and associated audio information on digital storage media. JPEG corresponds to the ISO/IEC international standard 10918-1~3, *Digital compression and coding of continuous-tone still images* or to the ITU-T Recommendation T. 81~4. It has been developed to compress still images and became widely used [17-25].

The basic scheme of the video part of MPEG-2 encoder is shown in the block diagram of Fig. 2. During preprocessing, the encoder performs color conversion to YCbCr with sampling reduction of 4:2:0 or 4:2:2. *Y* is the luminance component and *Cb* and *Cr* are the chrominance components. MPEG has two strategies for video compression. One is “intraframe coding” and the other is “interframe coding.” The former one is similar to still image compression by JPEG. For the latter one, MPEG subtracts macroblocks established with the motion vector and performs DCT coding of their difference. For the interframe coding, motion estimation yields the optimal motion vector of macroblocks. During the motion compensation, the prediction errors of moving pictures are computed between the best-matching macroblocks in different frames. The macroblock is the smallest coded unit. It consists of four 8×8 blocks of *Y*, one 8×8 block of *Cb*, and one 8×8 block of *Cr*.

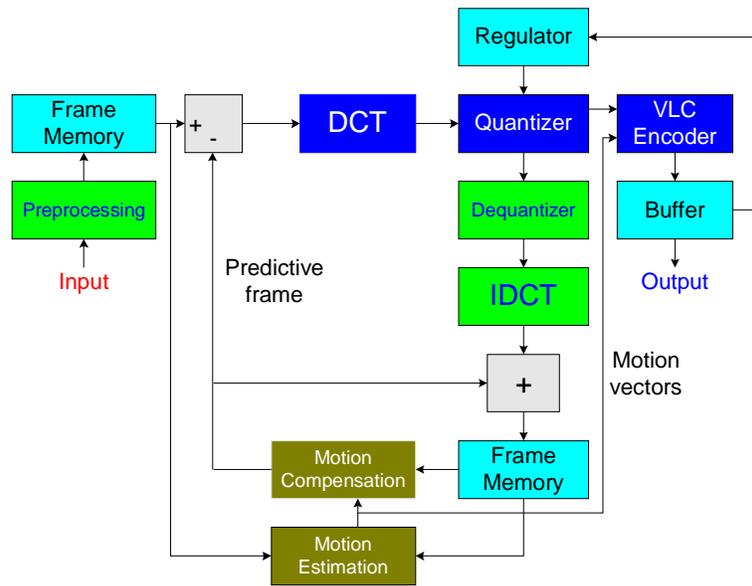


Fig. 2. Block diagram of MPEG encoder for video part.

GOP is the highest layer in MPEG. A GOP is composed of a series of consecutive pictures. A picture can be regarded as one of 3 types: intra (I), predicted (P), and bidirectionally predicted (B) pictures. I-pictures are coded without motion compensation, i.e. I-pictures are coded by the intraframe coding. P-pictures are coded using motion compensated prediction from past I or P-pictures. B-pictures are coded using motion compensated prediction from either past and/or future I or P-pictures. B-pictures provide the highest compression rate. A sequence of I, P, and B-pictures in a GOP is shown in Fig. 3. Figure 3 illustrates the GOP with $N = 6$ and $M = 3$. N is distance between I pictures and M is distance between consecutive I or P pictures.

After block-based DCT coding, a quantization matrix and a quantization scale factor decides the quantization coefficients. There are different quantization schemes for I, B, and P-pictures. They also depend on the type of macroblocks (inter or intra macroblocks), and values of the DCT coefficients [17]. The Variable Length Coding (VLC) is the final process for MPEG. MPEG adopts Huffman coding for lossless compression of motion vectors and quantized DCT coefficients for VLC.

MPEG-2 defines several profiles and levels for non-scalable mode and scalable mode. They have the upper bounds for picture resolution, frame rate, and bit rates [17].

