Received 14 June 2020; revised 27 September 2020 and 30 November 2020; accepted 29 December 2020. Date of publication 1 January 2021; date of current version 28 January 2021. The review of this article was arranged by Editor C. Surya.

*Digital Object Identifier 10.1109/JEDS.2020.3048634*

# Correlation Image Sensor for Algebraic Solution of Optical Flow

**SHIGERU ANDO [1] (Member, IEEE), MASANORI NAGASE [2], TAKASHI WATANABE [2] (Member, IEEE), TOMOHIKO KOSUGI [2], AND TETSUYA IIDA [2]**

1 Department of Information Physics and Computing, University of Tokyo, Tokyo 113-8656, Japan
2 Brookman Technology Inc., Hamamatsu 432-8003, Japan

CORRESPONDING AUTHOR: S. ANDO (e-mail: shigeru_ando@ipc.i.u-tokyo.ac.jp)

This work was supported in part by Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant JP24300070; in part by Toyota Motor Corporation; and in part by Ricoh Elemex Corporation.

**ABSTRACT** In this article, we describe several enhancements of a three-phase correlation image sensor (3PCIS) toward its uses for a direct algebraic method of optical flow detection, i.e., determining velocity field of image patterns for detecting, tracking, and three-dimensional recovery of objects in a scene. A recent theoretical study has provided an exact closed-form solution of the velocity from temporal Fourier integral measurements of a time-varying intensity distribution. An accurate analog multiplication and integration of incident light and reference signals required for this type of measurements is the inherent advantage of a 3PCIS. We fabricate and compare several types of correlation detection block. As continuous-value inputs of the algebraic solution, compensation operations of systematic sources of error are so designed to realize a near shot-noise-limited performance. For random telegraph noise and interference reduction, we show that a p-epi channel, deep p-well structure is the best 3P multiplier. Several experimental results of optical flow detection including an accuracy evaluation and the dense motion capture of a traffic scene are given.

**INDEX TERMS** Image sensor, correlation, lock-in, CMOS, optical flow, weighted integral method, Fourier coefficient.

## I. INTRODUCTION

The three-phase correlation image sensor (3PCIS) [1] is a device that simultaneously produces the temporal integrals of incident light $f(\boldsymbol{r}, t)$ on pixel coordinate $\boldsymbol{r} = (x, y)$ with and without the orthogonal reference signal $w_n(t)$ as

$$g_n(\boldsymbol{r}) = \int_0^T f(\boldsymbol{r}, t) w_n(t) dt, \quad g_0(\boldsymbol{r}) = \int_0^T f(\boldsymbol{r}, t) dt \qquad (1)$$

where $T$ is the frame interval and $n$ is the index of a function set $\{w_n(t)\}$. It accepts complex reference signals by means of a three-phase (3P), two-quadrant analog multiplier in each pixel. Several devices [2], [3] have been fabricated and its improvements have been proposed [4], [5]. Potential applications including projection stereos [6], spectral matching imagers [7], magneto-optic imagers [8], and target tracking [9] have been studied. As a device technology, the features of this device are a nonlinear function embedded in the transduction mechanism of light, and the dense and massive integration of fragile elements dealing with small analog signals. Also, it is a challenge to realize such a structure with extremely low energy requirements and an extended sensory acquisition capability analogous to the visual system of living creatures.

An important finding regarding the application of 3PCISs is a novel mathematical technique, the weighted integral method (WIM). It provides a direct algebraic method to determine the coefficients of a differential equation from weighted integral measurements of its variables [10], [11], [12], [13], [14]. One of the targets of the WIM is optical flow detection [15], [16], [17], [18]. The optical flow is the two-dimensional velocity field that describes the translational motion of image intensity distribution. It has been widely studied, motivated by a broad range of applications such as the detection, tracking, and three-dimensional recovery of objects in a scene [19], [21], [22], [23]. The well-known differential method [19], however, has a fundamental limitation

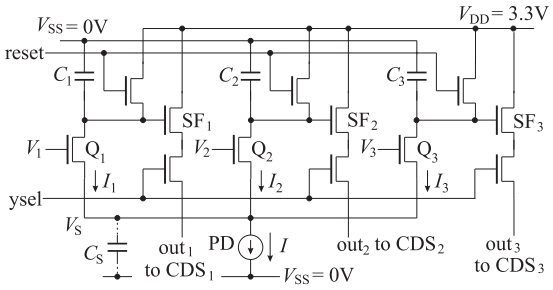
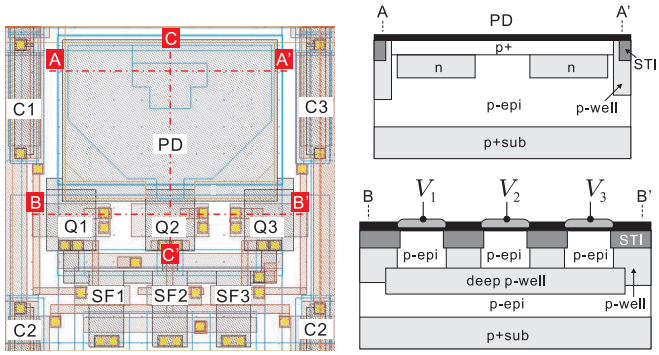



**FIGURE 1.** **Pixel circuit. $Q_1$, $Q_2$, $Q_3$: 3P multiplier, $C_1$, $C_2$, $C_3$: storage (integral of product) capacitors, and $SF_1$, $SF_2$, $SF_3$: source followers for readout.**




(a) pixel layout (b) pixel cross sections

**FIGURE 2.** **Pixel layout. PD is directly connected to the common source of the 3P multiplier $Q_1$, $Q_2$, $Q_3$ (p-epi channels shielded from the substrate by a deep p-well). MOS capacitors $C_1$, $C_2$, $C_3$ are additional storage capacitors to increase the saturation level of the integral and the SNR of correlation outputs. The fully depleted PD has a potential hill (no n-implant) inside and a monotonically slanted canal structure everywhere.**

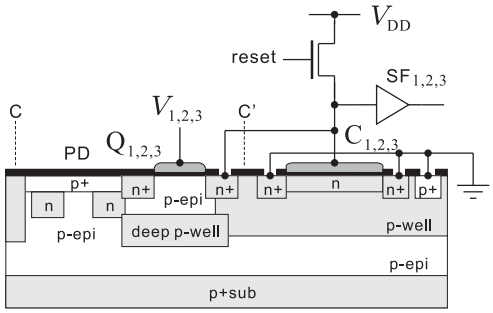



**FIGURE 3.** **Structure around the correlation detector block: PD, 3P multiplier with p-epi channels, and storage (integral of product) capacitors. The deep p-well shields the 3P multiplier from the via–substrate interference caused mainly by drifting photo-carriers and deep lobes of gate potential of surrounding pixels.**

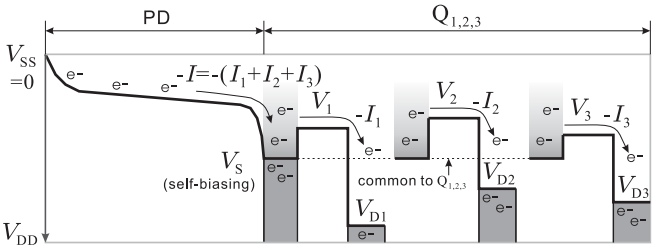



**FIGURE 4.** **Operation principle of correlation detector block of 3PCIS. Photoelectrons generated in PD are gathered and continuously flowing down to a common source of 3P multiplier, exceeding parallelly over the three gate potentials according to the thermal equilibrium conditions, and flowing down and accumulated in the corresponding drains. The common source voltage $V_S$ (floating) is automatically adjusted so that the outlet current $I_1+I_2+I_3$ is long-time balanced to the inlet PD current $I$.**

in the range of velocity and accuracy due to an approximation of temporal differential using two consecutive frames. In this respect, the combination of a 3PCIS and the WIM offers an effective and straightforward approach because the WIM provides an exact closed-form solution of the velocity from Fourier integral measurements (see Appendix A). The accurate analog multiplication of baseband signals and references required for this type of measurement is an inherent advantage of 3PCIS. Furthermore, the narrow frequency range of 3PCIS is no longer a shortcoming for it.

