

A 192x124 CMOS Image Sensor with Pixel-Parallel Temporal Computing Architecture

Shinichi Yoshimura, Toshinobu Sugiyama*, Kazuya Yonemoto**, Kazuhiko Ueda***

Sony-Kihara Research Center, Inc.,

*Imaging Device Company, SNC, Sony Corporation,

**DI Div., MNC, Sony Corporation,

***DSD Div., S&S Architecture Center, Sony Corporation

6-7-35 Kitashinagawa, Shinagawa, Tokyo 141-0001, Japan

Email: shinichi@krc.sony.co.jp

Abstract

A 192 x 124 pixel CMOS image sensor with pixel-parallel computational architecture enables video rate range sensing, motion detection, and digital image output. Each pixel has 4 current copier cells as frame memories and a chopper comparator for signal processing. This architecture allows computing a temporal response of the illumination at high frame rate during exposure.

I. Introduction

Real-time or video frame rate 3-D geometry sensing realizes motion capture of human gesture for computer graphics applications, object extraction with depth information for scene segmentation, etc. The conventional system with CCD camera and image processor is difficult to be utilized in these applications due to its capturing and computing speed. On the other hand, several CMOS range sensors have been developed to realize real-time range sensing [1-2]. They are designed to detect the timing of light stripe passing on the object, and distance to the object is computed based on the triangulation method. However, they have only a small number of pixels (less than 10k) and they cannot capture normal image. The goal of this project is to develop a CMOS image sensor with higher spatial resolution and this sensor must have the following functions; real-time range sensing, fast motion detection, and digital image output. Those functions are implemented on the pixel-parallel computation architecture, which operates at 48kfps maximum rate [3].

II. Architecture

The sensor architecture is shown in Fig.1 and the chip micrograph is shown in Fig.2. A pixel array contains 192(H) x 124(V) pixels.

124 pixel signals in a single column are read out from 192 column output circuits in parallel. The maximum data rate is 6.25MHz at 48kfps operation. The vertical scanner consists of 62 elements, and each of them feeds operating pulses to the 384 pixels simultaneously in both even and odd rows.

As shown in Fig.3, the pixel is composed of a sensing block, a frame memory block, a computing block, and an offset current generator. A sensing block has a photodiode (PD) and readout circuitry with current mirror. A hole-accumulated diode (HAD) is employed as a PD to reduce the dark current [4]. In a frame memory block, 4 current copier cells can store signal current for each frame in cyclic manner at maximum rate of 48kfps. They present the advantage of memorizing signal current accurately even if the memory transistors (Q1-Q4) have threshold voltage variations. The signals in frame memories are compared in a computing block and pixel output is in binary. At the chopper comparator, the offsets in the threshold of the inverters are compensated by its auto-zeroing function. That contributes to reduce fixed pattern noise (FPN) in this block. The offset current generator with an adjustable bias VL produces current offset during the comparison in the computing block. This offset makes comparator output stable by suppressing random signal variation existing even under the constant illumination.

III. Operation

For video frame rate 3-D sensing based on the triangulation with light stripe, thousands of frames are required during one scan of stripe to detect at which frame the pixel receives the brightest illumination. 4 current copier cells and a comparator operates peak detection. The

detected timing or frame of the illumination peak corresponds to the direction of the light stripe. Then the distance to the object is derived from geometric relation as illustrated in Fig.4. This peak detection is operated in pixel parallel, and entire depth information of the scene is obtained after a scan of light stripe. Fig.5 shows an example of timing diagram for pixel operation in a frame. At 'storing period', signal current from a sensing block is stored into one of four current copier cells (e.g. M1). At 'auto-zeroing period', signal currents in the two frame memories (e.g. M2 and M3) are readout with bias current I_z by adjusting VL to V_z . Another two frame memories, i.e. M4 and M1 are readout with bias current I_c by setting VL to V_c at 'comparing period'. The comparison result is readout to vertical signal line by Sd pulse. At the next frame, the combination of frame memories is shifted to meet the temporal sequence. By repeating the above operation with memory shifting, the comparator output transits from low to high, and the peak of illumination is detected as shown in Fig.4 (b). Even if the reflectance of the object surface is different, the relative variation of the brightness demonstrates the correct peak timing detection.

In order to detect the moving object, differentiation of the signal currents in 2 or 4 consecutive frames is evaluated in each pixel under the normal illumination. If the object moves, the variation of the brightness makes the transition of comparator output from low to high. This transition tells the timing when the motion is detected in each pixel during the operation.

For digital image output, 2 frame memories and a comparator work as an analog-to-digital converter (ADC). At the first frame, the fixed reference current is stored into one frame memory (e.g. M1) after resetting floating diffusion (FD) with VREF ($<VDD$). At the next frame, FD is reset with VDD and signal charges are transferred with TX. Signal current is stored in another frame memory (e.g. M2). In the following frames, signal charges are transferred without resetting FD, then integrated signal current becomes smaller and smaller and it is stored into the same memory (M2). At each frame, the reference current and signal current are compared. In the brighter illumination, a comparator output transits faster than darker illumination. Therefore, the detected frame number represents the inverse of illumination intensity (Fig.6).

