

# CMOS Image Sensors with Video Compression

Shoji Kawahito Yoshiaki Tadokoro

Akira Matsuzawa

Department of Information and Computer Sciences  
Toyohashi University of Technology  
Toyohashi 441  
Tel: +81-532-44-6755  
Fax: +81-532-44-6757  
e-mail: kawahito@signal.tutics.tut.ac.jp

Advanced LSI Technology Development Center  
Matsushita Electric Industrial Co., Ltd.  
Moriguchi 570  
Tel: +81-6-906-4906  
Fax: +81-6-906-3851  
e-mail matsu@vdr1.src.mei.co.jp

**Abstract—**

**This paper describes CMOS image sensors integrating video compression circuits. The on-sensor compression is particularly useful for the low-power design of moving picture compression hardware, which is demanded especially in the mobile computing and telephony. Recent progress of the CMOS image sensor technology allows us to realise the integration of high-performance image sensor and computational functions on a chip. An example of the CMOS image sensor with compression is described. Prospects of the CMOS image sensors with moving picture compression are discussed.**

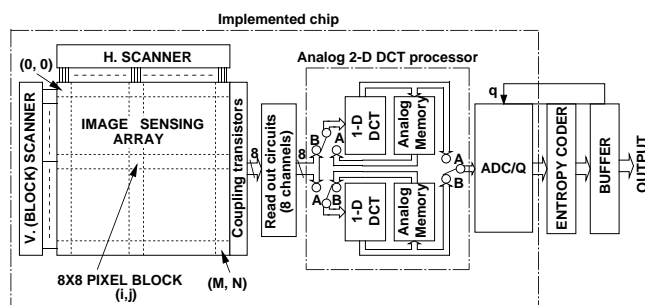
## I. INTRODUCTION

Recent progress in CMOS-based image sensors, especially in image quality, is creating new opportunities to develop a low-cost low-power one-chip digital video camera which have digitizing, signal processing and image compression functions[1][2][3]. Image compression is the most expensive hardware in digital video camera system, and on-sensor, or focal-plane compression leads to the cost-effective low-power implementation.

In this paper, CMOS image sensors for video compression are described. A few types of image sensor chips with compression have been developed such as a CCD-based image sensor with lossless image compression [4], a computational image sensor using conditional replenishment [5], and a CMOS image sensor using analog 2-D DCT based image compression[6]. Moving picture compression requires extremely high computational power which leads to the large power dissipation, and causes the difficulty to treat a moving picture on mobile terminals. The most important feature of the focal-plane image compression is the power reduction of the total video camera system. The CMOS image sensor with the analog 2-D DCT based image compression demonstrated that the on-sensor compression allows to reduce the power required for the image compression. This approach employs the intra-frame coding scheme, and the compressed data can be compatible

to the image coding standard of motion JPEG. This paper focuses on the CMOS image sensor with analog DCT based compression.

In sections II and III, the design and the experimental results, respectively, of the developed CMOS image sensor with analog DCT based compression are described. Section IV discusses the possibility of the CMOS image sensor with highly-efficient video coding such as MPEG2 compatible image compression.



**Fig. 1 CMOS image sensor with the 2-D DCT based ompression**

## II. DESIGN OF CMOS IMAGE SENSOR WITH 2-D DCT BASED COMPRESSION

### A. Architecture

Figure 1 shows the block diagram of the CMOS image sensor with focal-plane image compression. The photo diode array is divided into blocks, where each block consists of  $8 \times 8$  pixels. The accumulated signal charge in the pixel is converted into a voltage signal at the readout circuits. The analog 2-D DCT processor computes the 2-D DCT in parallel as a unit of column. To compute an  $8 \times 8$ -point 2-D DCT, 8-point parallel data are read out for 8 times from an image sensor block. In the ordinary image coding process, the DCT coefficient (the output of the 2-D DCT processor) is quantized by a digital quantizer. In Fig. 1, the quantization for the dynamic control of the output pulse rate using a quantization factor( $q$ ) is performed by the AD converter itself. The final step is

the entropy coding section. After the digitization, various image data coding method can be utilized. For example, for the efficient data compression of the moving picture, a simple inter-frame coding can be embedded in Fig. 1.

