World Automation Congress
Fourth International Forum on Multimedia & Image Processing

Seville, Spain
June 28th-July 1st, 2004

A High-Speed XGA 3-D Image Sensor And Its Applications

Yusuke Oike, Makoto Ikeda and Kunihiro Asada

A HIGH-SPEED XGA 3-D IMAGE SENSOR AND ITS APPLICATIONS

YUSUKE OIKE **MAKOTO IKEDA** **KUNIHIRO ASADA**

Dept. of Electronic Engineering
VLSI Design and Education Center
University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

ABSTRACT

We present a high-speed XGA 3-D image sensor based on the light-section method and its applications. The 3-D image sensor has the capability of 1024 × 768 3-D position data of a target scene in real time. It has been designed using a high-speed access method with an adaptive threshold circuit and column-parallel time-domain ADCs, and fabricated in 0.35 μm CMOS process. Sub-pixel position detection by the time-domain ADCs contributes to improve range accuracy within 0.1% range error. Simulation results show that the 3-D image sensor achieves 32.5k-fps position detection for the incident sheet beam. It corresponds to a real-time 3-D imaging system of 31.8 range maps/s. A range finding system with the 3-D image sensor was demonstrated. We show some examples of 3-D imaging applications, and also review performance requirements of a real-time system for the next step. The high-resolution and real-time 3-D imaging has the potential capability of future and attractive applications such as 3-D video contents.

KEYWORD: image sensor, 3-D image, range finding, object extraction, light section, XGA

1 Introduction

In recent years we often see 3-D computer graphics in movies and televisions, and handle them interactively using personal computers and video game machines. Furthermore new 3-D LCD panels are being developed, so we will require high-quality 3-D video contents in the near future. Some range finding methods were proposed for 3-D measurement, for example, the stereo-matching method, the time-of-flight method [1]–[5] and the light-section method [6]–[12]. However the state-of-the-art techniques [1]–[14] based on these methods have not achieved high-quality and real-time 3-D imaging for the future 3-D video applications yet.

The stereo-matching method provides a simple system configuration with two cameras, in other words, it provides a passive 3-D imaging system. It has been used for a kind of shape recording under an ideal measurement setup or a gesture recognition system with rough range accuracy. The time-of-flight techniques [1]–[5] are suitable for a real-time 3-D imaging system due to pixel-parallel range capturing. However the range accuracy is limited at a couple of centimeters. For example the range accuracy of Axi-vision Camera [5] is 1.7 cm at a distance of 2 m though it has the capability of real-time 3-D imaging with 1280 × 720 (HDTV) pixel resolution. It is suitable for an object extraction using range information.

To realize a high-resolution 3-D imaging system in short and middle range, the light-section method is most suitable. It has the potential capability of less than 1 mm range resolution at a distance of several meters. A system configuration of the light-section method is shown in Fig.1. The sheet beam is projected on a target object and the image sensor detects positions of the incident

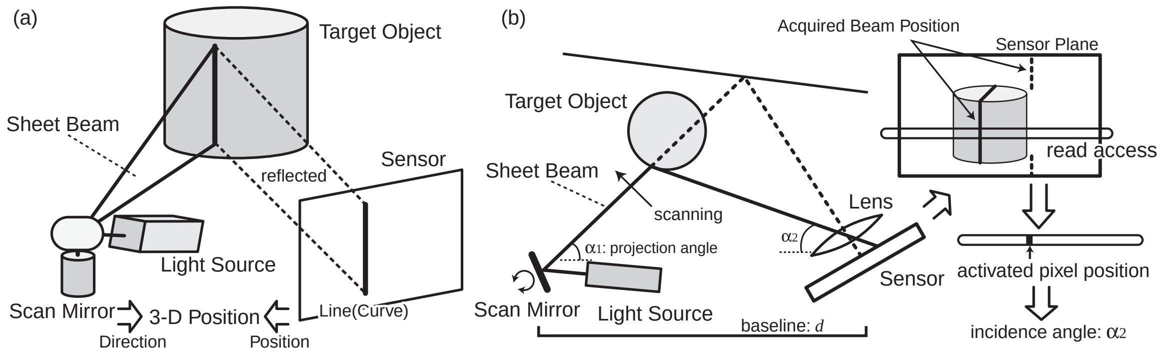


Figure 1: Light-section 3-D measurement system: (a) birds-eye view, (b) top view.

