SMART IMAGE SENSOR FOR HIGH SPEED IN-FOCUS DETECTION

Takashi Yoshida, Arimitsu Yokota, Hideki Kashiyama, Takayuki Hamamoto

Department of Electrical Engineering, Tokyo University of Science
{yoshida,yokota,kashi,hamamoto}@isl.ee.kagu.tus.ac.jp

ABSTRACT

We propose a smart image sensor for the detection of an in-focus position using multiple focus images. This smart image sensor can generate an all-in-focus image by detecting an in-focus position, pixel by pixel. It can be also applied to the estimation of the distance between the sensor and objects. A prototype of 64x64 pixels has been fabricated. The proposed sensor can operate at more than 1500 fps, therefore it can be expected that the system is able to generate all-in-focus images at a full frame rate or faster speed.

1. INTRODUCTION

We have been investigating methods for estimating the distance between the sensor and the objects by focal position in real time. In this paper, we propose a new smart image sensor that can detect the in-focus position from multiple focused images on the sensor, pixel by pixel.

For distance estimation, some smart image sensors have been proposed [1]-[3]. Most of the systems are based on active sensing methods. They scan 3-D space by using laser light and detect peak intensity when the laser light is passing on each pixel. The image sensors have been used for detection of the peak intensity. On the other hand, our system adopts a passive sensing method for which the scanning by the laser light is not necessary. In order to achieve a high-speed operation of more than 1500 fps, we integrate the function of in-focus judgment into the CMOS image sensor. All processing for the detection is executed within the sensor.

In this paper, we describe the detection algorithm for our sensor, and show some results from simulation. Then, we explain the circuit design of the new smart image sensor and its evaluation.

2. ADVANTAGES OF THE SENSING SYSTEM USING A SMART IMAGE SENSOR

We propose the sensing system for distance estimation based on in-focus judgment by using smart image sensor. The advantage of our system is its high-speed operation by peripheral circuits and instruments. Assuming that we use a normal image sensor to detect the in-focus position from multiple focus images, all images need to be sent to the back-end processor. This requires a wide bus between the image sensor and the back-end processor, and it tends to be a bottleneck of the system if image resolution is expanded or the system is operated at higher speed. Supposing the image resolution is 640x480 pixels with 8bit depth, and 50 different focuses are used for estimation, then total quantity Q_T in order to obtain one all-in-focus image is

$$Q_T = 640 \times 480 \times 8 \times 50 = 122.9\text{Mb} \quad (1)$$

In the case of 30fps(30 all-in-focus results per second), the system requires $Q_T \times 30 = 3.689\text{Gbps}$ throughput for the bus. Assuming that we use 32bit PCI bus (133.3MBps), the system cannot perform more than 8.68fps.

$$fr_C = 133.3 \times 8 / Q_T \approx 8.68\text{fps} \quad (2)$$

On the other hand, our proposal doesn't have this bottleneck, and it's relatively easy for the smart sensor to perform more than 1500fps[4]. It means that we can get 30 results per second under the same condition.

$$fr_P = 1500 / 50 = 30\text{fps} \quad (3)$$

Moreover, most of the processing can be done on the sensor without any interim outputs. It is reasonable to say that the smart image sensor with the function of in-focus judgment has a good advantage, especially for high-speed operation.

3. ON-SENSOR IN-FOCUS DETECTION

3.1 Detection Algorithm

In order to judge an in-focus position in the sensor, an algorithm should be considered to meet the following requirements for implementation.

- Judgment from one pixel or its neighbors
- Small memory
- Simple calculation

Then we adopt an evaluation value "d" based on a derivation of Laplacian, which uses the image intensity of 4 neighboring pixels of an attended pixel. Equation (4) shows the value "d". Although this evaluation value is

0-7803-8554-3/04/$20.00 ©2004 IEEE.

rather simple, it can be calculated by using small parallel analog circuits and its speed is very high.

$$d = \left| l_{F(x+1,y)} + l_{F(x-1,y)} - 2l_{F(x,y)} \right| + \left| l_{F(x,y+1)} + l_{F(x,y-1)} - 2l_{F(x,y)} \right| \quad (4)$$

l: pixel intensity, F: focal position, x, y: pixel address

Fig.1 shows example results of the "d" value versus focal position when the actual in-focus position is 24. At the edge part, the value "d" surges near the in-focus position because the difference of image intensity between a center and its neighboring pixels is considerable. With this value "d", we can detect in-focus position. On the other hand, the value "d" at the DC part doesn't show any typical peak point. Therefore we can also detect that the pixel is on the edge part or DC part. In a real system, we use threshold "d_{th}" in the sensor to decide whether the pixel is useful for reliable detection. This threshold is controllable from outside of the sensor.

