

Reconstruction of three-dimensional occluded object using optical flow and triangular mesh reconstruction in integral imaging

Jae-Hyun Jung,¹ Keehoon Hong,¹ Gilbae Park,¹ Indeok Chung,¹ Jae-Hyeung Park,²
and ByoungHo Lee^{1,*}

¹School of Electrical Engineering, Seoul National University, Gwanak-Gu Gwanakro 599, Seoul 151-744, Korea

²School of Electrical & Computer Engineering, Chungbuk National University, 410 SungBong-ro, Heungduk-gu, Cheongju, Chungbuk 361-763, Korea

*byoungHo@snu.ac.kr

Abstract: We proposed a reconstruction method for the occluded region of three-dimensional (3D) object using the depth extraction based on the optical flow and triangular mesh reconstruction in integral imaging. The depth information of sub-images from the acquired elemental image set is extracted using the optical flow with sub-pixel accuracy, which alleviates the depth quantization problem. The extracted depth maps of sub-image array are segmented by the depth threshold from the histogram based segmentation, which is represented as the point clouds. The point clouds are projected to the viewpoint of center sub-image and reconstructed by the triangular mesh reconstruction. The experimental results support the validity of the proposed method with high accuracy of peak signal-to-noise ratio and normalized cross-correlation in 3D image recognition.

©2010 Optical Society of America

OCIS codes: (110.2990) Image formation theory; (100.6890) Three-dimensional image processing.

References and links

1. A. Kubota, A. Smolic, M. Magnor, M. Tanimoto, T. Chen, and C. Zhang, "Multiview imaging and 3DTV," *IEEE Signal Process. Mag.* **24**(6), 10–21 (2007).
2. P. Benzie, J. Watson, P. Surman, I. Rakkolainen, K. Hopf, H. Urey, V. Sainov, and C. von Kopylow, "A survey of 3DTV displays: techniques and technologies," *IEEE Trans. Circ. Syst. Video Tech.* **17**(11), 1647–1658 (2007).
3. M. Okutomi, and T. Kanade, "A multiple-baseline stereo," *IEEE Trans. Pattern Anal. Mach. Intell.* **15**(4), 353–363 (1993).
4. B. Lee, J.-H. Park, and S.-W. Min, *Digital Holography and Three-Dimensional Display*, T.-C. Poon, ed. (Springer US, 2006), Chap. 12.
5. G. Lippmann, "La photographie intégrale," *C. R. Acad. Sci. Ser. IIC Chim.* **146**, 446–451 (1908).
6. 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**(7), 1598–1603 (1997).
7. J.-H. Park, S. Jung, H. Choi, Y. Kim, and B. Lee, "Depth extraction by use of a rectangular lens array and one-dimensional elemental image modification," *Appl. Opt.* **43**(25), 4882–4895 (2004).
8. M. Martínez-Corral, B. Javidi, R. Martínez-Cuenca, and G. Saavedra, "Formation of real, orthoscopic integral images by smart pixel mapping," *Opt. Express* **13**(23), 9175–9180 (2005).
9. G. Passalis, N. Sgouros, S. Athineos, and T. Theoharis, "Enhanced reconstruction of three-dimensional shape and texture from integral photography images," *Appl. Opt.* **46**(22), 5311–5320 (2007).
10. D.-H. Shin, and E.-S. Kim, "Computational integral imaging reconstruction of 3D object using a depth conversion technique," *J. Opt. Soc. Korea* **12**(3), 131–135 (2008).
11. J.-H. Park, G. Baasantseren, N. Kim, G. Park, J.-M. Kang, and B. Lee, "View image generation in perspective and orthographic projection geometry based on integral imaging," *Opt. Express* **16**(12), 8800–8813 (2008).
12. M. Zhang, Y. Piao, and E.-S. Kim, "Occlusion-removed scheme using depth-reversed method in computational integral imaging," *Appl. Opt.* **49**(14), 2571–2580 (2010).
13. S.-H. Hong, and B. Javidi, "Distortion-tolerant 3D recognition of occluded objects using computational integral imaging," *Opt. Express* **14**(25), 12085–12095 (2006).
14. D.-H. Shin, B.-G. Lee, and J.-J. Lee, "Occlusion removal method of partially occluded 3D object using sub-image block matching in computational integral imaging," *Opt. Express* **16**(21), 16294–16304 (2008).

15. K.-J. Lee, D.-C. Hwang, S.-C. Kim, and E.-S. Kim, "Blur-metric-based resolution enhancement of computationally reconstructed integral images," *Appl. Opt.* **47**(15), 2859–2869 (2008).
 16. C. M. Do, and B. Javidi, "3D integral imaging reconstruction of occluded objects using independent component analysis-based K-means clustering," *J. Disp. Technol.* **6**(7), 257–262 (2010).
 17. 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).
 18. M. Levoy, and P. Hanrahan, "Light field rendering," in *Proceedings of SIGGRAPH '96* (Association for Computing Machinery, New Orleans, 1996), pp. 31–42.
 19. B. Lee, S. Jung, S.-W. Min, and J.-H. Park, "Three-dimensional display by use of integral photography with dynamically variable image planes," *Opt. Lett.* **26**(19), 1481–1482 (2001).
 20. K. Hong, J. Hong, J.-H. Jung, J.-H. Park, and B. Lee, "Rectification of elemental image set and extraction of lens lattice by projective image transformation in integral imaging," *Opt. Express* **18**(11), 12002–12016 (2010).
 21. B. Horn, and B. Schunck, "Determining optical flow," *Artif. Intell.* **17**(1-3), 185–203 (1981).
 22. A. S. Ogale, C. Fermüller, and Y. Aloimonos, "Motion segmentation using occlusions," *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(6), 988–992 (2005).
 23. D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2010), pp. 2432–2439.
 24. S. Baker, D. Scharstein, J. P. Lewis, S. Roth, M. J. Black, and R. Szeliski, "A database and evaluation methodology for optical flow," in *Proceedings of IEEE Conference on International Conference on Computer Vision* (IEEE, 2007), pp. 1–8.
 25. H. Kim, J. Hahn, and B. Lee, "Mathematical modeling of triangle-mesh-modeled three-dimensional surface objects for digital holography," *Appl. Opt.* **47**(19), D117–D127 (2008).
 26. J. Hahn, Y. Kim, E.-H. Kim, and B. Lee, "Undistorted pickup method of both virtual and real objects for integral imaging," *Opt. Express* **16**(18), 13969–13978 (2008).
 27. R. Martinez-Cuenca, A. Pons, G. Saavedra, M. Martinez-Corral, and B. Javidi, "Optically-corrected elemental images for undistorted Integral image display," *Opt. Express* **14**(21), 9657–9663 (2006).
-