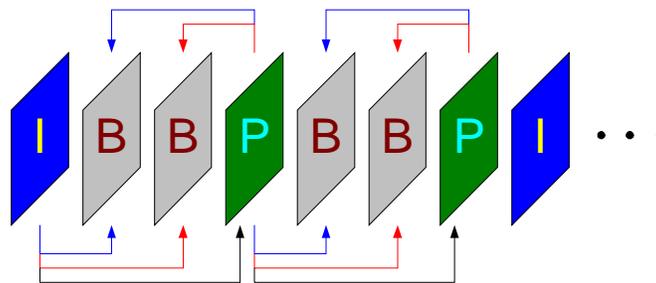


Fig. 3. I-, B-, P- pictures in GOP ($N = 6$, $M = 3$).

3. Compressing integral images using MPEG-2

The elemental images in Π are a consecutive image set of one common scene from the micro lens array. They contain different perspectives and depth information of 3D objects. Although JPEG is a representative coding skill for a still image, we apply a well-defined motion picture algorithm to compress the elemental images of Π . MPEG takes advantages of the correlation between the elemental images. We consider one elemental image as a frame in a moving picture for the compression by MPEG.

We consider 3 different scanning topologies for elemental images in Π . Figure 4 shows these three types of scanning topologies. Parallel scanning is a sequential scanning in parallel direction. It's suitable for Π with different size in horizontal and vertical directions. Perpendicular and spiral topologies are other scanning methods to minimize motion compensation between elemental images. The spiral scanning can be adopted if more focused images are located at the center of Π . It is noted that some elemental images can be blurred when captured if they are far from optical axis.

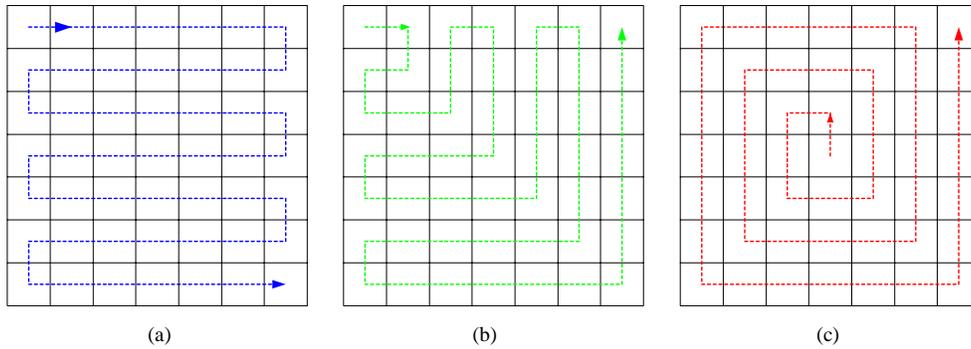


Fig. 4. Three different scanning topologies for converting an Π into a sequence of elemental images: (a) parallel scanning, (b) perpendicular scanning, (c) spiral scanning.

We evaluate the efficiency of compression by the following parameters. The PSNR (Peak-to-peak Signal to Noise Ratio) is defined as:

$$PSNR(I_o, I_u) = 10 \log_{10} \left(\frac{P^2}{MSE(I_o, I_u)} \right), \quad (1)$$

where P is the maximum value in one pixel; I_o is an original image; and I_u is the image obtained after decompression. MSE (Mean Square Error) is defined as:

$$MSE(I_o, I_u) = \frac{1}{M_s N_s} \sum_{n=1}^{N_s} \sum_{m=1}^{M_s} |I_o(n, m) - I_u(n, m)|^2, \quad (2)$$

where (M_s, N_s) are the number of pixels in the x and y axis of the image. Another metric for image quality is SNR (Signal to Noise Ratio) defined as:

$$SNR(I_o, I_u) = 10 \log_{10} \left(\frac{VAR(I_o)}{MSE(I_o, I_u)} \right), \quad (3)$$

where $VAR(I_o)$ is the variance of the original image I_o .