beam as shown in Fig.1 (a). The projection angle α_1 and the incidence angle α_2 are provided by the scanning mirror and the detected positions on the sensor plane respectively as shown in Fig.1 (b). Using triangulation, the distance of a target object is calculated by the baseline d and the angles α_1, α_2. A lot of position detections are required during beam scanning to get a range map of the target scene. In the light-section method, the range accuracy is proportional to the pixel resolution. On the other hand, the range finding speed is inversely proportional to the pixel resolution since much more position detections are required for a range map and the frame rate of position detection becomes slower. Therefore the pixel resolutions of [10]–[12] were limited at 64×64 to 192×124 pixels while they achieved a real-time 3-D imaging with around 1 mm range accuracy.

We proposed a high-speed access method [15] using an adaptive threshold circuit and column-parallel time-domain ADCs to realize a real-time 3-D image sensor with high pixel resolution. Then the first real-time 3-D image sensor with 640×480 (VGA) pixel resolution was developed [16], and the real-time VGA 3-D imaging system was presented [17]. In this paper, we present a high-speed XGA 3-D image sensor based on the light-section method. The new 3-D image sensor has the potential capability of 1024×768 3-D position data within 0.1% range error in real time. It has been designed using the high-speed access method and fabricated in 0.35 μm CMOS process. The preliminary tests for performance evaluation and some application examples of high-resolution 3-D imaging are shown.

2 Sensing Scheme and Circuit Configurations

The developed XGA 3-D image sensor employs the high-speed access method [15] using an adaptive threshold circuit and column-parallel time-domain ADCs. A real-time range finding with 1024 $\times$ 768 pixel resolution requires 30k lines/sec position detection of a projected sheet beam. Therefore a row access for the position detection must be carried out in 42 ns. It is difficult for a standard readout scheme to realize such a high-speed access. On the other hand, a pixel circuit should be compact as a standard CMOS imager to achieve high pixel resolution.

In the present sensor architecture, a pixel circuit is the same configuration of a standard CMOS image sensor, which consists of 1 photo diode and 3 transistors in pixel as shown in Fig.2. The pixel circuit has two operation modes of a 2-D imaging mode and a 3-D imaging mode. In the 2-D imaging mode, $MODE$ is set to low level. After an exposure period, pixels in the i-th row line are selected by SEL_i, and then the pixel value V_{pd} is read out through the source follower circuit as VAL_j as shown in Fig.2 (a). In the 3-D imaging mode, $MODE$ is set to high level. A column line VAL_j is connected to the ground level, and another column line COL_j is charged to V_{pc} by PRE. The voltage level of COL_j, which has been charged to V_{pc}, drops to the ground level according to the pixel value V_{pd} as shown in Fig.2 (b). The higher V_{pd} is, the faster COL_j drops. In other words, the outputs of bright pixels, which receive the strong incident beam, drop more slowly than the other

