

3D multi-view passive sensing and visualization using randomly distributed sensors

Bahram Javidi and Mehdi Daneshpanah

Department of Electrical and Computer Engineering,
University of Connecticut, Storrs CT 06269, USA

ABSTRACT

Three dimensional (3D) passive imaging systems are proven to be effective in a number of applications including Automatic Target Recognition (ATR). Such systems are traditionally designed around a regular, fixed grid of pickup locations such as lenslet arrays – a constraint that can not always be met for certain applications. With the recent advancements in this area, many applications call for more generic form of 3D imaging. Here, we overview out work in the area of multi-perspective imaging based on randomly distributed passive sensing. In particular, we propose a passive 3D imaging and visualization system with multiple view acquisitions from randomly distributed sensors. This method can further extend the applications of passive 3D imaging systems to areas such as 3D aerial imaging, collaborative imaging and etc. We discuss some of implications for improving performance of ATR algorithms.

Keywords: Three dimensional imaging, passive imaging, randomly distributed sensing

1. INTRODUCTION

Traditionally, two dimensional imaging systems have been the dominant method of sensing the world visually and to perform target recognition and identification tasks. Sophisticated Automatic Target Recognition (ATR) algorithms have been developed to tackle various recognition problems based on 2D images. Stereoscopic imaging and display provides 3D cues to some extent but is limited to fixed viewpoint capture and reconstruction.¹⁻³ In the recent years however, the interest in multi dimensional imaging, display and information processing systems has increased.⁴⁻⁸ Various active and passive imaging systems are actively being researched for different applications. In the case of three dimensional (3D) imaging systems, the information provided by the sensor provides a more accurate representation of the real world and thus is potentially valuable to ATR as an enabling technology.⁹⁻¹¹ Integral Imaging (II) is among the 3D image sensing techniques that bears promising results.¹²⁻¹⁵ This technique is also known as integral photography¹⁶ which is based on the original work of Lippmann with lenticular sheets¹⁷⁻¹⁹ and is classified under passive multi-perspective 3D imaging systems. The depth information can be extracted from ensemble of images captured from slightly different perspectives that are known as *elemental images*. Three dimensional reconstruction of the scene under investigation can be accomplished optically or digitally in continuous viewing angle, full parallax and full color without the need for coherent sources of illumination.²⁰ The method is relatively simple and low cost comparing to active techniques such as holography⁶ and laser radar.²¹

Integral Imaging based optical displays provide autosterioscopic 3D image or video²² with no need for special eyewear to perceive 3D cues. This is made possible essentially by recording the intensity and direction of light rays, i.e. the light field,²³ emanated from the scene in the capture stage and later back propagating the rays with the same parameters in the display stage.¹⁵ Developments in this venue include aberration reduction,²⁴ extension of depth of field,^{25,26} use of gradient index lens arrays²⁷ to handle the orthoscopic to pseudoscopic conversion, also resolution improvement methods including use of moving lenslet technique (MALT)²⁸ and electronically synthesized moving Fresnel lenslets.²⁹ Nevertheless, optical reconstruction approach suffers from low resolution,

Further author information: (Send correspondence to Bahram Javidi)

Bahram Javidi: E-mail: bahram@engr.uconn.edu

Mehdi Daneshpanah: E-mail: mehdi@engr.uconn.edu

low sampling rate, quality degradation due to diffraction, limited dynamic range and low overall visual quality partly due to limitation of electro-optical projection devices.¹⁵

On the other hand, computational reconstruction techniques deliver a more flexible framework to extract and exploit 3D cues by digital manipulation of elemental images.³⁰⁻³³ The reconstruction process amounts to calculation of the light ray distribution over a particular plane in the scene from the intensity information in the captured elemental images. The extracted 3D information can then be used to improve ATR and classification performance,⁹ arbitrary view display of the scene,³⁴ digital slicing,³⁵ digital refocusing,³⁶ 3D artwork and etc.

