3D Reconstruction from Stereo Camera Dynamic Image based on Particle Filter

Norikazu Ikoma, Wataru Ito, Toru Irie, and Hiroshi Maeda

Faculty of Engineering, Kyushu Institute of Technology, Kita-Kyushu city, Fukuoka 804-8550, JAPAN
E-mail: ikoma@comp.kyutech.ac.jp

Abstract: A new method for 3D reconstruction from dynamic stereo image by using state estimation with particle filter is proposed. Associations of feature points between two images and 3D position of the feature points are estimated simultaneously by particle filter. It is due to the applicability of particle filter for nonlinear and non-Gaussian state space model including unknown associations, non-Gaussian distributions appeared here, and nonlinearlity of projection. We assume that there are missing and error detection of feature points through the image processing for feature extraction. The novelty of our method is simultaneous estimation of the unknown associations and the 3D positions of feature points while most conventional methods are using 2 step estimation, i.e., firstly estimate the associations then secondly calculate the 3D positions depending on the estimated associations. Further improvement of estimation called Rao-blackwellization, which is a method for variance reduction of the estimate, is used at the implementation of particle filter with extended Kalman filter for nonlinear part of the model. Simulation experiment illustrates the efficiency of the method. At the concluding part, we mention about a possibility of this method to provide a basis for multiple sensor fusion problem in dynamic situation by extending the model into a general form.
Keywords: Stereo image, dynamic image, particle filter, nonlinear, non-Gaussian, sensor fusion.

1. Introduction

3D reconstruction in computer vision is one of the most popular problem in the image processing field and it is still challenging due to its fundamental difficulties [1]. It also has many possibilities for real applications such as robot vision, autonomous control of vehicle, scene analysis in geography, etc. Many works have been dedicated to this problem from many aspects, such as multiple view geometry [11], dynamic image situation[17], etc. Among these works, stereo camera, which uses two cameras to obtain two images of different views, is one of the earliest works in 3D reconstruction. It must be a major reason why stereo camera is popular that the fact of many mammals employing two eyes system provides an evidence of a system to reconstruct the 3D information from stereo camera.

In stereo camera, determining the associations between two images is a fundamental problem [21]. It is straightforward to calculate the 3D information from the stereo image when the associations are given. However, the associations are unknown in general and solving the association problem is not trivial despite that natural animals can easily do this task. From the theory of projection geometry (see e.g.,[11]), the associations are restricted to a line on the image plane called “epipolar line”. Epipolar line is obtained from the intersection of the image plane and a plane called “epipolar plane”, which is defined as a plane having two optical center of the cameras and the feature point that we are concerning. Thus the problem for solving the associations is reduced to a search problem on the epipolar line, i.e., one dimensional association problem.

For a solution to the association problem, the early paper, [21], proposed a method using the dynamic programming. In the paper, they firstly extract feature points from each image depending on local properties of the image such as 1st order difference along with the epipolar line. Thus feature points such as edges crossing the epipolar line are extracted in this process. Note that they also proposed a method using the information between adjoining epipolar lines, however for the sake of simplicity, we are focusing on intra-epipolar line method here. After that, to determine the associations between two sets of feature points, which correspond to left and right images respectively, dynamic programming is used with a cost function based on the similarity of intensity curves on the interval between two adjoining feature points on the epipolar line for each image. Thus the method uses only image (i.e.,2D) structure for solving the association problem. There is another advanced method to solve the associations problem including the occlusion and reversal positions [16], and it also uses only 2D structure for a solution.

Use of 3D structure for solving the association problem is interesting. However it seems to be difficult to solve the problem in this way since it involves simultaneous estimation of the associations and the 3D structure. On the other hand, when the associations are given, there are many researches to reconstruct the 3D information from multiple view or dynamic image. [4] proposed a factorization method of multiple view/dynamic image into 3D motion and structure in a scene of multiple moving objects with different motions. It uses singular value decomposition for reconstruction of motion and structure [17], and also uses block diagonalization for multiple body extraction. [3] uses extended Kalman filter to estimate 3D motion and structure from dynamic image of single object (i.e., single motion) scene. [15] deals with multiple motions by a method for simultaneous estimation of the motion of each feature point and the associations between feature points and objects in 2D space (i.e., on the image plane). All these methods assume the known associations of feature points between views for multiple view case or between frames for dynamic image case.

There is an another research field dealing with the unknown association. The field is target tracking, which is a classical field since the emergence of Kalman filter but is still interesting. In this field, simultaneous estimation of the trajectories and the associations has been actively investigated. Here, recent filtering technique called "particle filters" [8] [18] or "sequential Monte Carlo" (SMC) more generally [5] [19], which are computer intensive methods for state estimation (distribution estimation in complex situation more generally), allow us the simultaneous estimation due to the generality of the technique involving nonlinear and non-Gaussian situations. See [7] for more details. By using the particle filters, new approaches for the target tracking positively using the nonlinearlity in the problem are possible [12], [13], [20]. It is also possible to deal with the unknown association between target states and observed points [9], [10], [14].

In this paper, we propose to use the particle filters technique for 3D reconstruction problem in dynamic image with stereo camera. The method proposed here uses a state space model with system equation being representing the time evolution of feature points in 3D space and the unknown association of feature points, where the association is between the images of feature points of the model and the actually observed feature points. Here we assume the existence of missing and the error detection, which is a detection of feature point at not correct position due to the uncertainty of the observation process. Observation equation of the state space model consists of the camera projection model and the reflection of the association in the observation process. By doing the state estimation with the particle filters, we can obtain the estimation of the 3D positions of feature points as well as the association of feature points between left and right images indirectly through the estimation of the associations mentioned above.

Structure of the rest of this paper is as follows. We will firstly define a problem at section 2, then we will propose a model for the estimation of the 3D information and the association simultaneously at section 3. Technique of particle filters are reviewed at section 4 in order to be a self containing for this paper. Variance reduction method of the estimate called Rao-Blackwellization is also reviewed here. At section 5, we will report a result of simulation experiment to illustrate the efficiency of the method. Finally at section 6, we will make some concluding remarks including future works, especially that the proposed method will provide a new basis for multiple sensor fusion.

2. Problem Statement

We state the problem by defining camera model, stereo camera setting, and situation of the acquisition of dynamic image with missing and error detection in the following subsections.

2.1. Camera model

Firstly as a camera model, we employ central projection model as shown in Figure 1. In this figure, the origin of the world coordinate is placed at the camera center C. Each axis of the world coordinate is denoted by capital letters X, Y, and Z. f denotes focal length, which is distance between the camera center and the image plane. Image plane has a 2D coordinate called image coordinate, where each axis is denoted by small letters x and y. Z axis in this figure is called principal axis that starts from the camera center and is perpendicular to the image plane. An intersection point of the principal axis and the image plane, denoted by o in the figure, is called principal point.

