Local Estimation of High Velocity Optical Flow with Correlation Image Sensor

Hidekata Hontani, Go Oishi, and Tomohiro Kitagawa

Department of Computer Science
Nagoya Institute of Technology, Japan

Abstract. In this article, the authors address a problem of the estimation of high velocity optical flow. When images are captured by conventional image sensors, the problem of the optical flow estimation is ill-posed if only the temporal constancy of the image brightness is the valid assumption. When given images are captured by the correlation image sensors, though, you can make the problem of the optical flow estimation well-posed under some condition and can locally estimate the unique optical flow at each pixel in each single frame. The condition though would not be satisfied when the flow velocity is high. In this article, we propose a method that can estimate the normal component of high velocity optical flow using only the local image values in each single frame. The equation used for estimating the normal velocity is theoretically derived and the condition the equation holds is also revealed.

Keywords: optical flow, high velocity flow, real time estimation, time correlation image sensor.

1   Introduction

Ill-posedness is a fundamental problem for estimating optical flow. Most techniques of the optical flow estimation use one basic model that represents the temporal constancy of image brightness, $I(x, y)$: the image brightness of a particular point in the moving pattern is assumed temporally constant. Horn and Schunk derived one linear equation from this assumption[1]:

$$(u\partial_x + v\partial_y + \partial_t)I(x, y, t) = 0, \tag{1}$$

where ∂_* denotes the partial derivative with respect to the subscript variable, and a two-vector, $\boldsymbol{v} = [u, v]^T$, denotes the flow vector at (x, y). The problem of the flow estimation is ill-posed[2] because one single constraint (1) cannot determine the values of the two unknowns u and v, uniquely. Even if the corresponding edge is curved, as long as you use only the model (1), you cannot make the problem well-posed: Multiple pixels bring more unknowns than constraints. You can estimate only the component of the flow parallel to the spatial gradient of $I(x, y)$ if only the linear constraint equation (1) is available[2].

For determining the 2D flow uniquely, you need another model that constrains the flow based on a different aspect of the optical flow. Many models have

been proposed for that purpose[3] and they can be classified into two categories – local models and global ones. The former represents local coherence of the flow[4][5]: e.g., the unknown optical flow vector is assumed constant within some neighborhood[5]. The latter represents some global characteristics of the flow distribution, e.g. partly smooth distribution, and is employed by the variational methods[6][7][8] and by the discrete optimization ones[9][10]. You can find large varieties of the models for representing the characteristics of optical flow field and of the optimization algorithms, and the temporal constancy of the image brightness is still a basic assumption for all of these methods[3]. Those local/global models represent the prior knowledge of flow fields and are distinct from the temporal constancy model, which represents an aspect of measurement. In this article, the temporal constancy model is focused on.

Recently, a correlation image sensor has been developed by Wei et al. [11]. It has been demonstrated that, under certain conditions, you can estimate optical flows based only on the temporal constancy model when the input images are captured by the correlation image sensor[11]. Conventional image sensors capture images by integrating light intensity over the period the shutter is open, and this time integration eliminates the information of the temporal change of the light intensity. On the other hand, a correlation image sensor has three channels and measures not only the brightness but also the cross-correlation between the temporal change of the incident light intensity and reference signals supplied by you. The value of the cross-correlation measured at each pixel contains some information of the temporal change of the light intensity during the shutter is open, and is independent of the brightness. You can hence make the problem of the optical flow estimation well-posed and can uniquely estimate the optical flow by using the measurements obtained in a local small aperture based only on the temporal constancy model, under some conditions. The details will be described later.

Unfortunately, you can not always make the problem of the optical flow estimation well-posed even if the images are captured by correlation image sensors. The estimation would be ill-posed when the flow velocity is high and when the brightness pattern is heavily blurred by the high velocity motion. When the velocity is high, the linear constraint equation (1) does not hold well and the heavy motion blur smooths out the local structures of the pattern that are required for the optical flow estimation. This is why the proposed method is required for estimating high velocity flows. Figure 1 shows the outlines of the methods of optical flow estimation. The conventional methods estimate flows by using two images consecutively captured by a conventional image sensor (Fig.1(A)). The Wei's method[11] estimates flows by using three images simultaneously captured by a time correlation image sensor (Fig.1(B)), and the proposed method computes the normal components of the flows from the spatial gradient of the phase computed from two of the three images (Fig.1(C)).