Conventional Integral Imaging is historically developed based on lenticular sheets and lenslet arrays. However, limitations in the field of view, resolution-parallax compromise, high aberration and low imaging resolution of each lenslet have led to development of Synthetic Aperture Integral Imaging (SAII).^{37,38} In this technique a conventional 2D imaging sensor scans an area in a regular grid pattern and each elemental image is acquired in full frame. This enables one to obtain larger field of view, low aberration and high resolution elemental images. The sensitivity of the 3D digital reconstruction to uncertainties in sensor position measurement has also been analyzed.³⁹ However, in many practical applications maintaining a regular grid is not feasible.

In this paper, we overview our work in the area of randomly distributed 3D sensing^{40,41} that relaxes the position regularity constraint in Integral Imaging methods. In this framework, the sensors collecting perspective images can be located at randomly distributed coordinates in 3D space. For 3D reconstruction, a finite number of sensors with known coordinates are randomly selected from within the pick up volume. In particular, we build upon Synthetic Aperture Integral Imaging (SAII) where the sensor distribution is not controlled, that is, it is random; however, the locations of sensors in space shall be known. In this study the optical axes of the cameras are assumed parallel but each sensor has a different distances from the 3D object. A computational reconstruction framework based on the back projection method is developed using a variable affine transform between the image space and the object space. It can be shown that affine coordinate transformation corresponds to orthographic projection similar to what is needed in ray back-projection based II reconstruction.

2. RANDOM SENSOR ARCHITECTURE

In most passive multi perspective 3D imaging systems there is a predefined geometry such as planar or concave for collection of different images. The perspective images are captured at the predefined locations. For instance, in a conventional lenslet based II, the elemental images are captured on a regular, planar grid. In this overview paper, a generic framework for 3D multi perspective imaging with randomly or sparsely located sensors is proposed.⁴⁰ As shown in Fig. 1 each sensor is positioned independently and randomly in 3D space with a unique view from the target. The pickup location of the i th elemental image, \mathcal{P}_i , is measured in a universal frame of reference in Cartesian coordinates. It should be noted that the origin of the frame of reference is arbitrary. For mathematical convenience, we set the origin to be at the location of the 0-th perspective image. However, since the proposed mathematical image reconstruction framework merely relies on relative distance of elemental images in space, it stays consistent if the origin moves and all position measurements are adjusted accordingly. Note that a local coordinate system is also defined for each sensor with its origin coinciding the sensor's midpoint.

Let the sensor size be (L_x, L_y) , effective focal length of the i -th imaging optics, g_i , and the position of each sensor from the pickup stage to be known. In our analysis, we make no assumption on the distribution of elemental images in the space to achieve a generic reconstruction scheme. To demonstrate the feasibility of the proposed technique, the random pickup locations, $\mathcal{P}_i = (x_i, y_i, z_i)$, are chosen from three independent uniform random variables. Clearly, the actual distribution of elemental images is dictated by the specific application of interest. We have used uniform distribution to give all locations in the pickup volume an equal chance to get selected as sensor positions. The reference elemental image, i.e. the elemental image from which perspective the reconstruction is desired, is denoted as \mathcal{E}_0 .

3. THREE DIMENSIONAL TARGET RECONSTRUCTION

Different computational methods have been investigated for 3D reconstruction of the scene. In the Fourier domain, digital refocusing has been proposed³⁶ by applying Fourier Slice theorem in 4D light fields. This technique is relatively fast with complexity of $O(n^2 \log n)$, n being the total number of samples. However, this

