Wide-Area Recognition Using Hybrid Motion Stereo

Outlier Rejection for Motion Stereo Using Sequential Data

Shun Nishide, Tomohito Takubo, Kenji Inoue, Tatsuo Arai
Department of Systems Innovation
Graduate School of Engineering Science
Osaka University
Osaka, Japan
{*nishide,takubo*}@*arai-lab.sys.es.osaka-u.ac.jp*,{*inoue,arai*}@*sys.es.osaka-u.ac.jp*

Abstract— **Robots working in complex environment require accurate perception of a wide field of view. Using cameras, Hybrid Motion Stereo, which combines the computations by stereo vision and motion stereo, is capable of acquiring positional information of the whole field of view. However, the technique generates errors in computation by motion stereo when tracking feature points fails. This paper describes a method to reject the outliers to avoid erroneous recognition. The method uses computation results from several previous images, computing the spatial deviation and temporal deviation of points, and rejecting those considered to be deviated. The evaluation function is composed with a weighted average of the sequential results, each with a weight of the total number of points existing in the neighboring space, which represents the spatial deviation. Experimental results using the humanoid robot HRP-2 denote the effectivity of the method.**

Index Terms— **Mobile Robot, Environmental Recognition, Stereo Vision, Motion Stereo**

I. Introduction

Recently, researches on the robotic field have been activated for the aim of creating a human and robot interactive society[1][2]. While stationary robots are best suited for iterative tasks, they are ineligible for operating in human environments in which mobility cannot be disregarded. Therefore, mobile robots, especially those also capable of performing manipulation tasks, have been brought into researchers' attention.

One of the main issues in robotics is robot vision. Due to the characteristic of containing bounteous information, many applications for environmental recognition using cameras have been developed. Object recognition uses an image of a single camera, extracting the features of objects for identification[3]. Object tracking uses image features, such as color or brightness distribution, contained in the image of a single camera to determine the correspondence between sequential images[4]. 3D computation uses images captured at different views, calculating their disparity for recovering the 3D positions[5]. Such methods include the common stereo vision and motion stereo. The authors have focused on 3D computation using stereo vision and motion stereo as basic technologies to acquire broad 3D information.

Stereo vision and motion stereo each have their disadvantages; stereo vision is capable of computing only the overlapping views, and motion stereo requires camera movement for computation. The authors have proposed Hybrid Motion Stereo, which compensates for the disadvantages of stereo vision and motion stereo[6]. However, feature point tracking required for motion stereo computation occasionally fails, producing errors in computation by motion stereo. These points, or outliers, are required to be rejected in order to prevent false perception.

Methods for outlier rejection can be categorized into two groups: those developed by the statistical community and those by the computer vision society. Rejection techniques developed by the statistical community include M-Estimators and Least Median of Squares[7]. RANSAC[8], ALKS[9], and MUSE[10] are examples of techniques from the computer vision society. The difference between the two is the breakdown point of the proportion of the number of outliers to the total number of points. Least Median of Squares is tolerant to 50% of outliers while RANSAC, ALKS, and MUSE have been developed to exceed this limit. The methods developed by the computer vision society have the advantage over those by the statistical community for its ability to reject outliers from images with several populations of data. The features of these methods are their robustness using a set of data acquired at a single state of time. However the computation cost due to the process of fitting a model into the experimental data reduces their adaptabilities to the robotic field. Calling into account that images are constantly acquired, the authors propose a method for outlier rejection using a series of sequential data. The method described in this paper sets upon a simple computation of the weighted average reducing the computation cost as much as possible in order to retain real-time computation.

II. Hybrid Motion Stereo

Cameras are highly effective in computing 3D positions for its ability to acquire plentiful information with substantial precision. Indefectible camera modeling leads to computation errors which are minimized by Least Square Method in the process of stereo vision. Therefore, substantially high precision can be obtained in the area used for calibration. Errors in calculation increase as the calculating point recedes from the calibration area.

Computation by stereo vision can only be done in areas visible in more than two cameras (MCVA : Multiple Camera Visible Areas). Areas visible in only a single camera (SCVA :

Single Camera Visible Areas) require movement of the camera between two consecutive images for computation by motion stereo. Precision in computation by motion stereo is affected not only by the accuracy of camera modeling, but also by the precision of camera movement estimation. The authors have proposed a new technique, Hybrid Motion Stereo, which combines the computations in MCVA and SCVA to recover all the 3D positions of the visible points[6]. The proposed technique computes the camera movement in MCVA which can then be used to compute the 3D positions of points in SCVA. The concept of Hybrid Motion Stereo is shown in Fig. 1.