1. Introduction

Three-dimensional (3D) optical information acquisition and processing technology has emerged recently as an important issue with the rapid growth of 3D display market and 3D broadcasting [1,2]. Various methods for 3D image acquisition, analysis and reconstruction are proposed for more effective and realistic 3D display and broadcasting system. 3D optical information acquisition has shown much advance in technology, from the stereo camera based method with horizontal parallax to the two-dimensional (2D) lens array based method with full parallax [3,4]. Integral imaging (InIm) is one of the advanced 3D display and acquisition techniques, which was first invented by Lippmann in 1908 [5]. InIm is composed of a 2D lens array and pickup device such as photographic plate or charge-coupled device (CCD) in pickup phase, and uses another 2D lens array and display device for an elemental image set in display phase. It has the advantages of full parallax, quasi-continuous viewpoints, and expressing color image. Many kinds of 3D optical information processing using InIm were proposed such as depth extraction, 3D object recognition, and 3D image synthesis by many research groups [6–20].

Among them, 3D object reconstruction and recognition of an occluded object is one of crucial issues in 3D optical information processing based on InIm because these applications need a practical use of whole 3D optical information in elemental image set. Many papers about reconstruction of occluded object in InIm have been published, which are based on various image processing algorithms such as the computational integral imaging reconstruction (CIIR), blur metric of target depth plane, and light field rendering [12–18]. However, the previous methods for reconstruction of occluded region have some limitations. The pickup process in the previous methods has no consideration for the gap between lens array and pickup device, which leads to the fixation of central depth plane (CDP) and the limitation of the depth of focus. Moreover, the previous methods used the disparity based depth extraction method with pixel-unit accuracy and blur metric of target depth plane, which are only possible to extract discontinuous depth planes and cannot compute an accurate depth map of 3D volume. This problem is called the depth quantization problem and causes an imprecise reconstruction of the occluded object [7].

To improve the previous methods, we propose the precise depth extraction method using optical flow with sub-pixel accuracy and reconstruction method for the occluded object using point cloud representation and triangular mesh reconstruction with resolution enhancement.

Figure 1 shows the concept of the proposed method. First, we get the 3D optical information using the focal mode of InIm as elemental image set without forming CDP. After the pickup process, the elemental image set is converted to sub-image set which is appropriate for grasping total shape of 3D object and is suitable to compute the optical flow. For disparity estimation, we use the optical flow between sub-images to improve the conventional method which is based on the sum of square difference (SSD) or sum of SSD (SSSD) [7–17]. Optical flow is the motion vector estimation algorithm between neighbor frames in motion pictures [21–24]. In the proposed method, we assume a sub-image frame sequence and calculate the motion vector with sub-pixel accuracy using optical flow. The disparity map is extracted from the optical flows in sub-images, which can be transformed to continuous depth map with high precision.

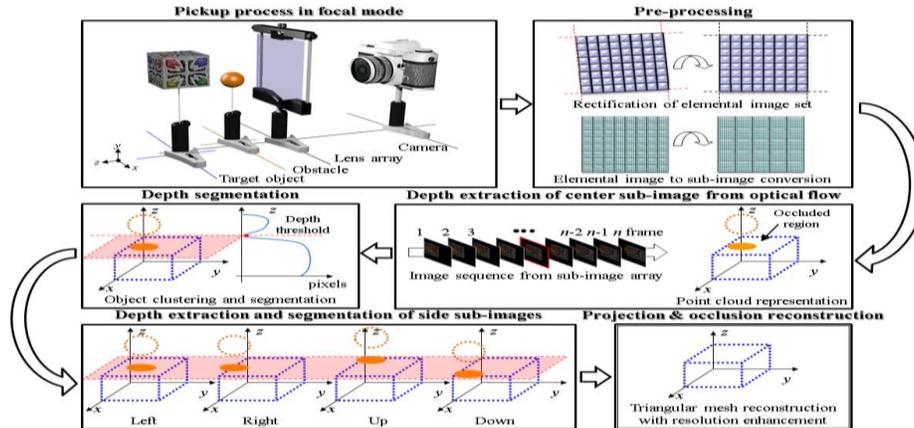


Fig. 1. Concept of the proposed method.

After extraction of depth map with sub-pixel accuracy, 3D information is clustered and segmented to obstacle and target object by the depth threshold using histogram. The segmented sub-images are remapped and projected to the coordinate of center sub-image by using the extracted depth map with sub-pixel accuracy. Each sub-image has different texture and depth information about occluded region of target object, which are represented as point clouds at the viewpoint of center sub-image. The projected point clouds may not only reconstruct the occluded region of target object, but also enhance the resolution of target object using triangular mesh reconstruction. In this paper, we explain the process of the proposed method and the analysis of the range of reconstruction of the occluded object. In experimental results, we present the reconstructed target object without occlusion in arbitrary view point and the result of 3D object recognition using the proposed method with peak signal-to-noise ratio (PSNR) and normalized cross-correlation (NCC).

2. Depth extraction using optical flow with sub-pixel accuracy

2.1. Pickup process based on focal mode in integral imaging and conversion between elemental image and sub-image

For the reconstruction of occluded 3D object, the pickup process without prior depth information has to take precedence in order for a practical depth extraction. The pickup process of InIm is categorized according to the gap between the lens array and digital pickup device as real mode, virtual mode, and focal mode. The gap is set to larger or smaller than the focal length of the lens array in the pickup scheme of real mode or virtual mode, whereas the gap in focal mode is set to the same as the focal length of lens array [4,19].

Figure 2(a) shows the pickup process in real mode, where the gap g is larger than focal length of lens array f . Each ray from 3D object is spatially modulated by the lens array with

lens equation and integrated on the CCD plane as an elemental image set. In this situation, f and g form the focused plane which is called by CDP. The depth of CDP is given by

$$D_{CDP} = \frac{fg}{g-f}. \quad (1)$$

The optical information of 3D object on D_{CDP} is recorded to the elemental image in focus. However, the rays from the 3D object on the other depth planes are captured as blurred image, which is hard to reconstruct and manipulate for depth extraction. In addition, the pickup scheme in real or virtual mode gives prior information for depth extraction of plane object because the range of real or virtual pickup is limited around CDP which is obviously calculated from f and g without optical information processing.

