Three-dimensional display by smart pseudoscopic-to-orthoscopic conversion with tunable focus

Manuel Martínez-Corral,[1,*] Adrián Dorado,[1] Héctor Navarro,[1] Genaro Saavedra,[1] and Bahram Javidi[2]

[1]3D Imaging and Display Laboratory, Department of Optics, University of Valencia, E-46100 Burjassot, Spain

[2]Department of Electrical and Computer Engineering, University of Connecticut, Storrs, Connecticut 06269, USA

*Corresponding author: manuel.martinez@uv.es

Received 19 February 2014; revised 23 April 2014; accepted 23 April 2014; posted 25 April 2014 (Doc. ID 206389); published 29 May 2014

The original aim of the integral-imaging concept, reported by Gabriel Lippmann more than a century ago, is the capture of images of 3D scenes for their projection onto an autostereoscopic display. In this paper we report a new algorithm for the efficient generation of microimages for their direct projection onto an integral-imaging monitor. Like our previous algorithm, the smart pseudoscopic-to-orthoscopic conversion (SPOC) algorithm, this algorithm produces microimages ready to produce 3D display with full parallax. However, this new algorithm is much simpler than the previous one, produces microimages free of black pixels, and permits fixing at will, between certain limits, the reference plane and the field of view of the displayed 3D scene. Proofs of concept are illustrated with 3D capture and 3D display experiments. © 2014 Optical Society of America

OCIS codes: (100.6890) Three-dimensional image processing; (110.6880) Three-dimensional image acquisition; (120.2040) Displays.

http://dx.doi.org/10.1364/AO.53.000E19

1. Introduction

Conventional photographic cameras do not have the ability to record the angular information carried by the rays of light passing through their objectives [1]. Instead, the irradiance received by any pixel is proportional to the sum of radiances of all the rays, regardless of their incidence angles. Much more interesting is a system with the capacity to register a radiance map with the spatial and angular information of all the rays proceeding from the 3D scene. Such maps have been named in different ways, such as integral photography [2], integral imaging [3], lightfield maps [4], or even plenoptic maps [5,6].

A clever way of recording the plenoptic map produced by 3D scenes was reported by Lippmann [2]. Lippmann proposed in 1908 a technique, called integral photography, based on inserting a microlens array (MLA) in front of the photographic film. This permitted the capture of a set of elemental images (EIs) storing many different, horizontal and vertical, perspectives of a 3D scene. The set of EIs will here be called the integral image of the 3D scene. The second stage of Lippmann's procedure was to insert the, properly developed, photographic film in front of a MLA similar to the one used in the capture. In this way, the light emitted by the EIs is integrated in front of the MLA, producing a 3D reconstruction of the original scene. Binocular observers placed in front of it can see such a reconstructed scene with continuous relief and full parallax. With the fast development of optoelectronic technology, the interest in this technique, now renamed integral imaging (InI), has been resurrected in the past decade since it can be implemented at video rate, and the images can be transmitted in real time [7].

1559-128X/14/220E19-07$15.00/0

In Fig. 1 we show the scheme of an InI capturing setup, and the corresponding captured plenoptic map. Note that any microlens captures the rays passing through its center, and with different angles. Thus, the pixels behind the microlenses register the angular information of the plenoptic map. In observing Fig. 1(b), it is interesting to note that, in the plenoptic map, points aligned with a slanted line correspond to a point source. The slope is proportional to the distance between the point and the microlenses. The intersection of the line with the abscissas axis gives the lateral position of the point. Points aligned in vertical correspond to rays passing through the center of one microlens. Then they correspond to the same EI. Points aligned in horizontal correspond to rays impinging the InI system with the same angle, thus they correspond to orthographic views (also called subimages) of the 3D scene. The capture of the EIs can be implemented not only with a MLA, but also by arranging a set of equidistant digital cameras with their optical axes parallel.

To implement an InI monitor, it is sufficient to project the EIs onto a pixelated screen, and fix a MLA in front of it. If the MLA and the EIs have the same pitch, the light emitted by the pixels of the display is integrated, producing a reconstruction of the 3D irradiance distribution of the original scene. This allows the observers to perceive the 3D nature of the scene, without suffering from the detrimental accommodation–convergence conflict [8].