3.2 Evaluation of our proposed method

We evaluate our proposal for detection of the best in-focus position at each pixel. We use 81 multiple focus images of which examples are shown in Fig.2. These are obtained by changing the focal position little by little repeatedly. In the images, the surface of two objects is flat and the distance from the sensor is different. All-in-focused image has been created with the accurate depth information and frequency analysis for comparison.

Fig.3(a) shows the result of the computer simulation when the threshold "d_{th}" is set to 0. In this example, PSNR is 25.4dB to the reference. By setting the "d_{th}" properly and interpolating in-focal data outside of the sensor, we could expect better quality. In Figs.3(b)(c)(d), the results of d_{th}=65 are shown. Fig.3(b) is only set to the "dth" and no interpolation has been made. The DC part is seen as black. In Figs.3(c) and (d), the DC part is interpolated. Fig.3(c) shows the in-focus position corresponding to the distance information. Therefore, the nearer object is shown brighter. Fig.3(d) is all-in-focus image with interpolation[5]. In this example, PSNR is 31.4dB.

4. OUTLINE OF OUR SENSING SYSTEM

Fig.4 shows the outline of the proposed in-focus detection system. It consists of the smart image sensor, a variable focus lens system, a control LSI and a PC only for user interface. The control LSI sends commands to the sensor and 6bit focal information to the sensor and the lens system. By this scheme, it is easy to synchronize the sensor and the lens. In addition, flexible operations can be performed when the programmable LSI such as FPGA is used as a controller.

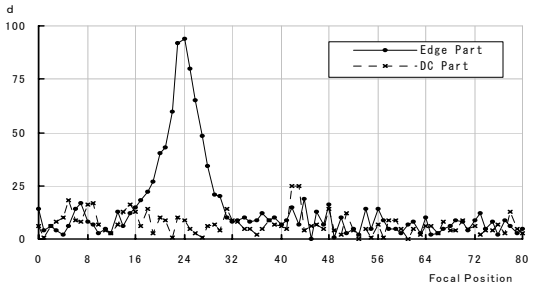


Fig.1 Example of d value versus focal position

(a) In focus right object (b) In focus left object (c) Out of focus

Fig.2 Sample of multiple focus images

(a) Threshold 'd' = 0 (b) Threshold 'd' = 65

(c)'d' = 65+interpolation Distance image (d)'d' = 65+interpolation All-in-focus image

Fig.3 Result of the computer simulation

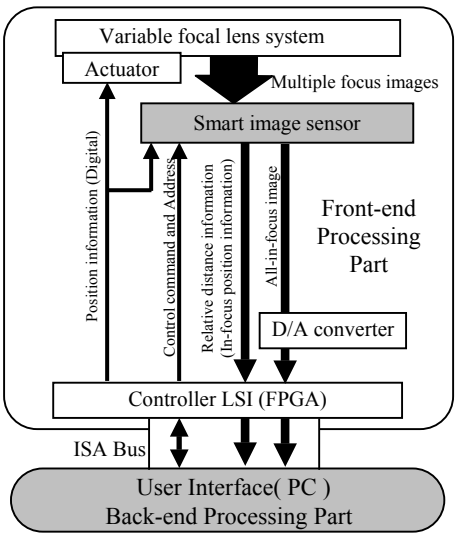


Fig.4 System outline

In this system, the interface computer initially sends the start signal to the controller LSI firstly. When the controller receives the signal, it sets the variable lens' focal position and sends the focal information to the sensor. After setting the focal position, the sensor captures an image and calculates the "d" value. Comparing the "d" value with former one, it holds the larger "d" value for every pixel and its focal information. This is the end of the process for 1 focal position. The system repeats this procedure until all images have been processed. After all of the multiple focus images are evaluated, the all-in-focus image is stored in the memory of the sensor. Finally, the all-in-focus image and in-focus data is output by the controller LSI and the image is shown at the display.

If "d_{th}" is set to divide the image into the DC part and the edge part, data at the DC part is interpolated into the controller LSI and/or back-end computer, then the interpolated data is redirected to the sensor memory. The system rescans the all-in-focus images with the interpolated data. Now we are planning to implement the DC interpolation into the controller LSI.

