Three-dimensional imaging with detector arrays on arbitrarily shaped surfaces

Mehdi Daneshpanah and Bahram Javidi*

University of Connecticut, 371 Fairfield Way, Storrs, Connecticut 06269, USA
**Corresponding author: bahram@engr.uconn.edu*

Received October 8, 2010; revised January 14, 2011; accepted January 15, 2011;
posted January 21, 2011 (Doc. ID 135725); published February 17, 2011

Conventional multiview imaging systems commonly utilize a planar lenslet array and a rigid, flat image sensor in the pickup stage to capture different views of the scene. In this Letter, we remove this constraint by proposing the concept of three-dimensional (3D) imaging with detector arrays that may conform to arbitrarily shaped surfaces or platforms. A nonplanar detector array configuration can be used in combination with a flexible lenslet array to capture different views. The orientation and optical axes of individual image sensing elements could vary. A point-by-point 3D reconstruction algorithm is described and the feasibility of the proposed approach is demonstrated through simulated imagery. © 2011 Optical Society of America

OCIS codes: 110.6880, 110.3010.

The century old integral photography idea [1] is being pursued in its modern form for three-dimensional (3D) imaging and display systems [2–8]. Among them, integral imaging (II) systems use a planar sheet of lenslet array to capture different two-dimensional (2D) views (elemental images) of the scene on an image sensor. The ensemble of elemental images encode the depth of each object point in form of disparity. The aperture size, focal length, pixel size, field of view, parallax, and relative arrangement of elemental sensors are interrelated parameters that involve design trade-offs [4,8]. A number of these trade-offs arise from the fact that current 3D systems are designed around the flat, rigid image detectors that are widely available. As an example, the common field of view of all elemental sensors has an inverse relationship with parallax. A large parallax is desirable to improve the 3D capture capability of the system. As a result, maintaining a large field of view with large parallax poses a challenge in II with flat image sensors [8].

However, with advancements in flexible electronics, both organic and inorganic material can be integrated on elastomeric substrates to create deformable electronic devices [9,10]. Several successful demonstrations of nonplanar detector arrays have already been undertaken, inspired by the most ubiquitous visual sensor, retina [11], as well as optical advantages provided by such detector arrays for compact optical design [12]. Such advanced detector arrays can be leveraged for 3D imaging to facilitate novel designs. Some conceivable applications include adding visual 3D sensing capability to endoscopic medical instruments, vehicles' bodies, advanced 3D borescopes, soldiers' helmets, or aircraft wings with minimal structural constraints.

In this Letter, we propose the concept of 3D imaging with nonplanar detector arrays that may conform to arbitrarily shaped surfaces or platforms, as shown schematically in Fig. 1. An array of lenslets embedded in an elastic scaffold can be integrated with a flexible optoelectronic detector array in an arbitrary nonplanar configuration. The resulting image sensor can conform to curved surfaces, with each elemental sensor having a unique optical axis. To eliminate the cross talk between images of neighboring lenslets, a matching parallax barrier array can be mounted in between the lenslet array and the image detector array.

Conventional II reconstruction algorithms fail to reconstruct the scene from imagery acquired by such an arbitrary sensor configuration [4,8]. Here, a generalized framework is developed that is based on a point-by-point computational image reconstruction using the ray intensity-angle information that is captured by the proposed sensor. The feasibility of the proposed 3D imaging scheme is demonstrated through simulated imagery with realistic device parameters.

Consider a Cartesian coordinate system as the global frame of reference described by unit vectors $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ and $\hat{\mathbf{z}}$, as shown in Fig. 1. Each elemental sensor in a flexible sensor array device can be characterized by the center position of its aperture, $\mathbf{p}_i$, as well as two orthogonal unit vectors, $\hat{\mathbf{u}}_i$ and $\hat{\mathbf{v}}_i$, in the horizontal and vertical directions of the detector array, as shown in Fig. 1. The ith unit normal vector is thus computed by $\hat{\mathbf{n}}_i = \hat{\mathbf{u}}_i \times \hat{\mathbf{v}}_i$, such that $(\hat{\mathbf{u}}_i, \hat{\mathbf{v}}_i, \hat{\mathbf{n}}_i)$ forms a right-hand local coordinate system for each elemental sensing unit. The normal Euclidean distance of a point in object space, $\mathbf{r}_o$, from the principle point of the ith lenslet can be written as $|(\mathbf{p}_i - \mathbf{r}_o).\hat{\mathbf{n}}_i|$. The magnification associated with area in the vicinity of point $\mathbf{r}_o$ with respect to the ith lenslet can thus be written as