The purpose of this study is to obtain a sufficient accuracy and signal-to-noise ratio (SNR) of pixel outputs and subsequent compensation operations suitable for optical flow detection required in automotive and traffic scene analysis. Special emphasis is placed on the reduction of stray carriers in a photodiode (PD), random telegraph noise (RTN), and the interference of 3P multipliers. The subsequent removal of fixed pattern noise (FPN) is also extended through the compensation of imbalances, nonlinearity, and frequency characteristics while comparing with the shot-noise-limited level. Several experimental results of optical flow detection are given including the dense motion capture of a traffic scene.

## II. DEVICE STRUCTURE

The 3PCIS we describe here is composed of a $704\times512$ array of black-and-white, $12\times12$ $\mu m^2$ correlation detection pixels with Bayer-arrayed 3P reference signals, and 120 fps max, triple CMOS readout lines, which are built using 0.18 $\mu m$ CMOS image sensor process.

### A. PRINCIPLE OF OPERATION [1]

Figs. 1 and 2 show the pixel circuit and pixel layout, respectively. The correlation detection block consists of a PD, 3P multiplier [1] $Q_1$, $Q_2$, $Q_3$, storage (integral of product) capacitors $C_1$, $C_2$, $C_3$, reset transistors for rolling exposure and the integral, and source followers $SF_1$, $SF_2$, $SF_3$ connected to three correlated double sampling (CDS) circuits per column for the voltage readout of the integrals. Figs. 3 and 4 illustrate the vertical structure and operation principle of the correlation detection block, respectively. Photoelectrons flow down to the common source of $Q_1$, $Q_2$, $Q_3$, and the fraction excited thermally above the gate potentials fall down to the drains to obtain the products of the PD current and gate potential differences (see Appendix B for the mathematical formulation).

### B. PHOTODIODE

As shown in Figs. 2 and 3, the PD is fully depleted and has a potential hill (no n-implant) inside to realize a monotonically

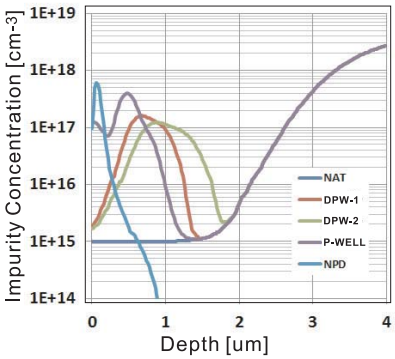



**FIGURE 5.** Impurity density profiles (TCAD simulation) near the PD and 3P multiplier. Channel structure of devices: proposed device: native (p-epi) + deep p-well (density profile is DPW-1 or DPW-2), compared type I device: p-well, and compared type II device: native (p-epi).

slanted canal structure everywhere [24] towards a single connecting port to the 3P multiplier. The structure eliminates the light-field dependence of detection, does not increase the common source capacitance $C_S$ of the 3P multiplier, and prevents carrier retention in the PD, which can generate poles of the frequency response of correlation outputs (see Section III-C) and flicker noise along with the drifting resistance of photoelectrons.

### C. 3P MULTIPLIER WITH RTN AND INTERFERENCE REDUCTION

A triplet of transistors $Q_1, Q_2, Q_3$ performs a 3P multiplication operation of the PD current connected to the common sources and 3P reference voltages (see Appendix B and [1] for detail). Each transistor has a low-impurity p-epi channel that is screened from the substrate by a weakly doped deep p-well. The p-epi channel with a wide width ($W = 1.8\mu$m) defined by shallow trench isolation (STI) is used to reduce both the barrier potential dispersion and RTN in it. The RTN is generated by temporally trapped and detrapped carriers at the interface between gate oxide and channel. The transition frequency increases according to the increase of electric field strength around the interface [25], [26], [27]. In CMOS image sensors, it is evident at SF transistors in low light levels [28]. In the 3P multiplier, they modulate the splitting ratio of the PD current, thus resulting in a severe source of correlation noise. Although the detailed mechanisms are unknown, constructing the channel with low-impurity p-epi reduces the field strength across the interface, hence, is expected to reduce the RTN.

On the other hand, the p-epi channel continuing to the substrate suffers from drifting photo-carriers and proximate gate potentials. The deep p-well formed by high–energy ion implantation [29] reduces this via–substrate interference. Fig. 5 shows impurity density profiles (TCAD simulation) near the PD and 3P multiplier. We have examined three structure types: 1) p-well, 2) native (p-epi) + deep p-well, and 3) native (p-epi) structures. Significant reductions in RTN and via–substrate interference in the native (p-epi) + deep p-well structure will be verified in Section IV.

**TABLE 1.** Compensation parameters for each pixel.

| item | parameters | bytes | measurement conditions |
|---|---|---|---|
| DC offsets | $b_1, b_2, b_3$ | $2\times3$ | outputs at dark |
| SF curvatures | $c_1, c_2, c_3$ | $2\times3$ | outputs at dark, mid., and bright |
| DC gain and 3P imbalance | $a_1.a_2.a_3$ ($\varepsilon_1.\varepsilon_2.\varepsilon_3$) | $2\times3$ | outputs at mid. and bright |
| AC gain and phase delay | $a_r, a_i$ ($\eta, a_{\rm AC}$) | $2\times2$ | outputs to modulated light |
| AC offsets | $b_r, b_i$ | $2\times2$ | for on-site FPN cancelation |
| total | | 26 | bytes/pixel |

### D. MOS CAPACITORS

Capacitors $C_1, C_2, C_3$ provide the storage for integrating the current outputs of the 3P multiplier. The use of MOS capacitors in addition to conventional floating diffusion increases the saturation level of the integral and reduces the ratio of shot noise in the correlation outputs so that sufficient accuracy can be obtained in optical flow detection.

### E. READOUT AND RESET

Three voltages on the storage capacitors $C_1, C_2, C_3$ are read out through triple CDS circuits for each column and output as 2 Vp-p differential signals [2]. Immediately after the readout, the charges are reset and the next integral of product begins. No electronic shutter is available.

### F. BAYER–ARRAYED CORRELATIONS

Three sets of 3P reference signals, Bayer-A, -B, and -C, are supplied to pixels in every $2\times2$ block in the form of $\begin{matrix}\mathrm{A}&\mathrm{B}\\ \mathrm{C}&\mathrm{A}\end{matrix}$. To operate in the single or dual correlation mode, all sets of reference signals or the sets of Bayer-B and -C reference signals are connected to each other (the optical flow uses the single correlation mode).

