A Bidirectional Multiple Charge Transfer Active Pixel Image Sensor for Low-Power Focal Plane Motion Vector Estimation

Non-member   Dwi Handoko   (Graduate School of Electronic Science and Technology, Shizuoka University)

Member   Shoji Kawahito   (Research Institute of Electronics, Shizuoka University)

Member   Yoshiaki Tadokoro   (Dept. of Information and Computer sciences, Toyohashi University of Technology)

Member   Akira Matsuzawa   (Advanced LSI Technology Development Center, Matsushita Electric Co. Ltd.)

This paper proposes a CMOS image sensor with high-speed non-destructive image sensing mode using bidirectional multiple charge transfer active pixels. The proposed image sensor is suitable for low-power focal plane motion vector estimation using iterative block matching while maintaining the image quality of video-rate (33ms/frame) pictures and the accuracy of the motion vector estimation. The performances of the bidirectional multiple charge transfer active pixel sensor are evaluated by circuit simulations and the proper operation of the sensor is confirmed by experiments.

Keywords: motion vector estimation, active pixel sensor, non-destructive image sensing, iterative block matching, bidirectional multiple charge transfer.

1. Introduction

Motion vector estimation (MVE) is the bottleneck of the video-encoding process since it has an enormous computational complexity. Therefore, in order to realize real-time video encoder, a specific LSI for MVE is required. However current hardware implementation of MVE requires a large-power large-size VLSI chip, which is not suitable for mobile consumer products.

On the other hand, recent advances of CMOS image sensors bring the ability to build an imager integrated with its signal processing and control functions. This benefit enables the development of a low-power camera system [1] [4].

The authors have proposed an on-sensor motion vector estimation technique based on iterative block matching and high-speed non-destructive intermediate image sensing in order to realize a low power video encoder [6]. This technique takes the benefit of small motion between successive intermediate high-speed pictures (high-speed pictures between 2 video-rate pictures) to obtain motion vectors of video-rate (33 ms/frame) pictures with a reduced computational complexity. However, in the conventional high-speed camera, the image quality of the pictures is degraded. The proposed non-destructive intermediate image sensing is important to obtain fully accumulated video-rate pictures while capturing intermediate pictures. However, in this technique, there still exist a problem on the image quality of intermediate high-speed pictures, which affects the accuracy of motion vector estimation.

This paper proposes a new CMOS image sensor with non-destructive intermediate image sensing mode using bidirectional multiple charge transfer active pixels. The proposed sensing device allows us to capture intermediate high-speed pictures with improved image quality as well as video-rate pictures.

2. Motion Vector Estimation

Block matching algorithm (BMA) is widely used for motion vector estimation. In a typical BMA, current picture is divided into macroblocks with size of $N \times N$ pixels. A macroblock of current picture is compared with the corresponding macroblock within the search area of $(2P + 1)^2$ in the previous picture, where P is the search range of block matching. The computational complexity of BMA is proportional to the search area [5]. On the other hand, a wide search range is required to cover large motion vectors.

The use of high-speed pictures reduces computational complexity of motion vector estimation using block matching algorithm, because a wide motion vector search range is not required due to small motion between high-speed pictures.

However, to estimate the motion vectors between video-rate pictures using high-speed pictures, a special consideration is necessary. The iterative BMA is suitable for this purpose [6]. The basic concept of the iterative BMA is shown in Fig. 1.

The conventional block matching estimates motion vector of neighboring pictures, while the iterative BMA, as shown in Fig. 1, iteratively estimates motion vector between a current intermediate high-speed picture and a previous video-rate picture as a fixed reference. The iterative BMA uses a previously obtained motion vector as a predictive vector to search the motion vector between intermediate pictures. This technique allows us to reduce the search range to a few pixels, while maintaining the estimation accuracy.

Let X, Y and $Y_k(k = 1, ...M - 1)$ be, respectively, a

Intermediate high-speed pictures

v_{k-1}

v_k

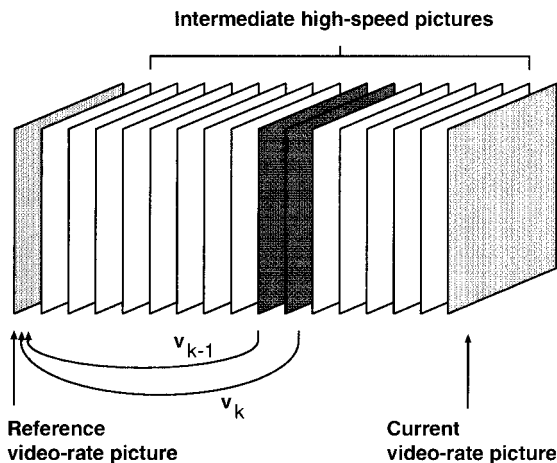**Reference video-rate picture**

Current video-rate picture

Fig. 1. Basic concept of the iterative BMA.

macroblock from reference picture, a macroblock from current video-rate picture and a macroblock from k-th intermediate high-speed picture. The iterative BMA is performed in the following steps:

(1) Estimate a motion vector $\mathbf{v}_1$ from Y_1 to X.

