7WB,<+F0@)8f3X2qElKL;YIt Bh223 2s8&5f=82q(2005.7.27)
;qNAHV9f223-8

Design and Analysis of High-Accuracy Passive 3D Measurement System

!{Mohammad Abdul Muquit*, Takuma Shibahara*, Naohide Uchida* and Takafumi Aoki*

*Graduate School of Information Sciences, Tohoku University

Keywords: 3D measurement, stereo vision, phase-based image matching, phase-only correlation, outlier detection

Contact : 6-6-05, Aramaki Aza Aoba, Sendai-shi, 980-8579 Japan
Phone: +81-22-795-7169, Fax: +81-22-263-9308,
E-mail: mukit@aoki.ecei.tohoku.ac.jp

1. Introduction

Recently the demand of high-accuracy 3D measurement is rapidly growing in a variety of computer vision applications, for instance, robot vision, human-computer interaction, biometric authentication, etc. Existing 3D measurement techniques are classified into two major types — active and passive. In general, active measurement employs structure illumination or laser scanning, which is not necessarily desirable in many applications. On the other hand, passive 3D measurement techniques based on stereo vision have the advantages of simplicity and applicability, since such techniques require simple instrumentation. However, poor reconstruction quality still remains as a major issue for passive 3D measurement, due to the difficulty in finding accurate correspondence between stereo images; this problem is generally known as "correspondence problem" [1]. As a result, application of passive stereo vision to capturing 3D surfaces of free form objects is still weakly reported in the published literature. The objective of this paper is to implement a passive 3D measurement system, whose reconstruction accuracy is comparable with that of practical active 3D scanners based on structured light projection.

The overall accuracy of passive 3D measurement is mainly determined by (i) the baseline length between two cameras and (ii) the accuracy of estimated disparity between corresponding points [1]. Conventional approaches to passive 3D measurement employ wide-baseline camera pairs combined with feature-based correspondence matching. However, in such approaches only a limited number of corresponding points can be used for 3D reconstruction. On the other hand, area-based correspondence matching (which requires narrow-baseline stereo cameras) makes possible to increase the number of corresponding points. However, the accuracy of 3D measurement becomes severely restricted when the baseline is narrow [2]. In this paper, therefore, we focus on the techniques for high-accuracy stereo correspondence in order to overcome the limitation of measurement accuracy in narrow-baseline stereo vision.

The key idea in this paper is to employ phase-based image matching for high-accuracy stereo correspondence. Our experimental observation shows that the methods using phase-based image matching exhibit better registration performance than the meth-

ods using SAD (Sum of Absolute Differences) in general [3, 4]. In our previous work, we presented an application of phase-based image matching to a generic correspondence search problem [5], sub-pixel resolution.

The goal of this paper is to implement the phase-based correspondence search technique in a practical 3D measurement system, and to analyze its impact on the system's performance (i.e., reconstruction accuracy and reliability). We demonstrate that the use of phase-based correspondence search makes possible to achieve fully automatic high-accuracy 3D measurement with a narrow-baseline stereo vision system. Another contribution of this paper is a highly reliable technique for detecting and correcting outliers (wrong or unreliable corresponding points) based on Phase-Only Correlation (POC) function — a correlation function used in the phase-based image matching to evaluate similarity between two images. By detecting and correcting outliers, we can achieve high-quality dense 3D reconstruction of objects. Through a set of experiments, we show that the proposed system measures 3D surfaces of regular shaped objects (a solid plane and a solid sphere) with sub-mm accuracy. Also, we demonstrate the system's applicability in 3D motion (rotation and translation) estimation of free form objects (human face) as an application to the field of human-computer interaction.

2. Phase-Based Image Matching for Stereo Correspondence

In the sections 2.1 and 2.2, we describe the high-accuracy image matching based on Phase-Only Correlation (POC) function[1], and its application to stereo correspondence problem. (See our papers [4, 5] for earlier discussions on the proposed techniques.) The section 2.3 describes the outlier detection and correction technique using the POC function.

2.1 Phase-Based Image Matching

The POC function between two given images $f(n_1, n_2)$ and $g(n_1, n_2)$ is the 2D Inverse DFT (2D IDFT) of $\hat{R}(k_1, k_2)$, where $\hat{R}(k_1, k_2)$ is the cross-phase spectrum (or normalized cross spectrum) between the 2D Discrete Fourier Transforms (2D DFTs) of $f(n_1, n_2)$ and $g(n_1, n_2)$. When two images are similar, their POC function gives a distinct sharp peak. (When $f(n_1, n_2) = g(n_1, n_2)$, the POC function becomes the Kronecker delta function.) When two images are not similar, the peak drops significantly. The height of the peak can be used as a good similarity measure for image matching, and the location of the peak shows the translational displacement between the two images.