IV. Experimental Results

Fig.7 shows the normal B/W image obtained by ADC operation (a), the depth image with range sensing operation (b), and the motion-detected result (c) to the same scene. The B/W image is grabbed at 48kfps during 67ms exposure. In this condition, more than 3,000 image resolution is achieved. In the depth image, the brighter regions are closer to the sensor than the darker areas. This scene is captured at 12kfps with 30Hz light stripe scanning. Only the 5 rotating blades in the scene (a) are extracted in the motion-detected result. The gradation of the extracted area represents the difference of the detected timing, and shows the direction of the motion.

Examples of depth resolution at 12kfps and 24kfps operation are depicted in Fig.8. At 400mm away from the sensor, the depth resolution of 1mm for 12kfps and 0.5mm for 24kfps are achieved.

Table 1 summarizes the prototype sensor characteristics.

Conclusions

The pixel-parallel temporal computing architecture with current copier cells and chopper comparator demonstrated real-time range sensing, motion detection and digital image output. The sensor is fabricated in a standard $0.35\mu\text{m}$ CMOS technology and operates at 3.3V power supply. The system with this sensor achieved the depth resolution of $500\mu\text{m}$ at 24kfps for range sensing and 12b digital image output resolution at 48kfps.

References

- [1] A. Gruss, et al., "Integrated Sensor and Range-Finding Analog Signal Processor," IEEE J. Solid-State Circuits, Vol.26, No.3, pp.184-191, Mar. 1991
- [2] A. Yokoyama, et al., "Realtime Range Imaging Using An Adjustment-free Photo VLSI," IEEE Proc. IROS, pp.1751-1758, Sep. 1994
- [3] S. Yoshimura, et al., "A 48kframes/s CMOS Image Sensor for Real-time 3-D Sensing and Motion Detection," ISSCC Digest of Technical Papers, pp.94-95, Feb. 2001
- [4] K. Yonemoto, et al., "A CMOS Image Sensor with a Simple FPN-Reduction Technology and a Hole Accumulated Diode," ISSCC Digest of Technical Papers, pp.102-103, Feb. 2000

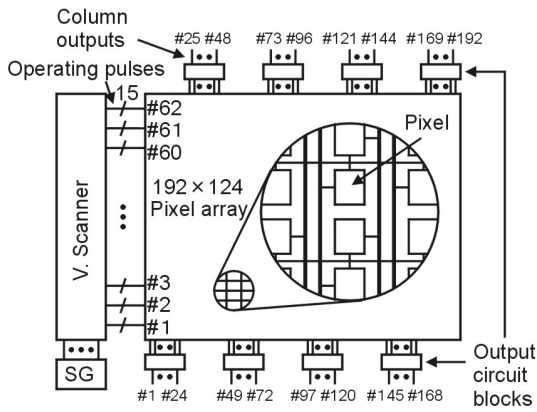


Fig.1: Sensor architecture

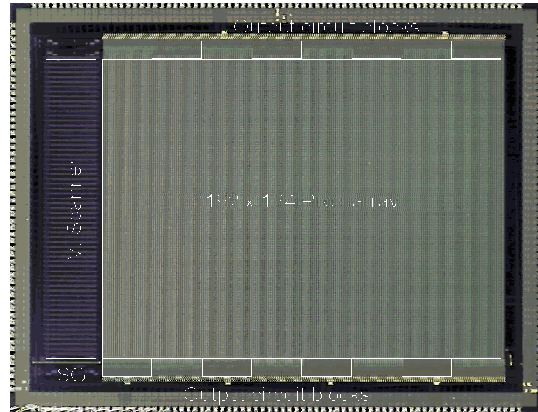


Fig.2: Chip micrograph

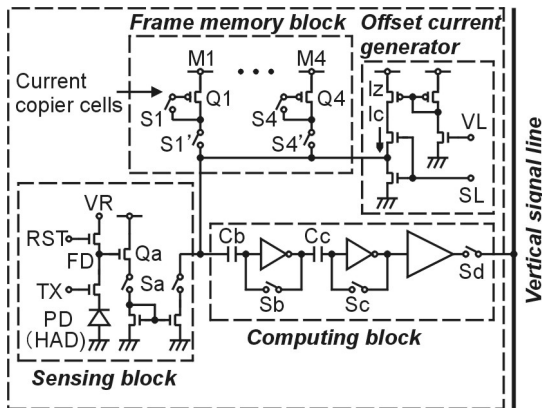
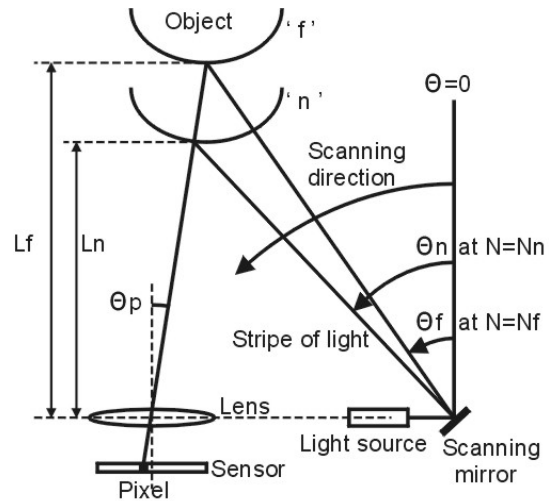
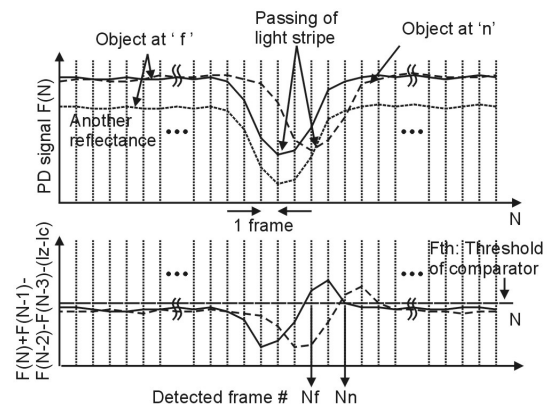