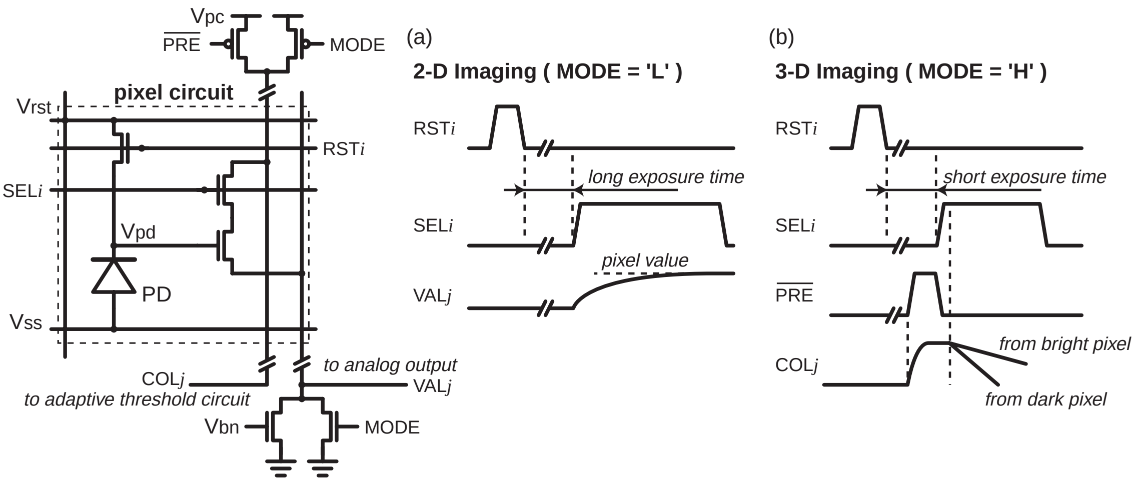


Figure 2: Pixel circuit configuration and operation: (a) 2-D imaging, (b) 3-D imaging.

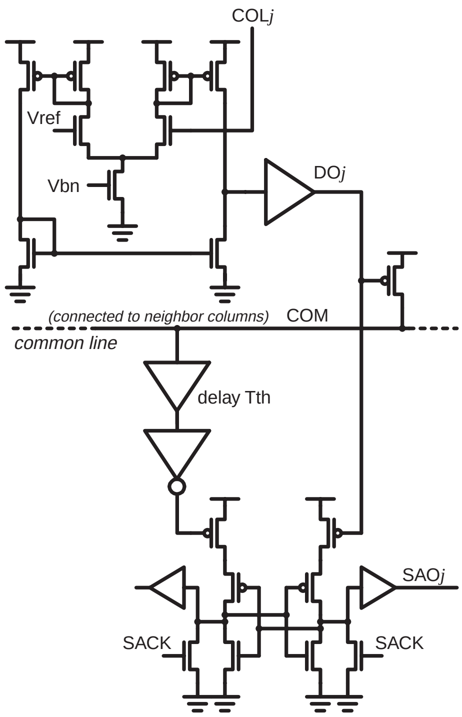


Figure 3: Schematic of adaptive threshold circuit in column.

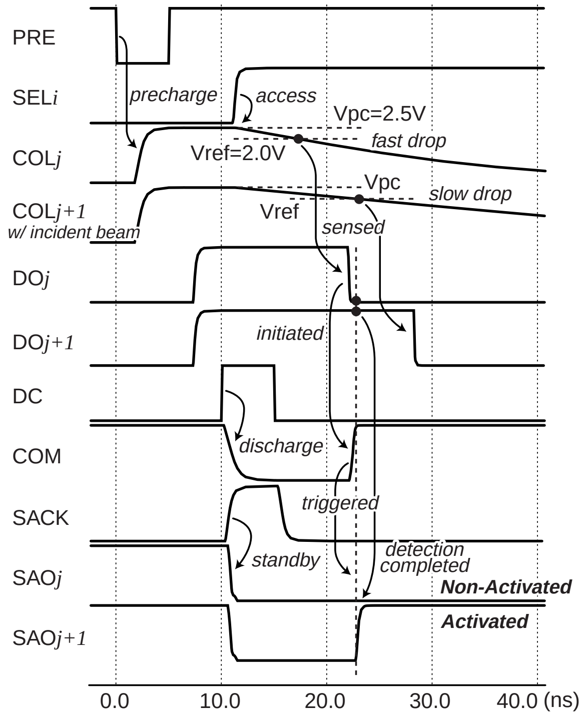


Figure 4: Simulation waveforms of activated pixel detection using 0.35 μm CMOS process.