In our previous paper [5], we derive the analytical peak model for the POC function between the same images that are minutely displaced with each other. We obtain a closed-form peak model of the POC function in continuous space, which is used for estimating the peak position with sub-pixel accuracy. While using digital images, we obtain a data array for each discrete index (n_1, n_2). It is possible to find the location of the peak that may exist between image pixels by fitting the closed-form peak model function. We implement a windowing technique and a spectral weighting technique to overcome the periodicity effect of DFT and to reduce aliasing and noise effect, respectively. Details on this technique is given in [4].

All these techniques for high-accuracy sub-pixel image matching are adopted in stereo correspondence search in our system. We use POC-based block matching of size 33×33 pixels, for which we can achieve about 0.05-pixel accuracy in displacement estimation.

[1] The POC function is sometimes called the "phase correlation function".

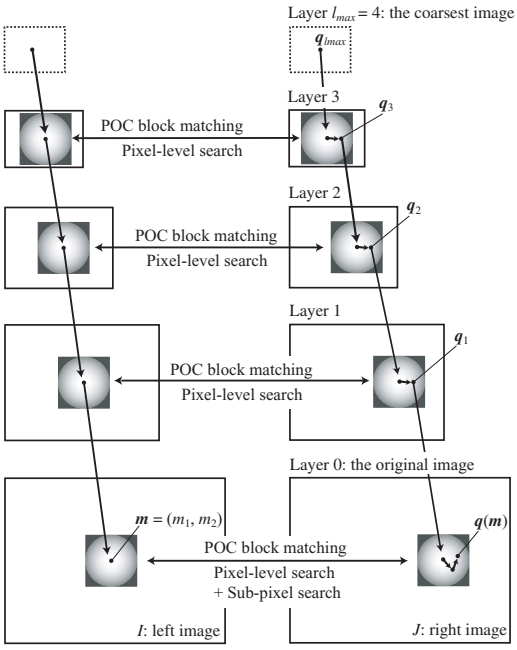

Fig. 1 ! ! Sub-pixel correspondence search using a coarse-to-fine strategy (e.g., $l_{max} = 4$).

2.2 Sub-Pixel Correspondence Search

This section discusses a high-accuracy stereo correspondence algorithm based on the sub-pixel image matching mentioned above. The algorithm described here is an improved version of the method reported in our previous paper [5]. Our algorithm employs (i) a coarse-to-fine strategy using image pyramids for robust correspondence search with POC-based block matching, and (ii) a sub-pixel window alignment technique for finding a pair of corresponding points with sub-pixel displacement accuracy. In the first stage (i), we estimate the stereo correspondence with pixel-level accuracy using hierarchical POC-based block matching with coarse-to-fine strategy. Thus, the estimation error becomes less than 1 pixel for every corresponding point. The second stage (ii) of the algorithm is to recursively improve the sub-pixel accuracy of corresponding points by adjusting the location of the applied window function with sub-pixel accuracy. As a result, the coordinates of corresponding points are obtained with sub-pixel accuracy.

Let $\boldsymbol{m} = (m_1, m_2)(\in \boldsymbol{Z}^2)$ be a coordinate vector of a reference pixel in the left image I (i.e., the reference image), where $\boldsymbol{Z}$ is the set of integers. The problem of sub-pixel correspondence search is to find a real-number coordinate vector $\boldsymbol{q}(\boldsymbol{m}) = (q_1, q_2)(\in \boldsymbol{R}^2)$ in the right image J that corresponds to the reference pixel $\boldsymbol{m}$ in I, where $\boldsymbol{R}$ is the set of real numbers, representing the coordinates in J with sub-pixel accuracy. For convenience, we use the symbol $C_I^{int}(\subset \boldsymbol{Z}^2)$ to denote the set of all integer coordinate vectors (with pixel-level accuracy) in the reference image I, and the symbol $C_J^{real}(\subset \boldsymbol{R}^2)$ to denote the set of all real-number coordinate vectors (with sub-pixel accuracy) in J. Then, the problem is to find the set of corresponding points $\boldsymbol{q}(\boldsymbol{m}) \in C_J^{real}$ for all reference points $\boldsymbol{m} \in C_I^{int}$. Figure 1 shows an overview of the sub-pixel correspondence search algorithm.

