計測自動制御学会東北支部 第 215 回研究集会 (2004.5.27)
資料番号 215-9

A Correspondence Search Technique for High Accuracy 3D Measurement

○ Mohammad Abdul Muquit*, Kenji Takita* and Takafumi Aoki*

*Graduate School of Information Sciences, Tohoku University
Aoba-yama 05, Sendai, 980-8579 Japan

Keywords: phase-only correlation, phase correlation, image matching, sub-pixel matching, block matching,

Contact : Phone: +81-22-217-7169, Fax: +81-22-263-9308,
E-mail: mukit@aoki.ecei.tohoku.ac.jp

1. Introduction

Area-based image matching is an important fundamental task in a variety of image processing applications, such as stereo vision, motion analysis, image sequence analysis, etc. (see [1] for a good survey on this topic). Although in some applications pixel-level image matching may be adequate, image matching with sub-pixel accuracy is becoming essential in recent applications. In stereo vision, for example, if we can estimate the disparity between two stereo images with 1/10-pixel accuracy, we can then measure the distance of a target object with 10 times the accuracy of the stereo vision system using pixel-level image matching.

In response to this need, many methods have been developed to estimate the translational displacement between two images with high accuracy [1)–4)], and among them the methods using the *Phase-Only Correlation* (POC) function are highly effective due to their high accuracy and robust performance [5]. (Note that the POC function is sometimes called the "phase correlation function".)

Our experimental observation shows that the methods using phase information in 2D Discrete Fourier Transform (2D DFT) exhibit better registration performance than the methods using SAD (Sum of Absolute Differences) in general. In our previous work, for example, we have proposed a high-accuracy image matching method based on the POC function [6),7)], which employs (i) an analytical function fitting technique to estimate the position of the correlation peak, (ii) a windowing technique to eliminate the effect of periodicity in 2D DFT, and (iii) a spectrum weighting technique to reduce the effect of aliasing and noise. Through a set of experiments, we have demonstrated that this matching method can estimate the displacement between two images with 1/100-pixel accuracy when the image size is about 100×100 pixels.

In practice, there are many applications that require high-accuracy matching of smaller image blocks. Typical applications include stereo vision 3D measurement [8] and super-resolution image reconstruction [9]. These applications require high-

accuracy image block matching for finding corresponding points in multiple images and for estimating the position of corresponding points with sub-pixel accuracy. However, the accuracy of the conventional matching methods, including our previously proposed method [7], degrades significantly as the image size decreases because the texture information contained in an image block decreases, correspondingly.

Addressing this problem, in this paper, we extend our previously proposed matching method [7] to improve the registration accuracy for small image blocks. The basic idea is to align the position of matching windows with sub-pixel accuracy in two image blocks. This *sub-pixel window alignment* technique is combined with a coarse-to-fine search technique to implement an efficient algorithm for finding the corresponding points in multiple images. Experimental evaluation shows that the proposed correspondence search algorithm makes possible to detect the position of corresponding points with approximately 0.05-pixel accuracy when using 11×11-pixel image blocks. This paper also describes an application of the proposed sub-pixel correspondence search algorithm to high-accuracy 3D measurement.

2. High-Accuracy Image Matching Using POC

Consider two $N_1 \times N_2$ images, $f(n_1, n_2)$ and $g(n_1, n_2)$, where we assume that the index ranges are $n_1 = -M_1, \cdots, M_1$ and $n_2 = -M_2, \cdots, M_2$ for mathematical simplicity, and hence $N_1 = 2M_1 + 1$ and $N_2 = 2M_2 + 1$. Let $F(k_1, k_2)$ and $G(k_1, k_2)$ denote the 2D Discrete Fourier Transforms (2D DFTs) of the two images. $F(k_1, k_2)$ and $G(k_1, k_2)$ are given by

$$\begin{aligned} F(k_1, k_2) &= \sum_{n_1 n_2} f(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_F(k_1, k_2) e^{j\theta_F(k_1, k_2)}, \end{aligned} \quad (1)$$

$$\begin{aligned} G(k_1, k_2) &= \sum_{n_1 n_2} g(n_1, n_2) W_{N_1}^{k_1 n_1} W_{N_2}^{k_2 n_2} \\ &= A_G(k_1, k_2) e^{j\theta_G(k_1, k_2)}, \end{aligned} \quad (2)$$

where $k_1 = -M_1, \cdots, M_1$, $k_2 = -M_2, \cdots, M_2$, $W_{N_1} = e^{-j\frac{2\pi}{N_1}}$, $W_{N_2} = e^{-j\frac{2\pi}{N_2}}$, and the operator $\sum_{n_1 n_2}$ denotes $\sum_{n_1=-M_1}^{M_1} \sum_{n_2=-M_2}^{M_2}$. $A_F(k_1, k_2)$ and $A_G(k_1, k_2)$ are amplitude components, and $e^{j\theta_F(k_1, k_2)}$ and $e^{j\theta_G(k_1, k_2)}$ are phase components.