Then, arbitrary point denoted in the world coordinate by $P = (X, Y, Z)$ is projected on the image plane at point $p = (x, y)$ in the camera coordinate with relations

$$x = \frac{f}{Z}X \ , \ y = \frac{f}{Z}Y \tag{1}$$

2.2. Stereo camera

We assume that two cameras in a stereo vision have the same camera model with the same camera parameters for the sake of convenience. Note that it is possible to relax this condition in the following arguments. The camera parameters are assumed to be known. Thus two cameras are denoted by central projection model with the same focal length f where f is known.

We also assume that relative position of the two cameras is known. For the sake of simplicity, we restrict the stereo vision problem into a special case called "the geometry of nonverged stereo" [1], where two cameras are placed in parallel with respect to their principal axes and their x-axes of the image

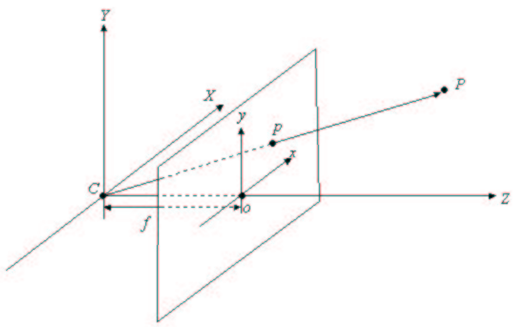

Figure 1: Central projection model.

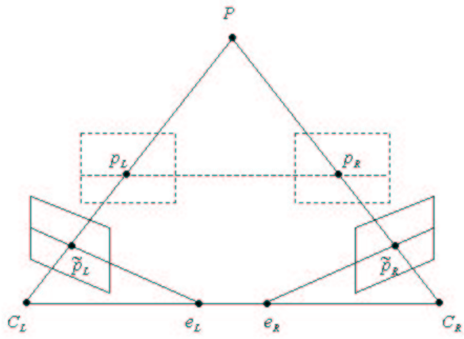

Figure 2: Epipolar constraint.

plane are on the same line. We simply refer this situation as “parallel stereo camera”. A transformation called “rectification” converts an image from arbitrary placed stereo camera into one of the parallel stereo camera.

Figure 2 illustrates stereo camera in general situation, i.e., two cameras are placed at arbitrary positions with camera centers C_L and C_R and the image planes are shown by rectangles with solid line in the figure. A point P in the world coordinate is observed by these cameras, then we obtain the images of the point as $\tilde{p}_L$ and $\tilde{p}_R$ for left and right cameras respectively. Then, epipolar plane is obtained to have P, C_L, and C_R on the plane shown by the triangle in figure 2. Epipolar line is then obtained by taking the intersection between the image plane and the epipolar plane. Where a line segment between C_L and C_R is called “baseline”, and an intersection point of the epipolar line and the baseline is called “epipole”. e_L and e_R denote epipoles for left and right images respectively.

Rectification is a transformation of the original images to the image plane shown by rectangles with dashed line in figure 2. Two image planes written with dashed lines are on the same plane in the world coordinate. Thus epipolar lines of two image are on the same line as denoted by a dashed line in the figure. Points observed by two cameras then become p_L and p_R.

2.3. Association problem

As observed in the previous subsection, the epipolar lines of left and right images are on the same line in parallel stereo camera situation. Thus the association problem between left image and right image is reduced into a problem in a space on the two horizontal line segments of left and right images with the same vertical position. Although there are many epipolar lines of different vertical positions in one pair of images in general, we here focus on one pair of epipolar lines of left and right images for the simplicity sake in the following arguments.

We formulate this situation as shown in figure 3. Where the baseline is taken as $2d$, origin of the world coordinate, O, is placed at the center of the baseline, i.e., at the central position between two camera centers C_L and C_R. X axis of the world coordinate is on the baseline. Principal points, o_L and o_R, are of the origin of each image plane for left and right image respectively. Thus the image of the feature point P on the left image, denoted by p_L is measured with origin o_L, and with origin o_R for right image p_R.

Then, we obtain relations

$$\left\{ \begin{array}{lcl} x^L & = & \dfrac{f}{Z}(X+d) \\ x^R & = & \dfrac{f}{Z}(X-d) \end{array} \right. \tag{2}$$

where x^L and x^R are observed values of x axes of the target feature point P for left and right images with origin o_L and o_R respectively, and $P = (X, Y, Z)$ is the position of the target feature point in the world coordinate. Note that we can omit Y coordinate in this formulation.

In general, many feature points on the epipolar line are obtained by a simple pre-process for the image using local properties of the image such as 1st order difference of intensity along with the epipolar line. Then the association problem is to obtain the correspondence between a set of feature points on the left image and a set of them on the right image. To discriminate each feature point, we put an index number on the feature points as

$$P_i = (X_i, Y_i, Z_i)\,, x_i^L, x_i^R \quad \text{for } i = 1, 2, \cdots . \tag{3}$$

Where we are temporally assuming that all feature points that we are concerning can be viewed by both

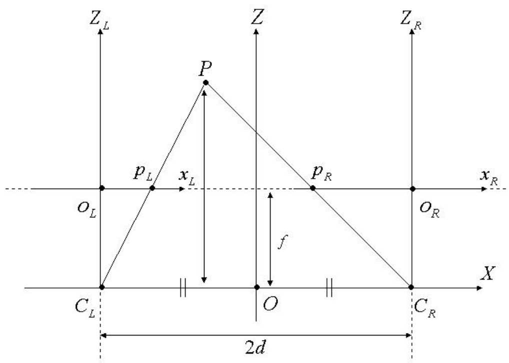

Figure 3: Parallel stereo camera.

left and right cameras without missing, and there is no error detection. In equation (3), the association between the feature point in the world coordinate, P_i, and the feature points on the two images, x_i^L and x_i^R, are known through the index i.

In real situation, some of the feature points might not be detected by simple image processing using local properties such as high value of 1st difference of intensity along with the epipolar line. This may occur in a probabilistic sense. Occlusion also provides the missing in a deterministic sense. On the other hand, we also should deal with an error detection, which is a detection of feature point where actual feature point does not exist. This is due to a small fluctuation of the intensity along with the epipolar line that comes from unstable conditions on image, for example light, shade, etc. Thus some of the observed feature points might be of the error detection.

3. Model

We propose a new model for 3D reconstruction in stereo camera dynamic image in this section. Since it is a first step for the new method for 3D reconstruction, we employ a simple situation with single motion in the scene. That is, there are objects with single motion in a scene, or camera system is moving with static scene. The objects in a scene are assumed to be rigid basically, although proposed method can deal with non-rigid objects when they are approximately rigid over short time interval. We also assume for the motion to be only translation, i.e., no rotation.

3.1. Feature points

We assume that there are N^* feature points with N^* be fixed and known. All feature points move with the same velocity due to a scene with one moving object or static scene with moving stereo camera system.