## III. COMPENSATION OPERATIONS

Compensation is unavoidable to achieve shot-noise-limited performance while accepting fluctuations of device parameters. The rather primitive structure of the proposed device enables us to obtain mathematical models that relate the fluctuations to the correlation outputs. Table 1 shows the list of 26-bytes-per-pixel compensation data.

### A. DC OFFSET NULLING AND SF LINEARIZATION

The overall offsets of source follower (SF), CDS, and readout circuits are compensated. The nonlinearity and its difference among the triple SFs are very harmful for balancing 3P outputs in a wide output range. As second-order compensation, the voltage gain characteristics from the gate to the source of SF are linearized.

### B. 3P MULTIPLIER BALANCING AND DC GAIN UNIFORMIZATION

The MOS transistor characteristics in the subthreshold region are expressed as

$$I_{\rm DS} = I_{\rm s} e^{\frac{q\kappa(V_{\rm GS}-V_{\rm T})}{kT}} = \left(I_{\rm s} e^{-\frac{q\kappa V_{\rm T}}{kT}}\right) e^{\frac{q\kappa V_{\rm GS}}{kT}}$$

where $I_{\rm DS}, I_{\rm s}, V_{\rm GS}, V_{\rm T}, q, \kappa, k,$ and $T$ are the drain current, saturation current, gate-source voltage, threshold voltage, elementary charge, gate sensitivity coefficient, Boltzmann constant, and absolute temperature, respectively. The compensation of this stage removes the imbalances of the parenthesized parts from their average. The imbalance of $\kappa$, caused mainly by the uneven gate-oxide thickness and channel dopant fluctuation, cannot be expressed by a simple compensation equation. Therefore, our approach is to reduce it using native MOS transistors with a low-impurity p-epi channel. This is also beneficial for reducing the RTN generated in them.

Let the imbalance ratios of $Q_1$, $Q_2$, and $Q_3$ be $1+\varepsilon_1$, $1+\varepsilon_2$, and $1+\varepsilon_3$, respectively. Let the photodiode current be $I$ and the split currents through $Q_1$, $Q_2$, and $Q_3$ be $I_1, I_2$, and $I_3$, respectively. Let the 3P reference voltages with respect to their average be $v_1, v_2$, and $v_3$. Then, the 3P correlation outputs after compensation are expressed as

$$\begin{bmatrix} <Iv_1> \\ <Iv_2> \\ <Iv_3> \end{bmatrix} = \frac{kT}{q\kappa} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} \begin{bmatrix} <I_1>/(1+\varepsilon_1) \\ <I_2>/(1+\varepsilon_2) \\ <I_3>/(1+\varepsilon_3) \end{bmatrix} \tag{2}$$

where $<\cdot>$ indicates the integral over a frame interval, and $<I_1>, <I_2>, <I_3>$ are the values after DC offset nulling and SF linearization (see Appendix B for the derivation).

The Nonuniformity of DC sensitivity, $a_{\rm DC}$, caused by the pixel lens and PD is also compensated at this stage using

$$a_1 = \frac{a_{\rm DC}}{1+\varepsilon_1}, \quad a_2 = \frac{a_{\rm DC}}{1+\varepsilon_2}, \quad a_3 = \frac{a_{\rm DC}}{1+\varepsilon_3},$$

which are multiplied by $<I_1>, <I_2>, <I_3>$, respectively.

### C. FREQUENCY COMPENSATION

The frequency characteristics from the modulated component of light to the correlation output are expressed as a first-order LPF with cutoff frequency

$$\omega_{\rm co} = \frac{1}{C_{\rm S} r_{\rm S}} = \frac{q\kappa \bar{I}}{kTC_{\rm S}},$$

where $\bar{I}$ is the DC component of $I$, $C_{\rm S}$ is the stray capacitance at the source, and $r_{\rm S} = -\partial V_{\rm GS}/\partial I$ is the equivalent source resistance of the 3P multiplier [1]. Therefore, under the standard illumination condition (half the saturation level), the cutoff frequency has the following simple relation with the scan frequency:

$$\eta \equiv \frac{\omega_{\rm co}}{\Delta_\omega} = \frac{3\kappa}{2\pi} \cdot \frac{V_{\rm max}}{kT/q} \cdot \frac{C}{C_{\rm S}}, \tag{3}$$

where $C = (C_1 + C_2 + C_3)/3$ and $V_{\rm max}$ is the saturation voltage across $C$. In our device, $\eta \sim 20$. Therefore,

$$A(j\omega) = \frac{a_{\rm AC}}{1 + j\omega/(\eta\Delta_\omega)} = \frac{1}{a_r + ja_i}, \tag{4}$$

where $\omega = n\Delta_\omega$ is the frequency of the modulated light and reference signals, $a_{\rm AC}$ is the gain coefficient of the pixel with respect to the average over the image, and $a_r + ja_i$ are the AC gain/phase compensation parameters. The finite Fourier integrals are obtained from the above as

$$g_0 = \int_0^T I(t)dt = \frac{<I_1>}{1+\varepsilon_1} + \frac{<I_2>}{1+\varepsilon_2} + \frac{<I_3>}{1+\varepsilon_3} \tag{5}$$

$$\begin{aligned} g_n &= (a_r + ja_i) \int_0^T I(t) e^{-jn\Delta_\omega t} dt \\ &= (a_r + ja_i) \int_0^T I(t) \Bigg( \cos(n\Delta_\omega t). \\ &\quad + \frac{j}{\sqrt{3}} \left( \cos\left(n\Delta_\omega t + \frac{2\pi}{3}\right) - \cos\left(n\Delta_\omega t + \frac{4\pi}{3}\right) \right) \Bigg) dt \\ &= \frac{a_r + ja_i}{v_{\rm ref}} \left( <Iv_1> + \frac{j}{\sqrt{3}} (<Iv_2> - <Iv_3>) \right), \end{aligned} \tag{6}$$

where

$$v_1 = v_{\rm ref} \cos(n\Delta_\omega t), \quad v_2 = v_{\rm ref} \cos\left(n\Delta_\omega t + \frac{2\pi}{3}\right),$$
$$v_3 = v_{\rm ref} \cos\left(n\Delta_\omega t + \frac{4\pi}{3}\right).$$

The optimum value of $v_{\rm ref} \sim q\kappa/kT$ is determined by comparing the nonlinearity and SNR of the correlation outputs, and $\kappa \sim 0.8$ is determined using a known modulated illumination.

The compensation by $a_r + ja_i$ assumes the standard illumination condition over the entire image. Frequency compensation depending on the intensity of each pixel is performed, if necessary, after readout by a PC.

## IV. EXPERIMENTAL EVALUATION

In this section, we examine the performance of the fabricated devices with relating to the operating conditions and structures. With the primary evaluations in [1], experiments are also for the validation of mathematical models for the correlation and compensation operations. As the SNR, shot noise limitedness is the final goal for the proposed device. According to the addition formula of the Poisson process, the sum and difference of the numbers of carriers split by the 3P multiplier, respectively, follow the Poisson and Skellam distributions with means $\mu_1 \pm \mu_2$ and variances $\mu_1 + \mu_2$, where $\mu_1$ and $\mu_2$ are means (= variances) of two Poisson processes. This means that the noise variances of the intensity and correlation in electron counts are equal to the total electron count (illumination/intensity level).

