International Journal of Computer Vision 63(3), 207–224, 2005

Sub-Pixel Estimation Error Cancellation on Area-Based Matching

MASAO SHIMIZU AND MASATOSHI OKUTOMI
Graduate School of Science and Engineering, Tokyo Institute of Technology, 2-12-1 Ookayama, Meguro-ku, Tokyo, 152-8550, Japan
mas@ok.ctrl.titech.ac.jp
mxo@ctrl.titech.ac.jp

Received July 29, 2004; Revised November 23, 2004; Accepted November 23, 2004

First online version published in April 2005

Abstract. Area-based image matching and sub-pixel displacement estimation using similarity measures are common methods that are used in various fields. Sub-pixel estimation using parabola fitting over three points with their similarity measures is also a common method to increase the matching resolution. However, few investigations or studies have explored the characteristics of this estimation.

This study[1] analyzed sub-pixel estimation error using two different types of matching model. Our analysis demonstrates that the estimation contains a systematic error depending on image characteristics, the similarity function, and the fitting function. This error causes some inherently problematic phenomena such as the so-called *pixel-locking* effect, by which the estimated positions tend to be biased toward integer values. We also show that there are *good* combinations of the similarity functions and fitting functions.

In addition, we propose a new algorithm to greatly reduce sub-pixel estimation error. This method is independent of the similarity measure and the fitting function. Moreover, it is quite simple to implement. The advantage of our novel method is confirmed through experiments using different types of images.

Keywords: area-based image matching, similarity measure, sub-pixel estimation, pixel-locking effect, estimation error reduction

1. Introduction

Area-based image matching is a common and basic method that is used in many fields such as motion estimation (Aggarwal and Nandhakumar, 1988), object tracking (Shortis et al., 1994), machine vision (Aghajan et al., 1993), video data coding, stereo image processing (Kanade and Okutomi, 1994), image-based measurement (West and Clarke, 1990), image registration (Irani and Peleg, 1991), remote sensing, and fluidics (Raffel et al., 1998). For estimating sub-pixel displacement, many applications have employed some methods to find a true peak of a similarity function by fitting a parabola or other functions over three indices near its extremum (Frischholz and Spinnler, 1993; Fusiello and Roberto, 2000; Raffel et al., 1998; Tian and Huhns, 1986). This method is called as the similarity interpolation method.

Nevertheless, in many cases, the only criterion considered is the computing speed without any consideration of the combination of similarity and fitting function, especially in industrial applications. It is also reported that the similarity interpolation method requires few computations in spite of the existence of systematic errors (Westerweel, 1998). Sub-pixel estimation error has been discussed on its magnitude and characteristics for a specific image under given conditions (Dvornychenko, 1983). Notwithstanding, they have not provided a method to reduce estimation errors.

Figure 1 shows a simple experiment. The target moves linearly in a horizontal direction at a constant speed. The measured horizontal positions are expected

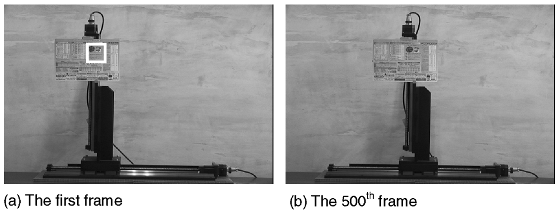

Figure 1. Images used for the target tracking. White rectangle indicates the tracking area.

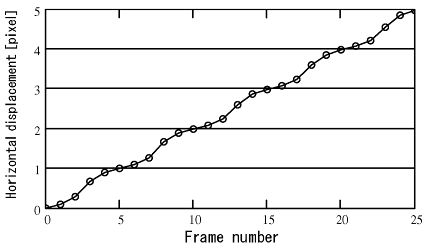

Figure 2. Measured target positions.

Figure 3. Images used for disparity estimation.

to increase linearly to the frame number (time). But the actual measured positions shown in Fig. 2 contain apparent systematic errors with a one-pixel period.

In the field of fluidics, the so-called *pixel-locking* effect has been reported (Westerweel, 1998). This effect is the phenomenon by which the estimated sub-pixel displacement inclines toward integer pixel positions. Its degree depends on images or similarity function. The only way to detect this effect is to make a histogram from many corresponding flow vectors.

Figure 4 represents a histogram of the horizontal component of displacement associated with 5040

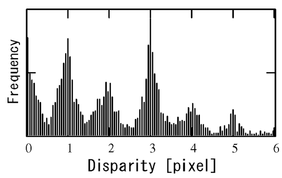

Figure 4. Histogram of estimated disparities.

points of flow-vectors estimated from real images shown in Fig. 3. The frequencies at integer pixel values are unnaturally high. This phenomenon indicates the possibility of systematic errors in all measurements using area-based matching.

This paper presents a novel estimation error cancel (EEC) method which can greatly reduce the estimation error. The method uses no iteration, thereby easy implementation of both software and hardware. It is independent of the similarity and fitting function, and the images used.

The EEC method assumes some properties in sub-pixel estimation error. Two different types of image matching models are examined to verify the assumptions. One is a one-pixel-matching model, which is used for analysis of the characteristics of sub-pixel estimation error. This model is minimal and it requires no specific image.

The other one is an edge image model, which is more practical for modeling and evaluation of the estimation error. The effect of EEC method will be shown after description of the relations between the two models.

This paper is organized as follows. The next section describes two types of similarity measure functions and two types of fitting functions. Section 3 describes a one-pixel-matching model for generalized sub-pixel estimation properties. Section 4 presents an edge image model after considering the optical imaging model; then relations between the two models are described. Section 5 describes the EEC method in detail. Additionally, the residual error is discussed. Experimental results are provided in Section 6 using four different types of images.

1.1. Related Works

The following methods have been generally used to estimate the sub-pixel displacement in images (Aggarwal and Nandhakumar, 1988; Tian and Huhns, 1986).

(1) *Similarity interpolation* estimates the true peak of a similarity function by fitting a parabola to the three indices near its extremum. This method requires few computations despite the existence of systematic errors (Westerweel, 1998).
(2) *Intensity interpolation* seeks a peak position of the similarity in high-spatial-resolution images obtained by image interpolation (Szeliski and Scharstein, 2002). This method generally requires much memory space and computation time. Image

interpolation methods with continuous functions or hierarchical coarse-to-fine approaches have been proposed to reduce these requirements (Fincham and Delerce, 1999; Hart, 1998; Lecordier et al., 1999). The equivalence of this method and gradient based method is shown in Davis et al. (1995).

(3) *Gradient-based method* computes the sub-pixel modification using image gradients. The method can estimate displacement of less than one pixel with iterations (Horn and Schunck, 1981; Lucas and Kanade, 1981). A study of convergence of the iterative method is shown in Mitiche and Mansouri (2004). An image gradient estimating filter is also affecting (Davis and Freeman, 1998). This method assumes that the displacement is very small. Therefore, it is generally used with coarse-to-fine search strategy (Tian and Huhns, 1986).

(4) *Phase correlation* detects phase differences in the Fourier domain as a shift in spatial domain (Tian and Huhns, 1986). To obtain sub-pixel phase differences, an interpolation in the Fourier domain is needed. A direct phase differences estimation using correlation values has been proposed (Foroosh et al., 2002), but the area of interest (AOI) is limited to rectangular shapes.

(5) *Geometric method* extracts feature points and find correspondence among them. The accuracy of this method depends on the feature extraction performance.

Herein, we examin method (1) because of its simplicity and effectiveness. The systematic error in this method has already been reported, but the only way proposed so far for reducing the error is to use outfocused images (Schreier et al., 2000). A more effective estimation error reduction method would widely affect computer vision and industrial applications, especially those system which require precise measurements.

2. Dissimilarity Measures and Sub-Pixel Estimations

Matching of two images finds the minimum or maximum value of a dissimilarity or similarity function corresponding to the displacement between the two images. Either the sum of absolute differences (SAD), sum of squared differences (SSD), cross correlation (CC), or zero-mean normalized cross-correlation is usually used as the measure for the matching region.

Let a given one dimensional displacement be d which can be considered to lie in the range $-0.5 \leq d \leq 0.5$ without loss of generality. It can be considered that the image matching in pixel unit has alreay been done. The similarity function $R(s)$ using SAD is can be expressed as

$$R_{\mathrm{SAD}}(s) = \sum_{i \in W} |f(i) - f(i-d+s)|, \qquad (1)$$

where $f(i)$ and $f(i-d+s)$ is the reference and observed image, respectively, and s is the shift value in pixel unit from the extremum position. $R(s)$ using SSD is represented as

$$R_{\mathrm{SSD}}(s) = \sum_{i \in W} (f(i) - f(i-d+s))^2. \qquad (2)$$

Further, $R(s)$ using CC is represented as

$$R_{\mathrm{CC}}(s) = \sum_{i \in W} (f(i) \times f(i-d+s)). \qquad (3)$$

We investigate using SAD and SSD becase the result using CC is equal to the case of SSD in our image model.

The sub-pixel displacement position can be estimated using the similarity values. Some symmetric functions are generally used for fitting over three similarity values to find the sub-pixel displacement. Figure 5(a) shows the first order symmetric fitting function (referred to as equiangular line fitting in this paper). The equiangular line fitting estimates the sub-pixel position as the intersection point of two lines. One line passes $(-1, R(-1))$ and $(0, R(0))$, while the other line passes $(1, R(1))$ with the inverse-sign gradient. The

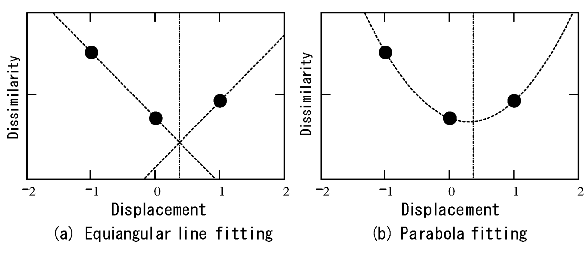

Figure 5. Examples of fitting functions.

sub-pixel estimation $\hat{d}_{\text{EL}}$ is

$$\hat{d}_{\text{EL}} = \begin{cases} \dfrac{1}{2}\dfrac{R(1)-R(-1)}{R(0)-R(-1)} & \text{if } R(1) < R(-1), \\ \dfrac{1}{2}\dfrac{R(1)-R(-1)}{R(0)-R(1)} & \text{otherwise,} \end{cases} \tag{4}$$

where $R(-1)$, $R(0)$, and $R(1)$ are the similarity values obtained from Eq. (1) or (2).

