

# High-resolution far-field integral-imaging camera by double snapshot

H. Navarro,<sup>1</sup> J.C. Barreiro,<sup>1</sup> G. Saavedra,<sup>1</sup> M. Martínez-Corral,<sup>1,\*</sup> and B. Javidi<sup>2</sup>

<sup>1</sup>Department of Optics, University of Valencia, E-46100 Burjassot, Spain

<sup>2</sup>Electrical and Computer Engineering Dept., University of Connecticut, Storrs, CT 06269-1157 USA

\*manuel.martinez@uv.es

**Abstract:** In multi-view three-dimensional imaging, to capture the elemental images of distant objects, the use of a field-like lens that projects the reference plane onto the microlens array is necessary. In this case, the spatial resolution of reconstructed images is equal to the spatial density of microlenses in the array. In this paper we report a simple method, based on the realization of double snapshots, to double the 2D pixel density of reconstructed scenes. Experiments are reported to support the proposed approach.

©2012 Optical Society of America

**OCIS codes:** (100.6890) Three-dimensional image processing; (110.6880) Three-dimensional image acquisition; (110.4190) Multiple imaging; (120.2040) Displays.

---

## References and links

1. G. Lippmann, "Epreuves reversibles donnant la sensation du relief," *J. Phys.* **7**, 821–825 (1908).
2. B. Javidi and F. Okano, eds., *Three-Dimensional Television, Video and Display Technologies* (Springer-Verlag, 2002).
3. S.-H. Hong, J.-S. Jang, and B. Javidi, "Three-dimensional volumetric object reconstruction using computational integral imaging," *Opt. Express* **12**(3), 483–491 (2004).
4. J.-H. Park, K. Hong, and B. Lee, "Recent progress in three-dimensional information processing based on integral imaging," *Appl. Opt.* **48**(34), H77–H94 (2009).
5. B. Javidi, R. Ponce-Díaz, and S.-H. Hong, "Three-dimensional recognition of occluded objects by using computational integral imaging," *Opt. Lett.* **31**(8), 1106–1108 (2006).
6. S. Yeom, B. Javidi, and E. Watson, "Photon counting passive 3D image sensing for automatic target recognition," *Opt. Express* **13**(23), 9310–9330 (2005).
7. F. Okano, J. Arai, H. Hoshino, and I. Yuyama, "Three-dimensional video system based on integral photography," *Opt. Eng.* **38**(6), 1072–1077 (1999).
8. N. Davies, M. McCormick, and L. Yang, "Three-dimensional imaging systems: a new development," *Appl. Opt.* **27**(21), 4520–4528 (1988).
9. J. P. Lüke, F. Perez Nava, J. G. Marichal-Hernandez, J. M. Rodriguez-Ramos, and F. Rosa, "Near real-time estimation of super-resolved depth and all-in-focus images from a plenoptic camera using graphics processing units," *Int. J. Digit. Multimed. Broadcast.* **2010**, 942037 (2010).
10. H. Navarro, R. Martínez-Cuenca, A. Molina, M. Martínez-Corral, G. Saavedra, and B. Javidi, "Method to remedy image degradations due to facet braiding in 3D InI monitors," *J. Display Technol.* **6**(10), 404–411 (2010).
11. J. Y. Son, B. Javidi, S. Yano, and K. H. Choi, "Recent developments in 3-D imaging technologies," *J. Display Technol.* **6**(10), 394–403 (2010).
12. S. H. Hong, J. S. Jang, and B. Javidi, "Three-dimensional volumetric object reconstruction using computational integral imaging," *Opt. Express* **12**(3), 483–491 (2004).
13. Y. T. Lim, J. H. Park, K. C. Kwon, and N. Kim, "Resolution-enhanced integral imaging microscopy that uses lens array shifting," *Opt. Express* **17**(21), 19253–19263 (2009).
14. A. Kubota, A. Smolic, M. Magnor, M. Tanimoto, T. Chen, and C. Zhang, "Multiview imaging and 3DTV," *IEEE Signal Process. Mag.* **24**, 10–21 (2007).
15. E. H. Adelson and J. Y. A. Wang, "Single lens stereo with plenoptic camera," *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(2), 99–106 (1992).
16. M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," *ACM Trans. Graph.* **25**(3), 924–934 (2006).
17. J. S. Jang and B. Javidi, "Three-dimensional integral imaging with electronically synthesized lenslet arrays," *Opt. Lett.* **27**(20), 1767–1769 (2002).
18. S. H. Hong and B. Javidi, "Improved resolution 3D object reconstruction using computational integral imaging with time multiplexing," *Opt. Express* **12**(19), 4579–4588 (2004).

