Effect of different traversal schemes in integral image coding

Nicholas Sgouros,* Ioannis Kontaxakis, and Manolis Sangriotis

Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Panepistimiopolis, Ilissia, P.C. 15784, Greece

*Corresponding author: nsg@di.uoa.gr

Received 4 October 2007; revised 26 December 2007; accepted 14 January 2008; posted 16 January 2008 (Doc. ID 88133); published 20 February 2008

Integral imaging (InIm) is a highly promising technique for the delivery of three-dimensional (3D) image content. During capturing, different views of an object are recorded as an array of elemental images (EIs), which form the integral image. High-resolution InIm requires sensors with increased resolution and produces huge amounts of highly correlated data. In an efficient encoding scheme for InIm compression both inter-EI and intra-EI correlations have to be properly exploited. We present an EI traversal scheme that maximizes the performance of InIm encoders by properly rearranging EIs to increase the intra-EI correlation of jointly coded EIs. This technique can be used to augment performance of both InIm specific and properly adapted general use encoder setups, used in InIm compression. An objective quality metric is also introduced for evaluating the effects of different traversal schemes on the encoder performance.

OCIS codes: 100.2000, 100.6890, 110.3000, 110.6880.

1. Introduction

A growing number of 3D visualization systems are used in specialized applications where 3D observation of objects and scenes is required. Moreover technological achievements in the areas of integrated optics with sensors and network infrastructures guarantee that these systems will soon be used in a large number of different everyday applications. In general, 3D viewing methods can be divided into stereoscopic and autostereoscopic. Stereoscopic methods require additional viewing aids such as glasses or helmets. In contrast, autostereoscopic methods do not require any additional eyewear or head-mounted equipment, but all optical components needed to correctly reproduce 3D objects and scenes are integrated into the display device [1–3].

InIm is a technique for producing highly realistic 3D color images, which was initially devised by Lippman [4] back in 1908. It is currently regarded as one of the most promising methods for autostereoscopic representation of 3D scenes [3]. Specifically, this technique provides natural viewing with full color support, enhanced detail, and adequate depth level. Moreover, multiple simultaneous viewers are supported, while most of the currently existent InIm setups for capturing and display provide omnidirectional parallax images [5,6]. A basic InIm capturing setup is assembled by using a charge coupled device (CCD) sensor and a lens array (LA), while the reproduction setup uses a liquid crystal display (LCD) and an appropriate LA. These setups are depicted in Figs. 1(a) and 1(b), respectively.

The InIm created during capturing is a 2D array of EIs that are formed on the CCD surface behind each lens of the LA of Fig. 1(a). The EIs form a 2D image grid, which is then projected through an appropriate LA to reproduce the originally acquired 3D scene to the viewer. A 3D display setup based on the InIm principle is shown in Fig. 1(b).

A. Related Work

One of the main issues taken into account in every InIm encoding scheme is the exploitation of the different types of redundancy present due to inherent properties of an InIm. High inter-EI image correla-

0003-6935/08/1900D28-10$15.00/0

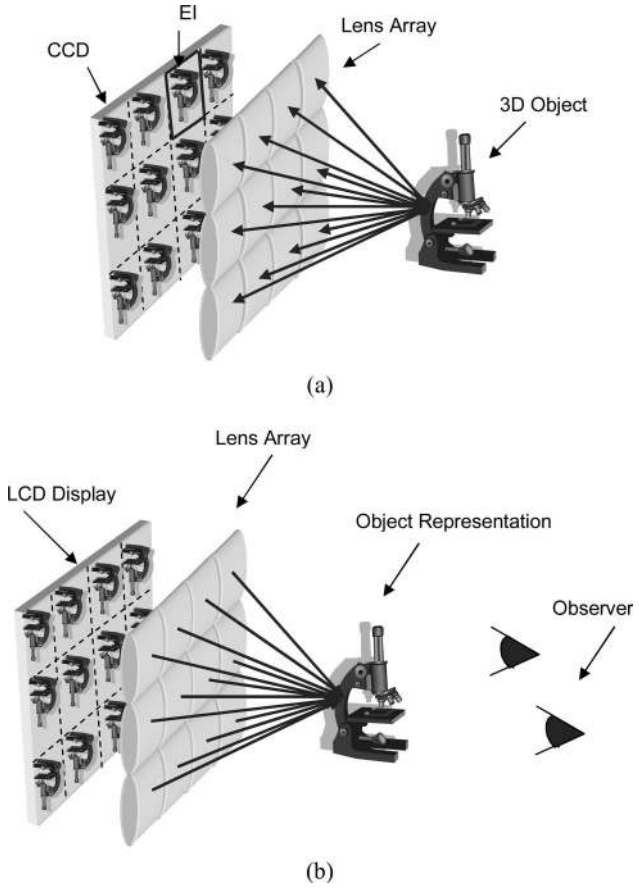


Fig. 1. Operating principle of (a) an InIm capturing setup, (b) an InIm display setup.

tion between adjacent pixels and high intra-EI correlation between adjacent EIs are combined and greatly increase the redundancy in an InIm.

The correlation coefficient values [7] for a square window of the central EI over the 3×3 EI neighborhood of Fig. 2(b) is depicted in Fig. 2(c) and shows the increased degree of redundancy between neighboring EIs.

Many InIm encoders are presented in the literature, which range from InIm-specific ones [8,9] to properly adapted generic 2D image and video coders [10,11]. Every one of these coding schemes includes a segmentation stage in which the 2D InIm structure is decomposed into a number of distinct EIs. These EIs are rearranged into a 1D image stream and encoded. After the 2D grid of EIs is rearranged in a 1D stream, only a small portion of highly correlated neighboring EIs in the 2D grid are finally placed in neighboring positions. The traversal scheme used in the reordering process influences this portion and to a great extent determines the rate–distortion performance of the encoder. In particular, the encoders proposed in [8,11] for omnidirectional parallax InIms introduce different InIm traversal schemes in order to retain high correlation between adjacent EIs in the constructed 1D stream.