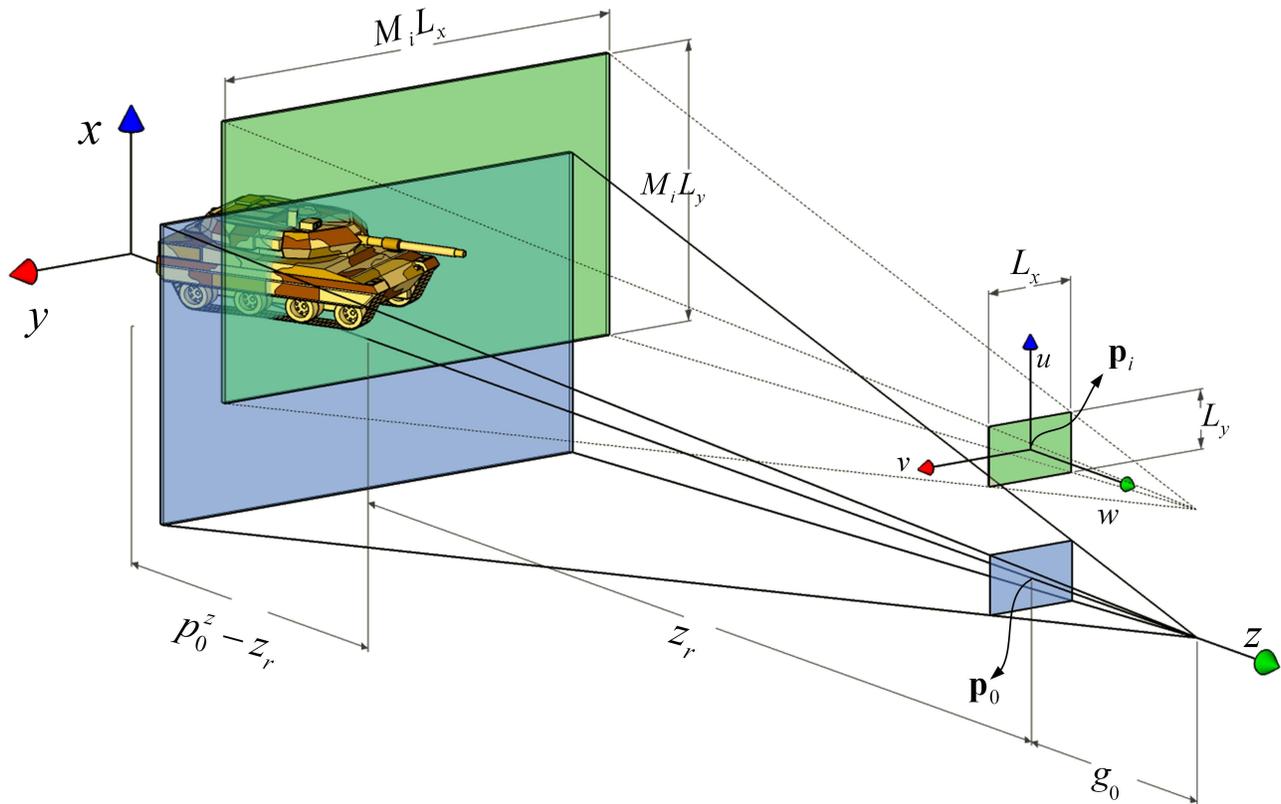


Figure 1. Illustration of two sensors that are randomly positioned to capture the scene. The 0-th (reference) and i th imagers are shown with their field of view in blue and green respectively.

method is intrinsically based on the assumption of periodic sampling of the light field and thus is not applicable to non-ordered randomly distributed sensors. In the spatial domain, a fast, ray tracing based reconstruction from the observers point of view is proposed³⁰ with complexity of $O(m)$, m being the number of elemental images. Although fast and simple, this method yields low resolution reconstructions. Yet another spatial domain reconstruction method is based on series of 2D image back projections.³¹ This method offers a higher reconstruction resolution comparing to³⁰ at the expense of an algorithm with complexity of $O(n)$, since usually $n \gg m$. For instance, m is typically in the range of 100-200 elemental images, while n can be as large as 10^7 pixels. In the context of this paper, we stay within the spatial domain as we tend to provide a generic reconstruction algorithm with minimum assumptions about the pickup geometry.

Computational reconstruction based on back-projection,^{31,32} has certain assumptions which are only valid for lenslet based integral imaging systems. In Ref. 40 we suggest the use of affine transform for orthographic back-projection of perspective images regardless of their pickup locations.

In essence, the relationship between the sensor's local frame of reference and the global frame of reference can be modeled by a linear affine transform which embeds the scaling (magnification) and translation (shifting) that occurs in ray back-projection. With such transform, each voxel in 3D object space can be located on a the sensor at an arbitrary location in space. The constituent parameters of such affine transform are the magnification between the image plane and desired plane of reconstruction as well as the relative position of the sensor to the frame of reference. Note that the choice of frame of reference is arbitrary and can be conveniently chosen to be at the location of one of the sensors.

