MULTI-VIEW STEREO CAMERA CALIBRATION USING LASER TARGETS FOR MEASUREMENT OF LONG OBJECTS

Takashi Yoshimi[1], *Takaharu Yoshimura*[2], *Ryuichi Takase*[1], *Yoshihiro Kawai*[1], *Fumiaki Tomita*[1]

[1]National Institute of Advanced Industrial Science and Technology
AIST Tsukuba Central 2, Tsukuba, Ibaraki 305-8568, Japan
E-mail: {tak–yoshimi,r–takase,y.kawai,f.tomita}@aist.go.jp
[2]Hitachi Metals Ltd.
6010 Mikajiri, Kumagaya-shi, Saitama-ken 360-0843, Japan
E-mail: {Takaharu_Yoshimura}@hitachi–metals.co.jp

ABSTRACT

A calibration method for multiple sets of stereo vision cameras is proposed. To measure the three-dimensional shape of a very long object, measuring the object at different viewpoints and registration of the data are necessary. In this study, two lasers beams generate two strings of calibration targets, which form straight lines in the world coordinate system. An evaluation function is defined to calculate the sum of the squares of the distances between each transformed target and the fitted line representing the laser beam to each target, and the distances between points appearing in the data sets of two adjacent viewpoints. The calculation process for the approximation method based on data linearity is presented. The experimental results show the effectiveness of the method.

1. INTRODUCTION

When measuring 3D objects or scenes that are too large to be taken by a single shot of a stereo camera system, 3D data measured by multiple sets of stereo camera systems should be registered to obtain all the data. In registering 3D data that are measured in different viewpoints, the calibration of the extrinsic parameters of the stereo cameras has significant influence on the results. In this paper, the intrinsic camera parameters are assumed to be previously calibrated.

The conventional method of calibrating the extrinsic camera parameters at different viewpoints is to measure the feature points appearing commonly in some camera fields of view (FOVs) and to calculate the transform matrix so that the corresponding points are identical [1, 2, 3, 4, 5]. Such a method may be less accurate when the common points are unevenly dispersed in the FOV, because the 3D position of a point far from the dispersed area is determined by extrapolation.

In measuring a long object, such as a ship or building, with high accuracy, a series of viewpoints lined along the object are needed to cover the whole measurement area.

In such an application, very small parts of adjacent FOVs overlap, and the points in the overlapped FOV cannot be observed by other cameras, which means points in the area must be extrapolated. Figure 1 illustrates the harmful effect: the registered object appears to be bent due to the calibration patterns.

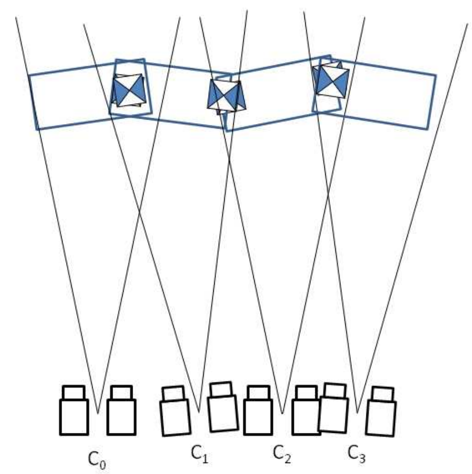


Fig. 1: Bending result in registration.

It is possible to prepare a calibration pattern of known sizes and measure it, although this approach is impractical when the object is very large. Not only is such a calibration pattern difficult to handle, but also it is hard to ensure the accuracy of the pattern itself, which changes according to environment conditions such as temperature.

In this paper, we employ calibration patterns generated with two laser diodes which are placed parallel to each other and project beams toward the longitudinal direction of the object. By measuring the 3D coordinates of the points illuminated by the lasers with stereo cameras, a sequence of points on the two straight lines of the beams are obtained. Then, the extrinsic parameters of the cameras are calculated so that an estimation function which totals the distances between the points and the lines produces the minimum value. We also present an effective approximation formula based on the linearity of the laser data.

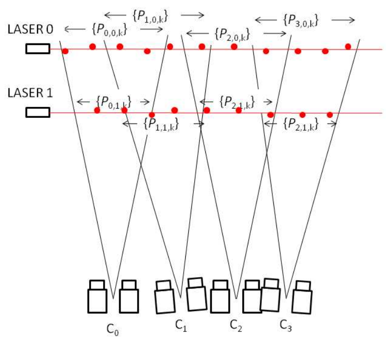


Fig. 2: Laser sampling points$\{P_{c,l,k}\}$.