The compression rate is defined:

$$r = \frac{\text{Original image size}}{\text{Compressed image size}}. \quad (4)$$

4. Experimental results

In this section, we present the experimental results for II compression. Figure 1 illustrates optical system for capturing 3D scenes. It is composed of 3D objects, a micro lens array, an imaging lens, and a CCD camera. The focal length of each micro lens is about 3 mm, the effective focal length of the imaging lens is 50 mm, and the f -number of the imaging lens is 2.5. The imaging lens is inserted between the lens array and CCD camera due to the short focal length of the lenslets.

Three objects in Fig. 5 are used to obtain the elemental images from the 3D scenes. The size of rectangular traffic sign plate is $2.3 \text{ cm} \times 2.3 \text{ cm} \times 0.35 \text{ cm}$, the diameter of circular traffic sign plate is 2.3 cm, and the size of toy car is $2.5 \text{ cm} \times 2.5 \text{ cm} \times 4.5 \text{ cm}$. The distance between CCD camera and imaging lens is 7.2 cm, and the distance between micro lens array and imaging lens is 2.9 cm. Two different 3D scenes noted as II-(1) and II-(2) are used in the experiments. For II-(1), as shown in Fig. 5(b), the rectangular plate and the circular plate are apart from the micro lens array by 10.5 cm and 12.3 cm, respectively. For II-(2), as shown in Fig. 5(c), the toy car and two traffic sign plates are apart from the micro lens array by 12 cm, 13 cm, and 14 cm, respectively. The detailed specifications for 3D scenes are presented in Table 1. Figure 6(a) and 6(c) show original elemental images of two different 3D scenes of II-(1) and II-(2), respectively. The original II requires 24 bits per pixel for storing. Eight bits are assigned to each color component.

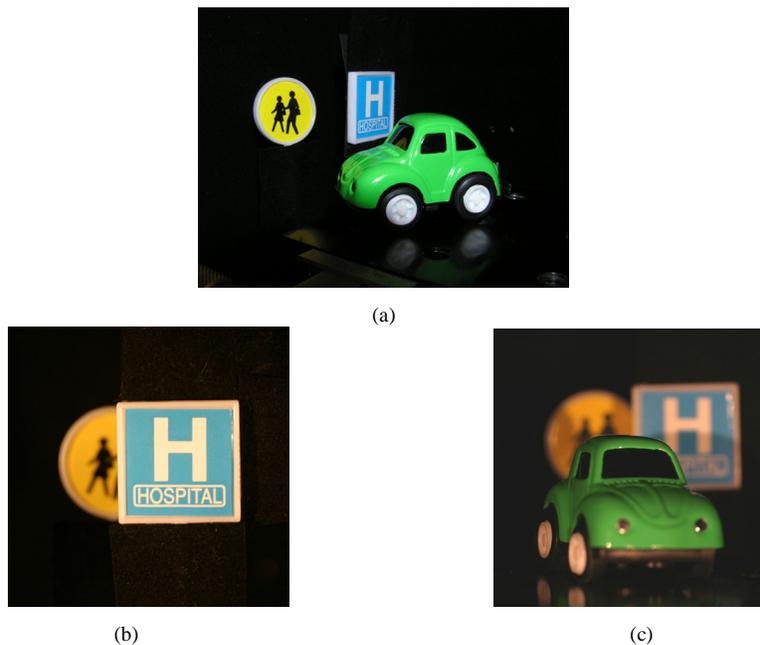


Fig. 5. 3D objects used in II compression experiments: (a) two traffic sign plates and a toy car, (b) 3D objects arranged in II-(1) scene, (c) 3D objects arranged in II-(2) scene.

Table 1. Specifications of II images used in the compression experiments

	II-(1): scene in Fig. 5(b)	II-(2): scene in Fig. 5(c)
Original file size (Kbytes)	6260	6267
Image size (pixels)	1,456 × 1,456	1,456 × 1,456
Number of elemental images	7 × 7 = 49	7 × 7 = 49

MPEG-2 Test Model (TM) 5 video codec from MPEG Software Simulation Group (MSSG) is adopted for MPEG compression [26]. JPEG compression is achieved by a built-in function in commercial software [MATLAB R12.1].

MPEG supports several parameters to be adjusted. During experiments most parameters were set at “default.” The profile and level ID were set at both “Main” in non-scalable mode. 4:2:0 sub sampling format is applied during preprocessing in MPEG-2.

