

## Comparisons of Object Recognition Performance with 3D Photon Counting & Gray Scale Images

Chung Ghu Lee

*Department of Electronic Engineering, Chosun University, 375 Seosuk-dong, Dong-gu, Gwangju 501-759, Korea*

Inkyu Moon\*

*School of Computer Engineering, Chosun University, 375 Seosuk-dong,  
Dong-gu, Gwangju 501-759 Korea*

(Received April 16, 2010 : revised September 28, 2010 : accepted November 15, 2010)

In this paper the object recognition performance of a photon counting integral imaging system is quantitatively compared with that of a conventional gray scale imaging system. For 3D imaging of objects with a small number of photons, the elemental image set of a 3D scene is obtained using the integral imaging set up. We assume that the elemental image detection follows a Poisson distribution. Computational geometrical ray back propagation algorithm and parametric maximum likelihood estimator are applied to the photon counting elemental image set in order to reconstruct the original 3D scene. To evaluate the photon counting object recognition performance, the normalized correlation peaks between the reconstructed 3D scenes are calculated for the varied and fixed total number of photons in the reconstructed sectional image changing the total number of image channels in the integral imaging system. It is quantitatively illustrated that the recognition performance of the photon counting integral imaging system can be similar to that of a conventional gray scale imaging system as the number of image viewing channels in the photon counting integral imaging (PCII) system is increased up to the threshold point. Also, we present experiments to find the threshold point on the total number of image channels in the PCII system which can guarantee a comparable recognition performance with a gray scale imaging system. To the best of our knowledge, this is the first report on comparisons of object recognition performance with 3D photon counting & gray scale images.

*Keywords* : Three-dimensional imaging, Integral imaging, Three-dimensional image recognition, Photon counting

*OCIS codes* : (110.6880) Three-dimensional image acquisition; (100.6890) Three-dimensional image processing; (030.5260) Photon counting

### I. INTRODUCTION

Three-dimensional (3D) integral imaging and its applications have been investigated for 3D sensing, visualization, display, and recognition of objects [1-12]. This technique records 2D elemental (or multi view) images of a 3D scene for 3D image reconstruction and depth extraction. Both optical and numerical reconstructions are possible [10-11, 13]. Recently, photon counting integral imaging has been explored for 3D object reconstruction and recognition for photon-starved

scenes. A variety of techniques, including maximum likelihood estimator and statistical sampling theory were applied for 3D imaging and recognition of photon-starved objects, respectively, using a fixed number of elemental images [13-15].

In this paper, we quantitatively compare the object recognition performance of a photon counting integral imaging (PCII) system with that of a conventional gray scale imaging system. Elemental image detection is based on a conventional photon counting model generating different Poisson numbers for each elemental image. A computational ray back propagation

---

\*Corresponding author: inkyu.moon@chosun.ac.kr

Color versions of one or more of the figures in this paper are available online.

algorithm and a parametric maximum likelihood estimator [16] are applied to the photon limited elemental images in order to reconstruct the 3D scene. The performance of photon counting image recognition is evaluated by measuring the correlation peaks between the reconstructed 3D scenes. The photon-limited 3D reference object is computationally reconstructed to synthesize a matched filter. Normalized correlation peak values are calculated for the varied and fixed total number of photons in the reconstructed sectional image in order to compare the photon counting recognition performance obtained by changing the total number of the image channels in the PCII system. These results are compared with conventional 2D image recognition using gray scale images. We quantitatively illustrate that the object recognition performance of PCII system can be similar to one obtained by using a general gray scale image as the number of the image viewing channels in integral imaging system is increased up to the threshold point. Also, we show that the threshold point on the total number of image channels in the PCII system, which guarantees a comparable recognition performance with gray scale imaging system, can be found using a statistical sampling distribution of the measured average correlation peak values.