## 1. Introduction

Stereoscopic or auto-stereoscopic monitors usually produce visual fatigue in the audience due to the convergence-accommodation conflict. An attractive alternative to these techniques is the so-called Integral Photography (or Integral Imaging—InI), proposed by Lippmann in 1908 [1], and resurrected about two decades ago due to the fast development of electronic sensors and displays [2]. The Lippmann idea was that one can store the 3D image of an object by taking many 2D photos of it from different views. This can be easily done by using a micro lens array (MLA) or a camera array. The key of his proposal is that each elemental image stores a different perspective of the object. When this information is projected onto a 2D display placed in front of an array of micro lenses, the perspectives are integrated into a 3D image. The reconstructed scene is perceived in 3D by the observer, independent of his/her position. Since an InI monitor truly reconstructs the 3D scene, the observation is produced without special goggles, with full parallax and potentially absent of any visual fatigue.

Although InI concept was initially intended for the capture and display of 3D pictures or movies, in the last decades other interesting applications have been proposed. Examples are the digital reconstruction of spatially incoherent 3D scenes [3,4], the reconstruction of partially occluded objects [5], or the sensing and recognition of low light level scenes [6]. Note that when the scene is far from the MLA so that the parallax of recorded elemental images is very small, it is necessary a depth-control lens [7,8] to image the 3D scene into the neighborhood of the MLA. In that case, the spatial resolution of reconstructed images is determined by the spatial density of lenses in the array. But to avoid diffraction effects, this density is in practice limited. Thus, the spatial resolution suffers in the reconstruction of the scene.

One solution to this problem is to use large-size image sensors. This would allow increasing the number of microlenses. Additionally, the use of computer-graphics post-processing algorithms also permits the improvement of the quality of reconstructed images [9]. Note, however, that algorithms cannot recover the information that was not captured by optical means. In this paper, we propose a technique that can increase the spatial resolution of the reconstructed scene. To obtain that, we propose a double snapshot recording. This allows synthesizing an integral image with doubled number of elemental images. This implies to double the number of pixels of reconstructed scenes.

To present the new approach, we have organized the paper as following. In Section 2, we explain the principles of the near-field integral imaging. In Section 3, we discuss the fundamental differences between near-field and far-field integral imaging. In Section 4 we describe our method for improvement of resolution, and perform a proof-of-concept experiment. Finally in Section 5, we summarize the main achievements of this paper.

## 2. Near-field integral imaging

The original Lippmann concept is based on the capture of many perspectives of a 3D scene by means of a multilens recording of a 3D scene. As shown in Fig. 1, the image sensor is placed behind the lenslets. The individual images obtained by any lens are noted as elemental images (EI). To store the 3D scene, it is useful to record wide viewing angle EIs, so that any part of the 3D scene appears in multiple EIs, and also to have a sufficient angular extension of the camera array, as seen from the object.

The proper selection of the capture parameters strongly depends on the application. For example, when the aim of the capture is the recording of EIs intended for being displayed in an InI monitor, one has to take into account that microlens pitch is the resolution unit in InI displays [10]. Besides, a density of 16-20 pixels/EI is considered enough for producing good

perspective resolution in displayed 3D images. Thus, for this kind of application a large number of EIs with moderate px/EI are required [11].

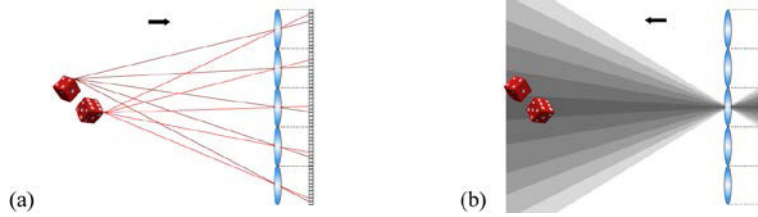


Fig. 1. Near-field integral imaging: (a) In the capture stage, the EIs of the 3D scene are recorded with large parallax; (b) In the reconstruction stage, the EIs are projected through the centers of the lenses. The resolution of reconstructed images is limited by the number of pixels/EI.