B. Overview

In this work we propose an image traversal scheme based on the Hilbert curve, in order to increase the efficiency of encoders where EI rearrangement is performed. This is achieved by maximizing a 2D locality property of the EIs that are encoded as a single entity and therefore increasing the correlation between jointly encoded EIs. As the encoded EIs are reassembled in a static InIm at the display and projected to the viewer, large variations in the quality of different areas in the InIm can reduce the 3D effect. For this reason a quality-assessment metric is introduced for evaluating the optimality of a number of different traversal schemes previously proposed in the literature [8,11], taking into account the variability in the quality of different regions of an InIm. A number of different test InIms are used for evaluating the effect of the different traversal schemes for the encoders proposed in [8,11].

This paper is organized as follows: in Section 2 we introduce previously proposed and novel traversal schemes and the metric that is used for evaluating their locality preservation property. The results of the metric are presented in the same section for the examined traversal schemes. In Section 3 the performance gain for different traversal strategies is calculated for the 3D-DCT [8] and MPEG-2 [11] encoders, while Section 4 concludes this paper.

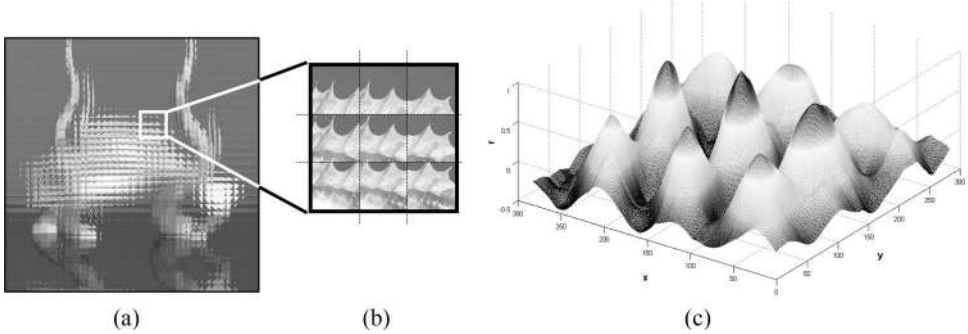


Fig. 2. (a) InIm of a fish, (b) 3×3 EI portion of the same InIm, and (c) correlation coefficient of a square part of the central EI in (b) over the area of 3×3 EIs.

2. Image Traversal Schemes

There are many types of curve [12] that can be used in order to traverse a set of points placed on a regular 2D grid and transform them to a 1D sequence of points. One of the most discussed properties of these curves is their ability to retain the locality properties of the reordered points. For this reason a number of different measures are used [13] to evaluate their efficiency. In omnidirectional InIm coding these curves are used in order to rearrange the EIs in a 1D stream of images preserving the highest possible correlation among jointly encoded EIs.

In this work we evaluated the performance of previously proposed traversal schemes [11] along with the one based on the Hilbert curve. The different traversal schemes are depicted in Fig. 3.

A. Locality Preservation Properties

The locality preservation properties of each of these space filling curves can be evaluated by estimating a large number of different metrics, such as the ones proposed in [13]. The Hilbert curve outperforms all previously proposed curves, achieving performance very close to the boundary values of these metrics. In this paper we propose an appropriate metric for evaluating the efficiency of the different traversal schemes, which is defined as the ratio between the EI distances in the 2D grid and their distance in the 1D stream as given by Eq. (1):

$$r_{ij} = \frac{d(e_i,\ e_j)^2}{|i-j|}, \quad \forall i,\ j{:}i > j. \tag{1}$$

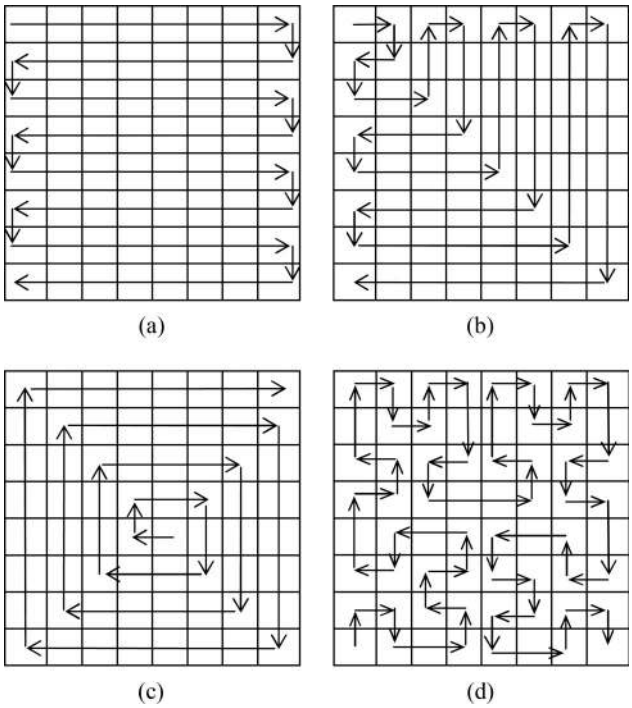


Fig. 3. Traversal schemes for rearranging the EIs of the InIm 2D structure into a 1D stream of EIs: (a) parallel, (b) perpendicular, (c) spiral, and (d) Hilbert.

