P25

CMOS Image Sensor for 3-D Range Map Acquisition Using Time Encoded 2-D Structured Pattern

Hiroki Yabe† and Makoto Ikeda‡
†Department of Electrical Engineering and Information Systems, The University of Tokyo
‡VLSI Design and Education Center (VDEC), The University of Tokyo
2-11-16 Yayoi, Bunkyo-ku,Tokyo, 113-0032 Japan
Email: {yabe, ikeda}@silicon.t.u-tokyo.ac.jp
Tel: +81-3-5841-6771 Fax: +81-3-5841-8912

***Abstract*—An image sensor in 0.18μm CMOS process for range map acquisition using a gray-encoded time-multiplexing 2-D structured pattern is presented. In order to achieve a high range map rate, we designed a dedicated image sensor that exports only the address of the projected encoded patterns after integrating 7 images. This mechanism eliminates unnecessary read out, thus improving 3-D acquisition speed. In this method the only information needed to reconstruct 3-D range map is whether the pixel is bright or not for the exposed structured patterns. We employed a 1-bit A/D converter in each pixel to determine the pattern. The 1-bit pixel data is stored in the in-pixel memory, and after acquiring 7 images, the 7-bit is read out in parallel. This reduces the read out overhead to 1/7. Using these techniques, the proposed chip achieved the read out speed of 292μs per range map. Measured results show 3-D range map acquisition of 92.7 range map per second with 1000 lumen DMD projector, with average and maximal range error of 4.6mm and 302mm at target distance of 800mm. Range error improves down to 2.5mm and 9.0mm, with brighter projector of 2800 lumen.**

I. Introduction

The 3-D surface reconstruction technique has a wide range of applications like robotics and human interfaces. The method frequently used for this purpose is stereo vision [1]. However, stereo vision is less robust because it depends highly on the illumination condition target objects, and calculating correspondence between two camera pixels is relatively complicated.

The structured light method is one of the 3-D surface reconstruction method based on projecting a light pattern onto a target object. Projecting pattern eases the correspondence problem and make this method more robust owing to lower dependence on the illumination condition. Among many patterns proposed so far [2], the gray-coded time-multiplexed structured light is a simple yet reliable way to reconstruct the 3-D surface of objects. Compared to methods that uses space-encoded patterns, where complicated calculation is required to mitigate the effect of shadows, occlusions and discontinuities [3], the gray-encoded pattern encodes address in time region, therefore not suffering from these problems and resulting in simpler decoding. However, to increase accuracy, the number of 2-D frames used for reconstruction must be increased, thus degrading overall range map rate.

In this paper, we demostrate an image sensor to improve the 3-D range map acquisition speed by reducing the read out time with in-pixel 1-bit A/D and 7-bit accumulators.

II. Image Sensor Chip Design

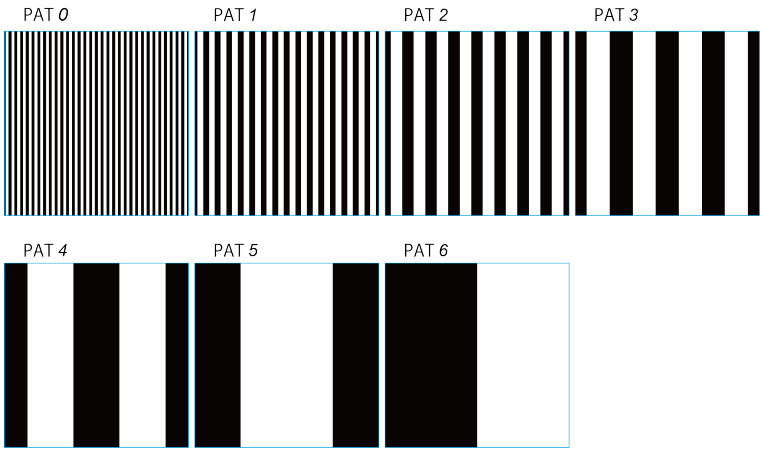
Fig. 1 shows the time encoded light pattern used for 3-D range map acquisition. The value of each pixel in PATi is the i-th bit of the address, i.e. gray coded x-coordinate. The use of gray code instead of binary code makes the recognition more robust because neighboring symbols in gray code differ by only one bit. In addition to this, the spatial frequency of gray-coded patterns become half of that of the binary code, results in less pattern recognition error.