In this article, we propose a method that estimates the normal components of high velocity flows based only on the measurements obtained in a small spatial window, e.g., a 3×3 pixels, in each *single* frame, even when the image brightness
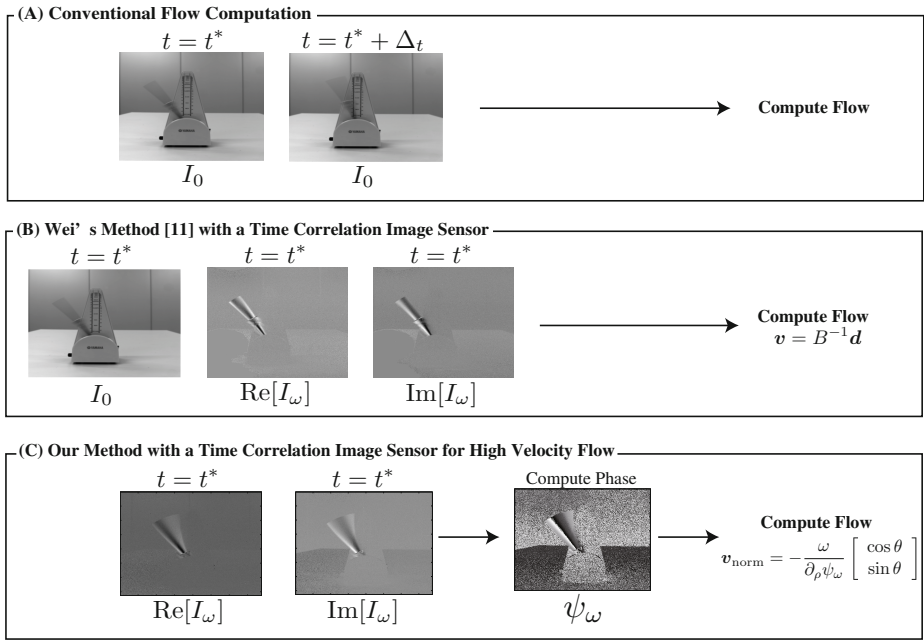
Fig. 1. Comparison between conventional methods(A), Wei's method(B)[11], and our proposed method(C). Conventional methods use two consecutive images captured at different times. The Wei's method[11] computes optical flows by using three images simultaneously captured by a time correlation image sensor. Our method computes normal components of high velocity flows by using two of the images captured by the time correlation image sensor.

pattern is heavily blurred by the motion. Analogous to the Wei's method, our proposed method uses the time correlation image sensor for measuring both the brightness and the complex Fourier coefficient corresponding to a specific frequency determined by the reference signals at each pixel. Then, the proposed method computes the spatial gradients of the phases of the complex Fourier coefficient, and estimates the normal component of the high velocity flow at each pixel by using the computed spatial gradient of the phase. The details of the correlation image sensors and of the algorithm of the Wei's method[11] are explained in the remainder of this section. Our proposed method is then described in the next section.

1.1   Correlation Image Sensor

A correlation image sensor has an array of pixel circuits, each of which measures the temporal cross-correlations between the light intensity and three reference signals. Let $f(x, y, t)$ denote the light intensity, where (x, y) denotes the location of the pixel circuit on the sensor and t denotes the time during the shutter of the

camera is open. Let $w_i(t)$ $(i = 1, 2, 3)$ denote the i-th reference signal supplied to every pixel circuit and let T denote the shutter speed, which is the time interval the image sensor is exposed to light. Each pixel has three channels and the pixel value of each channel is determined as follows[11]:

$$R_i(x, y) = \int_{-T/2}^{T/2} f(x, y, t)w_i(t) + \frac{1}{3}f(x, y, t)dt. \tag{2}$$

Conventional image sensors, on the other hand, measure just the temporal integration of $f(x, y, t)$ at each pixel and obtains the pixel values, $I_0(x, y)$, as follows:

$$I_0(x, y) = \int_{-T/2}^{T/2} f(x, y, t)dt. \tag{3}$$

The temporal integration in (3) eliminates the information of the temporal change of $f(x, y, t)$ during the shutter is open, and causes the motion blur.

Using a correlation image sensor, one can obtain not only $I_0(x, y)$ in (3) but also a complex Fourier coefficient of $f(x, y, t)$ corresponding to a specific frequency, ω, if one inputs the three sinusoidal waves shown in (4) as the reference signals:

$$w_i(t) = \cos\left(\omega t + (i - 1)\frac{2\pi}{3}\right). \tag{4}$$

In this paper, we assume that the Fourier coefficients of the temporal signal, $f(x, y, t)$, are well defined. Let f_ω denote the temporal Fourier coefficient: $f_\omega = A_\omega e^{j\phi_\omega}$, where A_ω denotes the amplitude and ϕ_ω denotes the phase. One can estimate A_ω and ϕ_ω as shown in (5) and (6) when $w_i(t)$ in (4) are the reference signals[11]:

$$\phi_\omega = \tan^{-1}\left(\frac{\sqrt{3}(R_2 - R_3)}{2R_1 - R_2 - R_3}\right), \tag{5}$$

$$A_\omega = \frac{2\sqrt{2}}{3}\sqrt{(R_1 - R_2)^2 + (R_2 - R_3)^2 + (R_3 - R_1)^2}. \tag{6}$$

In addition to the values of ϕ_ω and A_ω, one can estimate the value of I_0 shown in (3) as follows:

$$I_0(x, y) = R_1 + R_2 + R_3. \tag{7}$$

R_1, R_2, and R_3 are measured by a correlation sensor, and the values of I_0, A_ω, and ϕ_ω are calculated by a computer for each pixel in every frame. Using these values, you can represent the temporal signal $f(x, y, t)$ as follows:

$$f(x, y, t) = \frac{I_0}{T} + A_\omega \cos(\omega t + \phi_\omega) + \xi(t), \quad (-T/2 < t < T/2) \tag{8}$$

where $\xi(t)$ denotes any time-varying components except for the frequency, ω. A_ω and ϕ_ω carry some information of the temporal change of $f(x, y, t)$. Here, it

should be reminded that a target signal is (implicitly) assumed to be periodic when you compute its Fourier coefficients: A signal, $f_{\text{recon}}(x, y, t)$, reconstructed from the Fourier coefficients is periodic and satisfies $f_{\text{recon}}(x, y, t + mT) = f(x, y, t)$ $(-T/2 < t < T/2)$ where m is an integer.