In this equation i, j are the indices of the EIs in the constructed 1D stream and $d(\cdot)$ is their Euclidean distance in the 2D InIm structure. Specifically, if (w_i, c_i) and (w_j, c_j) are the row and column indices of the two EIs in the 2D InIm structure, then the distance is defined by Eq. (2):

$$d(e_i,\ e_j) \triangleq \sqrt{(w_i - w_j)^2 + (c_i - c_j)^2}. \tag{2}$$

Next, we calculate the mean value, $\bar{r}$, and the standard deviation, σ, of this quantity for each traversal scheme. A simulation performed for InIms with different numbers of EIs showed that the Hilbert curve achieves the minimum mean and standard deviation values for r_{ij} that are almost invariant regardless of the number of EIs contained in an InIm. In contrast, as shown in Fig. 4, the corresponding values for the other traversal schemes dramatically diverge as the number of EIs increases. In the same figure, a nonlinear increase of the standard deviation is observed for traversals other than the Hilbert curve, which increases the possibility for jointly encoding low correlated EIs and thus reducing the overall encoder performance.

Specifically, the performance of the different traversal schemes was evaluated over InIms having $2^p \times 2^p$ EIs, where $p = 2, \ldots, 5$. However, in order to traverse InIms with arbitrary sizes, generalizations of the previous traversal schemes have to be used. The proposed traversal schemes can be categorized as periodic or aperiodic. The aperiodic traversal schemes, such as spiral and perpendicular scans, can be subjectively expanded to arbitrary sized grids but do not preserve locality among neighboring points of the 2D grid. In contrast, the periodic parallel traversal can be straightforwardly extended for arbitrary sized grids, and the Hilbert traversal can be extended by using a number of generalizations proposed in the literature [14–16]. The Hilbert traversal outperforms the parallel traversal scheme even for arbitrary sized grids as shown in [14]. An extensive discussion of the properties of the Hilbert curve and its high clustering performance, along with different application fields, can be found in [17–19].

B. Encoder Efficiency

An accurate way to measure the degree of similarity and hence the redundancy in a group of jointly coded EIs in the 1D stream is the calculation of the correlation coefficient between the EIs of the group. The correlation coefficient constitutes a similarity metric between images and can be used for accurate motion estimation [20]. As already shown in Fig. 2, strong correlation exists between neighboring EIs in the 2D grid. However, when the EIs are rearranged in the 1D stream the correlation between consecutive EIs is affected by the choice of the traversal scheme. The 1D stream is further segmented in groups of EIs, and the EIs of each group are jointly encoded. High correlation among the EIs of the same group greatly affects the encoder perfor-

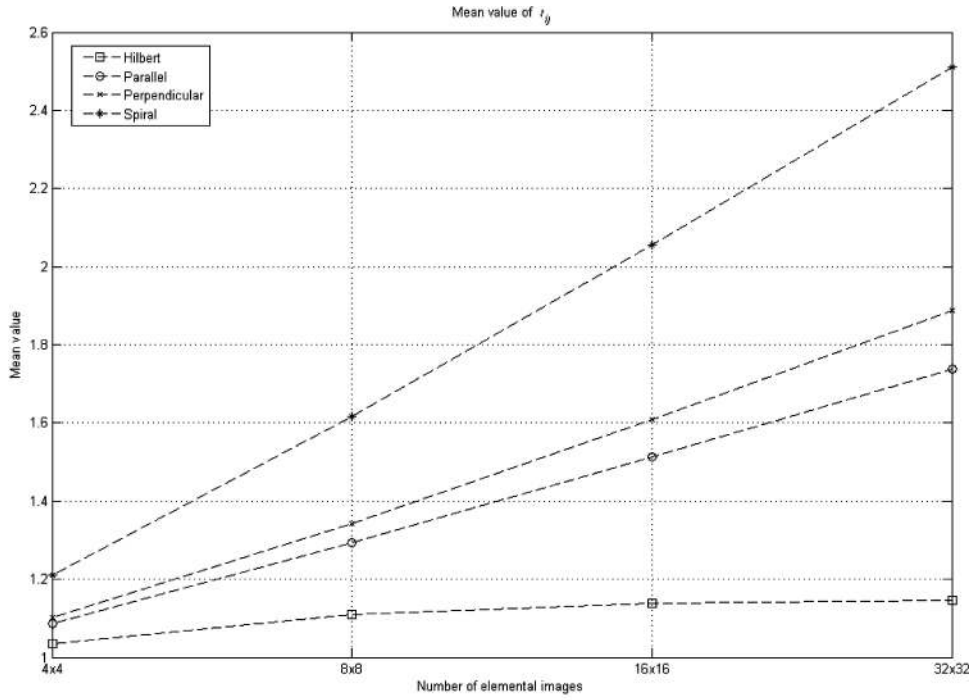


Fig. 4. (a) Mean value and (b) standard deviation of r_{ij} for different traversal schemes and InIm sizes.

mance as the information redundancy is greater among EIs of the group, and therefore they can be more efficiently encoded.

In Fig. 5 a comparison of the different traversal schemes for a characteristic InIm is presented. An InIm of 8×8 EIs is traversed by using the different traversals, and the rearranged EIs in the 1D stream are divided into groups of eight EIs. Figure 5(a) clearly shows that the Hilbert traversal scheme outperforms others in terms of correlation among EIs of the same group. This can be attributed to the high-locality-preservation property of the Hilbert curve. In fact the locality-preservation property of the Hilbert curve increases the probability of passing through neighboring EIs that contain parts of the same object in the scene and thus increases the