Note, however, that direct projection of the EIs usually is not possible. Among the problems one has to face is the pseudoscopic-to-orthoscopic (PO) conversion [9]. Others are the difference between the number of cameras in the capture and the number of microlenses in the display. Also important is to fix the scale of the displayed scene, and the position of the reference plane. In this sense, many algorithms have been proposed for the conversion of the EIs [10–18]. However, they have the drawback that an important part of the plenoptic map is lost in the conversion, or that neither the scale nor the position of the displayed scene can be flexibly modified.

The aim of this paper is to report a new algorithm for producing the images ready to be projected onto an InI monitor. The algorithm is based on the algebraic relation between the plenoptic field captured with an InI system and the one captured with a plenoptic camera. The algorithm we report is very simple and allows, together with the PO conversion, the flexible selection of the display resolution, the field of view (FOV), and the position of reference plane. To explain the utility of our approach, we have organized this paper as follows. First, in Section 2, we revisit the concept of a plenoptic camera and verify that plenoptic images can be used for InI display. In Section 3, we explain the conversion algorithm. In Section 4, we perform the InI capture experiment and apply the proposed algorithm to calculate the microimages that are displayed on the InI monitor. Finally, in Section 5 we summarize the main achievements of this research.

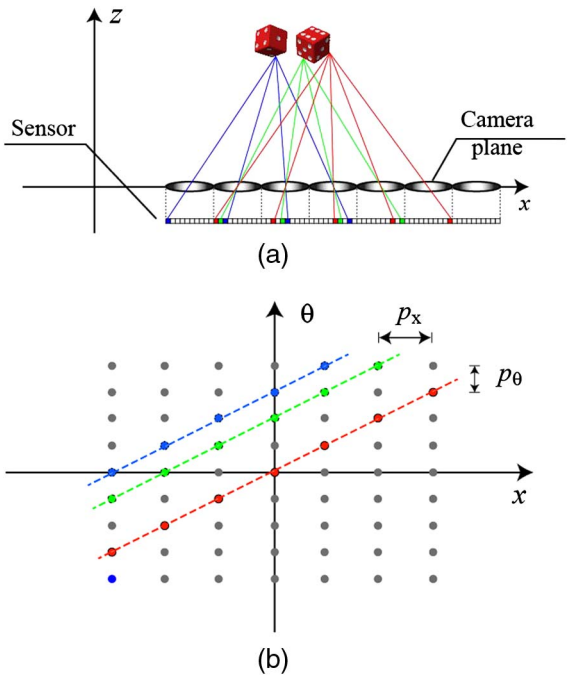


Fig. 1. (a) Scheme of the capture of the plenoptic field with an InI setup. (b) The captured, sampled, plenoptic field.

2. Plenoptic Camera and InI Display

Based on the concepts reported by Lippmann, some research groups [5,7,19] proposed a technique for recording, after a single snapshot, the sampled plenoptic field emitted by a 3D scene. From Adelson and Wang [5], this setup is known as the plenoptic camera.

As shown in Fig. 2, inserting a MLA in front of the sensor of a conventional camera sets a plenoptic camera. The system is adjusted in such a way that a reference plane is conjugated with the MLA through the camera lens [1], and the sensor is conjugated with the camera lens through the microlenses.

The position of the sensor can be calculated according to the correspondence equation:

$$z' = \frac{f^2}{z}. \tag{1}$$

The images recorded by the sensor are here called microimages. Note that the plenoptic camera does

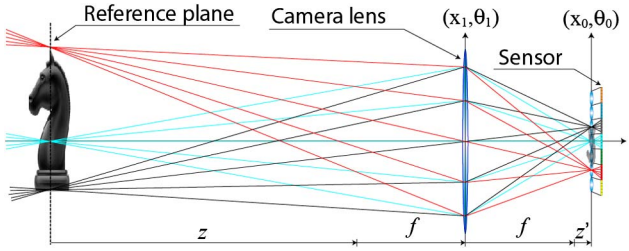


Fig. 2. Scheme of the capture of the plenoptic field with a plenoptic camera.

not capture the plenoptic field emitted by the 3D scene, but the one imaged by the camera lens.