Figure 5(b) shows parabola fitting. Parabola fitting estimates the sub-pixel position as the centerline of the fitted parabola. The sub-pixel estimation $\hat{d}_{\text{PB}}$ is

$$\hat{d}_{\text{PB}} = \frac{R(-1)-R(1)}{2R(-1)-4R(0)+2R(1)}. \tag{5}$$

3. One-Pixel-Matching Model

To examine the detailed properties of the sub-pixel estimation, the summation of the dissimilarity blocks us from finding the affection of the image property to the sub-pixel estimation error. Consequently, we employ a minimum set of pixels that can be used for matching and sub-pixel estimation.

3.1. *Minimum Set of Pixels*

Only one observed pixel can be used for the purpose. To estimate the sub-pixel displacement, three similarity values corresponding to the three positions are needed. Therefore three pixels are required as references. It follows that the problem is to find the best matching sub-pixel position for the one observed pixel against the three reference pixels.

The one observed pixel is considered as the best matched to the center of the reference pixels based on considering the similarity function in pixel unit. It allows that the displacement of the observed pixel against the center of the reference pixels is limited in -0.5 to $+0.5$ pixel. The analysis model is one dimensional because of its simplicity. Figure 6 illustrates the one-pixel-matching model (referred to as *opm* hereafter).

The characteristics of sub-pixel estimation error can be considered as a function of the intensity property of the model. Then we introduce a parameter and show its existing range in the following subsections.

The similarity values between the observed pixel and the reference pixels can be expressed using the model.

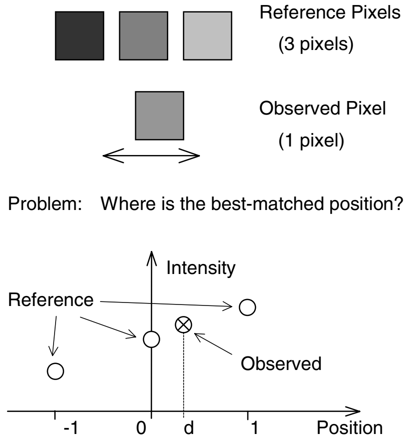

Figure 6. One-pixel-matching model.

The similarity values using the absolute difference are

$$\begin{aligned} R_{\text{AD}}(-1) &= |I(-1) - I(d)|, \\ R_{\text{AD}}(0) &= |I(0) - I(d)|, \\ R_{\text{AD}}(1) &= |I(1) - I(d)|, \end{aligned} \tag{6}$$

where $I(d)$ is the observed pixel intensity, and $I(-1)$, $I(0)$, and $I(1)$ are the intensity of reference pixels.

The similarity values using the squared difference[2] are

$$\begin{aligned} R_{\text{SD}}(-1) &= (I(-1) - I(d))^2, \\ R_{\text{SD}}(0) &= (I(0) - I(d))^2, \\ R_{\text{SD}}(1) &= (I(1) - I(d))^2. \end{aligned} \tag{7}$$

3.2. *Pixel Intensity Nonlinearity*

Sub-pixel estimation can be carried out using Eq. (4) or (5) with the similarity values of Eq. (7) or (8). The estimated value can be expected to vary with the intensity relation between three reference pixels. The purpose of this subsection is to introduce a parameter, which expresses an intensity layout of the reference pixels. This parameter can be used to investigate sub-pixel estimation errors in the following subsections.

We introduce the following measure E as a nonlinearity:

$$E = \frac{1}{\sqrt{6}}(2I(0) - I(-1) - I(1)). \tag{8}$$

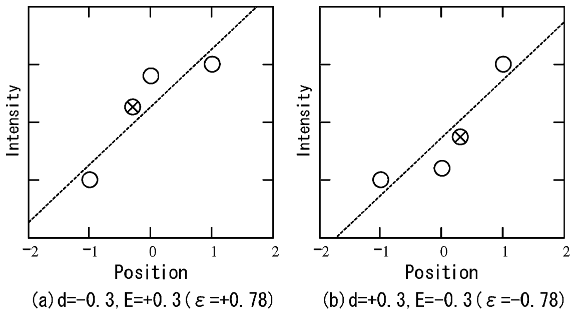

Figure 7. The analytical model. ○ and ⊗ denote intensity of the three reference pixels and observed one pixel, respectively. The dashed lines denote the least-squares estimated lines for the reference pixels.

The measure E describes the *signed* RMS error of three pixels from the least-squares estimated line for them, where $E > 0$ means that $I(0)$ is above the line, and $E < 0$ means that $I(0)$ is under the line (see Fig. 7). $E = 0$ when the reference three pixels are all on the line in terms of their intensities. $I(0)$ can be expressed from Eq. (8) as,

$$I(0) = \frac{1}{2}(I(-1) + I(1) + \sqrt{6}E). \tag{9}$$

Next, we derive the observed one pixel intensity $I(d)$ as a function of E and a given displacement d. To evaluate the sub-pixel estimation error, it is necessary to know the true input displacement of the one observed pixel against the three reference pixels. We assume that the observed pixel intensity $I(d)$ is a linearly interpolated value with three reference pixels,

$$I(d) = \begin{cases} I(0) + (I(0) - I(-1))d & \text{if } d < 0, \\ I(0) + (I(1) - I(0))d & \text{otherwise.} \end{cases} \tag{10}$$

The similarity values of Eq. (6) or (7) can be expressed with the nonlinearity E, a given displacement d, and pixel intensities $I(-1)$ and $I(1)$ using Eqs. (9) and (10).

Lastly, we should know the range of the nonlinearity E. As described before, the observed pixel intensity $I(d)$ should be best matched to the center of the reference pixels $I(0)$. For this condition, the following relations are considered as the range of E.

$$|I(1) - I(d)| > |I(0) - I(d)|,$$

and

$$|I(-1) - I(d)| > |I(0) - I(d)|. \tag{11}$$

Using Eqs. (9) and (10), the following condition can be obtained from Eq. (11) for $-0.5 \le d \le 0.5$, as

$$-\frac{I(1) - I(-1)}{\sqrt{6}} < E < \frac{I(1) - I(-1)}{\sqrt{6}}. \tag{12}$$

3.3. Sub-Pixel Estimation

The center pixel intensity $I(0)$ of the reference pixels is determined by $I(-1)$, $I(1)$, and E using Eq. (9). At the same time, the observed pixel intensity $I(d)$ is determined by the actual displacement d using Eq. (10). Substituting these equations into Eq. (6) or (7) leads to the similarity values of AD or SD, respectively. These similarity values provide sub-pixel estimation using Eq. (4) or (5) for the equiangular line (EL) or parabola (PB) fitting, respectively. These sub-pixel estimations are functions of E and d.

EL using AD:

$$\hat{d}_{\mathrm{AD+EL}}(d, \varepsilon) = \begin{cases} \dfrac{d + (1 + d)\varepsilon}{(1 + 2d)(1 + \varepsilon)} & \left(d < 0, \varepsilon \ge \dfrac{-d}{1 + d}\right), \\ \dfrac{d + (1 - d)\varepsilon}{1 + \varepsilon} & \left(d \ge 0, \varepsilon \ge \dfrac{-d}{1 - d}\right), \\ \dfrac{d + (1 + d)\varepsilon}{1 - \varepsilon} & \left(d < 0, \varepsilon < \dfrac{-d}{1 + d}\right), \\ \dfrac{d + (1 - d)\varepsilon}{(1 - 2d)(1 - \varepsilon)} & \left(d \ge 0, \varepsilon < \dfrac{-d}{1 - d}\right), \end{cases} \tag{13}$$

where ε is the following substitution for simplicity:

$$\varepsilon = \frac{\sqrt{6}}{I(1) - I(-1)}E. \tag{14}$$

The range described in Eq. (12) is derived as

$$-1 < \varepsilon < 1. \tag{15}$$

PB using AD:

$$\hat{d}_{\mathrm{AD+PB}}(d, \varepsilon) = \frac{1}{2}\frac{d + (1 - |d|)\varepsilon}{(1 - |d|) + d\varepsilon}. \tag{16}$$

EL using SD:

$$\hat{d}_{\mathrm{SD+EL}}(d,\varepsilon) = \begin{cases} \dfrac{2(d+(1+d)\varepsilon)}{(1+2d)(1+\varepsilon)^2} & \left(d<0, \varepsilon \geq \dfrac{-d}{1+d}\right), \\ \dfrac{2(d+(1-d)\varepsilon)}{(1+2d)+2\varepsilon+(1-2d)\varepsilon^2} & \left(d\geq 0, \varepsilon \geq \dfrac{-d}{1-d}\right), \\ \dfrac{2(d+(1+d)\varepsilon)}{(1-2d)-2\varepsilon+(1+2d)\varepsilon^2} & \left(d<0, \varepsilon < \dfrac{-d}{1+d}\right), \\ \dfrac{2(d+(1-d)\varepsilon)}{(1-2d)(1-\varepsilon)^2} & \left(d\geq 0, \varepsilon < \dfrac{-d}{1-d}\right). \end{cases} \tag{17}$$

PB using SD:

$$\hat{d}_{\mathrm{SD+PB}}(d,\varepsilon) = \frac{d+(1-|d|)\varepsilon}{1+2d\varepsilon+(1-2|d|)\varepsilon^2}. \tag{18}$$

Figures 8(a) through 8(d) show results of Eqs. (13) through (18). The horizontal axes correspond to the actual displacement d; the vertical axes correspond to the sub-pixel estimation error $\hat{d}(d,\varepsilon)-d$. Curves in Fig. 8 correspond to $\varepsilon = 0.99$, 0.8, 0.6, 0.4, 0.2, 0, −0.2, −0.4, −0.6, −0.8, −0.99 from top to bottom, respectively.