(2) For $k = 2$ to $M - 1$, estimate a motion vector $\mathbf{v}_k$ from Y_k to X, with $\mathbf{v}_{k-1}$ as a predictive vector so that $\mathbf{v}_k = \mathbf{v}_{k-1} + \mathbf{\Delta v}_k$, where $\mathbf{\Delta v}_k$ is the difference motion vector from Y_k to Y_{k-1}.

(3) For $k = M$, estimate a motion vector $\mathbf{v}_M$ from Y and X, with $\mathbf{v}_{M-1}$ as a predictive vector so that $\mathbf{v}_M = \mathbf{v}_{M-1} + \mathbf{\Delta v}_M$.

Vector $\mathbf{v}_M$ is the estimated motion vector between video-rate pictures. For N× N size of macroblock, $\mathbf{\Delta v}_k$ is obtained by block matching using cost function,

$$MAD_I(\Delta v_x^{(k)}, \Delta v_y^{(k)}) = \sum_{i=1}^{N} \sum_{j=1}^{N} |Y_k(i,j) -$$

$$X(i + v_x^{(k-1)} + \Delta v_x, j + v_y^{(k-1)} + \Delta v_y)|, \qquad (1)$$

where $Y(i,j)$ and $X(i,j)$ are the macroblock from the current picture and the macroblock from the reference picture at (i,j) coordinates, respectively. $(v_x^{(k)}, v_y^{(k)})$ and $(\Delta v_x^{(k)}, \Delta v_y^{(k)})$ are, respectively, component of vector $\mathbf{v}_k$ and $\mathbf{\Delta v}_k$.

Although the iterative BMA uses a fixed reference picture, the search range can remain small for all intermediate motion vector ($\mathbf{v}_k$) estimation between the k-th intermediate high-speed picture and the reference picture, because the iterative BMA only searches a difference vector ($\mathbf{\Delta v}_k$) by using a previously obtained intermediate motion vector ($\mathbf{v}_{k-1}$) as a predictive vector.

The motion vector estimation with a full search BMA has the most accurate motion vector but has the highest computational complexity among block matching algorithms.

As described above, the computational complexity of motion vector estimation is proportional to the number of search points given by $(2P + 1)^2$. For example, for P=32, the number of search points is 4225 in the case of full search BMA. In the case of iterative BMA, the

number of search points becomes $(2 \times \frac{32}{16} + 1)^2 = 25$ for the same search range of the video-rate picture, when 16 times higher frame rate is used. The total number of search points for 16 times iteration is 400. Therefore we can expect the reduction of the hardware complexity by a factor more than 1/10 using the iterative block matching.

According to the simulation results of moving picture encoding, the iterative BMA has almost the same PSNR (peak-signal-to-noise-ratio) and compression ratio of encoded image data, compared with the full search BMA [6].

3. Non-Destructive Image Sensing

The iterative BMA is suitable for single-chip integration of an image sensor and a video encoder. Though the iterative BMA requires high-speed image data for the motion estimation, the single chip solution relaxes the high-speed data transfer from the sensor to the motion estimation hardware by means of parallel multi-bit bus.

It is very important to obtain both the high image quality video-rate pictures and accurate motion vectors. The proposed non-destructive image sensing technique meets these requirements.

Fig.2(a) shows an image sensor with a typical 3-transistor type active pixel sensor (APS) array. Under a constant illumination, signal charge accumulation in the photodiode causes a gradual voltage decrease of the floating diffusion (FD) node of the sensor as shown in Fig.2 (b). The vertical scanner chooses one of horizontal line of pixels by activating the SELX signal and turning on the transistor M3. The signal voltage at the FD node is read out through a transistor M2, and then the FD node of the photodiode is reset by turning on the reset switch transistor M1 using the pulse signal R. Column correlated double sampling (CDS) circuits perform the subtraction of signal level from reset level to cancel fixed pattern noise of active pixel circuits. Thus the signal voltage V_S is given by

$$V_S = \frac{I_{photo} \times T}{C_{FD}} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (2)$$

where I_{photo}, T and C_{FD} are the photo current induced by the incident light input, the accumulation time and the stray capacitance at the FD node, respectively. The accumulation time is usually the frame period.