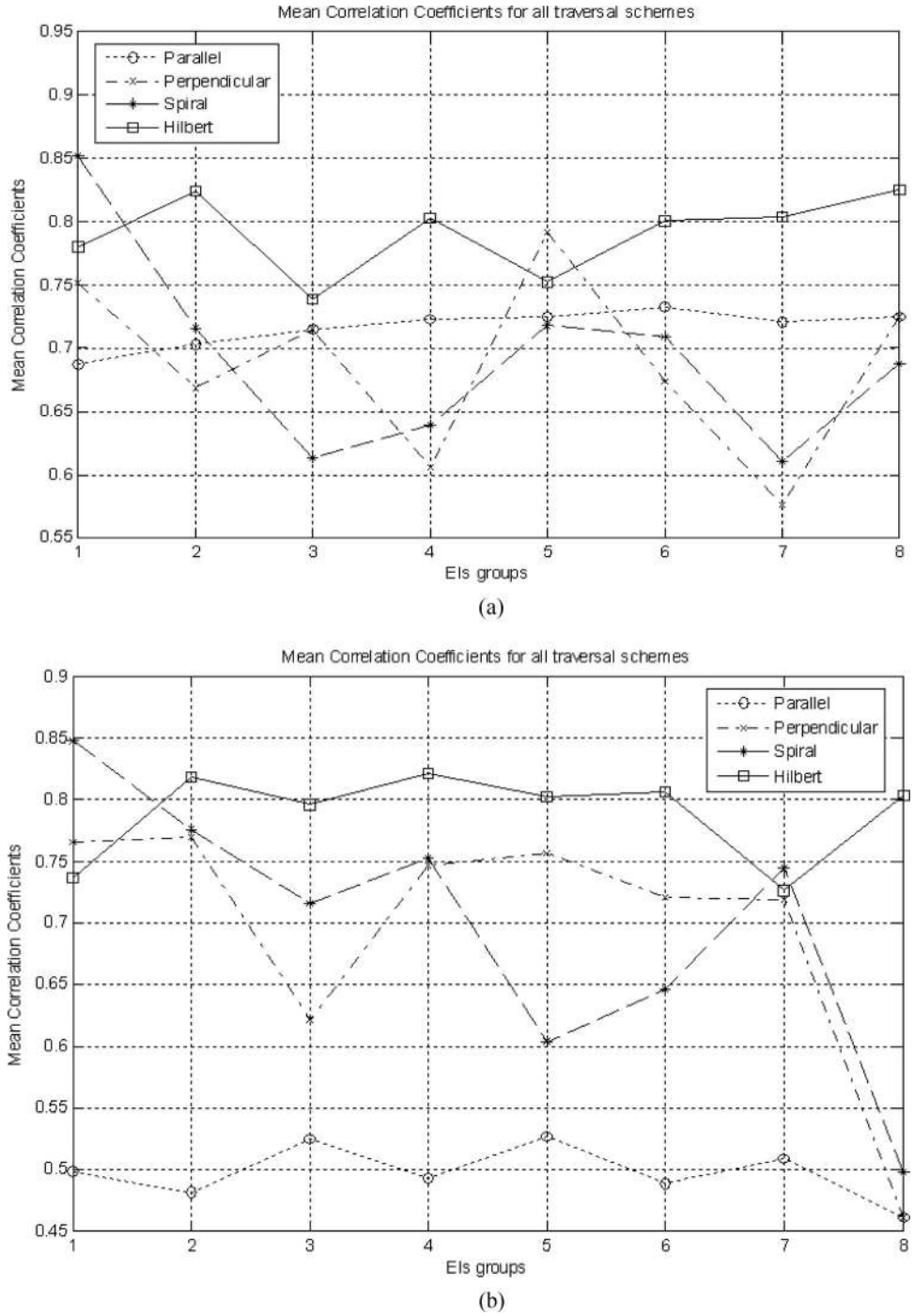


Fig. 5. Mean correlation coefficients for groups of eight EIs for all traversal schemes of (a) a characteristic InIm and (b) the same InIm rotated by 90°.

redundancy within the jointly encoded EIs of each group. The strong directionality of the other traversal schemes usually forces them to leave and reenter an object many times during the traversal, and as a result they jointly encode suboptimal EI groups in terms of similarity and hence information redundancy. However, the self-similarity of the Hilbert curve guarantees that each quadrant of the InIm will be exhaustively visited by the traversal, which increases the performance even in scenes with multiple objects.

A simple experiment was set up to demonstrate how the highly directional properties of the rest of the traversal schemes highly affect the encoders' performance. The InIm used for producing the results of Fig. 5(a) was rotated by 90°, and the correlation coefficients were recalculated for all groups of EIs. As is shown in Fig. 5(b), the Hilbert curve

retains the highest correlation among the EIs of each group. In contrast, the rest of the curves and especially the parallel scan are highly affected by the orientation of the objects of the scene.

3. Experimental Results

A number of different images are used to evaluate the effect of the different traversal schemes in both types of encoder. In the experiments a number of synthetic and physically acquired InIms were used. Four sample InIms of different complexities and arbitrary positions of objects and the original scenes used for computational generation or acquisition are shown in Fig. 6.

The synthetic InIms depicted in Figs. 6(a) and 6(b) were generated by using a ray-tracing technique. The LA that was used in the ray tracer was constructed by using constructive solid geometry (CSG) principles, and the original 3D scene, which is inset in Figs. 6(a) and 6(b), is a commonly available scene of the public domain software named POV-RAY [21]. The implementation details are contained in previous work of the authors [22,23]. The dimensions of all computationally generated images were 2048×2048 pixels. In Fig. 6(a) the InIm contains 8×8 EIs, while the one in Fig. 6(b) contains 256×256 EIs. For the real-world objects a simple capturing setup like the one described in [24] was used for physical acquisition. The dimensions of the physically acquired InIms in Figs. 6(c) and 6(d) are 1280×1280 pixels, and they contain 16×16 and 64×64 EIs, respectively. The original objects are also inset in Figs. 6(c) and 6(d).