From the sub-pixel estimation error shown in Fig. 8, the errors are zero when $\varepsilon = 0$ for EL using AD and PB using SD, but errors arise for other combinations. The condition of $\varepsilon = 0$ indicates that the intensities of three reference pixels (adjacent three pixels) changes linearly. All images contain low spatial frequency components rather than high frequency components, so this condition will occur with high probability. For this reason, these combinations are considered to be *good* ones.

3.4. Matching Window

Although, Fig. 8 reveals existence of *good* combinations of similarity function and fitting function, the errors in Fig. 8 will not explain the pixel-locking effect because the errors are always positive for any $\varepsilon > 0$, and negative for any $\varepsilon < 0$. This difference arises as a difference between the one-pixel-matching model and area-based matching. The area-based matching uses summation of the similarity measure for AOI. For example, the following equation is used in the case of the parabola fitting with SSD.

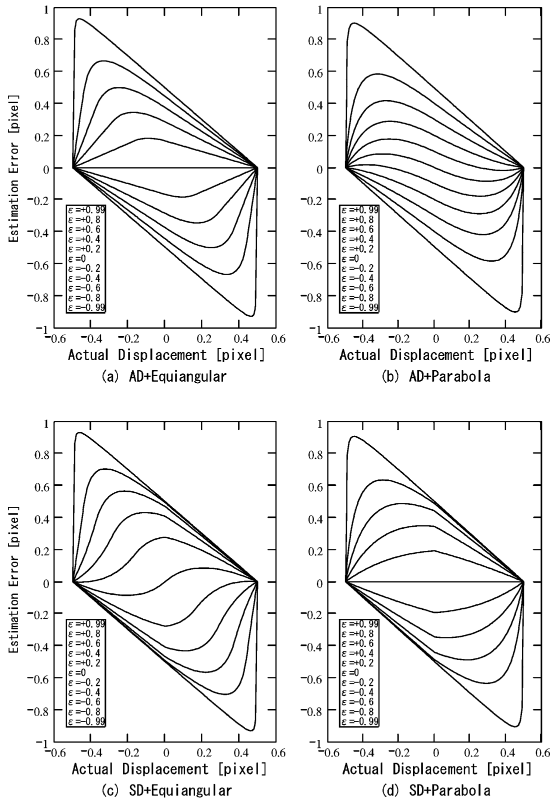

Figure 8. Sub-pixel estimation errors with the one-pixel-matching model.

$$\hat{d}_{\mathrm{SSD+PB}} = \frac{\sum_W R_{\mathrm{SD}}(-1) - \sum_W R_{\mathrm{SD}}(1)}{2\sum_W R_{\mathrm{SD}}(-1) - 4\sum_W R_{\mathrm{SD}}(0) + 2\sum_W R_{\mathrm{SD}}(1)}. \tag{19}$$

Equation (19) is regarded as *principal estimation* hereafter.

On the other hand, we can regard the weighted sum of Eq. (18) shown below as the sub-pixel estimation for AOI:

$$\bar{d}_{\mathrm{SD+PB}} = \frac{1}{n}\int_{-1}^{1} h(\varepsilon)\hat{d}_{\mathrm{SD+PB}}(d,\varepsilon) = \frac{1}{n}\int_{-1}^{1} h(\varepsilon)\frac{R_{\mathrm{SD}}(-1)-R_{\mathrm{SD}}(1)}{2R_{\mathrm{SD}}(-1)-4R_{\mathrm{SD}}(0)+2R_{\mathrm{SD}}(1)}, \tag{20}$$

where $h(\varepsilon)$ is the weighting function for ε, and n is the total pixel number of the region. Equation (20) is referred to as the *weighted average estimation* hereafter.

Next, the relation between the two estimations is investigated. Consider the region of $n-1$ pixels and the neighboring region of one pixel. We specifically consider the relationship between each estimation obtained from each region, and the estimation obtained from the merged n pixels region.

The *principal estimated* sub-pixel displacement from $n-1$ pixels region and one pixel region are

$$\begin{aligned}\hat{d}^{\langle 1,n-1\rangle}_{\mathrm{SSD+P}} &= \frac{\sum_{i=1}^{n-1} R^{\langle i\rangle}_{\mathrm{SD}}(-1) - \sum_{i=1}^{n-1} R^{\langle i\rangle}_{\mathrm{SD}}(1)}{2\sum_{i=1}^{n-1} R^{\langle i\rangle}_{\mathrm{SD}}(-1) - 4\sum_{i=1}^{n-1} R^{\langle i\rangle}_{\mathrm{SD}}(0) + 2\sum_{i=1}^{n-1} R^{\langle i\rangle}_{\mathrm{SD}}(1)} \\ &= \frac{\sum_{i=1}^{n-1} f_i}{\sum_{i=1}^{n-1} g_i},\end{aligned} \tag{21}$$

and

$$\begin{aligned}\hat{d}^{\langle n,n\rangle}_{\mathrm{SSD+P}} &= \frac{R^{\langle n\rangle}_{\mathrm{SD}}(-1) - R^{\langle n\rangle}_{\mathrm{SD}}(1)}{2R^{\langle n\rangle}_{\mathrm{SD}}(-1) - 4R^{\langle n\rangle}_{\mathrm{SD}}(0) + 2R^{\langle n\rangle}_{\mathrm{SD}}(1)} \\ &= \frac{f_n}{g_n},\end{aligned} \tag{22}$$

where $R^{\langle i\rangle}_{\mathrm{SD}}()$ denotes the similarity value obtained from the i-th pixel using SD. In addition, the following substitutions are used for simplicity:

$$\begin{aligned} f_i &= R^{\langle i\rangle}_{\mathrm{SD}}(-1) - R^{\langle i\rangle}_{\mathrm{SD}}(1), \\ g_i &= 2R^{\langle i\rangle}_{\mathrm{SD}}(-1) - 4R^{\langle i\rangle}_{\mathrm{SD}}(0) + 2R^{\langle i\rangle}_{\mathrm{SD}}(1). \end{aligned} \tag{23}$$

Additionally, using Eqs. (21) and (22), the *principal estimated* sub-pixel displacement from n pixels region that contains both $n-1$ pixels region and one pixel region is

$$\begin{aligned}\hat{d}^{\langle 1,n\rangle}_{\mathrm{SSD+P}} &= \frac{\sum_{i=1}^{n-1} f_i + f_n}{\sum_{i=1}^{n-1} g_i + g_n} \\ &= \frac{\hat{d}^{\langle 1,n-1\rangle}_{\mathrm{SSD+P}} \sum_{i=1}^{n-1} g_i + \hat{d}^{\langle n,n\rangle}_{\mathrm{SSD+P}} g_n}{\sum_{i=1}^{n} g_i}.\end{aligned} \tag{24}$$

On the other hand, the *weighted average estimated* sub-pixel displacement $\bar{d}^{\langle 1,n\rangle}_{SD+P}$ obtained from $\hat{d}^{\langle 1,n-1\rangle}_{\mathrm{SSD+P}}$ and $\hat{d}^{\langle n,n\rangle}_{\mathrm{SSD+P}}$ using region size weight is

$$\bar{d}^{\langle 1,n\rangle}_{\mathrm{SD}+P} = \frac{n-1}{n}\hat{d}^{\langle 1,n-1\rangle}_{\mathrm{SSD+P}} + \frac{1}{n}\hat{d}^{\langle n,n\rangle}_{\mathrm{SSD+P}}. \tag{25}$$

Therefore, the following δ can be found:

$$\begin{aligned}\hat{d}^{\langle 1,n-1\rangle}_{\mathrm{SSD+P}} &= \bar{d}^{\langle 1,n\rangle}_{SD+P} - \frac{1}{n}\delta, \\ \hat{d}^{\langle n,n\rangle}_{\mathrm{SSD+P}} &= \bar{d}^{\langle 1,n\rangle}_{SD+P} + \frac{n-1}{n}\delta.\end{aligned} \tag{26}$$

Substituting Eq. (26) into Eq. (24) yields

$$\hat{d}^{\langle 1,n\rangle}_{\mathrm{SSD+P}} = \bar{d}^{\langle 1,n\rangle}_{SD+P} - \frac{\sum_{i=1}^{n-1} g_i - (n-1)g_n}{\sum_{i=1}^{n} g_i}\frac{\delta}{n}. \tag{27}$$

The compensation term in Eq. (27) describes the difference between the two estimations. Because $\sum_{i=1}^{n-1} g_i - (n-1)g_n \ll \sum_{i=1}^{n} g_i$, and $\delta \ll n$, it can be considered as ≈ 0. This fact means that the *weighted average estimation* described in Eq. (20) can be used as the estimation for the region instead of the *principal estimation* approximately. While the explanation uses $n-1$ pixels region and one pixel region, it can be extended easily to any size region by changing the region size from $n = 2$ one by one. Although parabola fitting is used in the above discussion, it is essentially the same in the case of the equiangular line fitting.

3.5. Characteristics with Taking Account of Matching Window

The weighting function $h(\varepsilon)$ is assumed as a Gaussian-like distribution with mean zero as

$$h(\varepsilon) = \left(\cos\left(\frac{\pi}{2}\varepsilon\right)\right)^{n_{\mathrm{opm}}}, \tag{28}$$

where n_{opm} is a distribution parameter.