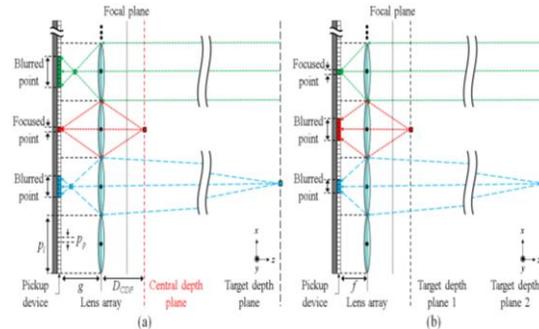


Fig. 2. Two types of pickup process in integral imaging: (a) real mode pickup process with central depth plane and the other depth plane, (b) focal mode pickup process.

As shown in Fig. 2(a), the captured image of the depth plane farther from CDP is more blurred than the depth plane closer to CDP. In pickup process of general cases, the 3D object is relatively located far from the focal plane because f of the lens array is short for the wide-viewing angle [4]. CDP should be formed at far from the focal plane, and g will be close to f . Hence, we use the pickup process in focal mode to focus the 3D object which is located far from the focal plane and to extract the practical depth map without prior information as shown in Fig. 2(b). The captured elemental image is rectified to extract the depth map without distortion using the rectification method based on extraction of lens lattice because the elemental image was tilted or rotated by the misalignment of pickup process [20].

After the pickup process, the rectified elemental image set is converted to sub-image array for the depth extraction using optical flow. The sub-image is a collection of pixels at the same position in all of the elemental image set, or equivalently at the same relative position with respect to the optic axis of the corresponding elemental lens. Previous optical information processing in InIm uses the elemental image to sub-image conversion frequently for orthographic geometry transformation [7,11,14]. To illustrate the difference of elemental image and sub-image, we generate the elemental image set from the pickup scheme in Fig. 1 using computer generated integral photography (CGIP) based on the OpenGL, which is converted to sub-image as shown in Fig. 3. The sub-image is more advantageous to comprehend whole shape of 3D objects than the elemental image since the sub-image has larger field of view in most pickup condition. The orthographic views have the advantage to find an accurate corresponding point in two different images because each sub-image has no distortion from perspective geometry. Moreover, the amount of computation in sub-image based method is smaller than elemental image based method in block searching algorithm for disparity extraction [7,14]. In the proposed method for depth extraction, each sub-image in different viewpoints is assumed as frame images of motion picture, and the motion vectors from the sub-image sequence using optical flow are set to the disparities. In this situation, the above-mentioned advantages of sub-image lead to the precise depth map.

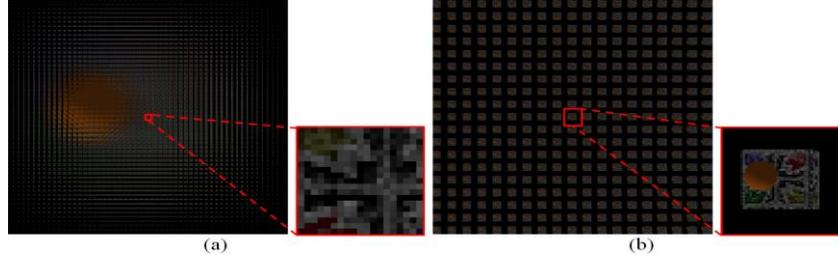


Fig. 3. An example of elemental image to sub-image conversion: (a) computer generated elemental image set and the center elemental image (56×55 lens array, 19×19 pixels/lens, $f = 3.3$), (b) sub-image array and the center sub-image (19×19 sub-image array, 56×55 pixels/sub-image).

2.2. Depth extraction with sub-pixel accuracy using optical flow

The detection of corresponding points on InIm is an essential process in depth extraction method. In previous researches, block searching and minimum difference function, such as the sum of absolute difference (SAD), SSD, and SSSD, are used for finding corresponding points between the reference and neighbor elemental image or sub-image [7–17]. However, the previous depth extraction methods have a fundamental problem which is the depth quantization problem [7]. In pickup and depth extraction process, the 3D optical information is handled with pixel unit of CCD which has finite size and causes the quantization error. Figure 4 shows the depth quantization problem in depth extraction based on the elemental image and sub-image in detail. The extracted depth plane is derived from the disparity of corresponding pixels between reference and neighbor images as follows:

$$D_{d_l, d_p} = \frac{f d_l p_l}{d_p p_p}, \quad (2)$$

where d_l is lens disparity in sub-image, d_p is pixel disparity, p_l is lens pitch, and p_p is pixel pitch. The depth quantization problem occurs from discontinuous pixel disparity in elemental image based depth extraction where the lens disparity is fixed in one lens pitch as shown in Fig. 4(a). If the lenses involved in depth extraction increases from the adjacent lenses to all lenses, the depth quantization problem will be reduced. However, it is not a fundamental solution of the depth quantization problem. If the target depth plane is located between the depth plane of $D_{1,2}$ and $D_{1,1}$, it is not possible to extract the precise depth from disparity with pixel unit.

Likewise, the lens disparity with pixel accuracy brings the depth quantization problem in depth extraction based on the sub-image as shown in Fig. 4(b). The sub-image is the collection of same pixel disparity d_p , and its disparity is represented by the lens disparity d_l . Therefore, the target depth plane which is located between $D_{2,4}$ and $D_{3,4}$ cannot be extracted exactly because of the discontinuous lens disparity with pixel accuracy. Consequently, the disparity extraction method with sub-pixel accuracy is essentially needed to extract the accurate depth map in InIm. We adopt the optical flow to estimate the disparity with sub-pixel accuracy in the proposed method.