A. Performance Evaluation

There are primarily two different approaches for evaluating the performance of image encoding techniques based on subjective or objective evaluation metrics [25,26]. The most used metric for objective 2D image quality assessment is the peak-signal-to-noise-ratio (PSNR), which is given by

$$\mathrm{PSNR}_{\mathrm{dB}} = 10 \times \log_{10}\left(\frac{255^2}{\mathrm{MSE}}\right), \tag{3}$$

where the mean squared error (MSE) is calculated by using Eq. (4):

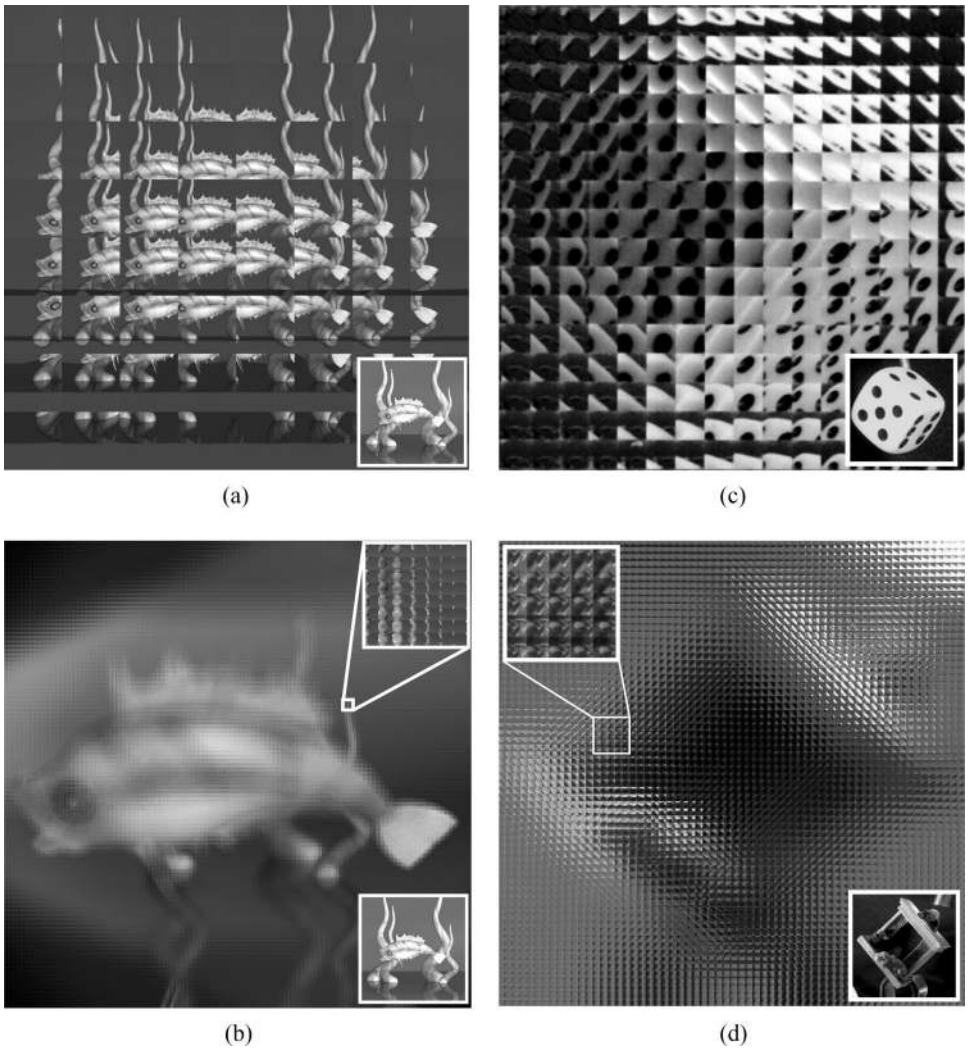
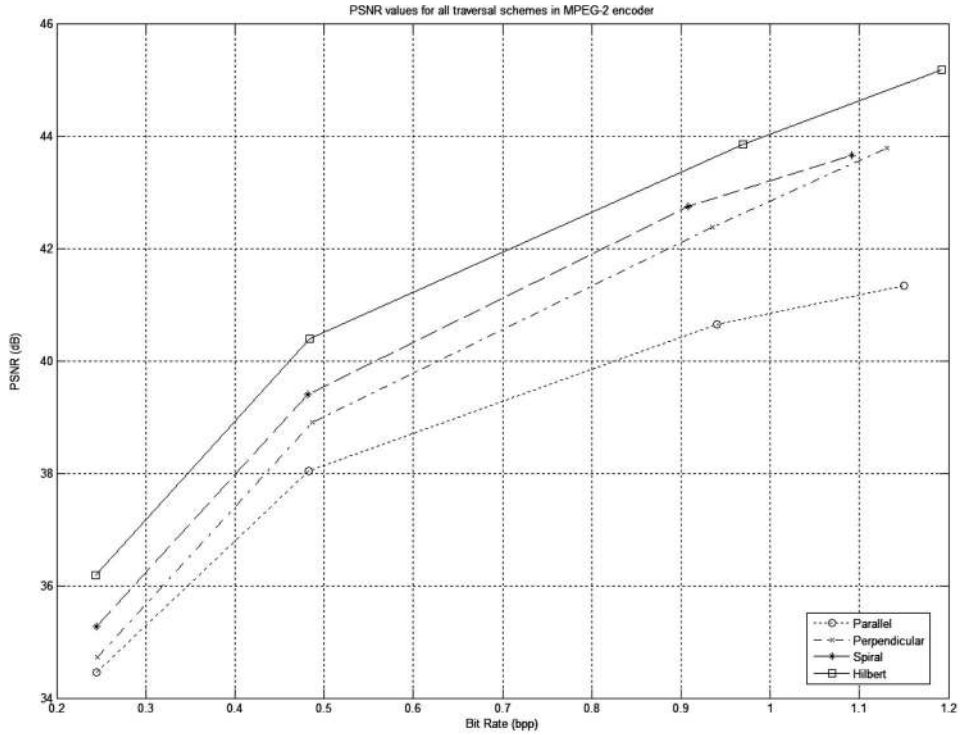
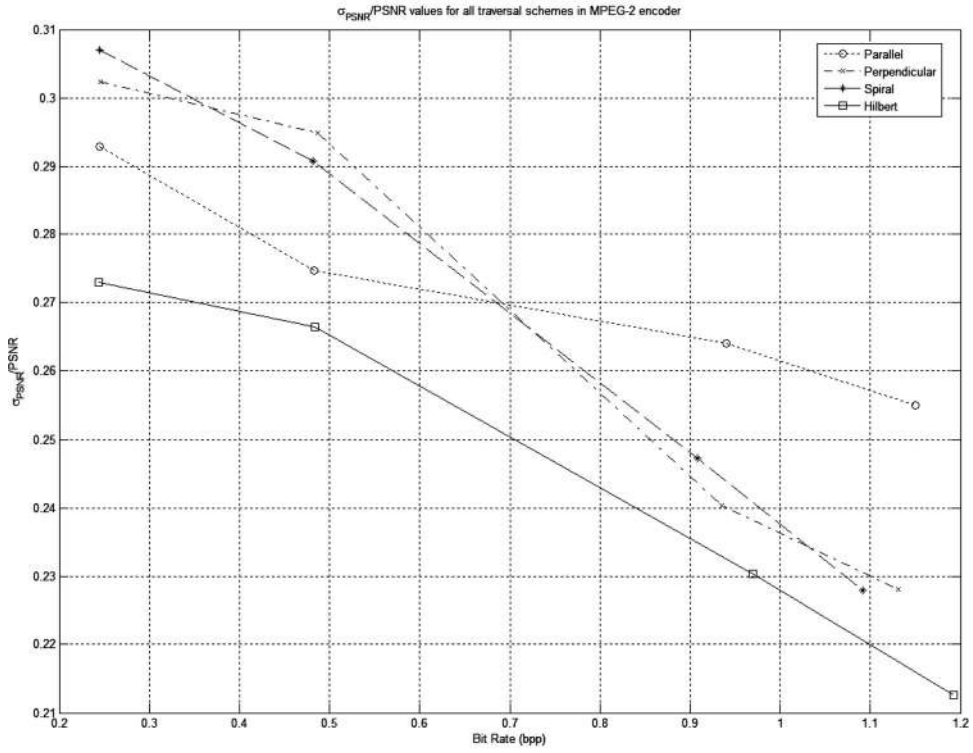