With the help of affine transform,⁴⁰ the back-projection of each elemental image can be computed at a desired plane of reconstruction with any arbitrary sensor location. Thus, the information collected by all sensors can be

back propagated to the desired plane independent of their relative position, hence removing the constraint on the regular pattern of sensor distribution. Using this technique, any random set of sensors can be chosen and used for reconstruction. Note that the randomness is in the constellation of the sensors and not in their position measurement. However, the effect of uncertainty on the position measurement has also been investigated,³⁹ the results of which reveal that multi-view imaging is rather robust to sensor position measurement error and the degradation in 3D reconstruction is inversely proportional to observation distance.

The suggested reconstruction technique, similar to its counterpart in,³¹ generates high resolution reconstructions; however, it is generalized to deal with elemental images captured in arbitrary locations in space. In addition, to avoid excessive computational burden due to magnification of the elemental images, the relative magnification with respect to difference in distances of sensors to the reconstruction plane has been used.⁴⁰ This approach can speed up the reconstruction process by 2 orders of magnitude.

4. EXPERIMENTAL RESULTS

Experimental results are demonstrated with toy models of a tank and a sports car resembling a 3D scene. The tank can be enclosed in a box of the size $(5 \times 2.5 \times 2 \text{ cm}^3)$ whereas the car model fits in a volume of $4 \times 3 \times 2.5 \text{ cm}^3$. Also, the tank and the car are placed approximately 19cm and 24cm away from the reference elemental image, respectively.