A. *Estimation of Camera Motion*

Using more than two cameras from a different point of view, the 3D positions of every point visible in more than two cameras can be computed by stereo vision, given the correspondence of the points in the image planes. Considering that the cameras are installed into the mobile robot, the camera coordinate systems fixed to the cameras move according to the robot's motions. The 3D coordinate calculated by stereo vision is based on the camera coordinate system. Therefore, the same immobile points in two consecutive images generate a slight displacement in the computed 3D positions.

Movements of the robot engender identical virtual motions of immobile points, which are the direct opposite of the camera motions. Let θ, ϕ, ψ be the rotational movements of the camera around the X, Y, and Z axes of the camera coordinate respectively. Let t_X, t_Y, t_Z be the translational movements of the camera along the X, Y, and Z axes of the camera coordinate respectively. Here, we will calculate the camera movement in the following order.

1) θ rotation around the X axis.
2) ϕ rotation around the Y axis.
3) ψ rotation around the Z axis.
4) Translational movements, t_X, t_Y, t_Z.

The movements of immobile points can then be denoted as follows:

1) $-\theta$ rotation around the X axis.
2) $-\phi$ rotation around the Y axis.
3) $-\psi$ rotation around the Z axis.
4) Translational movements, $-t_X$, $-t_Y$, $-t_Z$.

Using these variables, the virtual movements of immobile points between two consecutive images derive the following equation:

$$\begin{bmatrix} & & & -t_X \\ & \boldsymbol{R} & & -t_Y \\ & & & -t_Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix} = \begin{bmatrix} X_i^{'} \\ Y_i^{'} \\ Z_i^{'} \\ 1 \end{bmatrix}, \quad (1)$$

where,

$$\boldsymbol{R} = \begin{bmatrix} c_\phi c_\psi & s_\theta s_\phi c_\psi + c_\theta s_\psi & s_\theta s_\psi - c_\theta s_\phi c_\psi \\ -c_\phi s_\psi & -s_\theta s_\phi s_\psi + c_\theta c_\psi & s_\theta c_\psi + c_\theta s_\phi s_\psi \\ s_\phi & -s_\theta c_\phi & c_\theta c_\phi \end{bmatrix}. \quad (2)$$

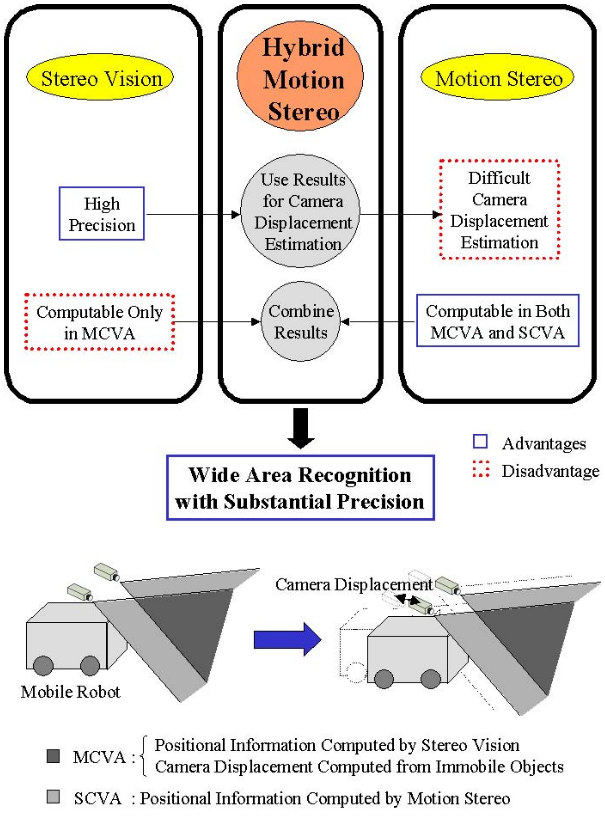

Stereo Vision
Hybrid Motion Stereo
Motion Stereo
High Precision
Use Results for Camera Displacement Estimation
Difficult Camera Displacement Estimation
Computable Only in MCVA
Combine Results
Computable in Both MCVA and SCVA
Wide Area Recognition with Substantial Precision
Advantages
Disadvantage
Camera Displacement
Mobile Robot
MCVA : Positional Information Computed by Stereo Vision
Camera Displacement Computed from Immobile Objects
SCVA : Positional Information Computed by Motion Stereo