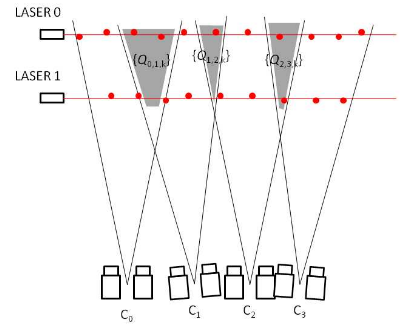


Fig. 3: Connection points $\{Q_{c,c+1,k}\}$.

As the estimation function depends only on the linearity of the laser beams, the parameters of the lines of the laser beams are not needed. The lasers can be installed approximately parallel to each other. This method can overcome the problems occurring when a large solid calibration pattern is used.

In Section 2, the basic theory is described. In Section 3, the experimental result is shown for both actual data and synthesized data. In Section 4, conclusions are given.

2. BASIC THEORY

2.1 Calibration data

First, all stereo camera systems are assumed to be previously calibrated so they can measure the 3D coordinates in the local camera coordinate system.

By projecting two laser beams in the longitudinal direction of the measurement field, the stereo camera systems can obtain two series of 3D point data. Let $\hat{P}_{c,l,k}$ be the 3D coordinates of the points of the laser beams in the local camera coordinate system and $P_{c,l,k}$ be the world coordinate, where c, l and k denote the indexes and N_c, N_l and $N_p(c,l)$ are the numbers of the stereo cameras, lasers and points, respectively. Figure 2 illustrates the alignments of $\{P_{c,l,k}\}$ in the measurement system.

Additionally, we introduce groups of feature points that lie in the overlapped FOV of adjacent two views. Let $\hat{Q}_{c,c+1,k}$ be the feature points in the camera coordinate system, and $Q_{c,c+1,k}$ be the points in the world coordinate system, where c and k denote the indexes of the stereo camera and the feature point, respectively. Although the feature points do not have to be the projected laser spots, they must be fixed when they are measured by the two stereo cameras. Figure 3 illustrates the alignments of $\{Q_{c,c+1,k}\}$ in the measurement system.

2.2 Evaluation function

Let S be the evaluation function of the whole system, which is defined as follows,

$$S = w_0 S_0 + w_1 S_1 + w_2 S_2, \tag{1}$$

where S_0 and S_1 are the functions to evaluate the linearity of laser spots $\{P_{c,0,k}\}$ and $\{P_{c,1,k}\}$, respectively, S_2 is the function to evaluate the distance between adjacent views, and w_l is the weight for S_l (l=0..2).

2.3 Recurrence calculation

Let $\boldsymbol{\phi} = \{\phi_\alpha\}$ be the vector of the unknown parameters, which are the exact rotation and transformation parameters of the camera coordinates. The number of all parameters is $(6 \times (N_c - 1))$. It is not $(6 \times N_c)$ because the set of parameters for one of the stereo cameras does not change.

To obtain the calibration parameters $\{\boldsymbol{\phi}\}$ that minimizes the function S, the equation $\partial S / \partial \boldsymbol{\phi} = 0$ is solved. This equation is solved by the Newton-Raphson method, applied in the following recurrence equation:

$$\left(\frac{\partial^2 S}{\partial \phi_\alpha \partial \phi_\beta} \right) \left(\delta \phi_\alpha \right) + \left(\frac{\partial S}{\partial \phi_\alpha} \right) = 0, \tag{2}$$

where $\{\delta\phi_\alpha\}$ is the shift vector for the parameters, and α and β are the indexes of the parameters $(1 \leq \alpha, \beta \leq 6 \times (N-1))$.