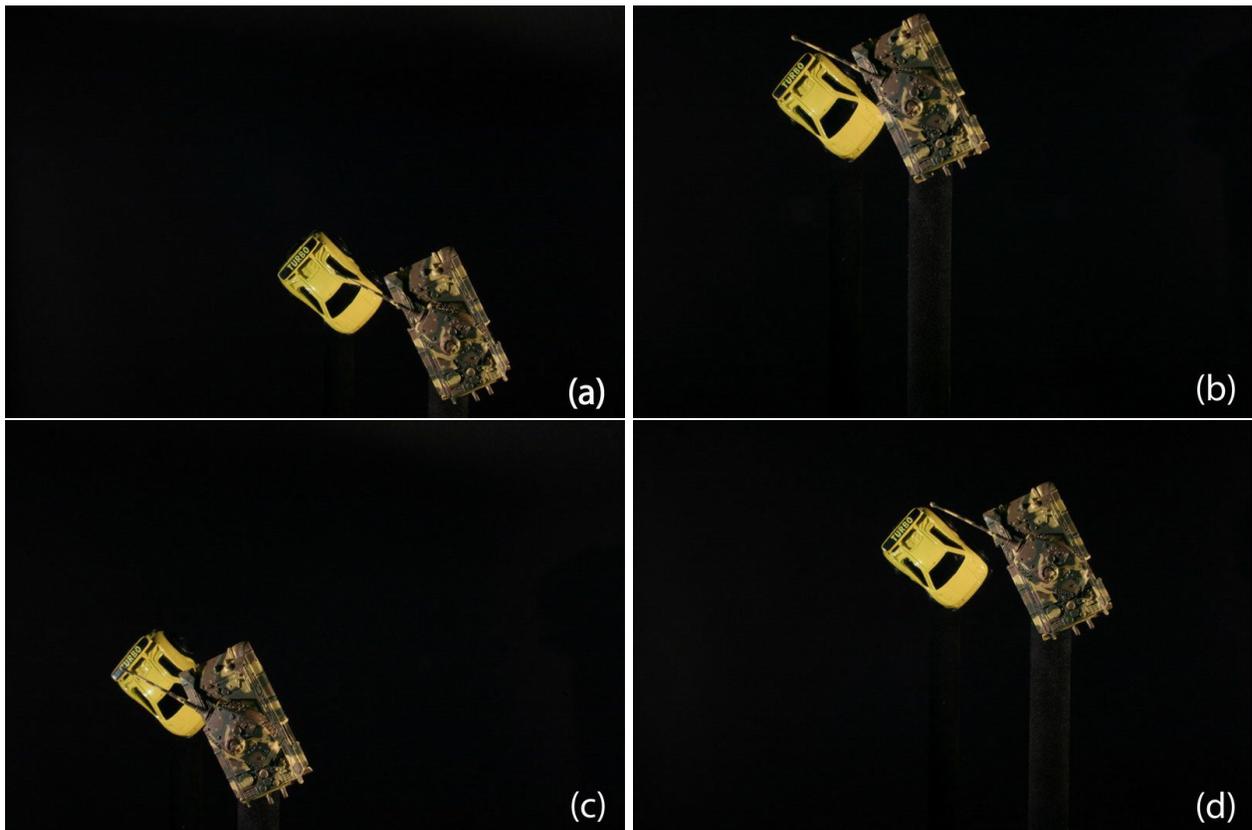


Figure 2. Four elemental images taken from position (a) $\mathcal{P}_1 = (-3.9, -3.9, 26.8)$, (b) $\mathcal{P}_2 = (1.4, 3.4, 25.0)$, (c) $\mathcal{P}_3 = (2.9, -3.2, 25.7)$ and (d) $\mathcal{P}_4 = (-2.9, 2.1, 26.4)$ all in cm .

To obtain random pickup positions, a set of 100 positions, $\mathcal{P}_i = (p_i^x, p_i^y, p_i^z)$, are obtained from three uniform random variable generators. Parallax in x and y is set to $(-4\text{cm}, 4\text{cm})$ and for z to $(25\text{cm}, 27\text{cm})$ assuming the desirable reconstruction range within $[19\text{cm}, 24\text{cm}]$ from the reference sensor. As mentioned in section 2, the choice of reference sensor is arbitrary as it merely sets the origin of the global coordinate system. As all the

measurements are relative the choice of reference sensor would not affect the affine transform back projection procedure described in section 3.

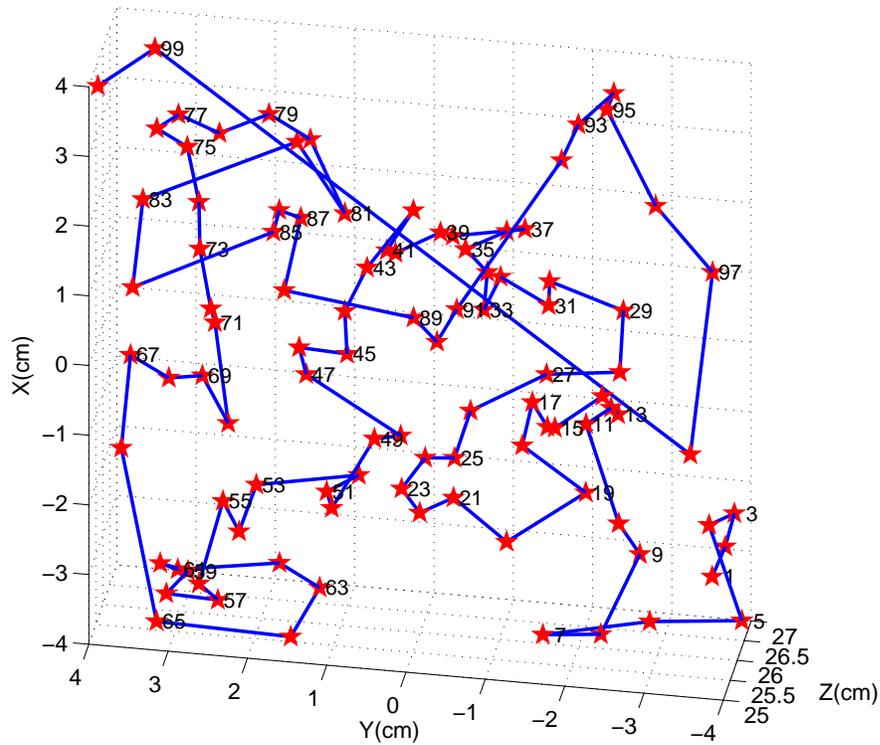


Figure 3. The sparse distribution of the sensors in the pickup volume. The blue line illustrates the trajectory of the sensor.

