Solution for Stereo Correspondences on Active Stereo Vision Robot

Masaaki Shibata
Department of Electrical Engineering and Electronics,
Faculty of Engineering, Seikei University
Kichijoji-Kitamachi, Musashino, Tokyo, JAPAN
Email: shibam@ee.seikei.ac.jp

Hajime Kawasumi
Department of Electrical Engineering and Electronics
Faculty of Engineering, Seikei University
Kichijoji-Kitamachi, Musashino, Tokyo, JAPAN

Abstract– **The paper describes a novel method for determining stereo correspondent points in the stereomate on active stereo vision robot. The appropriate stereo pairs are detected from the candidates of the combinations of feature points in the images. The substantiality of the candidate is distinguished during the robot motion, controlled with visual servo technique. The validity of the proposed method is confirmed in the physical experiments.**

I. INTRODUCTION

The paper describes a novel method for determining stereo correspondent points in the stereomate on active stereo vision robot.

Recently, the active stereo vision robot has been researched and developed for utilization in manufacturing, infrastructure or special applications [1-5]. Both advanced techniques of motion control and image processing have evolved its own performance and capability of the robot.

On the stereo vision system, the determination of corresponding points between a pair of stereo images is one of the fatal issues in general. The standard approach to relief such difficulty is the way to winnow the candidates of correspondence down. Besides, the way using the epipolar line is surely one of the effective methods for searching the candidates. Such the ways, however, are not so advantageous on the active stereo vision robot in particular, because just the epipolar line lies horizontally on the center of each image, thus the epipolar line does not give any clue.

Even if the feature points have been extracted as candidates from the stereomates, it is still even more difficult if there are several feature points similar to each other. In such the case, there is no way to determine the stereo correspondences in the static images of the stereomate unless the background pattern provides some clue.

In our approach, the issue mentioned above is resolved with using the active stereo vision robot. When several feature points have been extracted as the candidates for stereo correspondences, the appropriate stereo pairs are determined from the candidates during the robot motion.

II. STEREO CORRESPONDENCE ON ACTIVE STEREO VISION ROBOT

A. Issue on Stereo Vision System

First, the structure of the active stereo vision robot is briefed, and then the concerned issue is considered.

Fig.1. shows our active stereo vision robot (ASV robot) [5]. The robot consists of two cameras and the base, on which both cameras are settled. Fig.2. illustrates its kinematical configuration in the top view. The robot has three rotational joints at the left and right cameras and the center of the base, respectively. These vertical axes share a vertical plane and are parallel to each other at regular intervals. Both cameras and the base rotate in pan motion, and each motor attached on the joint is controlled independently. The cameras are set in same height without any tilt angle, so that the optical axes of them share a horizontal plane at any time.

Though such the structure of the robot has advantage in

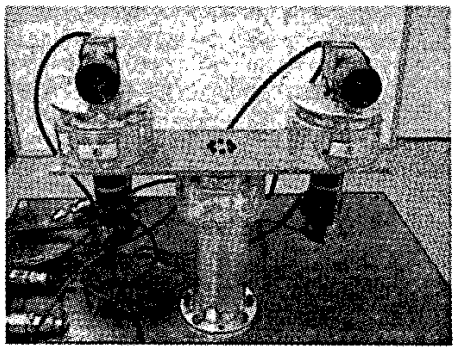


Fig.1. Active Stereo Vision Robot

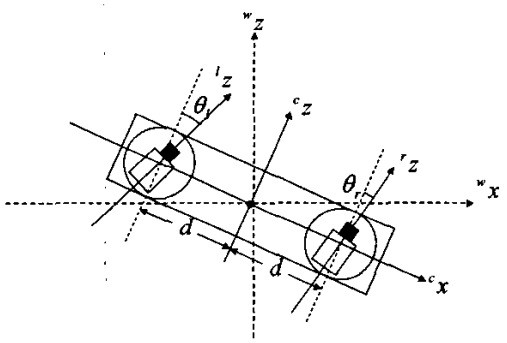


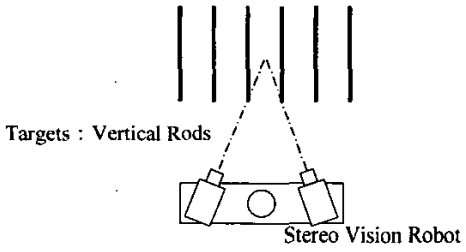
Fig.2. Kinematical Configuration of ASV Robot

estimation of the distance or the position of the target, it requires the expedient assumption, however, that the stereo matching pair should have been determined. The point of the issue is how the correct stereo pair could be distinguished.

