

# EXACT ALGEBRAIC METHOD OF OPTICAL FLOW DETECTION VIA MODULATED INTEGRAL IMAGING

## *Theoretical Formulation and Real-time Implementation using Correlation Image Sensor*

Shigeru Ando, Toru Kurihara

*Department of Information Physics and Computing, University of Tokyo, Tokyo, Japan*

Dabi Wei

*Department of Control and Systems Engineering, Tokyo Institute of Technology, Tokyo, Japan*

**Keywords:** Optical flow equation, Weighted integral method, Modulated imaging, Correlation image sensor, Computer vision, Velocity field measurement, Particle image velocimetry.

**Abstract:** A novel mathematical method and a sensing system that detects velocity vector distribution on an optical image with a pixel-wise spatial resolution and a frame-wise temporal resolution is proposed. It is provided by the complex sinusoidally-modulated imaging using the three-phase correlation image sensor (3PCIS) and the exact algebraic inversion method based on the optical flow identity (OFI) satisfied by an intensity image and a complex-sinusoidally modulated image captured by the 3PCIS. Since the OFI is free from time derivatives, any limitations on the object velocity and inaccuracies due to approximated time derivatives is thoroughly avoided. An experimental system was constructed with a  $320 \times 256$  pixel 3PCIS device and a standard PC for inversion operations and display. Several experimental results are shown including the dense motion capture of face and gesture and the particle image velocimetry of water vortices.

## 1 INTRODUCTION

Optical flow is the two-dimensional velocity field that describes the apparent motion of image patterns. The computation of optical flow has been widely studied since the 1970's motivated by broad range of applications such as detection and tracking of an object, separation from a background or more generally segmentation, three-dimensional motion computation, etc. One of the most established algorithm for optical flow determination both theoretically and practically will be the gradient-based method. It is based on the optical flow constraint (OFC) equation describing the intensity-invariance of moving patterns. According to B. McCane (McCane et al., 2001), Horn and Schunk's algorithm (Horn and Schunk, 1981) with regularization is still one of the best performer among numerous optical flow retrieval algorithms. The determination of optical flow from conventional image indeed is structurally very difficult due to the well-known aperture problem. It is stated clearly that optical flow cannot be retrieved locally for the information we get lo-

cally is one-dimensional; the conventional OFC determines the optical flow only in one direction (normal flow).

Recently, a novel imaging device, three-phase correlation image sensor (3PCIS), has been developed (Ando and Kimachi, 2003; Ando et al., 2007), which can capture, not only the intensity distribution as a conventional camera, but also produce simultaneously a complex-valued image consisting on quadrature demodulated components of time-varying intensity distribution. For the optical flow determination, the 3PCIS can offer an alternative and extended method to compute locally and therefore quickly the velocity field. A gradient-based algorithm for the images from 3PCIS was proposed and its basic properties were examined using simple images (Wei et al., 2007). It showed that, contrary to algorithms on conventional image, the computations are operated on only one frame, thus the temporal differentiation between frames which is one of the main sources of errors is not required and therefore the accuracy is not limited by the velocity (Ando et al., 2008).



where  $\Re$  and  $\Im$  denote the real and the imaginary part, respectively.

Two (real and imaginary) linear equations in it are sufficient to solve two real unknowns ( $v_x, v_y$ ). Combining with the gradient vector images  $\nabla g_0(x, y)$  and  $\nabla g_\omega(x, y)$ , we can compute the unknowns directly from one pixel and one frame. We hereafter call Eq.(7) the optical flow identity (OFI). The OFI is free from time derivative which is a severe source of error such as the finite differencing between frames and the temporal aliasing in large motion. For the spatial derivative in the OFI, the problem is far less susceptible by virtue of spatially dense image data and mostly ideal FFT-based gradient operation.

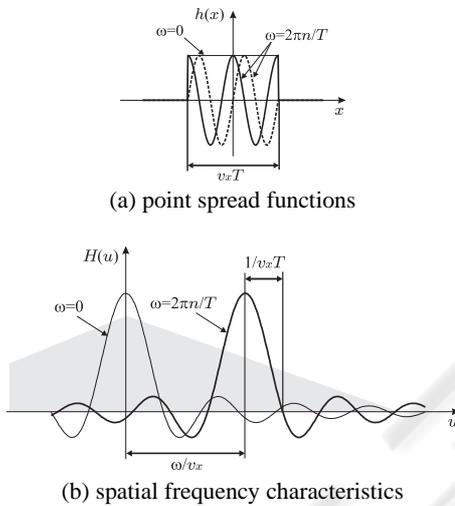


Figure 1: Point spread functions (a) and spatial frequency characteristics (b) in motion direction by means of complex sinusoidally modulated imaging.