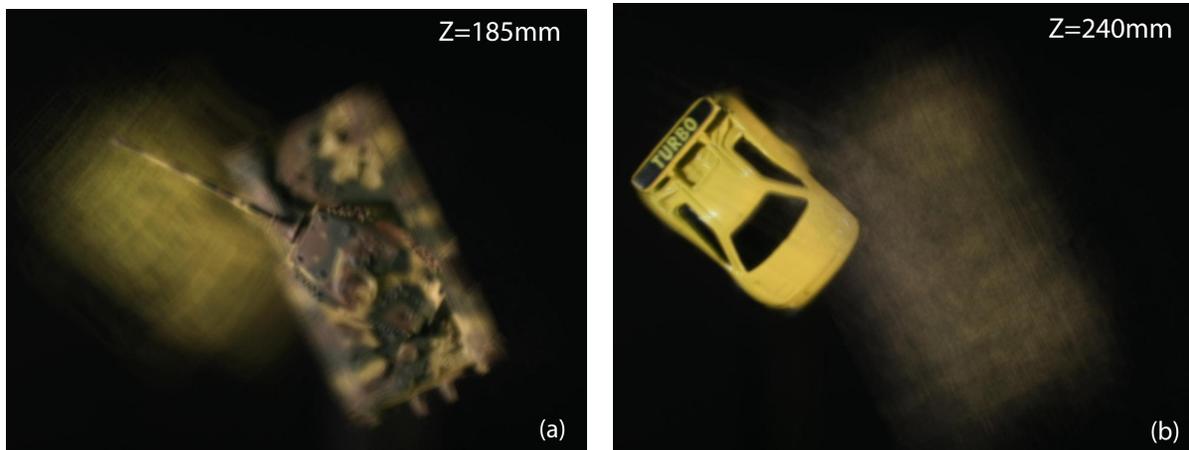


Figure 4. Two reconstructed images from viewpoint $\mathcal{P}_0 = (-0.2, 1.8, 25.9)$ at (a) $z_r = 185mm$, (b) $z_r = 240mm$.

The i th elemental image is then taken with a digital camera at its associated random pickup position, \mathcal{P}_i . The focal length for all lenses are set to be equal, i.e. $g_i = 25mm$, The CMOS sensor size is $22.7 \times 15.1 \text{ mm}^2$ with $7\mu m$ pixel pitch. The field of view (FOV) for each elemental image is then $48^\circ \times 33^\circ$ in the horizontal and

vertical directions respectively, which covers an area of $18 \times 12 \text{ cm}^2$ at 20 cm away from the pickup location in the object space.

A single camera is translated between the image acquisition points such that it only passes each location once while at each location a full frame image with size 3072×2048 pixels is captured. In our experiments, the camera is translated in x, y, z using off the shelf translation components. The trajectory of the sensors are illustrated in Fig. 3.

The 100 perspective images captured for this experiment are used to reconstruct the 3D scene at different distances from the viewpoint of the reference sensor, i.e. the sensor located at $\mathcal{P}_0 = (-0.2, 1.8, 25.9) \text{ cm}$. The reconstitution is done with varying $z_r \in [160 \text{ mm}, 300 \text{ mm}]$. Two of such reconstruction planes are shown in Fig. 4 at $z_r = 185 \text{ mm}$ and $z_r = 240 \text{ mm}$, respectively.

As illustrated in Fig. 4, the objects close to the reconstruction plane appear in focus while objects at a distance from desired reconstruction plane appear blurred. The blur is due to superposition of unmatched rays captured by the elemental images and it increases by disparity between the object and reconstruction planes. A movie of the reconstruction for several planes shows how objects enter and exit the focus region.⁴⁰ It is possible to infer the depth profile of the objects based on the statistical properties of the rays intersecting each point on a reconstruction plane.