A different case, the EIs are captured with the aim of reconstructing the 3D scene in a stack of planes parallel to the camera array. There are different methods for performing such calculations [10,12,13]. These reconstruction algorithms basically project the EIs (through pinholes placed at the centers of the lenses) and then process them to reconstruct the scene. Thus, the spatial resolution of the reconstructed images mainly depends on the pixels/EI whereas the angular resolution, i.e. the segmentation capacity of 3D reconstruction, is restricted by the number of lenses of the array. Note that the overall number of pixels/EI is constrained because it determines the data bandwidth associated to the transmission of InI information. Thus, it is convenient to record small number of EIs but with high number of pixels/EI. This scheme is used when the 3D object is near the lens array. Near means that the angular extension of the lens array as seen from the 3D scene is high. This can correspond to the case in which a small scene [a few cm in size] is captured with an array of microlenses. It can also correspond to the capture of macroscopic 3D scenes with an array of digital cameras arranged in a grid whose pitch is of some centimeters [14]. To highlight the differences between different InI architectures, we will name this type as Near-Field Integral Imaging (NInI).

### 3. Far-field integral imaging

The 3D scene may be far from the camera, so that the angular extension of the array of lenses is very small. In this case, a camera lens, also named as depth-control lens [7,8], is necessary to image the reference plane of the far 3D scene onto MLA. Some parts of the 3D scene are imaged in front of the MLA and other parts behind the MLA. Since this capture modality, shown in Fig. 2, is essentially different from the one described in the previous section, we denote it as Far-Field Integral Imaging (FInI). In computer-graphics community, the FInI cameras involve the same optical principles as the plenoptic cameras [15].

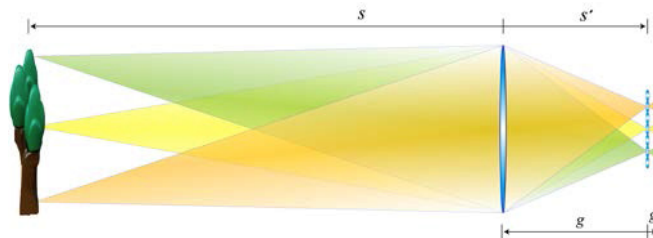


Fig. 2. Capture setup of far-field integral imaging. The reference plane is the plane that is conjugated with the MLA.

Note that using the camera lens has the effect of transposing the resolution constraints [16]. Thus, in FInI the MLA pitch determines the spatial resolution of reconstructed sections

of the 3D scene. The angular resolution, or segmentation capacity of the 3D reconstruction, is restricted by the number pixels/EI. To guarantee good spatial resolution in reconstructed sections of the 3D scene, a large number of small microlenses are required. Besides, to ensure sufficient segmentation capacity, only a density of around 15 pixels/EI is necessary.

To better understand this transposition property, it is worthy to note that from the captured EIs we can calculate the so-called sub-images, or view-images [4], by extracting and composing the pixels at the same local position in every elemental image. As we can see in Fig. 3, the sensor and the camera lens are conjugated through the microlenses. Then, all the pixels of a sub-image (for example the red pixels in the Fig. 3) only receive the light proceeding from the 3D scene and passing through a specific sub-aperture of the camera lens. The sub-aperture is defined as the conjugated of the red pixels through their corresponding microlens. Any sub-image sees a different perspective of the scene and has a high depth of field, as correspond to images obtained through smaller sub-apertures. The number of pixels of these sub-images is just equal to the number of microlenses of the array.

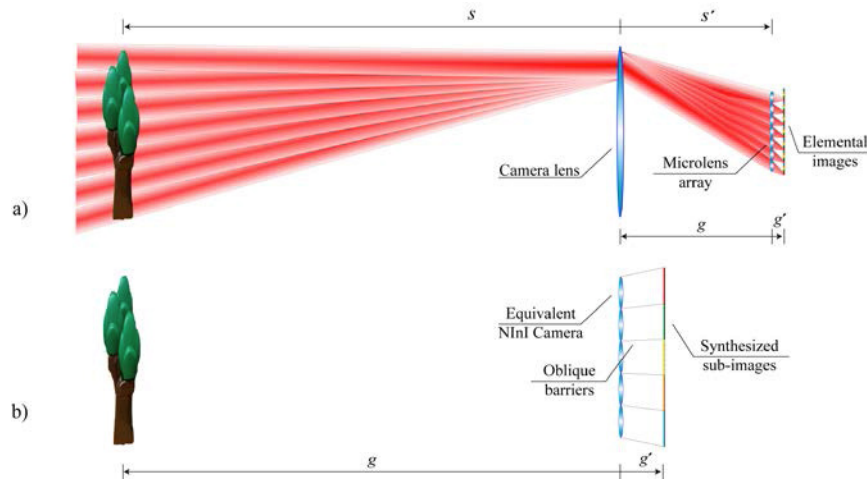


Fig. 3. (a) The sub-images are synthesized from the pixels placed at the same local positions in any elemental image; (b) The sub-images of the FInI cameras are equivalent to the elemental images that could be obtained by a NInI camera, but with the barriers arranged in conical form.

