MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28–30, 2000, The University of Tokyo, Japan

13—4

Projective Differential Stereo Vision with Robust Segmentation by M-estimation

Atsushi MIYAMOTO, Tsukasa KONDO, Shun'ichi KANEKO and Satoru IGARASHI *
Faculty of Engineering, Hokkaido University

Abstract

A novel method for 3-D measurement called 'Projective Differential Stereo Vision (PDSV)' is proposed, which includes shape reconstruction of 3-D solid objects from stereo images. The method enables us to retain measurement accuracy and applicability to much irregular objects having, for example, a lot inclination with respect to the image plane. In this method, evaluators for reliability of measured parameters are proposed as well. As a result of shape reconstruction by using the self-evaluators, measurements can be more reliable and systems using them can be more effective in real world measurement. Furthermore segmentation scheme of neighbors by a robust clustering based on the M-estimators in PDSV is proposed. More precise reconstruction of discontinuous shapes in depth and/or inclination is realized by extracting any region which fits to a plane well from neighbor region.

1 Introduction

One of main problems in three-dimensional shape measurement of solid objects by using binocular stereo images is a correspondence problem between a stereo pair of images. In the stereo vision, depth can be calculated on the basis of disparity, however, it also causes a troublesome problem to determine correspondence between pixels due to additive noise, projective distortion and/or intensity change. In this problem, when a window size representing neighbors is set up smaller, higher spatial resolution can be realized, while uncertainty of correspondence must increase [3]. Thus the strategy of designing shape and size of the window must be decided taking into account this trade-off, and a coarse-to-fine strategy [4] and an adaptive strategy [5] have been proposed so far.

In order to improve accuracy and stability of measurement in depth, we propose an extended method of 'Differential stereo vision (DSV)' method [1, 2]. In the proposed method, projective distortion of the corresponding region can be estimated as projective transformation on each right and left image. We call the method 'Projective differential stereo vision (PDSV)' and it is shown that measurement accuracy is prominently improved, resulting satisfactory reconstruction even at regions of steep slopes [6, 7]. PDSV has several additional advantages as follows: it is able to measure depth and inclination simultaneously, and also involves self-evaluator functions as effective extension of the original one [2]. It is verified that much more reliable measurement or reconstruction can be executed by use of the proposed self-evaluators even through experiments for several real solid objects.

Unreliable measurement result, however, is obtained in the surface for which it is difficult to approximate locally by a plane, for instance at pixels near the edge regions where two or more plane-like surfaces meet with rather distinct inclinations. For this problem, several algorithms [5, 8] which detect discontinuous parts in terms of the distribution of disparity have been proposed, however, their measurement accuracy are not enough yet, while they need to have some knowledge about the distributions in advance.

In this paper, a segmentation scheme is proposed for estimating window shape and size around any pixels for robust reconstruction by plane-based approximation of object surfaces. In order to realize robust parameter estimation, we introduce the M-estimation [9, 10] and a weighting algorithm originally proposed for getting more precise solutions [11]. This algorithm is limitedly applied to unreliable points selected by the self-evaluators.

2 PDSV

2.1 Concept of PDSV method

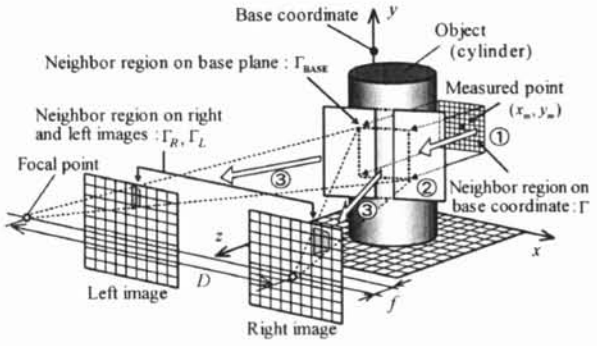
Fig.1 (a) shows the way to estimate a shape of the corresponding region in PDSV. The following assumptions are introduced:

[**Assumption 1**] A partial surface of an object is locally approximated by an inclined plane.

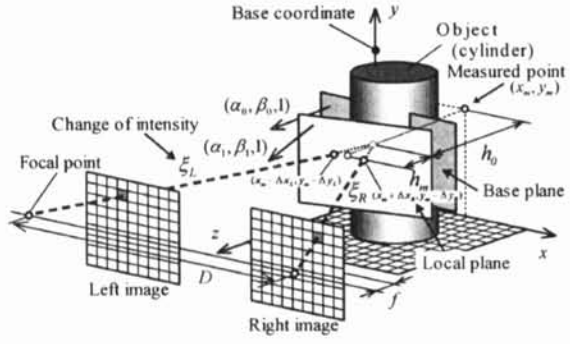
[**Assumption 2**] Difference in shape of the corresponding regions between the right and left images is represented by a projective transformation.

