

Integral-Imaging display from stereo-Kinect capture

Seokmin Hong, Amir Ansari, Genaro Saavedra, and Manuel Martinez-Corral

Department of Optics, University of Valencia, E-46100 Burjassot, Spain.

Seokmin.Hong@uv.es, Amir.Ansari@uv.es, Genaro.Saavedra@uv.es, Manuel.Martinez@uv.es

ABSTRACT

In this paper, we propose a new approach in order to improve the quality of microimages and display them onto an integral imaging monitor. Our main proposal is based on the stereo-hybrid 3D camera system. Originally, hybrid camera system has dissimilarity itself. We interpret our method in order to equalize the hybrid sensor's characteristics and 3D data modification strategy. We generate integral image by using synthetic back-projection mapping method. Finally, we project the integral image onto our proposed display system. We illustrate this procedure with some imaging experiments in order to prove an advantage of our approach.

Key words: 3D display, Integral Imaging, 3D data registration, Color transfer, Point clouds, Hybrid 3D cameras

Conventional photography is fully adapted to record a three-dimensional (3D) world's scene into a two-dimensional (2D) sensing device. Although 2D images can reflect the 3D nature of scenes, they still lack of important information. Fortunately, there are techniques that are able to record 3D information from 3D scenes. Among them, integral-imaging (InI) technique has been considered as one of the most potential technologies to record and display real world scenes. The main procedure of InI is performed by placing a microlens array in front of a 2D image sensing unit. This lens array records different perspectives of the 3D scene. This is because all of the reflected light from an object is transmitted into all the lenses, which distribute the light on different pixels of the 2D sensor depending on the incidence angle of the light. Here, we name as microimage which is recorded by any microlens. The whole array of microimages is named here as the integral image. When the integral picture is projected onto an InI display system, observers can see the 3D floating scene, which have full-parallax and quasicontinuous perspective views [1-3]. Many researchers and companies have applied this InI technique in many different fields [4-12].

In the meantime, many kind of depth sensing techniques have been developed to record 3D scenes [13-16]. Among them, infrared (IR) light sensing has been widely used during the last decades. Especially the Kinect device from Microsoft uses IR lighting technology for depth acquisition. Nowadays, two different versions of Kinect are released. It is well known that both sensors have many different features for obtaining a depth map. The Kinect v1 uses a structured IR light pattern emitter and IR camera to measure the depth distance from captured features at the scene [13-14]. In comparison, the Kinect v2 utilizes time-of-flight (ToF) technology, which consists of emitting IR flashes with high frequency. The IR light is reflected on most 3D surfaces and detected by a depth sensor on the Kinect v2 device. The depth distance is measured by the light's returning duration. [15-16].

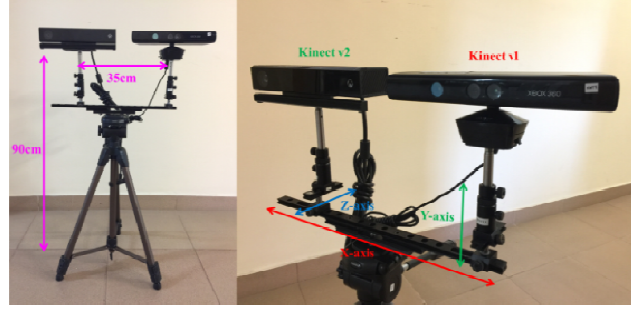


Figure 1 Proposed stereo-hybrid 3D camera system.

Since a few years ago, we have been utilizing a depth camera to capture 3D natural scenes and generate microimages by using synthetic back-projection method. These generated microimages can be projected onto an InI display system for 3D reconstruction [9-10]. But, even this innovative approach has several issues that must be solve. Among them, the depth-hole problem is critical. This is because monocular cameras can only detect the closest object from their sight. It means that occluded part from the sight of the camera has no information and as a result, that area forms a black shadow. Furthermore, there are also noisy patterns or depthless areas that are appearing at the captured depth map because of the limitation of IR light sensing technique. In order to solve these drawbacks, we utilize the stereo-hybrid 3D camera system. Fig. 1 shows our camera set up. In our stereo vision system, an extra perspective view is able to capture the occluded parts and, therefore, the hybrid system can complement each other to solve the lack of depth information.