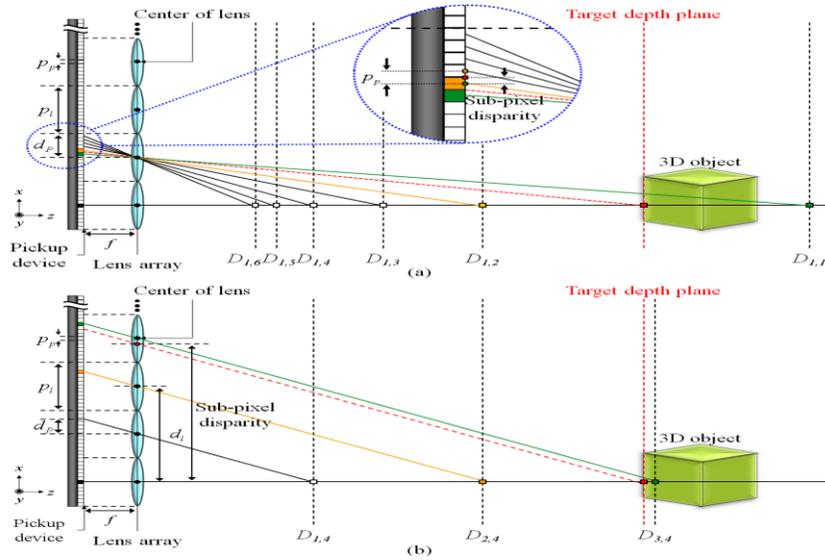


Fig. 4. Principles of depth extraction and depth quantization problem in (a) elemental image based method, (b) sub-image based method.

Optical flow is the distribution of apparent velocities of movement of brightness patterns in a visual scene caused by the relative motion between a camera and the scene [21]. Image sequence of motion picture facilitates the estimation of optical flow as either instantaneous image velocities or disparities. Let the image brightness at the point (x, y) at time t be denoted by $I(x, y, t)$. If the brightness of a particular point in the pattern is constant, the partial derivatives with respect to the spatial and temporal coordinates are derived by

$$\frac{dI}{dt} = \frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0. \quad (3)$$

The horizontal and vertical component of optical flow u, v are given by

$$u = \frac{dx}{dt}, v = \frac{dy}{dt}. \quad (4)$$

Equation (3) is an equation in two components of optical flow and cannot be solved without additional constraints. Various optical flow algorithms propose the additional conditions and constraints to solve Eq. (3) for estimating the actual flow [21–24]. The majority of recent methods strongly resemble the original formulation of Horn and Schunck [21]. We use the optical flow algorithm by Sun *et al.*, which is a modified method of the formulation of Horn and Schunck and can estimate the optical flow with high accuracy using weights according to the spatial distance, brightness, occlusion state, and median filtering [23]. This algorithm is ranked second in both angular and end-point errors in Middlebury evaluation which is well-known benchmarking method for stereo matching and optical flow [24].

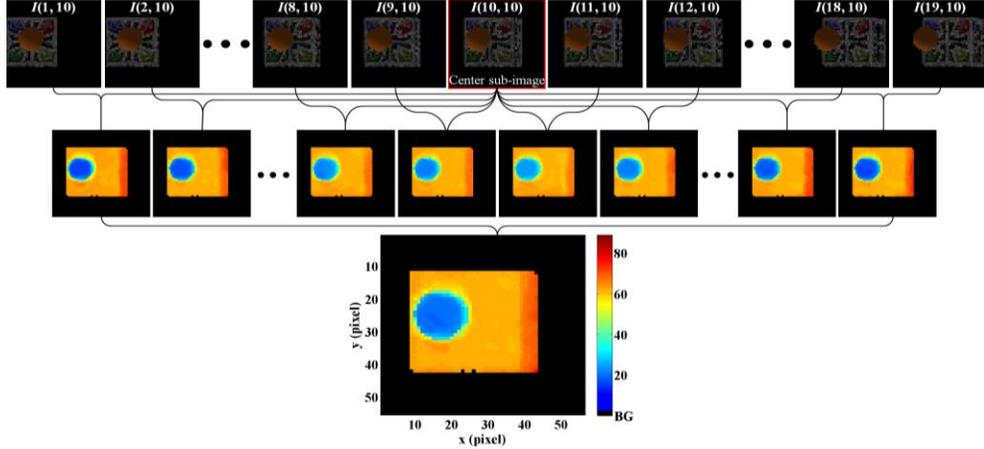


Fig. 5. The process of depth extraction using optical flows between center sub-image and x -directional sub-image sequence.

For the depth extraction in sub-image array, we set the center sub-image as the reference image and assume the x -directional sub-images or y -directional sub-images as the target sub-image sequence for the optical flow calculation. Figure 5 shows the process of depth extraction using optical flows between center sub-image and x -directional sub-image sequence at the same y -position from the elemental image set in Fig. 3. The disparity map from the optical flow between center sub-image and target image sequence is partially different in each sub-image because of different viewpoint. For the extraction of accurate depth map, we gather and average the depth information from optical flows of different viewpoint instead of disparity because the result of optical flow is represented by the unit of lens disparity d_l in Eq. (2). After extraction of depth map of center sub-image from x -directional sub-images, the depth map of center sub-image from y -directional sub-images is calculated in the same manner, and we define the depth map of center sub-image using the average of depth maps from x - and y -directional sub-images. The extracted depth of pixel (x, y) in center sub-image using optical flow is given by

$$D_{OF(p_c, q_c)}(x, y) = \frac{fp_l}{2p_p} \left[\frac{1}{n_x - 1} \sum_{p=1}^{n_x} \left| \frac{u_{(p_c, q_c), (p, q_c)}(x, y)}{p_c - p} \right| + \frac{1}{n_y - 1} \sum_{q=1}^{n_y} \left| \frac{v_{(p_c, q_c), (p_c, q)}(x, y)}{q_c - q} \right| \right], \quad (5)$$

where $u_{(p_1, q_1), (p_2, q_2)}(x, y)$ and $v_{(p_1, q_1), (p_2, q_2)}(x, y)$ denote the horizontal and vertical components of optical flow between (p_1, q_1) th and (p_2, q_2) th sub-images at pixel (x, y) , and (p_c, q_c) th sub-image is the center sub-image in n_x by n_y sub-image array.

Figure 6 shows the extracted depth map with background mask in conventional and proposed method. As shown in Fig. 6(a), the conventional method based on SSSD extracts only the quantized depth planes and shows the discontinuous and inaccurate depth map because of disparities with pixel accuracy. On the contrary, the proposed method gives more continuous and accurate depth map without depth quantization problem because it is calculated from the lens disparity d_l with sub-pixel accuracy using optical flow as shown in Fig. 6(b). Figure 6(c) illustrates the comparison of the ground truth and extracted depths using different methods in y - z plane, and the depths in ball and right edge of box show the improvement of the proposed method remarkably. Therefore, the depth map with high precision from the proposed method can facilitate the reconstruction of occluded object.