## 2.4 Modulation Frequency

Performance of the proposed method depends on frequency  $\omega$  of the modulated imaging with relating to spatial frequency spectrum of moving objects. This relation is described by the point spread functions of motion blur and the corresponding spatial frequency characteristics:

$$h(x) \equiv \text{rect}\left(\frac{x}{v_x T}\right) e^{j(\omega/v_x)x}$$

$$\xrightarrow{\mathcal{F}} H(u) \equiv |v_x T| \text{sinc}\left(v_x T\left(u - \frac{\omega}{v_x}\right)\right). \quad (10)$$

Fig. 1 shows them when  $\omega = 0$  and  $\omega > 0$ . For  $\omega = 0$ , the image is low-pass filtered with bandwidth  $2/v_x T$  of the original spectrum (illustrated as a shaded area in (b)). In another image using  $\omega = 2\pi n/T$  ( $n > 0$ ), it is band-pass filtered with the same bandwidth  $2/v_x T$

but its center at  $\omega/v_x$ . According to an increase of velocity  $v_x$ , two spectral bands are shifting toward a low frequency side. This usually decreases the power of gradient image, thus decreases the accuracy of velocity. When the velocity is small or  $\omega$  is too large, the spectral band locates in higher frequencies, hence the power of image and the accuracy of velocity decrease. Therefore, some optimum values of  $\omega$  will exist with relating to the spatial frequency distribution of moving object.

## 2.5 Aperture Problem

The aperture problem is arisen when the determinant of matrix  $A$  in Eq. (8) becomes zero. The condition is illustrated in Fig. 2. Namely, it vanishes when  $\nabla \Re g_\omega - (-1)^n \nabla g_0$  and  $\nabla \Im g_\omega$  are parallel. If the moving pattern is truly uni-directional (same direction in wide spatial frequency range), the direction of  $\nabla g_\omega$  and  $\nabla g_0$  coincides, thus the aperture problem happens in the same way as the conventional optical flow detection. When the directions in different frequency bands are different, however, the directions are different, thus the aperture problem can be avoided. In the proposed method, two frequency bands for  $g_0$  and  $g_\omega$  are used, and the aperture problem is only when the uni-directionalities are common. Therefore, the conditions for aperture problem are relaxed in this sense.

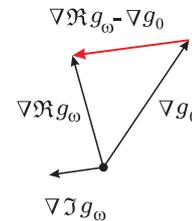


Figure 2: Condition for the aperture problem in the OFI when  $n$  of  $\omega T = 2n\pi$  is even. The velocity is indeterminate when  $\nabla \Re g_\omega - \nabla g_0$  is parallel to  $\nabla \Im g_\omega$ .

## 3 CORRELATION IMAGE SENSOR

### 3.1 Principle

As shown in Fig. 3(a), the 3PCIS consists of a photodiode PD, multiplier transistors  $Q_1, Q_2$ , and  $Q_3$ , capacitors of the same capacitance  $C$ , and readout transistors  $SW_1, SW_2$ , and  $SW_3$ . The photo-generated current  $I$  from the PD is split by  $Q_1, Q_2$ , and  $Q_3$  in proportion to their respective gate-source voltages  $V_1, V_2$ , and  $V_3$  into the drain currents

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} \simeq -\frac{q\kappa}{3kT} \begin{bmatrix} I(V_1 - \bar{V}) \\ I(V_2 - \bar{V}) \\ I(V_3 - \bar{V}) \end{bmatrix} + \frac{1}{3} \begin{bmatrix} I \\ I \\ I \end{bmatrix} \quad (11)$$

( $k$ : Boltzmann constant,  $T$ : absolute temperature,  $q$ : electron charge,  $\kappa$ : gate coefficient,  $\bar{V} \equiv (V_1 + V_2 + V_3)/3$ ), and accumulated at three independent capacitors  $C$ . By activating  $SW_1, SW_2$ , and  $SW_3$ , charges are transferred to external capacitors and the capacitors  $C$  are reset.