### B. Block access sensor

In order to interface to the subsequent analog 2-D DCT processor, an 8-channel parallel readout CMOS image sensor is designed, where 8 pixel signals are read out at the same time. Fig. 2 shows the sensor and the readout circuits. Each pixel contains one photo diode and two MOS transistors for row and column selection. This type of two-transistor selection scheme was used for TSL(Transversal Signal Line) MOS-type image sensor to obtain low smear and low sampling noise [7]. Row-selection signal is connected to 8 rows in the pixel block, and 8 pixels as a unit of column is read out in parallel. The size of the pixel is  $16.1\mu\text{m} \times 16.1\mu\text{m}$ , and the fill factor is 56.5%. The readout circuits consist of a front-end amplifier and the successive fully-differential amplifier based on the switched capacitor technique. The signal charge from a pixel is converted into a voltage with a 100fF capacitor. The conversion gain is  $1.6\mu\text{V}/e^-$ . The circuits are controlled by two phase non-overlapping clock. This scheme acts as correlated double sampling (CDS) circuits, where the  $1/f$  noise of the front-end amplifier are greatly reduced. The offset voltage deviation of the front-end amplifier is also reduced by the CDS circuits, so that a simple and small source-common amplifier can be used. However, the dominant random noise component is the  $kTC$  noise caused by the relatively large capacitance of the common signal line and the coupling transistor. The CDS technique shown in Fig. 2 is not always effective to reduce the  $kTC$  noise caused by the signal line capacitance.

A reference voltage input,  $V_{REF}$  is for voltage shift to increase the signal dynamic range in readout circuits and the analog 2-D DCT processor.

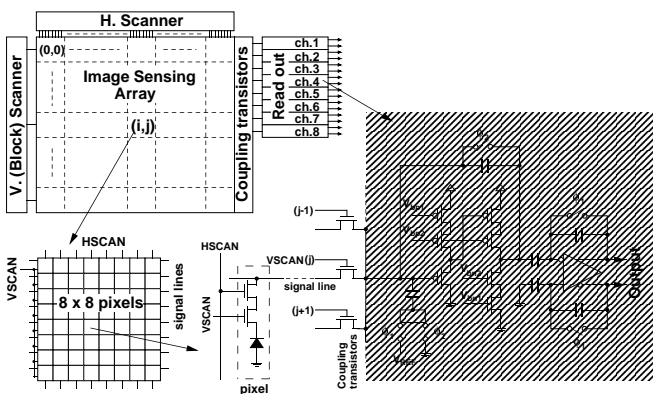


Fig. 2 Sensor readout circuits.

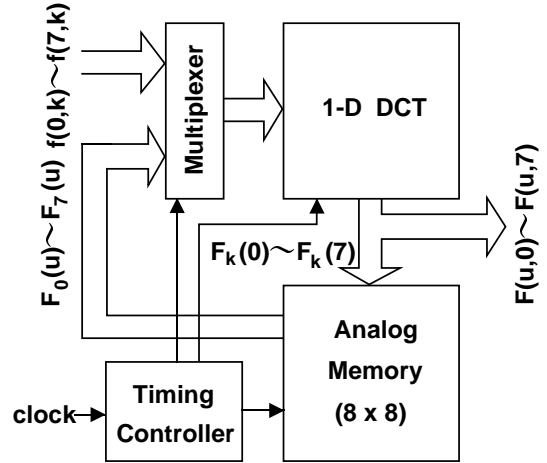


Fig. 3 Block diagram of the analog 2-D DCT processor.