Let position of i-th feature point in the world coordinate at time k, where Y coordinate is omitted, be denoted by

$$\mathbf{P}_i(k) = \Big(X_i(k)\ ,\ \tilde{Z}_i(k)\Big)\,. \tag{4}$$

Note that i varies in $\{1, 2, \cdots, N^*\}$. Where Z coordinate is log transform of its original value by $\tilde{Z}_i(k) = \log Z_i(k)$ for the sake of the range be in $\Re$ while the original range be positive.

Motion of i-th feature point is denoted by a difference equation

$$\mathbf{P}_i(k) = \mathbf{P}_i(k-1) + \mathbf{V}(k) \tag{5}$$

where

$$\mathbf{V}(k) = \Big(V_X(k)\ ,\ \tilde{V}_Z(k)\Big) \tag{6}$$

is a velocity vector shared by all feature points, and it is defined by a motion of the object in a single moving object scene or motion of the camera system with static scene.

3.2. Motion model

For the velocity vector of the feature points, there are several way to define it depending on the assumption for the actual motion. The simplest one is fixed velocity case denoted by

$$\mathbf{V}(k) = \mathbf{V}(k-1). \tag{7}$$

The second one is that the velocity is defined by a random walk process such as

$$\mathbf{V}(k) = \mathbf{V}(k-1) + \dot{\mathbf{V}}(k) \tag{8}$$

where $\dot{\mathbf{V}}(k) \sim N(\mathbf{0}, \mathbf{Q}_v)$ is a random factor to give a flexibility for the time change of the velocity vector. If the variance matrix $\mathbf{Q}_v$ has small values in its elements then the change will be small, vice versa.

The third one is more complicated where the velocity vector is defined by a second order difference equation with a form

$$\left\{\begin{array}{lclcl} \mathbf{V}(k) & = & \mathbf{V}(k-1) & + & \dot{\mathbf{V}}(k-1) \\ \dot{\mathbf{V}}(k) & = & \dot{\mathbf{V}}(k-1) & + & \mathbf{A}(k) \end{array}\right. \tag{9}$$

where a component $\mathbf{A}(k) = \Big(A_X(k)\ ,\ \tilde{A}_Z(k)\Big) \sim N(\mathbf{0}, \mathbf{Q}_a)$ is a random factor given to the first order difference of velocity, i.e. acceleration. This means that the acceleration is according to a random walk.

Higher order difference equations are possible to use if needed.

3.3. Projection model

All feature points are projected to the image planes of left and right cameras according to the each camera projection model. Here we employ the central projection model for both cameras, where relations in equation (2) hold, then i-th feature point at time k will be projected to each image by

$$\left\{ \begin{array}{lcl} x_i^L(k) & = & f e^{-\tilde{Z}_i(k)} \left(X_i(k) + d \right) \\ x_i^R(k) & = & f e^{-\tilde{Z}_i(k)} \left(X_i(k) - d \right) \end{array} \right. . \quad (10)$$

3.4. Observation model

Let $s \in \{L, R\}$ be a index of camera, where it denotes left camera if $s = L$, vice versa.

Assume that N_k^s feature points are observed from the image of s-camera at time k. Where observed feature points may contain error detections, and not necessarily all the feature points in the scene are observed. Thus, by letting $N_D^s(k)$ be the number of detected feature points and $N_C^s(k)$ be the number of error detections, it holds

$$N^s(k) = N_D^s(k) + N_C^s(k). \quad (11)$$

Note that $N_D^s(k)$ and $N_C^s(k)$ are unknown while N_k^s is known through the observation.

Notice that actually observed feature points are of unknown source, i.e., we do not known which feature point have produced a point on image actually observed at each time k. To represent this, association variable is introduced as

$$I_j^s(k) \in \{0, 1, \cdots, N^*\} \quad (12)$$

for j-th observed point of s-camera at time k. If the association variable takes 0 then the observed point is due to the error detection. If it takes positive value then the corresponding observation comes from $I_j^s(k)$-th feature point of equation (4).

By letting $o_j^s(k)$ be j-th observed point of s-camera at time k, the observation process is denoted by following formula

$$o_j^s(k) \sim \left\{ \begin{array}{ll} N(x^s_{I_j^s(k)}(k), \sigma^2) & , \quad I_j^s(k) > 0 \\ U(\mathcal{X}) \quad \text{or} & \\ N(\bar{x}, \alpha\sigma^2), \alpha >> 1 & , \quad I_j^s(k) = 0 \end{array} \right. \quad (13)$$

where $U(\mathcal{X})$ denotes uniform distribution with range $\mathcal{X}$, with $\mathcal{X}$ be the surveillance range of the observation, i.e., range of x axis of the image plane.

3.5. Uncertainty in observation process

We define here several random processes to represent the uncertainty in the observation process that will give conditions for the association between the feature points of the model and the actually observed points.

First of all, we define a process of missing observation that may occur in a probabilistic sense as follows. Let p_D be a probability of detection of each feature point of each camera. The detection process is assumed to be independent for each feature point and for each camera, and independent of the 3D positions of the feature points. It is also assumed to be independent with respect to time. Then, the number of detected points, $N_D^s(k)$, is according to a binomial distribution, $N_D^s(k) \sim B(N^*, p_D)$, with its probability function

$$p(N_D) = \left(\begin{array}{c} N^* \\ N_D \end{array} \right) P_D^{N_D} (1 - p_D)^{N^* - N_D}, \quad (14)$$

where we suppress time index k and camera index s for the sake of simplicity.

Secondly for the process of missing in a deterministic sense such as occlusion, we only suggest a direction of the modeling in this paper as follows. While the detection process in a probabilistic sense is represented by equation (14) as an unconditional distribution, the process in a deterministic sense will have a conditional probability distribution of N_D with 3D positions of the feature points be in the conditional part. There may be many choices for the conditional distribution, however we are not go into the details of the design for it.

Next, for the error detection, the number of error detection points, $N_C^s(k)$, is assumed to be according to a Poisson distribution, $N_C^s(k) \sim \text{Poisson}(\mu V)$, independently of time, camera, and the 3D positions of the feature points. Thus its probability function is, with suppressing indices of time and camera,

$$p(N_C) = e^{-\mu V} (\mu v)^{N_C} / N_C!, \quad (15)$$

where V is the volume of surveillance, i.e., width of the image, and μ is a spatial density to occur the error detection.

There is another uncertainty in the observation process that is the order of the observed points. We show here a case of completely no information about the order. Note that in stereo camera, there can be more restricted rules on the order, e.g., sorted in ascending order, however, we will not consider them in this paper. Permutation of the detected points and the error detection points are assumed to be uniformly distributed, i.e. there are $N^s(k)!$ combinations for it with the same probability $1/N^s(k)!$ for each case.

4. State estimation by particle filters

4.1. State space representation

The model described in the previous section can be written in a state space representation with a linear system equation and a nonlinear observation equation, which is conditioned by the association variables in equation (12).