Next, in Fig. 3, we depict a sketch of the plenoptic field captured by the plenoptic camera of Fig. 2. To make this representation we have used the typical sign criterion for the angles, which assigns positive values to the angles that correspond to rotations, or shearing, that follow the counterclockwise direction. Similar to the InI case, any column in the plenoptic map corresponds to a microimage. Again, we can compute subimages (see dotted envelope in the figure), which are obtained by grouping pixels that have the same relative position in their respective microimage. As is clear from Fig. 2, the central pixel of the central microimage collects a ray with slope $\theta_0 = 0$. The other central pixels collect rays with increasing slope. For this reason, the plenoptic map shown in Fig. 3 does not show a rectangular form but a sheared one.

As already reported by some groups [20,21], the microimages recorded with a plenoptic camera are ready to be projected onto an InI monitor to produce autostereoscopic images. This is because the image of the 3D scene produced by the camera lens is in the neighborhood of the MLA (with some parts in front and some parts behind). Naturally, one has to make sure that the two MLAs (the one used in the plenoptic camera and the one placed in front of the display) have the same number of microlenses and f-number.

To verify this property, we obtained in our laboratory a plenoptic picture of a 3D scene. Although there are already some commercial realizations of the plenoptic camera [22,23], we decided to prepare in our laboratory a plenoptic camera in open configuration. As sketched in Fig. 4, a camera lens of $f = 100$ mm was used to conjugate the object reference plane with the MLA. The lens array was composed by 96×60 lenslets of focal length $f_L = 0.93$ mm arranged in a square grid of pitch $p = 0.222$ mm (APO-Q-P222-F0.93 from AMUS). A digital camera (Canon 450D) with a macro-objective of 1:1 was used as the relay system that imaged the microimages onto the sensor. The plenoptic picture captured is shown in Fig. 5.

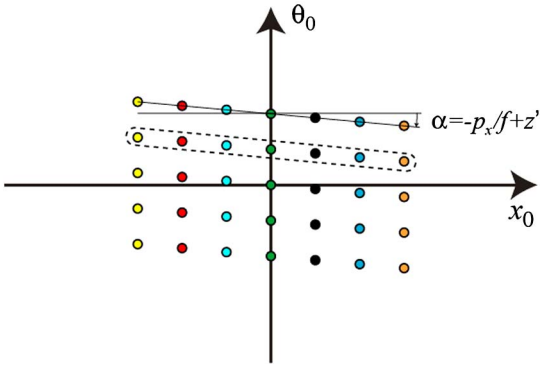


Fig. 3. Sampled plenoptic field captured with the plenoptic camera of the previous figure. Note that the pixels are not arranged in a rectangular grating, but are sheared by an angle $\alpha = -p_x/(f + z')$.

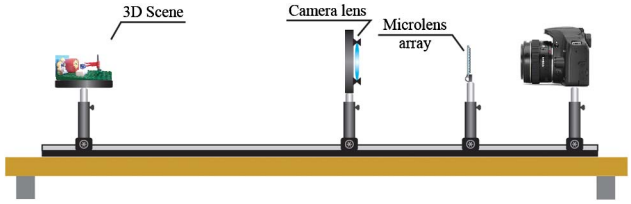


Fig. 4. Scheme of the plenoptic capturing setup.

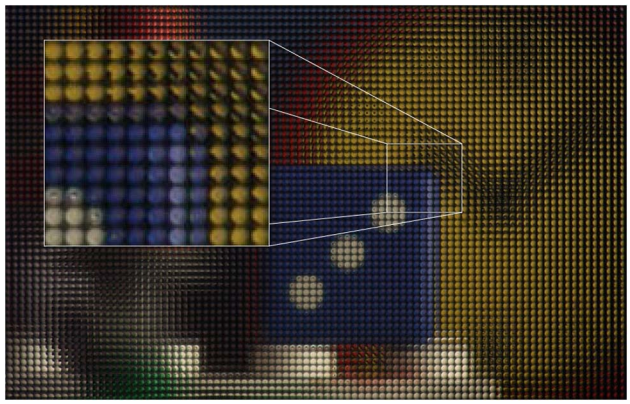


Fig. 5. Plenoptic picture of a 3D scene. The picture is composed of 96×60 microimages with 48×48 pixels each.