## II. OBJECT RECOGNITION WITH PCII SYSTEM

Fig. 1 shows a schematic setup of the integral sensing system with multiple image channels. The scene with very low intensity level light generates the photon counting elemental images. For 3D sensing of photon limited object, the system may utilize a single sensor with lenslet array, a sensor array, or a single moving camera in order to record the low intensity level light ray emanating from a 3D object. Each sensor captures its own photon counting two-dimensional (2D) elemental (perspective) image, which contains directional

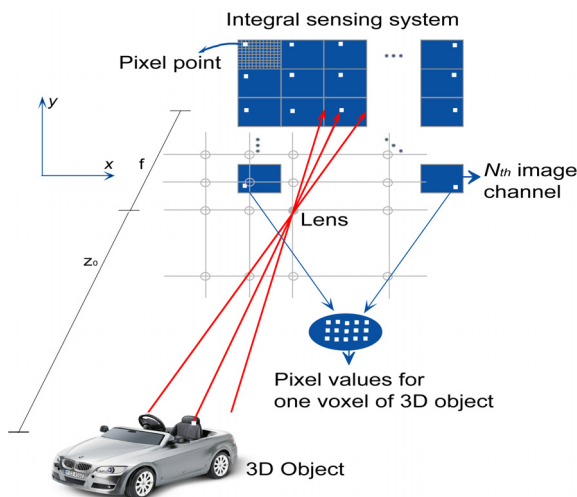


FIG. 1. setup of 3D integral imaging system with multiple elemental images (image channels).

information of the 3D object. The irradiance of one voxel on the surface of the photon limited 3D object is recorded on the corresponding pixel position of each photon limited 2D elemental image. Numerical 3D reconstruction of the original object can be performed by applying computational ray back propagation algorithm and parametric maximum likelihood estimator (MLE) to the recorded photon counted elemental images [13]. The pixel values on each elemental image corresponding to one voxel of the photon limited 3D object are assumed to be a random variable following a *Poisson* distribution function [17]. It is previously shown that one voxel value of the original 3D object can be retrieved by applying the parametric MLE to the elemental images pixel values as follows [13]:

$$MLE(\tilde{N}_v) = \arg \max_{\tilde{N}_v} \left[ \log \prod_{n=1}^{N_e} \exp(-\tilde{N}_v) (\tilde{N}_v)^{c_v(n)} \right] = \frac{1}{N_e} \sum_{n=1}^{N_e} c_v(n), \quad (1)$$

where  $C_v$  is *Poisson* random number with the mean parameter  $\tilde{N}_v$  at each voxel of the reconstructed image, subscript  $v$  denotes a voxel index,  $\tilde{N}$  is the expected number of photons in the photon limited elemental image,  $I_v$  is the normalized irradiance, and  $N_e$  is the total number of elemental images [13]. Thus, the original 3D object can be estimated by using this reconstruction procedure. It is interesting to note that a single photon at wavelength  $\lambda=500$  nm carries an energy of  $E=hc/\lambda = 3.97 \times 10^{-19}$  Joules, where  $h$  is Plank's constant and  $c$  is  $3 \times 10^8$  meters/second so that the received total energy  $E_t$  in the reconstructed sectional image of the PCII system can be equal to  $\frac{hc}{\lambda} \tilde{N} N_e$  [13]. Therefore, the received photon energy becomes proportional to the total number of the image channel in the PCII system,  $N_e$ . In this paper, it should be noted that the *Poisson* parameter (mean and variance) values for each pixel in the elemental images in the PCII system are very small. In other words, the probability of counting more than one photon per one pixel in the elemental image is very small.

In order to evaluate the object recognition performance of PCII system, a matched spatial filter is applied to the sectional images reconstructed by using the integral imaging technique as follows [18]:

$$C(x, y) = \left| F^{-1} \left\{ F[PII^{ref}(x, y; z=z_0)] \times F^*[PII^{inp}(x, y; z=z_0)] \right\} \right|^2 \quad (2)$$

where  $PII^{ref}(x, y; z=z_0)$  and  $PII^{inp}(x, y; z=z_0)$  are sectional images for reference object and input objects, respectively,  $z_0$  is the reconstruction distance, and  $F$  denotes the *Fourier* transform.

## III. EXPERIMENTAL RESULTS

Experiments to evaluate and compare object recognition performance of an imaging system based on a conven-

tional photon counting model are presented. We recorded  $9 \times 9$  elemental images for two toy cars by moving a CCD camera transversally in both x and y directions as shown in Fig. 1. Two objects denoted as car I and car II are used as two different classes of data for recognition (see Fig. 2).