Fig. 1. Concept of Hybrid Motion Stereo

$\boldsymbol{R}$ represents the rotation matrix of the camera coordinates between the images. s_α and c_α represents $\sin\alpha$ and $\cos\alpha$ respectively ($\alpha = \theta, \phi, \psi$). Vectors $\begin{bmatrix} X_i & Y_i & Z_i & 1 \end{bmatrix}^T$ and $\begin{bmatrix} X_i^{'} & Y_i^{'} & Z_i^{'} & 1 \end{bmatrix}^T$ represent the 3D positions of the ith immobile point of the primary and secondary images respectively.

In this paper, we will assume that the rotation matrix $\boldsymbol{R}$ in (1) equals to the identity matrix $\boldsymbol{I}$, that is, there are no rotational movements of the cameras. This results in

$$\boldsymbol{t} = \boldsymbol{x}_i - \boldsymbol{x}_i^{'}, \quad (3)$$

where $\boldsymbol{t} = \begin{bmatrix} t_X & t_Y & t_Z \end{bmatrix}^T$ represents the translational camera motions, $\boldsymbol{x}_i = \begin{bmatrix} X_i & Y_i & Z_i \end{bmatrix}^T$ and $\boldsymbol{x}_i^{'} = \begin{bmatrix} X_i^{'} & Y_i^{'} & Z_i^{'} \end{bmatrix}^T$ represents the computed 3D positions of the ith immobile point by stereo vision in the primary and secondary images respectively. Equation (3) estimates the camera movement using a single immobile point. Assuming that there are no rotational movements, the median of the movements of computed points provides a good estimate presuming the environment to be constructed with a majority of immobile objects.

B. *Outlier Rejection*

Motion stereo computation requires tracking between two consecutive images. The precision of motion stereo is directly affected by the accuracy of tracking, thus producing miscalculated points, or outliers, when tracking fails. Although techniques developed by the CV community, such as RANSAC or MUSE, are capable of robustly rejecting outliers, the computation cost required for fitting models implies it unsuitable in robotics using current computer technology. The authors propose a method using not only the results at a single moment but also previous results to detect and reject the outliers. The concept of the rejection method is illustrated in Fig. 2.

The method is relatively similar to ROR[11] which rotates the image to acquire a virtual image from a different camera pose to evaluate the accuracy of stereo matching. Since robotics call for computation as quick as possible, the authors have used the results from previously captured images for evaluation. The method creates an evaluation function using the acquired results instead of evaluating the matching process. The evaluation function can be created with a simple calculation, thus unaffecting the computation time required for 3D position computation.

Since the camera coordinate system fixed to the camera moves along with the robot, the coordinates of the prior results should be adjusted to the current coordinate system. Let $\boldsymbol{T}_j^{j+1}$ be the coordinate transformation matrix between the jth and $(j+1)$th images. The coordinates of a computed point in the ith image, $\boldsymbol{x}_i$, can be converted to the coordinates in the nth image by the following equation:

$$\boldsymbol{x}_n = \boldsymbol{T}_{n-1}^{n}\boldsymbol{T}_{n-2}^{n-1}\cdots\boldsymbol{T}_{i}^{i+1}\boldsymbol{x}_i. \tag{4}$$

Determination of outliers can be performed by using the two following properties. Spatial deviation, which segregates the outlier from the distribution of points, is a measure to determine outliers from a set of points at a single moment. Temporal deviation, which segregates the outlier from the computed sequential set, is a measure to determine outliers from a sequential set of images computing the same point. The authors have mixed the two measures to specify an evaluation function to reject the outliers.

By using prior results, outliers could be rejected more stably than using just the results of a single moment. However, the prior results also contain outliers which should be considered in the process of creating the evaluation function. The effects to the evaluation function of outliers in prior results should be minimized. The total number of computed points existing in the neighboring space around the evaluating point serves as the measure for spatial deviation, since no objects possess only a single feature point and the area around the outlier can be considered to be nondense. The authors have used weighted average to evaluate the temporal deviation.