2.3 Outlier Detection and Correction

The sub-pixel correspondence search technique described in the section 2.2 determines corresponding point for any given reference point. The robustness of correspondence search is one of the most important characteristics of our phase-based approach, where a large number of corresponding points are automatically detected without extracting image features. However, because of occlusion, image noise, projective distortion etc., corresponding points for some reference points may not be estimated correctly. For such reference points, the proposed technique outputs wrong or unreliable corresponding points (generally known as outliers), which degrade the accuracy of measurement.

We propose an outlier detection technique using the peak value of the POC function as a measure of correspondence reliability. When the peak value of the POC function between the local image blocks, centered at the reference point $\boldsymbol{m} = (m_1, m_2)$ and at the corresponding point $\boldsymbol{q}(\boldsymbol{m})$, is below a certain threshold, $\boldsymbol{q}(\boldsymbol{m})$ is regarded as an outlier. This tech-

nique improves both reliability and accuracy of the overall 3D reconstruction. In conventional stereo vision systems with calibrated cameras, the use of epipolar constraint [6, 7] is essential in correspondence search as well as outlier detection. In our approach, correspondence search and outlier detection are done using the POC function without epipolar constraint. This means that the proposed approach can be widely adopted in many applications using not only calibrated cameras but also uncalibrated cameras.

We also implement an outlier correction technique in our system. Since our approach can detect a large number of high-accuracy reliable corresponding points (usually known as inliers), it is reasonable to assume a basic neighborhood constraint for natural object surfaces – neighboring points on natural object surfaces generally have smooth change in disparity [8]. Therefore, the true position of an outlier will have similar disparity to those of its neighboring points. The key idea is to assume a tentative disparity for an outlier; the tentative disparity is calculated by taking median of disparities of neighboring points. This tentative disparity is updated again with sub-pixel resolution by POC-based block matching.

3. 3D Measurement System

We implement a multi-camera passive 3D measurement system based on the proposed correspondence search and outlier detection/correction techniques. In this section, the multi-camera system is presented in details, where the sections 3.1 and 3.2 describe the system architecture and the procedure of 3D measurement, respectively.

3.1 System Architecture

Figure 2 shows the proposed multi-camera 3D measurement system consisting of six calibrated cameras into three pairs of camera heads (i.e., left, front and right camera heads). Here, correspondence is taken only between every two cameras of the same camera head, where the two cameras are parallel to each other with a very narrow baseline (around 50 mm).

In general, the following two features must be considered in designing the optimal camera configuration for a 3D measurement system for dense surface reconstruction:

• The narrow-baseline camera configuration makes possible to find stereo correspondence automatically for every pixel, but a serious drawback is its low accuracy in the reconstructed 3D data when compared with wide-baseline configuration.

• The wide-baseline camera configuration makes possible to achieve higher accuracy, but automatic stereo correspondence is very difficult and is limited to a small number of edge points. This may be unacceptable in many practical applications of 3D measurement.

In our multi-camera system, we adopt narrow-baseline camera alignment, where the problem of low accuracy in 3D measurement is overcome by introducing the sub-pixel correspondence search technique. The use of phase-based image matching makes possible to achieve fully automatic high-accuracy 3D measurement with a narrow-baseline stereo vision system. This paper is the first demonstration of a passive 3D measurement system, whose reconstruction accuracy is comparable with that of practical active 3D scanners based on structured light projection.

In our system, we use simple off-the-shelf CCD cameras (JAI CVM10, 640×480 pixels, monochrome, 256 grey levels with a C-mount lens VCL-16WM), and a capture board (Coreco Imaging Technology, IC-PCI with AM-STD-RGB) for simultaneous imaging from the six cameras. Images are captured by the system in ambient light, and a volume of around 1500(W) × 1000(H) × 400(D) mm^3 is usually adopt-

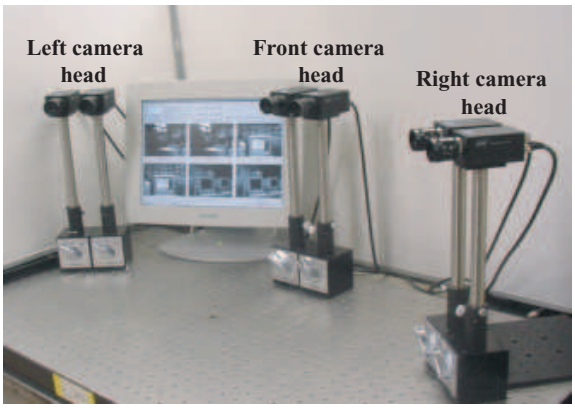

Fig. 2 ! ! Multi-camera 3D measurement system.

ed for measurement. Distance of target objects from the cameras is set around 800 $\sim$ 1200 mm, and the camera focus are adjusted to the distance.