We call each approximating plane by *a base plane*, and initial postures of base planes are not enough to approximate the object. By projecting the neighbor region Γ_{base} on the base plane onto the right and left

* Address: Kita 13, Nishi 8, Kitaku, Sapporo 060-8628 Japan. E-mail: miyamoto@mee.coin.eng.hokudai.ac.jp



(a)Estimation of corresponding region



(b)Calculation of local plane

Fig.1 Computational modeling in PDSV.

image planes respectively, the corresponding regions Γ_R and Γ_L can be calculated respectively as follows:

$$\Gamma_i = \mathrm{D}_i(\mathrm{C}_i(\Gamma_{\mathrm{base}})) \quad (i = R, L)$$
$$\Gamma_{\mathrm{base}} = \mathrm{P}(\Gamma),$$
$$\Gamma = \{(x_{m,t}, y_{m,u}) \equiv (x_m + t\cdot\mathrm{M}, y_m + u\cdot\mathrm{M}) \mid t, u \in [-N, N]\},$$

where P is a parallel projection transform along the z-axis onto the base plane. M is an interval between two adjacent points and $2\mathrm{N}+1$ is a side length of Γ. C_R and C_L are projection transforms onto binocular image planes respectively and D_R and D_L are correcting transforms of lens distortion. The distortion in shape of neighbor regions can be estimated by these simple geometrical calculations. By utilizing the linear relationship between brightness gradient and inclination of the surface, *a local plane*, which approximates the object surface more, can be calculated (Fig.1 (b)). And an iterative algorithm is designed for finding depth by using these two kinds of virtual planes.

2.2 Formalization of PDSV

Each local plane is estimated from the corresponding base plane by an iterative algorithm. The arrangement of coordinate system is shown in Fig.1 (b). The base plane $z(x, y)$ and the local plane $h(x, y)$ are respectively expressed by

$$z(x,y) = h_0 - \alpha_0(x - x_m) - \beta_0(y - y_m), \quad (1)$$
$$h(x,y) = h_m + h_0 - \alpha_1(x - x_m) - \beta_1(y - y_m), \quad (2)$$

where x_m, y_m are the xy coordinates of the point of interest.

Assuming that binocular images are subjected to intensity change ξ_R, ξ_L through an imaging process, and denoting the sum and difference of the right and left images by $f_+ = f_R + f_L, f_- = f_R - f_L$ and partial derivative of f_+ with respect to x,y coordinates by f_x, f_y respectively, we can have the following equations [6, 7].

$$f_- - \xi f_+ - E(h_m - \alpha X - \beta Y)g = 0, \quad (3)$$
$$E = \frac{d}{H_{m0}\{H_{m0} - 2(\alpha x_m + \beta y_m)\}}, \quad (4)$$
$$g = (H_{m0} - \alpha_0 x) f_x - \alpha_0 y f_y,$$
$$H_{m0} = H_0 - h_0 - \alpha_0 x_m - \beta_0 y_m,$$
$$X = x - x_m, \quad Y = y - y_m,$$
$$\alpha = \alpha_1 - \alpha_0, \quad \beta = \beta_1 - \beta_0.$$

According to the assumption that both depth and intensity within Γ do not change abruptly, we minimize the following integral over Γ.

$$J = \iint_\Gamma \{f_- - \xi f_+ - E(h_m - \alpha X - \beta Y)g\}^2 dx dy, \quad (5)$$

where $X = x - x_m, Y = y - y_m$. Note that (5) is nonlinear with respect to variables h_m, α and β. Hence we introduce the new variables defined by (8) and can derive the following normal equation.