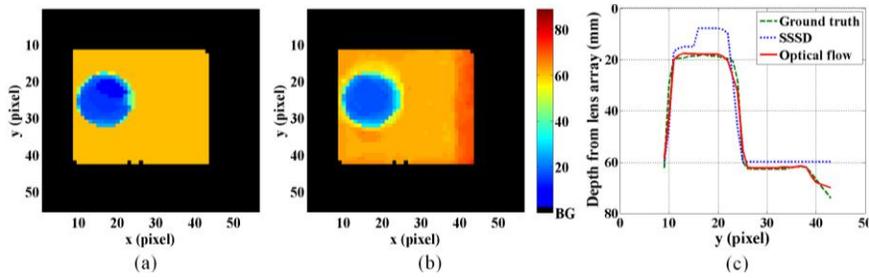


Fig. 6. Comparison of the depth map of center sub-image using SSSD and optical flow: (a) the extracted depth map using SSSD, (b) extracted depth map using optical flow, (c) comparison of the ground truth and extracted depth maps obtained by different methods in y - z plane ($x = 22$).

3. Occlusion reconstruction based on the depth map extraction with sub-pixel accuracy using optical flow

3.1. Depth segmentation using histogram of extracted depth map

For the reconstruction of occluded object, we find the occluded region and remove the obstacle using the extracted depth map. Generally, the obstacle is located in front of target object and has lower depth values than target object. If the extracted depth map has reliably high precision, we can segment the obstacle and target object using the number of pixels in same depth value. To simplify and reduce the computation, we use the histogram based depth segmentation method in various segmentation algorithms. Figure 7 shows the histogram of the number of pixels of depth map in Fig. 6(b). In this situation, the objects are segmented to three clusters which are background, obstacle, and target object using the inflection points of histogram. The pixels of obstacle or target object are gathered around the maximum values in inflection points of histogram of depths, while the background pixels have zero depth. Accordingly, the depth threshold for segmentation of the obstacle and target object is defined by the middle of depths in two inflection points as shown in Fig. 7.

Figure 8 illustrates the point cloud representation of the center sub-image which can express the depth information and texture information to the voxels of 3D space. The point clouds are identical to the vertices of the triangular mesh reconstruction. The texture information of obstacle in center sub-image is removed and filled with black pixels using the depth threshold from histogram while the depth information is remained.

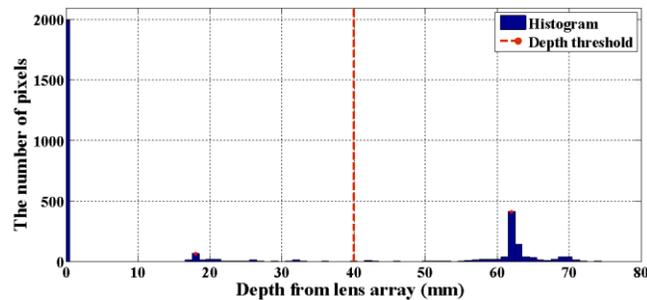


Fig. 7. Histogram of the number of pixels in depth map of center sub-image and depth threshold between the inflection points.

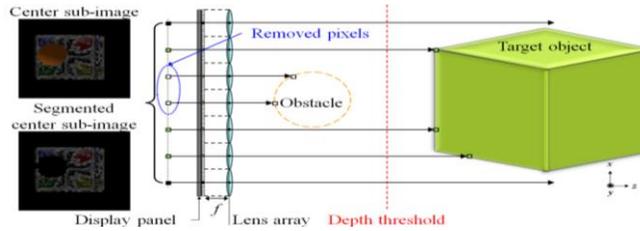


Fig. 8. Point cloud representation and depth segmentation of center sub-image using depth threshold.

3.2. Reconstruction of occluded region using triangular mesh reconstruction of point clouds

The last process of reconstruction of the occluded region in center sub-image is filling of occluded region using the optical information of sub-images in different views which are the end of right side, left side, upper side, and lower side from the center sub-image. The sub-images of side positions in every direction from the center sub-image have the most different optical information about the occluded region. Therefore, we calculate and represent the point clouds of the sub-images at side positions using optical flow as shown in Fig. 9(a).

After the depth extraction of sub-images at side positions, we segment and remove the obstacle using the depth threshold of center sub-image as shown in Fig. 9(a). In this process, the optical information of obstacle is removed, while the different information of occluded region in target object is remained at the different viewpoints. The red points in Fig. 9(a) are the occluded points at the viewpoint of center sub-image. On the other hand, they are the visible points in the sub-images of side positions. For reconstruction of occluded region, we represent and project the point clouds without obstacle at the side viewpoints to the viewpoint of center sub-image as shown in Fig. 9(b). The position of projected point clouds in center sub-image (x', y') is defined as

$$(x', y') = \left(x + \frac{p_p (p_c - p) D_{OF}(x, y)}{fp_l}, y + \frac{p_p (q_c - q) D_{OF}(x, y)}{fp_l} \right), \quad (6)$$

where the position of point clouds in side sub-image (p, q) is (x, y) , and its depth from optical flow is $D_{OF}(x, y)$.

However, the resolution of point cloud representation of center sub-image is limited by the number of lenses in the array, and the projected point clouds have the position information with sub-pixel accuracy. Therefore, we use the triangular mesh reconstruction for enhancing the resolution of center sub-image and reconstructing occluded region. As shown in Fig. 9(b), the red points in Fig. 9(a) are projected to the viewpoint of center sub-image from Eq. (6), which are located between the blue point clouds in occluded region of center sub-image. To apply the information of projected red points, we enhance the resolution of center sub-image with the scale factor S . The triangular mesh reconstruction can interpolate the blue points, which had information of the obstacle in the previous process, from the projected point clouds of target object such as the red points. The occluded region of center sub-image can be reconstructed and enhanced with the projected point clouds using triangular mesh reconstruction.