According to the classical photon counting detection model, the photon-counting elemental images of the toy cars were generated from the recorded elemental images, respectively [13]. Then, the sectional images of the 3D scenes for toy cars I and car II were reconstructed at a distance of  $z_0 = 100\text{cm}$  with their corresponding photon-counting elemental images. We vary the expected number of photons in each input scene to test the recognition performance of the PCII system. Figure 3 shows the sectional images reconstructed from the photon-counting elemental images of car I and car II, respectively.

The matched filter in Eq. (2) was applied to the reconstructed sectional images in order to inspect object recognition performance of PCII system. Figure 4(a) shows the correlation plots of the photon counting image of car I in Fig. 3(a) used as a reference with the true class photon counting image as an input data. The true class input data was generated with the same expected number of photons ( $\bar{N} = 1000$ ) and total number of the elemental images ( $N_e = 1$ ) as the reference but independent *Poisson* distributions were used in each case. Figure 4(b) shows the correlation plots of the photon counting image of the car I in Fig. 3(a) as a reference object with that in Fig. 3(b) as a false class

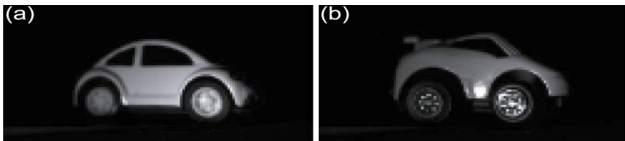


FIG. 2. Two toy cars used in the recognition experiments. Each elemental image size is  $125 \times 125$  pixels. (a) Car I. (b) Car II.

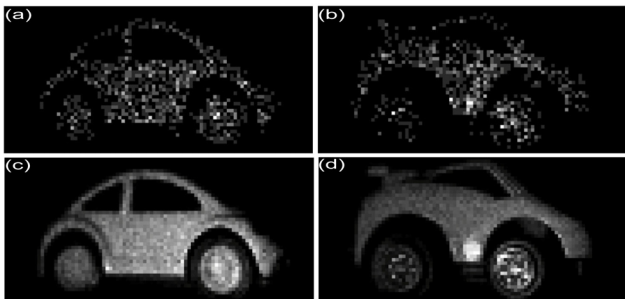


FIG. 3. Sectional images reconstructed from the photon counting elemental images of cars I and II. The expected number of photons in the photon limited elemental image  $\bar{N}$  was 1000. The total number of the elemental images  $N_e$  was varied from 1 or 81. (a) Reconstructed car I with  $N_e = 1$ , (b) reconstructed car II with  $N_e = 1$ , (c) reconstructed car I with  $N_e = 81$ , (d) reconstructed car II with  $N_e = 81$ .

input data. For comparison, both plots are normalized to the same value that is the autocorrelation of the reference object. The measured correlation peak values were 0.21 and 0.12, respectively. It is noted that it is difficult to make a discrimination between car I and II by using only a single photon counting elemental image with the expected number of photons  $\bar{N} = 1000$ .

Figure 5(a) shows the correlation plots of the photon counting image of the car I in Fig. 3(c) as a true reference object (true class) with the true class photon counting image as an input data. The true class input data was generated with the same expected number of photons ( $\bar{N} = 1000$ ) and total number of the elemental images ( $N_e = 81$ ) as the reference but with independent *Poisson* distributions in each case. Figure 5(b) shows the correlation plots of the

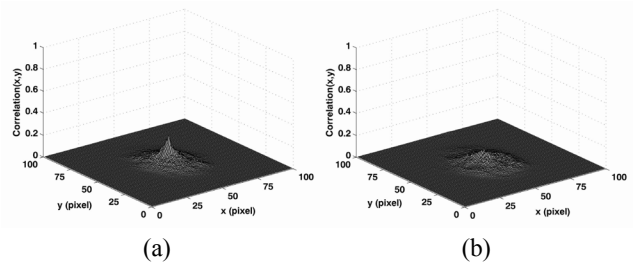


FIG. 4. The correlation plots obtained with the sectional images reconstructed from a small number of photons-counting elemental images. The expected number of photons in the photon limited elemental image  $\bar{N}$  was 1000. The total number of the elemental images (parallax)  $N_e$  was 1. (a) Correlation plot between the photon counting image of the car I in Fig. 3(a) as a true class target and the true class photon counted image as an input data, (b) correlation plot between the photon counting image of the car I in Fig. 3(a) as a true class target, and the photon counted image of car II in Fig. 3(b) as a false class input data.