2.4 Evaluating the linearity

S_0 and S_1 are calculated as the sum of the squares of the distances between each point and the fitted line. The following value can be substituted:

$$S_l \quad = \quad \lambda_1 + \lambda_2, \tag{3}$$

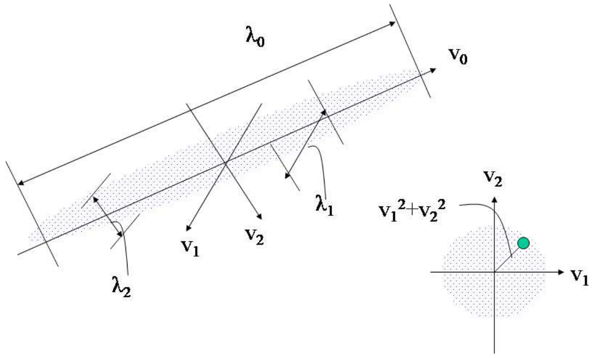


Fig. 4: Evaluating the distance from the fitted line.

where λ_1 and λ_2 are the second and the third largest eigenvalues of the covariance matrix of $\{P_{c,l,k}\}$. Because the two eigenvalues are equivalent to the variance in the eigenvector, the sum of them is equal to the sum of the squares of the distances from the center of gravity in $v_1 - v_2$ 2D space (Fig. 4), where v_1 and v_2 are the eigenvectors corresponding to λ_1 and λ_2.

Let $A = (a_{ij})$ be the covariance matrix of P. Then the equation to solve the eigenvalues is the following:

$$\lambda^3 + b_1\lambda^2 + b_2\lambda + b_3 = 0, \tag{4}$$

where

$$\begin{aligned} b_1 &= -(a_{00} + a_{11} + a_{22}) && (5)\\ b_2 &= (a_{00}a_{11} + a_{11}a_{22} + a_{22}a_{00} \\ &\quad -a_{12}a_{21} - a_{02}a_{20} - a_{01}a_{10}) && (6)\\ b_3 &= -\det(A). && (7)\end{aligned}$$

However, when the solutions of Eq.(4) are λ_k (k=0..2), Eq.(4) can be expressed by the following.

$$(\lambda - \lambda_0)(\lambda - \lambda_1)(\lambda - \lambda_2) = 0. \tag{8}$$

By using Eq.(4) and Eq.(8), b_1 through b_3 can be substituted with the eigenvalues also

$$\begin{aligned} b_1 &= -(\lambda_0 + \lambda_1 + \lambda_2) && (9)\\ b_2 &= \lambda_0\lambda_1 + \lambda_1\lambda_2 + \lambda_2\lambda_0 && (10)\\ b_3 &= -\lambda_0\lambda_1\lambda_2. && (11)\end{aligned}$$

Then, these equation can be approximated as follows:

$$\begin{aligned} b_1 &\simeq -\lambda_0 && (12)\\ b_2 &\simeq \lambda_0(\lambda_1 + \lambda_2), && (13)\end{aligned}$$

when the condition $\lambda_0 \gg \lambda_1 > \lambda_2$ is satisfied.

Thus, the next approximation is derived from Eq.(3):

$$\begin{aligned} S_l &= \frac{\lambda_0(\lambda_1 + \lambda_2)}{\lambda_0} \\ &\simeq -b_2/b_1. && (14)\end{aligned}$$

Equations (5), (6) and (14) derive a rational expression of S_l with $\{a_{ij}\}$, which causes a rational expression with (x_i), which is the 3D coordinates of $\{P\}$. Finally, we can calculate the values $\frac{\partial^2 S}{\partial\phi_\alpha\partial\phi_\beta}$ and $\frac{\partial S}{\partial\phi_\alpha}$ that appear in Eq.(2).

S_2 is calculated as follows:

$$\begin{aligned} S_2 &= \sum \left((x_{k0} - x_{k'0})^2 + (x_{k1} - x_{k'1})^2 \right.\\ &\quad \left. +(x_{k2} - x_{k'2})^2\right), && (15)\end{aligned}$$

where vector ${}^t(x_{k0}, x_{k1}, x_{k2})$ is the $Q_{c,c+1,k}$ captured by the stereo camera c and transformed to world coordinates, and vector ${}^t(x_{k'0}, x_{k'1}, x_{k'2})$ is the $Q_{c,c+1,k}$ captured by the stereo camera $c + 1$ and transformed to world coordinates.

2.5 Parameter initialization

To calculate the recursive equation Eq.(2), we need the initial values of the parameters. In this work, we apply the following steps.

1. Fit a plane to the set of both laser data $\{\hat{P}_{c,l,k}\}$ for each camera c. Let $\hat{\mathbf{u}}_c$ be the normal vector.