Search width and height for motion estimation are limited by f-code which specifies the maximum length of the motion vector. Search width and height were chosen heuristically when better results were produced. The bit rate was set at 4×10^5 bits/s, and the frame rate was set at 30 frames/s for II-(1) in Fig. 5(b) and II-(2) in Fig. 5(c). The quality factor of JPEG was set at 20 for II-(1) and 23 for II-(2) to provide the same compression rates between MPEG-2 and JPEG. The sizes of compressed files are 80 Kbytes for both methods. Figure 6(b) and 6(d) show decompressed elemental images of II-(1) and II-(2) after MPEG-2 compression. Table 2 shows detailed experimental results for II-(1) and II-(2). The PSNR and SNR of MPEG-2 are larger than those of JPEG. It is noted that MPEG-2 depends on high similarities between elemental images.

Table 2. Experimental results of II compression

	II-(1): scene in Fig. 5(b)	II-(2): scene in Fig. 5(c)
PSNR / SNR for JPEG (dB)	33.09 / 22.49	33.38 / 20.23
Maximum PSNR/SNR for MPEG-2 (dB)	34.78 / 24.18	34.26 / 21.11
Compressed file size (Kbytes)	80	80
Compression rate	78.25:1	78.34:1
Quality factor in JPEG	20	23
Bit rate (bits/s)	4×10^5	4×10^5
Frame rate in MPEG-2 (frames/s)	30	30
Search width and height for motion estimation in MPEG-2	1~3	9~15

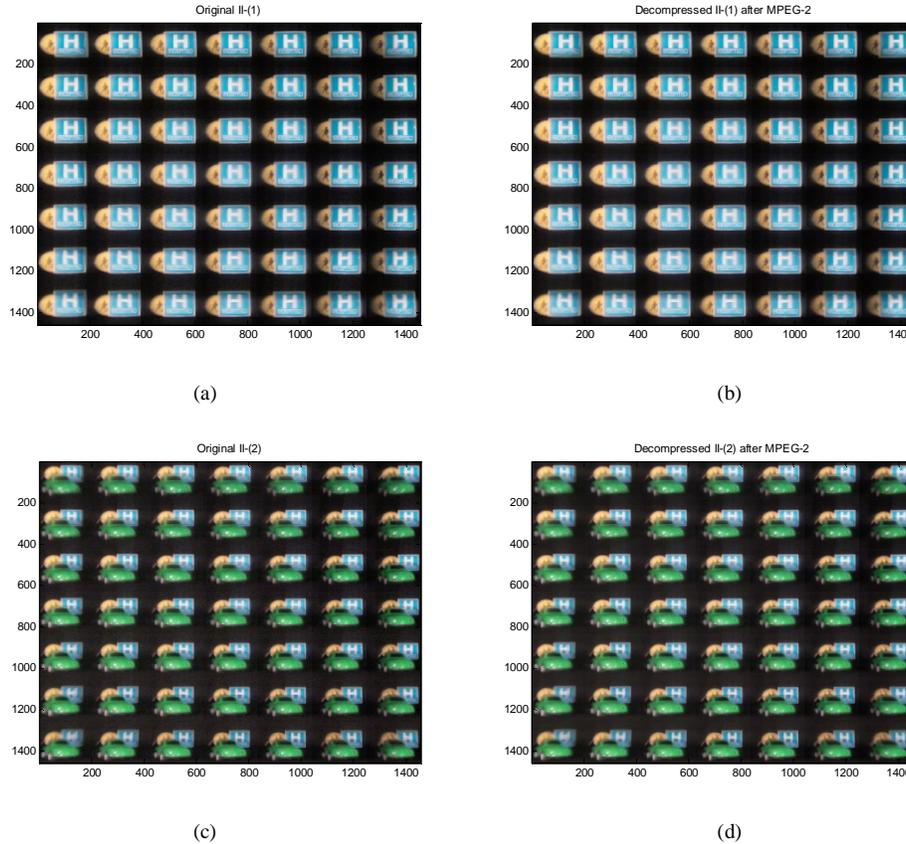


Fig. 6. Experimental results of 3D scene compression: (a) original II-(1) scene, (b) decompressed II-(1) scene after MPEG-2 compression ($r = 78.25$, spiral scanning, $N = 6$, $M = 3$), (c) original II-(2) scene, (d) decompressed II-(2) scene after MPEG-2 compression ($r = 78.34$, spiral scanning, $N = 6$, $M = 3$).