### C. Analog 2-D DCT processor

The 2-dimensional discrete cosine transform (2-D DCT) of  $8 \times 8$ -size data  $\{f(j,k) : j = 0, 1, \dots, 7, k = 0, 1, \dots, 7\}$  is often calculated by the following one-dimensional DCT (1-D DCT) equations as

$$F_k(u) = \frac{C(u)}{2} \sum_{j=0}^7 f(j,k) \cos \frac{(2j+1)u\pi}{16}, \quad (1)$$

and

$$F(u,v) = \frac{C(v)}{2} \sum_{k=0}^7 F_k(u) \cos \frac{(2k+1)v\pi}{16}, \quad (2)$$

where  $u = 0, 1, \dots, 7, v = 0, 1, \dots, 7$ , and

$$C(\omega) = \begin{cases} 1/\sqrt{2} & (\omega = 0) \\ 1 & (\omega \neq 0) \end{cases}$$

$F_k(u)$  is an intermediate result for the calculation of 2-D DCT. This technique is called row-column decomposition. We first compute the 1-D DCT of Eq. (1), and the intermediate results are stored in  $8 \times 8$ -cell memory. After transposing the memory data, we can get the final results of 2-D DCT by computing 1-D DCT of Eq. (2) using the stored intermediate results. Figure 3 shows the block diagram of the analog  $8 \times 8$  point 2-D DCT processor which consists of a 1-D DCT processor and an  $8 \times 8$ -cell analog memory. The first column ( $k=0$ ) of the input data  $f(j,k)$ , ( $j = 0, 1, \dots, 7$ ) is first given to the 8-point 1-D DCT processor, the intermediate results,  $F_0(u)$ , ( $u = 0, 1, \dots, 7$ ) are obtained by calculating Eq. (1), and they are stored in the first row of the analog memory array. Similarly, all the intermediate results,  $F_k(u)$  are obtained by repeating this process for  $k = 1, 2, \dots, 7$ . The first column of the intermediate results,  $F_k(0)$ , stored in the analog memory is given again to the 1-D DCT processor, and the first column of 2-D DCT results,  $F(0,v)$  are calculated according to Eq. (2). Similarly, all the 2-D

DCT results,  $F(u, v)$ , are obtained by repeating also for  $u = 1, 2, \dots, 7$ . To obtain all the 2-D DCT coefficients, the total of 16 steps is necessary. Using two 2-D DCT cores as shown in Fig. 1, the throughput can be doubled and the sensor signal can be read out continuously. In phase A, the lower-side DCT core processes the input image, while the upper-side DCT core processes the intermediate results stored in the analog memory. In the next phase (phase B), the upper-side DCT core processes the input image, while the lower-side DCT core processes the intermediate results. The phase A and the phase B are repeated alternatively.

The 1-D DCT processor performs weighted summation, according to Eqs. (1) or (2). Fig.4 shows the analog 1-D DCT processor based on the SC circuits. In Fig. 4, 32 additions and 32 multiplications are performed in parallel. The SC circuit technology is useful for the precision design of the coefficient multiplier. A fully differential (rail-to-rail) scheme is used to reduce the clock feedthrough error and digital pulse noises. The cosine coefficient matrix of  $8 \times 8$  point 2-D DCT is given by

$$C = k \begin{bmatrix} d & d & d & d & d & d & d & d \\ a & c & e & g & -g & -e & -c & -a \\ b & f & -f & -b & -b & -f & f & b \\ c & -g & -a & -e & e & a & g & -c \\ d & -d & -d & d & d & -d & -d & d \\ e & -a & g & c & -c & -g & a & -e \\ f & -b & b & -f & -f & b & -b & f \\ g & -e & c & -a & a & -c & e & -g \end{bmatrix} \quad (3)$$

where  $a = \cos(\pi/16)$ ,  $b = \cos(2\pi/16)$ ,  $c = \cos(3\pi/16)$ ,  $d = \cos(4\pi/16)$ ,  $e = \cos(5\pi/16)$ ,  $f = \cos(6\pi/16)$ ,  $g = \cos(7\pi/16)$ , and  $k$  is a scaling factor.