Fig. 6. Synthetic InIms of the fish scene with (a) 8×8 EIs, (b) 256×256 EIs. Physically acquired InIms (c) of a dice with 16×16 EIs and (d) a fireplace with 64×64 EIs.



(a)



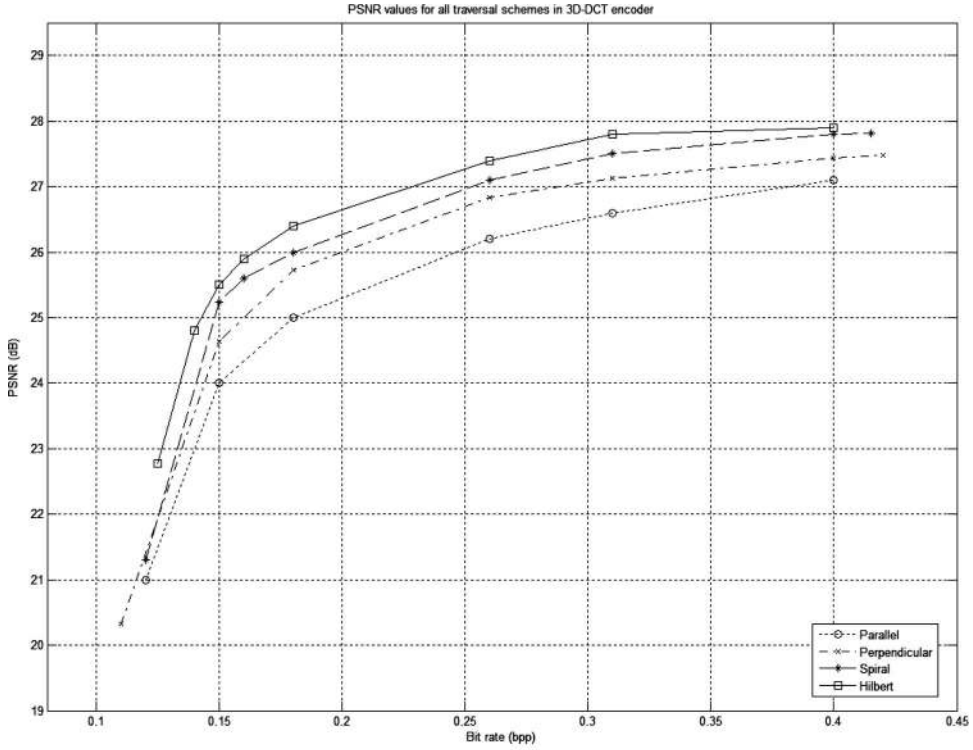
(b)

Fig. 7. (a) $\overline{\mathrm{PSNR}}$ and (b) $\sigma_{\mathrm{PSNR}}/\overline{\mathrm{PSNR}}$ for an encoded InIm as a function of bit rate for the MPEG-2 encoder.