Figure 7 and 8 show PSNR and SNR of decompressed images with different $N = 3, 6, 9, 12, 15$, and $M = 3$. As shown in the Figures, PSNR and SNR are mostly higher when $N = 6$ and $M = 3$ for three different scanning topologies shown in Fig. 4. The spiral scanning topology provides better result for II-(1) scene and the perpendicular and spiral scanning topologies provide better results for II-(2) scene. This is because elemental images for interframe coding are closer to each other when the perpendicular and spiral scanning topologies are used. Closer elemental images result in smaller and more exact motion compensation. This may be also due to the effect of having blurred elemental images which are located near the edge of 3D scenes.

Figure 9 and 10 show PSNR and SNR of decompressed images with different compression rates. Spiral scanning method are used with $N = 6$ and $M = 3$. Several compression rates of II are achieved by changing the bit rate with the fixed frame rate in MPEG-2 and changing the quality factor in JPEG. MPEG-2 TM codec supports adaptive quantization which can be controlled by bit rate or frame rate [27].

As shown in Figures, the quality of decompressed images is better for MPEG-2 with higher compression rates. Bit rate for MPEG-2 corresponding to each compression rate is 6×10^5 , 5×10^5 , 4×10^5 , 3×10^5 , and 2×10^5 bits/s. The quality factor for JPEG is 41, 31, 20, 10, 5 for II-(1) and 44, 34, 22, 12, 9 for II-(2). Figures 11 and 12 show movies of elemental images of II-(1) and II-(2) for original and decompressed elemental images after MPEG-2,

respectively. It is noted that the compression rate is far less than theoretical value ($30 \times 208^2 \times 24 / 2 \times 10^5 \approx 155.75$) when bit rate is 2×10^5 bits/s.

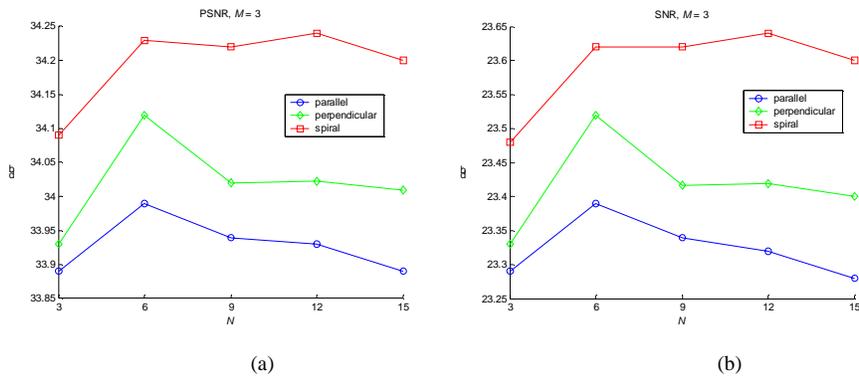


Fig. 7. MPEG-2 compression results for 3D scene in II-1 in Fig. 5(b): PSNR and SNR of three scanning topologies in Fig. 4 with varying N and fixed M : (a) PSNR, (b) SNR.