The cross-phase spectrum (or normalized cross spectrum) $\hat{R}(k_1, k_2)$ is defined as

$$\begin{aligned} \hat{R}(k_1, k_2) &= \frac{F(k_1, k_2)\overline{G(k_1, k_2)}}{|F(k_1, k_2)\overline{G(k_1, k_2)}|} \\ &= e^{j\theta(k_1, k_2)}, \end{aligned} \quad (3)$$

where $\overline{G(k_1, k_2)}$ denotes the complex conjugate of $G(k_1, k_2)$ and $\theta(k_1, k_2) = \theta_F(k_1, k_2) - \theta_G(k_1, k_2)$. The Phase-Only Correlation (POC) function $\hat{r}(n_1, n_2)$ is the 2D Inverse Discrete Fourier Transform (2D IDFT) of $\hat{R}(k_1, k_2)$ and is given by

$$\hat{r}(n_1, n_2) = \frac{1}{N_1 N_2} \sum_{k_1 k_2} \hat{R}(k_1, k_2) W_{N_1}^{-k_1 n_1} W_{N_2}^{-k_2 n_2}, \quad (4)$$

where $\sum_{k_1 k_2}$ denotes $\sum_{k_1=-M_1}^{M_1} \sum_{k_2=-M_2}^{M_2}$.

Now consider $f_c(x_1, x_2)$ as a 2D image defined in continuous space with real-number indices x_1 and x_2. Let δ_1 and δ_2 represent sub-pixel displacements of $f_c(x_1, x_2)$ to x_1 and x_2 directions, respectively. So, the displaced image can be represented as $f_c(x_1 - \delta_1, x_2 - \delta_2)$. Assume that $f(n_1, n_2)$ and $g(n_1, n_2)$ are spatially sampled images of $f_c(x_1, x_2)$ and $f_c(x_1 - \delta_1, x_2 - \delta_2)$, and are defined as

$$\begin{aligned} f(n_1, n_2) &= f_c(x_1, x_2)|_{x_1 = n_1 T_1, x_2 = n_2 T_2}, \\ g(n_1, n_2) &= f_c(x_1 - \delta_1, x_2 - \delta_2)|_{x_1 = n_1 T_1, x_2 = n_2 T_2}, \end{aligned}$$

where T_1 and T_2 are the spatial sampling intervals, and index ranges are given by $n_1 = -M_1, \cdots, M_1$ and $n_2 = -M_2, \cdots, M_2$. The POC function $\hat{r}(n_1, n_2)$ between $f(n_1, n_2)$ and $g(n_1, n_2)$ will be given by

$$\hat{r}(n_1, n_2) \simeq \frac{\alpha}{N_1 N_2} \frac{\sin\{\pi(n_1 + \delta_1)\}}{\sin\{\frac{\pi}{N_1}(n_1 + \delta_1)\}} \frac{\sin\{\pi(n_2 + \delta_2)\}}{\sin\{\frac{\pi}{N_2}(n_2 + \delta_2)\}} \quad (5)$$

where $\alpha \leqq 1$.

We use Eq. (5) — the closed-form peak model of the POC function — directly for estimating the peak position by function fitting. A spectral weighting technique and a windowing technique are used to reduce aliasing and noise effects and to overcome the periodicity effect in DFT, respectively.

3. Sub-Pixel Correspondence Search Algorithm

3.1 Algorithm Overview

The POC-based image matching technique described in the previous section is effective when the image size $N_1 \times N_2$ is relatively large [7]. We have demonstrated that the POC-based image matching can estimate the displacement between two images with 1/100-pixel accuracy when the image size is about 100×100 pixels.

In practice, there are many applications that require high-accuracy matching of small image blocks. Such applications include 3D measurement by stereo vision and super-resolution image reconstruction, where corresponding points in two images must be determined with high accuracy using block matching. In these applications, the accuracy of block matching (i.e., the accuracy of local displacement estimation) determines the total system performance directly. However, the accuracy of the conventional matching methods, including our previously proposed method, degrades significantly as the image size decreases, since the texture information contained in an image block decreases. Thus, it is difficult to implement a robust algorithm for high-accuracy correspondence search.