First, a state vector consists of feature points in the world coordinate and velocity vector as follows

$$\mathbf{\Theta}_k = (\mathbf{P}_1(k), \mathbf{P}_2(k), \cdots \mathbf{P}_{N^*}(k), \mathbf{V}(k))', \quad (16)$$

where $\mathbf{x}$' denotes transpose of vector $\mathbf{x}$. This is the case for the velocity is according to the constant model shown in equation (7), or is according to the random walk model shown in equation (8). If one choose the model in equation (9) as the motion model, then 1st order difference of the velocity, $\dot{\mathbf{V}}(k)$, appears in the state vector. Higher order differences of the velocity may appear if necessary.

Second, we form a association vector, $\mathbf{I}_k$, having elements $I_1^L(k)$, $I_2^L(k)$, $\cdots$, $I_{N^L(k)}^L(k)$ $I_1^R(k)$, $I_2^R(k)$, $\cdots$, and $I_{N^R(k)}^R(k)$. Then, a system equation of the state space representation can be written as

$$\mathbf{\Theta}_k = \mathbf{F}\mathbf{\Theta}_{k-1} + \mathbf{G}\mathbf{s}_k, \quad \mathbf{s}_k \sim N(\mathbf{0}, \mathbf{Q}) \tag{17}$$

where matrices $\mathbf{F}$ and $\mathbf{G}$ are appropriately defined in order to hold the models in equation (5) and the motion model. $\mathbf{s}_k$ is a system noise vector defined according to the motion model. It will be null vector if one use equation (7), will be $\dot{\mathbf{V}}(k)$ if one use equation (8), or will be $\mathbf{A}(k)$ if equation (9) is used.

By letting a observation vector, $\mathbf{O}_k$, having elements $o_1^L(k)$, $o_2^L(k)$, $\cdots$, $o_{N^L(k)}^L(k)$, $o_1^R(k)$, $o_2^R(k)$, $\cdots$, and $o_{N^R(k)}^R(k)$, we can write an observation equation, which is nonlinear with respect to the state vector $\mathbf{\Theta}_k$ and conditioned by the association vector, as follows,

$$\mathbf{O}_k = \mathbf{h}\left(\mathbf{\Theta}_k; \mathbf{I}_k\right), \tag{18}$$

where $\mathbf{h}(\cdot)$ is a vector valued function designed to hold the relation (13).

Now, we have the state space representation with system equation (17) and observation equation (18), conditioned by the association vector $\mathbf{I}_k$. By using 1st order Taylor expansion, we have a linear approximation of the observation equation as

$$\mathbf{O}_k \simeq \mathbf{H}\left(\mathbf{I}_k\right)\mathbf{\Theta}_k + \mathbf{E}\left(\mathbf{I}_k\right)\mathbf{e}_k, \;\; \mathbf{e}_k \sim N(\mathbf{0}, \mathbf{R}) \tag{19}$$

where $\mathbf{e}_k$ is a vector consisting of independent elements $e_j \sim N(0, \sigma^2)$, $j = 1, 2, \cdots, N^L(k) + N^R(k)$.

4.2. State estimation

State estimation is to obtain a conditional distribution of the state vector $\mathbf{\Theta}_k$ given a series of the observations up to a certain time say K. The estimation is called "filter" if $k = K$, it is called "prediction" when $k > K$, and is called "smoothing" if $k < K$. According to the aim of this research, filtering estimation is the one what we are interested in.

In the model proposed here, which is shown by system equation (17) and observation equation (18), if $\mathbf{I}_k$ is given for all time $k = 1, 2, \cdots$, then the state estimation can proceed by extended Kalman filter since the model consists of linear system equation and nonlinear observation equation, and all noises are Gaussian. However, we do not know the association vector $\mathbf{I}_k$, so we need to estimate the association vector as well as the state vector $\mathbf{\Theta}_k$. To achieve this, we augment the state vector to have the original state vector $\mathbf{\Theta}_k$ and the association vector $\mathbf{I}_k$ as

$$\Xi_k = \left(\mathbf{\Theta}_k{}', \mathbf{I}_k{}'\right). \tag{20}$$

Then the state space model and its state estimation problem become highly nonlinear, and there is no closed form solution in general. Hence we need to use some approximation method to solve the state estimation problem.

We will use "particle filters" for the approximation method of state estimation as described in the following subsections. Where, we will use a notation $\mathbf{\Theta}_{1:k} \equiv (\mathbf{\Theta}_1, \mathbf{\Theta}_2, \cdots, \mathbf{\Theta}_k)$. Thus the filtering distribution can be written as $p(\Xi_k|\mathbf{O}_{1:k})$.

4.3. Particle filters

We here review the particle filters [8] [18] or "sequential Monte Carlo" (SMC) more generally [5] [19]. See [7] for more details.

Particle filters approximate a target distribution by many number (say M) of weighted particles. For the target distribution, filtering distribution, denoted by $p(\Xi_k|\mathbf{O}_{1:k})$, is sufficient to our objective. However, for the sake of convenience, we will discuss on the estimation of a joint distribution up to current time k, denoted by $p(\Xi_{0:k}|\mathbf{O}_{1:k})$, instead of the filtering distribution. We can have the filtering distribution by a marginal of the distribution $p(\Xi_{0:k}|\mathbf{O}_{1:k})$.

Temporally assume that we can draw particles from the target distribution. Note that it is not a realistic assumption since our objective itself is to obtain the target distribution so it is not possible to draw from the distribution by a straightforward method in general. Under this assumption, particles are denoted by

$$\left\{\Xi_{0:k}^{(l)}\right\}_{l=1}^{M} \sim p(\Xi_{0:k}|\mathbf{O}_{1:k}). \tag{21}$$

Then, we can obtain Monte Carlo approximation of any estimate of the state in the form

$$I(f) = \int f(\Xi_{0:k}) p(\Xi_{0:k}|\mathbf{O}_{1:k}) d\Xi_{0:k} \tag{22}$$

which has no closed form solution to the integral, by

$$\hat{I}_M(f) = \frac{1}{M}\sum_{l=1}^{M} f(\Xi_{0:k}^{(l)}). \tag{23}$$

Note that in equation (22), we take $f(\Xi) = \Xi$ when we estimate the mean of the state, or take $f(\Xi) = \left(\Xi - \bar{\Xi}\right)^2$, with $\bar{\Xi}$ be the mean estimate, when one wants to estimate the variance of the state, for example. By the Monte Carlo approximation, not only obtain the approximation of the integral in equation (22) with no closed form solution, but also we can circumvent a problem so called "curse of dimension" involved in the numerical integration of equation (22).