From Fig. 3, we see that the sub-images synthesized from the FInI EIs, are just equivalent to the EIs that could be obtained with a NInI camera in which the microlenses were placed at the position of the camera lens. These NInI EIs are separated not by parallel barriers, but by conical barriers. Naturally, it is easy to apply the classical algorithms for 3D reconstructions by means of back-projection concept. The result of the 3D reconstruction will be a stack of 2D images whose resolution is determined by the density of pixels of the sub-images, that is, by pitch of the MLA in the FInI capture setup.

#### 4. Improving the spatial resolution in far-field integral imaging

In a typical FInI setup, the size of the MLA should be equal to the size of the image sensor. Let us take into account that in order to avoid diffraction effects, the pitch of the MLA cannot go below 100 microns. This fact is limiting the resolution of reconstructed images (in the case of typical matrix sensor size of 22.2x14.9 mm) to about 220x150 pixels. This is a high constraint that could be overcome by using larger image sensors. However, there is also a limit in this case, since camera lenses are designed free of aberrations only for a certain angular range.

To improve the resolution by optical means, we propose to increase the effective number of microlenses by recording two sets of elemental images. The second snapshot is done after

displacing the MLA half pitch along the diagonal direction. With the two sets of EIs, we compose a new set of synthetic EIs in which the pitch is reduced by factor  $\sqrt{2}$  (see Fig. 4). This is equivalent to improve the number of pixels of any 2D reconstructed section by factor 2. Note that the technique of moving the MLA was already proposed in the context of NInI [17,18]. However, as explained above, in such architecture there is no proportionality between the number of snapshots and the improvement of resolution.

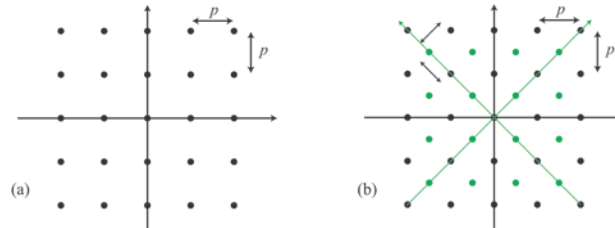


Fig. 4. (a) Grid of the array of microlenses in the case of single snapshot FInI; (b) Grid of the synthetic array in the case of the double snapshot.

To demonstrate our technique, we performed the proof-of-concept experiment shown in Fig. 5. A camera lens of  $f = 100$  mm was used to conjugate the reference plane with the MLA. The MLA was composed by lenslets of focal length 0.222 mm arranged in square grid of pitch  $p = 0.930$  mm (APO-Q-P222-F0.93 model from AMUS). A digital camera with a macro objective 1:1 was used to image the elemental images onto the sensor. To take the second snapshot we moved the MLA by means of micrometer actuators. However in a more advanced prototype the second snapshot could be realized, for example, by inserting, between the camera lens and the MLA, a folding phase plate with linear phase variation. This mechanism is similar to the semi-automatic moving mirror system used in reflex cameras.

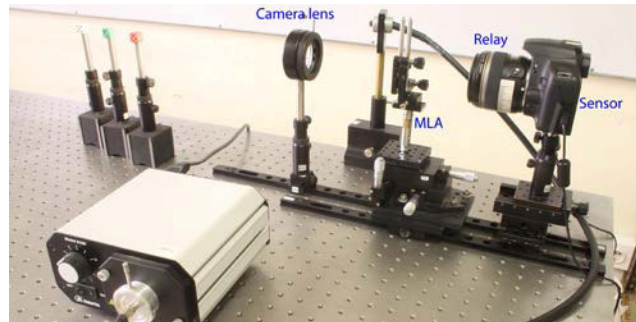


Fig. 5. Experimental setup

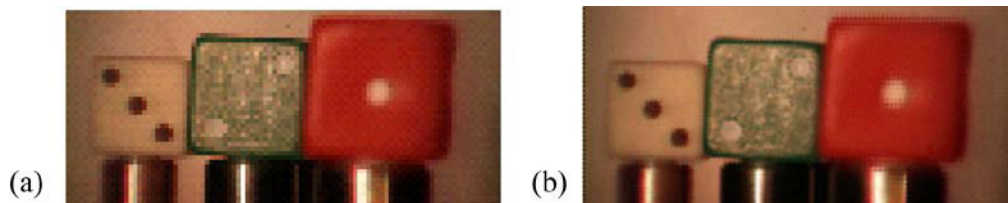


Fig. 6. Central subimage obtained (a) after single snapshot, (b) after double snapshot.