For inner product operation between the input vector and coefficient vectors, coefficients of  $\pm d$ ,  $(\pm b, \pm f)$ , and  $(\pm a, \pm c, \pm e, \pm g)$  are used for generating outputs,  $(F(0), F(4))$ ,  $(F(2), F(6))$ ,  $(F(1), F(3), F(5), F(7))$ , respectively. Therefore, some of coefficient capacitors can be common, and the number of coefficient capacitors are reduced to 32 from the straight-forward design using 64 coefficient capacitors. The 1-D DCT is computed by two steps using non-overlapping clock. For instance, at the  $F04$  output,  $F(0)$  and  $F(4)$  appear alternatively, synchronized to  $B0$  clock. Similarly,  $(F(2), F(6))$ ,  $(F(1), F(7))$  and  $(F(3), F(5))$  outputs appear alternatively at the outputs of  $F26$ ,  $F17$ , and  $F35$ , respectively. Fig. 4(b) shows an example of a basic SC cell. The coefficient is given by the ratio of a coefficient capacitor  $C_c$  to a feedback capacitor  $C_{fb}$ . Although the theoretical scaling factor  $k$  is  $1/2$ , a scaling factor of  $1/4$  is used to obtain sufficient signal dynamic range. Therefore,  $C_{fb}$  is chosen as  $C_{fb} = 4C_0$ , and the coefficient in 1-D DCT is given by  $C_c/C_0$ , where  $C_0$  is a unit capacitance. The  $C_0$  is chosen as  $0.5pF$ .

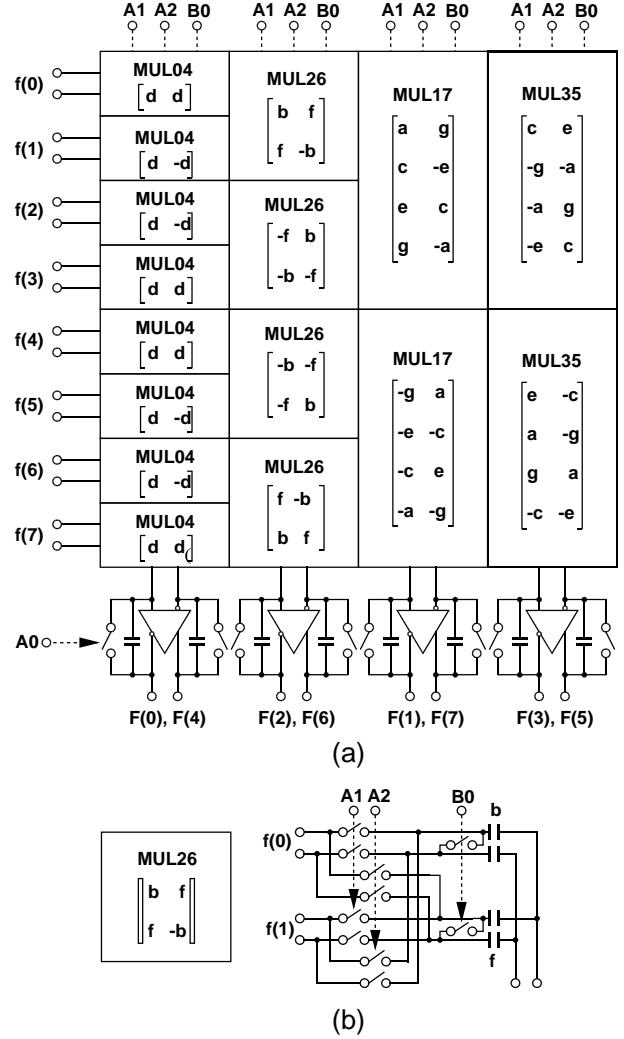


Fig. 4 1-D DCT circuits with a fully differential SC technique.