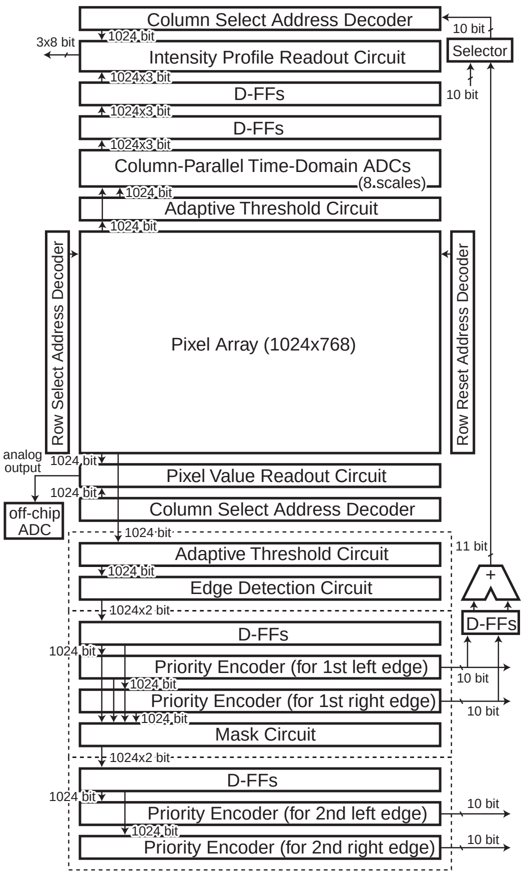
pixels. The pixel values are estimated in time domain during the row access. The slower-dropping outputs are detected as activated pixels.

To detect the slower-dropping outputs in time domain, each column has an adaptive threshold circuit as shown in Fig.3. The column output COL_j is compared with the reference voltage V_{ref} and sensed as the digital output DO_j. Fig.4 shows simulation waveforms of the designed XGA image sensor using 0.35 μm CMOS process. Here the precharge level V_{pc} is 2.5 V and the reference level V_{ref} is 2.0 V. A column output of a pixel with the strong incident beam drops more slowly as COL_{j+1} than the other outputs as COL_j in Fig.4. The fastest output signal D_{j+1} from the darkest pixel initiates the common trigger signal COM, which has been set to low level by DC at the same time of row select by SEL_i. The common line is connected to the neighbor columns, and then the trigger signal COM latches all outputs ($DO_0, DO_1...$) through a delay circuit. The delay time T_{th} represents the threshold margin ΔE_{th}. In the sensing scheme, the fastest transition timing of DO_j means the darkest pixel intensity E_{dark}. Therefore the adaptive threshold circuit detects a pixel with strong intensity over $E_{dark} + \Delta E_{th}$. The adaptive threshold scheme suppresses overall ambient light intensity and fluctuation of access speed in each row. The time-domain sensing scheme has the potential capability to acquire an intensity profile of activated pixels by using column-parallel time-domain ADCs [15, 16]. The intensity profile contributes to improve sub-pixel accuracy by gravity center calculation.

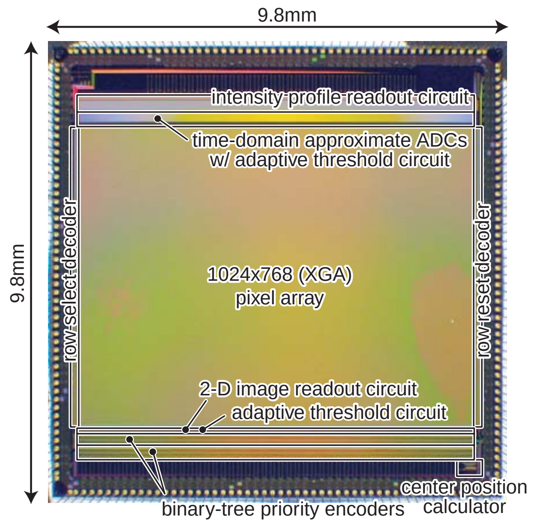
The activated pixels detected as SAO_{j+1} are transferred to the priority encoder via an edge detection circuit. The priority encoder consists of two parts, a priority decision circuit and an address encoder. The priority decision circuit keeps high speed in large input number due to a binary-tree structure and a compact circuit cell. The delay increases in proportion to $\log(N)$, where N is the number of inputs. It takes 11.1 ns with 1024 inputs in case of a circuit simulation using 0.35 μm CMOS process. The address encoder is implemented by dynamic logics and takes 14.0 ns in the same case.