The change of floating node voltage (V_{FD}) of a pixel in high-speed image sensing is shown in Fig. 3. The signal voltage swing of the FD node decreases due to shorter accumulation time in high-speed imaging. The accumulated signal is destructed for every signal readout, and the small signal level degrades the SNR.

Fig. 4(a) shows the change of V_{FD} of a pixel in high-speed non-destructive image sensing [6]. In this case the signal charge is accumulated during video-rate period while capturing intermediate high-speed pictures. The accumulated signal is not destructed in high-speed intermediate pictures readout. The accumulated signal is destructed in video-rate pictures readout only. The
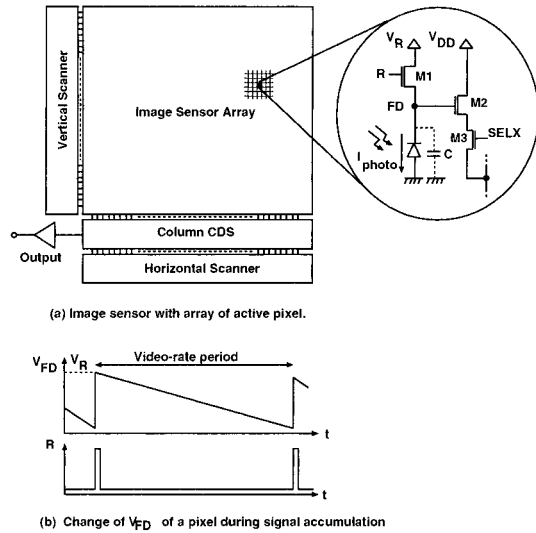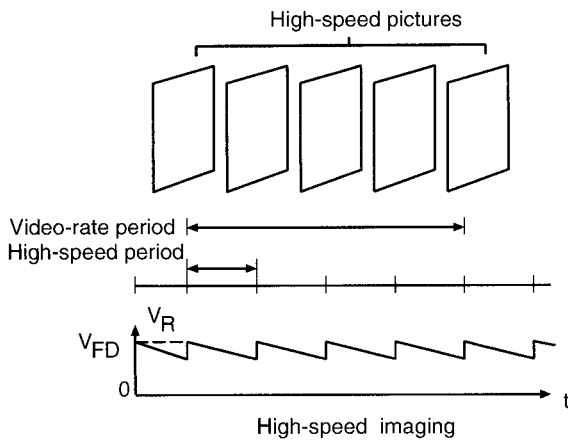
(a) Image sensor with array of active pixel.

(b) Change of V_{FD} of a pixel during signal accumulation

Fig. 2. Image sensor with 3-transistor active pixel sensor array.
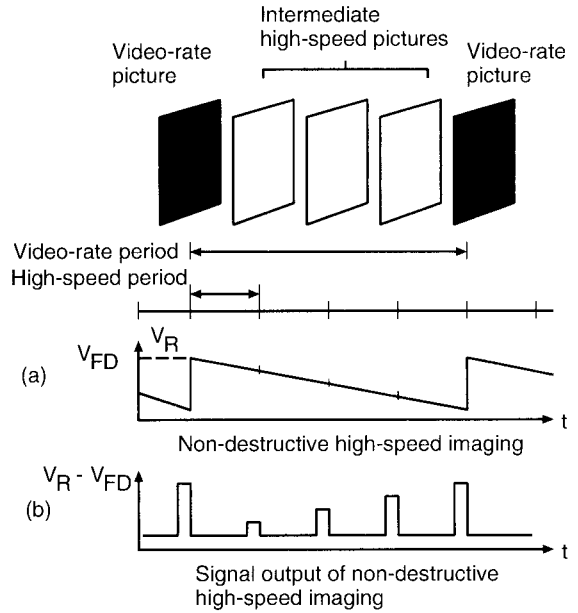

Fig. 3. High-speed image sensing.

Fig. 4. Principle of a high-speed non-destructive image sensing.
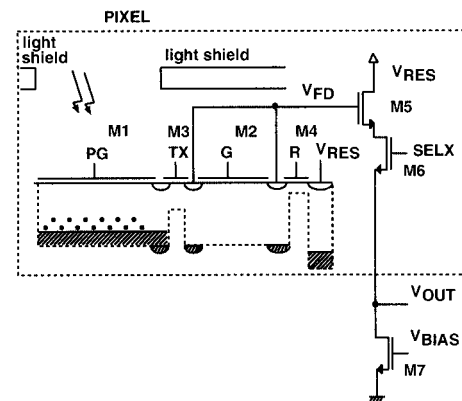
Fig. 5. Circuit configuration of a bidirectional multiple charge transfer active pixel image sensor.

high-speed image sensing with a non-destructive readout mode has a higher signal intensity of video-rate pictures compared to a usual high-speed image sensing with destructive readout mode. Fig. 4(b) shows the output signal level of a pixel in a high-speed imaging with a non-destructive readout mode.