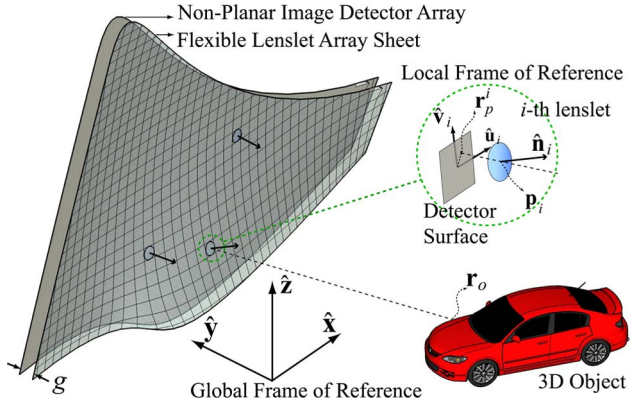

Fig. 1. (Color online) Proposed 3D image sensor, which is composed of many arbitrarily arranged elemental sensors. Inset shows the ith elemental sensor notations.

$$M_i(\mathbf{r}_o) = |(\mathbf{p}_i - \mathbf{r}_o).\hat{\mathbf{n}}_i|/g, \tag{1}$$

where g is the gap between the lenslet and the image detector. The detector area behind each lenslet is assumed to have a curvature within the depth of focus of the lenslet. The depth of focus of each lenslet can be written as $d_f \simeq 2g\mu/D$, within which the defocus is contained to pixel size. The radius of curvature, r, of the detector array behind each lenslet should not exceed the limit set by depth of focus, i.e.,

$$r \geq (D^2 + 4d_f^2)/2d_f. \tag{2}$$

Except for extremely sharp corners, this condition is not hard to satisfy in most practical applications.

Now, let the projection of an object point, $\mathbf{r}_o$, on the ith elemental sensor be denoted as $\mathbf{r}_p^i$ (see Fig. 1). Since point $\mathbf{p}_i$ is, in fact, parting line segment $\mathbf{r}_o\mathbf{r}_p^i$ by the magnification ratio $1{:}M_i(\mathbf{r}_o)$, the projected point $\mathbf{r}_p^i$ coordinates can be written as

$$\mathbf{r}_p^i = \frac{1 + M_i(\mathbf{r}_o)}{M_i(\mathbf{r}_o)}\mathbf{p}_i - \frac{1}{M_i(\mathbf{r}_o)}\mathbf{r}_o. \tag{3}$$

Given unit vectors $(\hat{\mathbf{u}}_i, \hat{\mathbf{v}}_i)$, it is straightforward to convert the position $\mathbf{r}_p^i$ in $\mathbb{R}^3$ (3D real space) to the local pixel location index, $[u_i, v_i]^{\mathrm{T}}$, in $\mathbb{C}_i^2$ (2D discrete space) on the detector array as

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} = \frac{1}{\mu}\begin{bmatrix} (\mathbf{r}_p^i - \mathbf{C}_i).\hat{\mathbf{u}}_i \\ (\mathbf{r}_p^i - \mathbf{C}_i).\hat{\mathbf{v}}_i \end{bmatrix}, \tag{4}$$

where μ is the pixel size and $\mathbf{C}_i$ is the midpoint of the area associated with the ith elemental sensor. Coordinates of this point can be computed by

$$\mathbf{C}_i = \mathbf{p}_i - g\hat{\mathbf{n}}_i. \tag{5}$$

Substituting Eqs. (3) and (5) in Eq. (4), one can find a unique transformation for the ith elemental image, $\mathcal{T}_i(.)$, which maps the object space points, $\mathbf{r}_o$, onto the corresponding pixel index of the ith elemental sensor, i.e., $\mathcal{T}_i{:}\mathbb{R}^3 \mapsto \mathbb{C}_i^2$, as

$$\mathcal{T}_i(\mathbf{r}_o) = \begin{bmatrix} u_i \\ v_i \end{bmatrix} = \frac{1}{\mu M_i(\mathbf{r}_o)}\begin{bmatrix} (\mathbf{p}_i - \mathbf{r}_o).\hat{\mathbf{u}}_i \\ (\mathbf{p}_i - \mathbf{r}_o).\hat{\mathbf{v}}_i \end{bmatrix}. \tag{6}$$