However, it is not possible in general to draw directly from the target distribution, we use importance sampling method by drawing the particles from an importance function $\pi(\cdot)$, which has a condition $p(\Xi) > 0 \Rightarrow \pi(\Xi) > 0$, as

$$\left\{\tilde{\Xi}_{0:k}^{(l)}\right\}_{l=1}^{M} \sim \pi(\Xi_{0:k}|\mathbf{O}_{1:k}), \tag{24}$$

and approximate the estimate of equation (22) by

$$\hat{\hat{I}}_M(f) = \sum_{l=1}^{M} f(\tilde{\Xi}_{0:k}^{(l)})\omega_k^{(l)}, \tag{25}$$

with weight

$$\omega_k^{(l)} = \frac{p(\tilde{\Xi}_{0:k}^{(l)}|\mathbf{O}_{1:k})}{\pi(\tilde{\Xi}_{0:k}^{(l)}|\mathbf{O}_{1:k})}. \tag{26}$$

Now the state estimation problem becomes to obtain a set of weighted particles $\left\{\left(\tilde{\Xi}_{0:k}^{(l)}, \omega_k^{(l)}\right)\right\}_{l=1}^{M}$ sequentially in time. That is, we will apply a update procedure to the weighted particles of time $k-1$ when the new observation $\mathbf{O}_k$ becomes available, then we obtain the updated weighted particles of time k. The procedures start from the initial weighted particles $\left\{\left(\tilde{\Xi}_0^{(l)}, \omega_0^{(l)}\right)\right\}_{l=1}^{M}$.

We can derive the update procedure as follows. Sequential estimation formula in general is of the form

$$\begin{aligned} p(\Xi_{0:k}|\mathbf{O}_{1:k}) &= p(\Xi_{0:k-1}|\mathbf{O}_{1:k-1}) \\ &\times \frac{p(\mathbf{O}_k|\Xi_k)p(\Xi_k|\Xi_{k-1})}{p(\mathbf{O}_k|\mathbf{O}_{0:k-1})}, \end{aligned} \tag{27}$$

here we have used Markov property of the system equation $p(\Xi_k|\Xi_{0:k-1}, \mathbf{O}_{1:k-1}) = p(\Xi_k|\Xi_{k-1})$ and independent property of the observation equation $p(\mathbf{O}_k|\Xi_{0:k}, \mathbf{O}_{1:k-1}) = p(\mathbf{O}_k|\Xi_k)$. Note that in equation (27), the denominator of right hand side, $p(\mathbf{O}_k|\mathbf{O}_{0:k-1})$, is the integral of the numerator with respect to $\Xi_{0:k}$. Thus we have another integration problem here, i.e., it has no closed form solution in general, as well as it has high dimensionality of integration.

To achieve the sequential estimation as shown in equation (27), at time k, we draw only particles for time k, and reuse the past particles. Then, the new draw becomes

$$\tilde{\Xi}_k^{(l)} \sim \pi(\Xi_k|\Xi_{0:k-1}^{(l)}, \mathbf{O}_{1:k}) \tag{28}$$

and we let the past particles be

$$\tilde{\Xi}_{0:k-1}^{(l)} \equiv \Xi_{0:k-1}^{(l)} \tag{29}$$

for $l = 1, 2, \cdots, M$. This means that we are using the importance function as

$$\begin{aligned} \pi(\Xi_{0:k}|\mathbf{O}_{1:k}) &\equiv \pi(\Xi_k|\Xi_{0:k-1}, \mathbf{O}_{1:k}) \\ &\times \pi(\Xi_{0:k-1}|\mathbf{O}_{1:k-1}). \end{aligned} \tag{30}$$

Next for updating the weights of particles, divide both sides of equation (27) by both sides of equation (30), then we obtain the weight update procedure

$$\tilde{\omega}_k^{(l)} \propto \omega_{k-1}^{(l)} \frac{p(\mathbf{O}_k|\tilde{\Xi}_k^{(l)})p(\tilde{\Xi}_k^{(l)}|\Xi_{k-1}^{(l)})}{\pi(\tilde{\Xi}_k^{(l)}|\Xi_{0:k-1}^{(l)}, \mathbf{O}_{1:k})}, \tag{31}$$

where we omit $p(\mathbf{O}_k|\mathbf{O}_{0:k-1})$ since it does not depend on $\tilde{\Xi}_k^{(l)}$, i.e., common constant for all particles $l = 1, 2, \cdots, M$. It can be shown that the weight without the constant is enough to obtain the set of weighted particles according to the posterior distribution by introducing a normalizing procedure

$$\omega_k^{(l)} = \tilde{\omega}_k^{(l)} / \textstyle\sum_{j=1}^{M} \tilde{\omega}_k^{(j)}. \tag{32}$$

Consequently, we have no need to evaluate the term $p(\mathbf{O}_k|\mathbf{O}_{0:k-1})$, which has the integral with no closed form and with high dimension, thus we can circumvent difficulty of the integration problem.

Finally, we will resample from the set of weighted particles $\left\{\left(\tilde{\Xi}_{0:k}^{(l)}, \omega_k^{(l)}\right)\right\}_{l=1}^{M}$ with probability proportional to the value of weight $\omega_k^{(l)}$. Specifically, for each $l = 1, 2, \cdots, M$, draw a random index variable $j^{(l)}$ according to multinomial distribution of values $1, 2, \cdots, M$ with probabilities $\omega_k^{(1)}, \omega_k^{(2)}, \cdots, \omega_k^{(M)}$ respectively, then we let $\Xi_{0:k}^{(l)} := \tilde{\Xi}_{0:k}^{(j^{(l)})}$. And let all weights be equal, i.e., $\omega_k^{(l)} := 1/M$. Now, we obtain particles, $\left\{\Xi_{0:k}^{(l)}\right\}_{l=1}^{M}$, that can be considered as drawn from $p(\Xi_{0:k}|\mathbf{O}_{1:k})$. The resampling step might not be necessary depending on the variance of the weights. When the resampling step is omitted, we do $\Xi_{0:k}^{(l)} := \tilde{\Xi}_{0:k}^{(l)}$ and keep the weight values $\omega_k^{(l)}$ until the next updating.

4.4. Rao-Blackwellization

Rao-Blackwellization (RB) [2] is a general variance reduction method for Monte Carlo estimation available for particle filters [6],[7]. RB uses a property of the target distribution possible to decompose into two parts, where one part has a closed form solution. In the context of particle filters in our model, the posterior distribution can be decomposed as

$$\begin{aligned} p(\Xi_{0:k}|\mathbf{O}_{1:k}) &= p(\mathbf{\Theta}_{0:k}, \mathbf{I}_{0:k}|\mathbf{O}_{1:k}) \\ &= p(\mathbf{\Theta}_{0:k}|\mathbf{I}_{0:k}, \mathbf{O}_{1:k})p(\mathbf{I}_{0:k}|\mathbf{O}_{1:k}), \end{aligned} \tag{33}$$

where $p(\mathbf{\Theta}_{0:k}|\mathbf{I}_{0:k}, \mathbf{O}_{1:k})$ approximately has a closed form solution since extended Kalman filter will apply. Thus the particle approximation of remaining distribution, $p(\mathbf{I}_{0:k}|\mathbf{O}_{1:k})$, is the main task in RB.