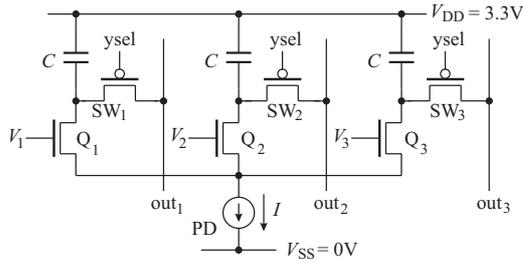


Figure 3: The pixel circuit of 3PCIS.  $V_1, V_2$ , and  $V_3$  are the reference inputs and  $I$  is an input from the PD. The charges stored in the capacitors  $C$  are the sums of the mean intensity  $\langle I \rangle$  and correlation  $\langle IV_i \rangle$  between  $I$  and  $V_i$  ( $i = 1, 2, 3$ ).



Figure 4: Lock-in camera using 3PCIS. This enables parallel correlation detection with two arbitrary analog orthogonal reference signals supplied in the three-phase form. The three correlation outputs are A/D-converted and transferred to a PC via the USB.

The  $200 \times 200$  pixel device was fabricated through the  $0.35\mu\text{m}$  2 poly-3 metal (2P3M) CMOS process provided by VLSI Design and Education Center (VDEC), Univ. Tokyo, and  $320 \times 256$  devices were fabricated through the  $0.35\mu\text{m}$  2P3M CMOS imager process by SHARP Corp., Japan. Several parameters and performances of these devices and cameras are summarized in Table 1.

### 3.2 Amplitude / Phase Recovery

Three outputs from each pixel of the 3PCIS can be converted into a background (time-averaged) intensity, a correlation amplitude, and a correlation phase

as follows. Let the time-varying intensity on a pixel at the coordinates  $(x, y)$  be

$$f(t) = A \cos(\omega t + \phi) + B + \xi(t). \quad (12)$$

here,  $\omega$  is the frequency of the modulated light,  $\phi$  is the phase,  $A$  is the amplitude,  $B$  is the stationary background intensity, and  $\xi(t)$  denotes any time-varying light components except for the frequency  $\omega$  and DC. As the reference signals of the 3PCIS, we input three sinusoidal waves whose frequency is  $\omega$  and whose initial phases are  $0, 2\pi/3$ , and  $4\pi/3$ . Then, the 3PCIS generates three outputs.

$$\begin{aligned} \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix} &= \left\langle f(t) \begin{bmatrix} \cos \omega t \\ \cos(\omega t + \frac{2\pi}{3}) \\ \cos(\omega t + \frac{4\pi}{3}) \end{bmatrix} + \frac{1}{3} \begin{bmatrix} f(t) \\ f(t) \\ f(t) \end{bmatrix} \right\rangle \\ &= \frac{AT}{2} \begin{bmatrix} \cos \phi \\ \cos(\phi - \frac{2\pi}{3}) \\ \cos(\phi - \frac{4\pi}{3}) \end{bmatrix} + \frac{BT}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \end{aligned} \quad (13)$$

Hence, the intensity is expressed as

$$B = \frac{1}{T}(R_1 + R_2 + R_3), \quad (14)$$

and the correlation amplitude  $A$  and the phase  $\phi$  are obtained as

$$\phi = \tan^{-1} \left( \frac{\sqrt{3}(R_2 - R_3)}{2R_1 - R_2 - R_3} \right) \quad (15)$$

$$A = \frac{2\sqrt{2}}{3} \sqrt{(R_1 - R_2)^2 + (R_2 - R_3)^2 + (R_3 - R_1)^2}. \quad (16)$$

The calculations are performed in real time by a PC from the A/D-converted outputs of the 3PCIS.

## 4 IMPLEMENTATION

An experimental system was constructed using a  $320 \times 256$  pixel 3PCIS and a PC (Core2 Duo<sup>(TM)</sup> @3.0GHz). Fig. 5 shows the block diagram of the camera and the algorithm in the PC.