### A. SUITABLE REFERENCE SIGNAL VOLTAGES

As shown by Eq. (6), $g_n$ is inversely proportional to the reference signal voltage $v_{\rm ref}$. Thus, the SNR of $g_n$ increases

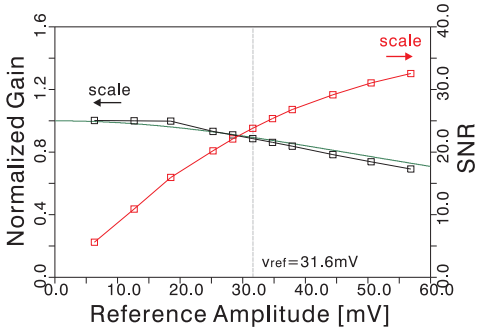



**FIGURE 6.** Graphs of normalized gain (left scale, green curve shows the theoretical values) and SNR at low modulation (right scale) of correlation output vs reference signal voltage $v_{ref}$. When $v_{ref} > kT/q\kappa$, the first-order approximation of $e^{q\kappa v_{ref}/kT}$ deteriorates; hence, the correlation gain and accuracy decrease. However, the 3P multiplier output and its SNR still increase. The thin dotted line shows the current setting of $v_{ref} = 31.6$ mV.

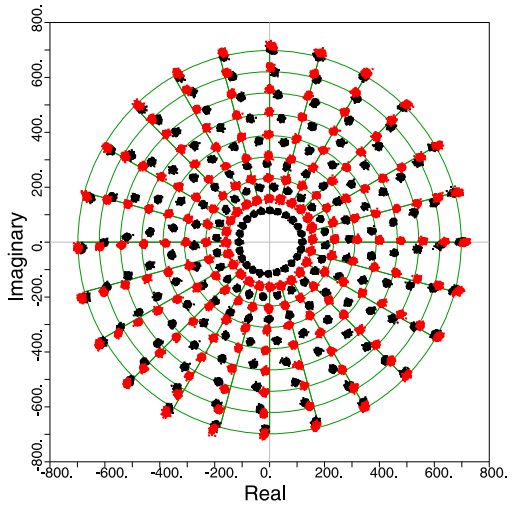



**FIGURE 7.** Correlation outputs (superposition of central 256 pixels) under fixed modulation ratio and varying illumination level (frame interval $T = 0.066$ s, modulation frequency $4/T = 60.58$ Hz). Raw outputs (black) show significant dependences of the attenuation and phase delay on the intensity level, but they are mostly removed in the intensity-dependent frequency-compensated results (red). Nodes of green mesh indicate the desired outputs.

with $v_{\rm ref}$. When $v_{\rm ref}$ becomes comparable to $kT/q\kappa$, however, the first order approximation of $e^{q\kappa v_{\rm ref}/kT}$ becomes inappropriate; hence, the correlation gain and accuracy decrease. Fig. 6 shows graphs of normalized gain (1 if $v_{\rm ref} \ll kT/q\kappa$) and SNR measured under a low modulation condition. The current setting of $v_{\rm ref} = 31.6$ mV was chosen as a trade-off between them. Experimental normalized gains (black) are close to the theoretical curve (green) from Appendix B. This shows an appropriateness of the exponential model of 3P multiplier.

### B. FREQUENCY CHARACTERISTICS AND NOISE LEVEL

Fig. 7 shows correlation outputs (black dots) under a fixed modulation ratio of 2/3 (sinusoidal amplitude/intensity) and varying illumination level and modulation phase (both in uniform steps). To describe the intensity dependence of

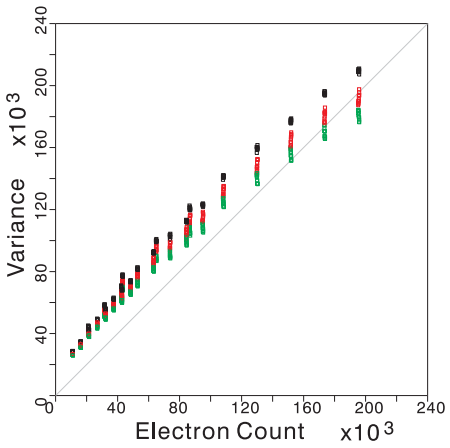
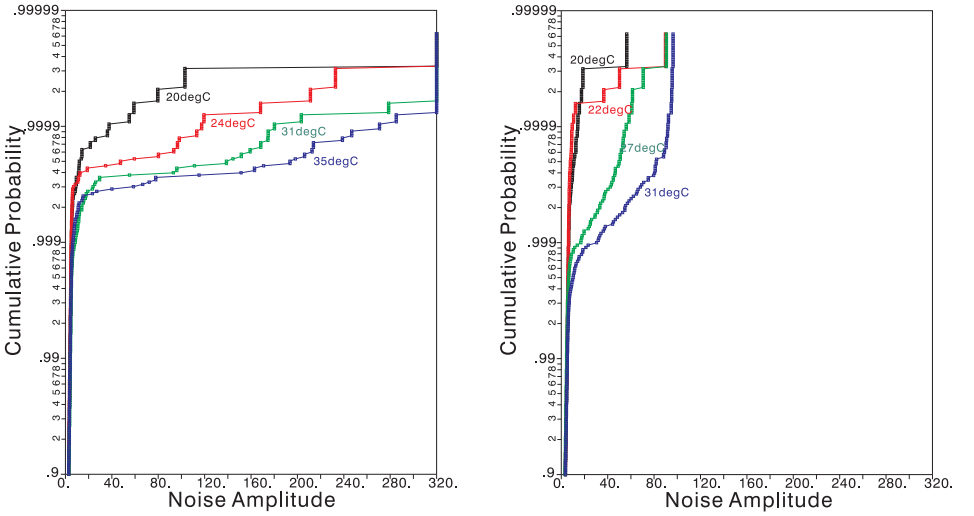



**FIGURE 8.** Graph of noise variances of intensity and correlation components vs intensity level (both are scaled as electron counts). Black, red, and green dots, respectively, indicate the intensity, the real part of correlation, and the imaginary part of correlation. The electron charge to output voltage coefficient is 6.36 $\mu$V/e$^-$. The 45 deg line shows the pure shot noise level.

amplitude attenuation and phase delay more clearly, the modulation frequency is increased to $4\Delta_\omega/2\pi = 60.58$ Hz from $\Delta_\omega/2\pi = 15.145$ Hz for optical flow detection. A significant decrease in amplitude and an increase in phase delay are evident toward the inner circle with decreasing intensity level. In this respect, Red dots of Fig. 7 shows the results of compensation based on the intensity-dependent LPF model given by Eqs.(3) and (4). Since the dependences are mostly removed towards the desired outputs (green), the validity of the LPF model of 3P multiplier is confirmed. Fig. 8 shows the dependences of the noise variances of the intensity and correlation components on the illumination level. According to the nature of the Poisson process and the splitting of events, the noise variances of these quantities should coincide with the total electron count. Assuming the conversion coefficient 6.36$\mu$ V/e$^-$, and an overall gain of the readout circuits of 0.8, the graphs are distributed parallel to the 45 deg line showing the pure shot noise level. The slight offset is caused by reset noise and/or interference from digital circuits.