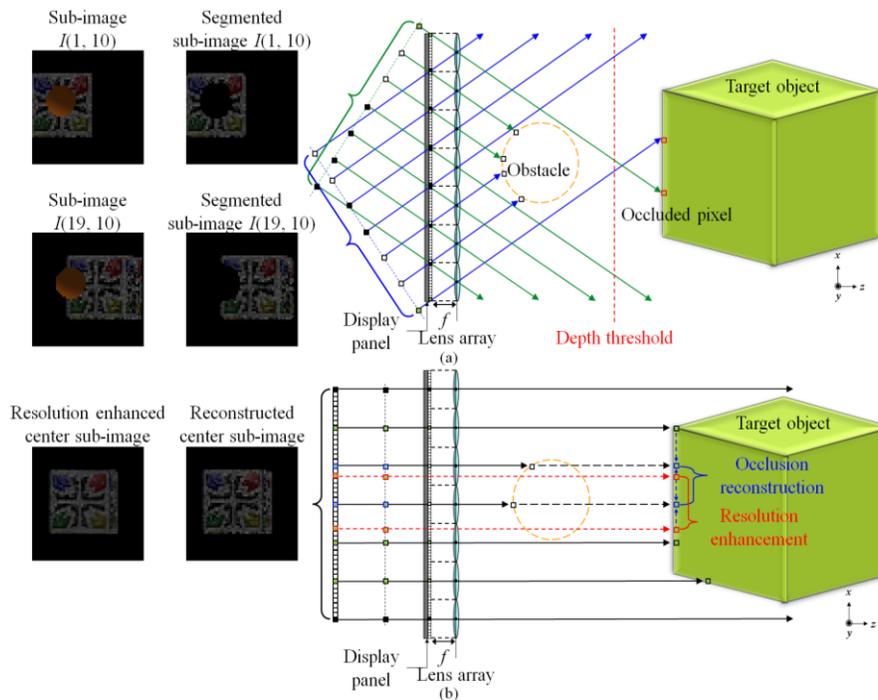


Fig. 9. Reconstruction process of occluded object in point cloud representation: (a) depth extraction and segmentation of side sub-images, (b) reconstruction of center sub-image with resolution enhancement using triangular mesh reconstruction.

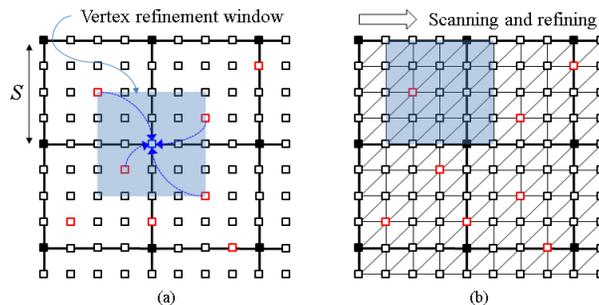


Fig. 10. Triangular mesh reconstruction of center sub-image: (a) reconstruction of the vertex of occluded region in center sub-image (blue point) using neighbor projected points (red points) in the refinement window ($S = 4$), (b) the refining of all other vertices using the scanning of the window and triangular mesh reconstruction.

Figure 10 shows more detailed process of triangular mesh reconstruction. The triangular mesh reconstruction is composed of vertex and facet [9,25]. In the proposed method, the vertex is the point cloud which has both depth and texture information, and the facet is the texture from the interpolation of three vertices. The point clouds of center sub-image are represented to the black-filled vertices with scale factor S as shown in Fig. 10(a). The red points are the projected points from the sub-images at the side position with sub-pixel accuracy, and the blue point is the removed point in center sub-image because it had the information of obstacle. To reconstruct the vertex of occluded region in center sub-image, we average the neighbor projected points which are included in the vertex refinement window. The size of vertex refinement window is the same as the factor of scale S . After the refinement of all vertices in the center sub-image, each vertex with resolution enhancement is filled with the average of red points and black-filled points in its refinement window, and linked with

neighboring three vertices for triangulation as shown in Fig. 10(b). Consequently, the scanning and refining of the vertex refinement window on the scaled center sub-image reconstruct the occluded region and enhance the resolution of center sub-image.

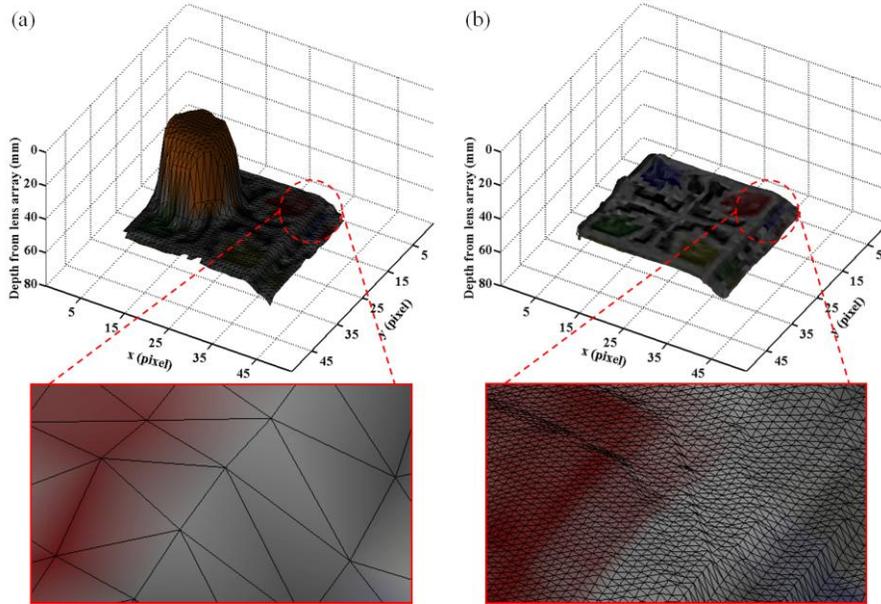


Fig. 11. Triangular mesh reconstruction of occluded region in center sub-image: (a) triangular mesh representation of occluded center sub-image, (b) triangular mesh reconstruction of occluded region in center sub-image with resolution enhancement ($S = 10$).

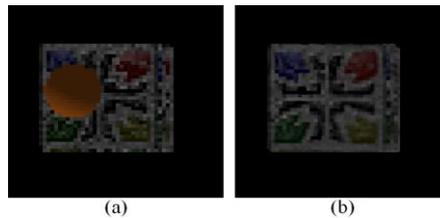


Fig. 12. Comparison of occluded sub-image array and reconstructed sub-image array in arbitrary viewpoints: (a) movie of occluded sub-image array ([Media 1](#)), (b) movie of reconstructed sub-image array with resolution enhancement ($S = 10$) ([Media 2](#)).