### 4.1 “Phase-Stamp” Imaging

Usual applications of 3PCIS are to 2-D active sensing schemes in which reflected lights of modulated illumination or returned lights modulated by a given perturbation on the object are observed (Kimachi et al., 2001; Kimachi et al., 2002; Ando et al., 2007; Kimachi, 2007). Contrastively, the phase-stamp imaging is a passive scheme. Moving objects are imaged

Table 1: Specification of 3PCISs and lock-in cameras used in experiments.

|                    |  |  |
|--------------------|--|--|
| process            | CMOS 0.35 $\mu\text{m}$ 2P3M               | CMOS 0.35 $\mu\text{m}$ 2P3M               |
| image size         | 200 $\times$ 200                           | 320 $\times$ 256                           |
| pixel size         | 40 $\mu\text{m}$ $\times$ 40 $\mu\text{m}$ | 24 $\mu\text{m}$ $\times$ 24 $\mu\text{m}$ |
| chip size          | 9.8mm $\times$ 9.8mm                       | 9.81mm $\times$ 9.81mm                     |
| correlation SNR    | —  | $\sim$ 30dB                                |
| phase SNR          | $\sim$ 44dB                                | $\sim$ 47dB                                |
| cutoff frequency   | $\sim$ 400 $\times$ scan frequency         | $\sim$ 200 $\times$ scan frequency         |
| correlation output | charge readout, 3 $\phi$                   | charge readout, 3 $\phi$                   |
| A/D conversion     | external (12bit, 3ch)                      | internal (10 bit, 3ch)                     |
| frame rate         | 1.875 $\sim$ 15 frames/s                   | 5.7 $\sim$ 183.1 frames/s                  |

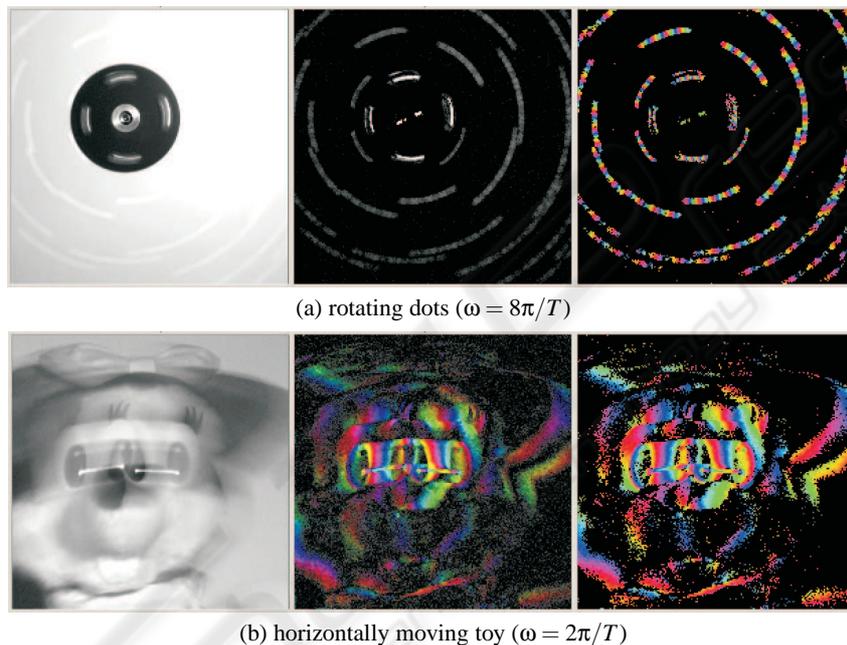


Figure 6: Examples of phase stamp imaging. From left to right of each, intensity image, correlation amplitude image, and correlation phase image (phase is indicated by hue).

while supplying sinusoidal reference signals to the 3PCIS.

The role of modulated imaging scheme can be readily understood when the objects moving are sparsely distributed points. When they are imaged with a sinusoidal reference signal with a fixed frequency, the passage time of a point is recorded as the phase of correlation image along its locus. Stationary objects and background do not appear in the correlation image although it is involved in the intensity image. An example is shown in Fig. 6(a). The objects are random dots on a rotating disk. The reference frequency is such that four cycles of the sinusoid is involved in a frame. In the correlation phase images, continuous phase distributions  $[0, 8\pi]$  are recorded along the loci.