3.2 3D Measurement Procedure and System Parameters

The measurement procedure is divided into four steps as follows:

• **Camera calibration**: Camera calibration is done in order to determine the projective matrices, which consist of the basic camera parameters, such as the relative rotation/translation of cameras with respect to the world coordinate, focal lengths, image centers, etc. Such parameters are needed to calculate the 3D coordinates of a point [1, 6]. In our system, we calibrate all the six cameras regarding the same predefined world coordinate system, so that the three camera pairs reconstruct a surface on the same world coordinate system.

• **Correspondence search**: Correspondence search is done using Procedure A. The technique makes possible to search correspondence for any given reference point, and in our system we usually determine correspondence for every 5th point with respect to the horizontal and the vertical image coordinates. For POC-based block matching, we set the parameters as: (i) the block size is $N_1 \times N_2 = 33 \times 33$ pixels (weighted by 2D Hanning window), (ii) the spectral weighting function is 2D Gaussian with $\sigma^2 = 0.5$, (iii) number of fitting points for the sub-pixel displacement estimation is 5×5, (iv) number of layers for the coarse-to-fine search is 5.

• **Outlier detection and correction**: Outlier detection and correction is done using Procedure B. The block size for POC-based block matching is same as in the correspondence search described above. The threshold α_{th} for the peak value of POC function is 0.3.

• **3D reconstruction**: The projective matrices of the cameras and the corresponding points are used to reconstruct the real-world 3D coordinates, where 4000 to 5000 points are reconstructed in our system.

Note that the tasks described above are done for each camera head independently, and thus the system outputs three sets of 3D points using the left, front and right camera heads. These data sets are mapped into the same world coordinate system using the projective matrices.

4. Experiments and Evaluation

In this section, we describe a set of experiments using simple and well characterized physical surfaces to evaluate the accuracy of the proposed 3D measurement system. In addition, a human face — a typical example of free form objects — is measured to demonstrate the system's capability of high-quality dense 3D reconstruction.

4.1 Impacts of Sub-Pixel Correspondence and Outlier Correction

We first evaluate the effects of the proposed techniques: the sub-pixel correspondence search (Procedure A) and the outlier detection and correction (Procedure B), on the quality of 3D measurement. For evaluation, we consider three different methods of 3D measurement:

- **Method I** employs a simplified version of Procedure A (where we skip the steps of [Sub-pixel estimation]), but does not employ Procedure B.
- **Method II** employs Procedure A, but does not employ Procedure B.
- **Method III** employs both Procedure A and Procedure B.

At first, we evaluate the accuracy of 3D reconstruction using two reference objects of geometrically regular shapes — a solid plane (a flat wooden board) of size 180×150 mm^2 and a solid sphere (a bowling ball) of radius 108.45 mm — both having sufficient machining accuracy. In order to evaluate measurement accuracy for the solid plane, we generate a best fitted plane for the measured points by the least-squares algorithm. Let (x_i, y_i, z_i) be the reconstructed 3D points, where $i = 1, 2, \cdots, K$. The plane fitting is to minimize the following function:

$$P(a, b, c) = \sum_{i=1}^{K} (z_i - ax_i - by_i - c)^2, \quad (1)$$

where (a, b, c) are fitting parameters. Accuracy of measurement is evaluated by the fitting error. Similar experiment is carried out using the solid sphere as a reference object. For sphere fitting, we minimize the following function:

$$S(c_1, c_2, c_3, r) = \sum_{i=1}^{K} (\sqrt{(x_i - c_1)^2 + (y_i - c_2)^2 + (z_i - c_3)^2} - r)^2, \quad (2)$$

where (c_1, c_2, c_3, r) are fitting parameters.

We use here only front camera pair for 3D measurement, where the camera baseline is 50.84 mm (estimated by the camera calibration) and the distance between the camera pair and the reference objects is around 900 mm. Table 1 compares the errors in 3D measurement by the Method I, II and III, when we use the solid plane as a reference object. Table 2 summarizes the similar experiment, when we use the solid sphere as a reference object.

Table 1 ! ! Error [mm] in 3D measurement of a plane.

	RMS error [mm]	Max. error [mm]
Method I	0.87	13.93
Method II	0.61	12.81
Method III	0.42	1.23

Table 2 ! ! Error [mm] in 3D measurement of a sphere.