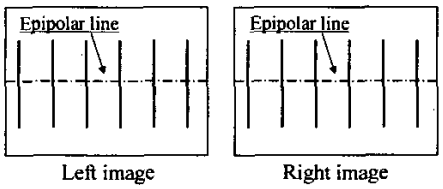
Moreover, some worse situations else may be supposed to the stereo pair detection. One is that several feature points as candidate for the stereo pair would appear on a horizontal epipolar line in each stereomate. Another is that not all the candidates can necessarily have their correspondence. The other is that no background gives any clue for matching, and some other difficulties might be apprehensive.

We are concerned with such difficulties mentioned above. Fig.3.(a) shows the target objects in the workspace. There exist some vertical thin rods in line at regular interval. They are assumed to be long enough and the same in terms of color, material, thinness and so on. On the captured images of the stereomate, only some vertical lines might exist, as shown in Fig.3.(b). The epipolar line just lies on the center of each image horizontally since the optical axes share a plane.

Sometimes human can also misconceive the stereo pair about such rods, meshes hanged in air or newly arranged tiles, so that he/she may lose a sense of distance toward such the worst targets.



(a) Target objects in the workspace



(b) Images captured by cameras

Fig.3. Disadvantageous Issue to Stereo Vision System

B. Strategy for Stereo Correspondence Distinction

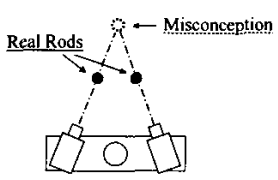
The wrong pairing of the stereo correspondences causes losing a sense of distance. The failure occurs in the configuration shown in Fig.4. Here, it is named *far side misconception* that the target position is estimated in wrong longer distance in fig.4.(a), while *near side misconception* in wrong shorter distance in fig4.(b).

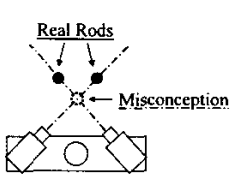
In order to distinguish whether a specified pair would be correct or not, the robot motion based on visual servo control technique is introduced. First, each camera is controlled to keep tracking the respective specified target on its optical axis. Then, the position of the target is estimated with the triangulation, no matter whether the position would be of truth. Next, the robot turns to change its whole posture. During the turning motion, it always keeps tracking the target.

The substantiality of the specified stereo pair is distinguished during the turning motion of the robot. Fig.5. illustrates the way of distinction of stereo correspondence. In case that the correct pair has been specified in fig.5.(a), the unique position of the target can be estimated properly. To the contrary, in case of the wrong pair in fig.5.(b), the position of the point estimated under misconception changes with responding to the robot motion.

For verifying the movement of the misconceived point, it is useful to formulate the estimated distance to the point. Fig.6. indicates the top view of the geometric configuration of the ASV robot. Assuming the distance l between the center of the robot base and the target, it is represented as follows.

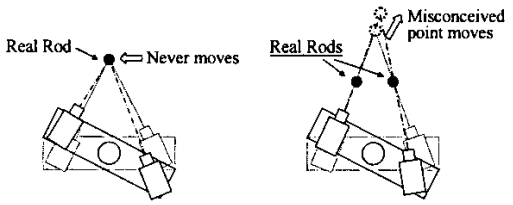
$$l = d \cdot \sqrt{\frac{4 + 2(\cos 2\phi_l + \cos 2\phi_r)}{1 - \cos(2\phi_r - 2\phi_l)}} - 1 \qquad (1)$$





(a) Far side misconception (b) Near side misconception

Fig.4. Mechanism of Losing a Sense of Distance



(a) Correct pair (b) Wrong Pair

Fig.5. Stereo Correspondence Distinction

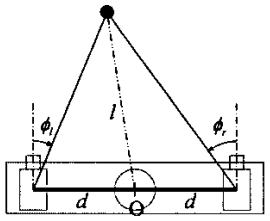


Fig.6. Geometric Configuration of ASV Robot

, where ϕ_l and ϕ_r are panned angles of the left camera and the right, respectively.