Figure 9 depicts the errors of the *weighted average estimation* of Eq. (20) with three different n_{opm}.

Figure 10 illustrates the histograms corresponding to errors of $n_{\mathrm{opm}} = 20$ (green line) in Fig. 9. The histograms were obtained numerically by the true input displacement from 0 to 6[pixel] step 1/1000[pixel]. They will be flat if sub-pixel estimation contains no error.

Figure 10(b) (PB using SAD) indicates high frequency at integer positions. It is called as pixel-locking effect. On the other hand, Fig. 10(c) (EL using SSD) indicates that high frequencies avoid integer positions. We call this *anti-pixel-locking* in this paper.

The following results derived from the study in this section.

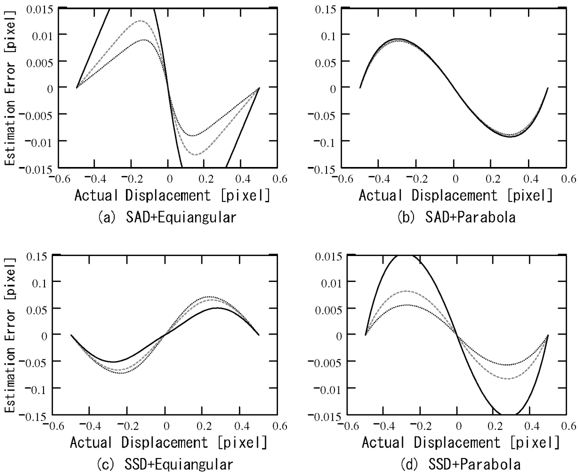

Figure 9. Sub-pixel estimation error obtained from the one-pixel-matching model with the matching window. $n_{\text{opm}} = 10$ (black lines), 20 (green lines), 30 (blue lines). The window region contains more high frequency components as n_{opm} decreases.

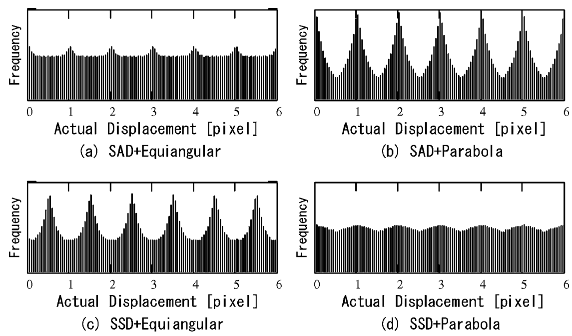

Figure 10. Histograms corresponding to the estimation errors depicted in Fig. 9. $n_{\text{opm}} = 20$ (green line) is used.

- Parabola fitting posits that the similarity or dissimilarity values follow a quadratic for displacement. Such (dis)similarities are SSD, CC, and ZNCC. This expectation is true for the image region with linearly changing intensity. On the other hand, equiangular line fitting posits that the values are on equiangular two lines. Such dissimilarity is SAD.
- These *good* combinations (EL using AD and PB using SD) show a similar simple tendency. Their estimation error decreases as the high spatial components of the image decrease (n_{opm} increases). The error magnitude is relatively small.
- Properties of sub-pixel estimation error are derived from the one-pixel-matching model and a simple $h(\varepsilon)$ distribution with no specific images. Therefore, we can use these results as the general matching characteristics for making our proposed EEC method.
- The pixel-locking effect can be well explained in the *weighted average estimation*. Figures 9(a), (b) and (d) show an obvious characteristic such that estimated values $\hat{d}$ are greater than true values when $d < 0$ and less than true values when $d > 0$. This fact causes the histogram of the estimated values $\hat{d}$ to have peaks at integer pixel locations as shown in Figs. 10(a), (b) and (d). This phenomenon is called pixel-locking. The degree of this effect differs with n_{opm} and the employed similarity measure and fitting function. It is noteworthy that estimation errors are rather large over any n_{opm} when SAD is used with parabola fitting.

On the other hand, the errors of EL using SSD shown in Fig. 9(c) have inverse-sign against the other combinations. The anti-pixel-locking effect is expected with this combination.

4. Edge Image Model

The image contains 2-D information, but each dimension can be treated independently in many cases. In this section, investigation is carried out in 1-D for simplicity, but it can be extended easily to 2-D. Additionally, the real 2-D images are used in the experiments mentioned later.

The unit of the position and the spatial frequency is set to the typical CCD imager pixel interval $p = 11.0 \times 10^{-6}$ [m]. This unit is indicated as a [pixel].

4.1. Imaging Characteristics

When using digital information of the image, knowledge of the transfer functions affected to the image data must be complete, but it is usually impossible to know all of them. Major factors affecting the imaging transfer function are an optical transfer function and a CCD aperture function. Approximating their combined function using a Gaussian function provides the edge image model.

4.1.1. Optical Transfer Function. As is well known, the point spread function of the optics can be represented as the following equation. If the lens has a circular aperture (Driscoll, 1978),

$$h(x) = \left[\frac{2J_1(kdx/2z')}{kdx/2z'}\right]^2, \tag{29}$$

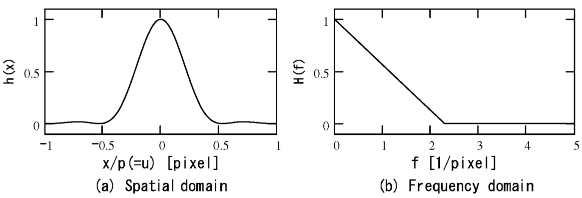

Figure 11. Examples of $h(x)$ and $H(f)$. ($\lambda = 600 \times 10^{-9}$ [m], $z'/d = 8$).

where x is the distance from the center, $J_1()$ is the Bessel function of the first kind, d is the effective diameter of the optics, z' is the focal length of the objective lens, and $k(= 2\pi/\lambda)$ is the wave number.

Its Fourier transform, which is usually referred to as the optical transfer function (OTF), is approximately

$$H(f) = \begin{cases} 1 - |f/f_c| & \text{if } |f/f_c| \le 1, \\ 0 & \text{otherwise,} \end{cases} \tag{30}$$
$$f_c = d/\lambda z',$$

where f is the spatial frequency and f_c is the cut-off frequency. The cut-off frequency depends on the wave length λ and the optical aperture z'/d. For example, $f_c = 6.88$ to 0.86[1/pixel] when $\lambda = 400 \times 10^{-9}$ to 800×10^{-9} [m] and $z'/d = 4$ to 16. Figure 11 shows an example.

4.1.2. CCD Aperture. Any CCD imager has an aperture as its light sensitive area, engendering an integration effect. Its impulse response is represented as

$$a(u) = \begin{cases} 1/a & \text{if } |u| \le a/2, \\ 0 & \text{otherwise,} \end{cases} \tag{31}$$

where a is the CCD aperture and u is the position in the image. Its Fourier transform is

$$A(f) = \mathrm{sinc}(af). \tag{32}$$

Figure 12 shows an example.

4.1.3. Approximation. Imaging characteristics depend on the wave length, the optical aperture, and the CCD aperture. In addition, there are other high-frequency blocking attributes like the optical lens aberration or the optical LPF in front of CCD. We selected the Gaussian function with mean zero and standard deviation σ as the approximated total imaging characteristics. In this subsection, the relations between the parameter σ and aparture a are investigated. Then the range of σ is examined.

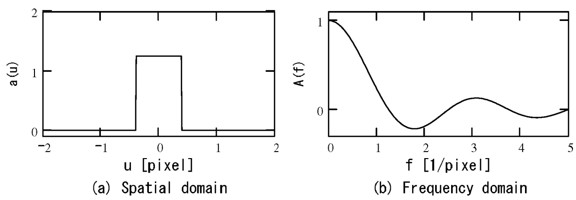

Figure 12. Examples of $a(u)$ and $A(f)$. ($a = 0.8$).

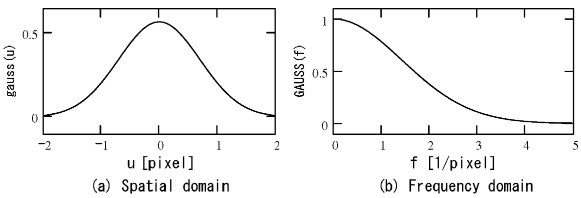

Figure 13. Examples of *gauss*(u) and *GAUSS*(f). ($\sigma = 0.7$).

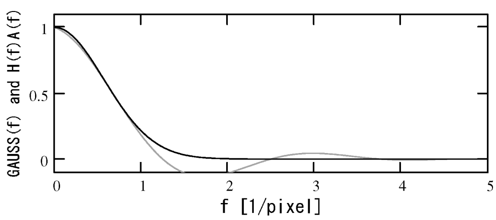

Figure 14. An example of the approximation at $a = 0.8$. Black line: *GAUSS*(f). Green line: $H(f)A(f)$.

The Gaussian function with mean zero and standard deviation σ, and its Fourier transform are (Fig. 13)

$$gauss(u) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{u^2}{2\sigma^2}}, \tag{33}$$

$$GAUSS(f) = e^{-\frac{1}{2}\sigma^2 f^2}. \tag{34}$$

The total transfer function $H(f)A(f)$ is approximated with the Gaussian function *GAUSS*(f), which has the same cut-off spatial frequency f_e,

$$H(f_e)A(f_e) = GAUSS(f_e) = 1/2. \tag{35}$$

Figure 14 shows this approximation.