The 1-D DCT processor performs the 8-point 1-D DCT by two clock pulses. In the computation of Eqs.(2) and (3), 16 and 16 steps, respectively, are required. Therefore,  $8 \times 8$ -point 2D-DCT can be performed by 32 clock pulses using a 2-D DCT core.

### III. IMPLEMENTATION OF THE CMOS IMAGE SENSOR WITH 2-D DCT BASED COMPRESSION

Figure 5 shows the photomicrograph of the implemented image sensor prototype integrating the  $128 \times 128$ -pixel CMOS imager array, the analog 2-D DCT processor using two cores for complementary operation, and the ADC/Q. The chip is fabricated with triple-metal double-polysilicon n-well CMOS technology. The chip size is  $5.4 \times 4.3mm^2$ .

The performance of the CMOS imager array is summarized in Table I. Conversion gain is relatively small compared with the recently reported active-pixel sensors (APS). The conversion gain can be increased by choosing a smaller capacitor value in the front-end readout amplifier, or by amplifying at the second stage. Because

of the parallel readout scheme and the small number of pixels, the operation clock frequency is only 62kHz at 30 frame/s. Therefore, the power dissipation due to scanners of the imager array is negligible, and the total power is dominated by the readout amplifiers. The design of the readout amplifier is optimized at 2MHz operation. The power dissipation due to the readout amplifiers is almost unchanged up to  $128 \times 128$ -pixel at 1000 frame/s, or 0.5 M pixel at 30frame/s. The dark current is not critical, because the output-voltage-referred value at 30 frame/s is  $270\mu V$  (0.02% to saturation). The saturation voltage is relatively large despite of relatively low supply voltage of 3V. This is due to the use of voltage shifting technique. The fixed pattern noise is due to the deviation of the offset voltage in the 8 readout channels. This offset deviation of channels was measured directly under the dark condition. The fixed pattern noise can be suppressed by using offset canceling technique in the second stage of the readout circuits.

The random noise is relatively large because of the large  $kTC$  noise caused by the large common row capacitance. The active pixel technique [8] can be applied to our block access sensor by adding a transistor for block selection in order to improve the noise performance.

Table I Performance of the CMOS image sensor.

No. of pixels	$128 \times 128$
Pixel size	$16.1\mu m \times 16.1\mu m$
Fill factor	56.6%
Conversion gain	$1.6\mu V/e^-$
Power supply	3V
Power dissipation	$7.2mW@3V$ (Max.1000frame/s)
Dark current	$503pA/cm^2$
Saturation	1.6V
Fixed-pattern noise	$6mV_{p-p}$ (0.38% to sat. level)
Random noise	-53dB to sat.

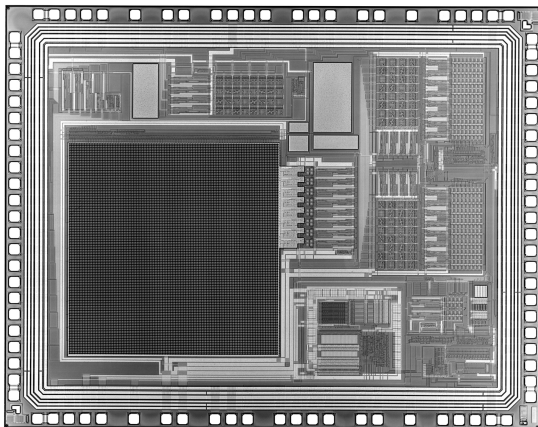
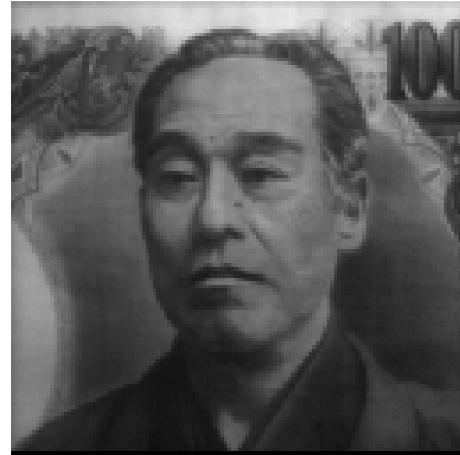
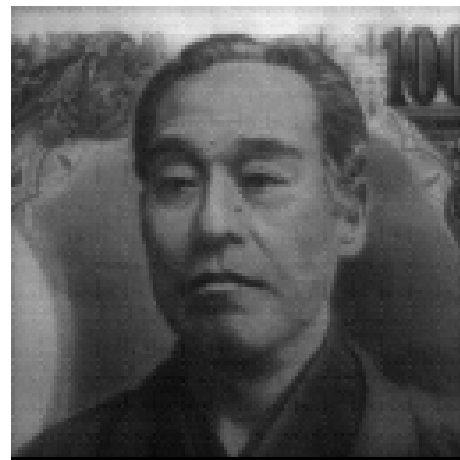