Assume that the robot has acquired a time-series result computed at times 1, 2, $\cdots$, n. Each result contains the computation results by motion stereo for all the computed points. Let $\boldsymbol{x}_{i1}$,$\boldsymbol{x}_{i2}$, $\cdots$, $\boldsymbol{x}_{in}$ be the set of computed results for the ith computed point, which have been converted to the nth camera coordinate system by (4). To evaluate whether $\boldsymbol{x}_{in}$ is an outlier or not, we will compute the weighted average, $\boldsymbol{\mu}_{in}$, using the following equation:

$$\boldsymbol{\mu}_{in} = \frac{\sum_{j=1}^{n}\omega_j\boldsymbol{x}_{ij}}{\sum_{j=1}^{n}\omega_j}, \tag{5}$$

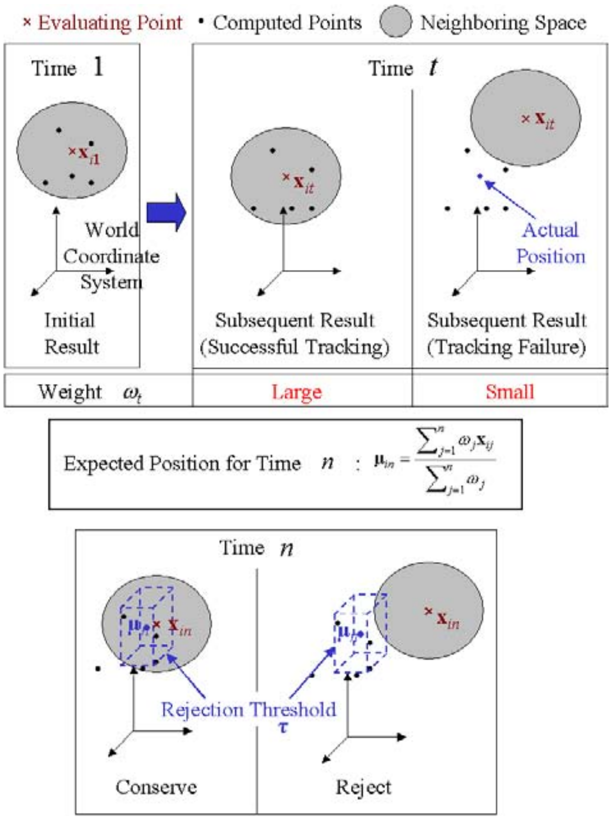


Fig. 2. Concept of the Rejection Method

where ω_j is the weight for $\boldsymbol{x}_{ij}$. The weight represents the reliability of the computation result, $\boldsymbol{x}_{ij}$, which is determined by spatial deviation. Defining the neighboring space around $\boldsymbol{x}_{ij}$ as $\boldsymbol{S}$, ω_j is calculated by the following equation:

$$\omega_j = \sum_k \delta_k(\boldsymbol{x}_{kj}), \tag{6}$$

where

$$\delta_k(\boldsymbol{x}_{kj}) = \begin{cases} 1 & if \quad \boldsymbol{x}_{kj} \in \boldsymbol{S} \\ 0 & if \quad \boldsymbol{x}_{kj} \notin \boldsymbol{S} \end{cases}. \tag{7}$$

The weighted average, $\boldsymbol{\mu}_{in}$, calculated in (5) represents the expected position at which point $\boldsymbol{x}_{in}$ should be located. Large deviation from the expected position denotes the computed point $\boldsymbol{x}_{in}$ to be an outlier. Therefore, setting a threshold τ for rejection, outliers can be specified as those satisfying the following inequality:

$$f(\boldsymbol{x}_{in} - \boldsymbol{\mu}_{in}) > \tau. \tag{8}$$

f represents a function to evaluate the deviation rate in which the 3D distance between $\boldsymbol{x}_{in}$ and $\boldsymbol{\mu}_{in}$ cannot be used directly due to the difference of distributions in computation errors along the three axes, i.e. the components of the computed results along the perspective axis possess larger errors compared to those along the lateral and longitudinal axes.