As the second step of our demonstration experiment, instead of using a projection-type display [24], we arranged the 3D InI monitor by using an iPad equipped with retina display (10.39 pixels/mm), and a MLA consisting of lenslets of focal length $f_L = 3.3$ mm and pitch $p = 1.0$ mm (Model 630 from Fresnel Technology).

To perform the display experiment, first we resized the plenoptic picture to 998×624 pixels. Since the number of pixels behind any microlens was not an integer, some aliasing was expected in the display. On the other hand, adjustment of the microlens apertures was not possible ($f/4.2$ in the camera and $f/3.3$ in the InI monitor). This will give rise to some distortion in the displayed images. In any case, for this demonstration experiment we considered it convenient to use commercially available elements, instead of customized ones. Next, the MLA was properly aligned in front of the iPad so that pixels were close to the focal plane. Also in this case, exact adjustment could not be made. This was due to the transparent plate that covers the retina display. This small misadjustment resulted in some braiding effect [25]. The InI monitor is shown in Fig. 6.

3. InI Display from Elemental Images

The main drawback of plenoptic cameras is their limited parallax, which is determined by the angle subtended by the camera lens as seen from the center of the reference plane (see Fig. 2). Thus, only in the case of photographing small, close, 3D scenes, can the plenoptic camera record microimages that provide

Fig. 6. Single-frame excerpt from video recording of the implemented InI monitor (Media 1). The video is composed of views of the monitor obtained from different horizontal positions. The small size of the effective display is caused by the small number of microlenses in the capture MLA.

satisfactory 3D-display results. Another problem is the lack of flexibility due to the fact that the two MLAs used for the capture and the display need to have the same number of microlenses and the same f-number.

These problems do not occur in the case of InI capturing setups, in which the parallax can be set, below certain limits, at will. One simply needs to arrange a grid of digital cameras with large enough viewing angle. However, the EIs captured with InI setups are far from being ready for direct projection into InI monitors. They need to be transformed computationally. The transformation algorithm must (a) solve the PO problem, (b) match the number of microimages to the number of microlenses of the InI monitor, and (c) move the 3D scene to the neighborhood of the MLA. In this paper we propose a new and simple conversion algorithm, the smart pseudoscopic-to-orthoscopic conversion (SPOC) 2.0, which satisfies all these requirements and has the additional ability of permitting the selection at will of the reference plane and FOV of the displayed 3D scene.

The algorithm is based in the transformation properties of the plenoptic map. Note that, since the plenoptic formalism is based on light-ray propagation, the conversion is easily analyzed by use of the ABCD matrices. According to ABCD formalism, a translation by distance t and a refraction in a lens of focal length f are represented, respectively, by the matrices

$$T_t = \begin{pmatrix} 1 & -t \\ 0 & 1 \end{pmatrix} \tag{2}$$

and

$$L_f = \begin{pmatrix} 1 & 0 \\ 1/f & 1 \end{pmatrix}. \tag{3}$$

Now, we can easily calculate the plenoptic field at the plane (x_1, θ_1) just before the refraction in the lens (see Fig. 2) from the plenoptic map at plane (x_0, θ_0). To this end, we simply need to calculate the inverse of the matrix, A, that connects such planes. The matrix is obtained as the product

$$A = T_{f+z'} L_f = \begin{pmatrix} 1 - \frac{f+z'}{f} & -(f+z') \\ 1/f & 1 \end{pmatrix}. \tag{4}$$

It is then straightforward to find that

$$\begin{pmatrix} x_1 \\ \theta_1 \end{pmatrix} = A^{-1} \begin{pmatrix} x_0 \\ \theta_0 \end{pmatrix} = \begin{pmatrix} x_0 - \frac{\theta_0}{f} \\ (f+z')x_0 - \frac{z'}{f}\theta_0 \end{pmatrix}. \tag{5}$$

The result of this transformation, for the case of $z' = 0$, which is shown in Fig. 7.