1.2 Making Optical Flow Problem Well-Posed

The optical flow estimation is an ill-posed problem if a conventional image sensor is used and if only the temporal constancy is the valid assumption. Under the constancy assumption, the following equation is satisfied for any $t \in [-T/2, T/2]$:

$$\frac{df}{dt}(x(t), y(t), t) = 0. \tag{9}$$

For small displacement, a first order Taylor expansion yields the optical flow constraint of the incident light intensity:

$$(u\partial_x + v\partial_y + \partial_t)f(x, y, t) = 0. \tag{10}$$

You cannot use directly the equation (10) for the flow estimation because no image sensor can measure the instantaneous light intensity, $f(x, y, t)$. The flow, (u, v), should be constrained by the pixel values, $I_0(x, y)$, if images are captured by conventional image sensors. Integrating the equation (10) over the period the shutter is open, one obtains the optical flow constraint equation proposed by Horn and Schunk[1] that represents the relationship between the flow, (u, v), and the partial differential coefficients of $I_0(x, y)$:

$$\int_{-T/2}^{T/2} (u\partial_x + v\partial_y + \partial_t)f(x, y, t)dt = (u\partial_x + v\partial_y + \partial_t)I_0(x, y) = 0. \tag{11}$$

This single linear equation is well satisfied in many cases[3], but is evidently insufficient for estimating unique values of the two variables, u and v, for each pixel: The estimation of (u, v) is ill-posed here, and you need other information of (u, v) for restricting the solutions. For example, some local methods assume that the flow is spatially constant in each local area[5] and some global methods assume optical flow fields are partly smooth[6]. In our proposed method, the information comes from the Fourier coefficient measured by a correlation image sensor.

The optical flow estimation can be a well-posed problem under some conditions even when only the temporal constancy is the valid assumption, if the images are captured by correlation image sensors[11]. You can derive a system of independent linear equations of (u, v) from the constancy assumption (10). Integrating the equation (10) with a modulating function, $w(t) = e^{-j\omega t}$, over the exposure time interval, one obtains the following equation, which describes

the relationship between the Fourier coefficient and the flow:

$$\int_{-T/2}^{T/2} \{(u\partial_x + v\partial_y + \partial_t)f(x,y,t)\} e^{-j\omega t} dt =$$

$$(u\partial_x + v\partial_y + j\omega)I_\omega(x,y) + [f(x,y,t)e^{-j\omega t}]_{-T/2}^{T/2} = 0, \tag{12}$$

where $I_\omega(x,y,t)$ is the complex Fourier coefficient, which can be obtained by a correlation sensor as

$$I_\omega(x,y) = \int_{-T/2}^{T/2} f(x,y,t)e^{-j\omega t} dt = A_\omega(\cos\phi_\omega - j\sin\phi_\omega). \tag{13}$$

It should be noted that the temporal partial differentiation is eliminated in (12) by using the integration by parts. As described in [11], setting the angular frequency of the reference signals as

$$\omega = \frac{2n\pi}{T}, \tag{14}$$

you can derive from (12) a system of two linear equations such that

$$B\boldsymbol{v} = \boldsymbol{d}, \tag{15}$$

where $\boldsymbol{v} = [u,v]^T$, $\boldsymbol{d} = [\omega\mathrm{Im}[I_\omega], -\omega\mathrm{Re}[I_\omega]]^T$, and

$$B = \begin{bmatrix} \partial_x\{\mathrm{Re}[I_\omega] - (-1)^n I_0\} & \partial_y\{\mathrm{Re}[I_\omega] - (-1)^n I_0\} \\ \partial_x\mathrm{Im}[I_\omega] & \partial_y\mathrm{Im}[I_\omega] \end{bmatrix}. \tag{16}$$

You can estimate the flow, $\boldsymbol{v}$, by solving (16) as $\boldsymbol{v} = B^{-1}\boldsymbol{d}$. The equation shown in (15) is employed for the estimation in the remainder of this article.

The information used for the estimation of the flow is very local. When you use a $\Delta \times \Delta$ difference operator (e.g. $\Delta = 3$ pixels) for computing each of the spatial differentiations in (16), you can estimate the flow vector for each pixel based only on the measurements, I_0 and I_ω, obtained at the neighboring $\Delta \times \Delta$ pixels in each single frame. Temporal difference operations are not required for the flow estimation because the temporal differentiation is eliminated in (12). The estimation of the optical flow is now well-posed if $\det B \neq 0$. This condition is well satisfied, e.g., around the boundaries of moving objects and in moving textured regions.

2   Local Estimation of Optical Flow

In this section, our proposed method for locally estimating high velocity optical flow is described. Before describing the method, we discuss the spatio-temporal characteristics of $f(x,y,t)$ and the aperture problem.
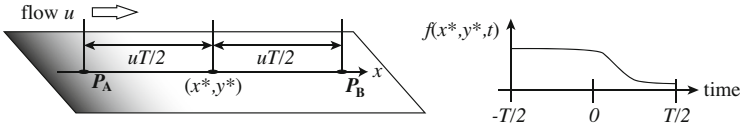
Fig. 2. A spatial trajectory of a point moving at the velocity, $\bar{u}$. The right graph shows the temporal change of $f(x^*, y^*, t)$.