3-D range map is reconstructed as follows:

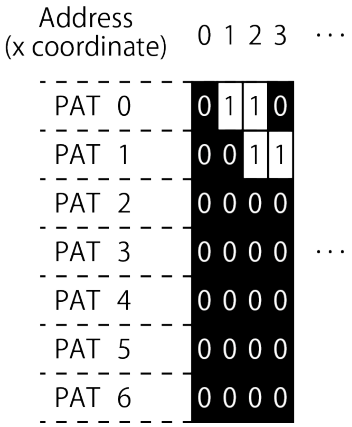
- Time encoded patterns are projected onto a target using a video projector.
- Images of the projected patterns are captured by an image sensor.
- Captured images are thre binarized.
- The address for each pixel is obtained from the binarize pixel values.
- 3-D surface shape is reconstructed by triangulation.

The depth resolution is determined by resolution of the 2-D structured pattern and number of horizontal pixels of the image sensor. 3-D range map acquisition speed is determined by the acquisition time of the 2-D structured patterns. Conventionally, the range map rate of 3-D acquisition using the time encoded 2-D structured pattern is limited to 1/(number of patterns), 1/7 in this study, of the 2-D frame rate. To improve the 3-D acquisition speed, we designed an image sensor that accumulates 7 patterns into in-pixel memory in pixel-parallel manner, and read them out in 7-bit parallel as digital values. This results in the read out time less than 1/7 compared with the conventional image sensors.

Fig. 2 shows the pixel circuit of the proposed image sensor. A pixel includes a comparator as the 1-bit A/D converter, and a 7-bit memory as the pattern accumulator. The threshold voltage of the comparator is adjusted via V_{REF}. The comparator is used to decide whether the pixel is bright or not, and it



(a) time encoded light pattern



(b) gray coded address

Fig. 1. The pattern used for this method

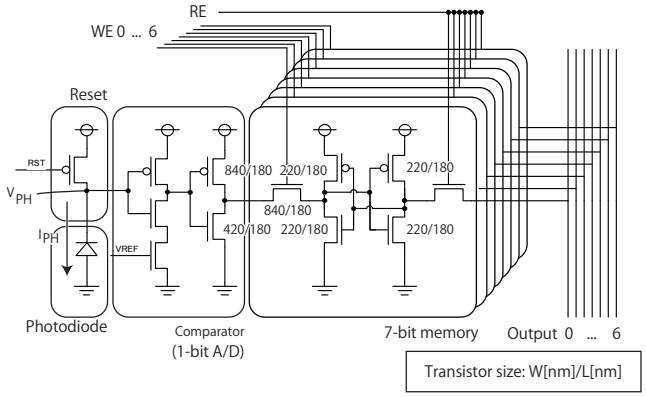


Fig. 2. Pixel circuit

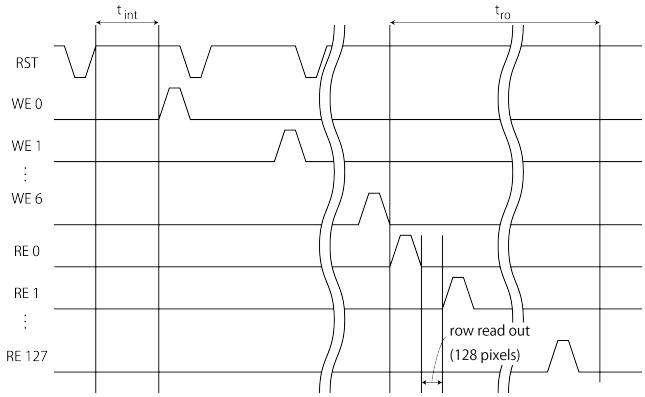


Fig. 3. Control signal wave form

outputs 1-bit digital data. This data is stored in the in-pixel memory instead of transferred immediately as in conventional image sensor. The containts of memory is read out through the 7 column lines in parallel when RE is asserted.