However, in the case of a 3-transistor type APS shown in Fig. 2(a), the sensor does not have an ability to cancel the fixed pattern noise (FPN) for intermediate high-speed pictures, because reset voltages for intermediate high-speed pictures are not available [6]. This problem can not be solved also by using a conventional APS with a photodiode and a transfer gate transistor or a conventional APS with a photogate and a transfer gate transistor. The FPN in intermediate high-speed pictures degrades the accuracy of motion vector estimation.

The proposed bidirectional multiple charge transfer active pixel image sensor has the ability to cancel FPN for both intermediate high-speed pictures and video-rate pictures.

4. Bidirectional Multiple Charge Transfer Active Pixel Image Sensor

4.1 Circuit Configuration Fig. 5 shows the schematic diagram of the bidirectional multiple charge transfer active pixel image sensor. It consists of a photogate transistor (M1), a transfer gate transistor (M3) that separates photogate and the floating diffusion (FD) node, a reset transistor (M4), an input transistor of in-pixel source follower (M5), a row selection transistor (M6) and a MOS transistor (M2) for backward charge transfer. The configuration is similar to a photogate type active pixel [3], except for an additional transistor M2. This transistor is light shielded. A current source transistor (M7) is common for one column of pixel array.

The sensor has two signal readout modes of non-destructive and destructive for intermediate high-speed pictures and video-rate pictures, respectively.

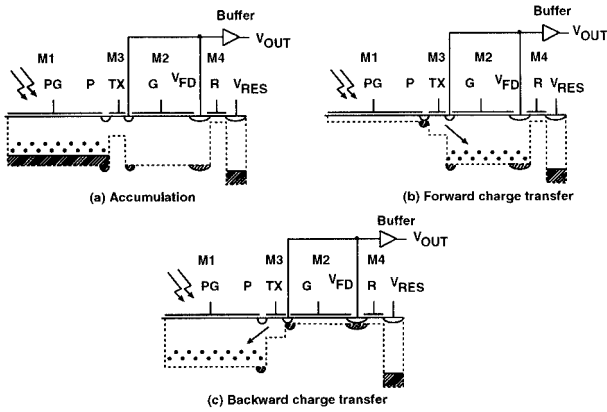**4.2 Non-destructive Readout Mode Operation** The operation of the sensor in a non-

(a) Accumulation   (b) Forward charge transfer

(c) Backward charge transfer

Fig. 6. Non-destructive readout operation.

(a) Reset to V_{RES} level   (b) Transfer offset charge of M1 to M2

(c) Transfer offset charge of M2 to V_{RES} node

Fig. 7. Reset operation.

destructive readout mode is illustrated in Fig.6. During signal accumulation period, both gates of photogate transistor (PG) and M2 (G) are set to V_{DD}. The gate of the transfer gate transistor (TX) is set to a middle voltage between V_{DD} and zero voltage during signal accumulation and readout operation. The gate of the reset transistor (R) is set to zero (Fig.6(a)).

To read out the signal in a non-destructive mode, the reset level of M2 is read out first when a pixel is selected. After that, the signal charge accumulated in M1 is forward transferred to M2 by setting the voltage of PG to zero (Fig.6(b)). Then the signal level of M2 is read out. After the signal is read out, the signal charge is backward transferred to M1 by setting the voltage of G to zero(Fig.6(c)). Thus the signal charge accumulation in M1 is continued to avoid SNR degradation. Furthermore, the CDS (correlated double sampling) operation allows us to cancel FPN for non-destructive mode signal.

4.3 Destructive Readout Mode Operation
After the signal of the video-rate picture is read out, the pixel is reset to initialize the FD node voltage for the signal accumulation of the next series of intermediate high-speed pictures. Fig. 7 shows pixel reset operation. First, the potential of photogate transistors M1 and M2 are set to V_{RES} by increasing gate voltages of the reset transistor (R) and the transfer gate transistor (TX) to V_{DD}. After that, TX is set to the medium voltage and R is set to the same level as TX. Next, the offset charge of M1 is transferred to M2 by setting PG to zero. Finally, TX is set to zero and the offset charge of M2 is transferred to V_{RES} node. In this way, the offset charges in M1 and M2 are transferred to V_{RES} node. This operation is not only effective to cancel the FPN, but also to reduce the kT/C noise which is a major component of the sensor random noise.