### C. RANDOM TELEGRAPH NOISE (RTN)

The RTN we refer to is the blinking of a small number of pixels at low rate (e.g., on/off per several frames) in a correlation image. The RTN is large and unpredictable, and thus can severely damage the image analysis. Its transition frequency increases significantly according to the increase of temperature [26] and the decrease of channel current [27]. Fig. 9 shows Gumbel plots of noise at several temperatures. The graphs show high-frequency (rate of occurrence), small-amplitude noise behaviors on the lower left, and low-frequency, large-amplitude noise behaviors on the upper right. The former are mostly shot noise and reset noise (see Fig. 8), and the latter are caused mainly by RTN. The RTN of the proposed device with a p-epi channel is much lower




(a) compared type I (p-well channel) (b) proposed type (p-epi+deep p-well channel)

**FIGURE 9.** **Gumbel plots of large-noise behaviors at several temperatures. The graphs show the high-frequency (occurrence rate), small-amplitude noise behavior caused mainly by shot noise and the low-frequency, large-amplitude one caused mainly by RTN. Lower RTN of the proposed type than the compared type I is clear. The results (b) also show the possibility of another source of low-frequency, medium-amplitude noise that increases proportionally to the temperature.**

than that of the compared device with a p-well channel (they are the same device except for the channel structure). The results in (b) show the possibility of another source of low-frequency, medium amplitude noise that increases proportionally to the temperature. This may be the RTN of SFs [27], [30]. If this is the case, the RTN of the compared device with p-well channel is considered to be the sum of the RTN of the 3P multiplier and that of the SFs. In contrast, the RTN of the proposed device with the p-epi channel will be caused only by the SFs. The RTN of 3P multiplier is much larger than those of SF because it affects the splitting ratio of current in 3P multiplier ($\sim 0.1$ pA) which is much less than that of SF (19.2 $\mu$A).

### *D. VIA–SUBSTRATE INTERFERENCES*

The via–substrate interferences referred to here are various false responses generated in the correlation image by some beyond-pixel interactions of the 3P multipliers. Although details of the phenomena have not been clarified, they can be significantly decreased by the appropriate shielding of the 3P multiplier from the substrate. Fig. 10 shows a comparison of three shielding structures: p-well (black), native (p-epi) + deep p-well (red), and native (p-epi) only (green). The leakage types are given in the caption. A significant reduction in interference upon adding the deep p-well structure is evident. The *x*-differential-like leakage will be relevant to a moving potential distribution such as a traveling wave induced by a 3P reference voltage array and drifting photocarriers. For this type of leakage, however, effectiveness of all structures is not definite.

## V. APPLICATION TO OPTICAL FLOW DETECTION

This section reports some experiments to confirm the optical flow detection capability of the 3PCIS. All are single-frame

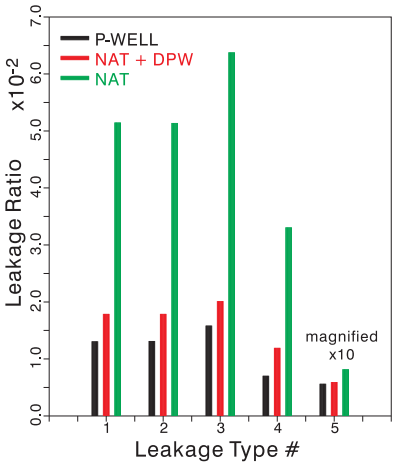



**FIGURE 10.** **Comparison of the via–substrate interferences of correlation outputs of the compared type I, the proposed type, and the compared type II devices with p-well, p-epi+deep p-well, and native (p-epi) channel multiplier structure, respectively. Leakage type 1: Bayer-BC outputs illuminated by Bayer-A frequency, leakage type 2: Bayer-A outputs illuminated by Bayer-BC frequency, leakage type 3: vertically rotating phase component of Leakage type 1, leakage type 4: 3P multiplier imbalance depending on the reference signal, and leakage type 5: *x*-differential-like response. All leakage ratios can be biased by inseparable random pixel noises.**

results without an intensity-dependent frequency compensation.

### *A. EVALUATION OF DETECTION ACCURACY*

Fig. 11 shows optical flow results of a rotating disk with a constant speed. Velocities with a wide range of magnitudes and directions are involved. The images (a), (b), and (c) are for the WIM, and (c), (d), and (e) are for the conventional differential method (DM) [20]. The intensity image

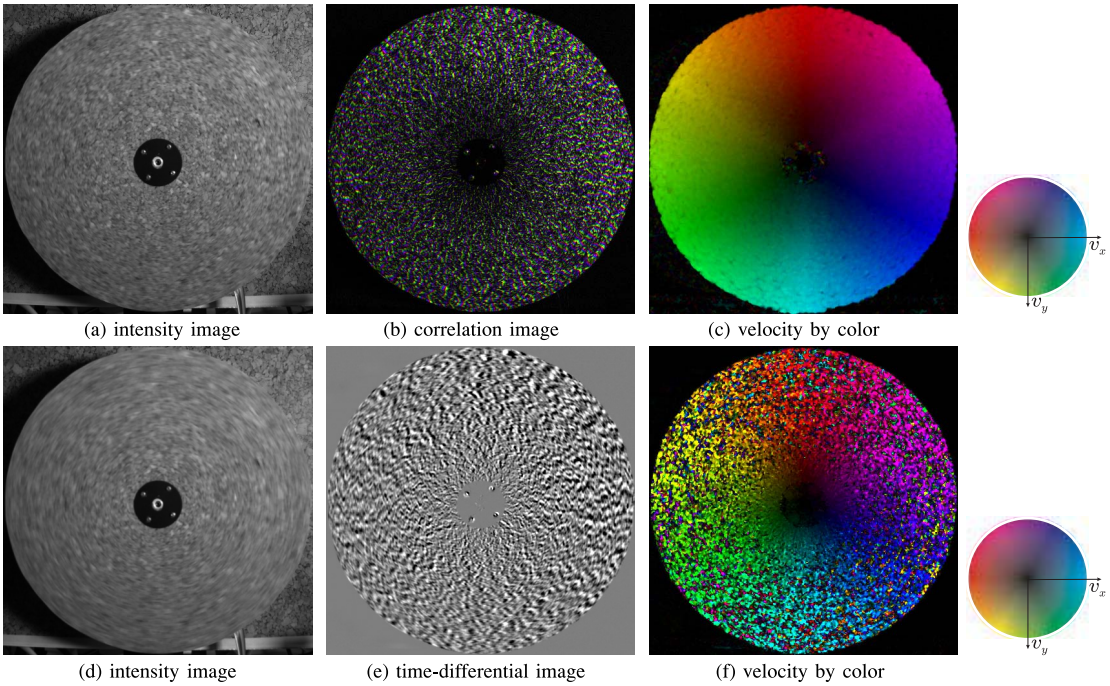



**FIGURE 11.** **Optical flow detection of a rotating disk. (a), (b), and (c) show source images and result of the weighted integral method, and (c), (d), and (e) show those of the differential method ((d) and (e) are the sum and difference images of two consecutive frames including (a)). Both are using 3 × 3 local least-squares algorithm. Velocity vs color correspondences are indicated by the diagrams near (c) and (f).**