Addressing this problem, this section proposes a robust algorithm for finding the corresponding points between two images with high accuracy. The problem considered here can be stated as follows: Given an arbitrary pixel in a reference image, the problem is to find the position of the corresponding pixel in the second image with sub-pixel accuracy. The proposed algorithm combines the three basic techniques: (i) the POC-based image matching technique described in the previous section, (ii) a coarse-to-fine search technique for pixel-level correspondence estimation, and (iii) a sub-pixel window alignment technique for sub-pixel correspondence estimation.

Consider two images $I_0(n_1, n_2)$ and $J_0(n_1, n_2)$, which are taken at different times, or are taken from different sensors. We try to find a pair of matching points between the images. Let a pair of matching image points be $\boldsymbol{p}_0$ and $\boldsymbol{q}_0$, which exist in $I_0(n_1, n_2)$ and $J_0(n_1, n_2)$, respectively. Usually, one of the points, say $\boldsymbol{p}_0$, is specified in advance in the image $I_0(n_1, n_2)$. Then, the corresponding point $\boldsymbol{q}_0$ is searched in the second image $J_0(n_1, n_2)$.

Our algorithm consists of two stages: a stage for pixel-level correspondence estimation and a stage for sub-pixel-level correspondence estimation. In pixel-level estimation, we detect the corresponding point $\boldsymbol{q}_0$ with pixel-level accuracy using coarse-to-fine POC, so that the estimation error becomes less than 1 pixel. In the sub-pixel estimation stage,

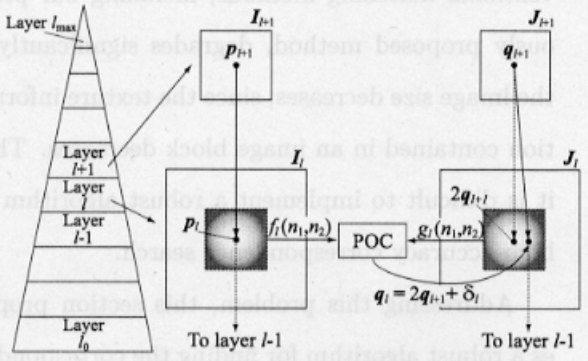


Fig. 1 Pixel-level correspondence search using a coarse-to-fine approach.

we recursively improve the sub-pixel accuracy of POC block matching by adjusting the location of the window function with sub-pixel accuracy. We call this technique *sub-pixel window alignment.*

3.2 Correspondence Estimation with Pixel-Level Accuracy

The pixel-level estimation stage mainly employs the coarse-to-fine correspondence search technique, where some coarser versions of the original input images are created. The POC-based block matching starts at the coarsest image layer and the operation gradually moves to finer layers. Figure 1 shows an overview of the technique. Let $\boldsymbol{p}_l$ and $\boldsymbol{q}_l$ be coordinate vectors of matching image points at the l–th layer. The goal in the pixel-level estimation is to find $\boldsymbol{q}_0$ that corresponds to the given reference point $\boldsymbol{p}_0$ at the layer 0 (the original image layer) with pixel accuracy.

Procedure for pixel-level correspondence estimation

Input:

images $I_0(n_1, n_2)$ and $J_0(n_1, n_2)$,

reference point $\boldsymbol{p}_0$ in $I_0(n_1, n_2)$

Output:

corresponding point $\boldsymbol{q}_0$ in $J_0(n_1, n_2)$

Step 1: For $l = 1, 2, \cdots, l_{max}$, create the l–th layer images $I_l(n_1, n_2)$ and $J_l(n_1, n_2)$, i.e., coarser versions of $I_0(n_1, n_2)$ and $J_0(n_1, n_2)$, recursively as follows:

$$I_l(n_1, n_2) = \frac{1}{4}\sum_{i_1=0}^{1}\sum_{i_2=0}^{1} I_{l-1}(2n_1 + i_1, 2n_2 + i_2),$$

$$J_l(n_1, n_2) = \frac{1}{4}\sum_{i_1=0}^{1}\sum_{i_2=0}^{1} J_{l-1}(2n_1 + i_1, 2n_2 + i_2).$$

In our experiments, we choose $l_{max} = 4$.

Step 2: For every layer $l = 1, 2, \cdots, l_{max}$, calculate the coordinate $\boldsymbol{p}_l = (p_{l1}, p_{l2})$ corresponding to the original reference point $\boldsymbol{p}_0$ recursively as follows:

$$\boldsymbol{p}_l = \lfloor \tfrac{1}{2}\boldsymbol{p}_{l-1} \rfloor = (\lfloor \tfrac{1}{2}p_{l-1\ 1} \rfloor, \lfloor \tfrac{1}{2}p_{l-1\ 2} \rfloor), \quad (6)$$

where $\lfloor x \rfloor$ denotes the operation to round the element of x to the nearest integer towards minus infinity.