	RMS error [mm]	Max. error [mm]
Method I	1.59	20.19
Method II	0.63	18.52
Method III	0.55	4.12

These results show that the proposed sub-pixel correspondence technique contributes to reducing the RMS (Root Mean Square) error significantly, and the outlier correction technique is effective for reducing both the RMS error and the maximum error. Figure 3 shows the 3D surfaces of the plane object and the sphere object reconstructed by the Methods I, II and III, respectively, which clearly visualizes the significant impacts of the proposed technique. We can observe that the Method I tends to produce stepwise error in the 3D data, and even Method II produces scattered points (i.e., outliers). The Method III, on the other hand, successfully reconstructs the smooth surfaces of the reference objects.

Table 3 summarizes the total number of reference points for which correspondence search is carried out, the number of detected outliers, the number of corrected outliers, and the resulting number of reconstructed 3D points. For both plane and sphere objects, about 8-9% of the total reference points are classified into outliers, but around 80-90% of the outliers are corrected. Thus the outlier correction technique makes possible to increase the number of reconstructed 3D points, and thus the technique may be useful

Table 3 ! ! Number of points in 3D reconstruction.

	Plane	Sphere
# of points in correspondence search	3898	4928
# of detected outliers	342	391
# of corrected outliers	309	312
# of reconstructed points	3865	4849

for many applications where dense reconstruction of 3D surfaces is necessary.

4.2 System Performance

In this section, we use the three camera heads (i.e., left, front and right camera heads) simultaneously and evaluate the overall accuracy of 3D measurement by changing the position of the reference objects. At first, the object is placed around 900 mm away from the cameras, and images are captured by all the cameras. Then, a micro-stage (with 7 μm displacement error) is used to move the object 4 times, where each time the displacement is 5 mm and images are taken at each position. Thus, we have a set of reconstructed object surfaces at 5 different positions. Let us denote the 5 different positions of the object by P1, P2, P3, P4, and P5 in order, where the distance of every movement: P1 $\rightarrow$ P2, P2 $\rightarrow$ P3, P3 $\rightarrow$ P4 or P4 $\rightarrow$ P5 is 5 mm. Table 4 and Table5 summarize the RMS fitting errors of the reconstructed object surfaces at the positions P1$\sim$P5, where RMS errors of the 3D data from the three camera heads are given in different columns as well as the overall RMS error when all the 3D data are combined in a common world coordinate system. The RMS error ranges from 0.35 mm to 0.46 mm for the plane object, and 0.53 mm to 0.70 mm for the sphere object, respectively.

We also evaluate accuracy of 3D movement estimation by calculating the displacements of the reference object (the plane or sphere) for the movements: P1 $\rightarrow$ P2, P2 $\rightarrow$ P3, P3 $\rightarrow$ P4 and P4 $\rightarrow$ P5. We compare the estimated displacement with the actual displacement of the object surfaces (5 mm each time with 7μm displacement error of the micro-stage). For the plane object, we calculate the distance between every two adjacent fitted planes, and evaluate the displacement error as shown in Table 6, where the error ranges from 0.00 mm to 0.10 mm. For the sphere object, we calculate the distance between the centers of every two adjacent fitted spheres, and evaluate the displacement error as shown in Table 7, where the error ranges from 0.01 mm to 0.11 mm.

Table 4 ! ! RMS error [mm] in 3D measurement of a plane at different positions.

Position	Left camera	Front camera	Right camera	Combined
P1	0.35	0.42	0.35	0.40
P2	0.35	0.46	0.35	0.40
P3	0.38	0.43	0.36	0.40
P4	0.36	0.43	0.37	0.40
P5	0.38	0.41	0.36	0.39

Table 5 ! ! RMS error [mm] in 3D measurement of a sphere at different positions.

Position	Left camera	Front camera	Right camera	Combined
P1	0.64	0.55	0.58	0.65
P2	0.54	0.53	0.57	0.65
P3	0.54	0.60	0.54	0.65
P4	0.67	0.58	0.58	0.64
P5	0.70	0.47	0.56	0.70

All these experiment results show that our proposed system reconstructs 3D objects with less than 1 mm error, where the distance between the object and the cameras is around 1 m. The accuracy is considered to be very high for passive 3D measurement system without using structured light projection or laser scanning.