2.1   Spatio-Temporal Structures of Light Intensity

A spatial trajectory of a point during the time the shutter is open is firstly considered. As shown in (3) and in (13), the pixel values, $I_0(x^*, y^*)$ and $I_\omega(x^*, y^*)$ measured at a point (x^*, y^*) are determined by the temporal change of the instantaneous light intensity, $f(x^*, y^*, t)$ $(t \in (-T/2, T/2))$. In this subsection, we assume that $\boldsymbol{v}(x^*, y^*) = [\bar{u}, 0]^T$ $(\bar{u} \neq 0)$ without loss of generality and that the velocity is spatially and temporally coherent in a local aperture around $(x^*, y^*, 0)$. The unit of the velocity is pixel per second and the time needed for moving to a neighbor pixel is $1/\bar{u}$ seconds. Then, the location of a point that passes through (x^*, y^*) at $t = 0$ is described as follows:

$$[x(t), y(t)]^T = [\bar{u}t + x^*, y^*]^T. \tag{17}$$

The spatial trajectory of the point between $[-T/2, T/2]$ is a line segment between two points, $P_A = (x^* - \bar{u}T/2, y^*)$ and $P_B = (x^* + \bar{u}T/2, y^*)$ as shown in Fig.2. Using $x^* = x(t) - \bar{u}t$, you can transform the temporal integrations shown in (3) and in (13) into spatial ones:

$$I_0(x^*, y^*) = \frac{1}{\bar{u}} \int_{-\bar{u}T/2}^{\bar{u}T/2} f(x^* - s, y^*, 0)ds, \tag{18}$$

$$I_\omega(x^*, y^*) = \frac{1}{\bar{u}} \int_{-\bar{u}T/2}^{\bar{u}T/2} f(x^* - s, y^*, 0)e^{-j\omega s/\bar{u}}ds, \tag{19}$$

where $s = \bar{u}t$. $I_0(x, y)$ is a spatial convolution between $f(x, y, 0)$ and a box filter, $b_u(x)$, where

$$b_u(x) = \begin{cases} 1/\bar{u}, & -\bar{u}T/2 < x < \bar{u}T/2, \\ 0, & \text{otherwise.} \end{cases} \tag{20}$$

This convolution represents the motion blur, of which scale is proportional to $\bar{u}$ and T[1]. $I_\omega(x, y)$ is a spatial convolution between $f(x, y, 0)$ and a complex partial sinusoidal wave, $g_u(x)$, where

$$g_u(x) = \begin{cases} \bar{u}^{-1}e^{-j\omega x/\bar{u}}, & -\bar{u}T/2 \leq x \leq \bar{u}T/2, \\ 0, & \text{otherwise.} \end{cases} \tag{21}$$

This convolution extracts the component of a specific spatial frequency, $\omega/\bar{u}$, from the profile of $f(x, y, 0)$ along the line segment $\overline{P_A P_B}$.

---

[1] $\int b_u(x)dx = T$ means the images become brighter when the exposure time is longer.

2.2   Temporal Phase and High Velocity Flow

The solution of (15) is exact if the first order Taylor expansion (10) of the constancy assumption exactly holds. When $\boldsymbol{v} = [\bar{u}, 0]^T$, the Taylor expansion of the constancy assumption yields $(\bar{u}\partial_x + \partial_t)f(x, y^*, t) = 0$. The term of ∂_y disappears from the constraint equation (10) because $dy(t)/dt = 0$. Let us assume that $f(x, y^*, 0)$ is linear with respect to x and that the first order Taylor expansion exactly holds: i.e., $f(x, y^*, 0) = ax + b$ $(a \neq 0)$. Then, setting $n = 1$ in (14) for simplicity, you obtain $I_0(x, y^*) = (ax + b) \otimes b_u(x) = (ax + b)T/\bar{u}$ and $I_\omega(x, y^*) = (ax + b) \otimes g_u(x) = -aT/j\omega$, and have the system of the equations in (15) as follows:

$$B\boldsymbol{v} = \begin{bmatrix} aT/\bar{u} & \epsilon \\ 0 & \delta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} aT \\ 0 \end{bmatrix}. \tag{22}$$

where ϵ and δ are not determined by the constraint equation. You can obtain a unique and exact solution, $\hat{\boldsymbol{u}} = [\bar{u}, 0]^T$, if $\epsilon \neq 0$ or $\delta \neq 0$. In other words, you do not have the aperture problem when $\partial_y \mathrm{Re}[I_\omega] \neq (-1)^n \partial_y I_0$ or $\partial_y \mathrm{Im}[I_\omega] \neq 0$. These conditions would be satisfied when I_0 or I_ω spatially changes along the direction perpendicular to the flow vector.

When the flow velocity is high, though, the solution of (15) might be inaccurate even if the light intensity of each moving point is temporally constant as assumed in (9). Higher velocity flow smooths out the spatial structures of $f(x, y, 0)$ heavily and generates more blurred images, $I_0(x, y)$. This results in that $\partial_x I_0(x, y) \simeq 0$ when $\boldsymbol{v} = [\bar{u}, 0]^T$ and that the calculated values of $\partial_x I_0(x, y)$ in (16) are unreliable because of noises and of quantization. In other words, the constraint equation of I_0 shown in (11) constrains little about the optical flow. This is one of the main reasons why you need some information other than the optical flow constraint for obtaining large displacement flow(e.g. [12]). When the equation (11) constrains u and v little, the estimation of the optical flow is again ill-posed even if the images are captured by correlation sensors. Now, the aperture problem returns and only the complex image, $I_\omega(x, y)$, contains information useful for locally estimating the optical flow.