Step 3: We assume that $\boldsymbol{q}_{l_{max}} = \boldsymbol{p}_{l_{max}}$ in the coarsest layer. Let $l = l_{max} - 1$. (Actually, we can omit the creation of the l_{max}–th layer images.)

Step 4: From the l–th layer images $I_l(n_1, n_2)$ and $J_l(n_1, n_2)$, extract two sub-images (or image blocks) $f_l(n_1, n_2)$ and $g_l(n_1, n_2)$ with their centers on $\boldsymbol{p}_l$ and $2\boldsymbol{q}_{l+1}$, respectively. For accurate matching, the size of image blocks should be reasonably large. In our experiments, we use 31×31 image blocks.

Step 5: Estimate the displacement between $f_l(n_1, n_2)$ and $g_l(n_1, n_2)$ with pixel accuracy using POC-based image matching, which is the simplified version of the matching algorithm described in Sect. 2.. We eliminate the function fitting and spectral weighting technique (Sect. 2.) for simplicity. Only the windowing technique (Sect. 2.) is employed. The displacement between the two image blocks is es-

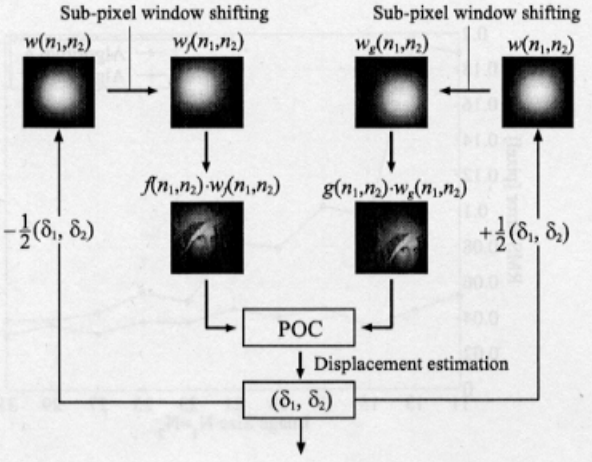


Fig. 2 Sub-pixel correspondence search using a sub-pixel window alignment technique.

timated by detecting the position of the maximum value of the POC function $\hat{r}(n_1, n_2)$ with pixel accuracy. Let the estimated displacement vector be $\boldsymbol{\delta}_l$. The l–th layer correspondence $\boldsymbol{q}_l$ is determined as follows:

$$\boldsymbol{q}_l = 2\boldsymbol{q}_{l+1} + \boldsymbol{\delta}_l. \quad (7)$$

Step 6: Decrement the counter by 1 as $l = l - 1$ and repeat from **Step 4** to **Step 6** while $l \geq 0$.

□

3.3 Correspondence Estimation with Sub-Pixel Accuracy

The sub-pixel estimation stage, on the other hand, employs the sub-pixel window alignment technique illustrated in Fig. 2.

Procedure for sub-pixel correspondence estimation

Input:

images $I_0(n_1, n_2)$ and $J_0(n_1, n_2)$,

reference point $\boldsymbol{p}_0$ in $I_0(n_1, n_2)$ given with pixel-level accuracy,

corresponding point $\boldsymbol{q}_0$ in $J_0(n_1, n_2)$ given with pixel-level accuracy

Output:

reference point $\boldsymbol{p}_{-1}$ in $I_0(n_1, n_2)$ given with sub-pixel accuracy,

corresponding point $\boldsymbol{q}_{-1}$ in $J_0(n_1, n_2)$ given with sub-pixel accuracy

Step 1: From the original images $I_0(n_1, n_2)$ and $J_0(n_1, n_2)$, extract $N_1 \times N_2$ sub-images $f(n_1, n_2)$ and $g(n_1, n_2)$ with their centers on $\boldsymbol{p}_0$ and $\boldsymbol{q}_0$, respectively. For mathematical simplicity, we assume that the index ranges of the two sub-images are given by $n_1 = -M_1, \cdots, M_1$ and $n_2 = -M_2, \cdots, M_2$, and hence $N_1 = 2M_1 + 1$ and $N_2 = 2M_2 + 1$.