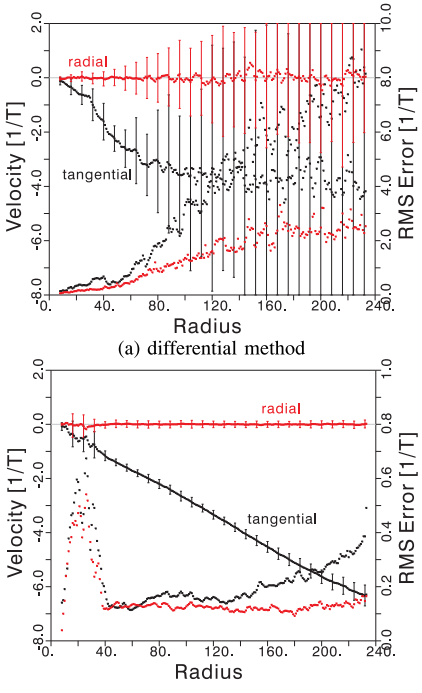



**FIGURE 12.** **Tangential (black) and radial (red) velocity components (left axis scale), and rms estimation errors (right axis scale) of the optical flow results in Fig. 11 (c) and (f). The horizontal axis is the distance from the rotation center. Vertical bars indicate the rms error. Note that the right axis scale (rms error) of (b) is 10-times magnified than (b).**

(a) shows significant motion blur near the margin, where, in contrast, the correlation image (b) captures strongly the moving textures as a complex-valued distribution. In the DM, the intensity image (d) and time-differential image (e) are obtained by the sum and difference of two consecutive frames. In both methods, spatial gradients were computed in the frequency domain, and $3 \times 3$ local least-squares algorithms were applied. In the DM result (f), most areas are filled with large clutters, and the accuracy is severely lost except for the central region with small velocities. In contrast, it is maintained everywhere in the WIM result (c). Dimmer blobs near the fringe show areas where the velocity estimate was ill-conditioned. They are mostly caused by the aperture problem or an excess velocity beyond the frequency band of the correlation image [17].

Fig. 12 (a) and (b) show the polar velocity components of the optical flow in Fig. 11 (c) and (f), respectively. In both results, the mean and variance are taken on each locus of circular motion, and plotted relating to the radius as the lateral axis. Although the DM shows a velocity limit at a very low level, $\sim 1.5$ pixel/frame, the accuracy of the WIM is maintained in wide velocity range. Except for the central area without texture, the rms error at a lower velocity is about 0.14 pixel/frame and the relative error at a higher velocity is about 4%. This shows the velocity can be obtained more reliably in wider area by using the 3PCIS combined with the WIM.

### *B. EVALUATION IN A REAL SCENE*

Fig. 13 shows an example of optical flow detection of a traffic scene. Some cars are passing through the intersection, and cars in the right-turn lane are moving forward slowly. The correlation image (b) is mostly dark in stationary regions and captures strongly various edges of moving objects.

(a) intensity image (b) correlation image (c) velocity by color

**FIGURE 13.** **Example of optical flow detection of a traffic scene at an intersection with a right-turn lane. (a) and (b) show raw images obtained by 3PCIS, and (c) is the optical flow result obtained by the Horn-Schunck iteration algorithm. The velocity vs color correspondence is the same as that in Fig. 11. Fast passing and slowly moving right-lane cars are indicated by red, and opposite and left-turn cars are indicated by light blue.**

In the optical flow detection result (c) obtained by the Horn–Schunck iteration algorithm (20 iterations), regions of forward passing cars and slowly marching right-lane cars are filled with a red hue, indicating leftward motion in the image plane with the velocity proportional to the brightness. Beyond the roofs of these cars, small light blue regions can be seen. They are responses to oncoming and opposite left-turning cars.

## VI. SUMMARY

We proposed a suitable design for the pixel and compensation operation of a 3PCIS for optical flow application. As inputs for the algebraic solution operation, it should be capable of near shot-noise-limited detection and accurate analog operation for Fourier integral measurements. We fabricated and compared several types of correlation detection block. For RTN and interference reduction, a p-epi channel, deep p-well structure was best for the 3P multiplier. Several experimental results were shown including the dense motion capture of a traffic scene.

## APPENDIX A
## ALGEBRAIC SOLUTION OF OPTICAL FLOW

Consider the irradiance $f$ on the surface of a moving object. Since it does not change with the motion of the object, the temporal derivative on the moving coordinate attached to the object surface is zero. Therefore,

$$\frac{\mathrm{D}f}{\mathrm{D}t} = \left(\boldsymbol{v}\cdot\nabla + \frac{\partial}{\partial t}\right) f(\boldsymbol{r}, t) = 0, \tag{7}$$

where $\boldsymbol{r} \equiv (x, y)$ is the pixel coordinate, $\boldsymbol{v}$ is the velocity of the moving surface, and $\nabla \equiv (\partial/\partial x, \partial/\partial y)$ and $\partial/\partial t$ are the spatial and temporal derivatives on the stationary coordinate.

The weighted integral method for the direct algebraic determination of optical flow is based on the identity relation in the frame interval $[0, T]$ using the finite Fourier transform:

$$\left(\boldsymbol{v}\cdot\nabla + \frac{\partial}{\partial t}\right) f(\boldsymbol{r}, t) = 0 \quad {}^{\forall}t \in [0, T]$$
$$\longleftrightarrow \int_0^T \left\{\left(\boldsymbol{v}\cdot\nabla + \frac{\partial}{\partial t}\right) f(\boldsymbol{r}, t)\right\} e^{-jn\Delta_\omega t} dt = 0 \quad {}^{\forall}n, \tag{8}$$

where $T$ is the frame time, $\Delta_\omega \equiv 2\pi/T$, and $n = 0, 1, 2, \ldots,$ is the order of Fourier coefficients. Then, the integration of the second equation by parts leads to

$$\begin{aligned} 0 &= \boldsymbol{v}\cdot\nabla \int_0^T f(\boldsymbol{r}, t) e^{-jn\Delta_\omega t} dt \\ &\quad + jn\Delta_\omega \int_0^T f(\boldsymbol{r}, t) e^{-jn\Delta_\omega t} dt + \left[f(\boldsymbol{r}, t) e^{-jn\Delta_\omega t}\right]_0^T \\ &= \boldsymbol{v}\cdot\nabla g_n(\boldsymbol{r}) + jn\Delta_\omega g_n(\boldsymbol{r}) + \left[f(\boldsymbol{r}, t)\right]_0^T, \end{aligned} \tag{9}$$

where

$$g_n(\boldsymbol{r}) \equiv \int_0^T f(\boldsymbol{r}, t) e^{-jn\Delta_\omega t} dt \tag{10}$$

is the $n$th order temporal Fourier coefficient of the incident light at the pixel location $\boldsymbol{r}$, i.e., the correlation image using the reference signal frequency $n\Delta_\omega$ ($n = 0$ indicates the intensity image). In the complex Eq. (9), the aperture problem is relaxed to the case when the real and imaginary parts of $\nabla g_n(\boldsymbol{r})$ have the same orientation. In contrast, the integral boundary term

$$\left[f(\boldsymbol{r}, t)\right]_0^T = f(\boldsymbol{r}, T) - f(\boldsymbol{r}, 0)$$

is the instantaneous intensity difference between the start and end of the frame, which is not observable by a 3PCIS. Including this and $\boldsymbol{v} = (v_x, v_y)$, Eq. (9) contains three unknowns for every location $\boldsymbol{r}$. Thus, solving it requires two or more equations with different frequency orders $n$.