Figure 11(a) shows the triangular mesh representation of center sub-image without depth segmentation and reconstruction of occluded region. As the enlarged figure, the resolution of point cloud of center sub-image without the proposed method is the same as the number of lens array. By contrast, the segmented center sub-image is filled with the sub-images of side positions, and the resolution is also enhanced using triangular mesh reconstruction as shown in Fig. 11(b). In the same manner, we can reconstruct all sub-images in arbitrary viewpoints using proposed method with different viewpoints of center sub-image (p_c, q_c) in Eq. (6). Figure 12 shows the movie of original sub-images and the reconstructed sub-images with the scale factor $S = 10$.

4. Analysis of the range of the reconstruction of occluded region using proposed method

In this paper, we propose the method for reconstruction of occluded object using optical flow and triangular mesh representation. However, the proposed method cannot reconstruct all occluded regions. The range of the reconstruction of occluded region is limited by the

specification of lens array and 3D objects. We analyze the range of the reconstruction of occluded region to maximize the range.

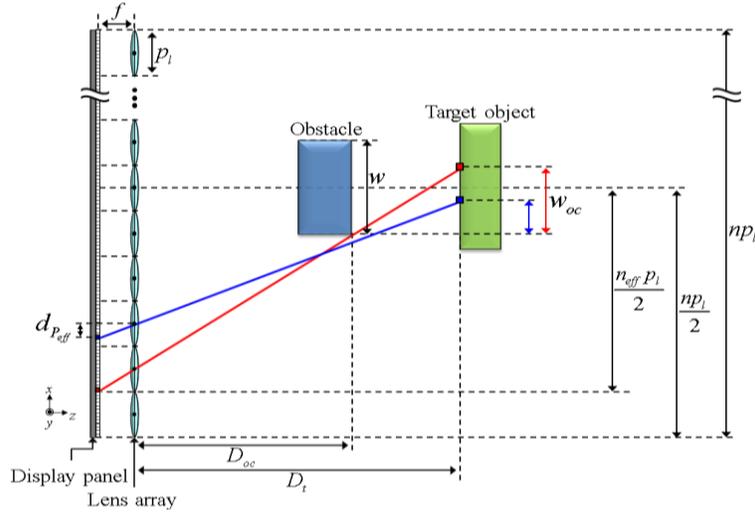


Fig. 13. Analysis of the range of reconstruction in occluded area using the proposed method (odd number n and n_{eff}): w_{oc} in red color is the maximum range of reconstruction with n_{eff} lens array, and blue color is the case of n which is smaller than n_{eff} lens array.

As shown in Fig. 13, the number of effective lenses n_{eff} is defined by the ray that meets an edge of an obstacle with maximum viewing angle, which has to be expressed in positive integer value. We derive the number of effective lens n_{eff} as follows:

$$n_{eff} = \left\lceil \frac{D_{oc}}{f} + \frac{w}{p_l} + 1 \right\rceil, \quad (7)$$

where $\lceil x \rceil$ is the ceiling function of x which is the least integer greater than or equal to x , the depth of obstacle is D_{oc} , the depth of target object is D_t , the width of occlusion is w , and the target object is located at the center of n lens array. If the number of lens array n is larger than n_{eff} , the optical information from the side lenses which are located at the out of n_{eff} is unnecessary for the reconstruction of occluded region. However, if n is smaller than n_{eff} , we have to define the effective pixel disparity $d_{p_{eff}}$ which is defined by the rays passing through the edge of obstacle in n th lens as follows:

$$d_{p_{eff}} = \left\lfloor \frac{f \lceil (n-1) p_l - w \rceil}{2 D_{oc} p_p} \right\rfloor, \quad (8)$$

where $\lfloor x \rfloor$ is the floor function of x which is the greatest integer less than or equal to x . From Eqs. (7) and (8), the maximum range of reconstruction of occluded region w_{oc} is given by

$$w_{oc} = \begin{cases} \frac{D_t p_l}{2f} - \frac{(n_{eff} - 1) p_l}{2} + \frac{w}{2}, & n \geq n_{eff} \\ \frac{D_t d_{p_{eff}} p_p}{f} - \frac{(n-1) p_l}{2} + \frac{w}{2}, & n < n_{eff} \end{cases}. \quad (9)$$

Consequently, w_{oc} can be calculated from four directions of obstacle, and the proposed method can totally reconstruct the occluded region which is smaller than two times of w_{oc} .

5. Experimental result

We have performed the experiments by using the elemental image set which is optically picked up from the two different 3D objects which are textured box and dice of 50 mm and 10 mm size, respectively. As shown in Fig. 14(a), the experimental setup is composed of the pickup camera, lens array, fiber light source, and objects. Figure 14(b) shows the formation of obstacle and target object. The specifications of the experimental setup are listed in Table 1. In many kinds of pickup schemes such as telecentric configuration, we use the camera lens as relay system with rectifying algorithm to simplify experimental setup [4,20,26,27].

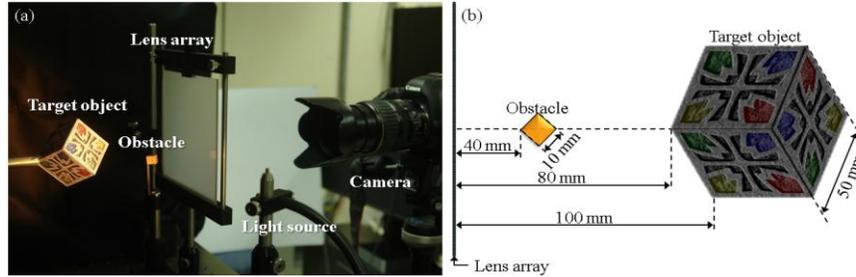


Fig. 14. Experimental setup: (a) experimental setup for pickup process, (b) formation of target object and obstacle.