5. Analysis and Simulation Results

The APS operation is verified using a circuit simulator. The simulation is conducted using 0.35 μm CMOS technology parameters. The equivalent APS circuit used in the simulation is shown in Fig. 8. Power supply voltage V_{DD} is 3.3V. Fig. 9 shows timing diagram and voltage level of gate voltages of the pixel.
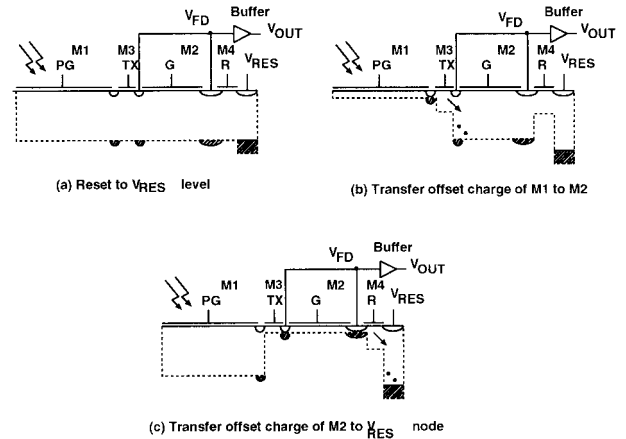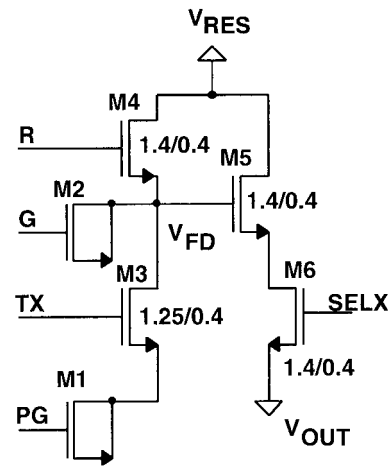

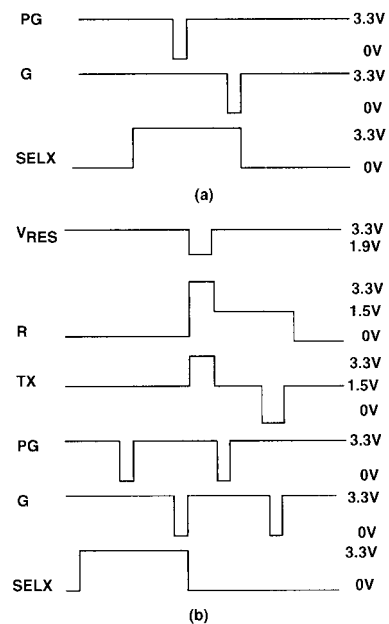Fig. 8. Equivalent pixel circuit used in the circuit simulation.

Fig. 9. Clock timing diagram of (a) non-destructive and (b) destructive readout operation.

5.1 Conversion Gain The sensor has a MOS capacitor in the FD node, to store signal charge tem-
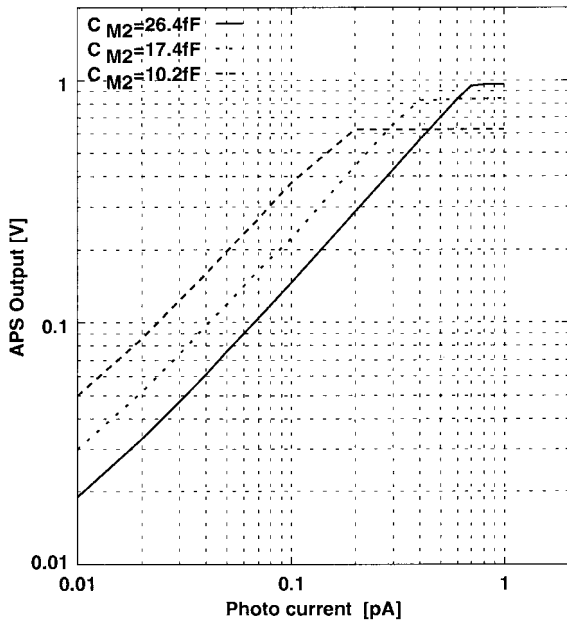
Fig. 10. Circuit simulation result of the photo conversion characteristics of the video-rate picture signal output to the signal photo current.
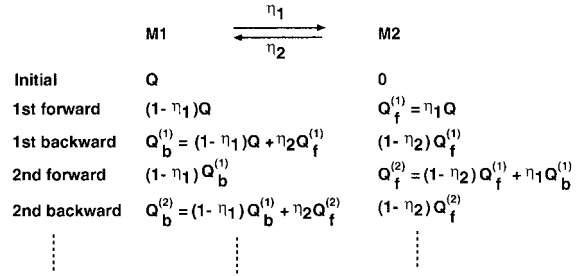

Fig. 11. Behavior of the transferred charge of the charge packet M1 and M2.