Thus, the estimated distance l is the function of the panned angles of the cameras. Therefore, the variation of l can suggest whether the supposed point would be substantial. Exactly, the ways for distinction of the substantiality with referring the variation of l or the estimated position are valid in the strictly theoretical sense.

However, even if the cameras realize the precisely controlled motion to track the target, the bottleneck of such approach is in the fact that the source of the control reference includes errors, that is, the quantization errors in the images captured by the cameras. In addition, some error or noise might come out practically.

Therefore, the validity of the mentioned approach should be considered and confirmed in the physical experiments.

III. Robot Motion Control

The ASV robot is controlled with the robust acceleration controller in order to achieve the desired motion. In this chapter, the total robot system and the controller design is explained.

A. Robot System

Fig.7. illustrates the total robot system. The cameras of the ASV robot send the captured images to the multiplexer (quad switcher), which arranges the stereomate of both images into a merged image. The merged image is sequentially forwarded to personal computer (PC) for image processing. The PC extract the feature points in the image, and then the positions data of the feature points are

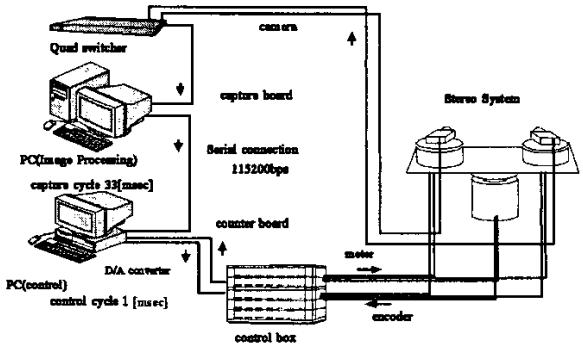


Fig.7. Robot System

sent to PC for robot control via serial connection. The latter PC generates the acceleration references to the joint actuators of the robot. The signals of the references are amplified by the servo driver in the control box, and the driver provides the currents for the joint actuators. Thus, the visual feedback system is established.

For realizing desired motion of the ASV robot, angular acceleration controller is introduced at each joint, based on the disturbance observer that suppresses the entire disturbance to joint actuator [6].

B. Motion Control Law

The robot changes its own posture for distinguishing substantiality of the stereo pair in the proposed approach. Each camera must always keep tracking the target during the base joint turning. The visual tracking control law is applied to the camera motion, while the position control law to the base motion.

First, the acceleration references to the camera joints are generated based on visual servo PD controller as follows.

$$\begin{cases} \ddot{\theta}_l^{ref} = \frac{1}{k} \cdot \left\{ Kp \cdot \left(u_l^{cmd} - u_l \right) - Kv \cdot \dot{u}_l \right\} \\ \ddot{\theta}_r^{ref} = \frac{1}{k} \cdot \left\{ Kp \cdot \left(u_r^{cmd} - u_r \right) - Kv \cdot \dot{u}_r \right\} \end{cases} \quad (2)$$

, where u_l and u_r are the target positions on the left image and the right, respectively, u_l^{cmd} and u_r^{cmd} the command positions, that are specified at the center of the respective image, Kp and Kv are the position feedback gain and the velocity, respectively. k is the scale

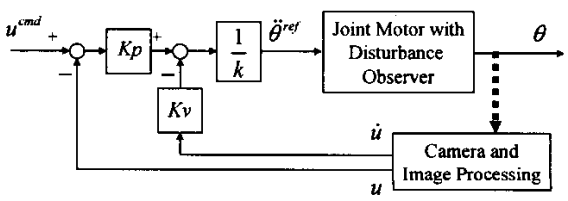


Fig.8. Image-based Camera Axis Controller

parameter between the joint angle and the position on the image, whose value is examined in the physical experiments. Fig.8. shows the block diagram of the image-based camera axis controller, based on (2).

Next, the acceleration reference to the base joint is generated based on PD controller as follows.

$$\ddot{\theta}_c^{ref} = Kp_c \cdot \left(\theta_c^{cmd} - \theta_c\right) + Kv_c \cdot \left(\dot{\theta}_c^{cmd} - \dot{\theta}_c\right) \qquad (3)$$

, where θ_c is the base joint angle, θ_c^{cmd} and $\dot{\theta}_c^{cmd}$ the position command and the velocity, Kp_c and Kv_c the position feedback gain and the velocity, respectively.