3 Chip Implementation

The XGA 3-D image sensor has been fabricated in 0.35 μm CMOS process. Fig.5 (a) shows a block diagram of the 3-D image sensor. It consists of a pixel array, a row reset decoder, a row select decoder and a pixel value readout circuit with a column select decoder. Moreover the position detector is implemented in the bottom part of the sensor, which consists of the adaptive threshold circuit and



(a) Sensor Block Diagram



(b) Chip Microphotograph

(c) Chip Specifications

Process	0.35μm 3-M 2-P standard CMOS
Die size	9.8 mm x 9.8 mm
# pixels	1024 x 768 pixels (XGA)
# FETs	3.2M FETs
Pixel size	8.4 μm x 8.4 μm
Photo diode	pn junction (n -dif. / p-sub)
# FETs/pixel	3 FETs
Fill factor	29.02 % (6.3 μm x 3.25 μm)

Figure 5: XGA 3-D image sensor: (a) block diagram, (b) microphotograph (c) specifications.

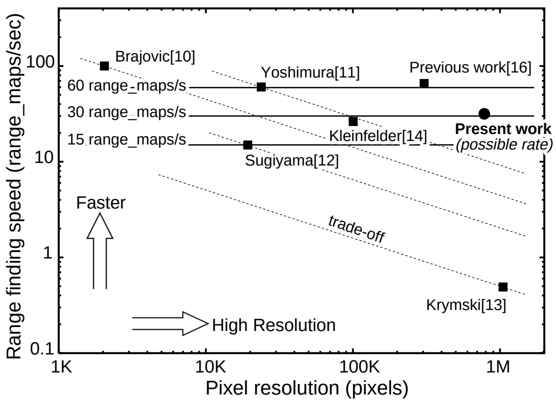


Figure 6: Range finding speed and pixel resolution.

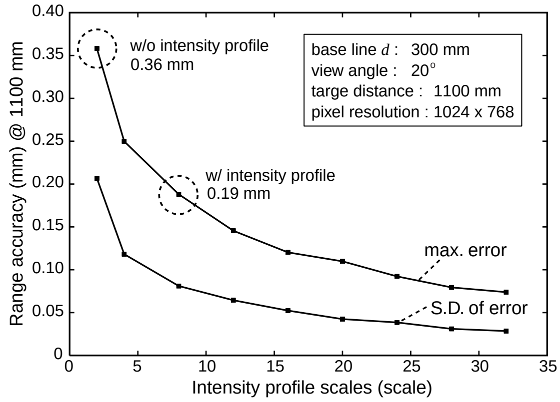


Figure 7: Possible range accuracy of XGA 3-D image sensor.

the priority encoders. The intensity profile detector with the column-parallel time-domain ADCs is implemented in the top part. The position detector of the bottom part is composed of three pipelined stages. The first stage is the adaptive threshold circuit and the edge detection circuit. It provides the left and right edge positions of consecutively activated pixels to the next stage. The second and third stages are the priority encoders, which provide the addresses of the left and right edges. The edge positions detected by the second stage are masked, and then the next position of activated pixels is encoded by the third stage. The intensity profile detector in the top part has the column-parallel time-domain ADCs to acquire an 8-scale intensity profile of activated pixels. The acquired intensity profile is read out selectively by the center position of activated pixels, which is calculated by the results of the position detector.