2. Fit a line to the set of laser data $\{\hat{P}_{c,l,k}\}$ for each camera c and laser l, and calculate $\hat{\mathbf{v}}_c$, which is the average of the two tangent vectors of the lines. Adjust $\hat{\mathbf{v}}_c$ to be perpendicular to $\hat{\mathbf{u}}_c$.

3. Calculate $\hat{\mathbf{w}}_c = \hat{\mathbf{u}}_c \times \hat{\mathbf{v}}_c$, so that $\hat{\mathbf{u}}_c\hat{\mathbf{v}}_c\hat{\mathbf{w}}_c$ forms a frame F_c.

4. Calculate the rotation matrix R_c which projects frame F_c to frame F_C, where C is the specified camera whose parameters are fixed.

5. Calculate the center of gravity (COG) $\hat{\mathbf{g}}_{c,c+1,k}$ and $\hat{\mathbf{g}}'_{c,c+1,k}$ for the connection points $\hat{Q}_{c,c+1,k}$, where $\hat{\mathbf{g}}$ is the COG at camera c and $\hat{\mathbf{g}}'_{c,c+1,k}$ is the COG at camera $c + 1$.

6. Calculate the translation vector $\mathbf{t_c}$ so that

$$R_c\hat{\mathbf{g}}_{c,c+1,k} + \mathbf{t_c} = R_{c+1}\hat{\mathbf{g}}'_{c,c+1,k} + \mathbf{t_{c+1}} \tag{16}$$

is satisfied for all c.

3. EXPERIMENT

We performed an experiment with actual calibration data and synthesized data to prove the proposed method. In this experiment, the weights w_0, w_1 and w_2 are 1.0 and the iteration stops when the following condition is satisfied:

$$\frac{S_{old} - S_{new}}{S_{new}} < \epsilon, \tag{17}$$

where ϵ is 1.0×10^{-6}.

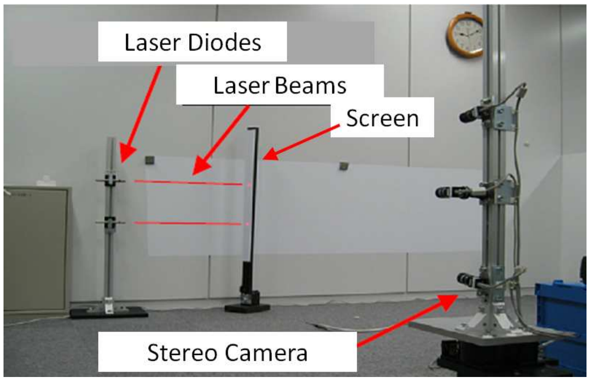


Fig. 5: Calibration system.

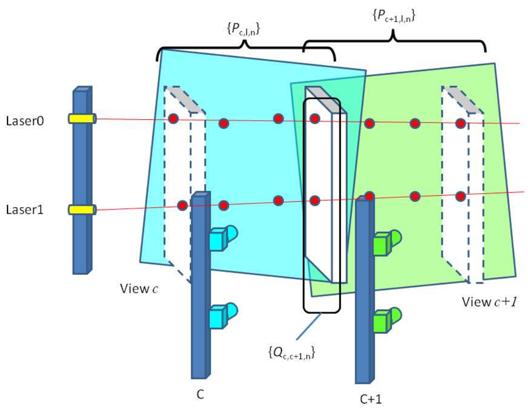
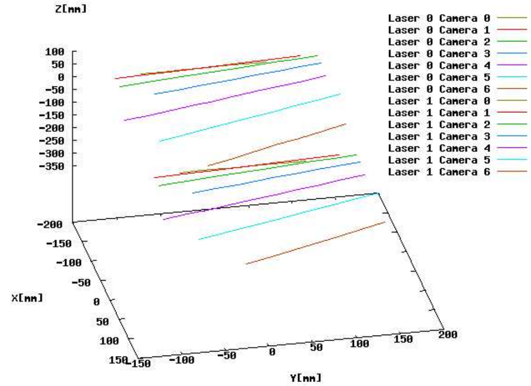
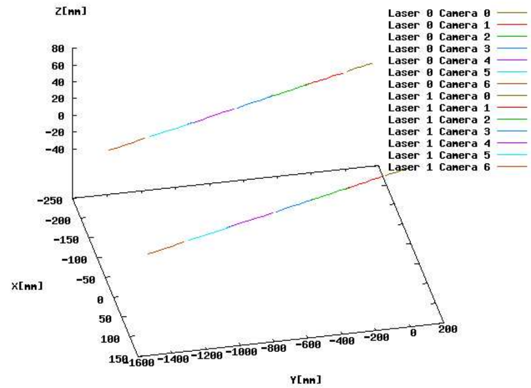


Fig. 6: Data acquisition using a movable screen.