Fig. 5 Implemented CMOS image sensor with compression.



(a) Captured image



(b) Reconstructed image

Fig. 6 Images captured and reconstructed by the implemented chip.

The image captured by the implemented CMOS imager at 30 frame/s is shown in Fig. 6 (a). A testing board for the 8-channel parallel readout CMOS imager array is used. First, the 8-channel rail-to-rail outputs of the CMOS imager are converted to the single-ended signals and are converted to digital signals using 8-channel 10-bit A/D converters. To reduce the non-linearity error due to the A/D converters, upper 8 bits of 10-bit A/D converter is finally used. The image is once stored in a frame memory, and is loaded into a computer to use it as a reference image for the encoding.

Fig.6(b) is the reconstructed image using the integrated analog 2-D DCT processor on the CMOS sensor chip. At 30 frame/s, the image is acquired and is encoded by the analog 2-D DCT processor in real time. The 4-channel parallel outputs of the processor are digitized using a part of the same 10-bit A/D converter array used for the imager. The results stored in the output frame memory are loaded into the computer. The inverse 2-D DCT is performed by software in the computer, and finally a reconstructed image is obtained. In Fig. 6(b), the gain error, the offset, and the cross talk compensations are carried

out. The PSNR of 36.7dB is obtained. With careful comparison between Fig.6(a) and (b), we can find the degradation in the reconstructed image. This PSNR value is not sufficient for high quality image compression, and we need an effort to improve the precision.

An inevitable error in the analog 2-D DCT processor using the SC technique is due to the mismatch in capacitors used for the coefficient multiplication. The capacitor mismatch causes AC coefficient deviations when we apply a DC kernel. Using a stochastic simulation method, the degradation of the quality of image due to the capacitor mismatch and the offset voltage of the operational amplifier is examined. A standard image is coded by the 2-D DCT with random errors corresponding to the capacitor mismatch and the offset voltage deviations, and then decoded by the ideal inverse 2-D DCT. The degradation of image quality is evaluated by calculating PSNR (peak signal-to-noise ratio) of the reconstructed image. The careful design of the capacitor in the precision analog LSI allows us to obtain the mismatch error of less than 0.1%[9]. If the offsets of the operational amplifiers are successfully canceled, and  $\sigma_{cap} = 0.2\%$  is obtained, the PSNR will be more than 50dB from this simulation. In general, it is known that the degradation of image quality is hardly detectable by human eyes when the PSNR is over 40dB. Therefore, high quality compression is done with keeping the PSNR of around 40dB, and medium quality compression is in the range of 35 to 40 dB. As a result, the analog 2-D DCT processor using the SC circuits technique possibly achieves sufficient precision to obtain a good image quality.