Then, from Eqs. (30), (32) and (35), the relation between σ and a becomes

$$H\left(\frac{\sqrt{2\ln 2}}{\sigma}\right) \times \mathrm{sinc}\left(a\frac{\sqrt{2\ln 2}}{\sigma}\right) = \frac{1}{2}. \tag{36}$$

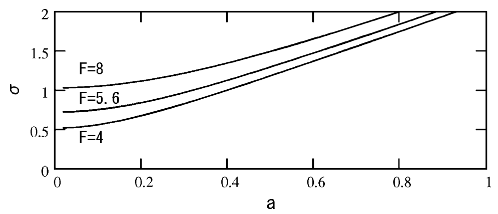

Figure 15. Relations between σ and a. $\lambda = 600 \times 10^{-9}$ [m], $z'/d =$ 4, 5.6, and 8 (=F value).

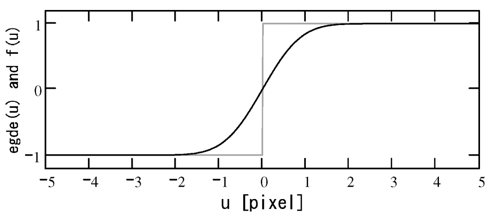

Figure 16. Edge image model (black line) ($\sigma_{\text{edge}} = 0.7$). Green line shows a pure edge.

Some relations are displayed in Fig. 15, computed with $z'/d = 4$, 5.6, and 8 under conditions of $\lambda = 600 \times 10^{-9}$ [m] and the interval between pixels is 11.0×10^{-6} [m].

Next, we investigate the range of σ. As shown in Fig. 15, σ varies with the optical aperture z'/d, and the CCD aperture a. σ decreases as the optical aperture z'/d decreases, but $z'/d = 4$ can be considered as a lower limit in practical use.

In addition, σ decreases as the CCD aperture a decreases, though the CCD makers try to increase a up to 1.0 so as to gain the sensitivity. It seems unrealistic that $a < 0.25$ in recent CCD. Therefore, the lower limit of σ is found at $z'/d = 4$ and $a = 0.25$. According to Eq. (36), the range of σ can be considered as $\sigma \geq 0.7$.

4.2. Edge Image Model

The following step function can be representative of the objective optical information because step edges in the images are the most basic element used for matching.

$$edge(u) = \begin{cases} -1 & \text{if } u < 0, \\ 1 & \text{otherwise.} \end{cases} \tag{37}$$

The pixel values can be described as the convolution of Eqs. (33) and (37),

$$f(u) = \int_{-\infty}^{\infty} edge(u - \xi) gauss(\xi)\, d\xi. \tag{38}$$

Equation (38) can be simplified as

$$f(u) = 2 \int_{0}^{u} gauss(\xi) d\xi, \tag{39}$$

where u is the pixel position. We use Eq. (39) as edge image model (Fig. 16). The characteristic of this model completely depends on the standard deviation σ of the Gaussian function. We call this the σ_{edge}, which is the function of the CCD aperture a as in Eq. (36). No matter how sharp the objective optical information is, the sampled pixel values become smoother to some extent. The range of σ_{edge} is $\sigma_{\text{edge}} \geq 0.7$.

4.3. Sub-Pixel Estimation

The similarity measures in Eq. (4) or (5) can be computed from Eq. (1) or (2) using Eq. (39). Therefore, the sub-pixel estimate $\hat{d}$ can be represented as a function of given input displacement d and σ_{edge}, as presented below.

EL using SAD:

$$\hat{d}_{\text{SAD+EL}}(d, \sigma_{\text{edge}}) = \frac{1}{2} \sum_{i \in W} \text{es}(d - i), \tag{40}$$

where

$$\begin{aligned} \text{es}(x) &= \text{erf}\left(\frac{x}{\sqrt{2}\sigma_{\text{edge}}} \right) \\ &= \frac{2}{\sqrt{\pi}} \int_{0}^{\frac{x}{\sqrt{2}\sigma_{\text{edge}}}} e^{-\xi^2} d\xi, \end{aligned} \tag{41}$$

and erf() is the error function.

PB using SAD:

$$\hat{d}_{\text{SAD+PB}}(d, \sigma_{\text{edge}}) = \frac{\sum_{i \in W} \text{es}(d - i)}{4 - 2\left|\sum_i \text{es}(d - i)\right|}. \tag{42}$$

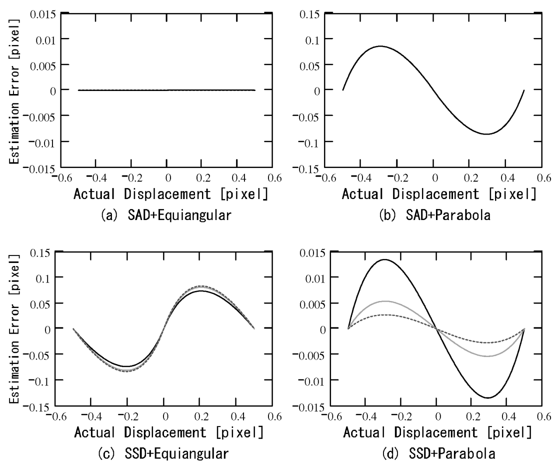

Figure 17. Sub-pixel estimation errors. $\sigma_{\text{edge}} = 0.7$ (black lines), 1.2 (green lines), and 1.7 (blue lines).

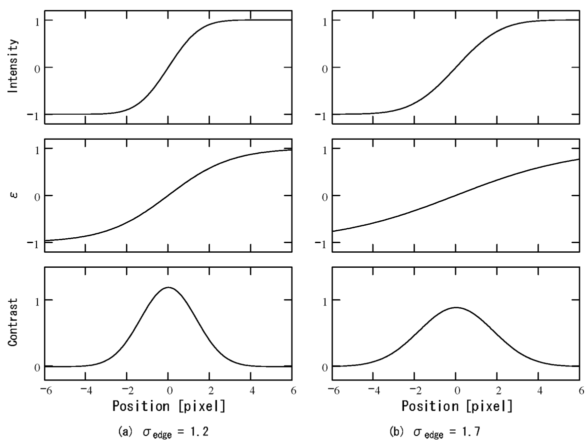

Figure 18. From top to bottom, the edge image model, ε, and contrast for the position.

EL using SSD:

$$\hat{d}_{\text{SSD+EL}}(d, \sigma_{\text{edge}}) = \begin{cases} \dfrac{1}{2} \dfrac{\sum_{i \in W} [(\text{es}(i) + \text{es}(d-i-1))^2 - (\text{es}(i) + \text{es}(d-i+1))^2]}{\sum_{i \in W} [(\text{es}(i) + \text{es}(d-i))^2 - (\text{es}(i) + \text{es}(d-i-1))^2]} & \text{if } d < 0, \\ \dfrac{1}{2} \dfrac{\sum_{i \in W} [(\text{es}(i) + \text{es}(d-i-1))^2 - (\text{es}(i) + \text{es}(d-i+1))^2]}{\sum_{i \in W} [(\text{es}(i) + \text{es}(d-i))^2 - (\text{es}(i) + \text{es}(d-i+1))^2]} & \text{otherwise.} \end{cases} \tag{43}$$

PB using SSD:

$$\hat{d}_{\text{SSD+PB}}(d, \sigma_{\text{edge}}) = \frac{1}{2} \frac{\sum_{i \in W} [\text{es}(i)(\text{es}(d-i+1) - \text{es}(d-i-1))]}{\sum_{i \in W} [(\text{es}(i+1) - \text{es}(i))(\text{es}(d-i) - \text{es}(d-i-1))]}. \tag{44}$$

Figures 17(a) through 17(d) show the estimation error $\hat{d}(d, \sigma_{\text{edge}}) - d$, as computed from Eqs. (40) through (44).

The errors shown in Fig. 17 agree with that in Fig. 9 very well, except 17(a) and 9(a). Still, we can lead the same conclusions in the Section 3.5 from Fig. 17, that is, the same *good* combinations exist, and the same pixel-locking and anti-pixel-locking effect can be seen.

4.4. Model Relations

To this point in our discussion, two completely different types of matching model have been introduced. The first one is the one-pixel-matching model, which illustrates some basic properties of sub-pixel estimation using no assumption regarding image. The second one is the edge image model, which is useful for mathematical operations. Our proposed EEC method uses the result obtained from the first model, but the second model will be used in the following sections. Therefore, the relationship between the two models should be mentioned.

The one-pixel-matching model uses the parameter of pixel intensity nonlinearity ε. Its distribution parameter n_{opm} is used for area extension. The edge model uses the parameter σ_{edge} as its blur strength. The parameter n_{opm} can be obtained from an edge model with parameter σ_{edge} if we consider that the one-dimensional edge model is one horizontal line of a two-dimensional slant edge image.

Figure 18 shows such an example. The first row depicts the edge image models with two different σ_{edge}. The second row depicts ε for the position. ε varies from -1 to 1. The first row of Fig. 19 shows this variation.

The variation sampled from the edge model is far different from a Gaussian-like distribution. This is because that one-pixel-matching model does not consider an image contrast which can be defined as $|I(1) - I(-1)|$. The contrast can be considered as a weight for the sum of similarities. The third row of Fig. 18 and the second row of Fig. 19 show the contrast. The third row of Fig. 19 illustrates the weighted ε histograms and dashed lines are their approximation with Eq. (28).

Figure 20 shows the relationship between the two parameters.

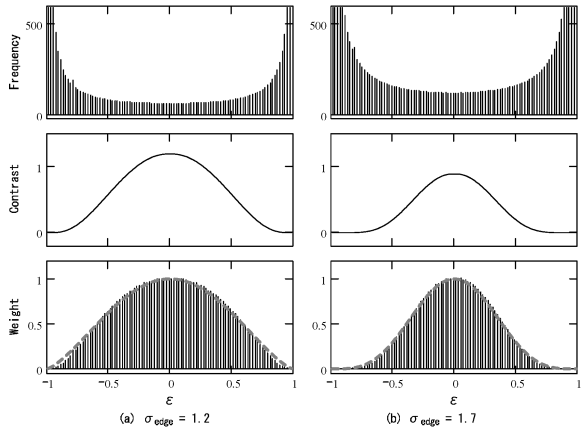

Figure 19. From top to bottom, ε frequency (histogram), contrast (weight), and total weight for ε.