However, hybrid 3D camera system has many different features itself. In our experiment, we attempt to equalize the hybrid sensor's characteristics and refer to 3D data modification method. First, we equalize the field-of-view (FOV) of cameras and the scale factors from one to another. In the reference paper [17], they used stereo-hybrid 2D camera and equalize both sensors' FOV. Eq.1 shows how to derive scale factors between two different sensors: $j_{u,v}$ is input image (pixel), $k_{u,v}$ is final corrected image (pixel), f, f' is Kinect v1 and v2's focal length (mm), $p_{u,v}, p'_{u,v}$ is pixels per 1mm at the imaging sensor (pixel), and $\sigma_{u,v}$ is scale factor from to another sensors respectively.

$$\begin{cases} k_u = \frac{j_u f' p'_u}{f p_u} = \sigma_u j_u \\ k_v = \frac{j_v f' p'_v}{f p_v} = \sigma_v j_v \end{cases} \quad (1)$$

Second, from the captured RGB and Depth information obtained from each device we are able to compose point clouds and dispose them into a virtual 3D space. However, both clouds data are still placed in individually, and not arranged properly. In order to make registration between them, we utilize Iterative-Closest-Point (ICP) algorithm. ICP algorithm calculates the movement between two sets of point clouds in order to minimize their distance [18]. And third, we adjust color transfer method in order to solve the lack of dissimilarity in color image. We referenced [19] and they analyzed both, input and target image's characteristics and processed color correction from input to target image. Fig. 2 and 3 shows its progressed result. And Media 1 shows before and after all of our progressed result clearly.

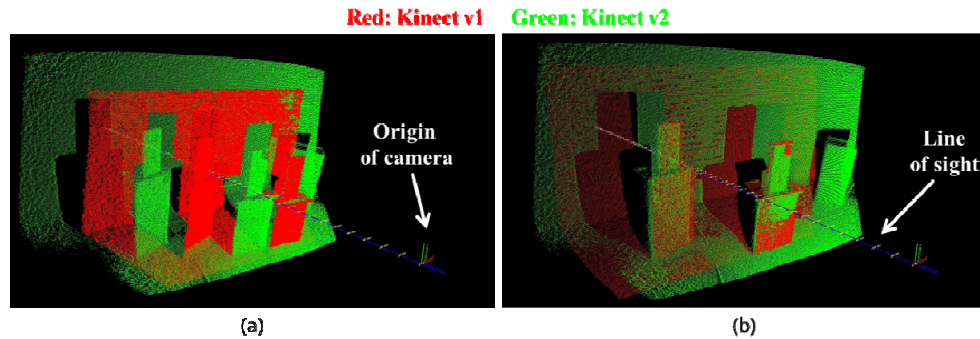


Figure 2 3D point clouds registration result: (a) is raw arranged 3D data, and (b) is the registration result. In this scene, red points are rescaled 3D information of Kinect v1, and green points are from Kinect v2.

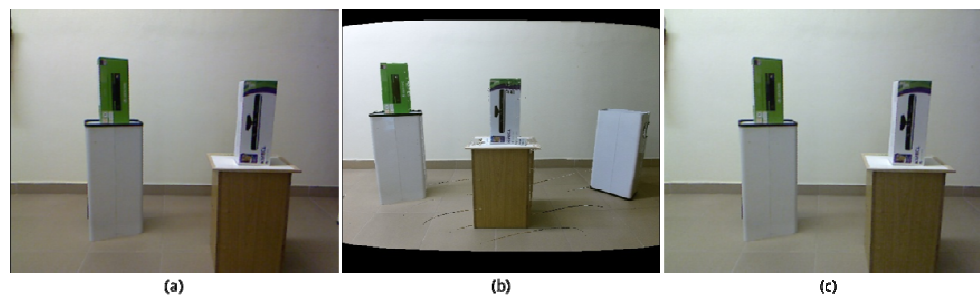


Figure 3 Processed result by using referenced color transfer algorithm: (a) rescaled RGB image captured by Kinect v1, (b) is RGB to IR camera mapping result in Kinect v2, and (c) is color transferred result from (a) to (b). As you can see the figure (c), (a)'s color tone is composed as similar as (b).