Table 1. Specification of the experimental setup

Setup	Specification	Characteristic
Lens array	Number of lenses	99 by 98
	Lens size	1 mm by 1mm
	Focal length	3.3 mm
Camera	Model	Canon 5D Mark II
	Resolution	5616 by 3744
	Lens system	Canon EF 28-135mm F3.5-5.6

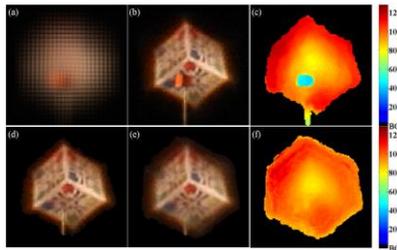


Fig. 15. Experimental result: (a) rectified elemental image set, (b) center sub-image, (c) depth map of center sub-image from the optical flow with sub-pixel accuracy, (d) reconstructed center sub-image, (e) reconstructed center sub-image with resolution enhancement ($S = 5$), (f) reconstructed depth map of center sub-image with resolution enhancement ($S = 5$).

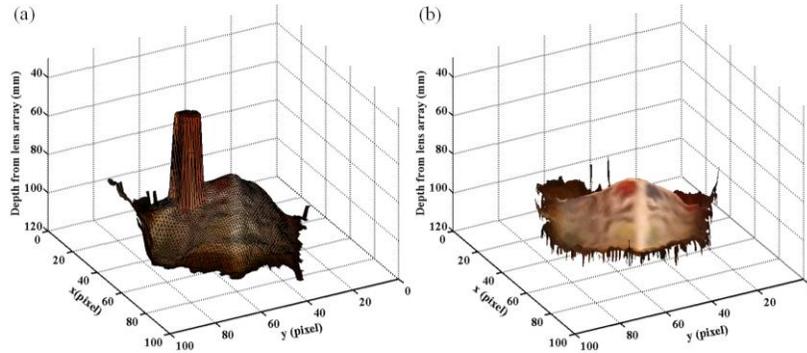


Fig. 16. Experimental result of triangular mesh reconstruction: (a) triangular mesh representation of occluded center sub-image, (b) triangular mesh reconstruction of occluded region in center sub-image with resolution enhancement using proposed method ($S = 5$).

The experimental results are shown in Figs. 15–17. Figure 15(a) is the captured and rectified elemental image set, and the center sub-image from the rectified elemental image set is shown in Fig. 15(b). As shown in Fig. 15(c), the depth map of center sub-image is accurately extracted by using the optical flow without depth quantization problem. After the depth extraction of center sub-image, the depth threshold for the segmentation is defined by 75 mm from the histogram of depth map of center sub-image. In the same manner, the optical flow method extracts the depth maps of sub-images at the side positions, and the sub-images are segmented and projected to the viewpoint of center sub-image. Figure 15(d) shows the result of the reconstruction of the vertex of occluded region in center sub-image, and the reconstruction with resolution enhancement using triangular mesh reconstruction is shown in Fig. 15(e) and 15(f). Figure 16 illustrates the occluded and reconstructed center sub-image with triangular mesh representation, and the movie of the reconstructed sub-images at arbitrary viewpoints is shown in Fig. 17(a) and 17(b). The range of reconstruction of occluded region w_{oc} is 5.6212 pixels from Eq. (9), and the occluded region is successfully reconstructed as shown in experimental results.

For thorough experimental validation, we have performed the experiment with other object sets which are wooden hand and candle of letter ‘L’ shape. The depths of objects are 80 mm and 40 mm, and the sizes are 65 mm by 95 mm and 10 mm by 20 mm, respectively. Figure 17(c) and 17(d) show the occluded and reconstructed sub-images at arbitrary viewpoints.

To verify the feasibility of the proposed method in 3D object recognition, we calculate PSNR and NCC between the center sub-image without obstacle and the reconstructed center sub-image. The test image set of CGIP is the center sub-image in Fig. 3(b) and its reconstructed image and template image which is generated without obstacle. The occluded and reconstructed images in the experiments, boxes and wooden hand with letter ‘L’, are compared with their template images. The PSNR and NCC results are listed in Table 2. We can see that the improvement of the accuracy of 3D object recognition is 4.62 dB in PSNR and 0.0202 in NCC on average, which confirms the feasibility of the proposed method.

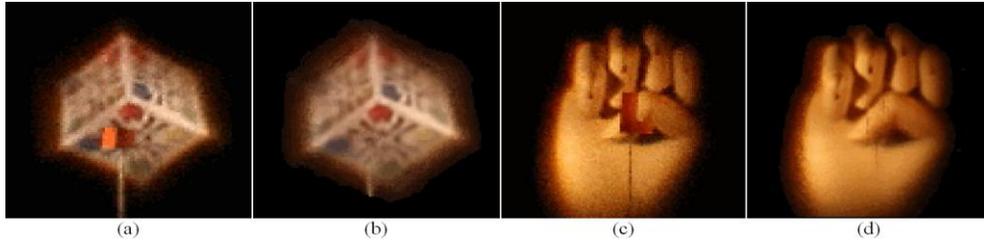


Fig. 17. Experimental result at arbitrary viewpoints: (a) movie of occluded sub-images of box objects (Media 3), (b) movie of reconstructed sub-images of box objects with resolution enhancement ($S = 5$) (Media 4), (c) movie of occluded sub-images of hand object (Media 5), (d) movie of reconstructed sub-images of hand object with resolution enhancement ($S = 5$) (Media 6).

Table 2. PSNR and NCC results of reconstructed object

Test image set	Occluded image		Reconstructed image	
	PSNR	NCC	PSNR	NCC
CGIP	28.78 dB	0.9558	36.93 dB	0.9874
Real objects (box)	26.22 dB	0.9743	28.51 dB	0.9907
Real objects (hand)	25.42 dB	0.9782	28.84 dB	0.9909

6. Conclusion

We proposed a reconstruction method for the occluded region of 3D object using the depth extraction based on the optical flow and triangular mesh reconstruction. A more precise depth extraction is proposed using the optical flow with sub-pixel accuracy, which alleviates the depth quantization problem. From the more accurate depth maps of sub-images at the side positions, the point clouds are represented and projected to the viewpoint of center sub-image without obstacle. Consequently, the point clouds are interpolated by the triangular mesh reconstruction, and the occluded region is reconstructed with resolution enhancement at arbitrary viewpoints. The feasibility of the proposed method was verified by 3D object recognition in PSNR and NCC.

Acknowledgment

This work was supported by the National Research Foundation and the Ministry of Education, Science and Technology of Korea through the Creative Research Initiative Program (2009-0063599).