We employed the conventional 6-T SRAM for the memory as shown in Fig. 2. transistor sizes of the SRAM and the comparators are optimized according both to guarantee the write margin and to read out speed, as shown in Fig. 2. The optimization is carried out by the corner simulations.

The timing diagram of the image sensor operation is shown in Fig. 3. First, pulling RST to low sets the photodiode voltage to the supply voltage. After integrating for $t_{\rm int}$ seconds, one of the WEi is set to high and the comparator output is written to the corresponding memory. This process is repeated seven times to accumulate all the time encoded patterns, and after that the accumulated address in the i-th row is read out by setting REi to high.

Fig. 4 shows the chip photo. The chip is fabricated in a 0.18μm standard CMOS process without photodiode option. The pixel size is 13.82μm$\times$13.82μm and the fill factor is 28%. A 128$\times$128 pixel array is designed in a 2.5mm$\times$2.5mm die.

III. Measurement Results

Fig. 5 shows measurement setup with a video projector and image sensor board. We employed 8000 fps 1000 lumen DMD projector to examine maximum 3-D acquistion speed and 2800 lumen video projector to evaluate range error. We achieved 3-D range map acquisition speed of 92.7 rps with integration time of 1.5ms for 1000 lumen DMD projector. Read out time was 292μs per range map at system clock of 60MHz. For this setup, since the integration time is larget than the read out time, 3-D acquisition speed is limited by the integration time, which can be solved by the brighter video projectors.

Fig. 8 shows the acquired 3-D range map of a sphere at the integration time of 1.5 ms. Fig. 6 shows acquired 3-D range map of a hand in front of a plane, with integration time of 20ms.

The measured standard deviation and maximal range error at the integration time of 1.5ms with 1000 lumen projector were 4.6mm and 302mm, respectively, due to the errorneous pixels, for the target distance of 760-870mm.

Fig. 7 shows measured range map error of a plane at the distance around 760-870mm, with integration time of 10ms

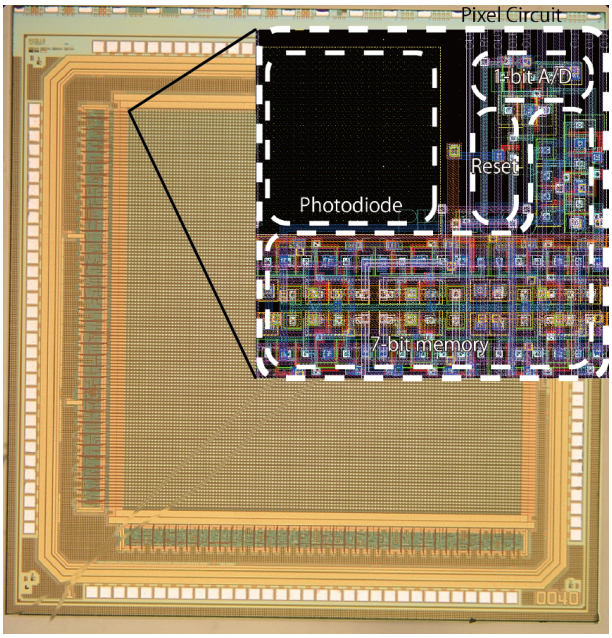


Fig. 4. Chip photo

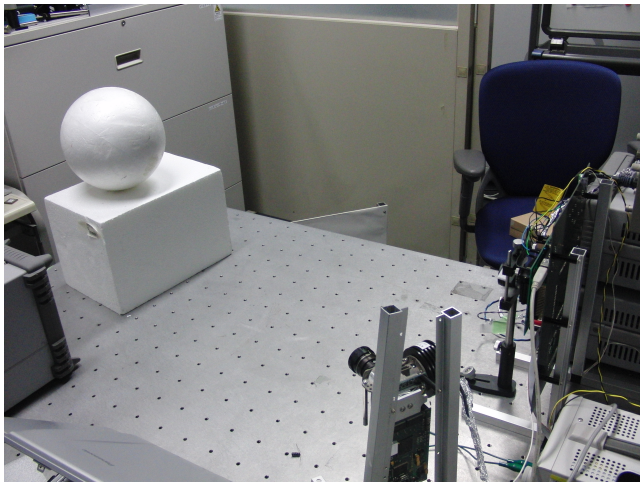
and 2800 lumen projector. For this case, measured standard deviation and max range error were 2.5mm and 9.0mm, respectively.