Let an isophase curve that passes through (x^*, y^*) be denoted by $\mathcal{C}$, where $\mathcal{C} = \{(x, y) | \phi_\omega(x, y) = \phi_\omega(x^*, y^*)\}$. Assuming some coherency on the spatial pattern of $f(x, y, 0)$ and on the spatial distribution of optical flow, you can assume $\mathcal{C}$ is continuous and smooth, and the temporal signals, $f(x, y, t)$, measured on $\mathcal{C}$ are similar. The normal component estimated by the proposed method is a component of the optical flow perpendicular to the isophase curve, $\mathcal{C}$. Let an orthogonal curve of the isophase curves that passes through (x^*, y^*) be denoted by $\mathcal{T}$ (see Fig.3). This curve, $\mathcal{T}$, is a pseudo trajectory of a moving point that passes through (x^*, y^*), and let assume, without loss of generality, $\mathcal{T}$ is a line parallel to the x-axis around (x^*, y^*). The coordinates of a point in $\mathcal{T}$ is represented by $(x^* + \rho, y^*)$ and the positive direction of ρ is identical with that of the flow. The temporal signal, $f(x^*, y^*, t)$ $(t \in (-T/2, T/2))$, is determined by the spatial profile of the light intensity along $\mathcal{T}$, $f^*_{\mathcal{T}}(\rho) = f(x^* + \rho, y^*, 0)$.
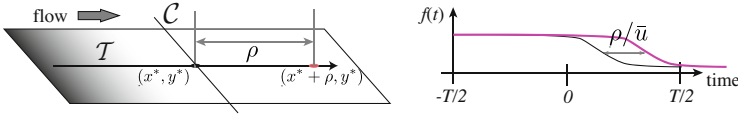
Fig. 3. The time delay observed at a neighboring pixel. The red curve in the right graph is measured at $(x^* + \rho, y^*)$ when the black one is measured at (x^*, y^*).

At the neighboring point, you measure a temporal signal identical to $f(x^*, y^*, t)$ with a time delay, $\rho/\bar{u}$:

$$f(x^* + \rho, y^*, t) = f(x^*, y^*, t - \rho/\bar{u}) = f(x^*, y^*, t) \otimes \delta(t - \rho/\bar{u}), \qquad (23)$$

where the convolution with the delta function is estimated with respect to the time t. Let again

$$I_\omega(x^*, y^*) = A_\omega e^{-j\phi_\omega} = \int_{-T/2}^{T/2} f(x^*, y^*, t) e^{-j\omega t} dt, \qquad (24)$$

where A_ω is the amplitude and ϕ_ω is the phase at (x^*, y^*). Substituting (24) for (23), you obtain the following equation, in which the phase of $I_\omega(x^* + \rho, y^*)$ is proportional to the spatial distance, ρ:

$$I_\omega(x^* + \rho, y^*) = A_\omega e^{-j\phi_\omega(\rho)} = A_\omega e^{-j(\phi_\omega - \rho\omega/\bar{u})}. \qquad (25)$$

In case the equation (25) holds, you can estimate the normal speed, $\bar{u} = \|\boldsymbol{v}_{\text{norm}}\|$, by using the value of the spatial derivative of the phase,

$$\frac{\partial \phi_\omega(\rho)}{\partial \rho} = -\frac{\omega}{\bar{u}}. \qquad (26)$$

Unfortunately, though equation (23) always holds, equation (25) does not always hold because of the fixed and bounded integration interval. At least the condition, $f(x^*, y^*, -T/2) = f(x^*, y^*, T/2)$, should be satisfied in order for equation (25) to hold. This is because, if $f(x^*, y^*, -T/2) \neq f(x^*, y^*, T/2)$, the reconstructed signal, $f_{\text{recon}}(t)$, is discontinuous at $t = -T/2 + mT$ and these discontinuous points are fixed and independent from the time-delay. For example, when a flying bright particle passes in front of the sensor with a high speed, $\bar{u}$, then the temporal change of the light intensity can be well represented by a delta function: $f(x^* + \rho, y^*, t) = A\delta(t - \rho/\bar{u})$, and you obtain $I_\omega(x^* + \rho, y^*) = Ae^{-j\omega\rho/\bar{u}}$ from (24). In this example, equation (25) holds and you can estimate the speed as $\bar{u} = \omega/(\partial_\rho\phi_\omega)$. On the other hand, for example, when the temporal change of the light intensity is represented by a step function: $f(x^* + \rho, y^*, t) = B \times h(t - \rho/\bar{u})$, where

$$h(t) = \begin{cases} 1, & \text{if } t \geq 0, \\ 0, & \text{otherwise,} \end{cases} \qquad (27)$$

then you obtain $I_\omega(x^* + \rho, y^*) = I_\omega^{(h)}(\rho)$, where

$$I_\omega^{(h)}(\rho) = \int_{-T/2}^{T/2} h(t - \rho/\bar{u})e^{-j\omega t}dt = \begin{cases} -(e^{-j\omega\rho/\bar{u}} - j)/\omega, & \text{if } \rho > 0, \\ -(e^{-j\omega\rho/\bar{u}} + j)/\omega, & \text{otherwise.} \end{cases} \quad (28)$$

In this case, $f(x^*, y^*, -T/2) \neq f(x^*, y^*, T/2)$ and the imaginary part of I_ω has a non-zero constant term. As a result, the equation (25) does not hold and you cannot estimate the velocity accurately by using the phase delay because $\bar{u} \neq \omega/(\partial_\rho \phi_\omega)$. From an application's point of view, a step function well represents the temporal changes of the light intensities measured near the boundaries of rapidly moving objects, and we need a method that can accurately estimate $\bar{u}$ even if the temporal change of the light intensity has a form like a step function.