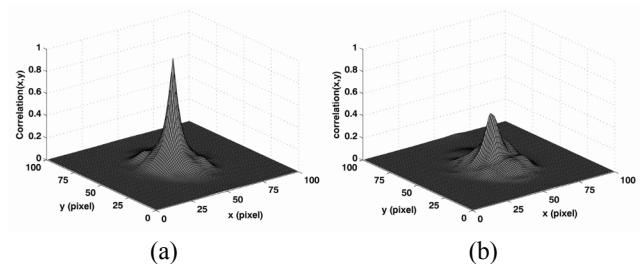


FIG. 5. The correlation plots obtained with the sectional images reconstructed from the small number of photons-counted elemental images. The expected number of photons in the photon limited elemental image  $\bar{N}$  was 1000. The total number of the elemental images (parallax)  $N_e$  was 81. (a) Correlation plot between the photon counted image of the car I in Fig. 3(c) as a reference and the true class photon counted image as an input data, (b) correlation plot between the photon counted image of the car I in Fig. 3(c) as a reference and the false class photon counted image of car II in Fig. 3(d) as an input data.

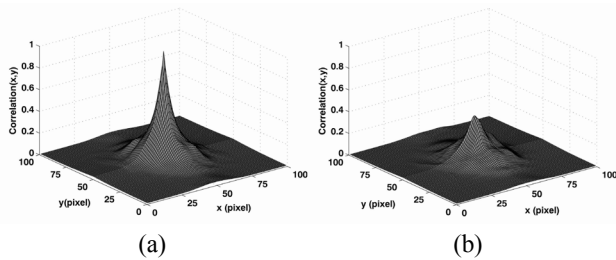


FIG. 6. The correlation plots obtained with the conventional gray scale images. (a) Auto-correlation plot of the gray scale intensity image of car I in Fig. 2(a). (b) cross-correlation plot between the gray scale intensity image of the car I in Fig. 2(a) as a reference object and the false class gray scale intensity image of car II in Fig. 2(b) as an input data.

photon counting image of car I in Fig. 3(c) as a reference object with that in Fig. 3(d) for car II as a false class input data. Both plots (Fig. 5(a) and 5(b)) were normalized to the same value that is the autocorrelation of the true reference object. The measured correlation peak values were 0.97 and 0.46, respectively. For comparison, we present the auto-correlation and cross-correlation plots in Fig. 6 with the gray scaled intensity images of car I and car II in Fig. 2(a) and 2(b), respectively. The gray scale images in Fig 2 have the same view point as the photon counting images in Fig 3. Both plots (Fig. 6(a) and 6(b)) were normalized to the same value that is the autocorrelation of the true reference object. The auto-correlation and cross-correlation peak values calculated by using the conventional gray scale intensity images were 1.00 and 0.41, respectively. It is noted that when the total number of the image channels in the integral sensing system was increased to 81, the correlation plots between reference (car I) and unknown input photon counted images (car I and car II) become very similar to those between reference and unknown input intensity images in Fig. 2.

Figure 7 shows the correlation peak values computed between the photon counting sectional image of reference car I, and the photon counting sectional image of true class car I or false class car II. The correlation peak values were normalized to the same value that is the autocorrelation of the true reference object. The total number of elemental image  $N_e$  was varied as 1, 9, 25, 49 and 81. In order to obtain the statistical parameters (maximum, mean and minimum values in Fig. 7) of the computed correlation peak values between reference and unknown input data, the sectional image of the 3D scenes for the reference car I was reconstructed at distance  $z_0 = 100$  cm with its own photon-counting elemental images. Then, the sectional images of the 3D scenes for the true class car I and false class car II were reconstructed at distance  $z_0 = 100$  cm with their own photon-counting elemental images, respectively. The process is repeated 30 times. Each time, the correlation peak values between reference and unknown input data were calculated.