III. Experiments Using the Humanoid Robot HRP-2

To evaluate the effectivity of the proposed technique, the authors have implemented the technique into the humanoid robot, HRP-2 (Fig. 3). The vision system of HRP-2 consists of three cameras equipped on the head. We will name each camera as stated in Fig. 3. The camera coordinate system Σ_C is set as stated: the X axis faces in the lateral direction to the right of the robot, the Y axis faces in the longitudinal direction downwards, and the Z axis faces in the perspective direction which the robot faces. The origin of the camera coordinate system is set at the neck joint.

A. Image Processing

Hybrid Motion Stereo requires image processing for stereo vision and motion stereo computation. Assuming that the cameras are calibrated, stereo vision requires the correspondence of the feature points between the images, while motion stereo requires the correspondence between sequential images. Since feature points of objects are often used for acquisition of 3D positions and orientations[12] the authors have implemented Hybrid Motion Stereo based on the computation of feature points. Feature points can be extracted from the image by determining those with big eigenvalues[13]. Methods for feature point tracking have been proposed by Lucas Kanade[14][15], which is capable of performing real-time tracking with substantial stability. The authors have implemented the functions in OpenCV[16], provided by Intel Corporation, for extracting and tracking feature points.

B. Configuration of the Cameras

The cameras of the robot are set so that a large area computable by stereo vision, manipulatable with both hands, can be obtained. The optic axes of each camera face inward, redounding high precision in manipulatable areas using stereo vision, but also creating large areas visible in only a single camera at distant locations. Fig. 4(a) shows the planer visible area 1500[mm] away from the robot. The light gray area and dark gray area each represent SCVA and MCVA. Defining each area as S and M, the area computable by stereo vision is limited to M, while Hybrid Motion Stereo is capable of computing $(M + S)$. Therefore, comparing Hybrid Motion Stereo and stereo vision, the rate of increase in computable area can be defined as $(M + S)/M$.

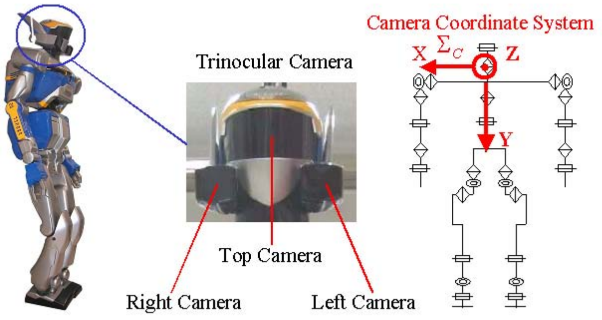
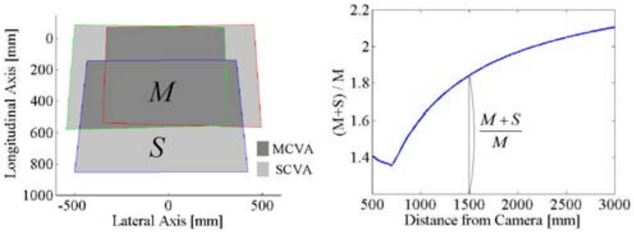


Fig. 3. Vision System of the Humanoid Robot HRP-2



(a) 1500 mm

(b) Proportion of MCVA and the Whole Field of View

Fig. 4. MCVA and SCVA

Since the configuration of the cameras form large SCVA in distant locations, the proportion, $(M + S)/M$ increases along with the distance from the cameras. The graph on Fig. 4(b) represents the increase rate of the proportion relative to the distance from the camera. As can be seen, Hybrid Motion Stereo is capable of computing more than 2 times as much area compared to stereo vision in distant areas.