From now, this modified point clouds data sets is possible to generate microimages. In order to generate microimages we followed our previous approach [9]. We placed the virtual pinhole array (VPA) at a certain distance from the point clouds. Fortunately, depth sensors can measure the real depth distance precisely and as a result, VPA's location reflects the correlation between real scene and synthetic scene. Moreover, this position decides the front and rear volumes at the displayed 3D scene. Due to this condition, we name this position as reference plane too. Then, the pixels of each point cloud are projected through the VPA and compose the integral image, as in [20].

In our experiment, we built an InI monitor using a Samsung SM-T700 (359 pixels/inch) device, and a MLA consisting of 181×113 lenslets of focal length $f = 3.3\text{mm}$ and pitch $p = 1.0\text{ mm}$ (Model 630 from Fresnel Technology). Each microimage is composed by 15×15 pixels, the gap between the microlenses and the display is fixed to $g = 49.5$ pixels, and thus, the full size of the integral image is 2715×1695 pixels (14.13 pixels/mm). After mounting and aligning the MLA in front of the tablet, the displayed scene is shown in 3D with full-parallax. Fig. 4 shows our experimental system's overview and Fig. 5 is the comparison result between Kinect v2's scene and our proposed method's result. The Fig 5 shows the advantage of our approach. We filled in several depth-hole areas and derived to smoother texture at the scene. Above all, some occluded area is recovered precisely.

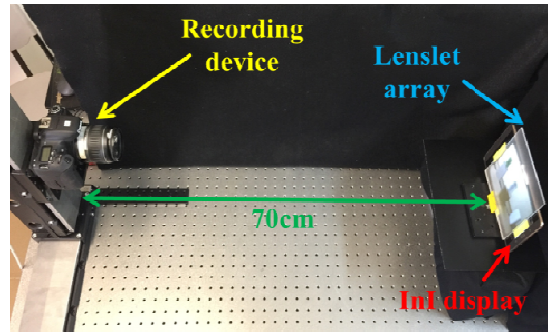


Figure 4 Overview of experimental system.

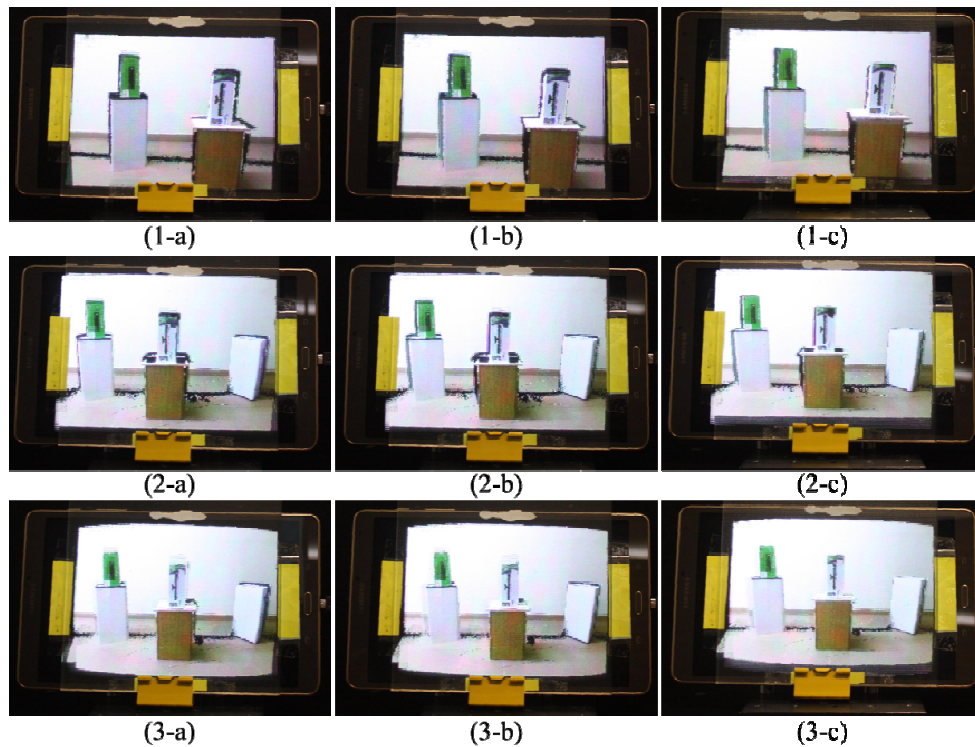


Figure 5 Comparison result between displayed integral image: (1-a, b, c) is Kinect v1, (2-a, b, c) is Kinect v2, and (3-a, b, c) is our proposed result. (a, b, c) shows different perspective position: (a) is left, (b) is right, and (c) is right-bottom perspective view in front of the InI monitor. Media 2, 3, and 4 shows its result well.