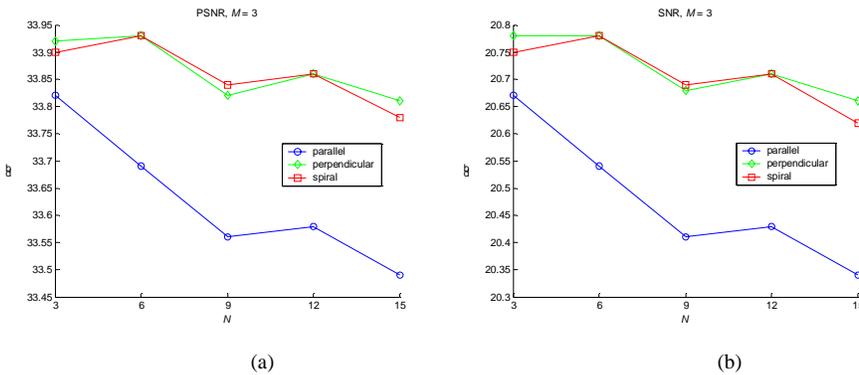


Fig. 8. MPEG-2 compression results for 3D scene in II-2 in Fig. 5(c): PSNR and SNR of three scanning topologies in Fig. 4 with varying N and fixed M : (a) PSNR, (b) SNR.

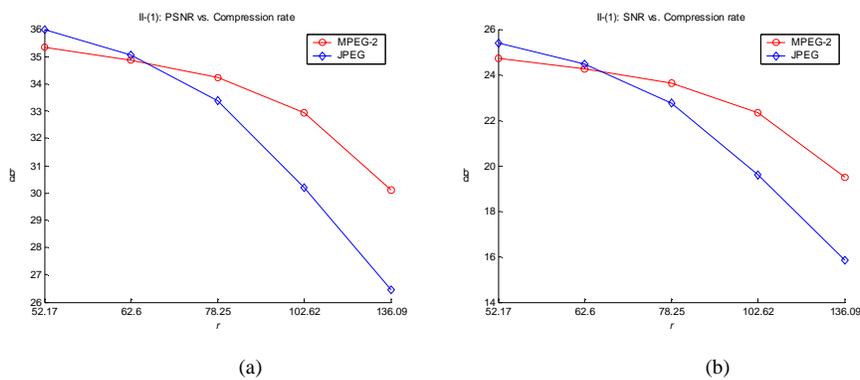


Fig. 9. Results of compression experiments for 3D scene II-1 in Fig. 5(b): image quality vs. Compression rates (spiral scanning, $N = 6$, $M = 3$): (a) PSNR, (b) SNR.

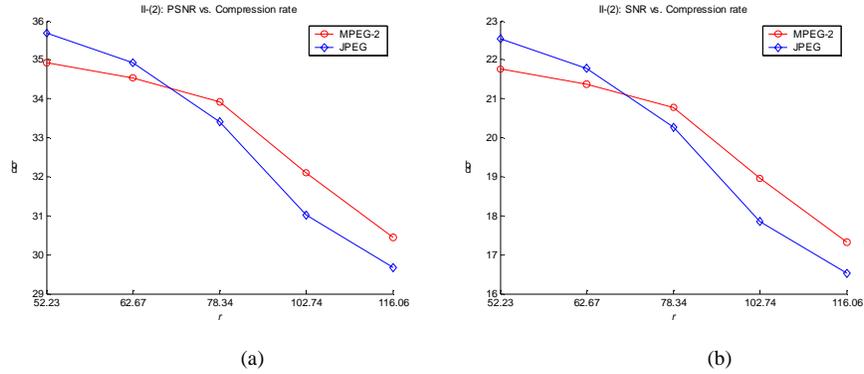


Fig. 10. Results of compression experiments for 3D scene II-(2) in Fig. 5(c), image quality vs. Compression rates (spiral scanning, $N = 6$, $M = 3$): (a) PSNR, (b) SNR.



Fig. 11. II-(2) Movie for original and decompressed elemental images after MPEG-2 (spiral scanning, $N = 6$, $M = 3$): (a) original (6.2 MB version, smaller version: 2.23MB), (b) $r = 52.17$ (6.2 MB version, smaller version: 2.23MB), (c) $r = 62.6$ (6.2 MB, smaller version: 2.23MB), (d) $r = 78.25$ (6.2 MB version, smaller version: 2.23MB), (e) $r = 102.62$ (6.2MB version, smaller version: 2.23MB), (f) $r = 136.09$ (6.2 MB version, smaller version: 2.23MB), smaller version movies contain only 18 frames.