Let the function of the temporal change of the light intensity be decomposed as follows:

$$f(x^*, y^*, t) = A \times b(t) + B \times h(t) + \eta(t), \quad (29)$$

where $h(t)$ is a step function described above, and $b(t)$ is any bounded function that has non-zero values only in a limited region as follows:

$$b(t) = \begin{cases} b_{\text{in}}(t), & \text{if } -\epsilon < t < \epsilon \\ 0, & \text{otherwise.} \end{cases} \quad (30)$$

Here, $0 < \epsilon < T/2$ and $b_{\text{in}}(t)$ can have non-zero values. $b(t)$ represents the temporal change of the light intensity that would be measured when a small or thin object rapidly passes in front of the sensor, and $h(t)$ would be measured at around the boundaries of rapidly moving large objects. The last term of the right hand side of (29), $\eta(t)$, denotes a component of $f(x^*, y^*, t)$ other than $b(t)$ and $h(t)$.

In the proposed method, we assume that the values, $I_\omega(x, y)$, measured by a correlation sensor are mainly determined by $b(t)$ and $h(t)$: The component of $\eta(t)$ corresponding to the frequency of ω has a power smaller than those of $b(t)$ and of $h(t)$, and is ignorable. This assumption does not hold in general but is satisfied by variety of signals including impulse signals, Gaussian pulses, and blurred step functions, which cover main signals measured when high speed objects are observed. Let $I_\omega^{(b)}(\rho)$ denote the Fourier coefficient of $b(t - \rho/\bar{u})$. Computing the coefficient, you obtain

$$I_\omega^{(b)}(\rho) = A_b e^{-j(\phi_\omega^{(b0)} + \rho/\bar{u})}, \quad (31)$$

where

$$\int_{-T/2}^{T/2} b(t)e^{-j\omega t}dt = A_b e^{-j\phi_\omega^{(b0)}}, \quad (32)$$

and A_b is a real coefficient. Under the condition that $\eta(t)$ is ignorable, you obtain

$$I_\omega(x^* + \rho, y^*) = \alpha e^{-j\omega\rho/\bar{u}} + \beta \quad (33)$$

by inserting (28) and (31) to (29), where $\alpha = A \cdot A_b \cdot e^{-j\phi_\omega^{(b0)}} - B/\omega$ and $\beta = -jB/\omega$. Here, it should be noted that the phase of $I_\omega(x^* + \rho, y^*)$ is

not proportional to ρ because of β. For eliminating β, we compute the spatial differentiation of $I_\omega(x^* + \rho, y^*)$:

$$\frac{\partial I_\omega(x^* + \rho, y^*)}{\partial \rho} = -\frac{j\omega}{\bar{u}}\alpha e^{-j\omega\rho/\bar{u}}. \tag{34}$$

Now, the phase of $\partial_\rho I_\omega$ is proportional to ρ as shown in (34). Let the phase of $\partial_\rho I_\omega$ be denoted by ψ_ω, where $\psi_\omega = \mathrm{Im}[\partial_\rho I_\omega]/\mathrm{Re}[\partial_\rho I_\omega]$. Then, you can compute the speed, $\bar{u}$ by using the following equation:

$$\partial_\rho \psi_\omega = -\omega/\bar{u}. \tag{35}$$

As described above, the normal speed, $\bar{u}$, can be obtained by computing the directional derivative of I_ω and of ψ_ω with respect to ρ, and the direction of ρ is parallel to the spatial gradient of the phase of $I_\omega(x, y)$. Let the phase of $I_\omega(x, y)$ be denoted by $\phi_\omega(x, y) = \tan^{-1}(\mathrm{Im}[I_\omega]/\mathrm{Re}[I_\omega])$ and the angle between the direction of the spatial gradient $\nabla\phi_\omega$ and the x-axis be denoted by θ. Then, the direction of ρ is parallel to $[\cos\theta, \sin\theta]^T$. The proposed method computes $\boldsymbol{v}_{\mathrm{norm}}$ as followings:

1. Compute the direction of $\nabla\phi_\omega = [\partial_x\phi_\omega, \partial_y\phi_\omega]^T$ as follows:

$$\cos\theta = \frac{\partial_x\phi_\omega}{\sqrt{(\partial_x\phi_\omega)^2 + (\partial_y\phi_\omega)^2}}, \text{ and } \sin\theta = \frac{\partial_y\phi_\omega}{\sqrt{(\partial_x\phi_\omega)^2 + (\partial_y\phi_\omega)^2}} \tag{36}$$

2. Compute the directional derivative of $I_\omega(x, y)$ with respect to ρ as

$$\partial_\rho I_\omega = (\partial_x I_\omega)\cos\theta + (\partial_y I_\omega)\sin\theta, \tag{37}$$

where $\cos\theta$ and $\sin\theta$ are given in (36). Let $\psi_\omega = \tan^{-1}\mathrm{Im}[\partial_\rho I_\omega]/\mathrm{Re}[\partial_\rho I_\omega]$.
3. Compute

$$\partial_\rho \psi_\omega = (\partial_x\psi_\omega)\cos\theta + (\partial_y\psi_\omega)\sin\theta. \tag{38}$$

The resultant two-vector, $\boldsymbol{v}_{\mathrm{norm}}$ is obtained as follows:

$$\boldsymbol{v}_{\mathrm{norm}} = -\frac{\omega}{\partial_\rho\psi_\omega}\begin{bmatrix}\cos\theta \\ \sin\theta\end{bmatrix}. \tag{39}$$

3   Experiments

The performance of the proposed method was evaluated by using artificial images and by using real ones captured by a correlation sensor.