Equation (6) provides a map between the object space, $\mathbb{R}^3$, and the discrete ith elemental sensor pixel index, $\mathbb{C}_i^2$. Such a map represents a perspective transformation between the two spaces and, as such, neither preserves the distances nor the angles. With the aid of Eq. (6), one can find the corresponding pixel to a particular object point on all elemental images, each represented by $\mathcal{E}_i$. For objects with Lambertian (diffusively scattering) surfaces, the intensity, $\hat{I}$, of each object point can be approximated by a weighted average of the corresponding pixels on all elemental sensors as

$$\hat{I}(\mathbf{r}_o) = N^{-1}(\mathbf{r}_o)\sum_{i=1}^{K} w_i(\mathbf{r}_o)\alpha_i(\mathbf{r}_o)\mathcal{E}_i(\mathcal{T}_i(\mathbf{r}_o)), \tag{7}$$

where $w_i(\mathbf{r}_o) = 1$, if object point $\mathbf{r}_o$ falls within the field of view of the ith lenslet and zero, otherwise

$$w_i(\mathbf{r}_o) = \begin{cases} 1 & \frac{(\mathbf{p}_i - \mathbf{r}_o).\hat{\mathbf{n}}_i}{|\mathbf{p}_i - \mathbf{r}_o|} \leq \cos(\theta_i) \\ 0 & \text{otherwise} \end{cases}, \tag{8}$$

where $\theta_i = \tan^{-1}(D_i/2g)$ is the half-angle determining the field of view of the ith lenslet with aperture diameter D_i.

Note that $N(\mathbf{r}_o) = \sum_{i=1}^{K} w_i(\mathbf{r}_o)$ is the total number of lenslets that have captured object point $\mathbf{r}_o$. Also, $\alpha_i(\mathbf{r}_o) = \rho_i^2(\mathbf{r}_o)/D_i$ compensates for variations in object point distance, $\rho_i(\mathbf{r}_o)$, and light collection efficiency due to aperture size, D_i.

Equation (7) provides a systematic way of reconstructing object space through a set of arbitrary points that can be chosen arbitrarily in a volume or on a surface or a line. In the special case that the chosen points lie on a grid representing a planar surface, the reconstructed images are similar to digital reconstruction in the II technique [8].

The density of reconstructed points is arbitrary. However, one should note that geometrical and wave optics requires that the angular separation of two reconstruction points, δ, to be larger than the diffraction limit or the geometrical resolution imposed by sensor pixelation, i.e., $\delta \geq \min(1.22\lambda/D, \mu/g)$.

To demonstrate feasibility, a 4000×4000 pixel nonplanar detector array with pixel size $\mu = 5\,\mu$m (100% fill factor) is modeled by a computer with the profile

$$z = 7\sin(0.4x) + 0.5\cos(0.3y), \qquad 10 \leq x, \qquad y \leq 10, \tag{9}$$

in which dimensions are in millimeters. x, y, and z denote the coordinate system parameters. The image forming optics is an array of 20×20 circular aperture lenslets with a diameter of $D = 1$ mm mounted at $g = 3$ mm from the detector, resulting in $\theta = 9.4°$. There are approximately 200×200 pixels associated with each elemental

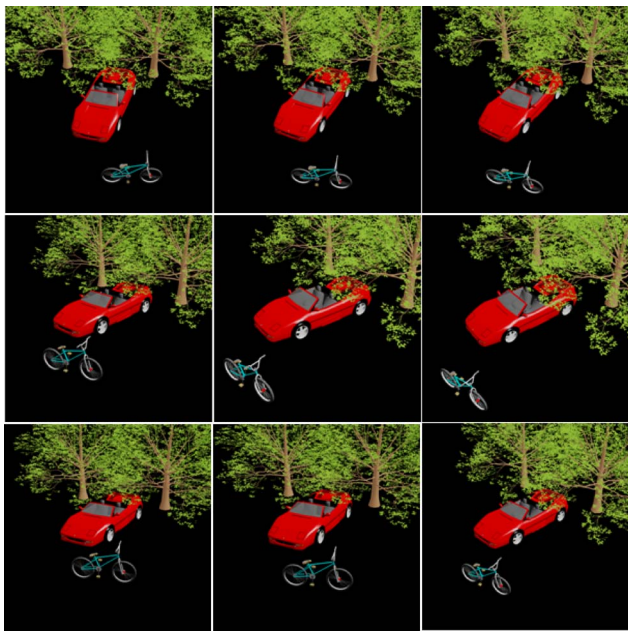

Fig. 2. (Color online) Subset of nine elemental images (out of 400) that capture different views of the objects in the scene.