Update procedure of weighted particles in RB is as follows. Weighted particles are then denoted by $\left\{\left(\tilde{\mathbf{I}}_{0:k}^{(l)}, \nu_k^{(l)}\right)\right\}_{l=1}^{M}$, where the particles are assumed to be drawn from a importance function as $\tilde{\mathbf{I}}_{0:k}^{(l)} \sim$

$\pi(\mathbf{I}_{0:k}|\mathbf{O}_{1:k})$. For a draw of new particle at time k, we also use a decomposed importance function $\pi(\mathbf{I}_{0:k}|\mathbf{O}_{1:k}) \equiv \pi(\mathbf{I}_{0:k-1}|\mathbf{O}_{1:k-1})\pi(\mathbf{I}_k|\mathbf{I}_{0:k-1},\mathbf{O}_{1:k})$ as same as the original particle filters. Update procedure of weight is obtained by a similar manner of equation (31) as

$$\tilde{\nu}_k^{(l)} \propto \nu_{k-1}^{(l)} \frac{p(\mathbf{O}_k|\tilde{\mathbf{I}}_{0:k}^{(l)},\mathbf{O}_{0:k-1})p(\tilde{\mathbf{I}}_k^{(l)}|\mathbf{I}_{k-1}^{(l)},\mathbf{O}_{0:k-1})}{\pi(\tilde{\mathbf{I}}_k^{(l)}|\mathbf{I}_{0:k-1}^{(l)},\mathbf{O}_{1:k})}. \quad (34)$$

Major difference from the original particle filters is that $p(\mathbf{O}_k|\tilde{\mathbf{I}}_{0:k}^{(l)},\mathbf{O}_{0:k-1})$ is obtained through the (extended) Kalman filter procedure for a given particle $\tilde{\mathbf{I}}_{0:k}^{(l)}$. Note for the other term, $p(\tilde{\mathbf{I}}_k|\mathbf{I}_{k-1},\mathbf{O}_{0:k-1})$, that it varies depending on the assumption for the time evolution of the association vector. If it simply occurs independently for each time, then $p(\tilde{\mathbf{I}}_k)$. It may depend on the association vector of previous time, then we have $p(\tilde{\mathbf{I}}_k|\mathbf{I}_{k-1})$. Further it can depend on the state vector $\mathbf{\Theta}_k$ such as occlusion case, then the term will be more complicated one.

Let us define some notations for the Kalman filter procedure, filtering distribution, $p(\mathbf{\Theta}_k|\tilde{\mathbf{I}}_{0:k},\mathbf{O}_{0:k})$, is Gaussian with mean vector $\theta_{k|k}$ and covariance matrix $\mathbf{C}_{k|k}$, and one-step-ahead prediction distribution, $p(\mathbf{\Theta}_k|\tilde{\mathbf{I}}_{0:k},\mathbf{O}_{0:k-1})$, is Gaussian with mean vector $\theta_{k|k-1}$ and covariance matrix $\mathbf{C}_{k|k-1}$. To proceed the Kalman filter for RB, we need to keep these vectors and matrices for each particle, thus we write $\theta_{k|k}^{(l)}$ and $\mathbf{C}_{k|k}^{(l)}$ for filtering distribution of l-th particle and $\theta_{k|k-1}^{(l)}$ and $\mathbf{C}_{k|k-1}^{(l)}$ for one-step-ahead prediction distribution of l-th particle. We also define notations for the prediction of observation denoted by $p(\mathbf{O}_k|\tilde{\mathbf{I}}_{0:k}^{(l)},\mathbf{O}_{0:k-1})$, which is Gaussian distribution as well, the mean vector is denoted by $\bar{\mathbf{O}}_{k|k-1}^{(l)}$, and covariance matrix is by $\mathbf{\Sigma}_{k|k-1}^{(l)}$.

Now we can write the extended Kalman filter procedure for RB. One-step-ahead prediction proceeds with

$$\theta_{k|k-1}^{(l)} = \mathbf{F}\theta_{k-1|k-1}^{(l)}, \quad (35)$$

$$\mathbf{C}_{k|k-1}^{(l)} = \mathbf{F}\mathbf{C}_{k-1|k-1}^{(l)}\mathbf{F}' + \mathbf{G}\mathbf{Q}\mathbf{G}'. \quad (36)$$

Next, we calculate the prediction of observation

$$\bar{\mathbf{O}}_{k|k-1}^{(l)} = \mathbf{H}\left(\tilde{\mathbf{I}}_k^{(l)}\right)\theta_{k|k-1}^{(l)}, \quad (37)$$

$$\mathbf{\Sigma}_{k|k-1}^{(l)} = \mathbf{H}\left(\tilde{\mathbf{I}}_k^{(l)}\right)\mathbf{C}_{k|k-1}^{(l)}\mathbf{H}\left(\tilde{\mathbf{I}}_k^{(l)}\right)' + \mathbf{R}. \quad (38)$$

Then we are able to update the weight according to equation (34).

Filtering procedure of Kalman filter is required for the next updating procedure since the procedure starts from the filtering distribution at time $k-1$. Firstly, Kalman gain is calculated by

$$\mathbf{K}_n^{(l)} = \mathbf{C}_{k|k-1}^{(l)}\mathbf{H}\left(\tilde{\mathbf{I}}_k^{(l)}\right)'\left[\mathbf{\Sigma}_{k|k-1}^{(l)}\right]^{-1}, \quad (39)$$

then the filtering procedure are as follows,

$$\theta_{k|k}^{(l)} = \theta_{k|k-1}^{(l)} + \mathbf{K}_n^{(l)}\left(\mathbf{O}_k - \bar{\mathbf{O}}_{k|k-1}^{(l)}\right), \quad (40)$$

$$\mathbf{C}_{k|k}^{(l)} = \left[\mathbf{I} - \mathbf{K}_n^{(l)}\mathbf{H}\left(\tilde{\mathbf{I}}_k^{(l)}\right)\right]\mathbf{C}_{k|k-1}^{(l)}. \quad (41)$$

Mean vector and covariance matrix of the filtering distribution are stored with the particle $\mathbf{I}_{0:k}^{(l)}$ and they are also resampled together with the particle if resampling proceeds.

4.5. Importance function

In particle filters, there are some choices on the importance function $\pi(\tilde{\mathbf{I}}_k|\mathbf{I}_{0:k-1},\mathbf{O}_{1:k})$. Key of the choice is effective use of the current observation, $\mathbf{O}_k$. Optimal importance function, in the sense of minimum variance of the weights, is investigated [7], and it is suggested that the distribution $p(\tilde{\mathbf{I}}_k|\mathbf{I}_{0:k-1},\mathbf{O}_{1:k})$ achieves the optimal. However it is impractical due to its huge computational cost. Thus the exploration of sub-optimal importance functions becomes a key for the particle filtering.