Fig.5 (b) and (c) show the chip microphotograph and specifications. The image sensor has 1024 $\times$ 768 pixels (XGA) in a 9.8 mm $\times$ 9.8 mm chip. The total number of transistors is 3.2M transistors. The pixel size is 8.4 μm $\times$ 8.4 μm with 29 % fill factor.

4 Performance Estimation

Fig.6 shows the range finding speed of the XGA 3-D image sensor estimated by a circuit simulation. The adaptive threshold circuit detects activated pixels in 30.0 ns after a precharge operation. Moreover the column-parallel time-domain ADCs need about 10.0 ns to acquire an intensity profile of activated pixels. The next priority encoder stage takes 25.1 ns and the intensity profile readout stage takes 27.8 ns. Therefore the cycle time of row access is 40.0 ns since each stage is pipelined. The XGA image sensor has the potential capability of 32.5k-fps position detection for the incident sheet beam. It corresponds to 31.8 range maps/s with 1024 $\times$ 768 3-D data. However it requires a high-speed sensor controller of 200 MHz, and also a strong beam intensity to activate pixels in a short exposure time of 30 μs.

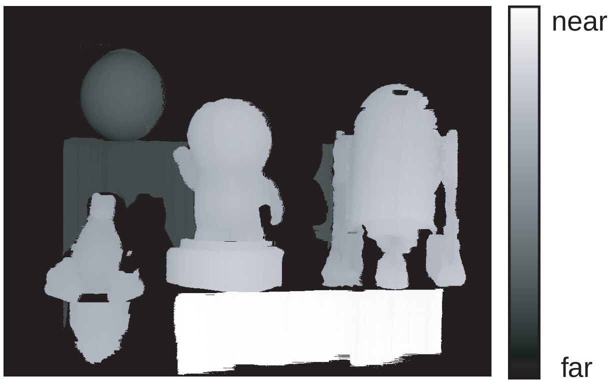
Fig.7 shows the possible range accuracy of the XGA 3-D image sensor in an ideal situation. A range accuracy of the light-section method depends on not only pixel resolution but also setup parameters, for example, a baseline distance between a camera and a beam source, a target distance, a view angle of camera and so on. In this simulation, the baseline distance is 300 mm, the target distance is 1100 mm, and the view angle is 20 degree. An intensity profile acquired by the time-domain ADCs can improve range accuracy according to the number of scales as shown in Fig.7. The maximum range error is 0.36 mm at a distance of 1100 mm in a normal position detection without an intensity profile. Furthermore the range accuracy achieves < 0.19 mm theoretically by using an 8-scale intensity profile provided by the time-domain ADCs.

5 Measurement Results and Applications

The developed XGA 3-D image sensor has been applied to a range finding system for the preliminary test. The measurement system is composed of a camera board with the sensor, a scanning mirror, a laser beam source of 300 mW and 665 nm wavelength, and a host computer. The host computer is equipped with digital parallel I/O boards of 2 MB/s for sensor control, an 8-bit A/D board for 2-D imaging, and a 12-bit D/A board for mirror scanning. The host computer controls the sensor and the sheet beam projector, acquires data from the sensor, and calculates 3-D position data. In the measurement setup, the viewing field of the camera is 400 mm $\times$ 300 mm at a distance of 1100 mm. The baseline between the camera and the sheet beam projector is 300 mm.

Fig.8 (a) shows a measured 2-D image of a target scene with 1024 $\times$ 768 pixels. Fig.8 (b) is a range maps reconstructed from the measured 3-D data. In the range map, brightness represents a distance from the camera. That is, the bright area is close to the camera and the dark area is far from the camera. A range finding system can be applied to various application fields. For example, an object extraction is realized promptly due to a range map as shown in Fig.8 (c). The object extraction method provides a depth-key system in stead of a chroma-key system. In this system, a blue-back screen is unnecessary. Therefore it can be applied to a realistic synthesizing system of real images and computer graphics, which has been reported in [5]. Fig.9 shows another application of the 3-D imaging system. The light-section 3-D measurement with high pixel resolution provides a precise wireframe model as shown in Fig.9 (a), which cannot be realized by the time-of-flight techniques [1]–[5] and the conventional light-section techniques [6]–[12]. A texture-mapped 3-D object is reconstructed by the wireframe model and the captured 2-D image as shown in Fig.9 (b).