(a) initial data



(b) final data

Fig. 7: Calibration result of registration for actual data.

3.1 Actual data

Figure 5 illustrates the actual experimental system. Two laser diodes are set on a fixed stand; a movable screen is positioned to the right of the lasers and a stereo camera is set on a rotating stage. Two cameras are mounted on a stage vertically to configure the stereo camera system. In this system, the rotating stage is employed to generate multiple viewpoints. The axis of the rotating stage is uncalibrated. The intrinsic parameters of the stereo camera were calibrated before the experiment.

The screen was moved along the laser beam, and the 3D coordinates of two laser spots were measured with stereo camera C (Fig.6). Those points were stored in $\{\hat{P}_{c,l,k}\}$. When the screen came to the overlapped FOV of camera C and $C+1$, those data were stored in $\{\hat{Q}_{c,c+1,k}\}$. The number of $\{\hat{P}\}$s must be greater than or equal to two for each laser beam, and the number of $\{\hat{Q}\}$s must be greater than or equal to one. The average measurement error tends to decrease with an increase in the number of $\{\hat{Q}\}$s. The calibration of the intrinsic parameters of the stereo camera and the measurement of the 3D points were performed based on Versatile Volumetric Vision (VVV) technology [6] developed at National Institute of Advanced Industrial Science and Technology (AIST).

The 3D coordinates of the laser spots on the screen were measured with the stereo cameras at 10 to 12 screen positions for each of six viewpoints.

Figure 7 illustrates the registered points of laser 0 and laser 1. In this figure, (a) shows the data before calibration and (b) shows the final result. Note that the scale of the Y axis is different: (a) Y=[-150,200] (b) Y=[-1600:200].

3.2 Synthesized data

To confirm the robustness of this method against measurement error, we generate synthesized data with given noises.

The data is assigned as follows.

1. Lasers are located parallel and the distance is 500 [mm].
2. The length of the laser segment in one FOV of a stereo camera is 1000 [mm].
3. The number of points (P) in one segment is 100.
4. The number of points (Q) in one overlapped area is 4.
5. The overlapped length is 100 [mm].
6. The number of sets of stereo camera is 9.

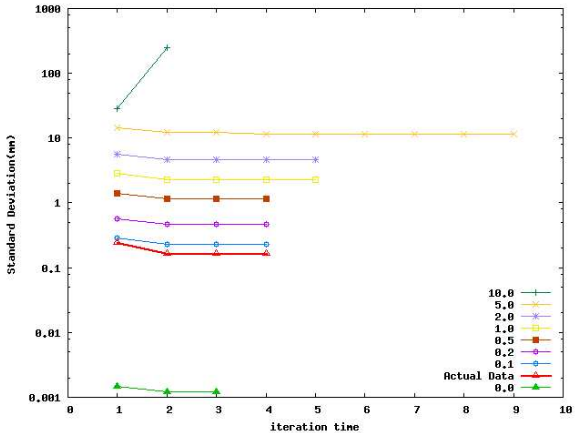


Fig. 8: Iteration result.

Table 1: Standard Deviation [mm] at convergence.

Noise Level	Approximated Value	Exact Value
0.0	8.51443×10^{-4}	9.00506×10^{-4}
0.1	0.231064079	0.231064077
0.2	0.462118253	0.462118258
0.5	1.15534162	1.15534173
1.0	2.31065027	2.31065116
2.0	4.62126000	4.62126708
5.0	11.5528605	11.5529710
10.0	NA	NA

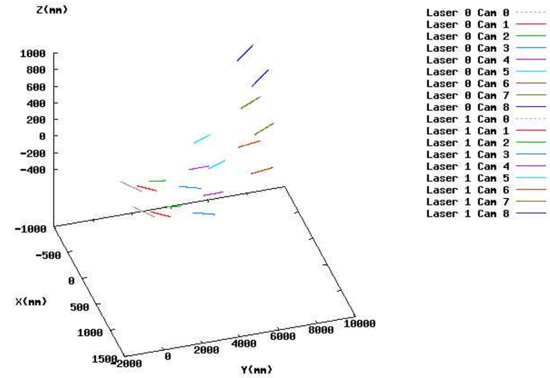
The error is given randomly, and the variance of the error is twice the value of Z of the camera coordinates. The variance simulates the situation that the Z-error is large in actual measurement. The simulated variances are 0.0, 0.1, 0.2, 0.5, 1.0, 2.0, 5.0 and 10.0 [mm].