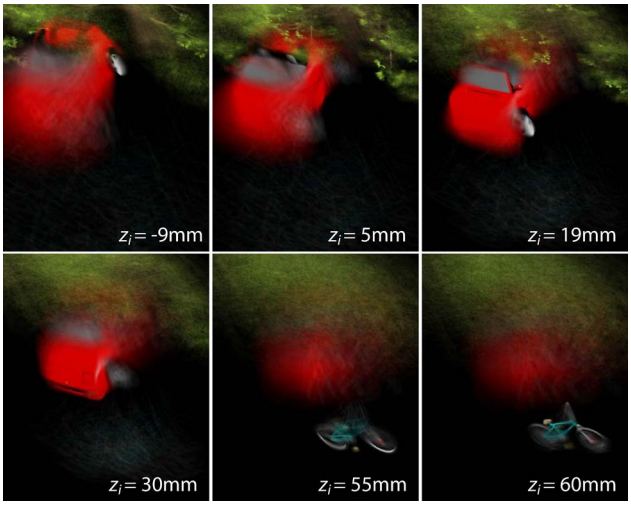

Fig. 3. (Color online) Reconstruction planes with $\hat{\mathbf{n}} = (1, 1, 1)$ at $z_j = -9$ mm, 5 mm, 19 mm, 30 mm, 55 mm, and 60 mm, where the car rear wheel, the passenger seat, the side mirror, the front of the car, the rear bicycle wheel, and the bicycle body are in focus, respectively.

image. This imaging configuration allows for mounting the sensor on a ball with a radius equal to $r_{\max} = 4.2$ mm, according to Eq. (2).

Elemental images are simulated by 3D rendering of a scene containing a model car, vegetation, and a bicycle. The length of the model car is 5 cm. The sensor is placed 30 cm from the objects. Figure 2 shows a subset of captured elemental images.

The reconstruction surface is generated to be planar with a normal vector of $\hat{\mathbf{n}}_r = (1, 1, 1)$ in the global frame of reference. Using Eq. (7), six reconstruction planes are generated, as shown in Fig. 3, in each panel of which a part of the scene is in focus. The reconstruction planes can be written mathematically as $\hat{\mathbf{n}}_r.\mathbf{r}_o = -z_j$, where z_j is the distance of the jth plane to the origin.

The proposed image reconstruction framework allows for desired coarse or fine reconstruction on arbitrary 2D manifolds, e.g., arbitrarily shaped surfaces. The algorithm complexity is only linearly proportional to the number of object space points being reconstructed. This approach can be extended to the sparse aperture 3D imaging case, where information from individual sensors in different locations and/or with different intrinsic parameters, e.g., focal length, field of view, or pixel size, can be combined to computationally reconstruct the scene in 3D.

In addition, many common computational tasks, such as object recognition, compression, or occlusion removal, can be applied to the scene under survey in the 3D domain. An important advantage of the proposed reconstruction framework is that the user is no longer limited to reconstructing the field on a plane parallel to the lenslet array (perpendicular to optical axis). One has the flexibility to reconstruct the field on a set of arbitrarily shaped planes or a point cloud with desired density, such as to maximize an objective metric in the task at hand. This extra degree of flexibility can be utilized to improve the effectiveness of such algorithms.

References

1. M. G. Lippmann, C.R. Hebd. Seances Acad. Sci. **146**, 446 (1908).
2. Y. Igarishi, H. Murata, and M. Ueda, Jpn. J. Appl. Phys. **17**, 1683 (1978).
3. F. Okano, J. Arai, K. Mitani, and M. Okui, Proc. IEEE **94**, 490 (2006).
4. R. Martinez-Cuenca, G. Saavedra, M. Martinez-Corral, and B. Javidi, Proc. IEEE **97**, 1067 (2009).
5. S. Maehara, K. Nitta, and O. Matoba, Appl. Opt. **47**, 594 (2008).
6. M. Daneshpanah, B. Javidi, and E. A. Watson, Opt. Express **16**, 6368 (2008).
7. M. Pollefeys, R. Koch, M. Vergauwen, A. Deknuydt, and L. Van-Gool, Proc. SPIE, **3958**, 215 (2000).
8. M. Cho, M. Daneshpanah, I. Moon, and B. Javidi, Proc. IEEE **PP**1 (2010).
9. R. Dinyari, S.-B. Rim, K. Huang, P. B. Catrysse, and P. Peumans, Appl. Phys. Lett. **92**, 091114 (2008).
10. J. A. Rogers, T. Someya, and Y. Huang, Science **327**, 1603 (2010).
11. X. Xu, M. Davanco, X. Qi, and S. R. Forrest, Org. Electron. **9**, 1122 (2008).
12. S.-B. Rim, P. B. Catrysse, R. Dinyari, K. Huang, and P. Peumans, Opt. Express **16**, 4965 (2008).