We propose to use a sub-optimal importance function as follows. Firstly, we can divide it into each camera due to the independent property. By denoting part of s-camera by $\mathbf{I}_k^s$, $\mathbf{O}_{1:k}^s$, etc., then importance function of s-camera can be decomposed as

$$\begin{aligned}\pi(\mathbf{I}_k^s|\mathbf{I}_{0:k-1}^s,\mathbf{O}_{1:k}^s) &= \pi(I_1^s(k)|\mathbf{I}_{0:k-1}^s,\mathbf{O}_{1:k}^s)\\ &\times \pi(I_2^s(k)|I_1^s(k),\mathbf{I}_{0:k-1}^s,\mathbf{O}_{1:k}^s)\\ &\times \pi(I_3^s(k)|I_1^s(k),I_2^s(k),\mathbf{I}_{0:k-1}^s,\mathbf{O}_{1:k}^s)\\ &\quad\vdots\\ &\times \pi(I_{N^s(k)}^s(k)|I_1^s(k),I_2^s(k),\cdots,\mathbf{I}_{0:k-1}^s,\mathbf{O}_{1:k}^s)\end{aligned} \quad (42)$$

where j-th term of the right hand side is proportional to the following pdf's; here we suppress the time index k in some parts for simplicity,

$$\pi(I_j^s|I_1^s,I_2^s,\cdots,I_{j-1}^s,\mathbf{I}_{0:k-1}^s,\mathbf{O}_{1:k}^s) \propto \begin{cases} N(o_j^s;\bar{x},\alpha\sigma^2) & , \quad I_j^s = 0\\ 0 & , \quad I_j^s \neq 0, \exists i<j \; I_j^s = I_i^s\\ N(o_j^s;x_{I_j^s}^s,\sigma^2) & , \quad \text{otherwise}\end{cases} \quad (43)$$

where $N(x;m,\sigma^2)$ denotes density function of Gaussian distribution with mean m and variance σ^2.

Practically, $\pi(\cdot)^\gamma$ is used as the importance function instead of $\pi(\cdot)$ itself, with $\gamma < 1$. This is aimed at relaxing the skewness of the original importance function. This is because that the original importance function might not be correct, i.e. not the same as the target distribution, so too skew distribution will produce degenerated particles in which most particles are far from the right zone in the state space.

5. Simulation experiment

A simulation experiments have been carried out to illustrate how the proposed method efficiently works.

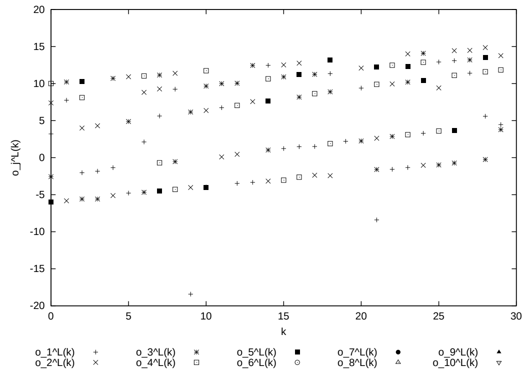

Figure 4: Feature points on left image.

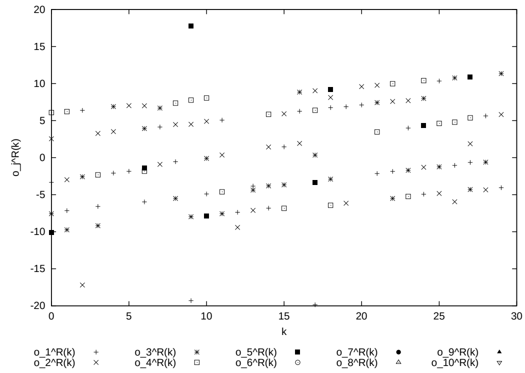

Figure 5: Feature points on right image.

Synthetic data, feature points on left and right image, are shown in Figure 4 and Figure 5. Where, number of feature points is $N^* = 5$, detection probability is set to $p_D = 0.7$, spatial density of error detection $\mu = 0.005$ is used with volume of surveillance $V = 40$. Gaussian observation noise with zero mean and variance $\sigma^2 = 10^{-2}$ is added to these feature points.

Conditions of state estimation by RB particle filters are as follows. Number of particles is set to $M = 1,000$. Observation noise variance is set to the true one $\sigma^2 = 10^{-2}$ with clutter factor $\alpha = 10^{16}$. System noise variance for V_X is set to 10^{-4}, and for V_Z is also set to 10^{-4}. Skew relaxing factor in importance function is set to $\gamma = 10^{-2}$.

Estimation results are shown in Figure 6 for X and Figure 7 for Z. Estimation results of associations are shown in Table 1. All associations for all time are correctly estimated in this experiment.

By looking at the estimation results of 3D positions of feature points in X-axis and Z-axis, shown in Figure 6 and Figure 7 respectively, we can see that estimated trajectories correctly track the true trajectories after 5 time step both for X-axis and Z-axis.

6. Conclusion

A new method for 3D reconstruction from dynamic image of stereo camera by using elaborated state space model and use of particle filters for the state estimation. The model consists of 3D positions of feature points moving with the same velocity of translation, and they are projected to the left and right images according to the camera models. Feature points observed on the images are assumed to have unknown associations to the feature points in the model. Thus the proposed method estimate the 3D information of the feature points and the associations simultaneously. We also assume that there are missing feature points, as well as error detections, which are detections of feature points at not existing

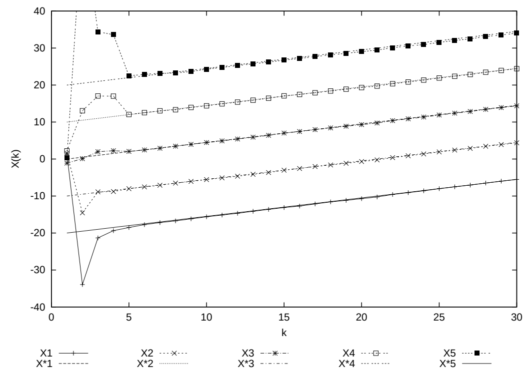

Figure 6: Estimation result of X.

positions. Simulation experiment illustrates the efficiency of the method, where we have obtained 3D trajectories of feature points very close to the true trajectories, and also obtained correct associations between observation and the model.

For future researches, there are several remaining problems in the proposed model, such as dealing with rotation, proper treatment of occlusion, permutation of the observation, and realistic camera model. By introducing them, we can apply the method to a real stereo image sequences. Current method is limited to a single motion case, so dealing with the multi motion case is interesting for the future research. In this case, we need to introduce another association variables between feature points and motion objects.