The iteration result is shown in Fig. 8. The X and Y axes indicate the iteration time and the standard deviation of the distance of the points from the fitted line, respectively. The value is calculated as

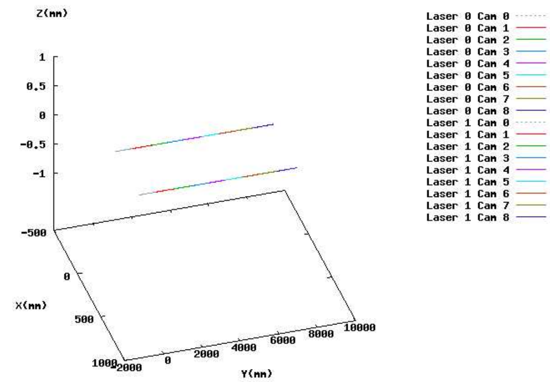
$$\sqrt{\frac{S_0 + S_1}{\sum_c \sum_l N_p(c,l)}}. \tag{18}$$

The figure shows that the proposed method works well even for data with noise, except for data with noise 10.0 [mm], in which the convergence failed. This figure also reflects the trend that the iteration time increases as the noise level increases. In this figure, the iteration result for the actual data is also plotted with a thick line. The result suggests that the noise level of the actual data is close to 0.1 [mm]. Figure 9 shows the registration result for the data with noise=0. The results for other noise levels are similar.

Table 1 shows a comparison of the evaluated values of the approximation in Eq.(14) and the exact calculation. The approximated calculation closely matches the exact value, even in measurement data with large noise. Therefore, the proposed approximation works robustly.



(a) initial data



(b) final data

Fig. 9: Calibration result of registration for synthesized data (noise=0.0).

4. CONCLUSION

In this paper, we proposed a method to calibrate a multi-view stereo camera system with laser beams. Based on the assured linearity of laser beams, this method may achieve better results than calibration methods using unarranged feature points, because it can avoid the problems caused by extrapolation. The proposed method is advantageous for measuring long objects because it can easily provide a virtual large-scale calibration pattern. The approximation method was proposed to decrease the complexity of calculation and the practicality was verified in the experiment.

The experimental result confirms that the method can measure a large object with high accuracy. By simulating the method with synthesized data of several noise levels, robustness against measurement error is also confirmed.

In the experiment described in this paper, a moving screen is used to obtain the points illuminated with the lasers. Such a procedure requires considerable time and effort, so it is not practical. Adding a smoke machine to the actual setup may help to obtain a large number of illuminated points with less time and effort.

We are planning to specify the better set of the weights

w_l used in Eq.(1) through additional experiments. We are also planning to apply the method to actual objects to verify its practical use.

5. REFERENCES

[1] Gerard Blais and Martin D. Levine. Registering multiview range data to create 3D computer objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):820–824, 1995.

[2] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge, 2003.

[3] Y. Kawai, T. Ueshiba, T. Yoshimi, and M. Oshima. Reconstruction of 3D objects by integration of multiple range data. In *Proc. ICPR92*, volume I, pages 154–157, 1992.

[4] K. Nishino and K. Ikeuchi. Robust simultaneous registration of multiple range images. In *Proc. the Fifth Asian Conference on Computer Vision*, pages 454–461, Jan 2002.

[5] K. Pulli. Multiview registration for large data sets. In *Proc. Int.Conf. on 3D Digital Imaging and Modeling*, pages 160–168, Ottawa, 1999.

[6] F. Tomita, T. Yoshimi, T. Ueshiba, Y. Kawai, Y. Sumi, T. Matsushita, K. Sugimoto, and Y.Ishiyama. R&D of versatile 3D vision system VVV. In *Proc. IEEE SCM '98*, pages 4510–4516, 1998.