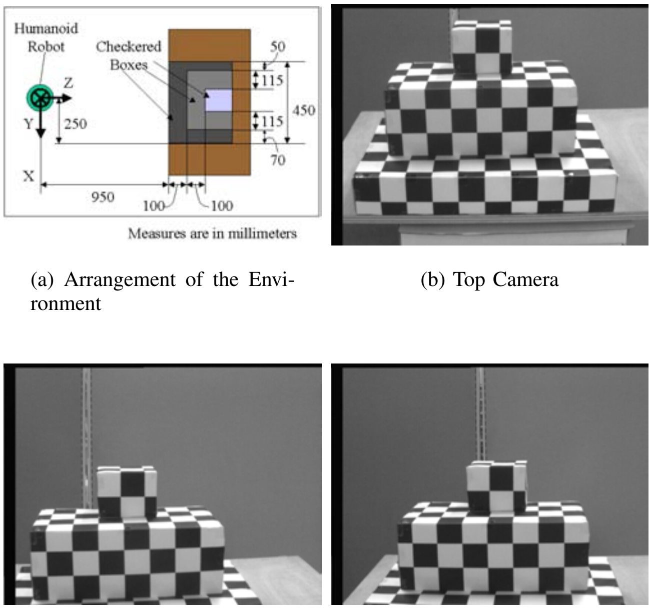
C. Experimental Conditions

The authors have constructed the experimental environment with three checkered boxes, each of a different size, stacked on one another. The setup and views of each camera are stated in Fig. 5. For simplicity, the movement of the cameras are created by rendering the robot to walk on a spot. The walk generates lateral cyclic motions induced by the movement of the center of gravity for stability retainment[17].

IV. Experimental Results

Feature points in SCVA can be computed only by motion stereo, while those in MCVA can be computed by both motion stereo and stereo vision. To evaluate the results of motion stereo, the authors have computed the 3D positions of feature points in MCVA and SCVA by motion stereo. Camera movements are estimated by using the median of the movements of feature points for each component, X, Y and Z. The total computation time for three images, excluding the image acquisition time, is approximately 0.1[s] for each frame. Outlier rejection is conducted using the results of 5 sequential images.

The computation results from overhead view are shown in Fig. 6. The results are shaded according to their height from the ground, white being the lowest and black being the highest. Feature points in the red oval in Fig. 6(a) represent the outliers which are required to be rejected. The neighboring area to compute the weight, or spatial deviation stated as $\boldsymbol{S}$ in (7), is set as a cuboid centering the evaluating feature point, with sides 10[cm] along the X and Y axes and 20[cm] along the Z axis. The side along the Z axis is defined larger than the others since the distribution of the points are scattered wider along the Z axis[18]. Since the Z axis, which faces the perspective



(a) Arrangement of the Environment

(b) Top Camera

(c) Left Camera

(d) Right Camera

Fig. 5. Setup of Environment and Views of Each Camera

direction to the camera, possesses the largest computation error, the rejection threshold, stated as τ in (8), has been defined heuristically as 50[mm] for the absolute difference of the Z components, stated as the function f in (8), between the computation result and the weighted average. As can be seen from Fig. 6, 14 out of the 15 outliers have been rejected, quoting the effectivity of the method to reject outliers both near and far.

Fig. 7 presents the perspective view of the results with points shaded according to their distance from the cameras, black being the nearest and white being the farthest. The results reconstruct quite accurate shapes of the three boxes. The distributions of the points are closer to the camera at lower places, i.e. those possessing a high value of Y components are shaded with darker colors, and tend to recede from the camera as the points get higher. The feature points in the red shaded area in Fig. 7 indicates the points incomputable by stereo vision. Nearly 30 percent of the whole points exist in the shaded area, quoting a 42 percent increase of points compared to pure stereo vision. Fig. 7 also remarks that the rejecting algorithm is capable of retaining the necessary results while rejecting the outliers. However, compared to the map of the constructed environment in Fig. 5(a), the distribution of the computed points is shifted 7 to 8 centimeters towards the robot in the perspective direction. It also possesses a few centimeters of deviation in the longitudinal direction, upwards. These errors are induced from camera movement estimation errors, which increase as the points used for camera movement estimation recede from the calibration area. Future works include a more stable method for estimating the motions of the cameras.

V. General Discussion

In this paper, a rejection method for Hybrid Motion Stereo, which produces outliers due to mistracking between consecutive images, has been proposed. Weighted average using spatial deviation and temporal deviation for creating the evaluation function proved to be effective in rejecting these outliers. The rejection threshold defined heuristically extracts the best results for the distribution of points in the experiment. Increasing the threshold results in a decrease of computed points, while decrease of the threshold results in insufficient rejection of outliers. Experimental results have shown that 14 out of 15 outliers have been rejected using the method. The method has been proved to retain the reliable results while rejecting outliers.