5. IMAGE SENSOR WITH THE FUNCTION OF IN-FOCUS DETECTION

5.1 Outline of the prototype sensor

We designed the smart image sensor by using 2-poly 3-metal 0.6um CMOS rule. Table 1 shows the specifications of the sensor. Fig.5 shows the block diagram of the sensor. This sensor consists of a transducer, a processor and memory elements. The pixels of each column share one processing circuit. Most of the sensor elements such as the 3 rows of Prepared Memory, Processor and Comparator for "d" value, memory array for "d" and an all-in-focus image are analog circuits. Only the focal information memory and the comparator for focal position are digital circuits. The sensor has 6bit vertical and horizontal address decoders for random access to the PD array and Memory arrays. The sensor is designed to be controlled from outside with a 4bit command, therefore it has a 4bit command decoder. The sensor can output analog raw images, the analog all-in-focus image, analog "d" values and digital focal position information, and it can be opperated at more than 1500 fps.

5.2 Processing flow in the sensor

The following is the detection process in the sensor. Firstly, an image is captured on the PD array. Then using the Address Decoder Y, 3 rows of the images are stored into the prepared memory. The "d" value is calculated by the center and 4 neighboring pixels. Therefore, not only a row for center, right and left pixels but also the upper and under rows need to be stored for other 2 neighbors. After this process, the Processor computes "d" values. With the column parallel architecture the "d" values of the pixels in one row are calculated at once. Then, the latest "d" value and older one need to be compared at the comparator, and if the latest value is larger, the "d" memory and position memory would be overwritten. At the same time, this comparator evaluate whether the "d" is larger than threshold "d_{th}". Threshold "d_{th}" is inputted from t_{in} in Fig.5. When the "d" value is overwritten, position memory also should be overwritten. The focal position signal is always sent from the controller LSI as a 6bit digital value. The sensor holds the information into the digital position memory. Digital comparator is required to

Table1 Specifications for the Sensor

number of pixels	64 x 64
chip size [mm^2]	2.78 x 7.49
element size [um^2]	
transducer	25.0 x 25.0
preparatory memory	62.3 x 25.0
ALU	99.8 x 25.0
analog comparator	90.5 x 25.0
digital comparator	106.9 x 25.0
d memory	18.5 x 25.0
f memory	41.6 x 25.0
i memory	12.6 x 25.0
number of transistors	
transducer [trs./pixel]	6
preparatory memory [trs./line]	18
ALU [trs./line]	57
analog comparator [trs./line]	12
digital comparator [trs./line]	73
d memory [trs./pixel]	6
f memory [trs./pixel]	24
i memory [trs./pixel]	5
fill factor [%]	24.06