Table I summarizes the overall imager specifications and measurement results.

TABLE I
SPECIFICATION AND MEASUREMENT RESULTS SUMMARY

Chip Specification	
Supply voltage	1.8V
Die size	2.5mm × 2.5m
Array size	128 × 128
Pixel size	13.82 μm × 13.82 μm
Fill factor	28%
Photo diode	Nwell/Psub (7.2 μm × 7.2 μm)
No. of transistor	48 per pixel
Read out time (t_{ro})	292 μs per range map
Minimal integration time (t_{int})	1.5ms
Maximal range map per second	92.7
3-D measurement results	
Standard deviation	2.5mm @ 760-870mm, 2800lm
	4.6mm @ 760-870mm, 1000lm
Maximal error	9.0mm @ 760-870mm, 2800lm
	302mm @ 760-870mm, 1000lm

In the measurement above, an 8000 fps 1000 lumen projector is used to project the patterns. Since the integration time is larger than read out time, the speed bottleneck is the integration time. The easiest way to decrease the integration time is to use the brighter projector. In order to investigate the highest feasible range map rate for this image sensor, we have conducted an experiment where we applied light from a variable light source onto the image sensor and counted



(a) With 8000 fps/1000 lumen DMD projector

(b) 2800 lumen video projector

Fig. 5. Measurement Setup

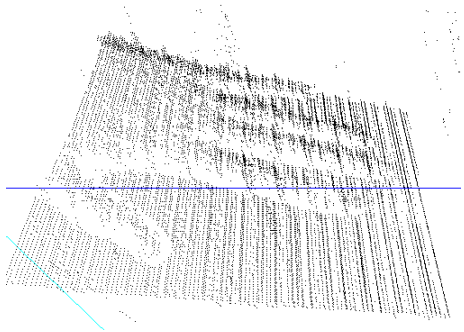


Fig. 6. Measured 3-D range map with 2800 lumen video projector, t_{int} = 20ms

the pixels that failes to recognize the light. Fig. 9 shows the integration time to realize error rate of 10% against illuminance at chip surface. Fit line is calculated by

$$t_{\mathrm{int}} = -K_1 \log\left(-\frac{K_2}{L} + 1\right) \tag{1}$$

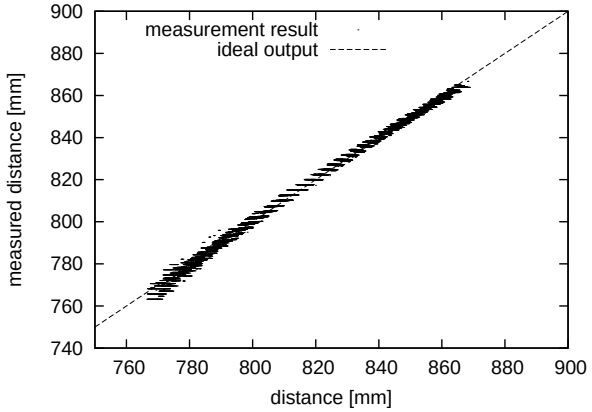


Fig. 7. Distance and error

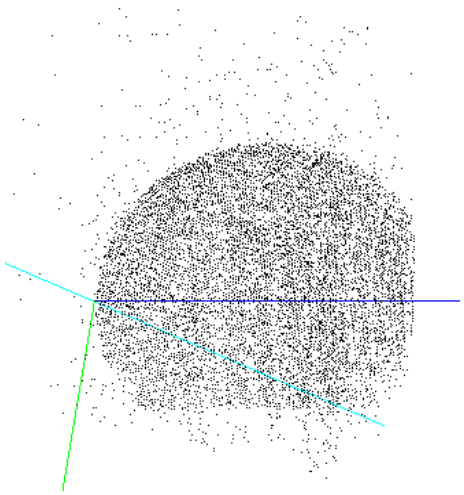


Fig. 8. Measured 3-D range map of sphere with 8000 fps / 1000 lumen DMD projector