The whole posture of the robot is changed by specifying both commands of the position and the velocity to the base joint.

IV. Experiments

The physical experiments have been executed. In the first half of the experiments, it is inspected whether the choice of combination of tracking targets would relate to the achievement of the distinction of stereo correspondences. Three kinds of the combinations are examined for the inspection.

(*Exp.1*) Correct combination.

(*Exp.2*) Wrong combination that leads far side misconception about the target

(*Exp.3*) Wrong combination that leads near side misconception about the target

In the second half, it is examined whether the depth of the actual distances of the targets would influence the estimation of their distances (*Exp.4*).

A. Experimental Setup

Fig.9. shows the practical targets and ASV robot. Three rods as the targets stand up in line at regular intervals in front of the robot. Fig.10. indicates the image merged in the multiplexer. The images captured by the left camera and the right appear at the upper left and the upper right, respectively. The cross (x) represents the respective center of each image, and the plus (+) on each rod represents the feature point defined at the middle position

Table 1 Controller Parameters

Camera Axis	Position Feedback Gain	Kp	200
	Velocity Feedback Gain	Kv	80
	Disturbance Observer Gain	g	200
Base Axis	Position Feedback Gain	Kp_c	100
	Velocity Feedback Gain	Kv_c	20
	Disturbance Observer Gain	g_c	100
Scale Parameter for	Left camera	k	570.6
	Right camera	k	569.3

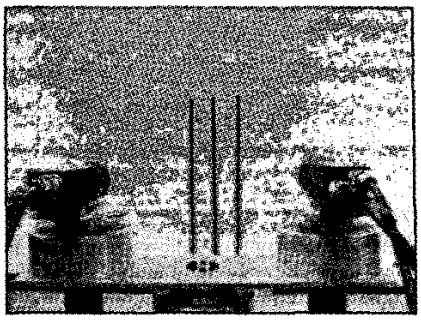


Fig.9. Experimental Setup

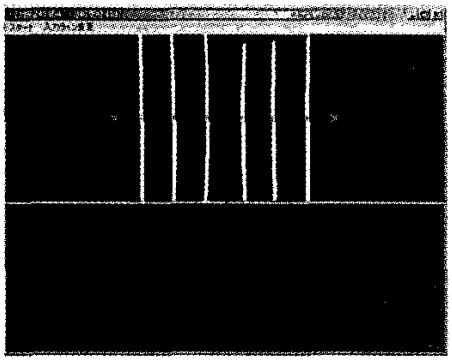


Fig.10. Image Merged in Multiplexer

of the rod. A specified target for each camera will be centered in controlling the robot motion. The values of the robot controller parameters are listed in Table 1.

B. First Experiments (Exp.1 – Exp.3)

Fig.11. illustrates the setup of the workspace for the first experiments. All distances among the robot and the rods are evenly set to L = 0.60[m], and each interval is evenly set to a = 5.0[cm].

The robot turns its own posture with the base axis rotation within the range from 0 to $\frac{\pi}{6}$ [rad] along clockwise direction for 5.0[sec] at constant angular speed. During the motion, each camera keeps staring a specified

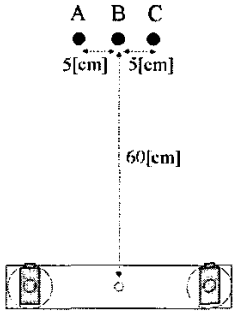


Fig.11. Workspace Setup for *Exp.1 – Exp.3*

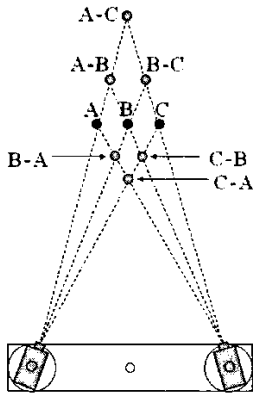


Fig.12. Real Rods and Misconceived Points

rod with visual servo control technique.