$$\begin{pmatrix} \langle f_+^2 \rangle & \langle f_+ \cdot g \rangle & \langle -X f_+ \cdot g \rangle & \langle -Y f_+ \cdot g \rangle \\ \langle f_+ \cdot g \rangle & \langle g^2 \rangle & \langle -X g^2 \rangle & \langle -Y g^2 \rangle \\ \langle -X f_+ \cdot g \rangle & \langle -X g^2 \rangle & \langle X^2 g^2 \rangle & \langle XY g^2 \rangle \\ \langle -Y f_+ \cdot g \rangle & \langle -Y g^2 \rangle & \langle XY g^2 \rangle & \langle Y^2 g^2 \rangle \end{pmatrix} \times \begin{pmatrix} \xi \\ E_{h_m} \\ E_\alpha \\ E_\beta \end{pmatrix} = \begin{pmatrix} \langle f_- \cdot f_+ \rangle \\ \langle f_- \cdot g \rangle \\ \langle -X f_- \cdot g \rangle \\ \langle -Y f_- \cdot g \rangle \end{pmatrix}, \quad (6)$$

$$\langle\ a\ \rangle = \iint_\Gamma a\, dx dy, \quad (7)$$

$$E_{h_m} = E h_m, \quad E_\alpha = E\alpha, \quad E_\beta = E\beta. \quad (8)$$

From (4) and (8), the coefficient E can be given by

$$E = \frac{d + 2(H_{m0} x_m + \alpha_0 d^2) E_\alpha + 2 H_{m0} y_m E_\beta}{H_{m0}^2 - (\alpha_0 d)^2}. \quad (9)$$

ξ can be directly calculated by (6), and (h_m, α_1, β_1) from (6), (8) and (9).

2.3 Reliability evaluation

From PDSV framework, self-evaluators can be derived. The evaluator J_{err}, which is based on the variance of error in the measurement, can provide estimation of the reliability. Regarding the error in equation (3) as noise in f_-, J_{err} for each measured result is calculated by using J_{res} which is a residual of J (5) in the optimal condition and the rule of propagation of error.

$$J_{\mathrm{err}}[\xi^2] = \frac{J_{11} J_{\mathrm{res}}}{\Gamma J_{\mathrm{det}}}, \quad J_{err}[h_m^2] = \frac{J_{22} J_{\mathrm{res}}}{E^2 \Gamma J_{\mathrm{det}}},$$
$$J_{\mathrm{err}}[\alpha^2] = \frac{J_{33} J_{\mathrm{res}}}{E^2 \Gamma J_{\mathrm{det}}}, \quad J_{\mathrm{err}}[\beta^2] = \frac{J_{44} J_{\mathrm{res}}}{E^2 \Gamma J_{\mathrm{det}}}, \quad (10)$$

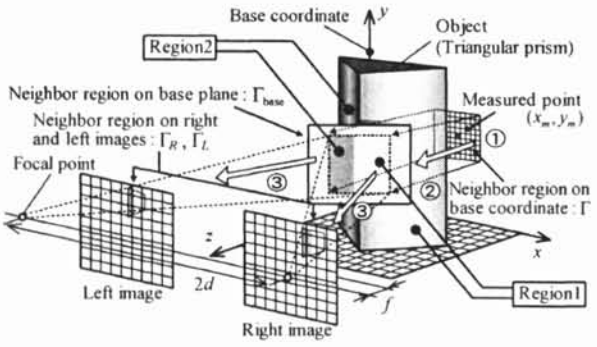


Fig.2 Division of neighbor region Γ.

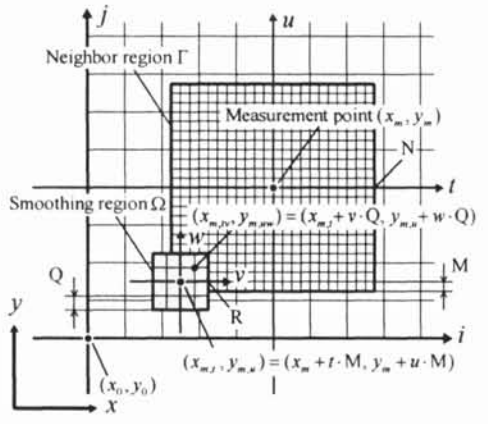


Fig.3 Neighbor region Γ and smoothing region Ω.

where J_{ii} is the determinant formed by removing both the ith row and column from J_{det}, which is the determinant of the coefficient matrix in the left hand side of the normal equation (6).

3 Introducing M-estimation for robust segmentation

3.1 Setting of the weight coefficient

In this section, we propose an extended method of PDSV abovementioned by a robust segmentation scheme based on the M-estimation, which is one of the methods for robust regression analysis [9, 10]. Fig.2 shows an example of segmentation in Γ. We divide Γ into the following two different sorts of regions: Region 1 for fitting a plane to the surface that includes measurement point and Region 2 as the remaining region.