Fig. 12. Circuit simulation result of the APS output in relation to intermediate frame number.

porarily while signal charge is read out and to control backward signal charge transfer. Since voltage conversion gain is inversely proportional to the capacitance of the FD node, the presence of M2 in the proposed APS may decrease the conversion gain. Therefore the size of M2 should be small enough to obtain high conversion gain. On the other hand the choice of smaller capacitance results in smaller signal saturation voltage.

Fig. 10 shows the circuit simulation result of the photo conversion characteristics of the video-rate picture signal output to the signal photo current. The size of the gate area of M1 is fixed to $9.3\mu m^2$ which corresponds to the capacitance of 35.0fF. The photo conversion characteristics are simulated for three gate area of M2 of 7.0, 4.6 and $2.7\mu m^2$ whose equivalent capacitances are 26.4fF, 17.4fF and 10.2fF, respectively. For the capacitance of M2 is 10.2fF, the conversion gain of 15.7 $\mu V/e^-$ and a sufficient saturation level can be achieved. This value is comparable or a little smaller than the conventional photogate APS[7][8].

5.2 Charge Transfer Efficiency Signal charge is bidirectionally transferred between M1 and M2. Fig. 11 depicts the behavior of the transferred charge of the charge packet in M1 and M2. η_1 and η_2 are the charge transfer efficiency of forward transfer (M1 to M2) and backward transfer (M2 to M1), respectively. $Q_f^{(i)}$ and $Q_b^{(i)}$ denote amount of the charge in M2 packet after a forward transfer, and M1 packet after a backward transfer, respectively. In general $Q_f^{(i)}$ and $Q_b^{(i)}$ can be written as,

$$Q_f^{(i)} = (1 - \eta_2)Q_f^{(i-1)} + \eta_1 Q_b^{(i-1)} \cdots\cdots\cdots (3)$$

$$Q_b^{(i)} = (1 - \eta_1)Q_b^{(i-1)} + \eta_2 Q_f^{(i-1)}. \cdots\cdots\cdots (4)$$

If $\eta_1 = \eta_2 \equiv \eta$, $e = 1 - \eta$, and η is nearly equal to 1, then $|e| \ll 1$, and $Q_f^{(1)} = (1 - e)Q$, $Q_b^{(1)} = (e + (1 - e)^2)Q \cong (1 - e)Q$. We can easily prove

$$Q_f^{(i)} = eQ_f^{(i-1)} + (1 - e)Q_b^{(i-1)}$$
$$\cong (1 - e)Q = \eta Q \cdots\cdots\cdots\cdots\cdots (5)$$
$$Q_b^{(i)} = (1 - e)Q_b^{(i-1)} + eQ_f^{(i-1)}$$
$$\cong (1 - e)Q = \eta Q. \cdots\cdots\cdots\cdots\cdots (6)$$

Therefore it can be concluded that signal charge transfer efficiency of n times bidirectional charge transfer approximately equals to one time transfer, due to the effect of the untransferred charge.

Fig. 12 shows the plotted circuit simulation result of image sensor output in relation to intermediate frame number. The frame rate of intermediate high-speed pictures is assumed to be 16 times higher than that of video-rate pictures. The linear response to readout instances of intermediate pictures is confirmed for four photo current levels of 0.02pA, 0.1pA, 0.5pA and 2.5pA. For large photo current of 2.5pA, the output of the sensor decreases after reached the saturation level. This is due to overflow of signal charge from the photogate transistor to the transistor M2.

5.3 Correlated Double Sampling The correlated double sampling circuits are important to suppress the fixed pattern noise of the sensor. Fig.13 shows the designed CDS circuits. The CDS circuits are common for one column of the sensor array. The CDS circuits are composed of capacitors and switch transistors only.
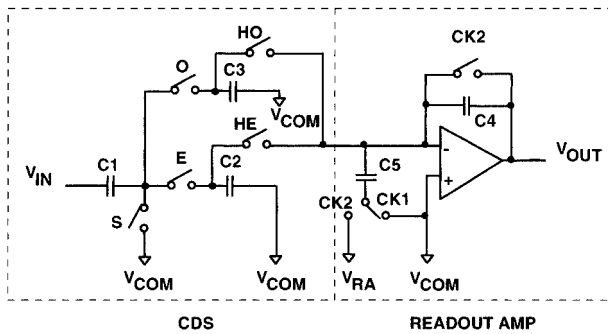
Fig. 13.   CDS circuit and readout amplifier circuit.