To summarize, we have reported an innovative approach in order to enhance the 3D image's quality. We propose a stereo-hybrid 3D camera system that overcomes the limitation of a mono perspective view. We captured one pair of 3D scene and modified it by using our strategy in order to get over the dissimilarity of hybrid sensors. The main advantage of our proposal is that the quality of 3D data is quite improved. Due to this combination, we filled in several depth-hole areas at the integral image and derived to abundant 3D data set. Above all, some occluded field is recovered precisely and thus, this output proves the benefit of our manipulation.

ACKNOWLEDGEMENTS

This work was supported by the Ministerio de Economía y Competitividad, Spain (Grant No. DPI 2015-66458-C2-1-R), and by Generalitat Valenciana, Spain (project PROMETEOII/2014/072). S. Hong acknowledges a predoctoral grant from University of Valencia (UV-INVPREDOC15-265754). And A. Ansari acknowledges a predoctoral contract from EU H2020 program under MSCA grant 676401.

REFERENCES

- [1] G. Lippmann, "Epreuves réversibles photographies intégrales", *Comptes Rendus de l'Académie des Sciences*, 146, 446- 451 (1908).
- [2] F. Okano, H. Hoshino, J. Arai, and I. Yuyama, "Real-time pickup method for a three-dimensional image based on integral photography", *Appl. Opt.* 36, 1598–1603 (1997).
- [3] S. Park, J. Yeom, Y. Jeong, N. Chen, J. Hong, and B. Lee, "Recent issues on integral imaging and its applications", *J. of info. Disp.* 15 (2014).
- [4] "Raytrix camera", <http://www.raytrix.de> (2010).
- [5] "Lytro camera", <https://www.lytro.com> (2011).
- [6] "PiCam: Pelican Imaging Camera", <http://www.pelicanimaging.com> (2013).
- [7] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera", *Tech. Rep. CSTR. 2*, vol. 2 (2005).
- [8] S. Hong, D. Shin, J. Lee, and B. Lee, "Viewing Angle-Improved 3D Integral Imaging Display with Eye Tracking Sensor" *J. Inf. Commun. Converg. Eng.* 12, 208-214 (2014).
- [9] S. Hong, D. Shin, B. Lee, A. Dorado, G. Saavedra, and M. Martinez-Corral, "Towards 3D television through fusion of kinect and integral-imaging concepts", *J. Display Technol.* 11, 894-899 (2015).
- [10] S. Hong, G. Saavedra, and M. Martinez-Corral, "Full parallax three-dimensional display from Kinect v1 and v2", *Optics Eng.* 56(4), 041305 (2016).
- [11] J. Zhang, X. Wang, Q. Zhang, Y. Chen, J. Yu, and Y. Liu, "Integral imaging display for natural scene based on KinectFusion", *Optik* 127, 791–794 (2016).
- [12] Y. Han, M. Lee, and B. Lee, "Air-touch interaction system for integral imaging 3D display", *First Intl. Workshop on Patt. Recog.*, 10011 (2016).
- [13] "Kinect for Windows Sensor Components and Specifications", <http://msdn.microsoft.com/en-us/library/jj131033.aspx> (2013).
- [14] "Primesense Calmine 1.08 & 1.09", <http://en.wikipedia.org/wiki/PrimeSense> (2013).
- [15] "Kinect for Xbox One Components and Specifications", <http://dev.windows.com/en-us/kinect/hardware> (2013).
- [16] "Intel Realsense", <https://software.intel.com/en-us/realsense> (2014).
- [17] H. Shin, S. Kim, and K. Sohn, "Hybrid Stereoscopic Camera System", *J. Broadcast Eng.* 16, 602-613 (2011).
- [18] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes", *IEEE transaction on Pattern Analysis and Machine Intelligence*, 14, 239-256 (1992).
- [19] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley, "Color Transfer between Images", *IEEE Computer Graphics and Applications*, 21, 34-41 (2001).
- [20] M. Martinez-Corral, B. Javidi, R. Martinez-Cuenca, and G. Saavedra, "Formation of real, orthoscopic integral images by smart pixel mapping", *Opt. Express* 13, 9175–9180 (2005).