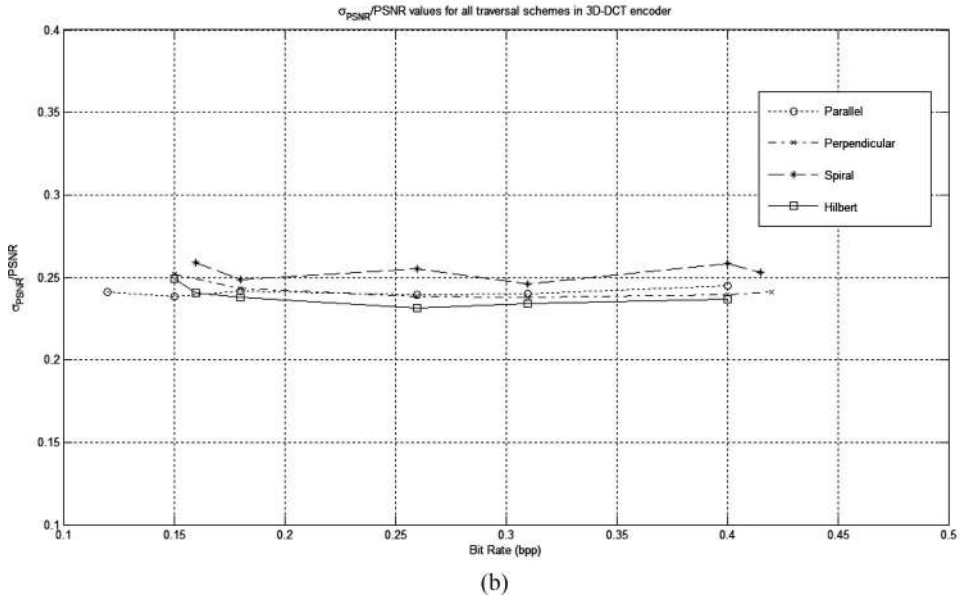
$$\mathrm{MSE} = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}[I(i,\,j) - I_r(i,\,j)]^2, \qquad (4)$$

where $I(i,j)$ are the intensity values of an 8 bit, grayscale, $N \times M$ pixel, original uncompressed InIm, and $I_r(i,j)$ are the intensity values of the decompressed InIm.

However, this measure estimates the overall image quality and does not take into account the special structure of the displayed InIm. In detail, as



(a)



(b)

Fig. 8. (a) PSNR and (b) $\sigma_{\mathrm{PSNR}}/\overline{\mathrm{PSRN}}$ for an encoded InIm as a function of bit rate for the 3D-DCT encoder.

neighboring EIs are simultaneously displayed to produce the 3D representation, they should exhibit similar quality characteristics in order to produce viewable results. In this case large variations of the PSNR values can diminish 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 ($\overline{\mathrm{PSNR}}$) and its standard deviation (σ_{PSNR}) are calculated by using the PSNR values found by applying Eq. (3) for each of the EIs in the InIm. The values for $\overline{\mathrm{PSNR}}$ and σ_{PSNR} for an InIm assembled of $K \times L$ EIs are given by Eqs. (5) and (6), respectively:

$$\overline{\mathrm{PSNR}} = \frac{1}{KL}\sum_{s=1}^{K}\sum_{t=1}^{L}\mathrm{PSNR}_{s,t}, \tag{5}$$

$$\sigma_{\mathrm{PSNR}} = \sqrt{\frac{\sum_{s=1}^{K}\sum_{t=1}^{L}\mathrm{PSNR}_{s,t}^{2}}{KL} - \left(\frac{\sum_{s=1}^{K}\sum_{t=1}^{L}\mathrm{PSNR}_{s,t}}{KL}\right)^{2}}. \tag{6}$$

B. MPEG-2 Evaluation

Whenever an MPEG-2 strategy is employed to encode a video sequence, the quality of the reproduced sequence varies between frames as a result of the encoding process. In the case of video sequences these fluctuations are not perceived by the observer because of the rapid succession of frames. On the other hand, InIms encoded by using an MPEG-2 encoder also exhibit this kind of PSNR variation between EIs, resulting in poor 3D image quality. The use of the Hilbert curve traversal for the InIm decomposition also manages to substantially decrease these variations relative to previously proposed traversal schemes. In the present context we introduce the relative standard deviation, given as $\sigma_{PSNR}/\overline{PSNR}$, which quantifies the fluctuation of the $\overline{PSNR}$ value for each traversal scheme and acts as an index of homogeneity characterization for the displayed InIm. The results for all traversal schemes for both $\overline{PSNR}$ and $\sigma_{PSNR}/\overline{PSNR}$ values for the representative InIm shown in Fig. 6(a) are depicted in Figs. 7(a) and 7(b), respectively.

As is shown in Fig. 7, the values of the $\overline{PSNR}$ in the case where the Hilbert curve traversal is used are higher than the other traversal schemes. In addition there is a significant decrease in the σ_{PSNR} values, indicating that as better localization between jointly encoded EIs is achieved both quality and PSNR fluctuations are improved.

C. 3D-DCT Evaluation

The same methodology was applied for the 3D-DCT encoder [7] targeting InIms that are assembled from a large number of low-pixel-count EIs. The $\overline{PSNR}$ and $\sigma_{PSNR}/\overline{PSNR}$ values for the 3D-DCT applied to the InIm of Fig. 6(d) are depicted in Figs. 8(a)–8(b), respectively. The results show that using this encoder for small-sized EIs has an impact on the variability of the $\overline{PSNR}$ values owing to the absence of the motion estimation procedure present in the MPEG-2 case. The PSNR variation was lower in all traversal schemes than in the MPEG-2 case, with the Hilbert scan achieving the lowest value among the other curves, as depicted in Fig. 8.