In the case of digital images, the integration (7) is replaced by the summation, furthermore, for each adjacent point $(x_{m,t}, y_{m,u})$ in Γ, weight w_{tu} is multiplied as follows:

$$\langle\ a\ \rangle = \iint_{\Gamma} w_{tu} \cdot a \, dxdy \tag{11}$$

In order to moderate the contribution according to the value of residual, the M-estimation procedure adjusts the weights in magnitude so as to decrease the effect from exceptional gross errors. This is an effective processing for making PDSV robust against irregularity occurred in the scene or somewhat mismatching of model plane.

3.2 Iterative algorithm

The M-estimation procedure is known feasible for real problems when proper initial conditions are provided. The following iterative algorithm is applied only to extracted points which are judged by some self-evaluators as unreliable ones. In the following, $s = 1, 2, \cdots$ indicates the iteration step.

1) Obtain $v_{tu}^{(s)}$ from the residual $r_{tu}^{(s)}$ at each points $(x_{m,t}, y_{m,u})$.

$$r_{tu}^{(s)} = \sigma(f_-) - \xi\sigma(f_+) - E(h_m - \alpha X - \beta Y)g, \tag{12}$$

$$\sigma(f_\pm) = \mathrm{sign}(f_\pm)\frac{1}{(2\mathrm{R}+1)^2} \times \sum_{(x,y)\in\Omega} |f_R(x,y) \pm f_L(x,y)|,$$

$$v_{tu}^{(s)} = \begin{cases} \left\{1 - \left(\frac{r_{tu}^{(s)}}{C_B}\right)^2\right\}^2 & (|r_{tu}^{(s)}| < C_B) \\ 0 & (|r_{tu}^{(s)}| \geq C_B) \end{cases}$$

where (12) represents the residual value which is used in J through (5). C_B is an important parameter for determining a band of residual [11]. In this step, a region Ω around each point is utilized for smoothing the variations of brightness values caused by various noise as shown in Fig.3.

$$\Omega = \{(x_{m,tv}, y_{m,uw}) \equiv (x_{m,t} + v \cdot Q, y_{m,u} + w \cdot Q) \mid v, w \in [-R, R]\},$$

where Q is an appropriate interval between two points that are selected for the computation (12), and $2\mathrm{R}+1$ is the side length of Ω.

2) Local clustering is carried out for each four neighboring point set, each of which satisfies the condition $|r_{tu}^{(s)}| < C_B$. This procedure defines some connected regions R_k in Γ.

3) $w_{tu}^{(s)}$ is calculated as a revised version of $v_{tu}^{(s)}$ through the following procedure:

$$w_{tu}^{(s)} = \begin{cases} (S_k/S_{\max})v_{tu}^{(s)} & (|r_{tu}^{(s)}| < C_B) \\ 0 & (|r_{tu}^{(s)}| \geq C_B) \end{cases},$$

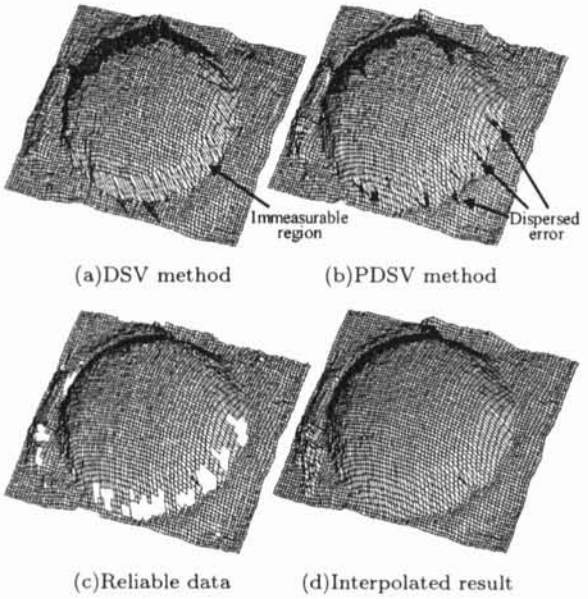
$$S_k = \sum_{(t,u)\in R_k} v_{tu}^{(s)},$$

$$S_{\max} = \max_k(S_k).$$

4) The local plane parameters $(h_m^{(s)}, \alpha_1^{(s)}, \beta_1^{(s)})$ are calculated by (6) with the modification of (11).

4 Experiments

We have set up a stereo pair two CCD cameras with the same focal distance 16mm spaced by 2d



(a)DSV method (b)PDSV method

(c)Reliable data (d)Interpolated result

Fig.4 Result of measurement (hemisphere).