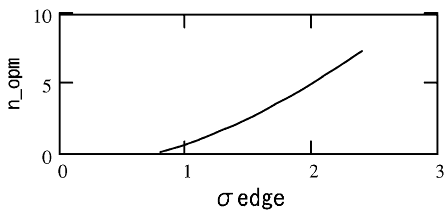

Figure 20. The relationship between n_{opm} and σ_{edge}.

5. Precise Sub-Pixel Estimation

5.1. Proposed Method

In the previous two sections, we have investigated two completely different types of matching models and their parameter relation. The both models show the same properties on the sub-pixel estimation error. The estimation error has the following features:

- The error is periodic with the pixel interval.
- The error magnitude is symmetric on $d = 0$.
- It is also nearly symmetric on $d = 0.25$ for the range $[0, 0.5]$, and on $d = -0.25$ for the range $[-0.5, 0]$.

These are considered as general features of sub-pixel estimation error in area-based image matching, through the studies in the two types of matching models. The same features can also be found in Dvornychenko (1983), Fincham and Delerce (1999), Westerweel (1998). Our proposed method utilizes these features to reduce the sub-pixel estimation error.

If we have another estimation function shifted by a half pixel from the original function, the resultant error has a negative phase versus the original estimation, meaning that this function can cancel the estimation error.

We employ the following interpolated images to achieve the cancellation function

$$\begin{aligned} g1(u) &= (f(u) + f(u+1))/2, \\ g2(u) &= (f(u-1) + f(u))/2. \end{aligned} \tag{45}$$

The shift values are $-1/2$ and $+1/2$, respectively. The reason why Eq. (45) is separated into two cases is that estimation results come into the range $[-0.5, +0.5]$. $g1(u)$ and $g2(u)$ are the same interpolated image functions with different positions. Linear interpolation is used in Eq. (45). Another interpolation function is investigated in the next subsection.

Using these shifted images, the similarity function for SSD can be represented as

$$\begin{aligned} Ri1_{\text{SSD}}(s) &= \sum_{i \in W} (g1(i) - f(i-d+s))^2 \\ &\qquad \text{if } -0.5 \le d \le 0, \\ Ri2_{\text{SSD}}(s) &= \sum_{i \in W} (g2(i) - f(i-d+s))^2 \\ &\qquad \text{otherwise.} \end{aligned} \tag{46}$$

Consequently, the sub-pixel estimations are

$$\begin{aligned} \hat{d}i1(d, \sigma_{\text{edge}}) &= \frac{Ri1(-1) - Ri1(1)}{2Ri1(-1) - 4Ri1(0) + 2Ri1(1)} - \frac{1}{2} \\ &\qquad \text{if } -0.5 \le d \le 0, \\ \hat{d}i2(d, \sigma_{\text{edge}}) &= \frac{Ri2(-1) - Ri2(1)}{2Ri2(-1) - 4Ri2(0) + 2Ri2(1)} + \frac{1}{2} \\ &\qquad \text{otherwise.} \end{aligned} \tag{47}$$

These estimations can be used to cancel the estimation errors as follows:

$$\tilde{d}(d, \sigma_{\text{edge}}) = \begin{cases} (\hat{d}(d, \sigma_{\text{edge}}) + \hat{d}i1(d, \sigma_{\text{edge}}))/2 & \text{if } -0.5 \le d \le 0, \\ (\hat{d}(d, \sigma_{\text{edge}}) + \hat{d}i2(d, \sigma_{\text{edge}}))/2 & \text{otherwise.} \end{cases} \tag{48}$$

Figure 21 shows estimation errors by conventional estimation, the cancellation function, and the proposed EEC method.

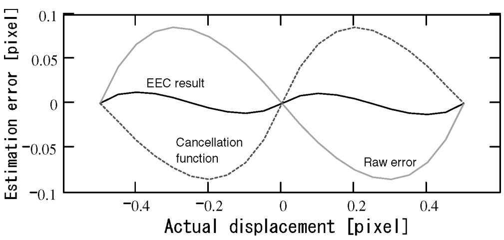

Figure 21. Illustration of the proposed method.

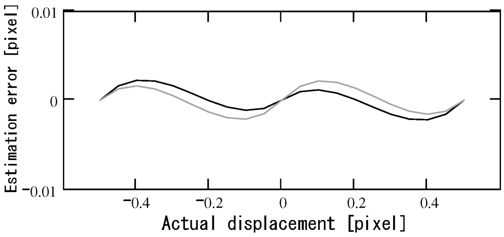

Figure 22. Residual errors of the proposed EEC method with two different interpolation functions. Black line: linear. Green line: cubic-convolution.

An essential feature of EEC method is that the systematic estimation error is cancelled by the same estimation error with inversed phase. The inversed phase estimation can be obtained from interpolated images. Therefore, the characteristics or properties of the images are irrelevant.

5.2. *Another Interpolation Function*

The following equations represent cubic-convolution interpolation (Keys, 1981): instead of the linear interpolation used in Eq. (45).

$$
\begin{aligned}
g1(u) &= \frac{-f(u-1)+5f(u)+5f(u+1)-f(u+2)}{8},\\
g2(u) &= \frac{-f(u-2)+5f(u-1)+5f(u)-f(u+1)}{8}.
\end{aligned}
\tag{49}
$$

Figure 22 shows differences of residual errors with the two interpolation functions. $\sigma_{\text{edge}} = 0.7$ and SSD are used. The figure reveals only a small difference, which become even smaller with increasing σ_{edge}.

5.3. *Residual Errors*

Residual errors can be examined using Eq. (48). In the case of PB using SAD, the error cancelled estimation is

$$
\tilde{d}_{\text{SAD+PB}}(d,\sigma_{\text{edge}}) = \frac{\left(\sum_{i\in W} es(d-i)\right)\left(5-\left|\sum_{i\in W} es(d-i)\right|\right)}{4\left(2-\left|\sum_{i\in W} es(d-i)\right|\right)\left(1+\left|\sum_{i\in W} es(d-i)\right|\right)}.
\tag{50}
$$

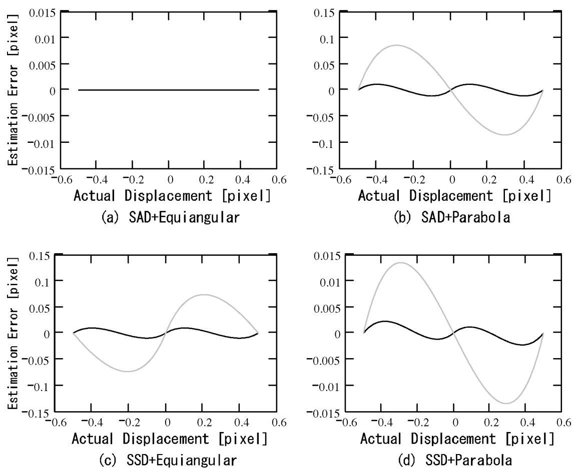

Figure 23. Sub-pixel estimation error using the proposed EEC method (black line) and conventional method (green line). $\sigma_{\text{edge}} = 0.7$.

Figure 23(b) shows residual errors of the EEC method compared with conventional errors in Eq. (42). In the case of the other combinations, only computation results are shown in Fig. 23.

5.4. *Overall Error Reduction Ratio*

Figure 24 shows RMS errors of the conventional and EEC method with four combinations. Figure 25 shows the error reduction ratio for each combination. The EEC method can reduce the estimation error up to 1/5 compared with the conventional method. Note that the equiangular line fitting with SAD (a) has almost no error.

5.5. *2-D Extension and Implementation*

In previous sections, the investigation used the 1-D image model. The 1-D image corresponds to one line in any 2-D image. Thus, it is quite straightforward to extend it to the 2-D image.

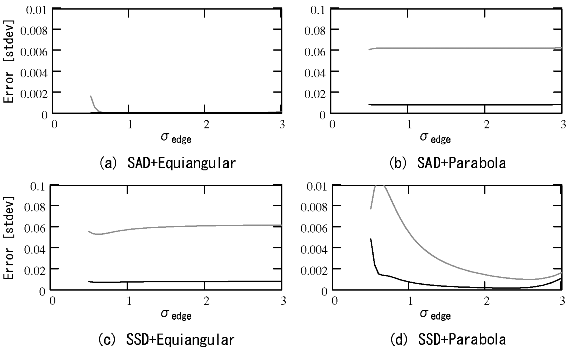

Figure 24. RMS errors.

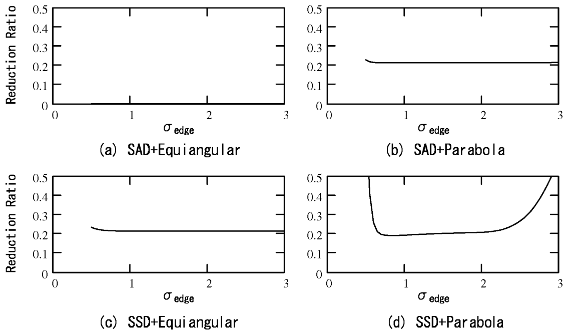

Figure 25. RMS error reduction ratio.

First, conventional sub-pixel estimation using original images is executed. Second, an interpolated image in Eq. (45) is generated by checking the sign of the sub-pixel estimation result $\hat{d}$. Third, another sub-pixel estimation using Eq. (47) is performed. Finally, the estimation errors can be cancelled with Eq. (48).