Note that the plenoptic map shown Fig. 7 is similar to the one that could be captured with an array of digital cameras (with the optical axes parallel to each other) if placed at the plane of the camera lens [4,20]. Next, in Fig. 8, we plot the transformed map, but for the case of $z' > 0$.

The map shown in Fig. 8 is similar to the one captured with an array of cameras with their optical

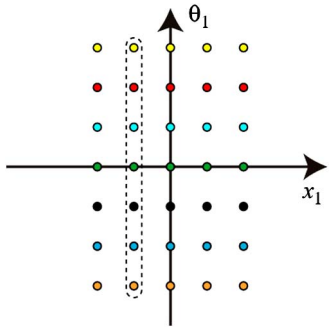


Fig. 7. Sampled plenoptic map at the plane of the camera lens, before refraction, for the case in which the MLA is placed just at the back focal plane of the camera lens, and, therefore, the reference plane is at the infinite (this field has been calculated from the one shown in Fig. 3).

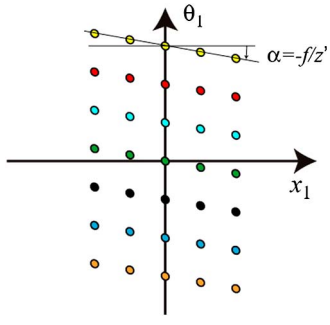


Fig. 8. Plenoptic map at the plane of the camera lens, before refraction, for the case in which the MLA is placed behind the back focal plane of the camera lens, and, therefore, the reference plane is at finite distance.

axes intersecting at a distance $f + z$ (see Fig. 2) in front of them. In this case the shearing angle is

$$\alpha = -\frac{z'}{f} = -\frac{f}{z}. \tag{6}$$

We can conclude that the EIs captured with an array of digital cameras, with their axes parallel to each other, can be easily converted into microimages ready for their projection onto an InI monitor. One simply needs to shear the plenoptic map proportionally to the desired value of z' and crop a rectangular portion from it. Then, by simply transposing rows and columns, the resulting microimages can be directly projected onto a pixelated screen to implement the InI monitor. The main advantages of this transformation over previously reported ones is the easy computation and that there is no loss of plenoptic information.

4. Experimental Results

To demonstrate the proposed approach we performed the capture of a 3D scene with an InI capturing setup. Instead of using an array of digital cameras, we used the so-called synthetic aperture method [26], in which all the EIs are picked up with only one digital camera that is mechanically displaced. The synchronized positioning, shooting, and recording of the EIs was controlled by a LabVIEW code. The digital camera (Canon 450D) was focused on the wooden panel. The camera parameters were fixed to $f = 18$ mm and $f/22.0$, so that the depth of field was large enough to obtain sharp pictures of the entire 3D scene. With this setup we obtained a set of $N_H = N_V = 11$ EIs with pitch $P_H = P_V = 10.0$ mm. Since the pitch is smaller than the size of the CCD sensor (22.2 mm × 14.8 mm), we cropped every EI to remove the outer parts; see Fig. 9(a). Note that by selecting the cropping factor, δ/Δ, one can tune at will the FOV of displayed images. In addition, we resized the EIs so that any image was composed by $n_H = n_V = 147$ pixels (note that the number of pixels of cropped images must equal the number of microlenses of the InI monitor). With the cropped images we composed the integral image (see Fig. 10).

By shifting linearly the crop center [see Fig. 9(b)] one can create a new integral image, which, after applying the conversion algorithm, permits us to tune the position of the reference plane of the displayed 3D images.

Next we applied the proposed conversion algorithm for the fast calculation of the microimage array. The resulting sets are composed by $N_H = N_V = 147$ microimages, having $n_H = n_V = 11$ pixels each. As displayed in Fig. 11 we created three sets of microimages. In Figs. 11(a) and 11(b) the reference plane was set at the cook's head. In Fig. 11(c) the reference plane was set at the window shutter. The difference between Figs. 11(a) and 11(b) is the FOV, i.e., the value of the cropping factor.