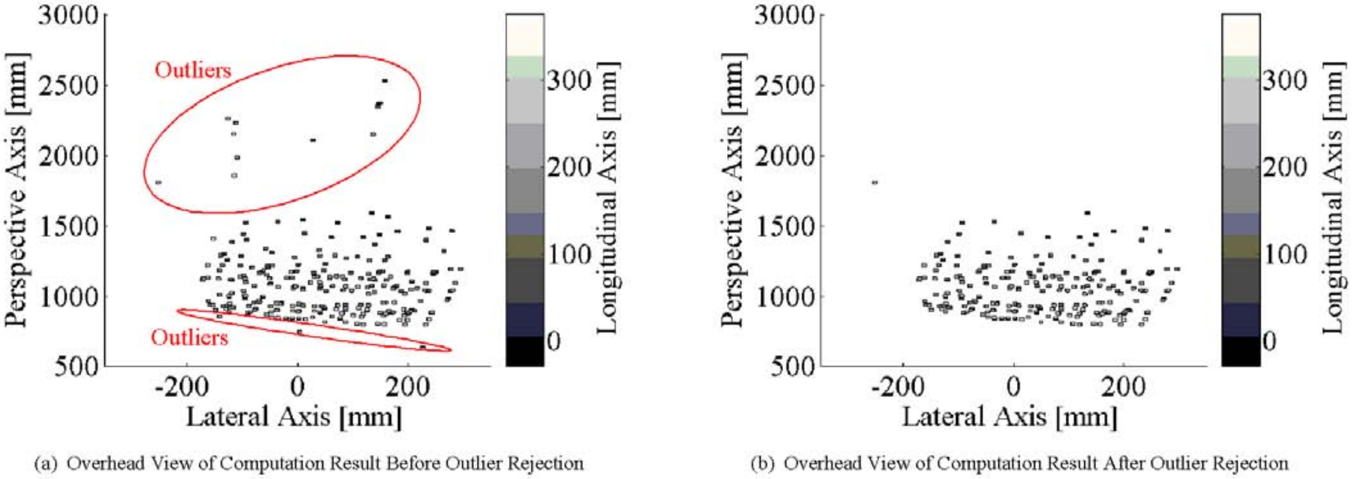
Although the method is capable of rejecting outliers, it is vulnerable to errors in estimated camera movements. Since Hybrid Motion Stereo uses consecutive images acquired between a short span, subtle errors in camera estimation induce large errors in motion stereo computation. Miscalculation in camera movements lead to the deviation of the whole distribution of points, which cannot be considered as outliers since the spatial density of the points cannot be distinguished from those computed when the camera movements have been accurately estimated. Errors in computation by stereo vision increase as the computing point recedes from the calibration area. Therefore, estimation of camera movement using stereo vision turns out to be quite unreliable when the points used are distant from the camera.

To minimize the errors in estimation of camera movement, an accurate calibration of the cameras or an alternative measure for estimation is required. Accurate calibration for off-the-shelf cameras implanted into mobile robots is an arduous task since the correspondence between the camera coordinate system and the robot coordinate system is difficult to obtain. As an alternative measure, odometry for wheeled robots or direct kinematics for biped robots can be used to evaluate the camera movement computed by stereo vision. Specifically, direct kinematics for biped robots possess large errors in camera movement estimation only in the stepping phase of the robot. Defining a proportion rate for camera movement estimation by stereo vision as α, the camera motions $\boldsymbol{\delta}$ could be estimated quite accurately by,

$$\boldsymbol{\delta} = \alpha \boldsymbol{C} + (1-\alpha)\boldsymbol{K} \qquad (0 \leq \alpha \leq 1), \tag{9}$$

when the most appropriate has been selected for the variable α, where $\boldsymbol{C}$ and $\boldsymbol{K}$ represent the camera movements estimated by stereo vision and direct kinematics respectively.

These future works for improving the precision of motion stereo is the first step for the completion of Hybrid Motion Stereo. Latter steps include consideration of rotational camera motions, creation of robot behavior from the constructed 3D map, and extracting moving objects in the environment. By overcoming these issues, we believe that Hybrid Motion Stereo is capable of applying to various practical situations.