Step 2: Set the initial window functions for $f(n_1, n_2)$ and $g(n_1, n_2)$, denoted by $w_f(n_1, n_2)$ and $w_g(n_1, n_2)$, as $w_f(n_1, n_2) = w_g(n_1, n_2) = w(x_1, x_2)|_{x_1=n_1, x_2=n_2}$, where for real variables x_1 and x_2, the window function $w(x_1, x_2)$ is defined as

$$w(x_1, x_2) = \frac{1 + \cos(\frac{\pi x_1}{M_1})}{2} \frac{1 + \cos(\frac{\pi x_2}{M_2})}{2}.$$

Instead of the Hanning window function in the above equation, other window functions such as Hamming, Gaussian, and Kaiser can also be used.

Step 3: Estimate the displacement between $f(n_1, n_2)$ and $g(n_1, n_2)$ with sub-pixel accuracy using the POC-based image matching described in Sect. 2.. Let the estimated sub-pixel displacement vector be represented as $\boldsymbol{\delta} = (\delta_1, \delta_2)$, where $\delta_1, \delta_2 < 1$. In our experiments, 5×5 data points around the maximum peak of a correlation array are used in function fitting ((i) in Sect. 2). As for the spectral weighting technique ((iii) in Sect. 2), the parameters U_1 and U_2 of the low-pass-type weighting function $H(n_1, n_2)$ defined in Eq. (7) are set as $U_1 = \lceil M_1/2 \rceil$ and $U_2 = \lceil M_2/2 \rceil$, where $\lceil x \rceil$ denotes the operation to round the element of x to the nearest integer towards infinity. These parameters are optimized empirically.

Step 4: Update the window functions $w_f(n_1, n_2)$ and $w_g(n_1, n_2)$ by shifting their centers as

$$w_f(n_1, n_2) = w(x_1 + \frac{\delta_1}{2}, x_2 + \frac{\delta_2}{2})\Big|_{x_1=n_1, x_2=n_2},$$
$$w_g(n_1, n_2) = w(x_1 - \frac{\delta_1}{2}, x_2 - \frac{\delta_2}{2})\Big|_{x_1=n_1, x_2=n_2}.$$

Step 5: Repeat from **Step 3** to **Step 4** for specific times. This is called the "sub-pixel window alignment" technique, which gradually reduces the sub-pixel error in the estimated displacement vector. In our experiments, the number of iterations is 5.

Step 6: As a result, the final matching points are estimated with subpixel accuracy as follows:

$$\boldsymbol{p}_{-1} = \boldsymbol{p}_0 - \frac{\boldsymbol{\delta}}{2},$$
$$\boldsymbol{q}_{-1} = \boldsymbol{q}_0 + \frac{\boldsymbol{\delta}}{2}.$$

□

Note that in some applications we must keep the reference point unchanged, i.e., $\boldsymbol{p}_{-1} = \boldsymbol{p}_0$. In this case, we use the following update equations for window functions:

$$w_f(n_1, n_2) = w(x_1, x_2)|_{x_1=n_1, x_2=n_2},$$
$$w_g(n_1, n_2) = w(x_1 - \delta_1, x_2 - \delta_2)|_{x_1=n_1, x_2=n_2}.$$

Accordingly, the final matching points are given by

$$\boldsymbol{p}_{-1} = \boldsymbol{p}_0,$$
$$\boldsymbol{q}_{-1} = \boldsymbol{q}_0 + \boldsymbol{\delta}.$$

The key point of our proposed technique is to align the image areas for calculation of the POC function so as to improve the sub-pixel accuracy of POC matching. Note that we can consider two possible ways of performing sub-pixel alignment of image areas: image shift and window shift. Sub-pixel image shift requires interpolation techniques, and its accuracy is limited by the accuracy of interpolation. It is much easier to perform sub-pixel shift of the window function, since the window function is defined on the continuous variables (x_1, x_2).

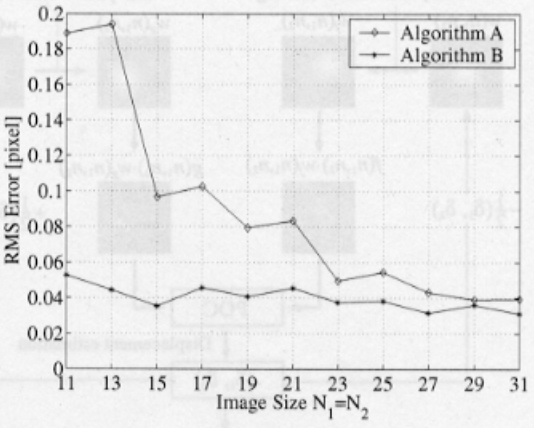


Fig. 3 RMS error versus image size.

We evaluate the following two new algorithms in comparison with simple POC-based image matching described in Sect. 2..