The image sensor output of the reset level of the even line is first sampled to the capacitor C_1 by turning on switches S and E. After that, by turning off switch S while switch E remains on, the signal level of APS of the even line is sampled and the difference of the signal level and the reset level of the APS is sampled to capacitor C_2. Charge stored in C_2 is then sampled and amplified by a readout amplifier. For image sensor output of the odd line, switch O and capacitor C_3 are used instead of switch E and capacitor C_2. The charge of capacitor C_2 is sampled while C_3 is used for the signal readout to the output, and vice versa. Hence the designed CDS circuits operate in interleave manner. This interleaving relaxes speed requirement of CDS circuits for high-speed imaging.

Fig. 14 shows a circuit simulation result of the signal voltage at FD node and the signal output voltage at CDS circuits output for readout instances of intermediate high-speed pictures. The photo current is 0.5pA. The operation frequency of the CDS circuit is 10MHz. From Fig.14, the signal voltage at the FD node is accurately preserved at the CDS circuits output.

6.   Experimental Results

The proposed image sensor chip with the size of 272 × 260 pixels is designed and fabricated using 0.35 μm CMOS technology. The pixel size is 9× 9 μm^2, and its fill factor is about 16%. The chip operates with a 3.3 V supply, and captures 480 frame/s intermediate high-speed pictures and 30 frame/s fully accumulated pictures.

In the measurement, the analog output is digitized by 12-b A/D converters. In Fig. 15, the measured digital output of a pixel is plotted in relation to intermediate frame number, under four different illumination levels. For moderate illumination, the output responses linearly to the intermediate frame number. While for strong illumination, the output decreases after reached the saturation level. The behavior agrees to simulation results described in section 5.

7.   Conclusions

A bidirectional multiple charge transfer active pixel image sensor for focal plane motion vector estimation has been described. The proposed image sensor captures non-destructive intermediate high-speed pictures
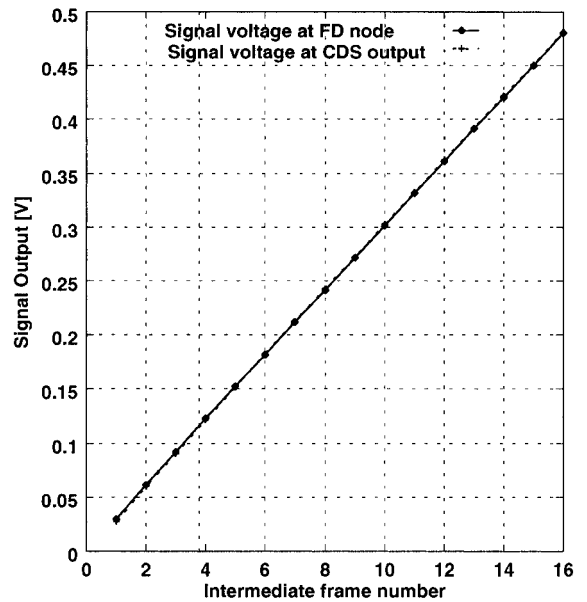

Fig. 14.   Circuit simulation result of the signal voltage at FD node and the signal output voltage at CDS circuits output as in relation to intermediate frame number.
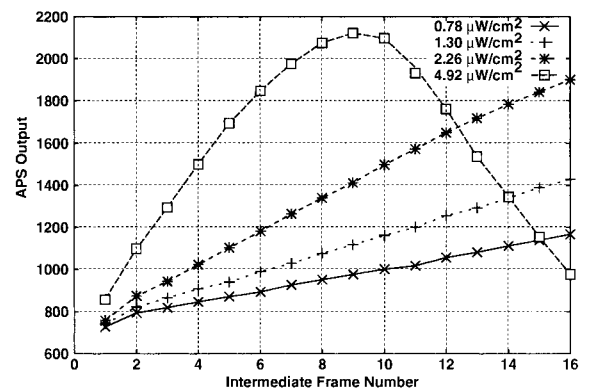

Fig. 15.   Measured digital output of intermediate pictures in relation to intermediate frame number under various illumination levels.

and destructive video-rate pictures with high image quality. The performances of the sensor were evaluated by circuit simulations. The experiment results confirm the proper operation of the sensor.

The practical implementation and the performance demonstration of motion vector estimation using the practical image sensor chip are left as a future subject.

(Manuscript received November 17, 1999, revised July 10, 2000)

References

(1) Bryan Ackland and Alex Dickinson: "Camera on a chip," Int. Solid State Circuit Conf., p.22, 1996.
(2) Shoji Kawahito, Makoto Yoshida, Masaaki Sasaki, Keijiro Umehara, Daisuke Miyazaki, Yoshiaki Tadokoro, Kenji Murata, Shirou Doushou and Akira Matsuzawa: "A CMOS image sensor with analog two dimensional DCT-based compres-

sion circuits for one-chip cameras," IEEE J. of Solid-State Circuits, Vol. 32, No. 12, pp.2030-2041, 1997.