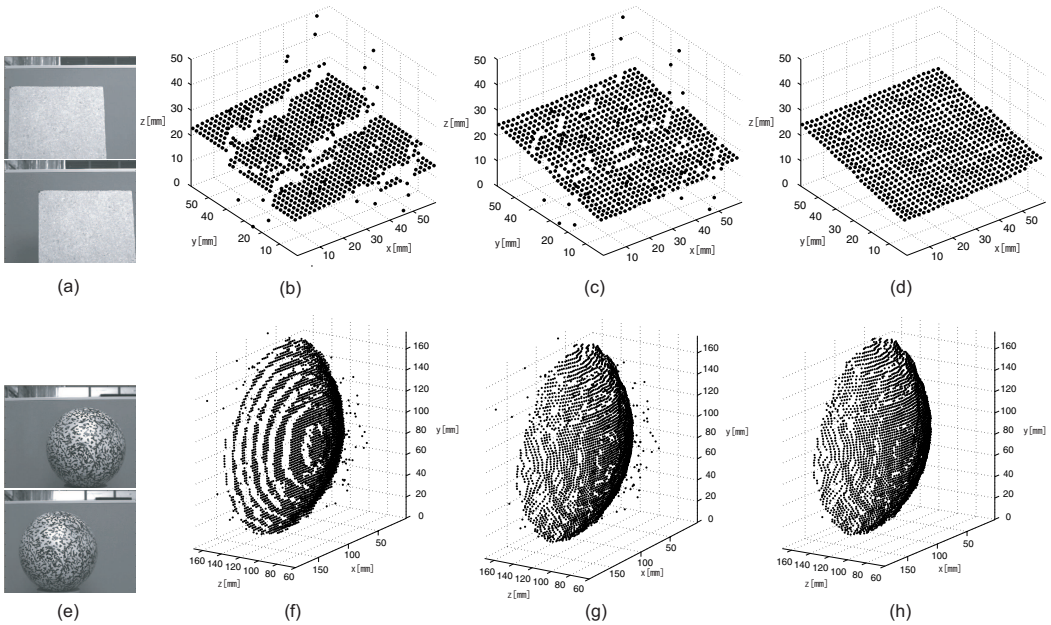

Fig. 3 ! ! Impacts of sub-pixel correspondence and outlier correction: (a) 2D images of a plane object, (b) reconstructed plane using Method I, (c) reconstructed plane using Method II, (d) reconstructed plane using Method III, (e)-(h) similar images for a sphere object; here, a portion of the reconstructed object is presented for convenience in visualization.

Table 6 ! ! Error [mm] in 3D movement estimation for a plane.

Movement	Left camera	Front camera	Right camera	Combined
P1 → P2	0.06	0.06	0.00	0.05
P2 → P3	0.08	0.08	0.02	0.05
P3 → P4	0.03	0.09	0.01	0.03
P4 → P5	0.09	0.10	0.05	0.10

Table 7 ! ! Error [mm] in 3D movement estimation for a sphere.

Movement	Left camera	Front camera	Right camera	Combined
P1 → P2	0.10	0.04	0.01	0.10
P2 → P3	0.03	0.08	0.06	0.08
P3 → P4	0.07	0.03	0.09	0.01
P4 → P5	0.11	0.10	0.03	0.02

4.3 Application to Human-Computer Interaction

Experimental analysis in section 4.1 and 4.2 shows that the proposed system performs dense 3D reconstruction with very high accuracy. In this section, at first we demonstrate 3D measurement of a typical example of free form objects — a human face. For objects with such irregular shape, we cannot evaluate the accuracy of 3D reconstruction directly, since the precise dimensions of the face are not known. Therefore, we verify the reliability of 3D reconstruction of the face by evaluating distance of every corresponding point from its epipolar line, which is computed by using the fundamental matrix obtained in camera calibration. In an ideal situation, every corresponding point should be on its epipolar line and the distance should be zero. Table 8 summarizes the RMS errors in the evaluated distance for the two reference objects and the human face. As for the Method III, the RMS error is 0.16 pixels for the plane, 0.35 pixels for the sphere and 0.27 pixels for the face. Thus, we can conclude that the reconstructed 3D face has high accuracy comparable with the reconstructed reference objects. Figure 4 displays the combined 3D