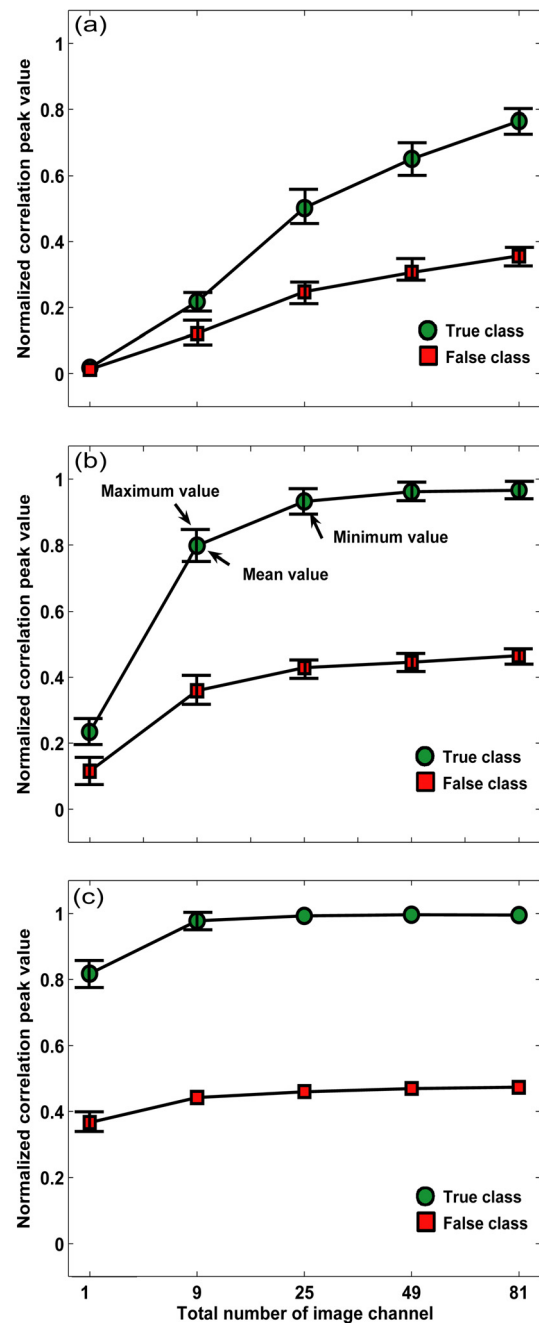


FIG. 7. Correlation peak values computed between the photon counting sectional images of reference car I and the photon counting sectional image of true class car I or false class car II, where the total number of elemental images (image channels)  $N_e$  was varied with 1, 9, 25, 49 and 81. (a)  $\bar{N}=100$ , (b)  $\bar{N}=1000$ , (c)  $\bar{N}=10000$ , where  $\bar{N}$  is the expected number of photons in the photon limited elemental image.

As shown in Fig. 7, the average correlation peak value between the photon counting sectional image of reference car I and the photon counting sectional image of true class car I approaches the auto-correlation peak value (=1.00) of the conventional intensity imaging in Fig. 2(a) when the

total number of the image channel in the system was increased for fixed a number of photons. It is also noted that the average correlation peak value between the photon counting sectional image of reference car I and the photon counting sectional image of false class car II approaches the cross-correlation peak value ( $=0.41$ ) calculated of the conventional intensity imaging of the car I and car II in Fig. 2(a) and 2(b) when the total number of the image channels was increased for fixed number of photons. These experimental results in Fig. 7 demonstrate that the increasing the number of the image viewing channels in PCII system can enable the recognition performance of PCII system to be similar to one of a conventional gray scale imaging system.

As the experiment to find the threshold point on the total number of image channels for the fixed averaged number of photons  $\bar{N}$  which guarantees a comparable recognition performance with gray scale imaging system, Table 1 and 2 show the average correlation peak values computed between the photon counting sectional image of reference car I and the photon counting sectional image of true class car I or false class car II changing the number of the image channels, respectively, where the expected number of photons in the photon limited elemental image was fixed as 10, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, or 10000. The correlation peak values in the Table 1 and 2 were normalized to the same value that is the autocorrelation of the true reference object. The

TABLE 1. Average correlation peak values computed between the photon counting sectional images of reference car I and the photon counting sectional image of true class car I