4. Conclusions

In this work, in order to take advantage of the high 2D correlation of neighboring EIs, an efficient omnidirectional InIm traversal scheme based on the Hilbert curve was presented for optimizing the performance of InIm encoders that rearrange the EIs in a 1D stream. A metric for evaluating the locality preservation properties of such curves is also presented to assess their performance for different InIm sizes. Finally, a quality-assessment objective metric was also introduced, in addition to the PSNR value, to describe the property of homogeneity that should be preserved when coding InIms. The results show that the Hilbert curve scan validates the locality measure predictions and achieves both greater image quality in terms of PSNR and lower PSNR fluctuation values than previously proposed traversal schemes.

References

1. S. Pastoor and M. Wöpking, "3-D displays: a review of current technologies," Displays **17**(2), 100–110 (1997).
2. M. Halle, "Autostereoscopic displays and computer graphics," Comput. Graph. **31**(2), 58–62 (1997).
3. J.-Y. Son and B. Javidi, "Three-dimensional imaging methods based on multiview images," J. Display Technol. **1,** 125–140 (2005).
4. G. Lippmann, "La Photographie integrale," C. R. Acad. Sci. **146,** 446–455 (1908).
5. J. S. Jang and B. Javidi, "Formation of orthoscopic three dimensional real images in direct pickup one-step integral imaging," Opt. Eng. **42,** 1869–1870 (2003).
6. J.-H. Park, Y. Kim, J. Kim, S.-W. Min, and B. Lee, "Three-dimensional display scheme based on integral imaging with three-dimensional information processing," Opt. Express **12,** 6020–6032 (2004).
7. R. C. Gonzalez and R. E. Woods, *Digital Image Processing,* 2nd ed. (Prentice Hall, 2002).
8. N. P. Sgouros, D. P. Chaikalis, P. G. Papageorgas, and M. S. Sangriotis, "Omnidirectional integral photography images compression using the 3D-DCT," in *Adaptive Optics: Analysis and Methods/Computational Optical Sensing and Imaging/ Information Photonics/Signal Recovery and Synthesis Topical Meetings on CD-ROM,* OSA Technical Digest (CD) (Optical Society of America, 2007), paper DTuA2.
9. R. Zaharia, A. Aggoun, and M. McCormick, "Adaptive 3D-DCT compression algorithm for continuous parallax 3D integral imaging," Signal Process. Image Commun. **17,** 231–242 (2002).
10. N. P. Sgouros, A. G. Andreou, M. S. Sangriotis, P. G. Papageorgas, D. M. Maroulis , and N. G. Theofanous, "Compression of IP images for autostereoscopic 3D imaging applications," in *Proceedings of 3rd International Symposium on Image and Signal Processing and Analysis,* S. Loncaric, A. Neri, and H. Babic, ed. (IEEE, 2003), pp. 223–227.
11. S. Yeom, A. Stern, and B. Javidi, "Compression of 3D color integral images," Opt. Express **12**, 1632–1642 (2004).
12. H. Sagan, *Space-Filling Curves* (Springer-Verlag, 1974).
13. C. Gotsman and M. Lindenbaum, "On the metric properties of discrete space filling curves," IEEE Trans. Image Process. **5,** 794–797 (1996).
14. J. Zhang, S.-I. Kamata, and Y. Ueshige, "A pseudo-Hilbert scan for arbitrarily-sized arrays," IEICE Trans. Fundam. Electron. Commun. Comput. Sci. **E90-A,** 682–690 (2007).
15. S. Kamata and Y. Bandoh, "An address generator of a pseudo-Hilbert scan in a rectangle region," in *Proceedings of IEEE International Conference on Image Processing* (IEEE, 1997), pp. 707–710.
16. K.-L. Chung and L.-C. Chang, "A novel two-phase Hilbert-scan-based search algorithm for block motion estimation using CTF data structure," Pattern Recogn. **37,** 1451–1458 (2004).
17. B. Moon, H. V. Jagadish, C. Faloutsos, and J. H. Saltz, "Analysis of the clustering properties of the Hilbert space-filling curve," IEEE Trans. Knowl. Data Eng. **13,** 124–141 (2001).
18. A. Lempel and J. Ziv, "Compression of two-dimensional data," IEEE Trans. Inf. Theory **IT-32,** 2–8 (1986).
19. D. Salomon, *Data Compression, the Complete Reference,* 3rd ed. (Springer, 2004).
20. B. Furht, J. Greenberg, and R. Westwater, *Motion Estimation Algorithms for Video Compression* (Kluwer Academic, 1997).
21. POV-Ray: The Persistence of Vision Raytracer, www.povray.org.
22. S. S. Athineos, N. P. Sgouros, P. G. Papageorgas, D. E. Maroulis, M. S. Sangriotis, and N. G. Theofanous, "Physical modelling of microlens array setup for use in computer generated

IP," in *Electronic Imaging 2005* (IS&T, SPIE, 2005), paper 5664A-75.
23. 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 the capturing optics," J. Electron. Imaging **15,** 043007 (2006).
24. G. Passalis, N. Sgouros, S. Athineos, and T. Theoharis, "Enhanced reconstruction of 3D shape and texture from integral photography images," Appl. Opt. **46,** 5311–5320 (2007).
25. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE Trans. Image Process. **13,** 600–612 (2004).
26. A. Stoica, C. Vertan, and C. Fernandez-Maloigne, "Objective and subjective color image quality evaluation for JPEG 2000 compressed images," in *Proceedings of IEEE International Symposium on Signals, Circuits and Systems* (IEEE, 2003), pp. 137–140.