[Algorithm A] The sub-pixel correspondence search algorithm that employs the coarse-to-fine search technique (in Sect. 3.2) followed by simple sub-pixel POC matching (in Sect. 2.).

[Algorithm B] The sub-pixel correspondence search algorithm that employs the coarse-to-fine search technique (in Sect. 3.2) followed by the sub-pixel window alignment technique (in Sect. 3.3).

The target we use in the evaluation is a wooden board with rich texture. Images are captured with a CCD camera (JAI CVM10 with Nikon Rayfact TM2514B1 lens), which is mounted on a micro stage that allows precise alignment of the camera position. The camera is moved and images are captured after each movement. We calculate the actual movement using the micro stage, and calculate the displacement using algorithm A and algorithm B. Figure 3 shows the impact of the sub-pixel window alignment technique on the reduction of the RMS

(Root Mean Square) error in displacement estimation. The sub-pixel window alignment technique is highly effective especially for smaller image sizes.

We also determine the disparity between two images using both algorithms for evaluation. Two images are taken at two camera positions with a distance of only 1 cm by moving the camera on the micro stage. The wooden board is placed 30 cm away from the camera and is set slightly panned with respect to the focal plane of the camera. The baseline is so narrow that almost no projective distortion is observed. This enables our stereo system to measure objects with various degrees of pan. The size of images taken with the camera is 640×480 and approximately 5000 points are chosen as reference points for correspondence search and disparity estimation. The block size used for matching is 15×15.

Figures 5 (a) and (b) show the measured disparity (i.e., length of a disparity vector), where (a) and (b) are the cases of algorithms A and B (described in Sect. 3.), respectively. In Fig. 5 (a), we can see stepwise disparity error up to a level, but the algorithm B (shown in Fig. 5 (b)) produces a smooth disparity change, which directly reflects the flat surface of the wooden board. Figure 5 (b) clearly shows that the sub-pixel window alignment technique achieves higher accuracy in 3D reconstruction.

4. Application to 3D Measurement

In this section, we implement a high-accuracy 3D measurement system using stereo vision by applying our proposed correspondence search algorithm. In stereo vision, area-based matching methods are frequently used to determine correspondences between stereo images (usually with pixel-level accuracy). The overall accuracy of 3D measurement system is mainly determined by (i) the baseline length between two cameras and (ii) the accuracy of estimated disparity between corresponding points. A conventional approach to high-accuracy 3D measurement is to employ a wide-baseline camera pair combined with edge-based correspondence matching. In this approach, only limited number of corresponding points located on edges can be used for 3D reconstruction. On the other hand, the area-based correspondence matching technique (which must be combined with narrow-baseline stereo cameras to avoid projective distortion between stereo

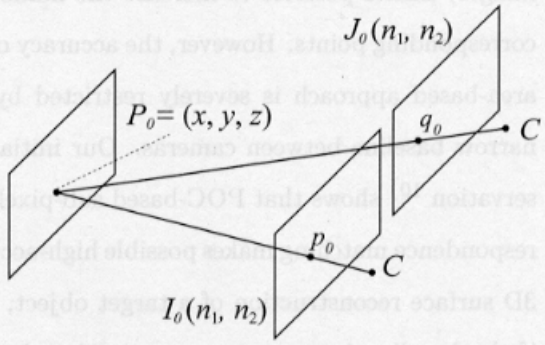
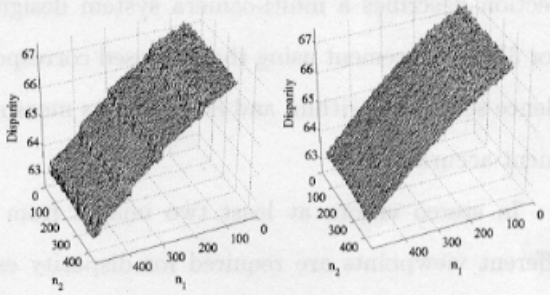


Fig. 4 3D measurement based on stereo vision



(a) (b)

Fig. 5 Disparity maps of the wooden board: (a) conventional pixel-level correspondence search algorithm, (b) algorithm A, and (c) algorithm B.

images) makes possible to increase the number of corresponding points. However, the accuracy of the area-based approach is severely restricted by the narrow baseline between cameras. Our initial observation [10] shows that POC-based sub-pixel correspondence matching makes possible high-accuracy 3D surface reconstruction of a target object, even if the baseline between two cameras is relatively short. Also, the proposed technique is robust to intensity variation between stereo images. This section describes a multi-camera system designed for 3D measurement using the proposed correspondence search algorithm, and evaluates its measurement accuracy.