(a) Overhead View of Computation Result Before Outlier Rejection

(b) Overhead View of Computation Result After Outlier Rejection

Fig. 6. Overhead View of Computation Result up to 3000[mm]

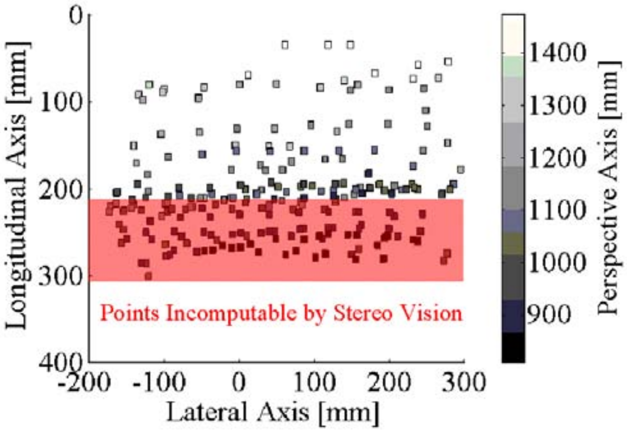


Fig. 7. Perspective View of Computation Result after Outlier Rejection

ACKNOWLEDGMENT

The authors would like to thank Prof. Yasushi Mae, at Fukui University, Japan, for providing valuable advices in the fruition of the proposed technique. This work was supported by MEXT under Grant-in-Aid for Creative Scientific Research (Project No. 13GS0018).

REFERENCES

[1] Y. Mae, N. Sasao, Y. Sakaguchi, K. Inoue, T. Arai, "Head Detection and Tracking for Monitoring Human Behaviors," First International Symposium on Systems & Human Science - For Safety, Security, and Dependability, pp. 239-244, 2003.

[2] K. Yokoyama, J. Maeda, T. Isozumi, K. Kaneko, "Application of Humanoid Robots for Cooperative Tasks in the Outdoors," In Proc of the 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems, Workshop on Explorations towards Humanoid Robot Applications, 2001.

[3] S. Belongie, J. Malik, J. Puzicha, "Shape Matching and Object Recognition Using Shape Context," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 24, pp. 509-522, 2002.

[4] C. Schlegel, J. Illmann, H. Jaberg, M. Schuster, R. Wörz, "Vision Based Person Tracking with a Mobile Robot," In Proc. British Machine Vision Conference, pp. 418-427, 1998.

[5] D. Murray, J. Little, "Using Real-Time Stereo Vision for Mobile Robot Navigation," Autonomous Robots, Vol. 8, No. 2, pp. 161-171, 2000.

[6] S. Nishide, T. Takubo, K. Inoue, T. Arai, "Wide-Area Recognition Using Hybrid Motion Stereo," In Proc. of the 2005 IEEE International Conference on Robotics & Automation, pp. 818-823, 2005.

[7] P. J. Rousseeuw, "Least Median of Squares Regression," Journal of the American Statistical Association, 79, pp. 871-880, 1984.

[8] M. A. Fischler, R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," Communications of the ACM, Vol. 24, No. 6, pp. 381-395, 1981.

[9] K. Lee, P. Meer, R. Park, "Robust Adaptive Segmentation of Range Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 2, pp. 200-205, 1998.

[10] J. Miller, C. Stewart, "MUSE: Robust Surface Fitting using Unbiased Scale Estimates," in Proc. of the 1996 IEEE Computer Vision and Pattern Recognition, pp. 300-306, 1996.

[11] A. Adam, E. Rivlin, I. Shimshoni, "ROR: Rejection of Outliers by Rotations in Stereo Matching," In Proc. of the 2000 IEEE International Conference of Computer Vision and Pattern Recognition, Vol. 1, pp. 2-9, 2000.

[12] E. Vincent, R. Laganiere, "Matching Feature Points for Telerobotics," Haptic Audio Visual Environments and Their Applications(HAVE), pp. 13-18, 2002.

[13] J. Shi, C. Tomasi, "Good Features to Track," In Proc. of the 1994 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94), pp. 593-600, 1994.

[14] J. Bouguet, "Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm," Intel Corporation, Microprocessor Research Labs. OpenCV Documents, 1999.

[15] B. D. Lucas, T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," In Proc. of Imaging Understanding Workshop, pp. 121-130, 1981.

[16] OpenCV. Open Source Computer Vision Library. http://www.intel.com/research/mrl/research/opencv/.

[17] S. Kajita, et. al, "Biped Walking Pattern Generation by using Preview Control of Zero-Moment Point," In Proc. of the 2003 IEEE International Conference on Robotics & Automation, pp. 1620-1626, 2003.

[18] S. Nishide, T. Takubo, K. Inoue, T. Arai, "Wide Area Recognition for Safety and Security Preserving Robots," Second International Symposium on Systems & Human Science - For Safety, Security and Dependability, 2005.