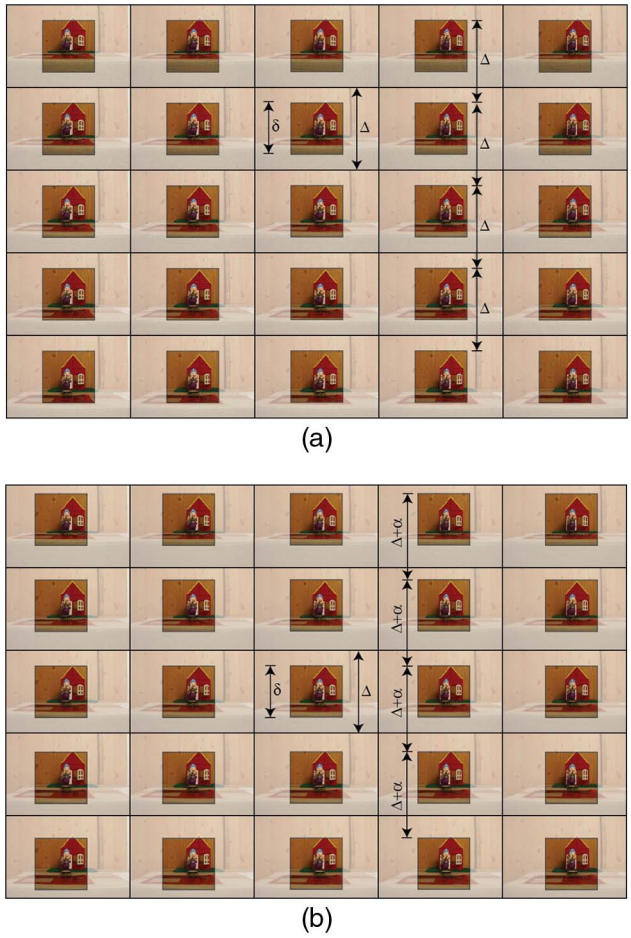
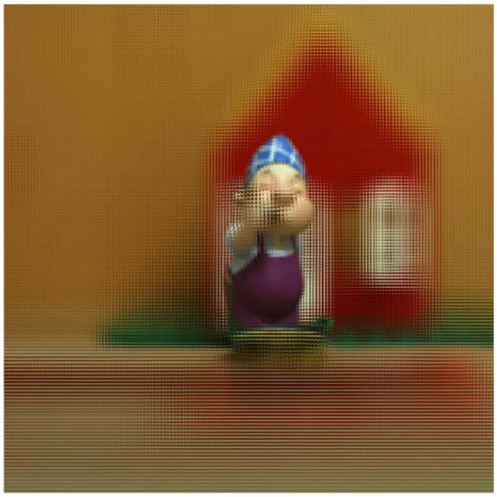


Fig. 9. Illustration of the cropping procedure. (a) When the crop center is set at the image center, the reference plane is set at infinity. Also, the cropping factor δ/Δ determines the FOV of the displayed images. (b) When the crop center is displaced linearly, by factor α, the reference plane is axially shifted.

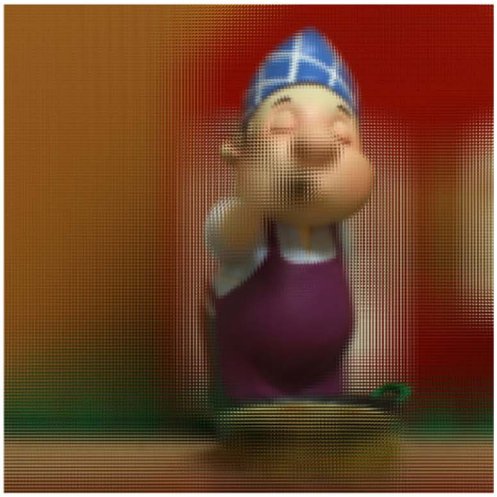
To implement the InI monitor, we used again the iPad equipped with retina display, and the same display MLA as in Section 2. To perform the display experiment, first, we slightly resized the plenoptic picture from 1617 × 1617 to 1527 × 1527 pixels. Next, we aligned the MLA in front of the iPad. The resulting InI display is shown in Fig. 12, for the three versions of the 3D scene.