When computation time is not important, an easier implementation can be taken: (1) Perform normal matching and sub-pixel estimation using an ordinary image pair. (2) Creating the hirizontal interpolation image using one of the image pair to shift the horizontal sampling location and then execute matching and sub-pixel estimation with another image. Subsequently, compensate 0.5 [pixel] to the result. (3) Create the vertical interpolation image using one of the image pair to shift the vertical sampling location and then executing matching and sub-pixel estimation with another image. Finally, compensate 0.5 [pixel] to the result.

With these results, Eq. (48) can be used to cancel the sub-pixel estimation error. In this case, the total computation time would be three times that of the conventional method. We used this simple implementation in the following experiment.

6. Experimental Results

6.1. Synthetic Images

We employ synthetic images to verify the sub-pixel estimation error for a given input displacement. The synthetic image consists of a horizontal and vertical sinusoidal intensity sweep pattern. The normalized intensity at the position (u, v) is

$$I(u, v) = \frac{1}{2} + \frac{1}{4}\left(\cos\left(\frac{\pi u^2}{R}\right) + \cos\left(\frac{\pi v^2}{R}\right)\right), \quad (51)$$

where R is the position at which the spatial frequency becomes 1[1/pixel]. $R = 1000$ is used in this case. The spatial frequency is $(u/R, v/R)$[1/pixel] at position (u, v). Figure 26 shows this synthetic image. The image size is 640×480[pixel].

The 20 images of $I_n(u - n/20, v - n/20)$, $n = 0$ to 19, are generated. I_{10} is used as the reference image. The size of AOI is 16×16[pixel].

To match the position in the image to σ_{edge}, a half period of the sinusoidal function is approximated with the image model. The correspondence is

$$\sigma_{\text{edge}} = 0.118312T, \quad (52)$$

where T is the period of the sinusoidal function.

Figures 27(a) and (b) show the approximation of the intensity of the synthetic image with the edge image model. The center of the matching window is $u = 169$[pixel] and $u = 99$[pixel]; the corresponding σ_{edge} is 0.7 and 1.2, respectively.

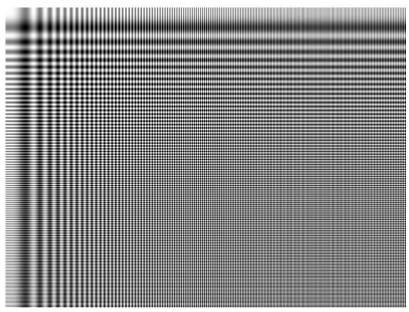

Figure 26. A synthetic image.

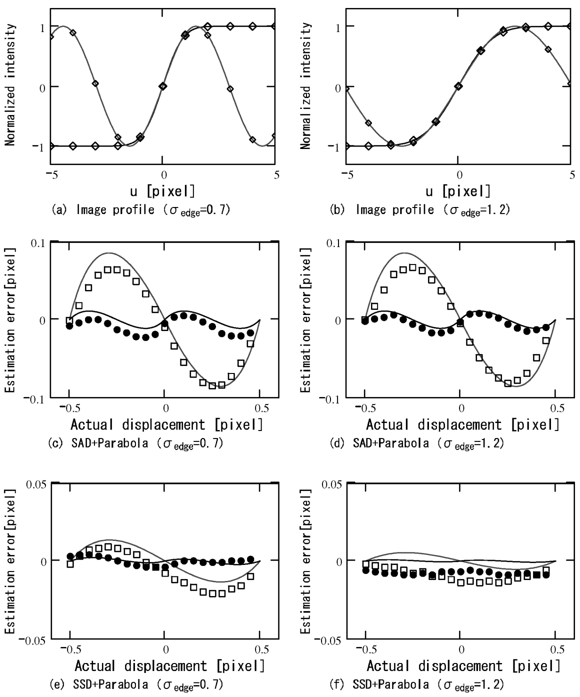

Figure 27. Experimental results using the synthetic images. (a) and (b) A partially normalized intensity pattern of the synthetic image (green line) and its corresponding edge image model (black line). Diamond marks denote the sampling position. (c) and (d) SAD. (e) and (f) SSD. (c) and (e) $\sigma_{\text{edge}} = 0.7$. (d) and (f) $\sigma_{\text{edge}} = 1.2$. Circle marks denote results using the proposed EEC method, whereas the black lines are the theoretical values. Squares denote results using the conventional method, whereas the green lines are the theoretical values.

Figures 27(c) to (f) display theoretical estimation errors and experimental results for the given input displacement. The reason that the results are not identical to the theory is inferred to be that the image is synthesized with 8-bit intensity resolution. In addition, the matching window contains multiple edges.

Figure 28 shows the standard deviation of the estimation errors over 20 estimates with respect to σ_{edge}. The black lines describe the theoretical values using the EEC method obtained from 100 different estimation values corresponding to different input displacement using Eq. (48). The green lines describe theoretical values using the conventional method. Circles and squares denote the experimental results using the EEC and the conventional method, respectively. As described in 4.2, the range of σ_{edge} is considered to be $\sigma_{\text{edge}} \geq 0.7$. In this region, the results agree well with theoretical values.

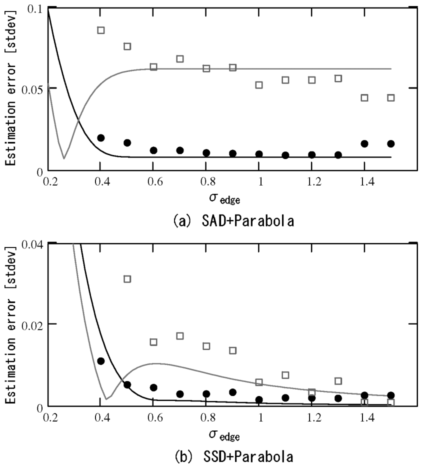

Figure 28. Experimental results expressed in standard deviation of estimated error. (a) SAD. (b) SSD. Circles denote results using the proposed EEC method, whereas the black lines are the theoretical values. Squares denote the results using the conventional method, whereas the green lines are the theoretical values.

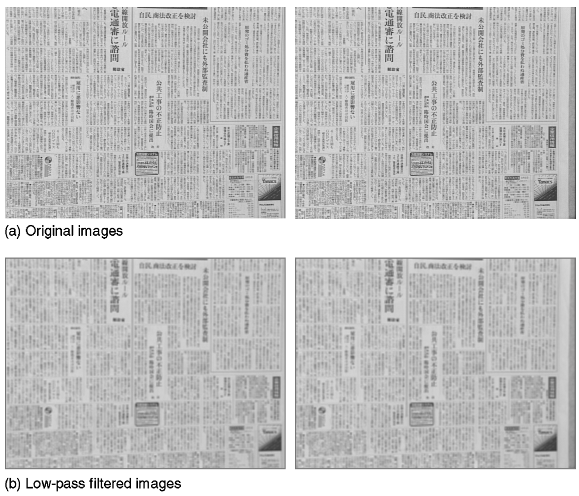

Figure 29. Images used for disparity estimation.

The experimental result using CC is almost the same as that using SSD.

6.2. *Disparity Using Real Images*

Figure 29 shows images used for the experiment. Images are captured using parallel stereo cameras. The objective target is a newspaper glued on a flat board that is not orthogonal to the optical axis. The image size is 640×480 [pixel] and the window size is 16×16 [pixel]. Disparities are obtained at 5040 positions.

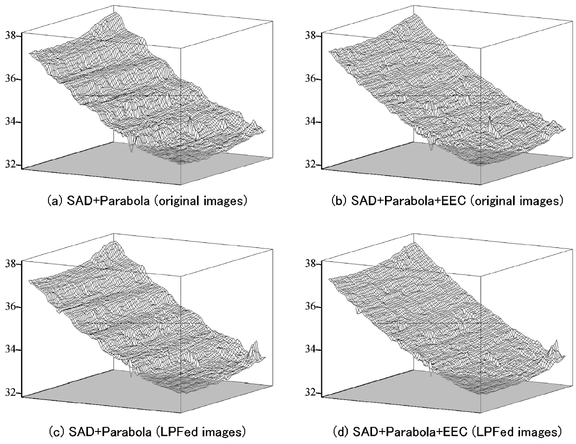

Figure 30. 3-D plots of the estimated disparities. (a) and (c) conventional, (b) and (d) proposed. The vertical axis corresponds to the disparity.

Figure 30 shows a 3-D plot of the disparities. The conventional method using SAD (30(a) and 30(c)) reveals the systematic sub-pixel estimation errors. The errors are decreased drastically in the proposed method (30(b) and 30(d)).[3]

Figure 31 shows the histograms of the horizontal component of the displacement. These results indicate that the proposed EEC method can produce an estimate with almost no systematic errors. Furthermore, the results show the following: (1) PB using SAD produces large errors regardless of the image properties; (2) EL using SSD produces errors which increase if LPF is applied to the image; (3) if the application uses a LPF, EL using SAD or PB using SSD are recommended; and (4) all above mentioned items agree with the analyses with the two types of image model.