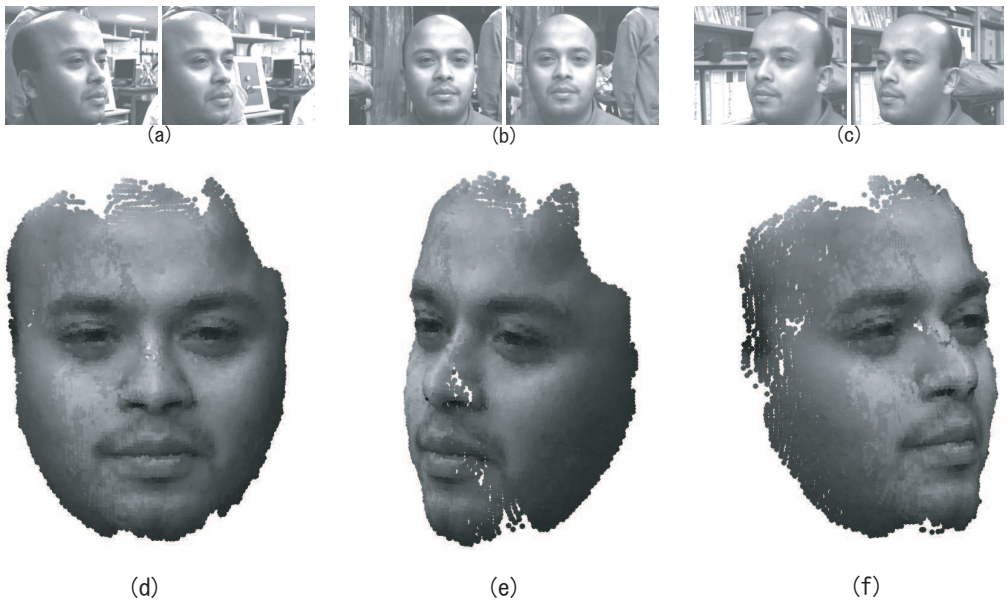

Fig. 4 ! ! Reconstructed 3D face data: (a), (b) and (c) are the 2D images captured by the left, front and right camera heads, respectively; (d), (e) and (f) are the reconstructed 3D images, where data obtained from the three camera heads are combined together.

Table 8 ! ! RMS errors [pixel] in the distances between the corresponding points and their epipolar lines.

	Plane	Sphere	Face
Method I	0.33	0.65	0.79
Method II	0.25	0.64	0.48
Method III	0.16	0.35	0.27

data of the human face from different view angles. To the best of the authors' knowledge, the quality of 3D reconstruction seems to be one of the best that is available with passive 3D measurement techniques reported to date.

The result in this paper clearly suggests a potential possibility of the proposed system to be widely used in many computer vision applications, e.g., face recognition, biometric authentication, human-computer interaction, virtual reality, etc. In this paper, we demonstrate a simple application of the system to the field of human-computer interaction. One of the basic tasks in this field is to estimate 3D motion of objects, which is useful for gesture/sign-language based instructions to computers, mobile robots etc. Here we show our system's applicability in determining 3D motion (3D rotation and translation) of a free form object through a simple experiment.

In this experiment, face images of the same person is taken at 5 different sessions with different poses, and the face object at each of the 5 sessions is reconstructed. Thus, we obtain 5 sets of reconstructed face of a person with different poses. Let these 5 sets of 3D data be $\boldsymbol{D}_i$ $(i = 1, 2, \cdots, 5)$. To deteremine the rotation $\boldsymbol{R}_{ij}$ and translation $\boldsymbol{t}_{ij}$ between two given sets of data $\boldsymbol{D}_i$ and $\boldsymbol{D}_j$ $(i = 1, 2, \cdots, 5; j = 1, 2, \cdots, 5; i \neq j)$, we use ICP (Iteratve Closest Point) algorithm [9]. Using the ICP algorithm $\boldsymbol{D}_j$ is rotated and translated to be aligned with $\boldsymbol{D}_i$. Let the aligned set of data be $\boldsymbol{D}'_j$. By estimating the rotation and translation needed for this alignment from $\boldsymbol{D}_j$ to $\boldsymbol{D}'_j$, we obtain $\boldsymbol{R}_{ij}$ and $\boldsymbol{t}_{ij}$. To evaluate the reliability of $\boldsymbol{R}_{ij}$ and $\boldsymbol{t}_{ij}$, we use the distance between $\boldsymbol{D}_i$ and $\boldsymbol{D}'_j$. In ideal case this distance should be zero, and we calculate the RMS error for the distance. Table 9 shows that the RMS errors remain below 1 mm, which clearly suggests a potential possibility of the system for estimating 3D motion of free form objects. Note that, in Table 9,

Table 9 ! ! RMS errors [mm] in the distances between the aligned datasets.