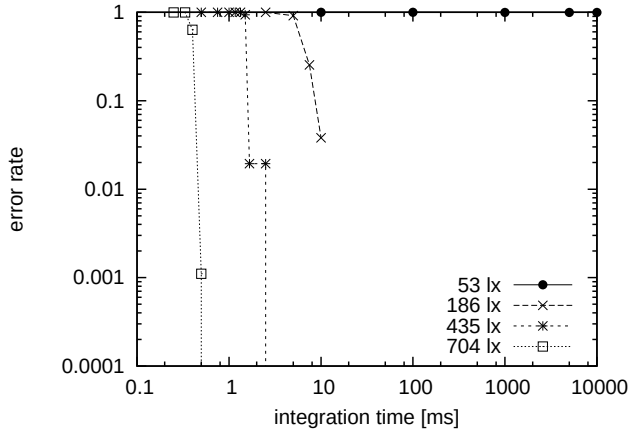


Fig. 9. Pixel error rate

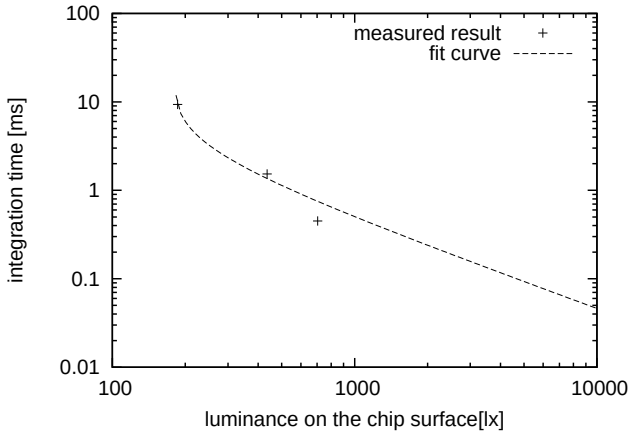


Fig. 10. Integration time that achieves the error rate of 10%

where L denotes illuminance, K_1 and K_2 are fitting parameters. In order to achieve the error rate of 10% when the integration is comparable to the read out time (292μs / 7 $\cong$ 40 μs), the light reflected on the target object should illuminate the image sensor chip at 10000 lux.

IV. Conclusion

An image sensor designed for 3-D surface reconstruction using gray-coded time-multiplexed structured light is proposed. By the use of in-pixel memory and parallel read out, we have achieved the read out time of 292 μs. The highest range map rate achieved was 92.7 μs due to the minimal integration time of 1.5ms. The standard deviation and maximum of the error was 10mm and 326mm, respectively, when a 1000 lumen DMD projector is used. This improves to be 2.5mm and 9.0mm when a 2800 lumen projector is used. Based on the analysis of the experiment using variable light source, higher range map rate will be realized by using brighter projectors.

Acknowledgment

This work is supported by VLSI Design and Education Center(VDEC), the University of Tokyo in collaboration with Cadence Design Systems, Inc., Mentor Graphics, Inc. and Synopsys, Inc.

The VLSI chip in this study has been fabricated in the chip fabrication program of VLSI Design and Education Center(VDEC), the University of Tokyo in collaboration with Rohm Corporation and Toppan Printing Corporation.

References

[1] M.Z. Brown, D. Burschka, and G.D. Hager. Advances in computational stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 25, No. 8, pp. 993 – 1008, 2003.

[2] Joaquim Salvi, Jordi Pagès, and Joan Batlle. Pattern codification strategies in structured light systems. *PATTERN RECOGNITION*, Vol. 37, pp. 827–849, 2004.

[3] Li Zhang, B. Curless, and S.M. Seitz. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *3D Data Processing Visualization and Transmission, 2002. Proceedings. First International Symposium on*, pp. 24 – 36, 2002.