The performance and the feature of the analog 2-D DCT processor are summarized in Table II. The effective area is  $1.1 \times 1.0mm^2$ , based on  $0.35\mu m$  double-metal double-polysilicon CMOS technology.

Because of the slow clock frequency due to the analog parallel processing, the analog 2-D DCT has low power. At  $760kHz$ , the power dissipation of the 2-D DCT processor is  $5.4mW$  at the supply voltage of 3V. The total power of two 2-D DCT cores is  $10.8mW$ . This is sufficiently low to attain  $50mW$  one-chip video camera. For high-speed design,  $12MHz$  operation is confirmed by simulation, which corresponds to the processing times of the 2-D DCT of  $2.7\mu s$  and  $1.3\mu s$ , for the single core and the complementary two core operations, respectively [11]. This is possible by increasing the power. The real-time encoding for larger format image is also possible using the analog 2-D DCT processor. For instance, the clock frequency required for the real-time encoding of  $1024 \times 1024$  pixel image at 30 frame/s is about 7.9MHz with two DCT cores.

The effective area of the analog 2-D DCT processor is about one fourth of the corresponding digital approach. For example,  $4mm^2$  2-D DCT processor based on  $0.3\mu m$  CMOS technology is reported based on the distributed arithmetic algorithm [10]. From the performance report-

ed in the same article, the equivalent performance at 3V and the scaling to the  $0.35\mu m$  technology, the power consumption and the 2-D DCT processing time are about  $151mW$  and  $0.43\mu s$ , or  $22mW$  and  $3\mu s$ , respectively. In the analog 2-D DCT processor, the power can be further reduced by half. The designed fully-differential opamp use a class-A folded cascode scheme and common-mode feedback circuits with MOS differential amplifier stages. The power of the opamp can be reduced by half without any performance degradation by the use of class-AB scheme and switched capacitor common-mode feedback circuits. As for the switching noise problem in the integrated CMOS image sensor with mixed-signal processing hardware, the analog 2-D DCT approach is better because the same processing speed can be obtained with the reduced clock frequency. The direct analog 2-D DCT encoding allows us to use efficient A/Dconversion techniques in order to reduce the total power. From the total comparison, the analog approach in the 2-D DCT processor has an advantage for the focal-plane compression.

Table II Feature of the analog 2-D DCT processor.

Core size	$1.1mm \times 1.0mm$
Unit capacitor	$0.5pF$
Power supply	3V
Power dissipation	$10.8mW$
PSNR (without compensation)	$31.4dB$
PSNR (with compensation)	$36.7dB$
Clock frequency	$760kHz @ 5.4mW$

#### IV. PROSPECTS OF CMOS IMAGE SENSORS WITH COMPRESSION

For more efficient moving picture coding, a hybrid (intra- and inter- frame) coding is essential.

The CMOS image sensor using conditional replenishment based image compression proposed by Aizawa et al [5] is the first trial to integrate a kind of inter-frame image coding on a CMOS image sensor. This approach is particularly useful for the high-speed camera. However, for usual video compression, inter-frame coding without motion compensation, despite of the relatively large hardware cost, is only effective when the global motion of the picture is little. We need motion-compensated interframe coding for focal-plane compression.

To meet the requirements for video compression on mobile terminals, we need much effort on the reduction of power dissipation of video encoder chip. The author believes that the solution to this problem lies in the integration of video compression hardware on a sensor chip. In the MPEG2 encoder chip, the power of more than 90% is consumed by that of motion vector estimation[12]. The focal-plane motion vector estimation is an interesting topic to reduce the total power required for the MPEG2 encoding. By integration the motion vector estimation hardware on an image sensor, the data access to the image becomes flexible. Especially, by increasing the frame

rate, the motion of objects can be restricted to the equivalent focal plane length of a few pixels. This greatly reduces the calculation complexity for block matching to obtain motion vectors. In our estimation, the calculation complexity is reduced to 1/30 using 16 times higher frame rates. The high speed imaging causes the degradation of the signal to noise ratio. However, the SNR can be recovered by using digital domain integration during the decimation process to the normal (30 frames/s) rate. The focal plane motion vector estimation will be one of most attractive topics in the field of CMOS image sensors and multimedia LSI chips for mobile products.