Fig. 10. Subset of the cropped EIs.

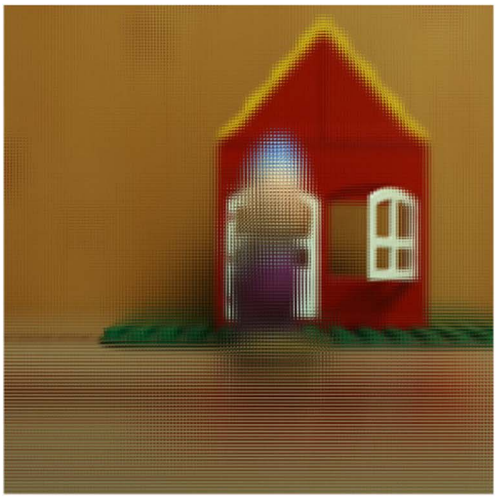
The three movies shown in Fig. 12 have optical features in common. They have the same viewing angle of 13.4°, lateral resolution of 147 × 147 display resolution units, and the same angular resolution of 1.3°.



(a)



(b)



(c)

Fig. 11. Microimages calculated from the EIs.

(a)

(b)

(c)

Fig. 12. Single-frame excerpts from video recording of the InI display. (a) The cook's head is displayed at the plane of the monitor and the house behind that plane (Media 2). (b) Same, but with a smaller FOV (Media 3). (c) The window shutter is displayed at the monitor plane, the house and the wooden panel within the monitor, and the cook in front of the monitor (Media 4).

The viewing angle could be improved to 17.6° by removing the plate that protects the pixels of the iPad. This viewing angle permitted easy perception of depth in a binocular observation. Since we cannot show here such perception, we have prepared movies with the pictures obtained by photographing the iPad while the camera was laterally displaced. Although the movies show only the horizontal parallax, the monitor, of course, displays also the vertical parallax.

If we compare Figs. 12(a) and 12(b) we verify that our algorithm permits creating very different 3D visualization experiences from the same stored information. In particular, we can switch from panoramic to foreground visualization. If we compare now Figs. 12(a) and 12(c), we see that we can also select at will the reference plane. In other words, we can decide the amount of the scene that is within and outside of the monitor. This can be very useful for content creators since they can decide *a posteriori* the level of immersion in the displayed scene.

5. Conclusion

We have reported a new algorithm for the calculation of the microimages to be displayed in an InI monitor. Our algorithm, here called the SPOC 2.0, is based on the smart cropping of the EIs captured with an array of digital cameras, and in the transposing relation between the plenoptic field captured with a plenoptic camera and the one recorded with an InI setup. The proposed cropping and mapping algorithm permits easy adaptation to the InI display requirements. The SPOC 2.0 algorithm surpasses previous PO algorithms in that it produces microimages without black pixels, and can tune the FOV and the reference plane. In addition, the number of microlenses and the number of pixels per microlens can also be set in order to adapt the calculated microimages to the geometric characteristics of the InI monitor.

This work was supported in part by the Ministerio de Economia y Competitividad (MINECO), Spain through Grant DPI2012-32994), and also by Generalitat Valenciana through Grant PROMETEO2009-077. A. Dorado acknowledges a predoctoral fellowship from the MINECO. H. Navarro acknowledges a predoctoral fellowship from the Generalitat Valenciana.