Since there are three rods, the targets tracked by both cameras would be specified among 9 combinations. Fig.12. illustrates the real rods and the misconceived points under all the combinations. The combinations are named as (Left camera target) - (Right camera target); the farthest misconception is occurred in the combination A-C, while the nearest in the combination C-A. The correct combinations are represented as A-A, B-B and C-C.

In *Exp.1* for correct combination, both cameras keep staring the central rod B, and the base axis turns. Fig.13. shows the profiles of the estimated distances of all combinations. In the figure, the estimated distances in correct combinations are just flat through all the robot motion. While the distances in incorrect combinations for far side misconception have dynamically varied, for near side misconception the distances have also varied, but not so much.

In *Exp 2* for incorrect combination, the left camera tracks rod A and the right rod B. Fig.14. shows the profiles in *Exp.2*. Though the specified combinations are wrong, the results suggest the similar one in *Exp.1*. While the distances in incorrect combinations for far side misconception have dynamically varied, for near side misconception the distances have also varied, but not so much.

In *Exp.3*, the left camera tracks rod C and the right rod A for near side misconception. Fig.15 shows the profiles in *Exp.3*, and the similar results have been obtained.

These results are summarized in brief as follows.

(1) The specified combination of tracking targets does not relate to the distance estimation; whether the combination would be correct or not.

(2) The estimated distance in correct combination keeps flat during all the robot motion.

(3) The distance in the far side misconception varies dynamically, and the distance in the near side misconception also varies, but it seems to change slightly.

C. Second Experiments (Exp.4.)

For confirming the influence by the differences of distances among the robot and the targets, the rods are replaced in slant arrangement in *Exp.4*. Fig.16. shows the workspace setup for *Exp.4*. The rods are rearranged on a line slant to the robot. Both cameras track the same target rod B during the robot motion.

Fig.17. shows the profiles of the estimated distances of all combinations. The distances in the correct combinations are appropriately estimated in flat. It is

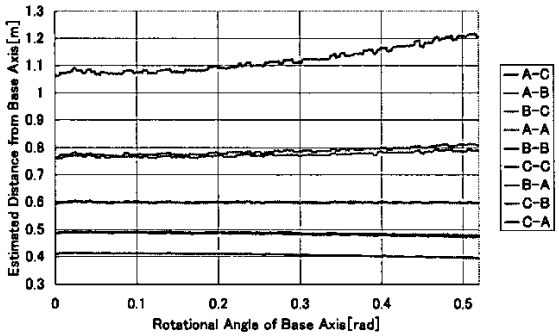


Fig.13. Estimated Distance Transitions in *Exp.1*

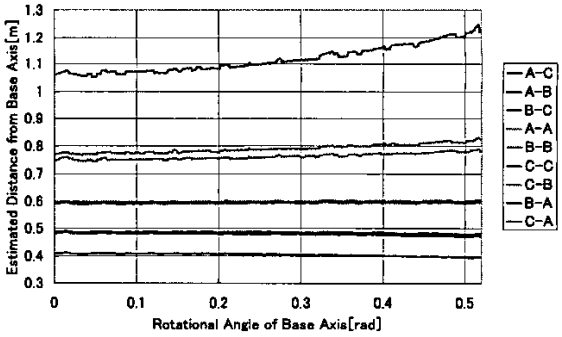


Fig.14. Estimated Distance Transitions in *Exp.2*

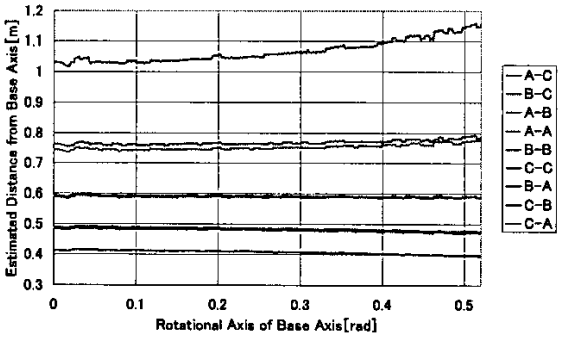


Fig.15. Estimated Distance Transitions in *Exp.3*

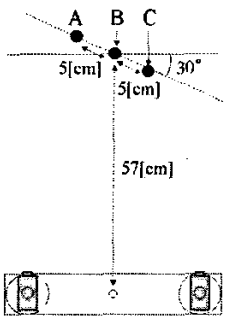


Fig.16. Workspace Setup for *Exp.4*

obvious that the differences of the actual distances among the targets should not affect the distinction.