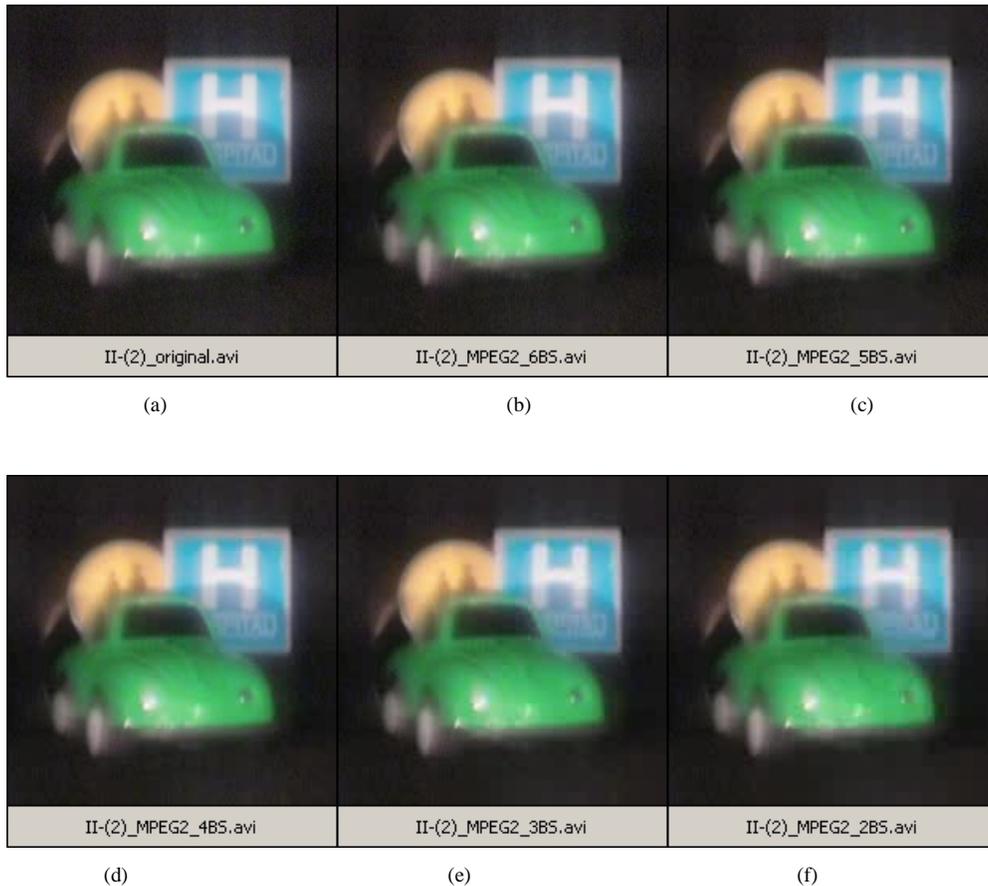


Fig. 12. II-(2) Movie for original and decompressed elemental images after MPEG-2 (spiral scanning, $N = 6$, $M = 3$): (a) original (6.2 MB version, smaller version: 2.23MB), (b) $r = 52.23$ (6.2 MB version, smaller version: 2.23MB), (c) $r = 62.67$ (6.2 MB version, smaller version: 2.23MB), (d) $r = 78.34$ (6.2 MB version, smaller version: 2.23MB), (e) $r = 102.74$ (6.2 MB version, smaller version: 2.23MB), (f) $r = 116.06$ (6.2 MB version, smaller version: 2.23MB), smaller version movies contain only 18 frames.

5. Conclusions

We have presented experiments on compression of 3D color integral imaging. MPEG-2 is adopted as our lossy compression scheme for II. We have investigated three scanning topologies of elemental images and their performance is evaluated. The image quality is compared between MPEG-2 and JPEG at the same compression rate. MPEG has shown to be a good compression alternative for II images. With fast developing MPEG standard, higher version of MPEG might be adopted for better performance.