(a) target scene (2-D image)
(1024 x 768 pixels)



(b) a range map

(c) object extraction (from 2-D image by range data)

Figure 8: Measured images and object extraction: (a) a 2-D image with 1024 $\times$ 768 pixels, (b) a range map, (c) object extraction using range information.

(a) a wireframe model

(b) a texture-mapped 3-D object

Figure 9: Reconstructed 3-D images: (a) a wireframe model, (b) a texture-mapped 3-D object.

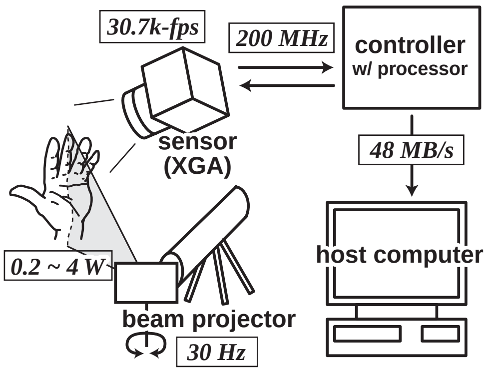


Figure 10: System requirements of real-time range finding.

A real-time range finding system with XGA pixel resolution has some system requirements as shown in Fig.10. First, a 3-D image sensor needs to detect an incident sheet beam on the sensor plane in 30.7k fps. Our developed 3-D image sensor has the capability of 32.5k-fps position detection. Second, the high-speed position detection requires 200 MHz sensor operation. It is difficult to be realized by direct control from a host computer. It will need a special controller implemented on FPGA/ASIC. It is better that the controller has the capability of triangulation processing. Furthermore a real-time 3-D imaging requires high-speed data rate over 48 MB/s between the controller and the host computer to transfer 1024×768 3-D data. In the high-speed position detection over 30k fps, an exposure time becomes 30 μs. Therefore a strong beam projection is required for pixel activation. A beam source of 4 W is necessary in the worst case because the sensitivity of the fabricated sensor is very low due to a PN-junction photo diode and no micro lens using a standard CMOS process. It has a potential to be suppressed up to around 200 mW by a special process for CMOS image sensors.

6 Conclusions

We presented a high-speed XGA 3-D image sensor based on the light-section method and its possible applications. It has been designed using a high-speed access method with an adaptive threshold circuit and column-parallel time-domain ADCs, and fabricated in 0.35 μm CMOS process. Simulation results show that these techniques realize a high-speed position detection of 32.5k fps with 1024×768 pixel resolution, and provide high sub-pixel accuracy within 0.1% range error. Some application examples of high-resolution 3-D imaging were demonstrated by a measurement system using the fabricated XGA 3-D image sensor. We have reviewed a possibility of a real-time 3-D imaging system using the XGA image sensor, and raised some practical problems as future works.

Acknowledgment

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

References

[1] R. Miyagawa and T. Kanade, "CCD-Based Range-Finding Sensor," *IEEE Trans. on Electron Devices*, vol. 44, no. 10, pp. 1648 – 1652, Oct. 1997.

[2] P. Gulden, M. Vossiek, P. Heide and R. Schwarte, "Novel Opportunities for Optical Level Gauging and 3-D-Imaging With the Photoelectronic Mixing Device," *IEEE Trans. on Instrumentation and Measurement*, vol. 51, no. 4, pp. 679 – 684, Aug. 2002.