Expected number of photons in the elemental image ( $\bar{N}$ )	Total number of image channels ( $N_c$ )				
	1	9	25	49	81
10	0.0469	0.0162	0.0471	0.0842	0.1977
1000	0.2300	0.7440	0.9254	0.9495	0.9731
2000	0.4227	0.8996	0.9606	0.9770	0.9946
3000	0.5565	0.9291	0.9666	0.9793	0.9863
4000	0.6011	0.9433	0.9824	0.9959	0.9911
5000	0.6726	0.9453	0.9777	0.9966	0.9998
6000	0.7414	0.9500	0.9828	0.9911	0.9932
7000	0.7423	0.9699	0.9894	0.9977	1.0000
8000	0.7459	0.9653	0.9849	0.9918	0.9959
9000	0.7902	0.9792	0.9924	1.0000	0.9994
10000	0.8023	0.9774	0.9955	0.9961	0.9964

TABLE 2. Average correlation peak values computed between the photon counting sectional images of reference car I and the photon counting sectional image of false class car II

Expected number of photons in the elemental image ( $\bar{N}$ )	Total number of image channels ( $N_c$ )				
	1	9	25	49	81
10	0.0367	0.0118	0.0281	0.0526	0.0988
1000	0.1154	0.3381	0.4349	0.4463	0.4636
2000	0.1862	0.4128	0.4472	0.4551	0.4734
3000	0.2502	0.4257	0.4490	0.4598	0.4708
4000	0.2734	0.4169	0.4582	0.4702	0.4745
5000	0.2956	0.4308	0.4489	0.4663	0.4772
6000	0.3303	0.4297	0.4531	0.4660	0.4756
7000	0.3252	0.4355	0.4583	0.4663	0.4801
8000	0.3319	0.4365	0.4558	0.4668	0.4769
9000	0.3542	0.4448	0.4575	0.4703	0.4765
10000	0.3552	0.4456	0.4601	0.4675	0.4732

sectional image tested in the experiments has  $125 \times 125$  pixels. It can be assumed that an arbitrary cut-off correlation value (for example 0.95 or 0.99) utilized in the conventional gray scale imaging system can be a possible decision criterion to evaluate the recognition performance in the 3D PCII system. Therefore, the threshold point on the total number of image channels for the arbitrary fixed averaged number of photons in the elemental image may be empirically found with the average correlation peak value measured in experiments. For example, if the average correlation peak value as the decision rule for an object recognition is set as 0.9500 or 0.9900 with the fixed averaged number of photons  $\bar{N}=6000$ , the threshold point on the total number of image channels can be 9 and 49, respectively as shown in Table 1. It seems very difficult to derive the closed-form to describe the relationship between the optical parameters ( $\bar{N}$ ,  $N_e$ ,  $z_0$ , and reconstructed sectional image size) in the PCII system in order to find the threshold point on the total number of image channels which guarantees a comparable recognition performance with gray scale image. However, it can be possible to empirically form the statistical distributions of the average correlation peak values for the training objects to classify the object of interest with the given optical parameters in PCII system. Then, the threshold point on the total number of image channels can be found using the statistical distribution of the average correlation peak values for the training objects.

The reconstructed sectional image can be characterized by the photon number per pixel, the total number of elemental images  $N_e$  and the expected number of photons  $\bar{N}$  in the photon limited elemental image. It is noted that the sectional image tested in the experiments with  $\bar{N}=6000$  and  $N_e=9$  has only 3.456 photons per pixel. It means that the received total energy in the reconstructed sectional image required for obtaining a comparable recognition performance with gray scale imaging system would be approximately  $2.14 \times 10^{-14}$  Joules with an average wavelength of 500 nm. It is also interesting to note that the average correlation peak value is kept as 0.9500 when the total number of photons in the reconstructed sectional image  $N_p$  is approximately fixed as 50000 (in other words, the received total energy in the reconstructed sectional image is approximately fixed as  $2.14 \times 10^{-14}$  Joules) even though the expected number of photons  $\bar{N}$  in the elemental image was decreased. These experimental results quantitatively illustrate that the recognition performance of the PCII system can be similar to one of conventional gray scale imaging system with less power consumption on condition that any noise sources in our analysis/experiments are not included. In addition, it is shown that the threshold point on the total number of image channels in the PCII system, which guarantees a comparable recognition performance with gray scale imaging system, can be found using the statistical distribution of the average correlation peak values measured in the experiments. In our next work, we

are planning to analyze the recognition performance in the PCII system with noise sources.