=71.5 mm. Input images are converted to 512×480 pixels of eight bit brightness. A random dot pattern is projected onto the object to give contrast on the surface for stable measurement.

4.1 Comparative experiment between DSV and PDSV (hemisphere)

Fig.4 shows measurement results of a hemisphere. It is difficult to measure the regions with steep slopes by DSV as shown in (a), while PDSV makes a large part of them measurable as shown in (b). Dispersed errors, however, occurs near the boundary slope because of lower contrast in such a shaded region. Fig.4 (c) shows the result removed unreliable points judged by the self-evaluator $J_{err}[h_m^2]$. This example shows the self-evaluators can work effectively. Fig.4 (d) shows interpolated result in which 5.2% points are repaired for (c).

4.2 Comparative experiment between PDSV and PDSVM (intersecting planes)

Fig.5 shows measurement results of an object consisting of two orthogonally intersecting planes. Fig.5 (b) shows that PDSV with M-estimation (hereafter, PDSVM) can reconstruct the shape sharply even at the edge region.

Fig.5 (c) to (h) show the shapes of neighboring region Γ_R(i) and Γ_L(ii) that correspond to the measured results (C) to (H) in (b). The black area is available for reconstruction of surface and gray area is clustered away (weights converge to zero). The weight modification can robustly segment Γ into reasonable subregions. This example also shows the proposed method can effectively estimate the distortion in Γ.

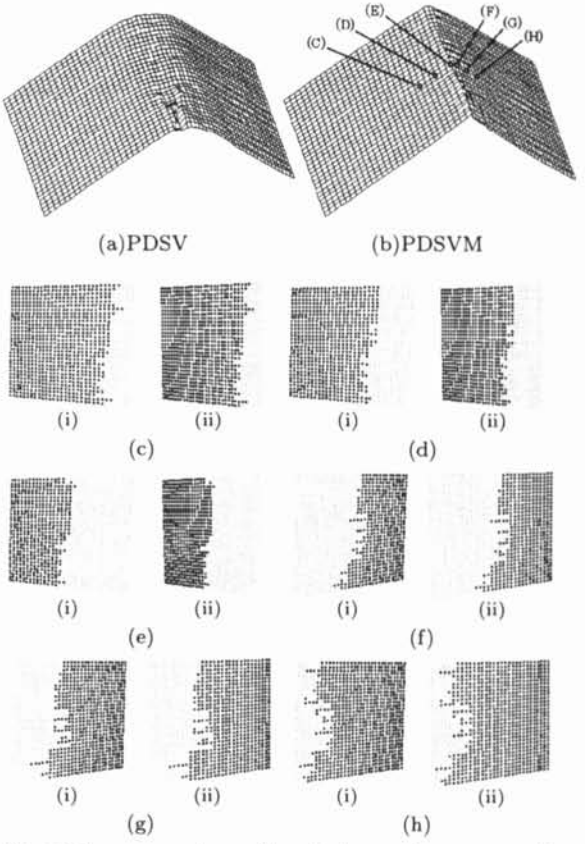


Fig.5 Measurement result and shape of corresponding region(two planes).

4.3 More examples

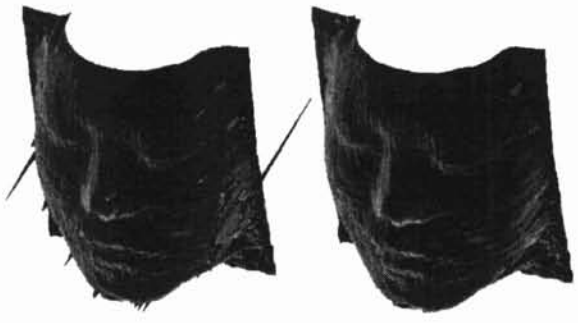
4.3.1 Mask

Fig.6 shows measurement results of a mask. Fig.6 (b) is the interpolated result of (a) judged by $J_{err}[h_m^2]$ like the example of the hemisphere. It is found that J_{err} can automatically detect ill-conditioned measurement values and data with large error in shade can be modified as shown in (b). And in spite of introducing local approximation by a plane in PDSV, curved surface can be reconstructed sufficiently.

4.3.2 Curved object

Fig.7 shows a measured object and Fig.8 (a),(b) shows measurement results of Fig.7 by PDSV and PDSVM respectively. The region with a large step (maximum difference in level is about 45mm) can be measured sharply by PDSVM.