For more complex objects, simple correspondence

between time and phase like above is lost because the loci of their components overlap each other. Nevertheless, rich information is involved in a particular manner described in the previous section. Fig. 6(b) shows an example. The correlation amplitude images are shown with the hue corresponding to the correlation phase. When the toy is moving, significant amplitudes appear along the boundary in the correlation image. In their perpendicular direction, the phase is changing regularly in proportion to the velocity. Contrarily in the intensity images, these regions are captured only as the blur.

## 4.2 Differential-of-Gaussian Filtering

This is to obtain gradients and smoothed versions of  $g_0(x, y)$  and  $g_\omega(x, y)$ . The filtering is performed in

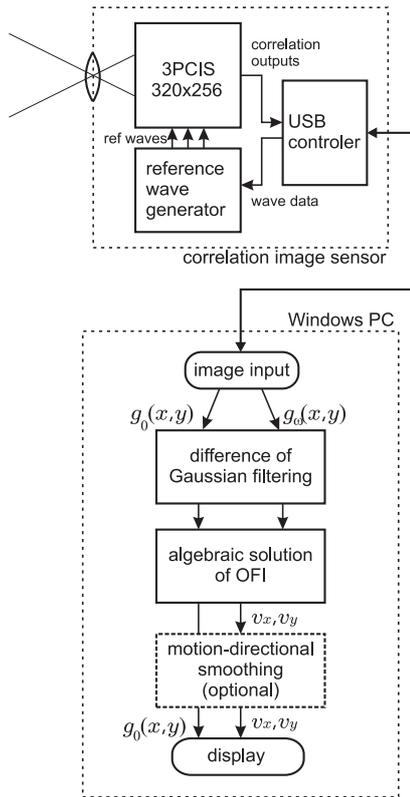


Figure 5: Block diagram of correlation camera using 3PCIS and algorithms performed by a PC.

Fourier domain using 2-D FFTs. The smoothing parameter  $\sigma$  (rms size of Gaussian kernel) is variable. In most applications,  $\sigma = 0.5$  [pixel length] is an appropriate choice to make a full use of spatial resolution. For fast moving object with low spatial frequency components,  $\sigma$  can be increased to fit the spectral passband to lower spatial frequency components. The use of consistent gradient operators(Ando, 2000) is an alternative choice for this stage.

### 4.3 Motion-directional Smoothing

During a frame time, each object keeps moving in its motion direction. This causes a smoothing effect in which the motion vector field is mostly continuous in the motion direction. Conversely speaking, we can smooth the motion vector field without a resolution loss so that noises are reduced and estimates are stabilized. In our system, it is performed in conjunction with the iterative local least squares estimate with varying spatial window(Lucas and Kanade, 1982; Ando, 1986). The window area is elongated along the previously estimated motion direction and the motion vector is estimated again.

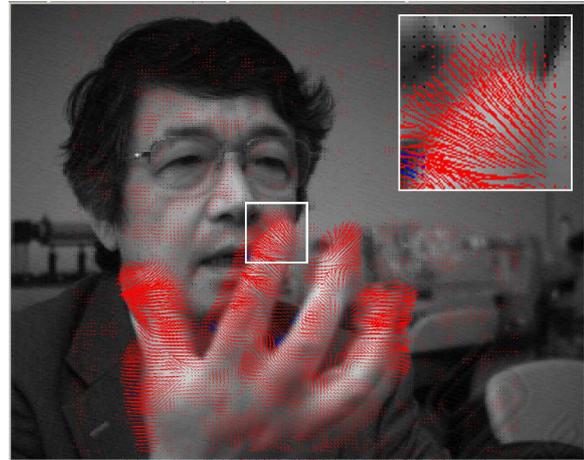


Figure 7: An example of optical flow detection of face and gesture. Detected optical flow vectors at every two pixels are indicated by directed lines. The length and direction are equal to a movement in a frame. The result is obtained in real time.



Figure 8: An example of optical flow detection of a traffic scene. The direction of motion is indicated by hue of line along with its length and direction.