#### IV. CONCLUSION

In summary, we have quantitatively compared object recognition performance of PCII system with that of object recognition using conventional gray scale images. For visualization of a scene with small number of photons, integral imaging technique and parametric maximum likelihood estimator can be applied to the photon limited elemental image of the object. To evaluate photon counting object recognition performance, normalized correlation peak values between the reconstructed reference object and unknown input objects are computed for the varied total number of photons or the fixed one in the reconstructed sectional image changing the total number of image channels in the PCII system. The results are compared with those measured between the reference object and unknown input images obtained by conventional gray scale imaging system. It is shown that photon counting object recognition performance can be similar to one of conventional gray scale imaging system as the total number of the image channels in PCII system is increased up to the threshold point. Also, it is shown that the threshold point on the total number of image channels in the PCII system can be found using the statistical distribution of the average correlation peak values measured in the experiments. The PCII system to require less power consumption may replace the conventional imaging systems that generate gray scaled images for object recognition purposes.

#### ACKNOWLEDGMENT

This study was supported by research fund from Chosun University, 2010.

#### REFERENCES

1. G. Lippmann, "La photographie integrale," *C. R. Acad. Sci.* **146**, 446-451 (1908).
2. H. E. Ives, "Optical properties of a Lippmann lenticuled sheet," *J. Opt. Soc. Am.* **21**, 171-176 (1931).
3. B. Javidi, F. Okano, and J. Son, *Three-dimensional Imaging, Visualization, and Display Technologies* (Springer, New York, USA, 2008).
4. T. Wei, D. Shin, and B. Lee, "Resolution-enhanced reconstruction of 3D object using depth-reversed elemental images for partially occluded object recognition," *J. Opt. Soc. Korea* **13**, 139-145 (2009).
5. R. Martinez-Cuenca, G. Saavedra, M. Martinez-Corral, and B. Javidi, "Progress in 3-D multiperspective display by integral imaging," *Proc. IEEE* **97**, 1067-1077 (2009).

6. T. Okoshi, "Three-dimensional displays," Proc. IEEE **68**, 548-564 (1980).
7. F. Okano, J. Arai, K. Mitani, and M. Okui, "Real-time integral imaging based on extremely high resolution video system," Proc. IEEE **94**, 490-501 (2006).
8. S. Park, B. Song, and S. Min, "Analysis of image visibility in projection-type integral imaging system without diffuser," J. Opt. Soc. Korea **14**, 121-126 (2010).
9. Y. Igarishi, H. Murata, and M. Ueda, "3D display system using a computer-generated integral photograph," Jpn. J. Appl. Phys. **17**, 1683-1684 (1978).
10. H. Arimoto and B. Javidi, "Integrate three-dimensional imaging with computed reconstruction," Opt. Lett. **26**, 157-159 (2001).
11. A. Stern and B. Javidi, "Three-dimensional image sensing, visualization, and processing using integral imaging," Proc. IEEE **94**, 591-607 (2006).
12. M. Levoy, "Light fields and computational imaging," IEEE Computer Mag. **39**, 46-55 (2006).
13. B. Tavakoli, B. Javidi, and E. Watson, "Three dimensional visualization by photon counting computational integral imaging," Opt. Express **16**, 4426-4436 (2008).
14. S. Yeom, B. Javidi, and E. Watson, "Photon counting passive 3D image sensing for automatic target recognition," Opt. Express **13**, 9310-9330 (2005).
15. I. Moon and B. Javidi, "Three-dimensional recognition of photon-starved events using computational integral imaging and statistical sampling," Opt. Lett. **34**, 731-733 (2009).
16. M. Guilanum, P. Melon, and P. Refregier, "Maximum-likelihood estimation of an astronomical image from a sequence at low photon levels," J. Opt. Soc. Am. A **15**, 2841-2848 (1998).
17. J. W. Goodman, *Statistical Optics* (John Wiley & Sons, inc., Hoboken, USA, 1985).
18. B. Javidi and E. Tajahuerce, "Three-dimensional object recognition by use of digital holography," Opt. Lett. **25**, 610-612 (2000).