In stereo vision, at least two images from different viewpoints are required for disparity estimation as illustrated in Fig. 4, where we assume parallel cameras for simplicity. The 3D coordinate $P_0 = (x, y, z)$ is a real world point, which is projected to p_0 and q_0 on the left and right camera images, $I_0(n_1, n_2)$ and $J_0(n_1, n_2)$, respectively. To determine the 3D coordinate P_0, we need to calibrate the cameras in advance and to know the corresponding points p_0 and q_0. The proposed correspondence search technique is used to find p_0 and q_0 in stereo images.

We evaluate the accuracy of 3D measurement based on our proposed correspondence search algorithm. In general, for the case of using area-based method, it is important to adjust the baseline length of the stereo cameras to reduce projective image distortion. We implement a high-accuracy multi-camera 3D measurement system shown in Fig. 6, where three pairs of narrow-baseline stereo cameras are employed with the POC-based correspondence search technique [10]. In our system, the

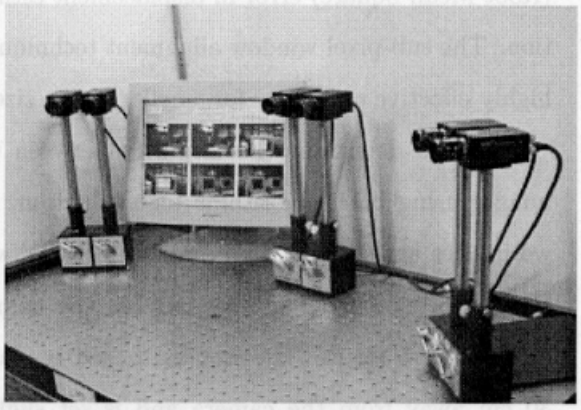


Fig. 6 The multi-camera system

baseline lengths between stereo cameras of each pair are around 5 cm. The proposed sub-pixel correspondence search technique is useful in such applications.

We calibrate the camera at each position with a conventional method [8], where the object with a checker pattern is used. We detect the corners on the checker pattern using the Harris corner detector [11], and determine the corresponding corners in stereo images. Using these information, we determine the basic camera parameters, such as the relative rotation/translation between the cameras and lens distortion parameters, which are needed to reconstruct 3D coordinates of the target object. After the camera calibration, we take two images at two camera positions, and determine the correspondences between the two images. We calculate disparity vectors for the pairs of corresponding points. While determining corresponding points between stereo images, unreliable correspondence occurs because of occlusion and projective distortion. We employ epipolar constraint [12] to evaluate the reliabilty of each point. For each reference point, we determine the Euclidean distance of the epipolar line from the corresponding point, and e-

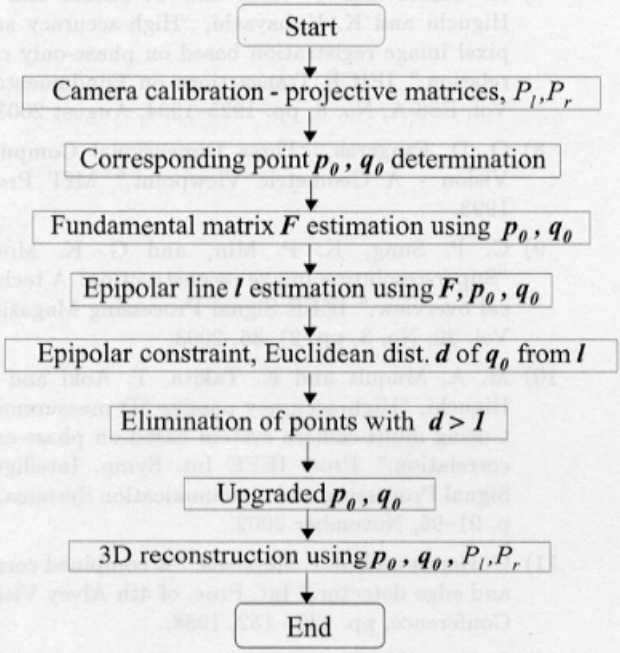


Fig. 7 Flowchart of the mechanism

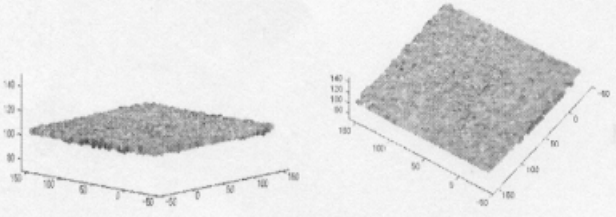


Fig. 8 The reconstructed plain surface

liminate the points with that lie more than 1 pixel away from the epipolar line. Fig. 7 shows a simple flowchart of the camera system's mechanism.