(3) Sunetra K. Mendis, Sabrina E. Kemeny, Russell C. Gee, Bedabrata Pain, Craig O. Staller, Quiesup Kim and Eric R. Fossum: "CMOS active pixel image sensors for highly integrated imaging systems," IEEE J. of Solid-State Circuits, Vol. 32, No. 2, pp.187-197, 1997.

(4) Zeng Li, Kiyoharu Aizawa and Mitsutoshi Hatori: "Motion vector detection on image sensor focal plane by high speed block matching", ITE Technical Report Vol.22, No.13, pp.7-12,1998.

(5) Borko Fuhrt, Joshua Greenberg and Raymond Westwater: "Motion Estimation Algorithms for Video Compression," Kluwer Academic Publisher, Norwell, MA, 1997.

(6) Dwi Handoko, Shoji Kawahito, Yoshiaki Tadokoro and Akira Matsuzawa: "On Sensor Motion Vector Estimation with Iterative Block Matching and Non-Destructive Image Sensing," IEICE Trans. Electron, VOl. E82-C, No. 9, pp.1755-1763, Sept. 1999.

(7) Marc J. Loinaz, Kanwar Jit Singh, Andrew J. Blanksby, David A. Inglis, Kamran A. Azadet, and Bryan D. Ackland: "A 200-mW, 3.3-V, CMOS Color Camera IC Producing 352 × 288 24-b Video at 30 Frames/s," IEEE J. of Solid State Circuits, Vol. 33, No. 12, pp. 2092-2103, 1998.

(8) Hon-Sum Philip Wong, Richard T. Chang, Emmanuel Crabe and Paul D. Agnello: "CMOS Active Pixel Image Sensors Fabricated Using a 1.8-V, 0.25-μm CMOS Technology," IEEE Trans. On Electron Devices, Vol. 45, No. 4, pp. 889-893, 1998.

Dwi Handoko (Non-member) received the B.E. and M.E. degrees in electronics engineering from Miyazaki University, Miyazaki, Japan, in 1994 and 1996, respectively. He is currently pursuing the D.E. degree at Shizuoka University, Shizuoka, Japan. His research interest is in smart image sensor. He is a student member of Institute of Image Information and Television Engineers of Japan.
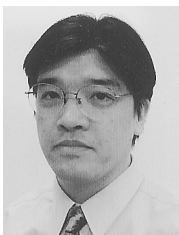
Shoji Kawahito (Member) received the B.E. and M.E. degrees in electrical and electronic engineering from Toyohashi University of Technology, Toyohashi, Japan, in 1983 and 1985,respectively, and the D.E. degree from Tohoku University, Sendai, Japan, in 1988. He is currently a Professor with the Research Institute of Electronics, Shizuoka University. From 1996 to 1997, he was a Visiting Professor at ETH Zurich. His research interests include integrated smart sensors and mixed analog/digital LSI circuits.
Dr. Kawahito received the Outstanding Paper Award at the 1987 IEEE International Symposium on Multiple-Valued Logic. His a member of the Institute of Electronics, Information and Communication Engineers of Japan and the Institute of Image Information and Television Engineers of Japan.

Yoshiaki Tadokoro (Member) received the B.E.,M.E.,and D.E. degrees in electronic engineering from Tohoku University, in 1967,1969, and 1976, respectively. From 1969 to 1978 he was an Instructor in the Department of Electronic Engineering, Tohoku University. From 1978 to 1986 he was an Assistant and Associate Professor, and he is currently a Professor in Toyohashi University of Technology. His recent interests have centered on digital signal processing and its applications to the communications and visual substitution system for the blind.
Dr. Tadokoro is a member of the Institute of Electrical Engineers of Japan, the Society of Instrument and Control Engineers of Japan, and the Information Processing Society of Japan.

Akira Matsuzawa (Member) received B.S., M.S., and Ph.D degrees in electronics engineering from Tohoku University, Sendai, Japan, in 1976, 1978 and 1997, respectively. In 1978, he joined Matsushita Electric Industrial Co., Ltd. Since then, he has been working on research and development of Mixed A/D LSIs and low power LSIs. He is currently a general manager in Corporate Semiconductor Development Division and his current interests are CMOS mixed signals LSIs, low-power DSP, power-management systems, low voltage SRAMs, CMOS RF circuits, high speed data converters, CMOS imagers, and analog boundary scan technology.
He is a member of the Institute of Electronics, Information and Communication Engineers of Japan.