We remark that the proposed method here is just for 3D reconstruction of stereo camera in dynamic image, however, the method is not limited to the specific case. That is, it essentially gives a basis for solving the association problem in multi sensor situation while the sensed information is reconstructed at the same time. Thus the extension of the method into multi sensor situation, not necessarily limited to camera image, but also including sensors with au-

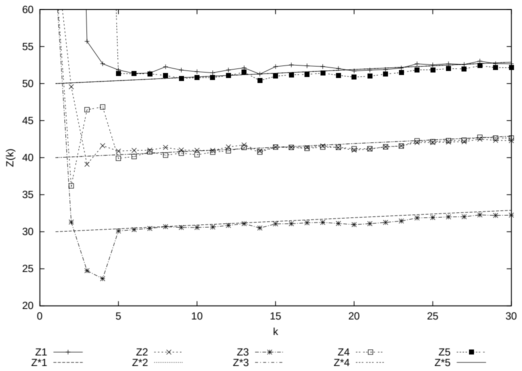

Figure 7: Estimation result of Z.

Table 1: Estimation results of associations.

k	Left image I_j^L						Right image I_j^R					
1	1	2	3	4	5	-	1	2	3	4	5	-
2	2	5	4	-	-	-	3	1	5	4	-	-
3	3	1	5	2	4	-	4	0	1	-	-	-
4	3	1	5	-	-	-	3	2	5	1	-	-
5	3	5	4	-	-	-	1	2	4	-	-	-
6	5	4	1	-	-	-	1	4	-	-	-	-
7	0	2	5	4	-	-	3	4	2	0	1	-
8	1	2	4	3	5	-	2	1	0	-	-	-
9	2	4	3	5	-	-	1	2	3	4	-	-
10	0	5	1	-	-	-	0	2	5	4	0	-
11	5	1	2	4	0	-	3	2	1	4	5	-
12	1	3	2	-	-	-	2	1	5	3	-	-
13	5	3	2	1	-	-	5	0	-	-	-	-
14	5	1	4	-	-	-	0	5	3	-	-	-
15	4	5	3	2	1	-	5	1	3	2	-	-
16	3	4	2	5	-	-	1	2	3	5	-	-
17	3	4	1	5	2	-	2	1	4	-	-	-
18	3	5	2	1	-	-	0	4	0	2	3	-
19	2	5	1	3	4	-	2	0	3	5	4	-
20	3	-	-	-	-	-	2	5	-	-	-	-
21	1	2	3	-	-	-	2	4	-	-	-	-
22	0	3	5	1	2	-	3	4	2	1	-	-
23	5	1	3	2	-	-	3	2	5	4	-	-
24	5	4	1	3	2	-	1	2	3	5	-	-
25	3	5	4	2	1	-	5	3	2	4	1	-
26	2	0	5	3	-	-	4	5	3	1	-	-
27	2	4	5	1	3	-	3	0	4	1	-	-
28	1	4	2	-	-	-	3	0	5	1	4	-
29	0	4	5	1	2	-	1	5	3	-	-	-
30	3	2	0	1	-	-	5	1	4	-	-	-

dio, ultrasonic wave, microwave, infrared light, and so on, is interesting for the future work.

Acknowledgment

This study was carried out under the ISM Cooperative Research Program (2003-ISM·CPR-2017).

References

[1] M.Z.Brwon, D.Bruschka, and G.D.Hager: Advances in Computational Stereo, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.**25**, No.8, pp.993-1008(2003).

[2] G.Casella and C.P.Robert, "Rao-Blackwellisation of Sampling Schemes", *Biometrika*, **83**, 1, pp.81-94, 1996.

[3] A.Chiuso, P.Favaro, H.Jin, and S.Soatto: "Structure from motion causally integrated over time", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, **24**, 4, pp.523-535, 2002.

[4] J. Costeria and T. Kanade, "A multi-body factorization method for motion analysis", *Technical report CMU-CS-TR-94-220*, Computer Science Department, Carnegie Mellon University, 1994.

[5] A.Doucet, J.F.G.de Freitas, and N.J.Gordon (eds): *Sequential Monte Carlo Methods in Practice*, New York, Springer (2001).

[6] A.Doucet, N.J.Gordon, and V.Krishnamurthy: Particle Filter for State Estimation of Jump Markov Linear Systems, *IEEE Trans.on Signal Processing*, Vol.**49**, No.3, 613-624 (2001).

[7] A.Doucet, S.J.Godsill, and C.Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering", *Statistics and Computing*, Vol.**10**, No.3, 197-208, 2000.

[8] N.J.Gordon, D.J.Salmond, and A.F.M.Smith: Novel approach to nonlinear / non-Gaussian Bayesian State Estimation, *IEE Proceedings-F*, **140**, No.2, pp.107-113(1993).

[9] N.J.Gordon and D.J.Salmond: Bayesian target tracking after group pattern distortion, *Proc. of the SPIE*, **3163**, pp.238-248(1997).

[10] N.J.Gordon and A.Doucet: Sequential Monte Carlo for Manoeuvering Target Tracking in Clutter, *Proc. of the SPIE*, **3809**, pp.493-500(1999).

[11] R.Hartley and A.Zisserman: Multiple View Geometry in Computer Vision, Cambridge University Pres (2000).

[12] N.Ikoma, N.Ichimura, T.Higuchi, and H.Maeda: Particle filter based method for maneuvering target tracking, *Proc. of the 2001 IEEE International Workshop on Intelligent Signal Processing*, pp.3-8 (2001).

[13] N.Ikoma, T.Higuchi, and H.Maeda: Tracking of maneuvering target by using switching structure and heavy-tailed distribution with particle filter method, *Proc. of the 2002 IEEE Conference on Control Applications*, pp.1283-1287 (2002).

[14] N.Ikoma and S.Godsill: Extended object tracking with unknown association, missing observations, and clutter using particle filters, *Proc. of the 2003 IEEE Workshop on Statistical Signal Processing*, to appear (2003).

[15] N.Ikoma, Y.Miyahara, and H.Maeda: Tracking of feature points in a scene of moving rigid objects by Bayesian switching structure model with particle filter, *Proc. of the 2003 IEEE International workshop on Neural Networks for Signal Processing*, to appear (2003).

[16] T.Irie, H.Maeda, and N.Ikoma: Synergetic Stereo Matching Algorithm for Occlusion and Reversal Position, *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol.**7**, No.2, to appear (2003).

[17] T. Kanade and D. D. Morris, Factorization methods for structure from motion, *Philosophical Transaction: Mathematical, Physical & Engineering Sciences, the Royal Society* **356**, pp.1153-1173, 1998.

[18] G.Kitagawa: Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics* **5**, No.1, pp.1-25(1996).

[19] J.S.Liu and R.Chen: Sequential Monte Carlo methods for dynamic systems. *Journal of the American Statistical Association*, **93**, pp.1032-1044(1998).

[20] S.McGinnity and G.W.Irwin: Manoeuvering target tracking using a multiple-model bootstrap filter. *Sequential Monte Carlo Methods in Practice*, Doucet, A, ed., Springer, pp.479-497(2001).

[21] Y.Ohta and T.Kanade: Stereo by Intra- and Inter-Scanline Search Using Dynamic Programming, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.**PAMI-7**, No.2, pp.139-154(1985).