The integral boundary term is directly related to the surface texture of the object; thus, can strongly vary with $\boldsymbol{r}$. In contrast, the optical flow velocity $\boldsymbol{v}$ remains locally constant on an object. In most visual scenes, the temporal spectrum of $f(\boldsymbol{r}, t)$ is concentrated at low frequencies. Therefore, we simultaneously solve the $n = 0$ and $n = 1$ equations

by eliminating the integral boundary term from them. This leads to

$$\boldsymbol{v} \cdot \nabla(g_1(\boldsymbol{r}) - g_0(\boldsymbol{r})) = -j\Delta_\omega g_1(\boldsymbol{r}). \tag{11}$$

To avoid the aperture problem, the local least squares method is expressed as

$$\boldsymbol{v}(\boldsymbol{r}) = \arg\min \int\int_{\Gamma(\boldsymbol{r}'-\boldsymbol{r})} \left|\boldsymbol{v}(\boldsymbol{r}') \cdot \nabla\big(g_1(\boldsymbol{r}') - g_0(\boldsymbol{r}')\big) + j\Delta_\omega g_1(\boldsymbol{r}')\right|^2 d\boldsymbol{r}', \tag{12}$$

where $\Gamma(\boldsymbol{r}' - \boldsymbol{r})$ is a small integral area (e.g., $3 \times 3$ pixels) around the pixel under consideration.

Another method is to use the Horn-Schunck global optimization algorithm [19]. It is based on the criterion:

$$\boldsymbol{v}(\boldsymbol{r}) = \arg\min \int\int_{\text{whole image}} \Big\{ \|\boldsymbol{v}(\boldsymbol{r}) - \bar{\boldsymbol{v}}(\boldsymbol{r})\|^2 + \lambda^2 |\boldsymbol{v}(\boldsymbol{r}) \cdot \nabla(g_1(\boldsymbol{r}) - g_0(\boldsymbol{r})) + j\Delta_\omega g_1(\boldsymbol{r})|^2 \Big\} d\boldsymbol{r}, \tag{13}$$

where $\bar{\boldsymbol{v}}(\boldsymbol{r})$ is the smoothed velocity field of $\boldsymbol{v}(\boldsymbol{r})$ and $\lambda^2$ is the Lagrange multiplier. Differentiating the criterion with respect to $\boldsymbol{v}(\boldsymbol{r})$ and equating it to zero, we obtain the iteration scheme

$$\boldsymbol{v}^{i+1}(\boldsymbol{r}) = \bar{\boldsymbol{v}}^i(\boldsymbol{r}) - \lambda^2 \Re\Big\{ \nabla\big(g_1^*(\boldsymbol{r}) - g_0(\boldsymbol{r})\big) \times \Big(\boldsymbol{v}^i(\boldsymbol{r}) \cdot \nabla(g_1(\boldsymbol{r}) - g_0(\boldsymbol{r})) - j\Delta_\omega g_1(\boldsymbol{r})\Big)\Big\}, \tag{14}$$

where the superscripts $i$ and $i+1$ indicate iteration counts.

## APPENDIX B
## 3P MULTIPLIER AND IMBALANCE CORRECTION

Let us express the drain currents of $Q_1, Q_2, Q_3$ as

$$I_i = I_{oi} e^{\frac{q\kappa(V_i - V_S)}{kT}}, \quad I_{oi} \equiv I_{si} e^{-\frac{q\kappa V_{Ti}}{kT}} \quad (i = 1, 2, 3),$$

where $V_{T1}, V_{T2}, V_{T3}$ are threshold voltages, and $I_{s1}, I_{s2}, I_{s3}$ are saturation currents. At the common source, it follows that

$$\begin{aligned} I &= I_1 + I_2 + I_3 \\ &= e^{-\frac{q\kappa V_S}{kT}} \Big\{ I_{o1} e^{\frac{q\kappa V_1}{kT}} + I_{o2} e^{\frac{q\kappa V_2}{kT}} + I_{o3} e^{\frac{q\kappa V_3}{kT}} \Big\}. \end{aligned}$$

By rewriting

$$\begin{aligned} &V_1 = \bar{V} + v_1, \quad V_2 = \bar{V} + v_2, \quad V_3 = \bar{V} + v_3, \\ &I_{o1} = I_o(1 + \varepsilon_1), \ I_{o2} = I_o(1 + \varepsilon_2), \ I_{o3} = I_o(1 + \varepsilon_3), \end{aligned}$$

where

$$\bar{V} = \frac{1}{3}(V_1 + V_2 + V_3), \quad I_o = \frac{1}{3}(I_{o1} + I_{o2} + I_{o3}),$$

it follows that

$$I_i = I \cdot \frac{I_o(1 + \varepsilon_i) e^{\frac{q\kappa}{kT} v_i}}{I_{o1} e^{\frac{q\kappa}{kT} v_1} + I_{o2} e^{\frac{q\kappa}{kT} v_2} + I_{o3} e^{\frac{q\kappa}{kT} v_3}} \quad (i = 1, 2, 3)$$

Dividing each equation by its imbalance ratio and neglecting second order terms of $\varepsilon_i$ and $q\kappa v_i / kT$ $(i = 1, 2, 3)$, we obtain

$$\begin{aligned} \frac{I_i}{1 + \varepsilon_i} &= \left( \frac{I_1}{1 + \varepsilon_1} + \frac{I_2}{1 + \varepsilon_2} + \frac{I_3}{1 + \varepsilon_3} \right) \times \frac{e^{\frac{q\kappa}{kT} v_i}}{e^{\frac{q\kappa}{kT} v_1} + e^{\frac{q\kappa}{kT} v_2} + e^{\frac{q\kappa}{kT} v_3}} \\ &\simeq \frac{1}{3} \left( \frac{I_1}{1 + \varepsilon_1} + \frac{I_2}{1 + \varepsilon_2} + \frac{I_3}{1 + \varepsilon_3} \right) + \frac{q\kappa}{3kT} I v_i \end{aligned}$$

By taking the temporal integral of the currents in the frame interval, this is expressed in the matrix-vector form as Eq. (2).

## REFERENCES

[1] S. Ando and A. Kimachi, "Correlation image sensor: Two-dimensional matched detection of amplitude-modulated light," *IEEE Trans. Electron Devices*, vol. 50, no. 10, pp. 2059–2066, Oct. 2003.

[2] S. Han, T. Sawada, T. Iwahori, S. Kawahito, and S. Ando, "Three-phase time-correlation image sensor using pinned photodiode active pixels," in *Proc. IS T/SPIE Electron. Imag.*, San Jose, CA, USA, 2010, Art. no. 75360S.

[3] S. Ando and S. Kawahito, "Development of 640×512 and 704×512 pixel image sensors," in *Proc. Joint Conf. Appl. Phys.*, Mar. 2012, p. 3-48.

[4] M. Habibi, "Analysis, enhancement, and sensitivity improvement of the correlation image sensor," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 3, pp. 708–718, Mar. 2012.

[5] M. Habibi and S. M. Sayedi, "Asynchronous demodulation technique for use in vision sensor image classification and segmentation," *Int. J. Circuit Theory Appl.*, vol. 39, no. 1, pp. 17–30, 2011.

[6] A. Kimachi and S. Ando, "Real-time phase-stamp range finder using correlation image sensor," *IEEE Sens. J.*, vol. 9, no. 12, pp. 1784–1792, Dec. 2009.

[7] A. Kimachi, S. Ando, M. Doi, and S. Nishi, "Three-phase quadrature spectral matching imager using correlation image sensor and wavelength-swept monochromatic illumination," *Opt. Eng.*, vol. 50, no. 12, 2011, Art. no. 127208.