D. Discussions

It is confirmed that the estimated distance in the correct combination keeps being flat during all the robot motion, while one in the incorrect has varied.

In the combination for near side misconception, the variation of the distance seems small. For comparing the variations among the combinations, the gradient of the transients of the distance are remarked. Table 2 shows the comparison among gradients of approximated lines of the distance transients. Among all experiments (*Exp.1 – Exp.4*), the gradients in correct combination are small, while ones in misconception are larger.

Even if the values in near side misconception are similar to that in correct combination, it is able to distinct the correct stereo correspondences. Because, about a concerned target, the distance in correct is longer than in near side misconceptions, that is namely, one in B-B is longer among the distances of B-* or *-B in near side misconceptions.

These experimental results have revealed the validity of the proposed approach. The correct stereo correspondences are found with inspecting the distance transients among the candidates.

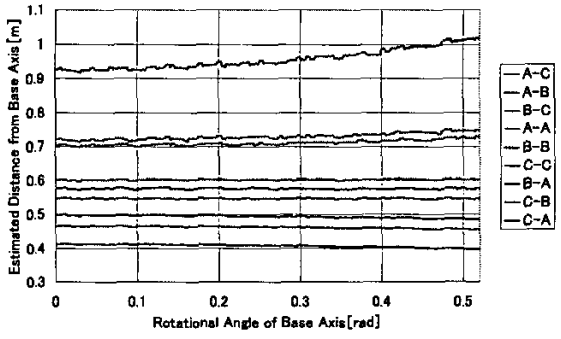


Fig.17. Estimated Distance Transitions in *Exp.4*

Table 2 Gradient of Approximated Line to Profile of Estimated Distance Transient

		Gradient of Approximated Line [cm/rad]			
		Exp.1.	*Exp.2.*	*Exp.3.*	*Exp.4.*
Correct Corresp.	A-A	0.17	-0.38	0.17	0.42
	B-B	-0.13	0.99	-0.44	0.36
	C-C	0.32	1.01	0.08	-0.29
Far Side Miscon.	A-C	27.60	32.98	24.90	20.95
	A-B	7.52	10.12	6.09	6.63
	B-C	4.64	6.02	5.00	5.48
Near Side Miscon.	C-A	-3.69	-3.02	-3.27	-3.25
	B-A	-2.98	-2.65	-2.60	-2.08
	C-B	-2.69	-2.19	-2.48	-2.20

V. CONCLUSIONS

The method for determining the stereo correspondent points on stereo images, captured in active stereo vision robot, has been proposed. In our approach, the consistency of the assumed candidates of the stereo pairs is examined and then distinguished in the robot motion

All the misconceptions have been easily distinguished by referring the transitions of the distances. Even if the tracking targets of the cameras are not of correct combinations for the stereo pair, the correct correspondences are simultaneously distinguished through a motion of the robot.

Consequently, the validity of the proposed method is confirmed in the physical experimental results.

REFERENCES

[1] S.Das, et al., "Performance Analysis of Stereo", IEEE Trans. on Patt. Anal. & Mach. Intell., Vol.17, No.12, pp.1213 - 1219, 1995

[2] Y.Kuniyoshi, et al., "Deferred imitation of human head movements by an active stereo vision head", IEEE Proc., 6th Int. Work. Rob. Human Comm. (ROMAN97), pp.88 – 93, 1997

[3] J.P.A.Barreto, et al., "Control performance Issues in a Binocular Active Vision System", IEEE/RSJ Proc. Int. Conf. Intell. Robots and Sys. (IROS98), pp.886 - 891, 1998

[4] M.Asada, et al., "Adaptive Binocular Visual Servoing for Independently Moving Target Tracking", IEEE Proc. Int. Conf. Robo. Auto. (ICRA2000), pp.2076 - 2081, 2000

[5] M.Shibata, et al., "3D Object Tracking on Active Stereo Vision Robot", IEEE Proc. Int. Works. on Advanced Motion Control (AMC2002), pp.567-572, 2002

[6] K.Ohnishi, et al., "Motion Control for Advanced Mechatronics", IEEE/ASME Trans. Mechatronics, Vol.1, No.1, pp.56-67, 1996