5. CONCLUSION

In this paper, we presented a summarized report on our on going research on new architectures for passive 3D multi-view imaging. Three dimensional passive imaging and display is a promising alternative to other forms of 3D imaging systems such LIDAR or holographic imaging. It can provide 3D imaging capabilities in situations where cost, weight and low complexity is important. The 3D information collected by passive sensors can be used to improve ATR performance and may reduce some algorithmic complexity. In the work presented in this paper, a multi-perspective imaging system is generalized to randomly distributed sensor locations. A finite number of sensors are randomly selected from within a volume and used to infer 3D information about the target. While the sensor distribution is random, that is, it is not controlled, the locations of sensors in space are assumed to be known at the reconstruction stage. The computational reconstruction is performed by using variable affine transform to back propagate the recorded 2D information onto the 3D space.

REFERENCES

- [1] Lypton, L., [Foundation of Stereoscopic Cinema, A Study in Depth], Van Nostrand Reinhold, New York, (1982).
- [2] Okoshi, T., [Three-Dimensional Imaging Techniques], Academic Press (1976).
- [3] Perlin, K., Paxia, S. and Kollin, J., S., "An autostereoscopic display," in *Proceedings of the 27th Ann. Conf. on Computer Graphics and Interactive Techniques*, 319–326, ACM Press/Addison-Wesley (2000).
- [4] Okoshi, T., "Three-dimensional displays," *Proceedings of the IEEE* **68**(5), 548-564 (1980).
- [5] Yang L., McCornick, M., Davies, N. "Discussion of the optics of a new 3-D imaging system," *Appl. Opt.* **27**, 4529-4534 (1988).
- [6] Benton, S., Ed., [Selected Papers on 3D Displays], SPIE Press, Bellingham, WA (2001).
- [7] Javidi, B., Hong, S. -H. and Matoba, O. "Multidimensional optical sensor and imaging system," *Appl. Opt.* **45**, 2986–2994 (2006).
- [8] Javidi, B., F. Okano, and J. Son, Eds., [Three Dimensional Imaging, Visualization, and Display Technology], Springer (2008).
- [9] Javidi, B., Ed., [Optical Imaging Sensors and Systems for Homeland Security Applications], Springer, NewYork (2006).
- [10] Yeom, S., Javidi, B. and Watson, E., A., "Photon-counting passive 3D image sensing for reconstruction and recognition of occluded objects," *Opt. Express* **15**, 16189–16195 (2007).
- [11] Frauel, Y., Tajahuerce, E., Matoba, O., Castro, A. and Javidi, B. "Comparison of passive ranging integral imaging and active imaging digital holography for three-dimensional object recognition," *Appl. Opt.* **43**, 452–462 (2004).