Fig.8 (c) shows distribution of $J_{err}[\xi^2](\times 50000)$ and (d) shows the result removed unreliable points judged by $J_{err}[\xi^2]$ with threshold 1.1×10^{-4}. From (c), it can be seen the distributions of unreliable regions are clearly apart from reliable one to set up threshold easily. From (d), it is found that removing only unreliable region (step) with remaining reliable region (plane and curve) can be realized in spite of



(a)Before reconstruction (b)After reconstruction

Fig.6 Measurement result(mask).

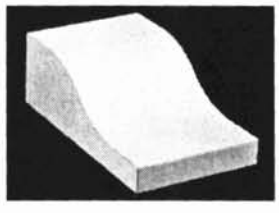


Fig.7 Curved object.

the transition of the difference level.

5 Conclusions

In this paper, PDSV and its robust version PDSVM have been proposed and the following results were obtained in summary.

- Introducing the original scheme for estimating the distortion in Γ, the accuracy of measurement can be heighly improved especially at the regions with steep slope.
- Segmentation in neighbor region realizes robust reconstruction of the shape in discontinuous surfaces in depth and/or inclination.
- It is possible to extract the unreliable result by the self-evaluators and to modify the result by the interpolation or the segmentation scheme.

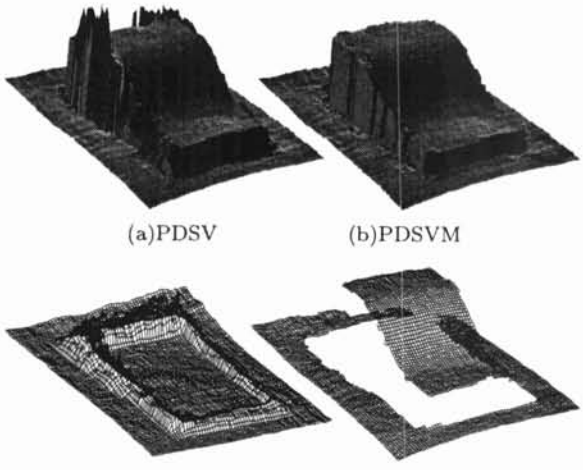
References

[1] S.Ando, "Profile Recovery System Using a Differential Identity of Binocular Stereo Images: A Parallel Algorithm and Its Experimental Evaluation", SICE Trans., Vol.23, No.4, pp.319-325, (1986-04)[in Japanese].

[2] S.Ando and T.Tabei, "Differencial Stereo Vision Systems with Dynamical 3-D Reconstruction Scheme", SICE Trans., Vol.24, No.6, pp.628-634, (1987-06)[in Japanese].

[3] S.T.Barnard and M.A.Fischler, "Stereo vision", in Encyclopedia of Artificial Intelligence. New York : John Wiley, pp.1083-1090, (1987).

[4] D.Marr and T.Poggio, "A theory of human stereo vision", Proc. Roy. Soc. London, Vol.B204, pp.301-328, (1979).



(a)PDSV (b)PDSVM

(c)Distribution of $J_{err}[\xi^2]$ (d)Group of reliable points

Fig.8 Measurement result (curve).

[5] T.Kanade and M.Okutomi, "A stereo matching algorithm with an adaptive window: Theory and experiment", IEEE Trans. PAMI, Vol.16, No.9, pp.920-932, (1994).

[6] A.Miyamoto, S.Kaneko and S.Igarashi, "3-D Measurement of shape using differential stereo vision system: Improvement of the precision by inclined base plane method", Proc. 8th ICPE, pp.528-533, (1997-08).

[7] A.Miyamoto, T.Kondo, S.Kaneko and S.Igarashi, "Proposal of PDSV Method Estimating Shape of Corresponding Region on Right and Left Images", J.JSPE, Vol.64, No.9, pp.1390-1394, (1998-09)[in Japanese].

[8] W.Hoff and N.Ahuja, "Surface from stereo : Integrating feature matching, disparity estimation, and contour detection", IEEE Trans. PAMI, Vol.11, No.2, (1989).

[9] P.J.Huber, "Robust statistics",John Wiley & Sons, (1981).

[10] P.W.Holland and R.E.Welsh, "Robust Regression using Iteratively Reweighted Least-Squares", COMMUN. STATIST. THEOR. METH, A6(9), pp.813-827(1977).

[11] A.Miyamoto, T.Kondo, S.Kaneko and S.Igarashi, "PDSV Method with Robust Division of Neighbor Region by Using M-Estimation", Proc. of PRMU Vol.99, No.217, pp.97-104, (1999-07)[in Japanese].