[3] R. Jeremias, W. Brockherde, G. Doemens, B. Hosticka, L. Listl and P. Mengel, "A CMOS Photosensor Array for 3D Imaging Using Pulsed Laser," *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. of Tech. Papers*, pp. 252 – 253, 2001.

[4] A. Ullrich, N. Studnicka and J. Riegl, "Long-Range High-Performance Time-of-Flight-Based 3D Imaging Sensors," *in Proc. IEEE Int. Symp. 3D Data Processing Visualization and Transmission*, pp. 852 – 856, 2002.

[5] M. Kawakita, T. Kurita, H. Hiroshi and S. Inoue, "HDTV Axi-vision Camera," *in Proc. International Broadcasting Convention (IBC)*, pp. 397 – 404, 2003.

[6] V. Brajovic and T. Kanade, "Computational Sensor for Visual Tracking with Attention," *IEEE Journal of Solid-State Circuits*, vol. 33, no. 8, pp. 1199 – 1207, Aug. 1998.

[7] T. Nezuka, M. Hoshino, M. Ikeda and K. Asada, "A Position Detection Sensor for 3-D Measurement," *in Proc. European Solid-State Circuits Conf. (ESSCIRC)*, pp. 412 – 415, 2000.

[8] A. Gruss, L. R. Carley and T. Kanade, "Integrated Sensor and Range-Finding Analog Signal Processing," *IEEE Journal of Solid-State Circuits*, vol. 26, no. 3, pp. 184 – 191, Mar. 1991.

[9] M. de Bakker, P. W. Verbeek, E. Nieuwkoop and G. K. Steenvoorden, "A Smart Range Image Sensor," *in Proc. European Solid-State Circuits Conf. (ESSCIRC)*, pp. 208 – 211, 1998.

[10] V. Brajovic, K. Mori and N. Jankovic, "100 frames/s CMOS Range Image Sensor," *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. of Tech. Papers*, pp. 256 – 257, 2001.

[11] S. Yoshimura, T. Sugiyama, K. Yonemoto and K. Ueda, "A 48k frame/s CMOS Image Sensor for Real-time 3-D Sensing and Motion Detection," *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. of Tech. Papers*, pp. 94 – 95, 2001.

[12] T. Sugiyama, S. Yoshimura, R. Suzuki and H. Sumi, "A 1/4-inch QVGA Color Imaging and 3-D Sensing CMOS Sensor with Analog Frame Memory," *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. of Tech. Papers*, pp. 434 – 435, 2002.

[13] A. Krymski, D. Van Blerkom, A. Andersson, N. Bock, B. Mansoorian, and E. R. Fossum, "A High Speed, 500 Frames/s, 1024 × 1024 CMOS Active Pixel Sensor," *IEEE Symp. VLSI Circuits Dig. of Tech. Papers*, pp. 137 – 138, 1999.

[14] S. Kleinfelder, S. Lim, X. Liu, and A. E. Gamal, "A 10k frame/s 0.18μm CMOS Digital Pixel Sensor with Pixel-Level Memory," *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. of Tech. Papers*, pp. 88 – 89, 2001.

[15] Y. Oike, M. Ikeda and K. Asada, "A CMOS Image Sensor for High-Speed Active Range Finding Using Column-Parallel Time-Domain ADC and Position Encoder," *IEEE Trans. on Electron Devices*, vol. 50, no. 1, pp. 152 – 158, Jan. 2003.

[16] Y. Oike, M. Ikeda and K. Asada, "640×480 Real-Time Range Finder Using High-Speed Readout Scheme and Column-Parallel Position Detector," *IEEE Symp. VLSI Circuits Dig. of Tech. Papers*, pp. 153 – 156, 2003.

[17] Y. Oike, H. Shintaku, S. Takayama, M. Ikeda and K. Asada, "Real-Time and High-Resolution 3-D Imaging System Using Light-Section Method and Smart CMOS Sensor," *in Proc. IEEE Int. Conf. on Sensors*, 15–4, 2003.