## V. CONCLUSIONS

In this paper, as an example of the integration of image sensors and image compression circuits on a chip, a CMOS image sensor with analog 2-D DCT based compression has been presented. One of important features of the integrated image sensor with compression is the low-power realization of the video compression hardware. The integration of motion vector detection on a sensor chip is promising approach for the thorough reduction of the power required for MPEG2 compatible compression, allowing us to treat digital moving pictures on mobile terminals.

## REFERENCES

- [1] E. R. Fossum, "CMOS image sensors :Electronic camera on a chip," Tech. Dig., IEEE Int. Electron Devices Meeting, pp. 17-25 (Dec. 1995).
- [2] E. R. Fossum, "Architectures for focal-plane image processing, Opt. Eng., vol. 28, 8, pp.865-871 (1989).
- [3] B. Ackland and A. Dickinson, " Camera on a chip," Dig. Tech. Papers, IEEE Int. Solid-State Circuits Conf., TA1.2, pp.22-25, 1996.
- [4] S. E. Kemeny, H. H. Torbey, H. E. Meadows, R. A. Bredthauer, M. A. Shell, and E. R. Fossum, "CCD Focal-Plane Image Reorganization Processors for Lossless Image Compression", IEEE J. Solid-State Circuits, vol. 27, no. 3, pp.398-405, 1992.
- [5] K. Aizawa, H. Ohno, T. Hamamoto, M. Hatori, and J. Yamazaki, " On sensor compression," IEEE Workshop on CCD and AIS, Dana Point, CA, April 1995.
- [6] S. Kawahito, M. Yoshida, M. Sasaki, K. Umehara, Y. Tadokoro, K. Murata, S. Doushou, A. Matsuzawa, " A compressed digital output CMOS image sensor with analog 2-D DCT processors and ADC/Quantizer," Dig. Tech. Papers, IEEE Int. Solid-State Circuits Conf., FA11.3, pp.184-185,1997.
- [7] T. Miyazawa, S. Nishizawa, M. Uehara, Y. Nakano, S. Nakamura, K. Tanaka, T. Horiuchi and K. Takemoto, " TSL solid-state imager," Proc. of TV Eng. Soc. Meeting Japan, pp.59-60, 1986.
- [8] S. Mendis, S. Kemeny, R. C. Gee, B. Pain, C. Staller, Q. Kim, E. Fossum, " CMOS active pixel image sensors for highly integrated imaging systems," IEEE J. Solid-State Circuits," vol. 32, no.2, 1997.
- [9] J. Shyu, G. Temes, and F. Krummenacher, "Random error effect in matched MOS capacitors and current sources" IEEE J. Solid-State Circuits, vol. SC-19, no. 6, pp.948-955,1984.
- [10] T. Kuroda, T. Fujita, S. Mita, T. Nagamatsu, S. Yoshioka, F. Sano, M. Norishima, M. Murota, M. Kato, M. Kinugawa, M. Kakumu and T. Sakurai, "A 0.9V 150MHz 10mW 4mm<sup>2</sup> 2-D discrete cosine transform core processor with variable threshold logic," IEEE Int. Solid-State Circuits Conf., FA10.3, pp. 166-167, 1996.
- [11] S. Kawahito, M. Yoshida, Y. Tadokoro, A. Matsuzawa, " An analog two-dimensional discrete cosine transform processor for focal plane image compression," IEICE Trans. Fundamentals, vol. J80-A, No.2, pp.283-291,1997.
- [12] M. Mizuno et al., " A 1.5W single-chip MPEG2 MPML encoder with low-power motion estimation and clocking", ISSCC97, pp.256-257, 1997.