[8] S. Ando, T. Nara, N. Ono, and T. Kurihara, "Real-time orientation-sensitive magnetooptic imager for leakage flux inspection," *IEEE Trans. Magn.*, vol. 43, no. 3, pp. 1044–1051, Mar. 2007.

[9] M. Habibi and M. Sayedi, "Geometric centre tracking of tagged objects using a low power demoduration smart vision sensor," *IET Circuits Devices*, vol. 4, no. 1, pp. 67–77, 2010.

[10] S. Ando and T. Nara, "An exact direct method of sinusoidal parameter estimation derived from finite Fourier integral of differential equation," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3317–3329, Sep. 2009.

[11] T. Nara and S. Ando, "Direct localization of poles of meromorphic function from measurements on incomplete boundary," *Inverse Problems*, vol. 26, no. 1, 2010, Art. no. 015011.

[12] S. Ando, T. Nara, and T. Levy, "Partial differential equation-based localization of a monopole source from a circular array," *J. Acoust. Soc. Am.*, vol. 134, no. 4, pp. 2799–2813, 2013.

[13] S. Ando, T. Nara, and T. Kurihara, "Spatial filtering velocimetry revisited: Exact short-time detecting schemes from arbitrarily small-size reticles," *Meas. Sci. Technol.*, vol. 25, no. 8, 2014, Atr. no. 085001.

[14] S. Ando, "Frequency-domain prony method for autoregressive model identification and sinusoidal parameter estimation," *IEEE Trans. Signal Process.*, vol. 68, pp. 3461–3470, Jun. 2020.

[15] S. Ando, D. Wei, and P. Masurel, T. Kurihara, "Complex-sinusoidally modulated imaging for optical flow detection," in *Proc. SIAM Conf. Imag. Sci.*, Minneapolis, MN, USA, 2006.

[16] D. Wei, P. Masurel, T. Kurihara, and S. Ando, "Optical flow determination with complex-sinusoidally modulated imaging," *Trans. Inst. Electron. Inf. Commun. Eng.*, vol. J90-D, no. 8, pp. 2009–2018, 2007.

[17] S. Ando, W. Dabi, and P. Masurel, "Optical flow detection via complex-sinusoidally modulated imaging and its realization with correlation image sensor," *Trans. Inf. Process. Soc. Japan*, vol. 49, no. 6, pp. 13–21, 2008.

[18] S. Ando, T. Kurihara, and D. Wei, "Exact algebraic method of optical flow detection via modulated integral imaging," in *Proc. Int. Conf. Comput. Vis. Theory Appl. (VISAPP)*, Lisbon, Portugal, 2009, pp. 480–487.
[19] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, nos. 1–3, pp. 185–203, 1981.
[20] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. 7th IJCAI*, 1982, pp. 674–679.
[21] C. Mead, *Analog VLSI and Neural Systems*. Boston, MA, USA: Addison-Wesley, 1989.
[22] A. A. Stocker, "Analog VLSI focal-plane array with dynamic connections for the estimation of piecewise-smooth optical flow," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 5, pp. 963–973, May 2004.
[23] D. Fortun, P. Bouthemy, and C. Kervrann, "Optical flow modeling and computation: A survey," *Comput. Vis. Image Understand.*, vol. 134, pp. 1–21, May 2015.
[24] M. Nagase, "Japanese patent application," J.P. Patent 1 835 02, 2017.
[25] Y. Higashi, N. Momo, H. S. Momose, T. Ohguro, and K. Matsuzawa, "Comprehensive understanding of random telegraph noise with physics simulation," in *Tech. Dig. Symp. VLSI Tech.*, 2011, pp. 200–201.
[26] J. Janesick, J. T. Andrews, and T. Elliot, "Fundamental performance differences between CMOS and CCD imagers; Part I," in *Proc. SPIE*, vol. 6276, 2006, pp. 1–19.
[27] R. Kuroda, A. Yonezawa, A. Teramoto, T.-L. Li, Y. Tochigi, and S. Sugawa, "A statistical evaluation of low frequency noise of in-pixel source follower-equivalent transistors with various channel types and body bias," in *Proc. IS T SPIE Electron. Imag.*, 2013, Art. no. 86590D.
[28] D. Leyris, F. Martinez, M. Valenza, A. Hoffmann, J. C. Vildeuil, and F. Roy, "Impact of random telegraph signal in CMOS image sensors for low-light levels," in *Proc. ESSCIRC*, 2006, pp. 376–379.
[29] K. Fujita *et al.*, "Advanced channel engineering achieving agressive reduction of $V_T$ variation for ultra-low-power applications," in *Proc. IEEE IEDM*, 2011, pp. 1–4.
[30] M.-W. Seo *et al.*, "A 1.1 $e^-$ rms temporal noise 87.5dB DR CMOS image sensors with low-noise transistors and column-parallel ADCs," in *Proc. IISW*, 2013, pp. 337–340.

**SHIGERU ANDO** (Member, IEEE) received the B.E., M.E., and Dr. Engr. degrees in mathematical engineering and information physics from the University of Tokyo in 1974, 1976, and 1979, respectively. He joined the Faculty of Engineering, University of Tokyo in 1979, served as a Associate Professor in 1987, as a Professor from 1995 to 2016, and currently a Professor Emeritus with the Department of Mathematical Engineering and Information Physics, and the Department of Information Physics and Computing. His interests are in image and signal processing, mathematical techniques in inverse problems, optical and acoustic sensing, and bio-inspired visual, auditory, and tactile sensor devices.

**MASANORI NAGASE** received the M.E. degree from Shizuoka University, Shizuoka, Japan, in 1989. He joined Matsushita Electric Industrial Company Ltd., Osaka, in 1989, Fujitsu VLSI Limited, Aichi, in 1991, Fujifilm Microdevices Company Ltd., Miyagi, in 2002, and Brookman Technology, Inc., Hamamatsu, Japan, in 2013, where he was involved in the research and development of CCD, CMOS image sensors, ion implantation, process integration, and TCAD. His current research interests include imager's pixels design and opt-device simulation.

**TAKASHI WATANABE** (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in physics from Osaka University, Osaka, Japan, in 1974, 1976, and 2000, respectively. He joined Sharp Corporation, Osaka, in 1976. In August 1976, he was worked to develop CCD signal processing ICs, CCD linear image sensors and area image sensors. In 1995, he led to develop CMOS image sensors and was concerned with first application to mobile phones in 2000. Since December 2007, he has been with Brookman Technology, Inc., Hamamatsu, Japan. Since April 2011, he has been a Visiting Professor with the Research Institute of Electronics, Shizuoka University. His interest is in device physics of image sensors.

**TOMOHIKO KOSUGI** received the B.S. degree from Chubu University, Aichi, Japan, in 2002. He was with Sanei Hytechs Company Ltd., from 2002 to 2009, and was engaged in the design of analog circuit. In 2010, he joined Brookman Technology, Inc., Hamamatsu, Japan, where he has been working on the development and design of CMOS image sensor and analog circuits.

**TETSUYA IIDA** received the B.S. and M.S. degrees from Shizuoka University, Hamamatsu, Japan, in 2008 and 2010, respectively. In 2010, he joined Brookman Technology, Inc., Hamamatsu, where he is engaged in the design of CMOS image sensor and analog circuits.