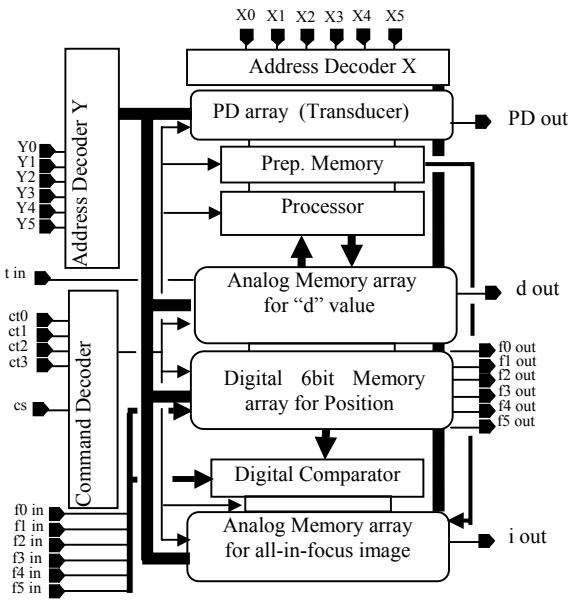


Fig.5 Block diagram of the sensor

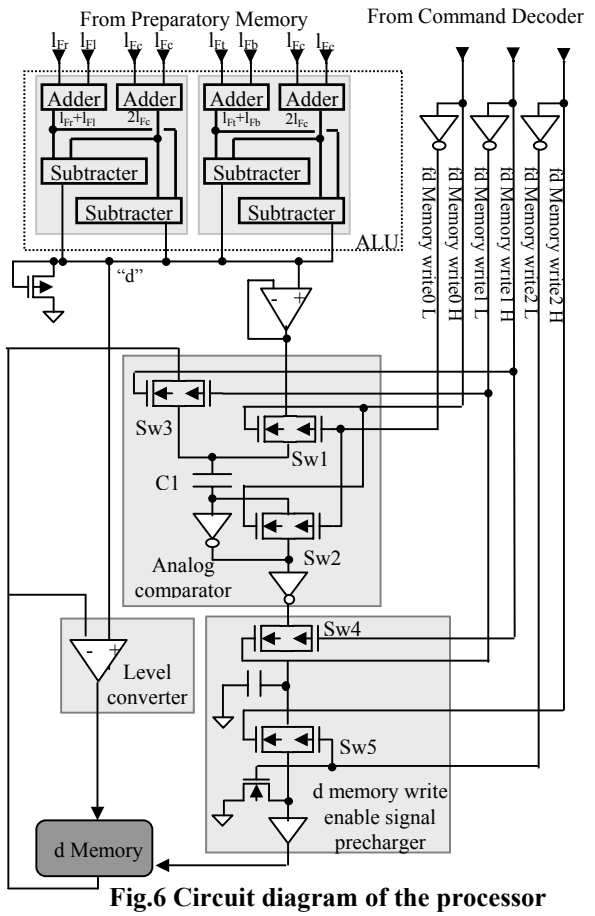
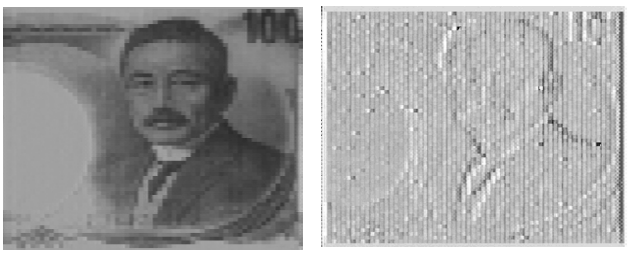


Fig.6 Circuit diagram of the processor



(a)Analog pixel value (b) "d" value

Fig.7 Example output from the sensor

compare the in-focus position and inputted position information when the sensor generates the all-in-focus image from the interpolated focal position.

5.3 Processing circuits for in-focus detection

Fig.6 shows the block circuit diagram of the processing part. This part contains ALU for the "d" value calculation, the comparator, the level converter and the "d" Memory write Enable Signal Precharger.

With the ALU in Fig.6, "d" value is calculated from the intensity value of the center pixel (l_{Fc}) and 4 neighbors (l_{Fr}, l_{Fl}, l_{Ft}, l_{Fb}) by 4 adders and 4 subtractors. Then, older "d" value from "d" Memory and newer "d" are compared at Analog Comparator. At first, the Sw3 is open and older "d" value is charged in the capacitor C1. In this moment, the Sw1 is closed and the Sw2 is open. Then when the Sw1 is open and the Sw2 is closed, the values are compared. When the "d" value from the ALU is larger, output of the Analog comparator would be HIGH; otherwise it would be LOW. Then at the Sw4 and Sw5, write enable signal is sent to the "d" Memory. Only if the current "d" value is larger than older one, the "d" memory is replenished by the current "d" value using the level converter. The level converter can adjust the memorized voltage by balancing the input value of "d" Memory with the output value.

5.4 Verification of prototype chip

The prototype of the sensor has been fabricated. The sensor verification is in progress. Fig.7 shows the analog pixel value output and "d" value output from the sensor when the in-focus image is captured. So far, we confirmed that the all test circuits of sensor elements (PD, Processor, Buffer Amp. and Analog Comparator) work sufficiently.

6. CONCLUSION

In this paper, we present a high-speed smart image sensor for in-focus detection by using multiple focused images. The proposed sensor has the circuits for detection of the in-focus position, pixel by pixel. With our proposal, it enables to get an all-in-focus image at a high-speed without any bottleneck.

We describe the processing algorithm for the sensor and simulate the performance. We explain the circuit design of the prototype, which has an analog column-parallel architecture. The prototype is now under verification. As the future work, other parts of the system such as controller need to be designed.

REFERENCES

[1] A. Makynen, et al., "A binary photodetector array for position sensing," Sensors and Actuators A, Vol.65, pp.45-53, 1998.

[2] M. de Bakker, et al., "A Smart Range Image Sensor," in Proc. of ESSCIRC, pp.208-211, 1998.

[3] Y. Oike, et al., "Smart Sensor Architecture for Real-Time High-Resolution 3-D Measurement and Its Implementation," ITE Technical Report Vol.26, No.41, pp.37-40, 2002.

[4] T.Hamamoto, et al., "Real-time objects tracking by using smart image sensor and FPGA" IEEE Int. conf. on Image Processing, III, pp.441-444 2002

[5] Serge Beucher, "Geodesic Reconstruction, Saddle Zones & Hierarchical Segmentation," Image Analysis & Stereology, Vol. 20, No.3, pp.137-141, 2001.