References

1. R. Ng, "Digital light field photography," Ph.D. dissertation (Stanford University, 2006).
2. G. Lippmann, "Epreuves reversibles donnant la sensation du relief," J. Phys. Theor. Appl. **7**, 821–825 (1908).
3. H. Arimoto and B. Javidi, "Integral three-dimensional imaging with digital reconstruction," Opt. Lett. **26**, 157–159 (2001).
4. M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," ACM Trans. Graph. **25**, 924–934 (2006).
5. E. H. Adelson and J. Y. A. Wang, "Single lens stereo with plenoptic camera," IEEE Trans. Pattern Anal. Mach. Intell. **14**, 99–106 (1992).
6. T. Georgiev and A. Lumsdaine, "The focused plenoptic camera and rendering," J. Electron. Imaging **19**, 021106 (2010).
7. F. Okano, J. Aral, H. Hoshino, and I. Yuyama, "Three-dimensional video system based on integral photography," Opt. Eng. **38**, 1072–1077 (1999).
8. H. Hiura, T. Mishina, J. Arai, K. Hisatomi, Y. Iwadate, T. Ito, and S. Yano, "A study on accommodation response and depth perception in viewing integral photography," *Proceedings of 3D Systems and Applications, Osaka (Japan)* (2013), paper P2-2.
9. F. Okano, H. Hoshino, J. Arai, and I. Yayuma, "Real time pickup method for a three-dimensional image based on integral photography," Appl. Opt. **36**, 1598–1603 (1997).
10. A. Aggoun, "Pre-processing of integral images for 3-D displays," J. Disp. Technol. **2**, 393–400 (2006).
11. M. Martínez-Corral, B. Javidi, R. Martínez-Cuenca, and G. Saavedra, "Formation of real, orthoscopic integral images by smart pixel mapping," Opt. Express **13**, 9175–9180 (2005).
12. K. S. Park, S. W. Min, and Y. Cho, "Viewpoint vector rendering for efficient elemental image generation," IEICE Transactions on Information and Systems **E90-D**, 233–241 (2007).
13. H. Deng, Q. Wang, and D. Li, "Method of generating orthoscopic elemental image array from sparse camera array," Chin. Opt. Lett. **10**, 061102 (2012).
14. D.-H. Shin, B.-G. Lee, and E.-S. Kim, "Modified smart pixel mapping method for displaying orthoscopic 3D images in integral imaging," Opt. Lasers Eng. **47**, 1189–1194 (2009).
15. J.-H. Park, K. Hong, and B. Lee, "Recent progress in three-dimensional information processing based on integral imaging," Appl. Opt. **48**, H77–H94 (2009).
16. E. Sahin and L. Onural, "A comparative study of light field representation and integral imaging," Imag. Sci. J. **58**, 28–31 (2010).
17. H. Navarro, R. Martínez-Cuenca, G. Saavedra, M. Martínez-Corral, and B. Javidi, "3D integral imaging display by smart pseudoscopic-to-orthoscopic conversion (SPOC)," Opt. Express **18**, 25573–25583 (2010).
18. J.-H. Jung, J. Kim, and B. Lee, "Solution of pseudoscopic problem in integral imaging for real-time processing," Opt. Lett. **38**, 76–79 (2013).
19. N. Davies, M. McCormick, and L. Yang, "Three-dimensional imaging systems: a new development," Appl. Opt. **27**, 4520–4528 (1988).
20. H. Navarro, J. C. Barreiro, G. Saavedra, M. Martínez-Corral, and B. Javidi, "High-resolution far-field integral-imaging camera by double snapshot," Opt. Express **20**, 890–895 (2012).
21. T. Iwane, "Light field camera and IP display," *Proceedings of 3D Systems and Applications, Osaka (Japan)* (2013), p. 32.
22. "3D lightfield camera," http://www.raytrix.de/.
23. "Lightfield based commercial digital still camera," http://www.lytro.com.
24. C.-W. Chen, M. Cho, Y.-P. Huang, and B. Javidi, "Improved viewing zones for projection type integral imaging 3D display using adaptive liquid crystal prism array," J. Disp. Technol. **10**, 198–203 (2014).
25. H. Navarro, R. Martínez-Cuenca, A. Molina-Martín, M. Martínez-Corral, G. Saavedra, and B. Javidi, "Method to remedy image degradations due to facet braiding in 3D integral imaging monitors," J. Disp. Technol. **6**, 404–411 (2010).
26. J. S. Jang and B. Javidi, "Three-dimensional synthetic aperture integral imaging," Opt. Lett. **27**, 1144–1146 (2002).