## 5 EXPERIMENTS

### 5.1 Natural Scenes

Facial motion or gesture are a typical target of optical flow detection that requires high spatio-temporal resolution and wide dynamic range of motion velocity. Fig. 7 shows an example of application result. Detected nonzero optical flow vectors at every two pixels are indicated by directed lines at the pixels. The length and direction are equal to a movement in a frame. Without motion-directional smoothing, the detection rate of 14 frame/s (a half of frame rate) is

achieved for  $128 \times 128$  motion vectors. Simultaneous processing of image input, calculation, and display will increase the speed up to the frame rate. With the motion-directional smoothing, the speed reduces to about 7 frame/s. Fig. 8 shows another result for a traffic scene. A man on a motorbike going rightward and a car coming from behind a waiting car are captured well by length and direction (indicated by hue) of the velocity vectors.

## 5.2 Rapidly Moving Object

By an appropriate selection of sinusoidal frequency, this system can be adjusted to a wide range of motion speed. Fig. 9 and Fig. 10 show examples to a rotating disk when the rotation is stopped, slow, and fast. For clarity, the direction of motion is indicated by hue of line along with its length and direction. The test chart on the disk is so designed that both the spatial frequency and temporal frequency when it is rotated is larger in the central zone. When the rotation speed is slow, most temporal frequencies are involved in correlation bandwidth of  $n = 1$  sinusoid ( $\omega$  is equal to the frame frequency). But when it is high, the frequency generated at the central exceeds the bandwidth. In this case, we doubled ( $n = 2$ ) the reference signal frequency to capture the high temporal frequency as shown in Fig. 10. Optical flow at there is captured successfully. But this causes lower cutoff in the marginal zone, hence the optical flow is not detected there.

## 5.3 Particle Image Velocimetry

The proposed sensor is very suitable for the particle image velocimetry (PIV). Fig. 11 shows an example of real-time measurement of velocity distribution of a water surface marked by aluminum powder is detected. The direction of motion is indicated by hue of line along with its length and direction. In conventional PIV, the velocity is obtained only particle-wise by tracking the centroid of each particle. Contrarily in this technique, the resolution is not limited by the particle size and density. Except for pixels where no fraction of particle is passing in a frame time, the optical flow can be detected based on the spatio-temporal intensity changes.

## 6 SUMMARY

A novel sensing scheme and algorithm for optical flow detection with maximal spatio and temporal resolution was proposed. An experimental system was

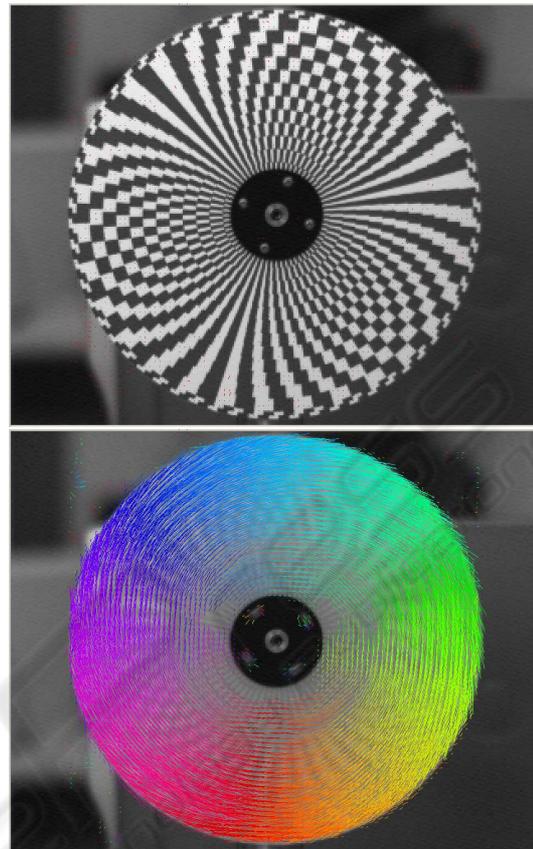


Figure 9: Optical flow of a rotating disk with wide ranging spatial frequencies. The upper image shows the object when it is stopped, and the lower image shows the detection result of optical flow. The direction of motion is indicated by hue of line along with its length and direction.