3.1   Simulation Experiments

The accuracy of the proposed estimation method was experimentally evaluated using artificial images. Setting the shutter speed $T = 1/30$ seconds and $n = 1$ in (14), we simulated image capturing of the correlation sensor. this paper, we
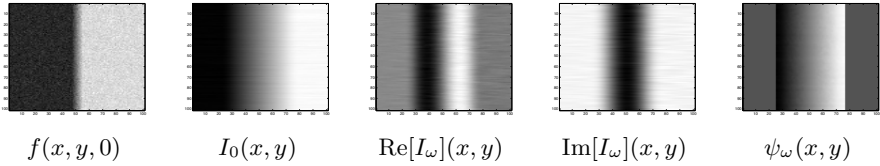
$$f(x, y, 0) \qquad I_0(x, y) \qquad \mathrm{Re}[I_\omega](x, y) \qquad \mathrm{Im}[I_\omega](x, y) \qquad \psi_\omega(x, y)$$

Fig. 4. Artificial images generated in the simulation experiments
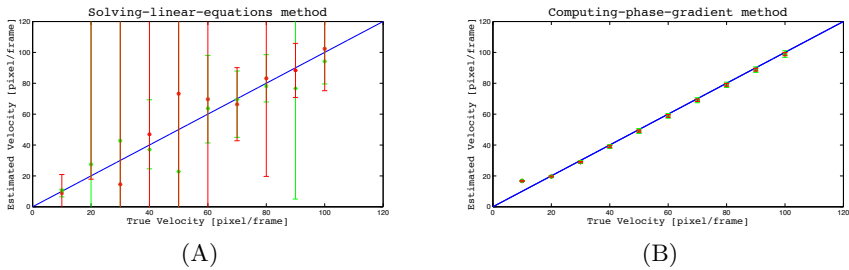

(A)   (B)

Fig. 5. Distributions of the estimates obtained by the Wei's method(A) and by the proposed method(B). In each graph, the horizontal axis shows the true normal velocity and the vertical one shows the estimated values. Mean values and one-sigma error bars are indicated in the graphs. The Wei's method cannot accurately estimate the flow when $\bar{u}T$ is larger than $\sigma = 10$.

report the results obtained when we set the light intensity, $f(x, y, 0)$, as a blurred step function contaminated by a Gaussian noise as follows:

$$f(x, y, 0) = (B \cdot h(t)) * G(\sigma^2) + \xi, \tag{40}$$

where $G(\sigma^2)$ is a Gaussian blurring filter of which scale is σ^2 and ξ is a Gaussian noise of which variance is s_f^2. Examples of the simulated images, $f(x, y, 0)$, $I_0(x, y)$, $\mathrm{Re}[I_\omega](x, y)$, $\mathrm{Im}[I_\omega](x, y)$, and $\phi_\omega(x, y)$, which were obtained when $\boldsymbol{v}_{\mathrm{norm}} = [\bar{u}, 0]^T = [50, 0]^T$ pixel/sec and when $\sigma = 10$ pixel, are shown in Fig.4.

For a comparison purpose, two flow estimation methods were applied to the identical set of simulated images: One was the proposed method and the other was the Wei's method[11] that solves the system of the linear-equations (15). The graphs shown in Fig.5(A) and (B) show the distributions of the estimated velocities. In the graphs, the horizontal-axis indicates the true velocity, $\bar{u}$, the vertical-axis indicates the estimated values, and the blue line shows the ideal estimates. The red colored results were obtained when the noise level, s_f, was 5% of the step size of B in (40) and the green ones were obtained when the noise level was 10% of B. In case the $f(x, y, 0)$ is given as shown in (40), the Wei's method works well only when the moving distance of $f(x, y, 0)$, $\bar{u}T$, and the spatial blurring scale, σ, is comparable: If $\bar{u}$ is so large that $\bar{u}T \gg \sigma$, then the spatial gradients of I_0 supply little information about the target motion and the solution of the system of the linear equations (15) is unreliable. As shown in

Fig. 6. A camera with the correlation image sensor (left) and the experimental setting (right)

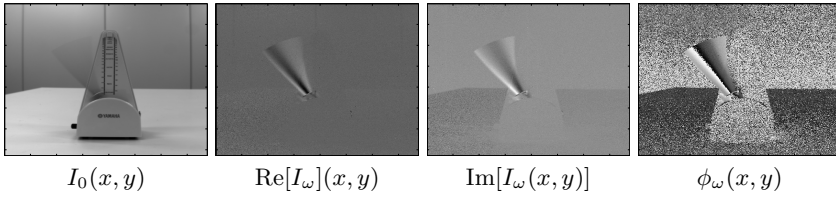| $I_0(x,y)$ | $\mathrm{Re}[I_\omega](x,y)$ | $\mathrm{Im}[I_\omega(x,y)]$ | $\phi_\omega(x,y)$ |

Fig. 7. Examples of images captured by the correlation image sensor

the graph (A), the Wei's method[11] estimates the flows accurately only when the true velocity is equal or less than $\sigma = 10$. On the other hand, the proposed method estimates the flows accurately when $\bar{u}$ is enough large, as shown in the graph (B).