We measure a plain surface containing random texture using the mulit-camera system. The plain object is palced around 90 cm away from the cameras. Fig. 8 reconstructs the 3D coordinates of the surface using our proposed correspondence search algorithm. We evaluate the reconstruction accuracy of each camera pair separately.

The reconstruction accuracy is evaluated by fitting an ideal plane to the 3D surface data in Fig. 8. The error is evaluated by calculating the distances

Table 1 Reconstruction Accuracy (Plain Surface)

Error [mm]	Pair1–2	Pair3–4	Pair5–6	All
RMS	0.41	0.49	0.47	0.64
Max	1.53	2.25	2.34	4.15

Fig. 9 Reconstructed human face

of the measured points from the ideal plane. Table 1 shows the evaluated RMS errors for each camera pair. The multi-camera system can reconstruct plain surface with sub-mm accuracy. This result clearly demonstrates that the proposed method can be used for high accuracy 3D measurement.

Figures of a reconstructed human face using the proposed camera system are shown in Fig. 9, where the reconstructed face is shown from different view angles regarding the 3D coordinate. In Fig. 9, we use around 20000 points for reconstruction. We emphasize that no interpolation is used in the reconstructed object presented here. Fig. 9 shows that the system reconstructs 3D objects with very high accuracy.

5. Conclusion

In this paper, we have proposed a high-accuracy correspondence search algorithm, which employs (i) a POC-based image matching technique, (ii) a coarse-to-fine search technique for pixel-level correspondence estimation, and (iii) a sub-pixel window alignment technique for sub-pixel correspondence estimation. Experimental evaluation of matching accuracy shows that the proposed method makes possible to estimate image displacements with approximately 0.05-pixel accuracy when using 11×11 image blocks. We have also implemented a high accuracy multi-camera system using the proposed technique for 3D measurement. Our system can achieve sub-mm accuracy in 3D measurement, even when the baseline length between stereo cameras is 5 cm and the camera-object distance is about 90 cm. Also, we are planning to apply the proposed technique to improve the performance of POC-based industrial vision sensors developed by our business-academia collaboration (see [13] for example).

参考文献

1) L. G. Brown, "A survey of image registration techniques," ACM Computing Surveys, Vol. 24, No. 4, pp. 325–376, december 1992.

2) Q.Tian and M. N. Huhns, "Algorithms for subpixel registrations," Computer Vision, Graphics and Image Processing, Vol. 35, No. 2, pp. 220–233, August 1986.

3) G.A. Thomas, "Television motion measurement for DATV and other applications," BBC Research Department Report, September 1987.

4) Foroosh H., J. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," IEEE Trans. on Image Processing, No. 2707, 2002.

5) C. D. Kuglin and D.C. Hines, "The phase correlation image alignment method," Proc. int. Conf. on Cybernetics and Society, pp. 163–165, 1975.

6) T. aoki, K. takita, T. Higuchi and K. Kobayashi, "Phase-based image matching and its application to intelligent vision systems," Proc. Int. Symp. New Paradigm VLSI Computing, pp. 95–100, December 2002.

7) K. Takita and T. Aoki and Y. Sasaki and T. Higuchi and K. Kobayashi, "High-accuracy subpixel image registration based on phase-only correlation," IEICE Transactions on Fundamentals, Vol. E86-A, No. 8, pp. 1925–1934, August 2003.

8) O. D. Faugeras, "Three Dimensional Computer Vision - A Geometric Viewpoint," MIT Press, 1993.

9) C. P. Sung, K. P. Min, and G. K. Moon, "Super-resolution image reconstruction: A technical overview," IEEE Signal Processing Magazine, Vol. 20, No. 3, pp. 21–36, 2003.

10) M. A. Muquit and K. Takita, T. Aoki and T. Higuchi, "High-accuracy passive 3D measurement using multi-camera system based on phase-only correlation," Proc. IEEE Int. Symp. Intelligent Signal Processing and Communication Systems, pp. 91–95, November 2002.

11) C. Harris and M.J. Stephens, "A combined corner and edge detector," Int. Proc. of 4th Alvey Vision Conference, pp. 147– 152, 1988.

12) G. Xu and Z. Zhang, "Epipolar Geometry in Stereo, Motion and Object recognition," Kluwer Academic Publishers, 1996.

13) http://www.aoki.ecei.tohoku.ac.jp/poc/.