- [12] Okano, F., Hoshino, H., Arai, J. and Yuyama, I., "Real-time pickup method for a three-dimensional image based on integral photography," *Appl. Opt.* **36**, 1598-1603 (1997).
- [13] Okano, F., Arai, J., Hoshino, H. and Yuyama, I., "Three-dimensional video system based on integral photography," *Opt. Eng.* **38**, 1072-1078 (1999).
- [14] Martinez-Cuenca, R., Saavedra, G., Martnez-Corral, M. and Javidi, B., "Extended Depth-of-Field 3-D Display and Visualization by Combination of Amplitude-Modulated Microlenses and Deconvolution Tools," *IEEE J. Display Technol.* **1**, 321-327 (2005).
- [15] Stern, A. and Javidi, B. "Three-dimensional image sensing, visualization, and processing using integral imaging," *Proceedings of the IEEE* **94**(3), 591-607 (2006).
- [16] Igarishi, Y., Murata, H. and Ueda, M., "3D display system using a computer-generated integral photograph," *Jpn. J. Appl. Phys.* **17**, 1683-1684 (1978).
- [17] Lippmann, M., G., "La photographie intégrale," [*Comptes-rendus de l'Académie des Sciences*] **146**, 446-451 (1908).
- [18] Sokolov, A., P., [*Autostereoscopy and Integral Photography by Professor Lippmanns Method*] Moscow State Univ. Press, Moscow (1911).
- [19] Ives, H., E., "Optical properties of a Lippmann lenticuled sheet," *J. Opt. Soc. Am.* **21**, 171-176 (1931).
- [20] Son, J.-Y., Saveljev, V., V., Choi, Y.-J., Bahn, J.-E., Kim, S.-K. and Choi, H., "Parameters for designing autostereoscopic imaging systems based on lenticular, parallax barrier, and integral photography plates," *Opt. Eng.* **42**, 3326-3333 (2003).
- [21] McManamon, P., "Laser radar development," 15th Annual AESS/IEEE, Dayton Section Symposium, 53-60 (1998).
- [22] Travis, A., R., L., "The display of three dimensional video images," *Proc. of the IEEE* **85**(11), 1817-1832 (1997).
- [23] Levoy, M., Hanrahan, P., "Light field rendering," in *Proc. of ACM SIGGRAPH*, 31-42, New Orleans, (1996).
- [24] Martinez, R., Pons, A., Saavedra, G., Martinez-Corral, M. and Javidi, B., "Optically-corrected elemental images for undistorted integral image display," *Opt. Express* **14**, 9657-9663 (2006).
- [25] Martinez-Cuenca, R., Saavedra, G., Martnez-Corral, M. and Javidi, B., "Enhanced depth of field integral imaging with sensor resolution constraints," *Opt. Express* **12**, 5237-5242 (2004).
- [26] Castro, A., Frauel, Y. and Javidi, B., "Integral imaging with large depth of field using an asymmetric phase mask," *Opt. Express* **15**, 10266-10273 (2007).
- [27] Arai, J., Okano, F., Hoshino, H. and Yuyama, I. "Gradient-index lens-array method based on real-time integral photography for three-dimensional images," *Appl. Opt.* **37**, 2034-2045 (1998).
- [28] Jang, J., S. and Javidi, B., "Improved viewing resolution of 3-D integral imaging with nonstationary micro-optics," *Opt. Lett.* **27**, 324-326 (2002).
- [29] Jang, J., S. and Javidi, B., "Three-dimensional integral imaging with electronically synthesized lenslet arrays," *Opt. Lett.* **27**, 1767-1769 (2002).
- [30] Arimoto, H., Javidi, B., "Integral three-dimensional imaging with digital reconstruction," *Opt. Lett.* **26**, 157-159 (2001).
- [31] Hong, S. -H., Jang, J., S. and Javidi, B., "Three-dimensional volumetric object reconstruction using computational integral imaging," *Opt. Express* **12**, 483-491 (2004).
- [32] Hong, S. -H. and Javidi, B., "Improved resolution 3D object reconstruction using computational integral imaging with time multiplexing," *Opt. Express* **12**, 4579-4588 (2004).
- [33] Stern, A. and Javidi, B., "Three-dimensional image sensing and reconstruction with time-division multiplexed computational integral imaging," *Appl. Opt.* **42**, 7036-7042 (2003).
- [34] Hwang, Y., Hong, S. -H. and Javidi, B., "Free View 3-D Visualization of Occluded Objects by Using Computational Synthetic Aperture Integral Imaging," *IEEE J. Display Technol.* **3**, 64-70 (2007).
- [35] Saavedra, G., Martinez, R., Martinez, M., Navarro, H., Daneshpanah, M. and Javidi, B., "Digital slicing of 3D scenes by Fourier filtering of integral images," *Opt. Express* **16**, 17154-17160 (2008).
- [36] Ng, R., "Fourier slice photography," in *Proceedings of ACM SIGGRAPH*, **24**, 735-744 (2005).
- [37] Jang, J., S. and Javidi, B., "Three-dimensional synthetic aperture integral imaging," *Opt. Lett.* **27**, 1144-1146 (2002).

- [38] Levoy, M., "Light fields and computational imaging," *IEEE Computer* **39**, 46–55 (2006).
- [39] Tavakoli, B., DaneshPanah, M., Javidi, B. and Watson E., A., "Performance of 3D integral imaging with position uncertainty," *Opt. Express* **15**, 11889–11902 (2007).
- [40] DaneshPanah, M., Javidi, B. and Watson E., A., "Three dimensional imaging with randomly distributed sensors," *Opt. Express* **16**, 6368-6377 (2008)
- [41] –, "High-speed science," *Nature Physics*, **4**(6), 435 (2008)