3.2 Experiments Using a Real Correlation Image Sensor

Capturing real image sequences of a metronome by the correlation image sensor (Fig.6), we evaluated the performance of our proposed method. Figure 7 shows examples of the images captured by the sensor. Using these images, the flows were computed by our proposed method and by the Wei's method[11], and the results were compared with gold standards manually made from $I_0(x,y)$. An example of the flow estimated by each of the methods is shown in Fig.8. The color shows the direction of the flow. The flow obtained by the Wei's method is contaminated by the textures behind the swinging pendulum. On the other hand, the proposed method accurately estimated the flow. It should be noted that these flows are computed in real time for each single frame. The results of the quantitative evaluation are shown in Fig.9. As shown in the graph (A), the Wei's method underestimated the flows of the pendulum when it moved fast, but the proposed one accurately estimated the high-speed flows, as shown in the graph (B).
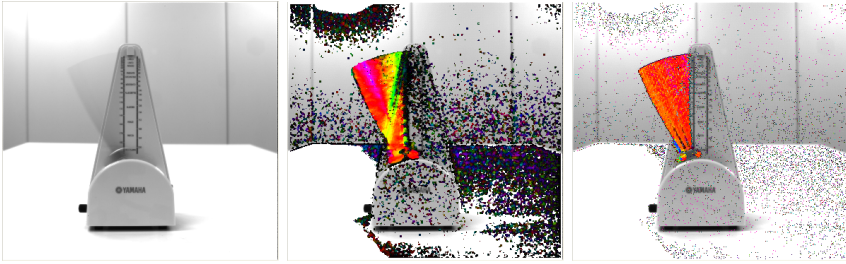
Fig. 8. Examples of estimated optical flow. Left: $I_0(x, y)$. Middle: Wei's method[11]. Right: Proposed method.
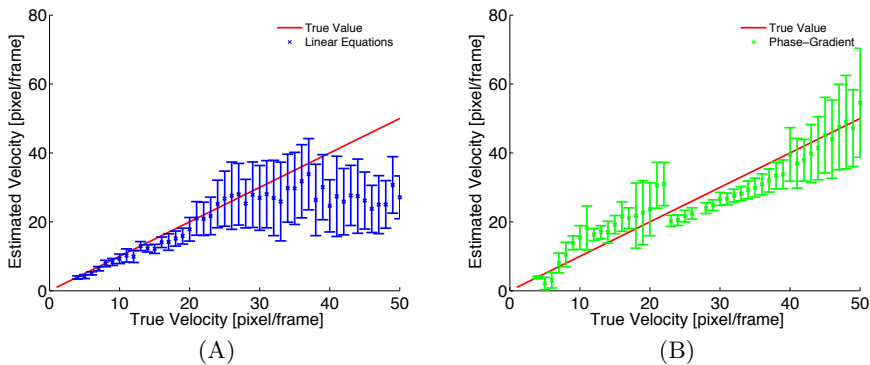

Fig. 9. Distributions of the velocities estimated by the Wei's method (A) and by the proposed method (B)

4   Conclusion

In this article, we proposed a method that can estimate the normal component of high velocity optical flow using the local measurements obtained by the correlation image sensor. When given images are captured by the correlation image sensor, you can make the optical flow estimation well-posed if the corresponding system of the linear equations is not degenerated. In our proposed method, the normal component of the flow is estimated by using the spatial gradient of the phase, $\psi_\omega(x, y)$ when the system of the equations is degenerated. The equation used for estimating the normal velocity was theoretically derived in this paper. Our future works include to develop a vision system that can compute a wide rage of optical flows in real time by combining the Wei's method and our proposed one.

References

1. Horn, B.K., Schunck, B.G.: Determining optical flow. Artificial Intelligence 17(1), 185–203 (1981)
2. Bertero, M., Poggio, T.A., Torre, V.: Ill-posed problems in early vision. Proceedings of the IEEE 76(8), 869–889 (1988)
3. Sun, D., Roth, S., Black, M.J.: Secrets of optical flow estimation and their principles. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2432–2439. IEEE (2010)
4. Bigun, J., Granlund, G.H.: Optical flow based on the inertia matrix of the frequency domain. In: Proceedings from SSAB Symposium on Picture Processing, pp. 132–135. Lund University, Sweden (1988)
5. Lucas, B.D., Kanade, T., et al.: An iterative image registration technique with an application to stereo vision. In: IJCAI, vol. 81, pp. 674–679 (1981)
6. Ayvaci, A., Raptis, M., Soatto, S.: Sparse occlusion detection with optical flow. International Journal of Computer Vision 97(3), 322–338 (2012)
7. Garg, R., Roussos, A., Agapito, L.: Robust trajectory-space tv-l1 optical flow for non-rigid sequences. In: Boykov, Y., Kahl, F., Lempitsky, V., Schmidt, F.R. (eds.) EMMCVPR 2011. LNCS, vol. 6819, pp. 300–314. Springer, Heidelberg (2011)
8. Steinbrucker, F., Pock, T., Cremers, D.: Large displacement optical flow computation withoutwarping. In: 2009 IEEE 12th International Conference on Computer Vision, pp. 1609–1614. IEEE (2009)
9. Cosker, D., Li, W., Brown, M., Tang, R.: Optical flow estimation using laplacian mesh energy. In: IEEE International Conference on Computer Vision and Pattern Recognition (CVPR). University of Bath (2013)
10. Trobin, W., Pock, T., Cremers, D., Bischof, H.: Continuous energy minimization via repeated binary fusion. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 677–690. Springer, Heidelberg (2008)
11. Wei, D., Masurel, P., Kurihara, T., Ando, S.: Optical flow determination with complex-sinusoidally modulated imaging. In: 2006 8th International Conference on Signal Processing, vol. 2. IEEE (2009)
12. Brox, T., Malik, J.: Large displacement optical flow: descriptor matching in variational motion estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence (99), 1 (2011)