After the first snapshot we obtained the conventional EIs. The set was composed by  $99 \times 44$  EIs with  $41 \times 41$  pixels each. After the second snapshot we could calculate the synthetic EIs set, which was composed by  $140 \times 62$  EIs with  $41 \times 41$  pixels each. The number of EIs in our approach was then double that of the conventional one. We applied the pixel mapping procedure to calculate the two subimage sets. Due to the transposition nature of this

transformation, both subimage sets were composed by 41x41 images, but the number of pixels of any subimage was 4356 pixels, in the conventional set, and 8680 in the synthesized one. In Fig. 6, we present the central subimage for both cases. It can be seen that the spatial resolution obtained after the double snapshot is clearly superior to the conventional one.

We applied the reconstruction algorithm to the two sets of subimages. In particular, in Fig. 7 we show the reconstructed field in the planes of the dices. In the left column we show the images reconstructed from the conventional subimages. In the right-hand column we show the reconstructions obtained from the subimages synthesized after the double shot. The improvement in resolution is apparent.

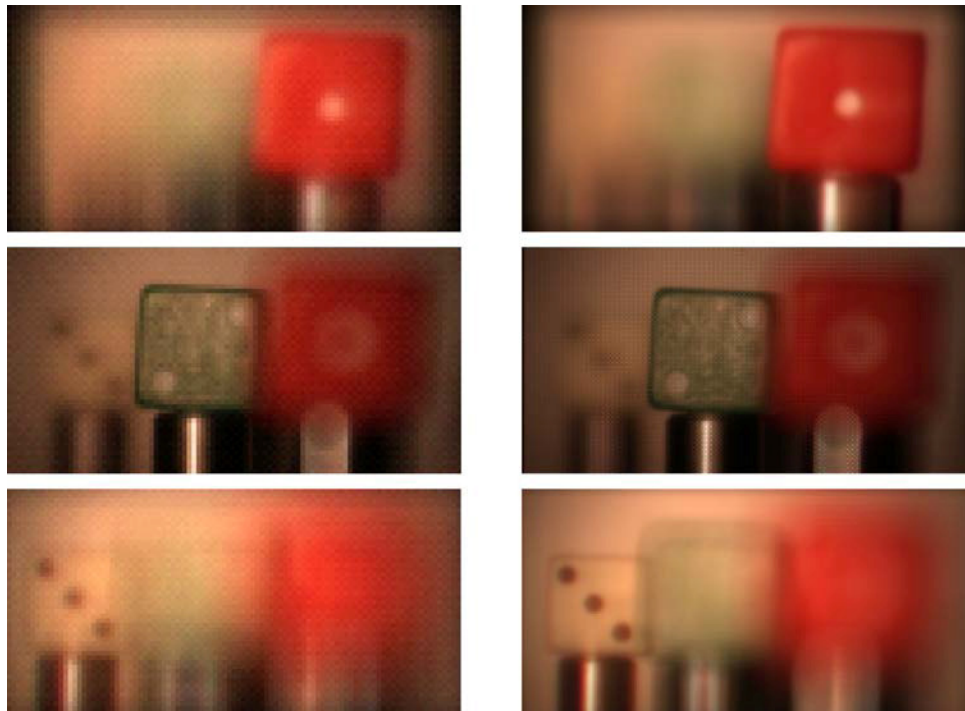


Fig. 7. The left column shows the images reconstructed from the single-shot set of subimages. The right column shows the images obtained from our proposed double shot.

## 5. Conclusions

We have shown that it is possible to double the resolution of reconstructed images in far-field integral imaging. We have performed a proof-of-principle experiment that demonstrates the improvement in resolution. Naturally, in an advanced prototype more pixels are required for reconstructed images. This can be obtained by use of large-size matrix detectors. Also the quality of reconstructed images can be improved by computer graphics post-processing. But our technique provides a method for increasing the final resolution of reconstructed images. A variety of applications in InI can benefit from the proposed approach [19].

## Acknowledgments

This work was supported in part by the Ministerio de Ciencia e Innovación, Spain (Grant FIS2009-9135), and also by Generalitat Valenciana (Grant PROMETEO2009-077).