	$\boldsymbol{D}_1$	$\boldsymbol{D}_2$	$\boldsymbol{D}_3$	$\boldsymbol{D}_4$	$\boldsymbol{D}_5$
$\boldsymbol{D}_1$	-	0.65	0.66	0.72	0.84
$\boldsymbol{D}_2$	0.71	-	0.67	0.53	0.72
$\boldsymbol{D}_3$	0.58	0.52	-	0.56	0.64
$\boldsymbol{D}_4$	0.72	0.58	0.67	-	0.63
$\boldsymbol{D}_5$	0.71	0.64	0.63	0.90	-

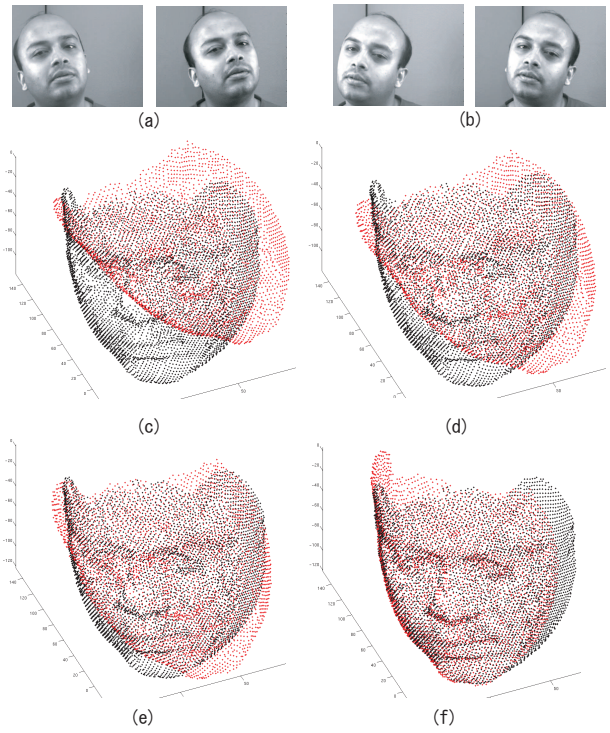

Fig. 5 ! ! 3D alignment by ICP: (a), (b) 2D images of a face with different poses taken by the front camera head; (c) two datasets at initial position, (d) after 5 iterations, (e) after 12 iterations and (f) after 57 iterations.

the value in the i-th row and the j-th column refers to the error obtained when dataset $\boldsymbol{D}_j$ is aligned regarding dataset $\boldsymbol{D}_i$. Figure 5 shows an example of data alignment regarding two different datasets.

5. Conclusions

In this paper, we have proposed a high-accuracy multi-camera passive 3D measurement system, which employs (i) a phase-based sub-pixel correspondence search technique and (ii) an outlier detection and correction technique. We have successfully implemented a passive 3D measurement system with reconstruction accuracy comparable to practical 3D scanners using structured light projection. Through some experimental evaluations, we show that the system achieves sub-mm ($\sim$ 0.5 mm) accuracy in 3D measurement, even with narrow baseline ($\sim$ 50 mm) stereo camera heads. In addition, we show that the system performs dense reconstruction of free form objects with high quality and estimates 3D motion of object with reliability, which holds a potential possibility to be applied in the field of human-computer interaction. A main goal of our current research project is to develop a passive real-time 3D capture system based on the proposed approach.

Referneces

1) O. D. Faugeras, Three Dimensional Computer Vision - A Geometric Viewpoint, MIT Press, 1993.
2) M. Okutomi and T. Kanade, "A multiple-baseline stereo," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.15, no.4, pp.353–363, April 1993.
3) C. D. Kuglin and D. C. Hines "The phase correlation image alignment method," Proc. Int. Conf. on Cybernetics and Society, pp.163–165, 1975
4) K. Takita, T. Aoki, Y. Sasaki, 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.
5) K. Takita, M.A. Muquit, T. Aoki, and T. Higuchi, "A sub-pixel correspondence search technique for computer vision applications," IEICE Transactions on Fundamentals, vol.E87-A, No.8, pp.1913–1923, August 2004.
6) G. Xu and Z. Zhang, Epipolar Geometry in Stereo, Motion and Object recognition, Kluwer Academic Publishers, 1996.
7) K. Kanatani, "3-D interpretation of optical flow by renormalization," International Journal of Computer Vision, vol.11, no. 3, pp.267–282, 1993.
8) I.J. Cox, S.L. Hingorani, S.B. Rao, and B.M. Maggs, "A maximum likelihood stereo algorithm," Computer Vision and Image Understanding, vol.63, no.3, pp.542– 567, 1996.
9) Z. Zhang, "Iterative point matching for registration of free form curves," Technical Report RR-1658, INRIA-Sophia Antipolis, Valbonne Cedex, France, 1992.