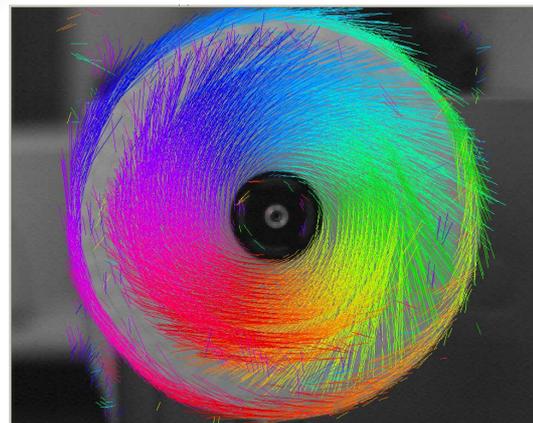


Figure 10: Optical flow of fast rotating disk. The object is same as Fig. 9. The reference signal frequency is increased to capture the high temporal frequency of the central zone. The frequency is below cutoff in the marginal zone, hence the optical flow is not detected.

constructed with a  $320 \times 256$  pixel 3PCIS. Several experimental results are shown using this system.

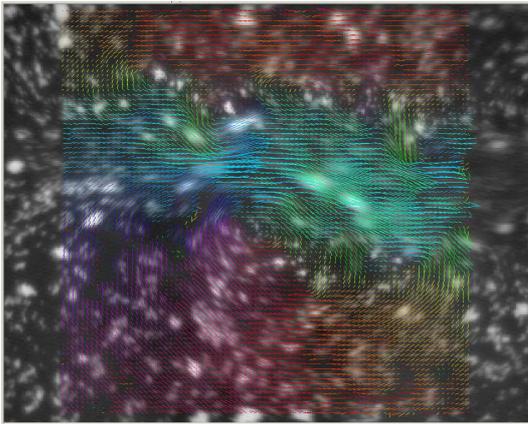


Figure 11: An example of real-time particle image velocimetry. Velocity distribution of a water surface marked by aluminum powder is detected. The direction of motion is indicated by hue of line along with its length and direction.

Device development of this work was supported by VLSI Design and Education Center, University of Tokyo, and Advanced Technology Research Laboratory, Sharp Corporation.

## REFERENCES

- Ando, S. (1986). A velocity vector field measurement system based on spatio-temporal image derivative. In *Trans. SICE*, vol.22, pp.1330-1336.
- Ando, S. (2000). Consistent gradient operators. In *IEEE Trans. Pattern Anal. Machine Intell.*, vol.22, no.3, pp.252-265.
- Ando, S. and Kimachi, A. (2003). Correlation image sensor: Two-dimensional matched detection of amplitude modulated light. In *IEEE Trans. Electron Devices*, vol.50, no.10, pp.2059-2066.
- Ando, S., Nara, T., Ono, N., and Kurihara, T. (2007). Real-time orientation-sensitive magneto-optic imager for leakage flux inspection. In *IEEE Trans. Magnetics*, vol.43, no.3, pp.1044-1051.
- Ando, S., Wei, D., and Masurel, P. (2008). Optical flow detection via complex-sinusoidally modulated imaging and its realization with correlation image sensor. In *Trans. IPSJ, Computer Vision and Image Media*, vol.49, no.6, p.13-21.
- Horn, B. K. P. and Schunk, B. G. (1981). Determining optical flow. In *Artificial Intell.*, vol.17, pp.185-203.
- Kimachi, A. (2007). Real-time heterodyne imaging interferometry: focal-plane amplitude and phase demodulation using a three-phase correlation image sensor. In *Applied Optics*, vol.46, no.1, pp.87-94.
- Kimachi, A., Imaizumi, T., Kato, A., and Ando, S. (2002). Spectral matching imager using correlation image sensor. In *Trans. IEEJ*, vol.122-E, no.4, pp.200-206.

- Kimachi, A., Kurihara, T., Takamoto, M., and Ando, S. (2001). A novel range finding system using correlation image sensor. In *Trans. IEEJ*, vol.121-E, no. 7, pp.367-375.
- Lucas, B. D. and Kanade, T. (1982). An iterative image registration technique with an application to stereo vision. In *Proc. 7th IJCAI*, pp.674-679.
- McCane, B., Novins, K., Crannitch, D., and Galvin, B. (2001). On benchmarking optical flow. In *Comput. Vis. Image Underst.*, vol.84, no.1, pp.126-143.
- Wei, D., Masurel, P., Kurihara, T., and Ando, S. (2007). Optical flow determination with complex-sinusoidally modulated imaging. In *Trans. IEICE*, vol.190-D, no.8, pp.2009-2018.