Fig.3: Unit pixel architecture



(a) Triangulation method with light stripe



(b) Peak detection of swept light stripe

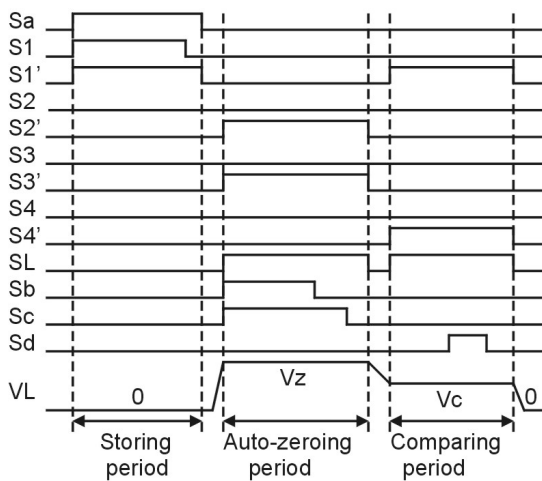


Fig.5: Example of timing diagram for pixel operation

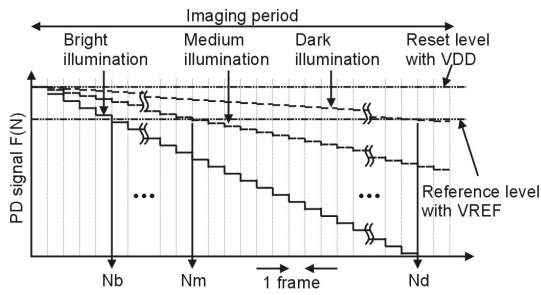


Fig.6: AD conversion for digital image output

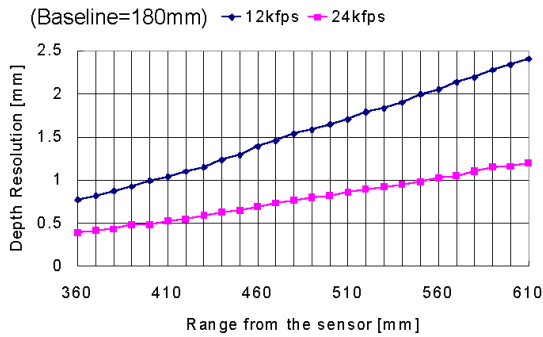


Fig.8: Example of depth resolution

Technology	0.35 μ m CMOS (1P, 3M)
Chip size	11.98mm(H) x 9.24mm(V)
Array size	192(H) x 124(V)
Pixel size	46.4 μ m(H) x 54.0 μ m(V)
Pixel fill factor	25%
Maximum frame rate	48kfps
Power supply	3.3V
Chip power	0.5W~2W
Depth resolution	500 μ m at 400mm (24kfps@30Hz)
ADC resolution	~12bits (48kfps@15Hz)

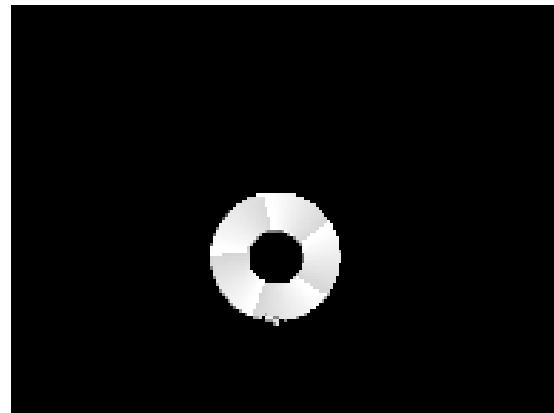
Table 1: Prototype specifications



(a) Normal B/W image (ADC: 48kfps@15Hz)



(b) Depth image
(Peak detection: 12kfps@30Hz)



(c) Motion detected image
(Differentiation: 12kfps@30Hz)

Fig.7: Examples of reproduced images