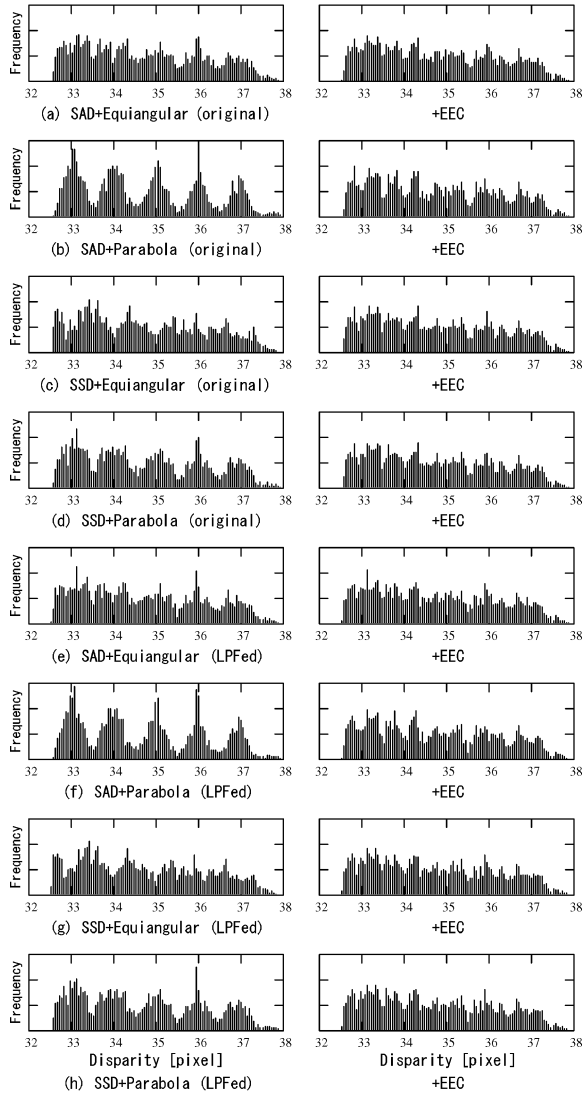

Figure 31. Histogram of the estimated disparity using images showed in Fig. 29.

6.3. Precise Target Position

Figure 1 (in Section **1**) shows images used for the experiment. The target moves linearly in a horizontal direction at a constant speed. The measured horizontal positions are expected to increase linearly to the frame number (time). Compared to Fig. 2, the positions can be measured very precisely with the EEC method as shown in Fig. 32(a).

Figure 32(b) shows sub-pixel errors against the target position. With the EEC method, the errors are reduced to about 1/5.

6.4. Other Real Images

Figure 3 (in Section 1) shows images used for the experiment. The images have been captured at slightly different camera positions and orientations. The image size is 640×480[pixel] and the window size is 16×16[pixel]. Flow-vectors are obtained at 5040 different positions.

Figure 33 shows histograms of the horizontal component of flow-vectors. The conventional method using SAD (33(a)) or SSD (33(c)) clearly shows the pixel-locking effect. This tendency is greatly reduced in the

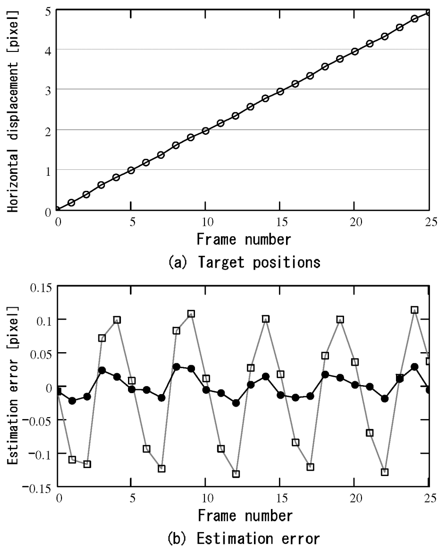

Figure 32. (a) Measured target positions with EEC method. (b) Estimation errors of the conventional method (rectangle) and EEC method (circle) against the target position.

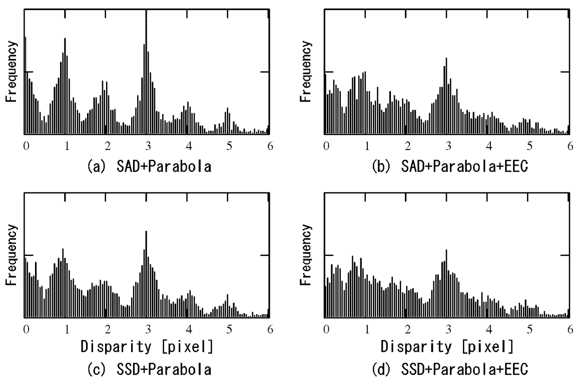

Figure 33. Histograms of the horizontal component of flow-vectors. (a) and (c) conventional. (b) and (d) EEC.

proposed EEC method (33(b) and (d)). The experimental result using CC is almost identical to that using SSD.

7. Conclusions

The combination of SAD and the sub-pixel estimation using parabola fitting is commonly used because of its short computation time. Nevertheless, those results will contain sub-pixel estimation errors regardless of σ_{edge}. We have proved and verified that equiangular line fitting using SAD produces almost equal accuracy to that of parabola fitting using SSD.

Moreover, an alternative sub-pixel estimation method has been proposed in this paper after analysis of conventional estimation errors. The proposed method achieves 1/5 of the estimation errors of the conventional method. It is independent of the similarity measure and the fitting function to be used.

Experiments using four types of images demonstrate the effectiveness of the proposed method.

Notes

1. Portions of this work have appeared in conference proceedings (Shimizu and Okutomi, 2001).
2. The considered similarity values are only for a one pixel region. Therefore, these values are termed here as the absolute difference (AD) or the squared difference (SD).
3. The disparities are not completely flat because the imaging condition is not perfect in the parallel stereo camera. In addition, the objective lens has aberrations.

References

Aggarwal, J.K. and Nandhakumar, N. 1988. On the computation of motion from sequences of images—a review. *Proceedings of the IEEE*, 76(8):917–935.

Aghajan, H.K., Schaper, C.D., and Kailath, T. 1993. Machine vision techniques for sub-pixel estimation of critical dimensions. *Optical Engineering*, 32(4):828–839.

Davis, C.Q. and Freeman, C.Q. 1998. Statistics of subpixel registration algorithms based on spatiotemporal gradients or block matching. *Optical Engineering*, 37(4):1290–1298.

Davis, C.Q., Karu, Z.Z., and Freeman, D.M. 1995. Equivalence of subpixel motion estimators based on optical flow and block matching. In *IEEE International Symposium for Computer Vision*, Coral Gables, Florida, pp. 7–12.

Driscoll, W.G. 1978. *Handbook of Optics*. New York: McGraw-Hill.

Dvornychenko, V.N. 1983. Bounds on (deterministic) correlation functions with application to registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2):206–213.

Fincham, A. and Delerce, G. 1999. Advanced optimization of correlation imaging velocimetry algorithms. In *Proceedings of the 3rd International Workshop on PIV*, additional print, Santa Barbara, USA.

Foroosh, H., Zerubia, J.B., and Berthod M. 2002. Extension of phase correlation to subpixel registration. *IEEE Transactions on Image Processing*, 11(3):188–200.

Frischholz, R.W. and Spinnler, K.P. 1993. A class of algorithms for real-time subpixel registration. In *Europto Confference*, Munich.

Fusiello, A. and Roberto, V. 2000. Symmetric stereo with multiple windowing. *International Journal of Pattern Recognition and Artificial Intelligence*, 14(8):1053–1066.

Hart, D.P. 1998. The elimination of correlation errors in piv processing. *Proceedings of the 9th International Symposium on*

Applications of Laser Techniques to Fluid Mechanics, Lisbon, Portugal, page 13.3.

Horn, B.K.P. and Schunck, B.G. 1981. Determining optical flow. *Artificial Intelligence*, 17:185–204.

Irani, M. and Peleg, S. 1991. Improving resolution by image registration. *CVGIP: Graphical Models and Image Processing*, 53(3):231–239.

Kanade, T. and Okutomi, M. 1994. A stereo matching algorithm with an adaptive window: Theory and experiment, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(9):920–932.

Keys, R. 1981. Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(6):1153–1160.

Lecordier, B., Lecordier, J.C., and Trinite, M. 1999. Iterative subpixel algorithm for the cross-correlation piv measurements. In *Proceedings of the 3rd International Workshop on PIV*, Santa Barbara, USA, pp. 37–43.

Lucas, B. and Kanade, T. 1981. An iterative image registration technique with an application to stereo vision. In *Proceedings of the 7th International Joint Conference on Artificial Intelligence*, Vancouver, 674–679.

Mitiche, A. and Mansouri, A.-R. 2004. On convergence of the Horn and Schunck optical-flow estimation method. *IEEE Transactions on Image Processing*, 13(6):848–852.

Raffel, M., Willert, C.E. and Kompenhans, J. 1998. *Particle Image Velocimetry*. Heidelberg: Springer-Verlag.

Schreier, H.W., Braasch, J.R., and Sutton, M.A. 2000. Systematic errors in digital image correlation caused by intensity interpolation. *Optical Engineering*, 39(11):2915–2921.

Shimizu, M. and Okutomi, M. 2001. Precise sub-pixel estimation on area-based matching. In *Proceedings of the 8th IEEE International Conference on Computer Vision*, Vancouver, Canada, pp. 90–97.

Shimizu, M. and Okutomi, M. 2002. An analysis of sub-pixel estimation error on area-based image matching, In *Proceedings of the 14th International Conference on Digital Signal Processing*, Santorini, Greece Vol. II, pp. 1239–1242 (w3B4).

Shortis, M.R., Clarke, T.A., and Short, T. 1994. A comparison of some techniques for the subpixel location of discrete target images. In *Proceedings of the SPIE*: *Videometrics III*, Boston, MA, USA, Vol. 2350, pp. 239–250.

Szeliski, R. and Scharstein, D. 2002. Symmetric sub-pixel stereo matching, *Proceedings of the 7th European Conference on Computer Vision*, II:525–540, Copenhagen.

Tian, Q. and Huhns, M.N. 1986. Algorithms for subpixel registration. *Computer Vision, Graphics and Image Processing*, (35):220–233.

West, G.A.W. and Clarke, T.A. 1990. A survey and examination of subpixel measurement techniques. In *Proceedings of the SPIE: Close-Range Photogrammetry Meets Machine Vision*, Zurich, Switzerland, Vol. 1395, pp. 456–463.

Westerweel, J. 1998. Effect on sensor geometry on the performance of PIV interrogation. In *Proceedings of the 9th International Symposium on Applications of Laser Techniques to Fluid Mechanics*, Lisbon, Portugal, page 1.2..